

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Network Traffic Prediction Based on
Artificial Intelligence in Vehicular Ad-Hoc Networks**

SANAZ SHAKERSEPASGOZAR

Département de génie informatique et génie logiciel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie informatique

Décembre 2022

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

**Network Traffic Prediction Based on
Artificial Intelligence in Vehicular Ad-Hoc Networks**

présentée par **Sanaz SHAKERSEPASGOZAR**

en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Alejandro QUINTERO, président

Samuel PIERRE, membre et directeur de recherche

Soumaya CHERKAoui, membre

Abderrezak RACHEDI, membre externe

DEDICATION

*To my treasured parents,
for their endless support, love, and encouragement throughout my life...*

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my research director Professor Samuel Pierre for his continued support, guidance, and encouragement. I am thankful for giving me the opportunity to learn and grow as your Ph.D. student in Mobile Computing and Network Research Laboratory (LARIM). His motivation, patience, humanity, and kindness have deeply inspired me and it was an honor to work with him.

I would like to give special thanks to Professor Alejandro Quintero, Professor Soumaya Cherkaoui, and Professor Abderrezak Rachedi for being on my jury team and kindly accepting to participate.

I am thankful to my friends and colleagues at LARIM for having a friendly and informative environment. A debt of gratitude is owed to Dre. Franjeh El Khoury for her kindness, guidance, companionship, sympathy, and being a true friend all the while. I wish to say thanks to Eric, Amir, Sepehr, Lamia, Farnoush, Nasrin, Olson, Claudy, Jean, Loïc, Parvin, Marc, Carria, Nick, Somayyeh, Gaelle, and Hadjer as my valuable friends who accompanied and helped me on this journey with their helpful comments. The memories of the LARIM will remain in my mind and my heart forever.

From the bottom of my heart, I wish to express my utmost gratitude to my wonderful parents, Majid and Rozita for all the sacrifices, support, and care that gave me strength along my way. Their presence filled my heart full of love every moment even from the distance. Also, I am thankful to my brother Omid, my sister Sepideh and her husband Keramat, also my dearest nephew Sasha for their attention, support, and love. Without their support, it was impossible to pass this meandrous way.

Last, but certainly not the least, I would like to say thank you to my friend, my supporter, my love, and my spouse Alireza for his patience, unconditional love, and companionship along this path which we made it together.

RÉSUMÉ

Le système de transport intelligent "Intelligent Transportation System" (ITS) est une nouvelle solution pour le transport dans les villes intelligentes et un réseau Ad-hoc véhiculaire "Vehicular Ad-hoc Network" (VANET) est l'une de ses technologies les plus importantes, qui tente de rendre l'environnement plus sûr en offrant une meilleure communication sans fil. La prédiction, le contrôle et l'évitement du trafic peuvent améliorer les communications sans fil dans les villes intelligentes et, par conséquent, améliorer la qualité des services "Quality of Services" (QoS) et affecter les performances du réseau. Les VANET sont établis sur des véhicules intelligents, et peuvent avoir des communications de véhicule à véhicule "Vehicle-to-Vehicle" (V2V) et de véhicule à bord de la route "Vehicle-to-Roadside" (V2R). En outre, ils peuvent être utilisés pour des applications liées ou non à la sécurité, telles que le soutien à la sécurité routière, la gestion des places de stationnement, la mise à disposition d'autoroutes pour les véhicules d'urgence, etc.

En raison de la croissance des données générées par les ressources dans les villes intelligentes, les données sont de plus en plus volumineuses. Le stockage, le traitement et l'analyse de ces données sont importants pour prédire et contrôler le trafic sur le réseau. La prédiction du trafic a certaines limites comme l'utilisation de la bande passante, l'utilisation des ressources et le délai de transmission en communication et en informatique. L'utilisation de méthodes d'intelligence artificielle (IA) peut aider à prédire le trafic, ce qui permet de réduire les retards et les pertes de paquets et d'améliorer les performances des réseaux VANET.

Par conséquent, deux types de trafic sont présents dans le réseau VANET, à savoir le trafic routier et le trafic réseau qui peut se produire dans l'environnement. Chaque type de trafic (i.e., le trafic routier et le trafic réseau) a des raisons différentes de se produire, divers paramètres et plusieurs effets sur la communication et le transport. Par exemple, dans le cas du trafic routier, certains paramètres, tels que le nombre de véhicules, la vitesse du véhicule, les conditions météorologiques, les limitations de la route et autres, peuvent avoir un effet sur le trafic routier. Dans le réseau, d'autres paramètres, tels que la vitesse des expéditeurs, le retard, la perte de paquets et la bande passante peuvent être des paramètres efficaces pour le trafic réseau.

Dans la littérature, de nombreuses méthodes ont été proposées pour traiter séparément le trafic routier et le trafic réseau, et certaines prennent en compte l'effet des conditions météorologiques sur le trafic routier. Cependant, le trafic routier et le trafic réseau, compte tenu de la croissance du nombre de véhicules, des dispositifs intelligents et du volume de

données qu'ils produisent, peuvent s'affecter mutuellement. Par exemple, dans une situation de trafic routier, les véhicules peuvent perdre leur connectivité, des retards et des pertes de paquets peuvent se produire sur le réseau. Bien que, les méthodes émergentes de l'IA et leurs branches puissantes, y compris l'apprentissage automatique "Machine Learning" (ML) et l'apprentissage profond "Deep Learning" (DL), peuvent fournir la précision, la sécurité et la complexité de temps dans les problèmes de prédiction.

Dans cette thèse, une nouvelle méthode de prédiction du trafic réseau est proposée tandis que l'effet du trafic routier sur le trafic réseau est considéré. Le problème de prédiction peut apparaître dans les problèmes de classification comme lorsque l'objectif est de prédire des valeurs continues telles que la vitesse, ou dans les problèmes de régression où les algorithmes visent à prédire des valeurs discrètes incluant des cas vrais ou faux. Des algorithmes d'IA ont été mis en œuvre dans cette thèse pour prédire le flux de trafic du réseau, afin de couvrir les problèmes de prédiction dans les styles de régression et de classification. Un ensemble de données VANET réel basé sur les communications V2V et V2R, qui a recueilli des données en installant des antennes sur le dessus des véhicules, a été utilisé.

Le modèle de prédiction du flux de trafic du réseau basé sur l'évolution du trafic routier est proposé. Le modèle proposé est une intégration des algorithmes ML et DL, appelé algorithme de prédiction du trafic du réseau par "Random Forest-Gated Recurrent Unit-Network" (RF-GRU-NTP), dans le but de prédire le flux de trafic du réseau en tenant compte des paramètres de la route et du réseau en même temps. Le modèle proposé comprend trois phases et, dans chaque phase, différents algorithmes d'intelligence artificielle ont été mis en œuvre. Dans la première phase, le trafic réseau basé sur les communications V2R a été prédit à l'aide d'algorithmes ML, tels que "Random Forest" (RF), "K-Nearest Neighbors" (KNN), "Support Vector Machines" (SVM), et "Naive Bayes" (NB). Ils ont été évalués sur la base de métriques de classification et la RF a obtenu les meilleurs résultats alors que la "réception de paquets" était notre cible. Dans la deuxième phase, le trafic routier basé sur la communication V2V a été prédit, la "vitesse de l'expéditeur" étant notre cible, et nous avons supposé qu'une situation de trafic se produit lorsque la "vitesse de l'expéditeur" tombe en dessous de 60 Km/H comme seuil, et qu'au-delà, il s'agit d'une situation de non-traffic. Trois algorithmes d'apprentissage profond, dont "Long Short-Term Memory" (LSTM), "Bi-directional Long Short-Term Memory" (Bi-LSTM) and "Gated Recurrent Unit" (GRU), ont été mis en œuvre pour prédire le trafic routier. Leurs résultats ont été évalués sur la base de métriques de régression et les meilleurs résultats ont été obtenus par la GRU.

Dans la dernière phase, le flux de trafic du réseau est prédit en tenant compte du trafic routier en implémentant les algorithmes DL et ML. À cette fin, les ensembles de données V2V et

V2R ont été combinés et la RF a été mise en œuvre pour la sélection des caractéristiques. L'objectif principal de la mise en œuvre de la RF était de trouver les caractéristiques les plus efficaces de la route et du réseau sur le flux de trafic du réseau, à savoir la "réception des paquets" et la "vitesse du récepteur", qui peuvent affecter la "vitesse de l'émetteur" et le flux de trafic du réseau. Ensuite, l'algorithme RF-GRU-NTP proposé a été mis en œuvre pour la prédiction du trafic réseau, et les résultats de l'évaluation par régression ont révélé que le modèle proposé est capable de surmonter l'algorithme de base avec l'erreur de prédiction et le temps d'exécution les plus bas pendant le processus d'apprentissage.

De plus, cette thèse propose une analyse comparative de cinq algorithmes d'IA communs pour la prédiction du trafic réseau dans VANET. L'objectif principal de l'analyse de prédiction de classification proposée est de prédire l'état de "réception de paquets" comme paramètre de trafic routier. Un ensemble de données VANET réel basé sur la communication V2R est utilisé. Les données sont étiquetées en classe 0 pour la non réception des paquets, ce qui montre la situation de trafic et la classe 1 pour la réception des paquets qui dépeint la situation de non-traffic. Quatre algorithmes ML différents, dont RF, KNN, SVM et NB, ont été mis en œuvre. En outre, une classe de réseaux neuronaux artificiels (ANN) de type feed-forward appelée Multi-Layer Perceptron (MLP) a été mise en œuvre. L'objectif principal est de comparer les performances des algorithmes IA en fonction de différentes mesures de classification, afin de trouver le meilleur algorithme de classification pour la prédiction du trafic réseau en fonction du paramètre "réception de paquets". Les résultats de la simulation ont révélé que, bien que le KNN et le RF aient eu de bonnes performances en termes de précision et d'erreur, le KNN a pris un temps d'exécution plus élevé (6,47 minutes) par rapport au RF (0,73 minute). Par conséquent, sur la base de toutes les mesures d'évaluation, nous pouvons conclure que le RF, en tant qu'algorithme ML traditionnel, pourrait avoir les meilleures performances en termes de précision, d'erreur de prédiction et de temps d'exécution.

En ce qui concerne la sécurité des données dans le réseau, un algorithme d'apprentissage fédéré pour la prédiction du trafic réseau (Fed-NTP) est proposé. Dans cette partie de la thèse, un algorithme sécurisé appelé apprentissage fédéré "Federated Learning" (FL) basé sur LSTM pour la formation locale est proposé dans le but de prédire le trafic du réseau en utilisant simultanément les paramètres de la route et du réseau. L'algorithme FL fonctionne de manière décentralisée, ce qui signifie qu'il n'est pas nécessaire d'envoyer les données brutes au serveur, mais que le modèle formé se déplace entre les clients et les serveurs du réseau. Un jeu de données VANET réel (de type réseau ad hoc 802.11) a été adopté, incluant les communications V2V et V2R. Pour le paramètre de trafic du réseau, la "vitesse de l'émetteur" a été choisie avec un seuil supposé (< 60 Km/h) pour le trafic du réseau et la "vitesse du récepteur" a été considérée comme un paramètre du trafic routier capable d'affecter le flux

de trafic du réseau.

Les résultats de l'évaluation par régression ont montré que l'algorithme Fed-NTP proposé est plus précis que les autres algorithmes de base (algorithmes centralisés et décentralisés) et qu'il peut préserver la confidentialité des données en conservant le client et en transférant le modèle formé au serveur. Le langage de programmation python a été utilisé pour mettre en œuvre toutes les méthodes d'IA proposées dans cette thèse. Différentes métriques de régression et de classification ont été utilisées pour évaluer les performances des méthodes proposées en termes de précision, d'erreurs de prédiction, de temps d'exécution et de sécurité des données. Les résultats de la simulation ont révélé que l'algorithme RF-GR-NTP proposé pouvait prédire le flux de trafic du réseau avec précision, tout en tenant compte des paramètres de la route et du réseau avec un temps d'exécution minimal. Les résultats de la comparaison de différents algorithmes d'IA pour la prédiction du flux de trafic du réseau ont montré que le RF, bien qu'étant un algorithme d'apprentissage traditionnel, est capable de prédire le flux de trafic du réseau avec une plus grande précision et un temps d'exécution plus faible. De plus, les résultats de la simulation ont montré que l'algorithme Fed-NTP proposé est capable de prédire le trafic réseau en tenant compte des paramètres de la route et du réseau simultanément tout en préservant la confidentialité des données.

ABSTRACT

Intelligent Transportation System (ITS) is a novel solution for transportation in smart cities and a Vehicular Ad-hoc network (VANET) is one of its most important technologies, which tries to make the environment safer by having better wireless communication. Predicting, controlling and avoiding traffic can improve wireless communications in smart cities and consequently, enhance the Quality of Services (QoS) and affect the network performance. VANETs are established on vehicles that are intelligent, and can have communications like vehicle-to-vehicle (V2V) and vehicle-to-roadside (V2R). Moreover, it can be employed for safety and non-safety applications, such as supporting the safety of the road, managing parking spaces, providing freeways for emergency vehicles and so on.

Due to the growth of data generated by resources in smart cities, we are facing big data. Storing, processing and analyzing big data is an important issue for prediction and control traffic over the network. Traffic prediction has some limitations like usage of bandwidth, usage of resources and transmission delay overhead in communication and computing. Using Artificial Intelligent (AI) methods can help traffic prediction, which leads to reduced delay and packet loss and improve the performance of VANET.

Therefore, we are facing two types of traffic in the VANET including road traffic and network traffic that could happen in the environment. Each type of traffic (i.e., road and network) has different reasons for happening, various parameters and several effects on communication and transportation. For instance, in road traffic, some parameters, such as the number of vehicles, the speed of vehicles, the weather condition, the road limitations and others, could be effective in road traffic happening. While in the network, other parameters including the speed of senders, the delay, the packet loss and the bandwidth, could be effective parameters in network traffic happening.

In the literature, so many methods have been proposed that address road and network traffic separately, and some consider the effect of weather conditions on road traffic. However, road and network traffic considering the growth of the number of vehicles, smart devices and their produced volume of data, can affect each other. For instance, in a traffic situation on the road, the vehicles can lose their connectivity, delays could happen in the network and packet loss can occur. Although, emerging AI methods and their powerful branches including machine learning (ML) and deep learning (DL), can provide accuracy, security and time complexity in prediction problems.

In this dissertation, a novel method for network traffic prediction is proposed while the effect

of road traffic on network traffic is considered. The prediction problem can appear in the classification problems like when the target is to predict continuous values such as speed, or in regression problems where the algorithms aim to predict discrete values including true or false cases. AI algorithms have been implemented in this dissertation to predict network traffic flow to cover the prediction problems in both regression and classification styles. A real VANET dataset based on V2V and V2R communications, which gathered data by installing antennas on top of the vehicles, has been used.

The network traffic flow prediction model based on road traffic happening is proposed. The proposed model is an integration of ML and DL algorithms called a Random Forest-Gated Recurrent Unit-Network Traffic Prediction algorithm (RF-GRU-NTP) with the aim of network traffic flow prediction considering road and network parameters at the same time. The proposed model includes three phases and, in each phase, different AI algorithms have been implemented. In the first phase, the network traffic based on V2R communications has been predicted using ML algorithms, such as Random Forest (RF), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Naive Bayes (NB). They have been evaluated based on classification metrics and the RF got the best results while the “packet receiving” was our target. In the second phase, road traffic based on V2V communication has been predicted where the “sender speed” was our target, and we assumed that the traffic situation is happening when the “sender speed” drops below 60 Km/H as a threshold, and above it would be a non-traffic situation. Three DL algorithms including Long Short-Term Memory (LSTM), Bi-directional Long Short-Term Memory (Bi-LSTM) and Gated Recurrent Unit (GRU), have been implemented in order to road traffic prediction. Based on regression metrics, their results have been evaluated and the best results belonged to the GRU.

In the last phase, the network traffic flow is predicted considering road traffic happening by implementing DL and ML algorithms. For this purpose, V2V and V2R datasets have been combined and the RF implemented as a feature selection. The main purpose of implementing RF was to find the most effective features in road and network on network traffic flow that were “packet receive” and “receiver speed”, which can affect the “sender speed” and the network traffic flow. Then, the proposed RF-GRU-NTP algorithm was implemented for network traffic prediction, and the regression evaluation results revealed that the proposed model is able to overcome the baseline algorithm with the lowest prediction error and execution time during the learning process.

Moreover, a comparative analysis of five commons AI algorithms for network traffic prediction in VANET is proposed in this thesis. The main target of the proposed classification prediction analysis is to predict the status of “packet receiving” as a road traffic parameter. A real

VANET data set based on V2R communication is used. The data is labeled into class 0 for not receiving the packets, which show the traffic situation and class 1 for receiving packets that depict the non-traffic situation. Four different ML algorithms including RF, KNN, SVM and NB have been implemented. Furthermore, a class of feed-forward Artificial Neural Network (ANN) called Multi-Layer Perceptron (MLP) has been implemented. The main target is to compare the performance of AI algorithms according to different classification metrics for finding the best classification algorithm for network traffic prediction based on the “packet receiving” parameter. The simulation results revealed that though both the KNN and RF had good performance in terms of accuracy and error, although, the KNN took a higher execution time (6.47 minutes) in comparison with the RF (0.73 minutes). Therefore, based on all evaluation metrics, we can conclude that the RF as a traditional ML algorithm could have the best performance in terms of accuracy, prediction error and execution time.

Regarding the security of data in the network, a federated learning algorithm for network traffic prediction (Fed-NTP) is proposed. In this part of the dissertation, a secure algorithm named federated learning (FL) based on LSTM for local training is proposed with the aim of network traffic prediction using road and network parameters simultaneously. FL algorithm works in a decentralized way, which means there is no need to send the raw data to the server instead the trained model will move between clients and servers in the network. A real VANET dataset (802.11 ad-hoc network type) has been adopted including V2V and V2R communications. For the network traffic parameter, the “sender speed” has been chosen with an assumed threshold (< 60 Km/h) for network traffic happening and the “receiver speed” has been considered as a road traffic parameter that is able to affect the network traffic flow. The regression evaluation results showed that the proposed Fed-NTP algorithm performed more precisely than the other baseline algorithms (centralized and decentralized algorithms) while it can preserve the privacy of data by keeping in the client and transferring the trained model to the server.

The python programming language has been used to implement all the AI proposed methods in this thesis. Different regression and classification metrics have been used to evaluate the performance of the proposed methods in terms of accuracy, prediction errors, execution time and security of data. The simulation results revealed that the proposed RF-GR-NTP algorithm could predict network traffic flow accurately, while considering road and network parameters with the minimum execution time. The comparison results of different AI algorithms to network traffic flow prediction showed that the RF in spite of being a traditional learning algorithm is able to predict network traffic flow with higher accuracy and lower execution time. Moreover, the simulation results showed that the proposed Fed-NTP algorithm is able to perform network traffic prediction considering road and network parameters

simultaneously while the privacy of data is preserved.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGEMENTS	iv
RÉSUMÉ	v
ABSTRACT	ix
TABLE OF CONTENTS	xiii
LIST OF TABLES	xvi
LIST OF FIGURES	xvii
LIST OF SYMBOLS AND ACRONYMS	xix
LIST OF APPENDICES	xxiv
CHAPTER 1 INTRODUCTION	1
1.1 Definitions and Basic Concepts	2
1.2 Problem Statement	7
1.2.1 Need to Extract the Effective Parameters in Traffic Flow	7
1.2.2 Shortage of Intelligent Methods for Network Traffic Prediction Based on Road Parameters	8
1.2.3 Safety of Data	8
1.2.4 Underestimating the Importance of Applying Traffic Prediction	8
1.2.5 Considering the Pros and Cons of AI Methods to Find a Suitable Al- gorithm for Prediction Problems	9
1.2.6 Congestion and Traffic Flow	9
1.2.7 Research Questions	10
1.3 Research Objectives	11
1.4 Global Research Methodology	11
1.5 Research Contributions	16
1.6 Outline of Dissertation	17
CHAPTER 2 LITERATURE REVIEW	18

2.1	The Importance of Traffic Prediction in Vehicular Networks	18
2.2	Traffic Prediction and Control Based on Road Information	21
2.3	Traffic Prediction Based on Network Information	31
2.4	Traffic Prediction and Control Using Intelligent Methods	39
CHAPTER 3 ARTICLE 1: NETWORK TRAFFIC PREDICTION MODEL CONSIDERING ROAD TRAFFIC PARAMETERS USING ARTIFICIAL INTELLIGENCE METHODS IN VANET		
		51
3.1	Introduction	51
3.2	Background and Related Work	54
3.2.1	Network Traffic Prediction	54
3.2.2	Road Traffic Prediction	56
3.3	Methodology	57
3.3.1	Phase 1 - Network Traffic Flow	58
3.3.2	Phase 2 - Road Traffic Prediction	62
3.3.3	Phase 3 - Network Traffic Prediction Considering Road Parameters	66
3.4	Data Preparation and Performance Evaluation	69
3.4.1	Dataset	69
3.4.2	Performance Evaluation of The Network Traffic Prediction	70
3.5	Conclusion	79
CHAPTER 4 ARTICLE 2: A COMPARATIVE STUDY OF ARTIFICIAL INTELLIGENCE ALGORITHMS FOR NETWORK TRAFFIC PREDICTION IN VANET		
		82
4.1	Introduction	82
4.2	Related Work	83
4.3	Proposed Model for Network Traffic Prediction	85
4.3.1	Dataset	85
4.3.2	Implementation	86
4.3.3	Selected AI Algorithms	86
4.3.4	Evaluation Metrics	88
4.4	Simulation Results	89
4.5	Conclusion	92
CHAPTER 5 ARTICLE 3: FED-NTP: A FEDERATED LEARNING ALGORITHM FOR NETWORK TRAFFIC PREDICTION IN VANET		
		94
5.1	Introduction	94
5.2	Background and Related Work	97

5.3	Methodology	99
5.3.1	LSTM Algorithm	100
5.3.2	Federated Learning	102
5.3.3	Proposed FeD-NTP	102
5.4	Data Preparation and Performance Evaluation	105
5.4.1	Dataset	105
5.4.2	Implementation	105
5.4.3	Evaluation Metrics	106
5.4.4	Experimental Results	106
5.5	Conclusion	108
CHAPTER 6 GENERAL DISCUSSION		111
CHAPTER 7 CONCLUSION		114
7.1	Summary of Contributions	114
7.2	Limitations	115
7.3	Future Work	116
REFERENCES		118
APPENDICES		134

LIST OF TABLES

Table 2.1	List of parameters	29
Table 2.2	List of parameters	30
Table 2.3	Data input and policy relationship for types of ML and DL al- gorithms	42
Table 3.1	Summary of related work in road and network traffic prediction	58
Table 3.2	Classification reports	73
Table 3.3	Execution time for the RF, NB, SVM and KNN	74
Table 3.4	Duration time for GRU, Bi-LSTM and LSTM	76
Table 3.5	Evaluation metrics for Bi-LSTM, GRU and LSTM	76
Table 3.6	Evaluation metrics for Bi-LSTM, RF-GRU-NTP and LSTM . .	78
Table 3.7	Duration time for RF-GRU-NTP, Bi-LSTM and LSTM	81
Table 4.1	Classification reports	90
Table 5.1	Summary of Related Work in Traffic Prediction	99
Table 5.2	Models Evaluation	107

LIST OF FIGURES

Figure 1.1	VANET system architecture.	3
Figure 1.2	AI and its branches [9].	5
Figure 1.3	Neural Network [9].	6
Figure 2.1	Layers of VANET [7].	19
Figure 2.2	Control congestion context within VANET [14].	22
Figure 2.3	The architecture of VGRAN.	23
Figure 2.4	Predictive trajectories and positions diagram.	25
Figure 2.5	Graphical view of the proposed scheme with intelligent agents.	27
Figure 2.6	A wireless mesh network.	37
Figure 2.7	Relation between the notations.	40
Figure 2.8	Behavior prediction using different AI-based models.	45
Figure 2.9	Timeline diagram of different artificial intelligence techniques.	46
Figure 2.10	AI-based model's contributions based on: (a) number of papers and (b) percentage-wise.	47
Figure 3.1	The architecture of VANET.	59
Figure 3.2	The architecture of VANET.	60
Figure 3.3	The workflow of network traffic prediction (phase1).	62
Figure 3.4	The structure of the GRU.	63
Figure 3.5	The workflow of road traffic prediction (phase 2).	65
Figure 3.6	The structure of the proposed RF-GRU-NTP model (phase 3).	67
Figure 3.7	The structure of the confusion matrix.	70
Figure 3.8	Confusion matrices of the RF, NB, SVM and KNN.	71
Figure 3.9	ROC curve for KNN, RF, SVM and NB.	73
Figure 3.10	PR curve for the KNN, RF, SVM and NB in network traffic prediction.	74
Figure 3.11	The MAE for GRU, LSTM and Bi-LSTM (phase2).	75
Figure 3.12	Test data vs prediction for the LSTM, GRU and Bi-LSTM (phase2).	77
Figure 3.13	Effective parameters in network traffic.	78
Figure 3.14	The MAE for GRU, LSTM and Bi-LSTM (phase3).	79
Figure 3.15	Test data vs prediction for the LSTM, RF-GRU-NTP and Bi- LSTM (phase 3).	80

Figure 3.16	Time for fitting the LSTM, RF-GRU-NTP and Bi-LSTM (phase 3).	81
Figure 4.1	The architecture of proposed comparison model.	85
Figure 4.2	The architecture of MLP network.	88
Figure 4.3	ROC curve for KNN, RF, SVM, NB and MLP.	89
Figure 4.4	PR curve for KNN, RF, SVM, NB and MLP.	91
Figure 4.5	Execution time for network traffic prediction using KNN, RF, SVM, NB and MLP.	92
Figure 5.1	VANET architecture and training in centralized learning. . . .	100
Figure 5.2	Structure of LSTM algorithm.	101
Figure 5.3	Sequence diagram for federated learning.	103
Figure 5.4	The architecture of proposed Fed-NTP model.	104
Figure 5.5	Network traffic prediction results for the LSTM, GRU, FedGRU and Fed-NTP.	108
Figure 5.6	The MSE loss for the LSTM, GRU, RNN, FedGRU and Fed-NTP algorithms.	109

LIST OF SYMBOLS AND ACRONYMS

AC	Access Category
ADAM	Adaptive Moment Estimation
AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
AMFN	Attentive Mechanism Federated Network
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ANTSC	An intelligent Naïve Bayesian Traffic Stable Cluster
AP	access points
ARIMA	Auto Regressive Integrated Moving Average
ARIMAX	Auto Regressive Integrated Moving Average with Explanatory variables
AR-LSTM	Auto Regressive Long Short-Term Memory
as1	First-Order Regressive Integrated Moving Average
AUC	Area Under the Curve
AWS	Amazon Web Services
BIC	Bayesian information criteria
Bi-LSTM	Bi-directional Long Short-Term Memory
BMH-DMM	Bursty Multi-node Handover-Distributed Mobility Management
BP-ANN	Back Propagation Artificial Neural Network
BPNN	Back-Propagation Neural Network
BS	base station
CALM	Continuous Air interface for Long and Medium range
CCU	Carrefour Control Unit
CD	Congestion Degree
CE	Cross-Entropy
CGASVR	Collected Genetic Algorithm Support Vector Regression
CMSVM	Cost-Sensitive Support Vector Machines
CNN	Convolutional Neural Network
CPS	Congestion Prevention System
CPSS	Cyber-Physical-Social System
CTPM	Congestion-based Traffic Prediction Model
DARIMA	Delta Auto Regressive Integrated Moving Average
DBSCAN	Density-Based Spatial Clustering of Applications with Noise

DL	Deep Learning
DLASTM	Delta Long Short-Term Memory
DLL	data link layer
DNN	Deep Neural Network
DNN-BTF	Deep Neural Network-Based Traffic Flow
DoS	Denial of Service
DRL	Deep Reinforcement Learning
DSA	Dynamic Spectrum Access
DSR	Decision Tree Regressor
DSRC	Dedicated Short-Range Communications
DT	Decision Tree
EAM	Evaluation Automatic Module
EATD	Earliest Arrival Time Destination
EDRL	Enhanced Deep Reinforcement Learning
EMD	Empirical Mode Decomposition
FedDA	Dual Attention-Based Federated Learning
FedGRU	Federated Learning-based Gated Recurrent Unit
Fed-NTP	Federated learning algorithm for Network Traffic Prediction
FeD-NTP	Federated learning algorithm for Network Traffic Prediction
FedSTN	Federated Deep Learning Spatial-Temporal Networks
FFNN	Feed-forward Neural Network
FL	Federated Learning
FPR	False Positive Rate
FRBSs	Fuzzy-Rule-Based Systems
GA	Genetic Algorithm
GACE	Genetic Algorithm Cross-Entropy
GAN	Generative Adversarial Network
GCN	Graph Convolution Neural network
GCN-DHSTNet	Graph Convolutional Network- Deep Spatio-Temporal Neural network
GCP	Google Cloud Platform
GIN	Graph Isomorphism Network
GNN	Graph Neural Network
GPS	Global Positioning System
GRNN	General Regression Neural Network
GRU	Gated Recurrent Unit
HA	Historical Average

HetVNET	Heterogeneous Vehicular Network
HFRBS	Hierarchical Fuzzy Rule-Based System
HMM	Hidden Markov Model
HOV	High Occupancy Vehicle
HS	Harmony Search
IGSARBFNN	Improved Gravitation Search Algorithm for Neural Network
IIoT	Industrial Internet of Things
IoT	Internet of Things
IPT	Intelligent Processing Technology
IR	Imbalanced Ratio
ITS	Intelligent Transportation System
JS	Jensen-Shannon
KNN	K Nearest Neighbor
LASSO	Last Adaptive Absolute Shrinkage and Selection Operator
LLC	Logical Link Contro
LM	Levenberg–Marquardt
LMD	Local Mean Decomposition
LP	Linear Prediction
LR	Linear Regression
LSTM	Long Short-Term Memory
LTE	Long-Term Evaluation
MA	Moving Average
MAC	Medium Access Control
MAE	Mean Absolute Error
MANET	Mobile Ad Hoc Network
MAPE	Mean Absolute Percentage Error
MEC	Mobile Edge Computing
MILP	Mixed Integer Linear Program
ML	Machine Learning
ML-CC	Machine Learning Congestion Control
MLSTM	Multivariate Long Short-Term Memory
MOTabu	Multi-Objective Tabu
MRE	Mean Relative Error
MSE	Mean Square Error
MSVM	Multi-Support Vector Machine
MTL	Multi Task Learning

NB	Naïve Bayes
NF	Normalizing Flow
NLP	Natural Language Processing
NM	Naive Method
NN	Neural Networks
NPAH	Neighbor Predictive Adaptive Handoff
NRMSE	Normalized Root Mean Square Error
OBU	onboard unit
PCA	Principal Component Analysis
PHFRBS	Paralle Hierarchical Fuzzy Rule-Based System
PHY	physical layer
PM	Probabilistic Model
PoI	Point-of-Interest
PR	Precision-Recall
PRHMM	Predictive Routing based on a Hidden Markov Model
PRTMS	Predictive Road Traffic Management System
PSO	Particle Swarm Optimization
QoS	Quality of Service
RBF	Radial Basis Function
RBFNN	Radial Basis Function Neural Network
RCCS	Remote Core Cloud Server
ResNet	Residual Network
RF	Random Forest
RF-GRU-NTP	Random Forest- Gated Recurrent Unit- Network Traffic Prediction
RL	Reinforcement Learning
RLCN	Recurrent Longterm Capture Network
RMSE	Root Mean Square Error
RMSProp	Root Mean Square Propagation
RNN	Recurrent Neural Network
ROC	Receiver Operating Curve
RSU	roadside unit
SAE	Stacked Auto Encoder
SAE-LM	Atacked Auto Encoder Levenberg–Marquardt
SARIMA	Statistical Auto-Regressive Integration and Moving Average
SCN	Semantic Capture Network
SDA	Stack Denoising Autoencoder

SDAPM	Stacked Denoising Autoencoder Prediction Model
SDN	Software-Defined Network
SGD	Stochastic Gradient Descent
SSE	Square Sum Error
STFSA	Spatial-Temporal Feature Selection Algorithm
StoS	Sequence to Sequence
S-TPCA	Smart Traffic Prediction and Congestion Avoidance
SVM	Support Vector Machines
SVR	Support Vector Regression
TIC	Theil Inequality Coefficient
TM	Traffic Matrix
TPR	True Positive Rate
TQL	Time Queue Length
TSE	Traffic State Estimation
TVMLR	Time-Varying Multiple Linear Regression
Uni-LSTM	unidirectional Long Short-Term Memory
UPTDN	Urban Public Traffic Dynamic Network
V2I	Vehicle to Infrastructure
V2R	vehicle-to-roadsid
V2V	vehicle-to-vehicle
V2X	Vehicle-to-Everything
VAE	Variational Auto Encoder
VANET	Vehicular Ad hoc Network
VAPE	Variance Absolute Percentage Error
VDGNN	Variational Directed Graph Neural Network
VFL	Vertical Federated Learning
VGRAN	Variational Graph Recurrent Attention Neural Networks
VLSTM	Vanilla Long Short-Term Memory
WAN	Wide-Area Network
WAVE	Wireless Access in Vehicular Environment
WOA	Whale Optimization Algorithm
XAI	Explainable Artificial Intelligence
XGB	Extreme Gradient Boosting
XGboost	Extreme Gradient Boosting
5G	Fifth Generation
6G	Sixth Generation

LIST OF APPENDICES

Appendix A	Classification Performance Metrics	134
Appendix B	GRU algorithm equations	135
Appendix C	Regression Performance Metrics	136
Appendix D	Minimum Variance	137
Appendix E	Importance Score	138
Appendix F	MLP algorithm equations	139
Appendix G	LSTM algorithm equations	140

CHAPTER 1 INTRODUCTION

Vehicular Ad-hoc Networks (VANETs) are established on vehicles that are intelligent and can have communications like vehicle-to-vehicle (V2V) and vehicle-to-roadside (V2R). VANET can be employed for safety and non-safety applications such as supporting the safety of the road, managing parking spaces, providing freeways for emergency vehicles and so on. Predicting, detecting and controlling congestion can lead to a decrease in urban traffic flow, energy consumption, car accident and injury. Traffic flow prediction with high accuracy is significant in current transportation systems and it can help have the best path planning and make a better choice in selecting the best route for travelers and decrease traffic congestion. Distinguishing where and when the traffic will happen is promising for managing transportation and it could be able to assign resources such that ways which might be at the risk of congestion so it could reduce the traffic flow [1].

According to a large number of vehicles, the volume of data is growing up so, managing and storing big data is one of the main challenges that should be solved. The use of intelligent ways via machine learning techniques is the optimum solution that can address this critical problem. Computational predicting strategies like Bayesian modeling, fuzzy logic, hybrid modeling, Neural Networks (NN) and statistical modeling have been used for predicting traffic flow. All these approaches, especially NN, are verified to be effective in detecting the statistical rules that are hidden in the traffic data and forecasting the traffic flow in the future [2]. The significant point that should be considered in all these ways, is the accuracy of prediction.

Some works for predicting the traffic flow focused on sending messages with vehicles to vehicles by using Road-Side Units (RSUs) and Base stations (BSs) to a center about the events that happen on the road and to make aware other vehicles of the congestion of roads. However, the accuracy of messages is challenging because there might be some malicious or selfish vehicles that can astray the other vehicles. So, the messages should be validated to prevent sending false messages to both vehicles and servers [3].

Neural Networks can learn complicated nonlinear systems, so they are suitable to predict traffic flow. Earliest, the back-propagation neural network was used to forecast the short-term traffic flow. Although there are some different models of NN that are used in the field of traffic forecasting consisting of radial basis function neural network, recurrent neural network, time-delayed neural network model, the reverse propagation algorithm, multi-layer feedback neural network and other similar models [4].

1.1 Definitions and Basic Concepts

Vehicular Ad-Hoc Network (VANET)

One of the most important technologies for the Intelligent Transportation System (ITS), is VANET that which tries to make the environment safer and have better transportation by using wireless communications. By exploiting some principles of Mobile Ad Hoc Networks (MANETs) and some improvements to their weaknesses, VANETs are generated and accordingly, they inherit many attributes of MANET. These attributes in VANET lead to facing some challenges like scalability, security, high mobility, connectivity and so on that could affect the performance of VANET [5].

The main difference between them is that the role of nodes in MANET is played by vehicles in VANET so the mobility in this type of network is going to be higher. Generally, VANET contains a group of moving vehicles and the communications in VANET are in the form of V2V, V2R and Vehicle to Infrastructure (V2I). All these communications are wireless and try to improve the safety of roads and decrease traffic accidents. Another concern about VANET is the reliability of the network and to provide this important issue, we should consider the Quality of Service (QoS) [5]. To decrease the conflict between the restricted capacity of the cellular network and big data aggregation, the VANET could be a promising solution [6]. The structure of VANET includes some portions like onboard units (OBUs) that are installed in the vehicles, infrastructure systems like RSUs or access points (APs), vehicles and communication channels that are wireless. The architecture of VANET tries to create a wireless connection between vehicles with each other and with the infrastructures on the road to have the latest information about the road. In VANET three modes of communication are represented in Fig. 1.1 and explained as follows:

- Vehicle-to-Vehicle (V2V): in this kind of communication, the vehicles can make a connection with their neighbor directly.
- Vehicle-to-Roadside unit (V2R): this communication mentions the information replaced among vehicles and a roadside unit like the sign of warnings or traffic lights on the road where the server for storing the data is the roadside unit. Also, they could be an intermediate hop for enhancing the range of communication to the destination. At long distances, the V2R communications are reachable diverse V2V communications.
- Vehicle-to-Infrastructure (V2I): vehicles by this type of communication, can connect to the internet and take advantage of various internet services.

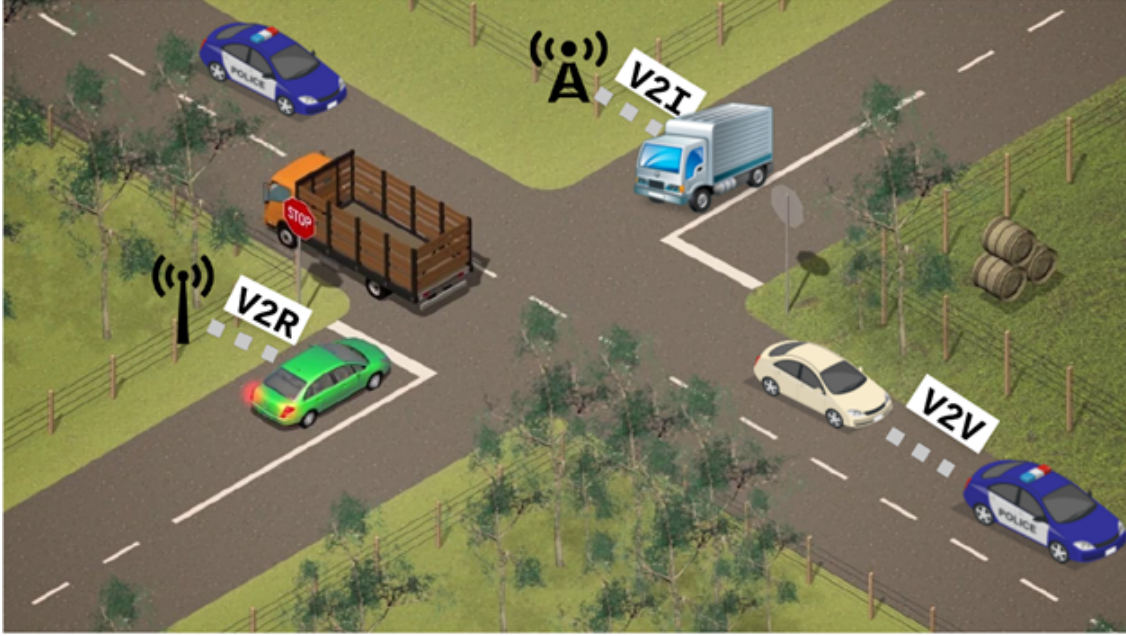


Figure 1.1 VANET system architecture.

Also, there is another communication called hybrid which is a combination of V2V, V2R and V2I that the vehicles can communicate with the infrastructures or RSUs via multi-hops or single-hop way which is relying on the distance. It supports communications for vehicles that are far [7]. As well, it could be used in heterogeneous networks. Also, VANET consists of six layers including an access layer, Network and transport layer, Facilities layer, Security layer, Management layer and Application layer. However, VANET can support two types of standards which are Wireless Access in Vehicular Environment (WAVE) standard and Continuous Air interface for Long and Medium range (CALM) [7], [8].

VANET has some challenges that should be considered. The security of VANET is an important concern because it might be subjected to various attacks such as Denial of Service (DoS) attacks, Gray hole, packet dropping attacks, Sybil attacks and etc that various novel methods have been proposed in order to keep the safety of VANET [9]. Based on the movement of VANET components, mobility management is another challenge in VANET that various solutions have been proposed for that [10]. We can mention the three challenges in mobility management including hand-off, localization and communication management. If we implement convenient mobility management, so the demand for services in the VANET would perform very well by keeping active communications [11]. Finally, due to the large amount of data that is produced in wireless networks such as VANET, storing and processing data is a crucial matter [12].

Traffic problem

When the volume of information is more than the capacity of the network, congestion has occurred which leads to traffic and it could decrease the efficiency of the network, enhance the number of packet losses and increase the delay and failed messages. The congestion control solutions that are proposed have some constraints. Congestion control should guarantee the delivery of the messages while keeping their security, so the first step is detecting the congestion and then controlling it. Control data congestion is an effective solution that could decrease the delay and packet loss and enhance the reliability of the network [13]. One of the challenging subjects in VANET is congestion control [14] that can be categorized into closed-loop and open-loop solutions [15] and by using various measurement methods, the congestion can be detected [16].

The precise prediction of traffic data could be an efficient solution for avoiding traffic congestion. When congestion prediction has occurred, it could be beneficial for controlling congestion and data management. One way for traffic prediction is to detect congestion and analyze the data of traffic. By traffic prediction, we can assign more resources to the areas that are at risk of congestion, so it could be useful for managing transportation [1]. Traffic prediction could be one of the principles of data management and resource allocation. Different types of technologies have been used for traffic data prediction but some of them have limitations and there are still some challenges. Data analysis by using data mining methods, machine learning algorithms and artificial intelligence methods proposed efficient models for congestion prediction.

Data mining

Due to the high growth of vehicles and the large amount of data that is generated by them, one of the optimum methods is data mining which can be effective in predicting, detecting and controlling congestion. By using data mining techniques, we can find a pattern from a huge amount of data. Data mining by using machine learning algorithms and statistical models, try to find the patterns that are not visible while the volume of data is large. There are some methods for data mining: classification, association, prediction, clustering, sequential patterns and decision trees. Clustering is one of the most common techniques of data mining in this technique, some features which are more similar to each other and are closer will be grouped in one cluster. So, by using this method the analysis of data becomes easier.

Artificial Intelligence

Artificial Intelligence (AI) can be used for developing traditional data-driven strategies in various areas of scientific research. For instant, by using Vehicle-to-Everything (V2X) system and AI, we can have information from different sources to predict the congestion on the road for more safety in the environment and decrease delay and energy consumption. In some fields such as data science, natural language processing, computer vision, medical and robotics, applying AI can be used for designing algorithms that can predict. AI can enable irregular applications like autonomous transport facilities, data storage in vehicles, real-time traffic flow prediction and management, autonomous transport facilities and congestion control in Vehicular Ad-hoc Networks (VANETs). An overview of the techniques that are used in the area of AI research is shown in Fig. 1.2 [9].

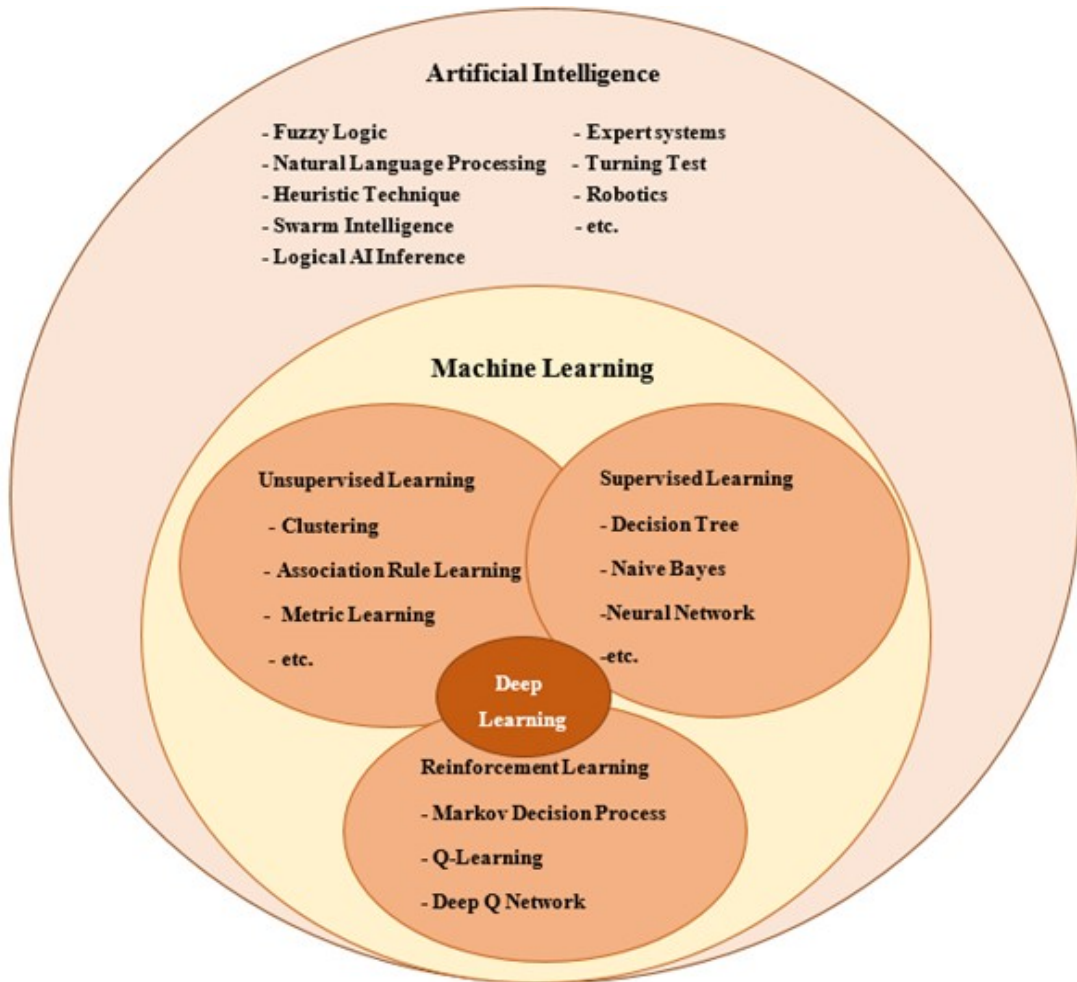


Figure 1.2 AI and its branches [9].

Neural Networks

An Artificial Neural Network (ANN) is used as a computing system that is including layers of neurons that can learn complex patterns to predict unexpected situations. In NN we have some layers that by applying some non-linear functions on their inputs, we can have the outputs that would be the inputs for the next layer. The first layer is called the input layer and after that, it would be some layers as hidden layers (if the number of them increased so we should use Deep Neural Network) and at the end, the out layers are deployed. A neural network model is depicted in Fig. 1.3

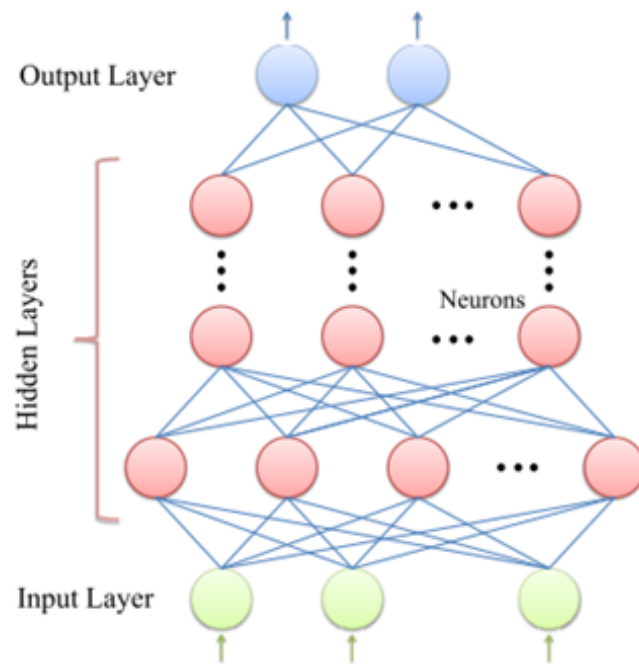


Figure 1.3 Neural Network [9].

By using the neural networks, the implementation of flow control will be more complicated so before the action, it needs a training data set. The training set can be formed by traffic patterns that are gained from the simulation results [16]. In the next phase, the parameters that are given from the training phase are used for a set of samples that is invisible and could be an evaluation criterion for the accuracy of prediction.

Machine learning algorithms

Machine learning (ML) protects a huge part of AI. ML techniques are divided into three types: Unsupervised learning (training would be based on unlabeled data), supervised learning (training would be based on labeled data) and Reinforcement Learning (RL) (it learns from the performance of the learning agent). Also, there are some types of ML schemes like transfer learning and online learning, that have sub-categorized these three types of ML schemes. ML is including two essential stages: training and testing. In the training phase, the model has been trained and then at the testing phase, based on the trained model the prediction will be done [9].

Deep learning algorithms

Deep Learning is a subclass of ML algorithms that are able to solve complicated problems. In order to build a computational model, the DL will use multiple layers with the aim of representing more detail of data. In last recent years, so many DL techniques have been proposed in various types like Natural Language Processing (NLP), speech and audio processing, visual data processing and etc which have shown promising solutions in different fields such as transportation and healthcare [17].

1.2 Problem Statement

1.2.1 Need to Extract the Effective Parameters in Traffic Flow

The main idea is to estimate when, where and how the traffic could be happening and avoid it. By monitoring the parameters that affect the traffic, we can find a pattern in their behavior of them on the road and network. We can say that the traffic forecast is reliable when we have extracted the effective parameters in traffic flow and utilized them in the implementation of forecasting algorithms, otherwise, the prediction can not be very accurate and reliable. Since we are facing a large number of data that are generated by vehicles, extracting valuable data is necessary for network optimization. We believe that using AI methods and specifically classification algorithms in extracting valuable data for predicting traffic even in unexpected situations, can eliminate the memory issue. So, the model can learn from the behavior pattern with high accuracy to predict traffic. Machine learning techniques can make a pattern from the behavior of effective parameters based on the history of the network for predicting traffic using extracted essential data. Some resources such as sensors, cameras, vehicles, etc are able to provide us with information in order to predict and control the traffic and the important

point is discovering the relation between hidden layers of data that can affect the traffic in the vehicular network.

1.2.2 Shortage of Intelligent Methods for Network Traffic Prediction Based on Road Parameters

Due to the increase in the number of vehicles, we can expect that we will have congestion and traffic flow on roads. Delays, increasing travel time and traffic congestion are some examples of factors that can increase fuel consumption, make pollution and greenhouse gas emissions. From another point of view, according to the large amount of data that is generated by vehicles, the network can also be affected by road traffic and suffer from disorder and network traffic flow. Hence, considering these two types of traffic simultaneously and extracting the effective parameters of traffic flow is an open problem that needs to be addressed.

1.2.3 Safety of Data

In last recent years by growing the volume of data that has been produced in smart cities, keeping the security of data is an important task. On the other hand, most AI techniques need data with the aim of doing training and testing processes. Therefore, transferring data in the network may lead to data leakage. Considering the importance of data security in vehicular networks, no matter how accurate a traffic model is, if it can not guarantee data security, it will not be a perfect prediction model. In this regard, using the federated learning (FL) algorithm gives the opportunity to keep the privacy of data by sending the trained model instead of local data over the network for the training process, while we have an accurate traffic flow prediction. However, to the best of our knowledge number of research works that proposed FL-based architecture to solve the network traffic problem in VANET is limited currently.

1.2.4 Underestimating the Importance of Applying Traffic Prediction

Traffic control and avoidance mechanism are important, however, they should not make us underrate traffic prediction. The first step to avoiding traffic happening is identifying the parameters which lead to occur traffic and predicting it. Basically, controlling and avoiding the traffic without having accurate predictions is not quite reliable.

1.2.5 Considering the Pros and Cons of AI Methods to Find a Suitable Algorithm for Prediction Problems

AI algorithms are promising solutions for prediction problems in terms of regression and classification. However, AI has various branches that each of them has some advantages and disadvantages. For instance, the KNN algorithm is easy to implement with a single hyperparameter but it is not good for a large volume of data due to its time complexity. Therefore, it is a very complicated task to implement an algorithm for traffic prediction which has the most advantages and the least disadvantages for the desired problem.

1.2.6 Congestion and Traffic Flow

Congestion is a phenomenon that could happen in the network when the volume of data is more than the capacity of the network and it could happen on the road when the number of vehicles is high. So, to manage this situation we should predict and control the congestion and also the traffic with the aim of decreasing the traffic flow. Although a lot of research has been done in this field, it is still a challenging issue that should be addressed.

Rapid growth in the number of smart vehicles and the volume of data that are generated by them leads to creating some restrictions on analyzing and storing the produced data while the security and integrity of data should be kept. For analyzing data, we must clean data by using data mining techniques to extract valuable data that could be useful to predict and decrease the congestion in the network. However, there are some limitations in terms of the capacity of the network, energy consumption, latency, etc.

Choosing a service that can protect the data is critical and some of the previous works selected cloud computing which has some drawbacks such as limited features, low capacity space, low speed and availability. A distributed network that can make a connection between VANET and cloud computing with low latency can solve the problem and ease the limitations. Some other works, utilized fog computing as an optimum framework that can cover most of the weaknesses of cloud computing and communication challenges in order to data storage and processing.

Concerning the importance of congestion and traffic happening in the network and its drawback on vehicular networks, so many novel methods have been proposed by many researchers. In most of them, network traffic and road traffic have been considered separately. Different parameters exist which can lead to creating traffic on the road such as road conditions, speed limitations, number of intersections and etc. On the other hand, network traffic could have different effective parameters for happening for instance delay, packet receiving, the speed

of senders, network platform and so on. Also, some research focused on road traffic considering the weather condition and their effect on traffic happening. However, network traffic prediction considering road traffic happening is still an open problem.

Different techniques and algorithms are utilized for solving traffic problems. The proposed methods have some advantages and disadvantages in terms of the recent techniques and algorithms specifically Artificial Intelligence, using the new methods can be a promising solution to push the performance of the network to the next level with higher performance. Machine learning techniques are able to learn from the model which exists and then used for new data and have two essential approaches including supervised learning and unsupervised learning. Supervised learning algorithms are based on using a labeled dataset to train. They could cover regression problems which are able to discover the relations between variables to predict numerical values or classification problems for categorical predictions. while unsupervised learning is used for analyzing an unlabeled dataset in order to discover the hidden patterns of data. Using an AI method with high accuracy while keeping the privacy of data is still an open problem. Therefore, considering recent proposed methods and technologies, proposing an intelligent method for network traffic prediction considering road traffic happening while the privacy of data is preserved are considered in this dissertation.

1.2.7 Research Questions

Regarding the aforementioned open challenges, the following unsolved major questions may raise: Considering the influential role of AI algorithms for prediction problems in vehicular networks and growing the number of smart vehicles in smart cities, how can we deal with network traffic prediction considering the most effective parameters in VANET?

More specifically:

- Regarding the high volume of produced data in vehicular networks, how we can propose a model which is able to precisely predict network traffic with high accuracy considering road and network traffic simultaneously?
- Based on different AI algorithms for prediction problems, which algorithm is the fittest for network traffic prediction in terms of regression or classification problems?
- While AI techniques need the data to do the training process, how can we utilize them for traffic prediction problems without passing the data over the network and keeping its privacy in VANET?

1.3 Research Objectives

The main objective of this dissertation is to design an AI-based method to predict traffic in VANET while data is generated by smart resources in the network. The following sub-objectives are considered for acquiring the main objective of this dissertation:

1. Devising an accurate network traffic prediction AI-based model considering road and network traffic parameters in VANET;
2. Proposing an intelligent network traffic prediction algorithm in classification problems based on AI algorithms;
3. keeping the privacy of data by implementing a deep learning algorithm in a federated way for network traffic prediction considering road and network parameters in VANET;
4. Evaluating the performance of the proposed methods.

1.4 Global Research Methodology

When we are facing a big volume of produced data in smart cities, extracting valuable information is a challenge that should be addressed. By extracting hidden and useful information about the behavior of the network we can make a pattern to predict the traffic in the network.

One of the main concerns about data is its structure. The collected data is divided into three categories. 1) Structured data: this type of data can be stored in a database, 2) Unstructured data: this kind of data is consisting of text and multimedia, and 3) Semi-structured data: they are not lie in the database but they have some features that make their analysis easier. Some processes lead to having reliable data sets such as aggregation, data collection, cleansing and preprocessing [18]. In this dissertation, we aim to use machine learning techniques for extracting the information.

The produced data in a smart city is a complex set with a high volume of data so it's difficult to analyze. Data mining is an optimum solution that can extract the data that we need to more understanding about the process which can lead to getting the desired conclusions and enhance the accuracy of prediction. Machine learning is a growing technology that has too many algorithms and tools for extracting knowledge from data [18].

Choosing a technique whose algorithms have the minimum process time and the maximum precision is a real challenge. Some advanced learning methods could be a promising solution for solving the problem of extracting effective parameters as follows:

- Representation Learning: it is a good solution for learning the valuable representation of data that can lead to extracting meaningful information. The target of representation learning is the enhancement of statistical and computational efficiency.
- Deep learning: it usually uses supervised or/and unsupervised strategies for learning hierarchical representations automatically. In the case of complex patterns, deep learning plays a vital role to provide predictive solutions.
- Distributed and parallel learning: in this framework, the learning is performed in a distributed manner. It leads to keeping energy and time by avoiding the requirements of gathering the data into one station while we have central processing.
- Transfer learning: it allows the tasks, distribution, and domains to be various while we have one or more sources, we can extract knowledge from them and employ knowledge to the main task. Transfer learning can use the experience to solve new problems.
- Active learning: when we are evolved with a large amount of data that is unlabeled, learning from them is hard and time-consuming. It aims to get high exactness while it uses minimum labeled data to decrease the cost of getting labeled data.
- Kernel-based learning: their main features are that they can map samples from the main space while the products can be calculated by kernel functions. The Kernel functions support the non-linear means to communicate or side information when we have big data [19].

Different DL algorithms are used in this dissertation in terms of extracting a complex pattern from a large volume of data. The main advantage of deep learning is that it can analyze and learn a large amount of data, creating a useful tool for data analytics while there is a large amount of raw data that they are unlabeled and un-categorized, automatically [20].

Previous works considered one type of data (road information or network information) for detecting, controlling and predicting the network traffic while we aim to combine two types of data for congestion prediction. Due to the volume of data that we will receive, just some part of them can help us to predict the network traffic so, by deploying machine learning algorithms we are able to get meaningful data that are influential on the traffic flow on road and consequently in the network.

Association rules are based on the if-then statement which means it shows the relationship between parameters. Association rules by using a machine learning algorithm can analyze the data for making behavior patterns in the database. For instance, if the weather is foggy

then the traffic would be more, and traffic will happen. Some parameters create our data set and could be classified for congestion prediction. The proposed method classified the data set into two groups:

- Road information: the data of this part is achieved by the Roadside Units (RSUs) that are installed on the road. They can receive the information that is broadcasted by vehicles, cameras, and sensors about the traffic flow, accidents, construction, speed of vehicles, the number of intersections, the road width, the weather and so on.
- Network information: it is including delay, the number of packet loss, bandwidth, connection, speed of senders and etc.

Traffic prediction in the network is a way of the network controlling so the accuracy of predictions is vital. For traffic prediction, we are motivated to use some parameters in data extracted to predict traffic flow. When we want to predict the traffic flow in the network, the accuracy of prediction is one of the main factors and we should not overlook the delay parameter. Traffic prediction is an efficient way that can enhance the QoS of VANET that need to be designed. For prediction, we need to train the data set so we will use a supervised machine learning algorithm.

According to the type of data that we have, the requirements that we need and the existing restriction for traffic prediction, AI algorithms are promising solutions. The following are some of their advantages:

- They consist of Neural Networks (NN) which can be learned by themselves and the output that they can produce is not restricted to the input which they received.
- They don't need the database for storing the input, and they can do this operation in their network so that when some data is lost, there is no effect on the performance of the network. So, they have distributed memory.
- They can learn from the training data set and employ them in unexpected situations. So, they have the ability of training machines.
- The output production and fault detection at the network do not depend on a neuron or piece of information.
- They can compute complex computations and accomplish multiple tasks while there is no effect on the performance of the network. So, they have parallel processing ability.

In the first specific objective of this dissertation, a model with the aim of network traffic flow prediction considering effective parameters in road traffic happening was proposed. To reach the first specific objective, three phases have been considered and in each phase, different ML learning and DL learning algorithms have been implemented using Python. The main focus in the first phase was on network traffic prediction based on packet receiving. Different ML algorithms including Random Forest (RF), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Naive Bayes (NB) algorithms have been implemented on the VANET dataset based on V2R communication. In order to evaluate their performances, various classification metrics have been used and the results revealed that RF got the best results for network traffic prediction considering packet receiving as a network parameter.

In the second phase, the main focus was on road traffic prediction based on sender speed as a road parameter. Various DL algorithms such as Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM) and Bi-directional Long Short-Term Memory (Bi-LSTM) have been implemented on the VANET dataset based on V2V communications. Different regression evaluation metrics have been used in terms of performance evaluation. The results showed that the GRU got the best results for road traffic prediction and predicting sender speed as a road parameter in terms of accurate prediction and duration time for implantation.

Finally, network traffic prediction considering the effect of the parameters on road traffic has been done in the third phase. The proposed RF-GRU-NTP model combined ML and DL algorithms with the aim of network traffic prediction considering road traffic happening. the RF algorithm has been implemented as a feature selection to find the most effective parameters in the traffic. A combined VANET dataset including V2V and V2R communication has been used and the results have been compared with LSTM and Bi-LSTM algorithms which shows the proposed model has the best result in network traffic prediction.

In the second specific objective, a comparative analysis of different AL algorithms for network traffic prediction in VANET has been done. Four traditional ML algorithms (i.e., KNN, RF, SVM and NB) and one AI algorithm based on NN called Multi-Layer Perceptron (MLP) have been implemented in Python with the aim of prediction packet receiving for network traffic prediction. A real VANET dataset based on V2R communications has been utilized and the packet receiving as a traffic parameter considers in terms of network traffic flow prediction. Various evaluation metrics including Receiver Operating Curve (ROC) curve, Precision-Recall (PR) curve, calculating the execution time and classification report have been implemented with the aim of finding the fittest algorithm for network traffic flow prediction.

Regarding the classification report, different metrics such as accuracy, precision, recall and

F1-score have been with the aim of measuring the performance of the prediction methods. Accuracy shows how accurate is the prediction model. Precision shows how many of the positive predictions, are truly predicted. Recall shows the number of times that the predicted model predicts the actual positive data correctly. In cases where the value of recall and precision are the same, the F1-score becomes an important parameter that will combine recall and precision as one parameter. The simulation results showed that the KNN and RF, with the slightly different, had the best performance while the KNN took remarkably more time for executing. So, based on all the evaluations, the RF got the most performed results for network traffic prediction in VANET based on packet receiving as a traffic parameter.

In the third specific objective, we went beyond the accuracy in traffic flow prediction based on road and network parameters and the privacy of the data has been concerned. A Federated learning algorithm for Network Traffic Prediction (Fed-NTP) has been proposed with the aim of predicting the network traffic flow precisely without sending the local data over the network. LSTM algorithm in a decentralized way by using a federated learning algorithm has been implemented on a real VANET dataset to network traffic flow prediction. The federated learning (FL) algorithm performs the training process on the model which has been sent by the workers in the network. The main point is that the local data will keep in the local place (workers) and just the trained model will transfer to the server over the network which reduces the risk of information leakage.

Python and PyTorch are used to implement the proposed model. Also, the PySyft library to design the FL algorithm has been used which is able to provide the FL requirements and can be used for private DL algorithms. The proposed Fed-NTP model has been implemented on a VANET dataset including V2V and V2R communications. With aim of defining the network traffic, the “sender speed” is considered a network parameter in V2R communication and the “receiver speed” has been considered a road parameter in V2V communication. The traffic happening assumed in a situation that the “sender speed” drops below 60 Km/h while there were not any speed limitations. The three pas values of each parameter have been taken for the prediction of the next value of “sender speed” as the network parameter and prediction of the network traffic flow.

In the first step, the server which is RSU in our work, will initial the model based on random parameters and send them to the workers that are vehicles. Then the LSTM has been performed to do the training process locally and send the trained model by the vehicles to the RSU. At that stage when the RSU will receive the trained models from the vehicles, it will aggregate all the models and send them back to be trained again. During this process, the local data will keep by the vehicles and the RSU just receive the trained model which

can guarantee the privacy of data besides the accurate prediction.

The performance of the proposed model has been compared with three centralized and one decentralized algorithm including LSTM, GRU, RNN and FedGRU respectively. Their performances have been analyzed by using different regression evaluation metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE) and $R^2 - SCORE$. RMSE shows the difference between the predicted value and the actual one. MAE shows the sum of differences between actual and predicted values. MSE illustrates the squared difference between actual and predicted values. MAPE shows the sum of the average difference between actual and predicted values. $R^2 - SCORE$ shows the efficiency of the predicted model and fits it in comparison between actual and predicted values. the simulation results revealed that the proposed model not only is able to predict the network traffic flow more accurately but also can preserve the privacy of data.

1.5 Research Contributions

Rapid growth in the number of smart vehicles and the volume of data that are generated by them leads to creating some restrictions on analyzing and storing the produced data while the security and integrity of data should be kept. On the other hand, the large number of smart vehicles can create road traffic and consequently network traffic happening which causes increasing energy consumption, delay, accidents and so on. Intelligent methods and AI algorithms are advantageous solutions for prediction problems in VANET which are proposed in this dissertation. Furthermore, the contributions of this dissertation are as follows:

1. **Proposing a network traffic prediction model considering road traffic parameters based on AI methods in VANET:** road traffic could affect the network traffic as well. By integrating the effective parameters in both road and network we can have more accurate network traffic prediction in terms of the most effective parameters. A combination of ML and DL algorithms could be a powerful method for having an efficient performance in network traffic prediction in VANET.
2. **Analyzing different AI algorithms for network traffic prediction in VANET:** Finding the most accurate and efficient algorithm for network traffic prediction is still a challenge. Comparing the prediction results of traditional learning algorithms and a DL algorithm, in order to network traffic flow prediction in classification problems has been done. Based on all evaluation results the RF as an ML algorithm could take advantage of other AI algorithms in prediction packet receiving as a traffic parameter.

3. **Proposing a federated learning algorithm for network traffic prediction with the aim of preserving the privacy of data:** Implementing the FL algorithm helps to keep the privacy of data by sending the trained model instead of local data to the server and over the network. The proposed Fed-NTP model not only can predict the network traffic flow considering the road and network parameters but also is able to preserve the privacy of data by keeping the local data in the local places(vehicles) instead of sending them over the network.

1.6 Outline of Dissertation

This dissertation is consist of seven chapters. Related works are reviewed in Chapter 2. In this chapter network different prediction methods for congestion and traffic on the road and network also their evaluation metrics are briefly explained and discussed.

In Chapter 3 a network traffic prediction model considering road parameters based on AI algorithms in VANET is proposed. Moreover, a proposed RF-GRU-NTP model is including three phases and in each phase, different ML and DL algorithms have been implemented. Different classification and regression evaluation metrics have been used with the aim of evaluating the performance of the proposed model and comparing the simulation results.

In Chapter 4 a comparative analysis has been made between various AI algorithms for network traffic flow prediction in VANET. Then in Chapter 5, a Fed-NTP model is proposed for network traffic flow prediction using federated learning algorithms in order to keep the privacy of data. A general discussion is presented in Chapter 6. Finally, in Chapter 7 the contribution of this dissertation, the restrictions and future works are presented.

CHAPTER 2 LITERATURE REVIEW

2.1 The Importance of Traffic Prediction in Vehicular Networks

The growing number of vehicles has led to a large volume of data and one of the main challenges is finding an optimized solution to predict and control the congestion and traffic to tackle this big amount of data. Smart traffic prediction can support the critical advantages in solving the problem of traffic congestion in the cities [21]. On the other hand, using intelligent methods can help us to have a pattern to predict the behavior of vehicles in a new situation that they haven't experienced before. Traffic prediction is a way that can reduce the noise in the environment, decrease the energy consumption and traffic flow in an urban environment, make the road safer and decrease the number of killed or injured and so on.

ITS is a system that relies on advanced technologies with the aim of providing smart and resilient safety services which are intended to predict and control traffic in order to prevent accidents and save lives. Data packet is one of the most important requirements of this kind of application. VANET has been designed to support wireless communications between vehicles and road supplies to exchange traffic data [22] and it consists of six layers that are shown in Fig. 2.1 and are explained in the following:

- Access layer: this layer is composed of the physical layer (PHY) with the aim of a physical device connected to the communication medium and the data link layer (DLL) that has two sub-layers: Medium Access Control (MAC) is responsible for managing the communication medium and Logical Link Control (LLC) which is for controlling the flow of data and managing the errors.
- Network and transport layer: the routing protocols are responsible to provide the reliability of communications.
- Facilities layer: the main task of this layer is presenting the information to the system users. Coding and decoding the messages based on the formal language that is used will be done in this layer.
- Security layer: the target of this layer is supporting the safety of the network against threats.
- Management layer: the duty of this layer is the management of communication service, networking, applications of VANET, general congestion control and heritage of system

protection.

- Application layer: different types of VANET applications that are responsible for security, checking the performance and reliability are in this layer [7].

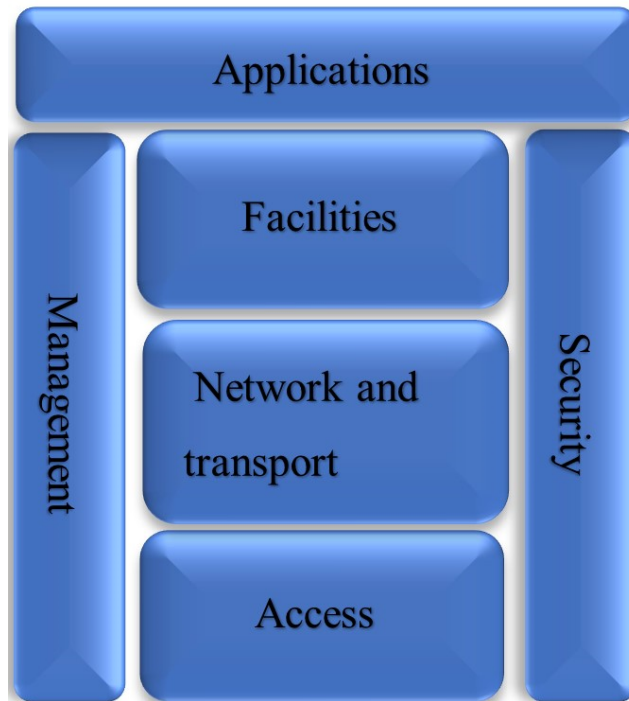


Figure 2.1 Layers of VANET [7].

VANET is including vital entities such as Road Side Units (RSUs), On-Board Units (OBUs) and vehicles. They are able to transmit information with wireless communications which are supported by dedicated short-range communication (DSRC) and accessed by the standard IEEE 802.11p and wireless access in vehicular environments (WAVE) is supported by it [23]. So much research has been done [24–26] based on the movements of vehicles and applied different methods to predict traffic with the aim of improving the delay, throughput and performance of the network [27]. Generally, VANET supports various kinds of standards for communications which two of them are explained as follows:

- Wireless Access in Vehicular Environment (WAVE) standard: this type of standard covers the attributes of Dedicated Short-Range Communications (DSRC), the IEEE 802.11p for PHY and MAC layers and the IEEE 1609 family for upper layers. DSRC for supplementary technology uses Dynamic Spectrum Access (DSA). When the mobility of vehicles is high and their speed is 10 up to 200 Kmph, the communication range

of 300 m, reaching up to 1000 m, and the rate of data is more than 27 Mbps, the DSRC/WAVE will support this situation.

- Continuous Air interface for Long and Medium range (CALM): it's a mixture of UMTS-3G, wireless technologies including GSM2G/GPRS-2.5G and Infrared communication and wireless systems while the band is 60 GHz and it is adjusted to IEEE 802.11p [7], [8].

One important concern about VANET is its safety against threats like jamming attacks. The Black hole and Wormhole are the two difficulties in discovering the routing attack in the VANET. For protecting the VANET against this kind of routing attack, a Swarm Intelligence algorithm was proposed. Some researchers used Artificial Neural Network (ANN), K Nearest Neighbor (KNN), Feed-forward Neural Network (FFNN) and SVM for protecting the network against different types of attacks such as Denial of Service (DoS) attack, Gray hole, packet dropping attack, Sybil attack and etc [9].

The main components of VANET are vehicles that are moving mostly. The nodes in the VANET have action and during this movement, their connectivity should have been kept. However, the security of mobility is a critical issue. Mobility management has three sub-functions: management of location, hand-off and forwarding. Location management traces the movements of vehicles inside or outside the network and it keeps the information about the location of vehicles. Hand-off consists of two steps including hand-off triggering and connection re-establishment. When the situation has been changed, so the hand-off will happen. Repeated hand-off in the VANET is just due to the movement of vehicles that are in the network and they lead to appear packet loss and hand-off delay. Common mobility management methods that are about the host and network mobility, do not support the VANET scenarios. So distributed mobility management is proposed but there are some challenges [10].

Due to increasing the number of smart devices in smart cities, a large amount of data will also be generated which can be called big data. Big data is including structured and unstructured data that are generated by different sources. In general, when we have a large amount of data, processing, analyzing and managing them is not possible with traditional methods. So, when we have a large and complex data set, we are facing big data. The authors in [12] considered five characteristics of big data as known 5V's as follows:

- Velocity: due to analyzing the large data that is generated, the speed of transmission and data accessibility in real-time should be considered. By using big data technology, we can analyze the data without putting them into databases.

- Volume: it refers to the amount of data that is generated by different sources like social media, cell phones, vehicles, cameras, sensors, etc. by using distributed systems, we can store some parts of data in different locations and then fetch them together.
- Value: it refers to the value of the extracted data.
- Variety: it is about the different types of data that we can use. Big data technology lets us store structured and unstructured data concurrently.
- Veracity: it mentions the quality or reliability of the data.

2.2 Traffic Prediction and Control Based on Road Information

Finding an optimum way for congestion prediction can help reduce traffic flow and fuel consumption. In [28], proposed a Congestion Prevention System (CPS) that is based on VANET to find the best direction that can help the driver reduce the travel time that leads to a decrease the fuel consumption. For achieving this goal, they focused on analyzing the trustful traffic information. Increasing the number of vehicles and lack of information on drivers are the most common reasons for congestion. They consider some parameters for finding the best path like the maximum speed at each part of travel, the number of vehicles that exist on the road due to the capacity of the road and the number of intersections and traffic lights. The method that the authors proposed is intending to reduce travel time, decrease the downtime at intersections and roundabouts and manage the traffic flow on the road network. By achieving these goals their model would be able to reduce the waiting time, congestion of roads, fuel consumption, greenhouse gas emission and the stress of driving and make a safer environment. In this model, they needed traffic information from infrastructures for analysis. In the first step, they recognized the road map and divided the map into smaller parts based on some parameters for having simpler analysis and each part has Roadside Unit (RSU). After that, they collected the movement information on vehicles and saved them in a database. For the next step, they predicted the time of travel due to the chosen route. For calculating the difference between the travel time of the road and the time that vehicles need to arrive at this road, they used a recursive function called “getNbrVehicle”. They determined the best way based on Dijkstra’s algorithm. The results showed that they could achieve their objectives.

The context of congestion control in VANET is depicted in Fig. 2.2 [14]. We can categorize the congestion control solution into two groups including the closed-loop solution which when the congestion has been detected, then it would be controlled and the open-loop solution

which avoids the congestion before it occurred [15]. For congestion detection, we can use measurement methods that can monitor the count of messages in the line, the amount of channel usage and the time of channel occupancy [16].

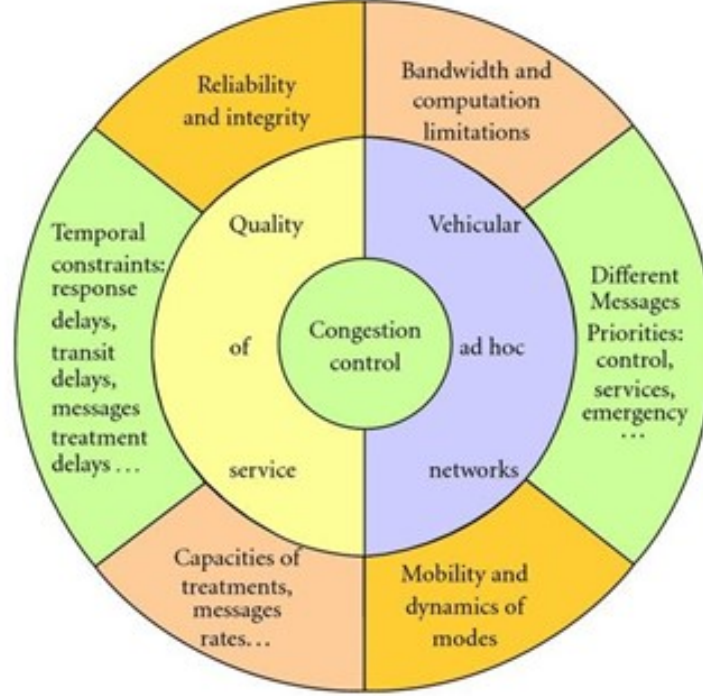


Figure 2.2 Control congestion context within VANET [14].

Intelligent transportation system (ITS) is one of the most critical applications in the industrial Internet of Things (IoT) which tries to enhance the safety and effectiveness of transportation networks. A Bayesian framework called variational graph recurrent attention neural networks (VGRAN) was proposed in [29] with the aim of traffic prediction. The proposed framework was based on capturing time-varying from the installed sensors in the road by using dynamic graph convolution operations. The temporal traffic correlations and stochasticity of sensor attributes are considered in the proposed model to make it more flexible. Also, they utilized the variational autoencoder (VAE) [30] with the aim of encoding stochastic hidden features in representation and time-varying sensor readings structures. The proposed model has two parts including a variational directed graph neural networks (GNN) named (VDGNN) module which has been used for learning spatial and temporal variables through the sensors and a normalizing flow (NF) based seq2seq module to obtain the dependencies of long-short-term traffic sequences. The architecture of the proposed VGRAN is shown in Fig. 2.3. The authors implemented the proposed model on n large-scale traffic datasets which

the experimental results based on MAE, MAPE and RMSE as evaluation metrics, showed that the proposed model overcame the other baseline algorithms. But there are still some challenges that should be considered about the proposed model which is the proposed model is not able to obtain severe events such as accidents or congestion.

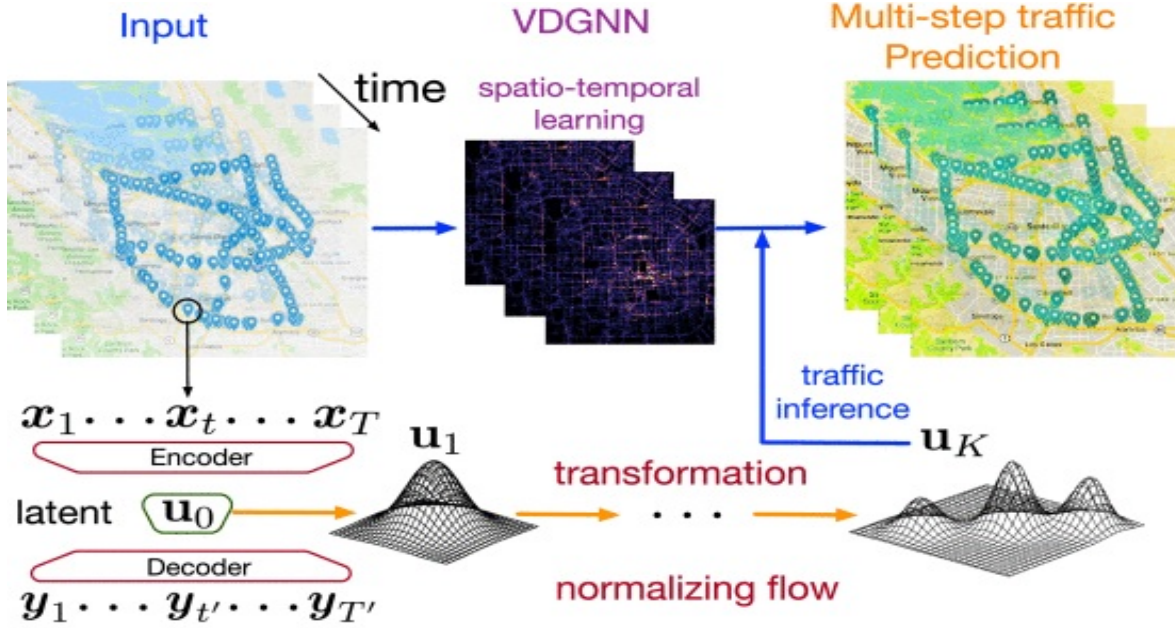


Figure 2.3 The architecture of VGRAN.

Predicting congestion is a way that can help reduce traffic, waiting time and fuel consumption. In [31], the authors proposed a technique for congestion prediction in the urban environment. It was event-based route selection and depended on the information that the sensors of the network collected. They focused on special events like bad weather, crashes, low control at the intersection and so on. In this work, they combined a central point, a sensor network, a knowledge base and various modules of software for managing the knowledge or predicting the traffic. The video sensors collected the traffic data from the intersections. The central node calculated the Time Queue Length (TQL) and then reserve all data in a buffer and executed each module. They generated traffic patterns for training the radial basis function network (RBF) network and used the data collected in one week in different conditions of traffic. They used a multi-layer neural network called RBF that the k-means clustering algorithm used in it, to predict the parameters of traffic. The technique that they presented leads to estimating two parameters of traffic which are total queue length at intersections and the set of roads that can affect traffic congestion. The results showed that by using the proposed method, traffic congestion, delay, and fuel consumption will be reduced.

In [32], the authors tried to minimize traffic congestion by predicting travel time. They proposed an architecture that considers predicted transit-times by sending feed-back to RSUs and since the RSUs have some experiences of transit-time, they could predict the transmit-time for the future to enhance the accuracy and decrease the travel time. On the other hand, reservation and routing of vehicles consider a continues-time horizon. They proposed two algorithms that were about the routing of vehicles which can support various trade-offs between the quality of solution and complication of computational. Their main focus was on the Earliest Arrival Time Destination (EATD) problem. They tried to solve the problem of optimization by specifying the way that would permit the vehicle to arrive at the destination in a minimum time. So, they presented a new formulation for Mixed Integer Linear Program (MILP) that employed the travel time that is predicted to decide the EATD problem. It is an NP-complete problem and for this matter, the new formulation can solve some samples of that. Also, they developed the Time-Varying Multiple Linear Regression (TVMLR) prediction methods. Their heuristic proposed model is a developed version of Dijkstra's algorithm. But their method didn't consider the road that might be reserved, car accidents and some kind of situation that might lead to rescheduling the routing of vehicles.

In [33], they proposed a helpful system called Whale Optimization Algorithm (WOA) and a classifier named Multi-Support Vector Machine (MSVM) that can affect the improving performance of detecting traffic congestion. By using data normalization, the pre-processing operation is applied to the selected data with the aim of scaling a set of numerical data that are heterogeneous. When the data is pre-processed, for selecting the optimal attributes, WOA is used. The WOA's output can determine which properties were essential in the classification of the road traffic congestion and these improved properties were given as the input for MSVM in order to classify the road traffic congestion. The proposed method included four phases:

1. Data collection: they collected an unstructured or imbalanced dataset from a tunnel and consider the speed of vehicles, occupancy and flow rate of cars for five minutes.
2. Pre-processing: they did pre-process on unstructured data to convert them to a structured form.
3. Data optimization: by using WOA they selected optimal features and it was a new metaheuristic approach that emulated the Hump-back whales.
4. Classification: the feature values that are optimized, used as input for multi-class classifier and the MSVM for classifying the various type of traffic congestion phases.

The results showed that the proposed strategy can improve the accuracy of classification in traffic congestion.

In [34] a road traffic risk assessment model for the avoidance of road traffic accidents for smart vehicles was proposed. The proposed method was based on the Hidden Markov Model (HMM) for evaluating the probability of steering angle in different intervals and calculating the risk of road traffic in various areas. They assumed that all the vehicles are following the same traffic rules which leads to a steady driving process. They considered the steering angle and speed of vehicles as constant parameters to predict the position of vehicles in a certain period of time in the expected ways as illustrated in Fig. 2.4.

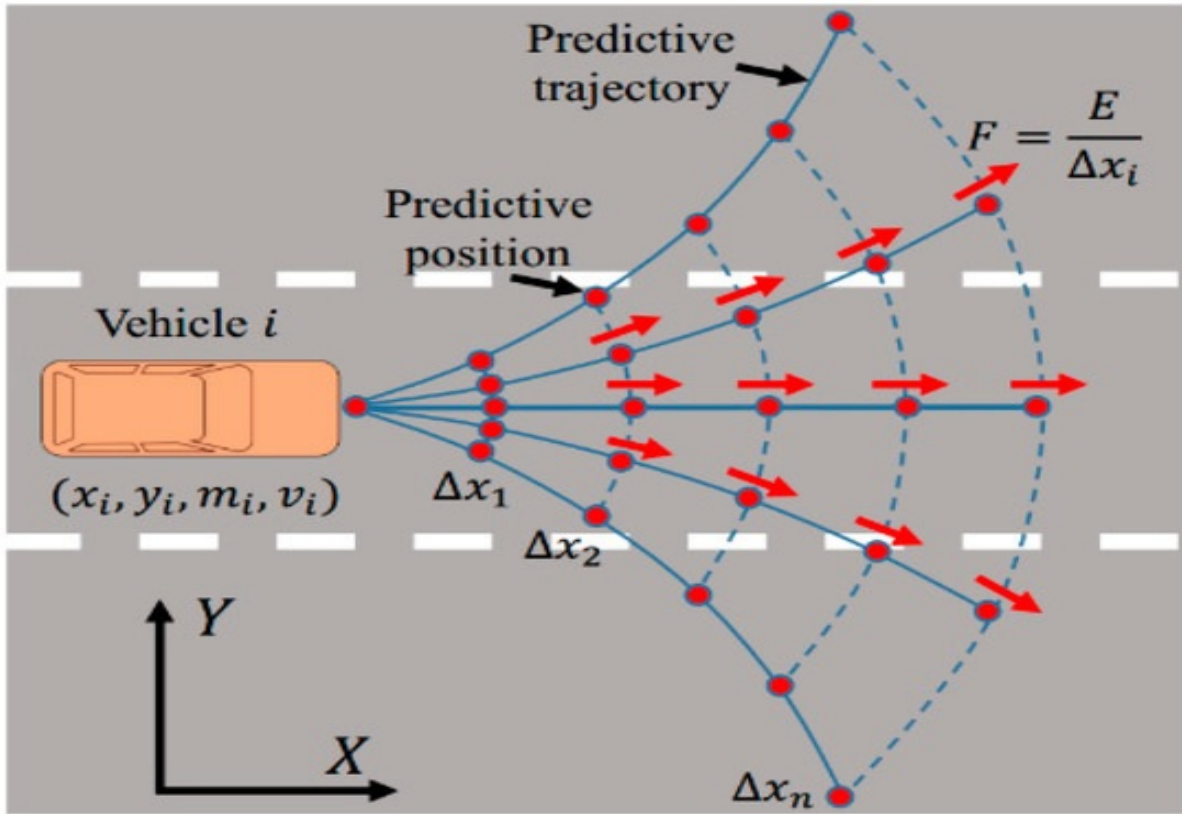


Figure 2.4 Predictive trajectories and positions diagram.

The predictive trajectory is shown in the blue curve and the predictive positions are shown in red points. They assumed that the predicted trajectories and the points are generated while the steering angle is stable. The identical force in all positions is shown in the red arrow. The proposed HMM method was able to evaluate the risk of road traffic on a larger scale.

The intelligent transportation system is one of the main issues due to the growing number of vehicles and in [35], the authors proposed a novel predictive road traffic management system

(PRTMS) which is based on VANET architecture. To estimate the density of traffic at intersections, they used a new scheme of communications based on some linear prediction (LP) algorithm that was improved. They used an IEEE802.11p based on Vehicle to Infrastructure (V2I) communications to collect the information of the journey and broadcast the control signals for applying the multi-junction traffic flow control. A Central Controller Unit (CCU) via re-routing the vehicles can reduce the level of congestion. They assumed that at each intersection RSUs are installed via Wide-Area Network (WAN) are connected to the CCU and the vehicles have On Board Unit (OBU) that sends their information about the trip to the RSUs. At the next step, all RSUs sent the information about traffic flow to the CCU and it will predict the traffic at each intersection and based on that re-route of the vehicles to decrease the traffic congestion will be done.

The proposed system implemented a two-layered control system based on V2I by using VANET. The first level of control was generated by the RSUs and the data was accumulated from vehicles at this level. At the second level, the CCU stored the data from the controller that was local then operated them and produced control signals that were transmitted to the OBUs through RSUs. The PRTMS proposed system has three tasks that CCU was responsible for them including collecting the information from the RSUs, predicting traffic flow and predicting congestion for re-route the vehicles. They used a Modified Linear Prediction (MLP) for the best development of traffic flow and for the aim of getting a pattern of traffic, the data was collected in real-time. The results showed that the proposed PRTMS model was able to increase the accuracy of prediction and decrease the time of trip by re-routing the vehicles.

In [36], they proposed a distributed predictive road traffic management for VANET that can predict the traffic of the roads and leads to a decrease in the traffic of the network without using the central server. Their results showed that waiting and driving times would have decreased. The construction of their model has two elements: vehicles and the Road-Side Unit and they investigated wireless communication for vehicles and RSUs and wired communications between the RSUs. Based on the messages that were transferred between vehicles and RSUs at a certain time, the congestion can be detected, and the optimal route will be suggested. Based on their method when the vehicles sent requests to the RSUs, the distribution of rerouted vehicles will be checked all the time. They predicted the traffic via distributed action developed on Linear Prediction (LP) model. They compared their model with a centralized system and a conventional system without prediction and the results showed that the proposed method can modify the distribution of vehicles and decrease the waiting time and the whole time of the journey.

In [37] a system for traffic prediction by using smart agents was proposed. BY predicting the traffic accurately, the traffic could be managed by setting the strategy for traffic lights to avoid congestion on the roads. For this purpose, they used artificial neural networks for estimating the speed of vehicles as a parameter for identifying road congestion. The main task of intelligent agents is collecting, processing and storing the collected road data to predict road traffic for preventive traffic measures including changing the time of traffic lights and alternative road suggestions. The architecture of the proposed model is shown in Fig. 2.5

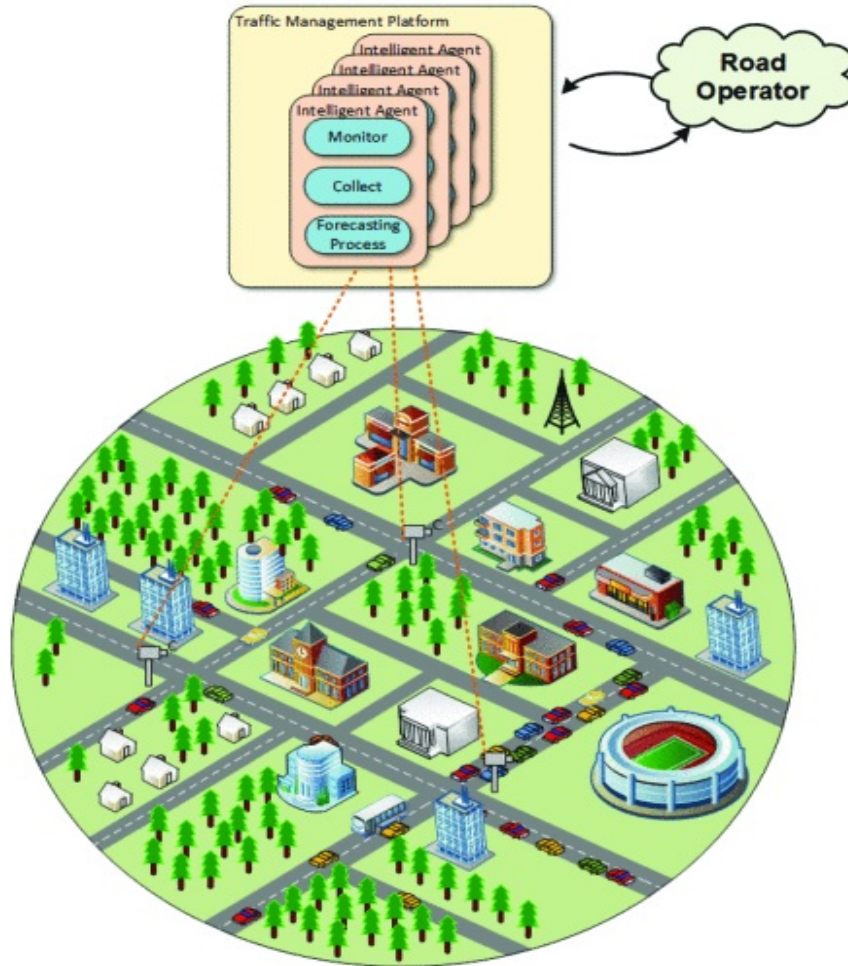


Figure 2.5 Graphical view of the proposed scheme with intelligent agents.

With the aim of validating the proposed model, they designed an ANN algorithm based on General Regression Neural Network (GRNN) [38] in predicting process. The simulation results showed that the proposed model can achieve high accuracy of almost 60% based on mean absolute percentage error (MAPE).

In [39], they proposed a smart traffic monitoring and controlling framework that can execute the communication of information. This traffic management paradigm employed a system of smart signaling with the algorithm for finding an optimum path for controlling congestion. Also, it has some modules for preventing car crashes, smart vehicle parking, and so on. Their proposed model was based on the technique of traffic management efficiently to address challenges, directly and indirectly, concerning traffic management. Their model has eight parts including Smart vehicles, Confidential Party, Governmental Authority, Carrefour Control Unit (CCU), Segment Control Unit, Master Control Unit, Law Enforcement Agencies and Emergency Management Departments. They considered three parameters for defining the warning time including driver response time, vehicle speed and distance between vehicles. They could predict the collision by using these parameters in mathematical models. They proposed algorithms for finding the optimal path and accident prevention. In their model, the emergency vehicles would broadcast a message to the CCUs, and they will clear the way for them.

In [40], three different machine learning algorithms such as long short-term memory (LSTM), autoregressive LSTM (AR-LSTM) and a convolutional neural network (CNN) have been implemented to predict traffic flow based on the got information from radar and meteorological sensors in essential weather conditions. The dataset that they used was including 170 million records considering 2019, 2020 and 2021 which gathered data from parking sensors and radars data. Their main target was to develop a regression prediction model to forecast traffic flow in Barra and Costa Nova considering environmental parameters such as the speed of vehicles. They considered 70% of the dataset for training, 20% for testing and 10% for validation. They used the MAE for evaluating the performance of LSTM, AR-LSTM and CNN. The evaluation results showed that CNN had better performance in terms of error and execution time.

A Smart Traffic Prediction and Congestion Avoidance System (s-TPCA) was proposed in [41] to identify the traffic situation with the aim of avoiding and predicting congestion. The proposed framework consisted of three phases including traffic identification, prediction and congestion avoidance which its main task was checking the fitness function to identify the traffic severity and predict the level of traffic in the future. The proposed system intends to predict traffic congestion based on the time arrivals of vehicles by using Poisson distribution as a prediction technique on two levels of traffic including low and moderate. Table 2.1 shows the major parameters for identifying the traffic. Also, the used primary variables in the proposed algorithm are shown in Table 2.2.

The used dataset and considered parameters showed that the proposed technique is able

Table 2.1 List of parameters

Parameter	Notation
Velocity	v
Average number of vehicles	m
Total number of vehicle	T_p
Arrival time	A_t
Arriving intervals	A_i
Maximum speed of the vehicle	Speed
Travel time of the vehicle	t
Length of the vehicle	L

to overcome the baselines in terms of scalability and sustainability. A probability genetic prediction model was proposed with the aim of traffic prediction and congestion avoidance. The simulation results showed the relation between observed and predicted frequency and decrease fuel consumption.

The Congestion-based Traffic Prediction Model (CTPM) with the aim of traffic prediction based on congestion propagation pattern has been proposed in [42]. The proposed model was able to integrate the congestion propagation data and it can be used in conjunction without the need to replace well-functioning methods. The evaluation results showed that their model was able to improve the accuracy in comparison with the LSTM and Extreme Gradient Boosting (XGB) models.

In [43] a precise time series predictive model for road traffic accident prediction based on autoregressive integrated moving average (ARIMA) and autoregressive integrated moving average with explanatory variables (ARIMAX) was proposed. They used four evaluation metrics including Bayesian information criteria (BIC), root mean square error, mean absolute percentage error and higher coefficient of determination (R-Squared) to evaluate the performance of the proposed method. The experimental results showed that the ARIMAX model overtook the ARIMA model. The proposed model showed that for having an accurate and resilient predictive model, considering more parameters such as incorporating human, environmental related factors (road factors) and vehicles could be effective.

Table 2.2 List of parameters

Notation	Parameter
T_p	Total population
A_i	Arriving intervals
A_t	Arrival time
G_c	Genetic criteria
TC	Total cost
μ_{ij}	Membership function
fn_i	Fitness value
G_cM	Mutation results
y^i	Traffic pattern

One of the practical solutions for avoiding traffic congestion is predicting future traffic in intelligent transportation systems. An ensemble learning method for predicting hourly traffic flow was proposed in [44]. So much research has been done on two types of traffic flow including short-term and mid-term future and different parameters are considered in each of them. For instance in mid-term traffic prediction, the effective parameters are holidays, weather conditions and cyclical temporal features while in short-term prediction models not only the mentioned parameters are considered but also the past 3 to 8 hours of traffic flow have been used.

They implemented three ML algorithms including KNN, RF and Recurrent Neural Network (RNN) with the aim of predicting hourly traffic flow. Then the Extreme Gradient Boosting (XGboost) model has been implemented as a stacking super model which used the output of three used ML in the previous step (KNN, RF and RNN) as an input. In the final step, the bagged averaging algorithm was utilized for predicting the final output. The traffic data of Tehran-Saveh in both directions has been used as a traffic dataset. The dataset that they used was related to a rural road that has nonroutine trips which makes it more complicated than the urban road traffic data. The evaluation results based on RMSE and MAPE evaluation metrics showed that the XGBoost can obtain the lowest prediction error in both short-term and mid-term traffic prediction models. The XGBoost model was able to have more accurate

traffic flow prediction in the first and fourth quartiles of traffic data in comparison with other baselines.

According to the all above mentioned road traffic predictions, we can conclude that despite of developing in ML and DL algorithm, there is still some challenges such as the security of data, execution time and accuracy that need to be addressed in various algorithms. From another point of view, we can say that the main reasons for road traffic are the growing number of vehicles, the weather condition, road congestion, traffic lights, intersections, etc which can affect the network traffic as well.

2.3 Traffic Prediction Based on Network Information

Another dynamic and distributed strategy for congestion control in VANET is proposed in [5]. This strategy is closed-loop and it has two parts: 1) Congestion Detection: the congestion has been detected by measuring the level of channel usage and 2) Congestion Control: they used Multi-Objective Tabu (MOTabu) algorithm for controlling the congestion.

They tried to improve the reliability and safety of VANET and to getting this aim, they used a Meta-heuristic algorithm (Tabu Search) that has a solution for NP-hard problems. The Tabu Search algorithm can get the best values for transmission rate and range while jitter and delay are in the minimized situation and the optimized value would transfer the data over channels of communication. In this work, they focused on decreasing the delay and jitter and transfer the messages in a shorter time. They developed a multi-objective Tabu Search and they considered delay and jitter as objectives. In this work for defining an initial solution, they used the current state method. The Tabu Search would create a new solution as a neighborhood set and for getting the new solution, they applied a memory mechanism (short-term, mid-term and long-term).

For traffic propagation, the authors applied the TwoRayGround model in highway scenarios and the Nakagami propagation model in urban scenarios. Because the diversity in the urban scenario is more than in the Highway scenario, the performance metric in that one has a larger quantity. Implementation of their strategy has some limitations in real VANET. All the vehicles should have an On-Board Unit (OBU) for calculating the time, the vehicle should estimate the level of channel usage for congestion detection and have tuned the transmission rate and range. Also, they should have GPS to estimate the delay and jitter. The results showed that the average delay, the number of packet losses, the packet loss ratio and the number of retransmissions decreased, and the average throughput improved. The MOTabu strategy is a scalable distributed (applied in each vehicle) strategy that considers

both parameter delay and jitter.

A new architecture for decreasing traffic congestion and managing urban traffic was proposed in [45]. Their proposed model has four categories for managing the urban traffic with the 5G networks, software-defined network (SDN), mobile edge computing (MEC) and convergence of VANET. The tiers of the new model are as follows.

1. Environment Sensing Layer: there were some sensors with high resolution that were installed, and the vehicles reported their status.
2. Communication layer: for addressing the problem of communication, they considered 5G and SDN that can support the scalability and flexibility of the network.
3. Mobile Edge Computing Server Layer: they added this layer to their model for awareness of traffic flow and to increase the rescue for accidents. The focus in this layer is a critical task in real-time. The MEC layer is installed on the Base Stations of the 5G. The framework has three parts: MEC hosting infrastructure, MEC application platform, and application layer. Also, they added four services to their architecture to simplify the management of urban traffic. The first one was a vehicle localization service that was installed on the MEC server for supporting a trustworthy service. The second service was a traffic data process that is installed at the MEC server to filter the extra data and sort them for a short time, and the Remote Core Cloud Server (RCCS) layer focuses on data synchronizing, prefetching and storage. The third one is the Traffic Prediction service which has been installed at the MEC server for the propagation of the messages to the drivers based on predicted information and on the RCCS layer due to predict the traffic flow by using a deep learning algorithm and the fourth service is Traffic Light Control that deployed just on the MEC service due to the control of the traffic lights at the real-time.
4. Remote Core Cloud Server Layer: its framework was like MEC servers and the focus in this layer was on storage and analysis of the big data.

The SDN-based heterogeneous networks can provide high-bandwidth and low latency and the MEC server can improve the rapid response. The rescue system can be improved by applying three parameters: 1) Remote video diagnosis and initial assessment 2) Early-warning and no-go area 3) a Green rescue lane for emergency vehicles. The proposed novel architecture may diminish traffic congestion and enhance the capability to manage urban traffic.

A Machine Learning Congestion Control algorithm (ML-CC) proposed by N. Taherkhani S. Pierre in [13] for solving the problem of congestion at intersections. They proposed a central-

ized and localized data congestion control by using Roadside Units (RSUs) at intersections that will send warning messages to the vehicles that are stopped before the red traffic light (congestion area) for avoiding collisions. In this work, they considered some communication parameters like transmission range, transmission rate, contention window size and AIFS that are adjusted based on minimum transferring delay for the messages of every single cluster. For doing this adjustment they chose the proper value of these parameters that the DSRC standard has defined their value.

Their strategy has three parts to detect the congestion, cluster the messages and control the congestion of data. To detect the congestion that happened on the channel, they measured the level of channel usage. The strategies for congestion control have two groups of the solution. The first one is an opened-loop that prevents the congestion before happening and the second one is a closed-loop solution that after detection of the congestion, would be controlled. The strategy that they proposed was a closed-loop and hybrid strategy and it has three parts. The first unit is for detecting the congestion, the second unit is to control the data and the third one controls the congestion.

In this work they used unsupervised machine learning algorithms (K-means algorithm) for clustering a large amount of data caused by the mixture of messages over communications like V2V and V2I and the data that are unlabeled in VANETs and this type of algorithm doesn't need to a training data set. By using the K-means algorithm, they filtered, gathered and clustered the messages. They considered an urban scenario and the results showed that their method can enhance the performance of VANET and reduce the packet loss ratio, delay and the possibility of collision. ML-CC strategy could increase the packet delivery ratio and the average throughput. Improvement in the accuracy of traffic flow prediction can affect intelligent traffic management and reduce congestion.

In [2], the authors tried to develop and optimize traffic learning by using the Taguchi method via layer-by-layer features categorized with the unsupervised algorithm. They proposed an optimized structure for traffic flow prediction that is trained with a stacked autoencoder (SAE) and Levenberg–Marquardt (LM) algorithm to implement a short-term traffic prediction. Their method was the stacked autoencoder Levenberg–Marquardt (SAE-LM) model with the purpose of increasing the precision of short-term traffic flow. The proposed model was a kind of deep architecture of the neural network which used the Taguchi method for learning the features of traffic flow.

To implement and evaluate their method, they used real collected data. They implemented some constraints to the hidden layers (they considered five hidden layers) and they assumed that the neurons are inactive most of the time. They used the LM algorithm for training the

input that the last autoencoder generated it.

For assessing the performance of the proposed method, they used two performance values including the MAPE and the variance absolute percentage error (VAPE). The lower value of them demonstrates better performance. The results showed that the proposed model is more accurate at congestion prediction and has less computational time than the previous methods. Based on the comparison that they did between their method and the others, they found that the percentage of error value in the proposed SAE-LM model, most of the time is almost 0%. The data on traffic flow might be irregular because of some unpredictable situations and the proposed model was practical in a deal with this problem.

The authors in [46] proposed a new deep-learning-based internet traffic flow forecasting method called the stacked denoising autoencoder prediction model(SDAPM). For learning general internet traffic flow, they used a Stack Denoising Autoencoder (SDA) forecasting model. The SDAPM includes two parts: 1) The data preprocessing algorithm that was used for processing initial sample data to the format that was expected and 2) SDAPM Generation Algorithm that was used for training the SDAPM. They also added an isotropic Gaussian noise to the sample of data for providing an improvement inability of the generalization and the ability of the model's adaption.

They set some parameters for SDAPM and then initialized the SDAPM that is consist of four parts: 1) Define the dimension of input and label dimension, 2) Slot the date of sample, 3) Put the number for the layers of model and neurons per layer and 4) For each layer, set binary masking noise. Also, in the third step they did pre-training for the hidden layers of SDAPM that has three steps: 1) Assume the first layer as the input layer, 2) Every hidden layer has been trained like an autoencoder for getting the output of the front layer as an input and 3) Define a value of weight for each hidden layer and constant vector. For the last step they fine-tuned each hidden layer and constant vector to minimize the loss in the validation of value and it included three parts: 1) the Prediction layer would compute the value of prediction traffic flow, 2) Compute the difference between the value of prediction and label value for minimizing the average cost value and 3) Stock the present parameters in SDAPM and compute the validation loss value for finding the best parameter set of each hidden and prediction layer. For training the model they used the train data set and validation set also, to assess the performance of the model, they used a test set. The results showed that the proposed model has a good effect on the improvement of traffic flow prediction.

In [47] a novel smart traffic prediction method was proposed based on graph convolution neural network (GCN)-GRU called AGG. The GCN algorithm was used for capturing the spatial features of traffic and the GRU was utilized with the aim of capturing the temporal

features of traffic. Also, they used an attention mechanism for capturing the importance of various features. The proposed model not only was able to capture spatial-temporal features but also was able to collect temporal global change information. They implemented the proposed algorithm on a real dataset and compared its performance with various algorithms such as Historical average (HA), ARIMA, SVR and GRU. The simulation results showed that GCN-GRU algorithms got the best prediction effect based on various evaluation metrics.

In [48], the authors designed a smart agent that absorbs the policy of traffic light control. They focused on using intelligent agents just as controllers of the traffic light. They considered some aspects of traffic light control and simulation. the first one is the action that might happen from the side of vehicles and the second one is the synchronization in arterials that it is like a greenway and for optimization of that, we need to find the bandwidth. For optimization, the target considers two parameters that are critical, low latency and high throughput. In this work, the authors tried to use an intelligent agent as a traffic light controller. Their duty is optimization the traffic flow via learning a control policy. Intelligent agents for making the decision need a formal framework. Marcove Decision Process is one of the common ways used.

To design their controllers, they implemented a traffic light simulation module in a vehicle simulator that is homemade. For simulation, they considered Urban Traffic Control Simulator and Agent Learning. For learning the agent, they regarded some variables like the number of vehicles that are on the horizontal road, the number of vehicles on the vertical road, the present condition of the light and the number of decision phases. For optimization of changing the light and delay of the vehicles, the agent must learn and for this purpose, it used a policy gradient reinforcement learning algorithm OLPOMDP. As a result, by using the traffic light control agent, they learned the control policy. The results showed that their learned policy not only executed as well as pre-timed near-optimal controllers but also, they had better performance in dissimilar traffic situations.

A method for improving mobility management in VANET is proposed in [11]. They proposed a neighbor selection that is based on the Q-learning technique as a Neighbor Predictive Adaptive Handoff (NPAH) algorithm for guaranteeing solid communication irrespective time of application service. It can determine the communications links that are not strong in the service. Handoff and transmission have some challenges and for addressing them they consider two independent part for NPAH neighbor discovery that is based on assessing the consumption of bandwidth and accessibility of links and neighbor selection that use Q-learning for cost factors and usage of storage.

They selected some neighbors based on minimum cost and distance by using the Q-learning

technique. Their results showed that their model could decrease packet loss and outage times and improve the factor of delivery. Mobility management has some challenges like localization, handoff, and management of communication that they should be addressed.

In their network model, vehicles sent their request to their neighbors and Base Stations (BSs) for information. The BS for accessibility, storing and recovery of the information, is connected to the cloud network and can update the information that vehicles need. The NPAH is a distance-based neighbor discovery based on Q-learning performance. The vehicles that are participating in the communication process will operate the distance-based neighbor discovery and it leads to selecting an optimal route for vehicles while re-routing is performed and before the handoff. It has two stages the first one should provide the reliability of service discovery and the second one should guarantee minimal delay and outage at the handoff time. For getting service from BS, the vehicles trusted their neighbors. Based on the storage that neighbors consumed if it is low, the amount of handoff request service would be high and contrariwise.

They evaluate the NPAH model by using a Network Simulator in a topology of urban mobility and they used some metrics like delivery factor, consumption of storage, packet loss and outage time. They found that these parameters are affected by the number of vehicles. They compared the performance of the proposed method with the other methods that already exist like predictive routing based on a hidden Markov model (PRHMM) [49] and Bursty Multi-node Handover scheme for mobile internet using the partially Distributed Mobility Management (BMH-DMM) [50] and the results showed that their model can improve the factors for congestion prediction in comparison with the other methods.

Analyzing various existing methods for traffic prediction has been done in [51]. They implemented plenty of DL and ML algorithms such as Statistical Auto-Regressive Integration and Moving Average (SARIMA), Poisson's regression, Decision Tree Regressor (DTR), Multi-Layer Perceptron (MLP), Linear Regression (LR) and LSTM. They proposed an architecture based on CNN and LSTM named Convo-LSTM model. They implemented the proposed model on wireless mesh networks which are easy to adapt and configure. The wireless mesh networks based on mesh topology are shown in Fig. 2.6

the proposed system has been implemented on a multivariate dataset and They evaluated the proposed Convo-LSTM based on five various intervals including hourly, daily, weekly, monthly, and yearly. The simulation results revealed that the proposed system was able to overcome other compared algorithms.

In [52], the authors presented a novel method for optimizing the parameters of a hierarchy of fuzzy-rule-based systems (FRBSs). It is a combination of a Genetic Algorithm (GA)

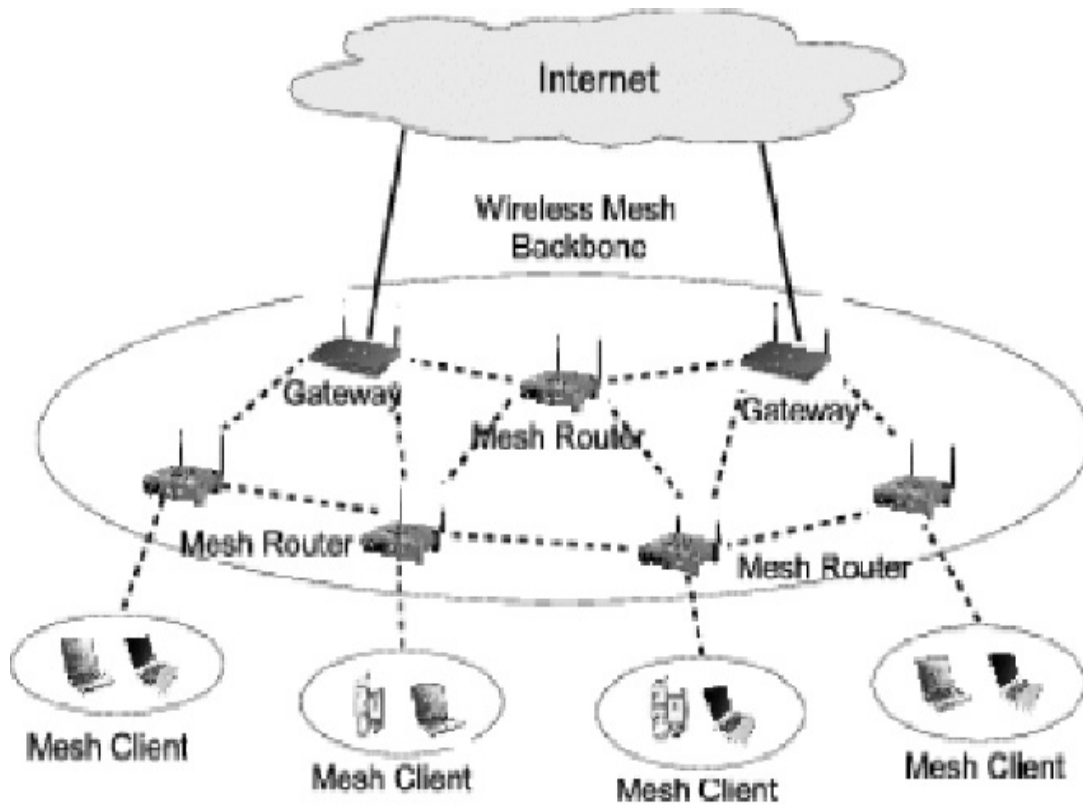


Figure 2.6 A wireless mesh network.

and the Cross-Entropy (CE) method that in this paper they called GACE, and they used this method to predict congestion. Their goal from this combination was to get a synergy between exploitation and exploration intending to improve the parameters of the system. They proposed a way of using the Hierarchical Fuzzy Rule-Based System (HFRBS) due to traffic prediction on a freeway. To optimize the units in the hierarchy, they combined a Cross Entropy (CE) algorithm with the Genetic Algorithm (GA).

The proposed model divided the population into two subpopulations. GA is used in one of them while CE is executed in the other one. Then both are met, and they exchange the current population. Based on the system's structure there are three types of HFRBS: Serial HFRBS, Parallel HFRBS and Hybrid HFRBS that in the proposed method they focused on Parallel types (PHFRBS).

In this type, the structure has some layers and the input of each layer is the output of the previous layer. They proposed a PHFRBS optimized with a hybrid GA and CE(GACE) that is used for predicting short-term traffic congestion. The dataset that they used contains a large number of free flow examples that leads to having an unbalanced dataset and because

of this issue they used an imbalanced ratio (IR) and for balancing the dataset, they used the K-nearest method. The complete data set was used for testing the results and the reduced data sets were used as a training dataset for GACE. The results showed that the algorithm can forecast short-term congestion at a point with a lower error.

VANET is a kind of solution for decreasing the conflict between the constraint volume of the cellular network and big data gathering. Heterogeneous networks are consisting of cellular networks and VANET and choosing an efficient network is an important issue for having Quality of Service for vehicles that in [6] they proposed an intelligent network that is approved by big data. They focused on solving the big data problem and the main target of this paper is to implement an intelligent network access system based on traffic status, user requirements and conditions of the network.

They proposed architecture for resourcing the real-time service. It has some layers including the access layer where vehicles are like sensors for the selection of traffic information and the aggregation layer which sensing records that are uploaded over the network and also collected by the center controller for analysis of the traffic data and the center controller suggests network selecting based on some traffic parameters like delay, bandwidth, cost, and traffic distribution, and the last layer is the application layer that they installed an application on the terminal in the vehicle that its task is executing the network examine the handover.

They considered data rate, delay, and cost as factors that can affect QoS and due to them, they analyzed the utility function. They considered some assumptions in designing their model. They regarded both cellular networks and VANET, but they didn't notice V2V communication. They attended to critical factors like traffic density and influencing factors like bandwidth, delay and cost also they considered two types of real-time services that were video and voice communication. They suggested the Analytic Hierarchy Process (AHP) that could be a suitable algorithm for having an optimum network while there are multi-criteria decisions and various conditions. They used a pair-wise comparison matrix for executing priority among parameters. The result of the simulation showed that the proposed algorithm can support high network usage and can assure the QoS of vehicles.

Based on reviewing the previous research that has been done in network traffic prediction, we can conclude that network traffic is happening generally by dropping the speed of senders, decreasing the packets, losing communication and so on which can lead to reducing the throughput and efficiency of the network and increasing the network traffic flow. Almost all past research has investigated the issue of road and network traffic separately while these two issues are not irrelevant and can affect each other.

2.4 Traffic Prediction and Control Using Intelligent Methods

In [53], they investigate three metrics for the performance of VANET including capacity which means the level of traffic that in a specific time the network is able to manage QoS, insurance which is a group of provisions like rate of error, delay, priority and so on and for the safety constraints they considered the redundancies that Intelligent Processing Technology (IPT) can suggest a process in a minimum time and keep the exactness simultaneously. By using IPT in VANET some applications like QoS Guarantee, traffic optimization, and network security can be reinforced. They classified the intelligent Processing Technologies into four categories: 1) machine learning technology that can lead to learning the behavior of vehicles in a situation that they haven't experienced before 2) fuzzy logic technology that can emulate the pattern of behavior of human decision and its practical 3) bionic optimization Technology and 4) data mining, game theory, and technology.

In [54] a new method called Flow2graph for predicting network traffic flow in the Industrial Internet of Things (IIoT) was proposed. The proposed method was a time-varying-aware network traffic prediction which has been designed based on learning dynamic shapelets using the sense of the next change in network traffic flow. Each time series could be divided into many segments and all the fragments can be selected through the learning process. This mechanism is shown in Fig. 2.7

They adopted a graph isomorphism network (GIN) with the aim of getting the time-varying features of the traffic sequence. The robustness, accuracy and adaptability of the proposed model, have been proved by comparing it with a series of mostly used models in traffic prediction.

Routing is one of the impressive factors that can affect traffic flow. An intelligent Naïve Bayesian for estimation of the traffic flow to a stabled cluster form in VANET (ANTSC) was proposed in [55]. The proposed model can improve the routing by knowing the traffic flow by using some parameters like speed, direction, level of connectivity and distance between nodes and their neighbors by using intelligent techniques. Their technique is based on finding some vehicles as cluster heads with the longest lifetime and it will consider the information that will be transferred between vehicles like speed, direction, level of connectivity and so on.

They proposed an algorithm for clustering that is intelligent with the aim of enhancing the lifetime of the cluster. They considered the urban scenario and focused on traffic flow at intersections. They chose a head for each cluster that has more traffic flow due to the lifetime and consistency of the cluster. They assumed that each vehicle is equipped with GPS and grouped the vehicles that have different traffic flows in one cluster for more stability and the

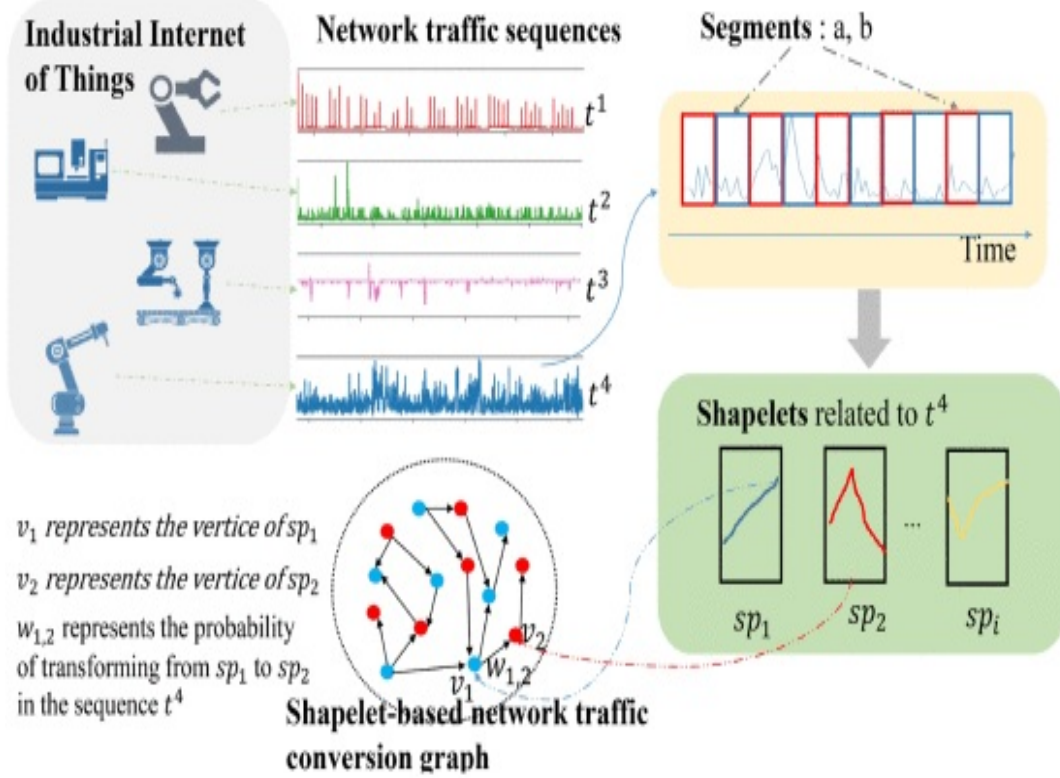


Figure 2.7 Relation between the notations.

size of the cluster depends on the density of vehicles. They used one kind of machine learning algorithm (Naïve Bayes) in the Bayesian network and the results showed that by using this method the lifetime of the cluster head can be improved caused by choosing the cluster head from the line that has the heaviest traffic flow. In the proposed technique, by increasing the range of transmission they could reduce the number of clusters per vehicle. They showed that their method can increase the stability and lifetime of the clusters.

Recently, Deep Neural Networks (DNNs) can forecast the traffic flow with big data but there are some challenges that in [1] the authors proposed a model to cover these problems. They proposed a model for predicting the traffic flow that is based on DNN (DNN-BTF) to enhance the exactness of the prediction. For validation of their model, they used data from the open-access database PeMS, on a long-term horizon forecast task. The proposed model presented a forecast model with reliable exactness and it analyzed the internal mechanism of DNN on data of traffic flow. The results showed that deep learning based on the prediction of traffic flow still needs more studies.

Static traffic signals and dynamic traffic flow are two main factors that can lead to traffic

congestion. The authors in [56] proposed a Jordan Sequential Network for more accuracy and prediction the road traffic by using an Artificial Neural Network to learn the old pattern and predict future traffic. They used dynamic and non-stationary data that are collected from various detectors. For accompanying traffic signals. They needed short-term forecasting and they used traffic volume as a parameter for prediction. The major focus of this work is on the accuracy of artificial neural networks for different types of structures and finding the best of them in terms of accuracy. Jordan's Network has four phases including an input layer, a hidden layer, an output layer and a context layer. The context layer has a connection with itself and its input is the output of the output layer. The input layer and context layer provided the input of the hidden layer and in the proposed model they got the best results when there is just one hidden layer while the number of neurons that exist in the hidden layer is the same as the square of the neurons which exist in the input layer. For the training data set, they predicted the output by taking four inputs and then the predicted output is stocked in the context layer.

Because of using the sigmoid activation function, they should normalize the values for using them as the input so the output value would be normalized and after the process, they should de-normalize it again. The prediction has two processes including feed-forward and backpropagation that after feed-forward, they get output and at this step they compared the predicted value with the expected value. The back-propagate process will be done if the error parameter is bigger than the determined value for re-adjusting the weight of the NN for achieving the highest accuracy.

The results showed that the accuracy and the speed of the proposed model are better than the other methods. For accuracy they considered three parameters: 1) learning rate: they kept the learning rate at the middle value of the range 2) structure and 3) the number of instances. The results showed that the proposed model could cover the targets with accuracy.

A new method based on enhanced deep reinforcement learning (EDRL) algorithm with the aim of improving network traffic analysis and prediction has been proposed in [57]. The mechanism of the proposed model is based on network traffic for admission control, anomaly detection and congestion control while the traffic data is big and built upon the deep architecture model-based internet traffic flow prediction. they compared the performance of the proposed model with various algorithms such as neural networks, principal component analysis (PCA), linear time series models, statistical models, linear regression (LR), support vector machine (SVM) and LSTM. The relation between methods and inputs, alongside the problems and policy for various ML and DL algorithms, are shown in Table 2.3

The experimental results showed that the proposed algorithm performed well in terms of

Table 2.3 Data input and policy relationship for types of ML and DL algorithms

Method	Input data Know Answer	Input data Policy/Problem
Supervised learning	Learned output with supervision	Learning reward-based output with supervision
Reinforcement learning method	Maximize reward-based output	Feedback trained maximized reward-based output
Deep learning method	Deep learning-based output	Deep learning and feedback trained-based output
Deep reinforcement learning method	Deep-based maximize reward-based output	Deep and feedback trained-based maximize reward-based output
Enhanced deep reinforcement learning method	Accurate deep-based maximize reward-based output	Accurate output, based on deep and feedback trained maximize reward

mean accuracy, mean false positive, mean precision and mean false negative not only in numerical but also in graphical results as well.

In [58], the authors proposed a method to forecast the degree of congestion for optical networks by using Back Propagation Artificial Neural Network (BP-ANN) with the aim of predicting and monitoring the Congestion Degree (CD) in the networks which are optical. The BP-ANN is including input, hidden and an output layer that in a suitable situation of backpropagation, the signal will be transmitted from the input layer to the output layer. Until the error of BP-ANN output is decreased to the desired amount, the back-propagation process will be continued. They considered four factors for each service including bandwidth, holding time, source-destination and node pair. So, there would be 4000 input layer neurons. Their method is practical for big data and they could improve the accuracy of prediction by improving the quality and number of data and the exactness of the network model.

A network traffic hybrid prediction model is proposed in [59]. The new model that the authors proposed is a combination of two grey models and Elman neural networks. When

the fluctuations in the network are low, the grey model has good accuracy for prediction and can forecast the general process of network traffic and the Elman neural network model that is a Back Propagation neural network that has two layers with feedback structure while it can predict well for time series that are nonlinear. The proposed model has the advantages of both the grey and Elman NN model. The simulation results showed that their model has more accuracy and fewer prediction errors. In this model, they optimized the weight coefficients of the hybrid prediction model by improving the Harmony Search (HS) algorithm which leads to having the best results for prediction. For the simulation, they used real data that are collected from an actual network of telecommunication and the results showed that the convergence speed is higher, and the fitness value is better when they used the improved Harmony Search algorithm.

In [60], the authors used the neural network and theory of chaotic time series as the base for proposing a model for prediction based on the Radial basis function (RBF) neural network optimized by using an enhanced gravitation search algorithm. They proposed an improved gravitation search algorithm for neural network (IGSARBFNN) that is derived from the non-linear time series and they analyzed the network traffic prediction features in the form of in-depth and RBF neural network parameter improvement. By using a linear function, they improved the gravity Eco-efficiency and by presenting the memory and group communication of the particle swarm optimization (PSO) algorithm, they enhanced the speed selection formula. For monitoring the traffic flow of the network, they used the SNMP protocol data collection. The results showed that when they have to change the rule of network traffic, by using the proposed method they can have a highly accurate reflection.

An algorithm based on multitask learning with the aim of network traffic prediction considering spatial and temporal network features was proposed in [61]. Based on the evolution of the industrial Internet of Things (IIoT), software-defined networks (SDNs) algorithms can provide flexible access to the applications. They combined a convolutional neural network (CNN) and long short-term memory (LSTM) network and utilized the CNN and LSTM algorithms in order to extract the spatio-temporal features. Also, for having more accurate predictions they implemented a traffic prediction algorithm based on multitask learning (MTL). MTL is one type of ML algorithm that can improve fitting problems by sharing the hard parameters. The comparison results showed that the proposed approach has less deviation than the baseline algorithms.

Enhancement of short-term traffic prediction models based on unidirectional and bidirectional LSTM was proposed in [62]. The unidirectional LSTM (Uni-LSTM) is able to identify a wider sequence of traffic in time series data with high performance. The authors implemented an

extended version of Uni-LSTM called bidirectional LSTM (BiLSTM) which the input data will be trained in both forward and backward directions. Their main target was using AI techniques based on Uni-LSTM and BiLSTM models with the aim of predicting the speed and traffic flow. The proposed model was implemented on a real historical dataset extracted from sensors installed on sidewalks on three different freeways in Australia. Also, the authors explored the effect of adding extra layers in the training process on improving the accuracy of speed and traffic flow prediction. They used 60% of the dataset for training and 40% of it for testing and the mean absolute percentage error (MAPE) metric was used for performance evaluation. The MAPE will calculate the average absolute difference between predicted and true values. The experimental results showed that both Uni-LSTM and BiLSTM are able to obtain high prediction results. However, the BiLSTM has better performance in comparison with Uni-LSTM achieving 92.6% accuracy in one of the three freeways up to 60 minutes into the future. Also, they implemented a combined model including LSTM and BiLSTM as one model with the aim of predicting 15 minutes in the future. The simulation result showed that the 4-layered BiLSTM overcome the other models in terms of speed and traffic flow prediction in two different datasets.

In [63] a cyber-physical-social system (CPSS-UPTDN) which is an urban public traffic dynamic network proposed as a general framework for developed public transportation systems based on big data by using AI methods. They tried to make a simulation system that was matched with the artificial system with the aim of comparing their results. The proposed method was able to mix the multi-source big data and AI methods. The experimental results showed the effective management of the proposed model in urban public transportation.

In [64] the authors reviewed the different techniques for predicting traffic actors' behavior in the next action for intelligent vehicles. They summarized different AI techniques which are used for behavior prediction for smart vehicles. Fig. 2.8 shows the different AI techniques which are used for traffic behavior prediction.

Prediction is a real challenge in traffic or even the behavior of smart vehicles and most research papers in 2019, 2020 and 2021 proposed different methods to solve it. Based on the development of AI techniques, most researchers tried to solve complex problems by implementing AI techniques. Fig. 2.9 shows the chronology of various AI techniques which are used in this field.

For solving the problem in traffic actors prediction for smart vehicles, different probabilistic models based on recurrent neural network (RNN) and generative adversarial network (GAN) have been used which are shown in Fig. 2.10

Some of the AI techniques are explained as follows.

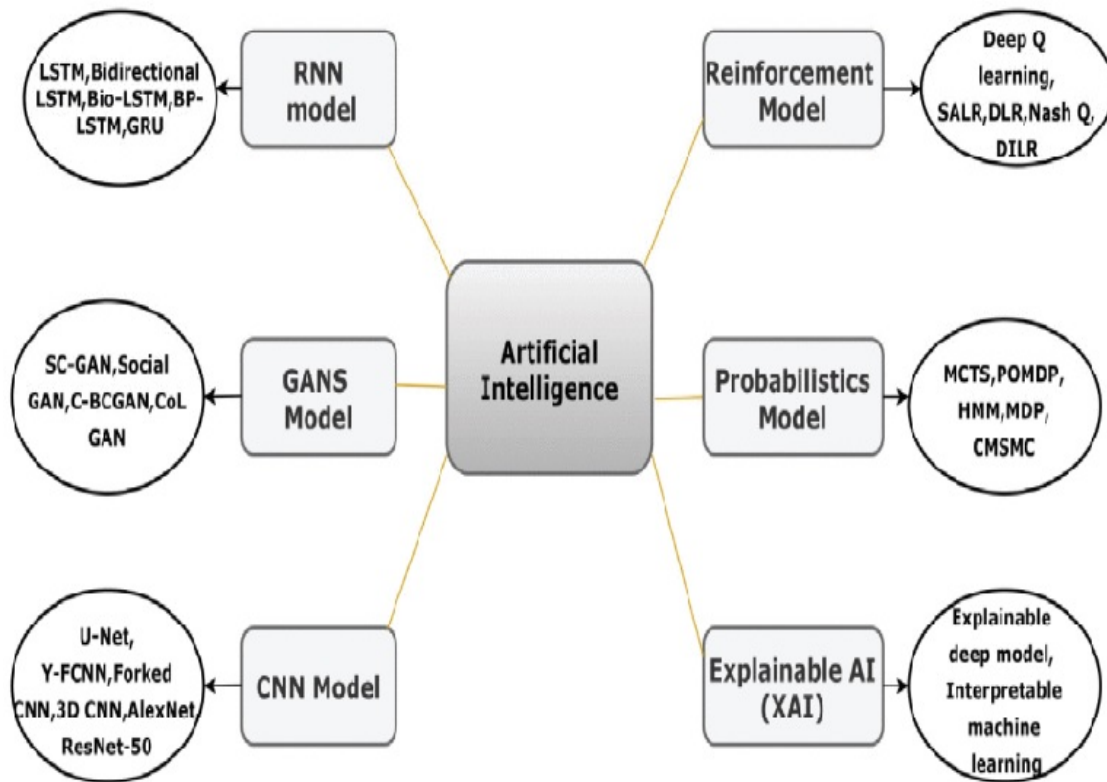


Figure 2.8 Behavior prediction using different AI-based models.

- The Probabilistic Model (PM) is a way to link the symbolic and metric action illustrations while a reasonable way for dealing with semantic formulation has been offered. This method is able to achieve a symbolic situation to predict the future which is the base of probabilistic decision-making theory.
- The Convolutional Neural Network (CNN) Model is including convolution layers that wrap a filter over the data by using learnable weights and the layers which are fully connected and are able to map the input to the required output.
- The Generative Adversarial Networks (GANs) Model was introduced in 2014 by Goodfellow in [65]. It is an unsupervised learning algorithm based on the mini-max concept, which is a competition between the efficiency of discriminator and generator networks. However, the GAN algorithm is untrustworthy and hard to implement since the Jensen-Shannon (JS) divergence has been used as its loss function.
- Reinforcement learning (RL) is a resilient learning system that is able to learn complex policies when the environment has a high dimension. It is based on an interaction

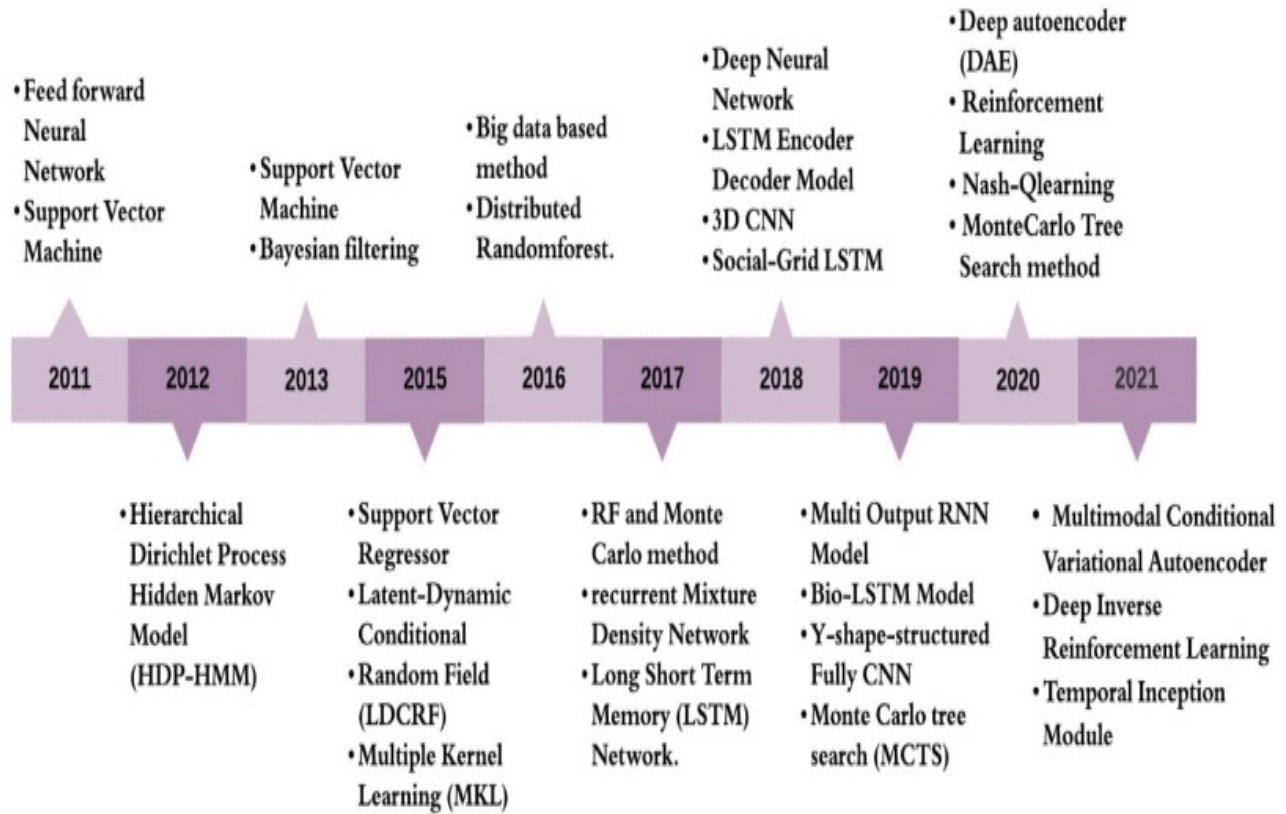


Figure 2.9 Timeline diagram of different artificial intelligence techniques.

between smart agents and the environment with the aim of improving output. RL algorithms are used in different fields such as robotics, video games and intelligent vehicles.

- One critical dimension of learning algorithms is The ‘explainability’ that is able to explain the internal structure of black-box algorithms. Explainable Artificial Intelligence (XAI) is a bunch of methods with the aim of converting black-box AI algorithms to white-box ones.

They reviewed different AI techniques which are used in the last decent years and identified the challenges and their effectiveness. Also, they believed that the availability of real-world datasets would be a serious challenge in research about behavior prediction problems.

The volume and dispensation of traffic flow can be explained by the Traffic Matrix (TM). An accurate TM prediction strategy can lead to handling essential tasks such as short-time traffic scheduling or re-routing, traffic accounting, network design, anomaly detection, etc. In [66] three various DL algorithms including LSTM, BiLSTM and GRU have been

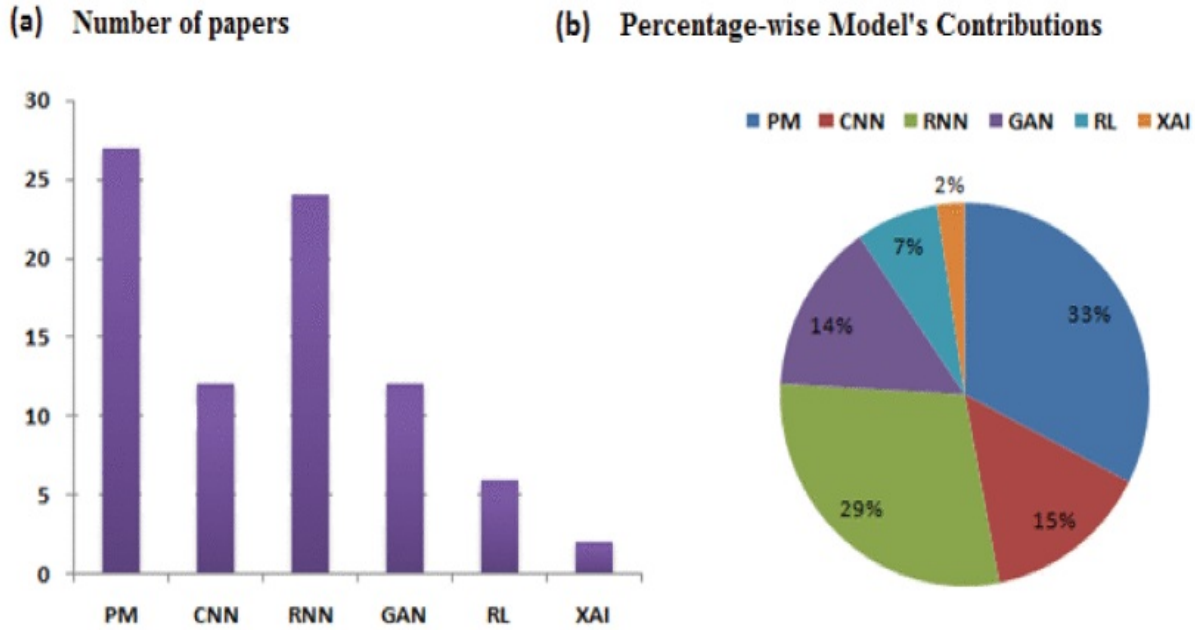


Figure 2.10 AI-based model's contributions based on: (a) number of papers and (b) percentage-wise.

implemented for TM prediction mechanisms of Software-defined Network (SDN) network. The proposed application had two modules including TM Observation and TM Prediction. The current information of the network will be shown by the TM observation, and the future's information is shown by the TM prediction. The three mentioned DL algorithms have been implemented on the TM prediction module and two different datasets have been used. They evaluated the performance of the proposed model by using RMSE as an evaluation metric. The experimental results revealed that the proposed model is able to generate precise traffic information for the future and the GRU acted as the most performed algorithm for complex modern networks.

In [67] the authors proposed a model to predict the future road accident based on previous data. They implemented various supervised ML algorithms such as logistic regression (LR), decision tree (DT), random forest (RF) and naive Bayes (NB) also some unsupervised ML algorithms including Density-based spatial clustering of applications with noise (DBSCAN) and hierarchical clustering. A rule generation model for identifying the effective parameters in a severe accident was proposed based on ML approaches. They used a real dataset including 28,102 reports about traffic accidents from 2016 to 2019 including different data sources such as road, weather, victim, driver and vehicle information alongside many other parameters. They focused on the RF algorithm which had the best result in accident prediction in terms

of false positive rate (FPR), specificity, sensitivity and accuracy. The evaluation results showed that the proposed model was effective to find the most suitable pattern in a fatal and non-fatal traffic accident.

One of the most complicated parts of prediction crowd flow movement is measuring the spatial-temporal dependencies while the external parameters such as surrounding point-of-interest (PoI) distribution and weather are involved. In [68] a novel deep spatio-temporal neural network model based on Graph Convolutional Network (GCN) LSTM called GCN-DHSTNet was proposed. The main target of the proposed model was to predict the crowd flows in each region at the same time. The proposed GCN-DHSTNet was able to show the dependencies of spatio-temporal while the effect of various parameters such as time granularity has been considered. Implementing the proposed model on two different datasets and evaluating it by using RMSE and MAPE as evaluation metrics, showed that the proposed model was able to overcome the baselines.

Due to the large growth of the number of vehicles, intelligent vehicles and the volume of data that is produced by them predicting, controlling and avoiding the traffic on the road and network is a real challenge in intelligent transportation systems. On the other hand, artificial intelligence algorithms had significant development in terms of making predictions in big and complex traffic data.

The previous works can be divided into two different approaches: first, the authors focused on road traffic and they investigated the different effective environmental parameters such as the number of vehicles, the weather condition, the road limitations, various driving rules and so on. Then they proposed various methods to detect, predict and control road traffic by implementing different traditional and developed learning algorithms. In the second approach, the authors addressed the network traffic flow and tried to monitor some parameters such as measuring the channel level usage, the packet loss, dropping the sender speed, throughput and so on which leads to network traffic flow and disrupting the network performance.

However, the main point that can be concluded from the study of previous works is that they did not consider the impact of these two different types of traffic (road traffic and network traffic) on each other at the same time. Vehicular networks are one of the most important parts of intelligent transportation systems which is able to support wireless communication respecting traffic analysis and safety. Although, the impact of road traffic on network traffic in VANET is also an important issue that should be considered.

The main part of this dissertation is published and submitted papers. The next three chapters contain the published/submitted papers.

The main objective of this dissertation is to propose a method for traffic prediction by using the AI method in the VANET environment. As mentioned in Chapter 1, we defined three sub-objectives in this dissertation with regard to the main objective. In each sub-objectives, we adopted various AI algorithms such as DL and ML algorithms to solve regression and classification problems in terms of traffic prediction in VANET.

The article presented in Chapter 3, proposes a network traffic prediction model based on AI algorithms in both regression and classification problems. his article entitled "Network Traffic Prediction Model Considering Road Traffic Parameters Using Artificial Intelligence Methods in VANET" has been published as a journal paper in "IEEE Access" journal. The proposed model predicts network traffic flow considering effective road and network parameters at the same time. The Random Forest-Gated Recurrent Unit- Network Traffic Prediction algorithm (RF-GRU-NTP) model consists of three phases. In the first phase, the main focus is on network traffic prediction based on V2R communication and ML algorithms have been implemented to predict the network traffic flow based on “packet receiving” as a classification problem. The classification evaluation metrics show that the RF got the best results in terms of prediction error and execution time. In the second phase, the center of attention is road traffic prediction based on V2V communication by implementing DL algorithms based on “sender speed” as a regression problem. The regression evaluation metrics prove that the GRU overcomes other baselines and gets better results in terms of prediction error and execution time. Finally, in the third phase, the proposed RF-GRU-NTP algorithm predicts the network traffic flow considering effective road and network parameters based on V2V and V2R communications. In this phase, the RF uses as a feature selection to find the most effective parameter and then the GRU is implemented to predict the network traffic flow in VANET.

The article presented in Chapter 4, analyses the performance of the most five common AI algorithms for network traffic prediction in VANET as a classification problem. This part of the dissertation was published as a conference paper with the title "A Comparative Study of Artificial Intelligence Algorithms for Network Traffic Prediction in VANET" by the IEEE publisher. Four different ML algorithms including RF, KNN, SVM and NB also a class of feed-forward Artificial Neural Network (ANN) called Multi-Layer Perceptron (MLP) are implemented and analyzed. Different classification evaluation metrics are employed and the results show that the RF with 96% accuracy and 0.73 minute execution time, has the best performance for traffic prediction in VANET.

In Chapter 5, a Federated learning algorithm for Network Traffic Prediction (Fed-NTP) based on Long Short-Term Memory (LSTM) algorithm to predict network traffic flow with

high accuracy and while preserving privacy is proposed. This chapter of the dissertation is submitted to "IEEE Access" journal with the title of "Fed-NTP: A Federated Learning Algorithm for Network Traffic Prediction in VANET". The LSTM algorithm is implemented in a decentralized way by using the federate learning (FL) algorithm on the VANET dataset and predicting network traffic based on the most effective features of network traffic flow in the road and network. By implementing the proposed FED-NTP model, there is no need to transfer the local data to the server over the network for training and just the trained model will transfer to the server which is the RSU in this work and the local data will remain in the client's place (vehicles). The simulation results show that the proposed model has better performance in terms of prediction error.

Therefore, the main objective and the three sub-objective of this dissertation are met and are explained comprehensively in more detail in the following three chapters. Network traffic prediction considering the road traffic happening and involving the most effective parameters in the road and network traffic can lead to an accurate traffic prediction. Although, based on the volume of produced data that is produced by intelligent vehicles in smart cities, preventing the leakage of information and security of data is an important issue that is considered in our objectives. Thus, the contributions of this dissertation lead to improving the accuracy of network traffic prediction, decreasing the execution time and preserving the privacy of data in the VANET environment.

CHAPTER 3 ARTICLE 1: NETWORK TRAFFIC PREDICTION MODEL CONSIDERING ROAD TRAFFIC PARAMETERS USING ARTIFICIAL INTELLIGENCE METHODS IN VANET

Sanaz Shakersepasgozar, Samuel Pierre

Mobile Computing and Network Research Laboratory (LARIM)
Department of Computer and Software Engineering, Polytechnique de Montréal
Montreal, QC H3T 1J4, Canada
Email : {sanaz.shaker-sepasgozar, samuel.pierre}@polymtl.ca.

Status: published in *IEEE Access*, 18 January 2022.

Abstract Vehicular Ad hoc Networks (VANETs) are established on vehicles that are intelligent and can have Vehicle-to-Vehicle (V2V) and Vehicle-to-Road Side Units (V2R) communications. In this paper, we propose a model for predicting network traffic by considering the parameters that can lead to road traffic happening. The proposed model integrates a Random Forest-Gated Recurrent Unit- Network Traffic Prediction algorithm (RF-GRU-NTP) to predict the network traffic flow based on the traffic in the road and network simultaneously. This model has three phases including network traffic prediction based on V2R communication, road traffic prediction based on V2V communication, and network traffic prediction considering road traffic happening based on V2V and V2R communication. The hybrid proposed model which implements in the third phase, selects the important features from the combined dataset (including V2V and V2R communications), by using the Random Forest (RF) machine learning algorithm, then the deep learning algorithms to predict the network traffic flow apply, where the Gated Recurrent Unit (GRU) algorithm gives the best results. The simulation results show that the proposed RF-GRU-NTP model has better performance in execution time and prediction errors than other algorithms which used for network traffic prediction.

Keywords: Vehicular network, network traffic prediction, road traffic prediction, regression methods, classification methods, machine learning algorithms, deep learning algorithms.

3.1 Introduction

One of the important technologies for the Intelligent Transportation System (ITS) is VANET that tries to make the environment safer and have better transportation using wireless com-

munications [5].

The traffic flow prediction with high accuracy is a significant issue in current transportation systems. It can help have the best path planning, make a better choice in selecting the greater route for travelers and decrease the traffic flow. Distinguishing that where and when the traffic will happen is a promising solution for managing transportation [1]. However, the new perspective of network traffic flow is that the traffic in the road could affect network traffic. According to the V2V communications in VANET, vehicles can send packets to each other to forecast the road traffic. By increasing the number of vehicles and traffic on the road, the number of packets sent would grow, leading to network traffic.

Previous studies worked on road traffic and network traffic separately, and we investigated them in the literature review. However, most of them addressed the traffic problem on the road or in the network independently, while in this paper, we will discover the relation between road and network traffic parameters together with the aim of network traffic prediction. Intelligent ways via machine learning (ML) techniques are the optimum solutions that can address traffic prediction problems with the aim of traffic flow prediction. There are some computational approaches like Bayesian modeling, fuzzy logic, hybrid modeling, Neural Networks (NN), and statistical modeling, which most of them, specially the NN, are promising solutions aiming to improve the accuracy of prediction in data traffic flow [2].

The significant point that should consider in all these ways, is the accuracy of prediction. ML techniques are divided into three types: Unsupervised Learning (training would be based on unlabeled data), Supervised Learning (training would be based on labeled data), and Reinforcement Learning (it learns from the performance of the learning agent). Moreover, some types of ML schemes like Transfer Learning and Online Learning are sub-categorized by these three types of ML schemes [9].

Another promising solution in the case of a large and complex dataset is deep learning (DL) algorithms for prediction problems. It has different types of algorithms that Recurrent Neural Network (RNN) [69, 70] and Convolutional Neural Network (CNN) [71] are the two famous algorithms that are used in many studies. Generally, the RNN has two modules called Long Short-Term Memory (LSTM) [72] and Gated Recurrent Unit (GRU) [73, 74], where the LSTM algorithm is similar to RNN by intention to address the vanishing problem. One of the most critical features of these algorithms is that they can learn dependencies for a long time with the aim of prediction in time-series datasets, and the GRU algorithm is like LSTM with more minor complications due to the number of its gate that leads to making it faster than LSTM [75]. Furthermore, to extract more features and bidirectional dependencies, Bi-directional Long Short-Term Memory (Bi-LSTM) algorithm can be used.

In this kind of algorithm, the sequence of the process can be done in two directions (forward and backward) using two different hidden layers [76].

In this work, we propose a network traffic prediction method considering road traffic parameters. To the best of our knowledge, it is the first time that network traffic is predicted due to road traffic happening. We try different machine learning and deep learning algorithms with the aim of network traffic prediction and divide our work into three phases. The first phase is about network traffic prediction. The second phase is about road traffic prediction. The third phase is a combination of the two previous phases with the aim of network traffic prediction considering the parameters, which are effective in road traffic that could affect the network traffic as well. We deploy two datasets from the Global Positioning System (GPS): the first one is based on V2R communication that is used for network traffic prediction and the second one is based on V2V communication that is used for road traffic prediction. The significant contributions of this work are as follows:

- We predict the network traffic flow using ML algorithms applied on a real dataset derived from V2R communications based on the packets sent by the vehicles to the Road-Side Units (RSUs). Furthermore, we implement variant ML algorithms like the RF, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Naive Bayes (NB) algorithms, where the best performance belongs to the RF.
- We concentrate on the impact of the sender's speed as a road parameter and predict the road traffic flow considering its effect on the network traffic flow. We take advantage of a real dataset built on V2V communications intending to predict the speed of senders using deep learning algorithms to aim for road traffic prediction. We tried different DL algorithms like the LSTM, Bi-LSTM, and GRU, where we got the best performance by the GRU algorithm.
- We investigate the effect of the last two previous steps based on road and network parameters on network traffic happening with the aim of network traffic prediction using both machine learning and deep learning algorithms. At this step, we combine both datasets (V2V and V2R), then we do feature selection using the RF algorithm to find the most influential parameters in the network traffic. After that, by implementing the GRU algorithm, we predict network traffic based on the adequate parameters in the road and network together. The results show that the novel proposed RF-GRU-NTP model would predict the network traffic affected by road traffic happening in the VANET environment more accurately than pure algorithms. Also, duration time for executing the proposed RF-GRU-NTP model aiming predict network traffic flow takes

in average 70% less time than LSTM and Bi-LSTM.

The originality of this paper lies on the fact that it combines machine learning and deep learning techniques into a single network traffic flow prediction model by considering road parameters and network parameters simultaneously.

The rest of this paper is organized as follows. Section II presents related work and provides an overview of the proposed prediction approaches in road traffic prediction and network traffic prediction. Section III explains the methodology and describes the implementation of the proposed prediction model phases. Section IV provides the evaluation results of the experimental validation for each phase. Finally, the conclusion of the study and our direction for future work are placed in section V.

3.2 Background and Related Work

Intelligent traffic prediction can support the critical advantages of solving traffic problems in the cities [21]. On the other hand, intelligent methods can help us predict the traffic flow considering different effective parameters. ML and DL methods are promising solutions for analyzing data and getting more accurate prediction results [77]. Several studies have been proposed different models to predict network traffic and road traffic independently using learning algorithms. Moreover, some researchers attempt to predict the road traffic flow using weather conditions that affect traffic flow. From a certain point of view, we can divide the previous works into two parts: road traffic prediction and network traffic prediction, where most of them used different machine learning and deep learning algorithms.

3.2.1 Network Traffic Prediction

In [78], the authors proposed a framework aiming to improve the prediction of network-wide link-level traffic deploying LSTM. They collected statics by Software Defined Network (SDN) or through SNMP measurements to forecast future throughputs of the network. They implemented their model on a real dataset, and they considered one hour, including a large amount of traffic of packets. They tried three different variations of LSTM, including Vanilla LSTM (vlstm), Delta LSTM (dlstm) and Multivariate LSTM (mlstm), and three types of Auto Regressive Integrated Moving Average model (ARIMA), including simple ARIMA model, Delta ARIMA model (darima), First-Order Autoregressive ARIMA (as1). Their experimental results showed that all variants of the LSTM had better performance than ARIMA based models for modeling the network traffic. However, they could have tried with other deep learning algorithms, which are more accurate than ARIMA baselines, to compare.

Intending to optimize network resource allocation and network traffic prediction, in [79], the authors deployed deep learning algorithms using a real dataset with five minutes time step for data-driven. They proposed the Evaluation Automatic Module (EAM) algorithm, which made the learning process automate and generalize the model for prediction aiming to have the best performance. Their proposed model is composed of two parts: in the first part, they used Artificial Neural Network (ANN) based on GRU algorithm to train the prediction model; and in the second part, they used EAM algorithm for model evaluation in every single iteration during the learning process. In addition, they used the Mean Absolute Error (MAE) metric for performance evaluation of the proposed model. The resource allocation part compared their results with static planning, which calculates the highest bandwidth related to the links in the traffic matrix. The experimental results showed that the proposed model had good accuracy in prediction and allocation resources.

A new methodology was aimed to improve the network traffic prediction proposed in [80]. They tried to predict network traffic intelligently using sequence mining, and for this purpose, they implemented the LSTM and Adaptive Neuro-Fuzzy Inference System (ANFIS) as a time series model. They selected real network data to implement their proposed model. Moreover, they clustered similar data using fuzzy c-means clustering. As well, they clustered the objects into five clusters as input for the LSTM and ANFIS algorithms with the aim of network traffic prediction. They evaluated the proposed model using metrics that diagnosed the prediction errors, which showed their model could decrease the prediction error and increase the network's performance.

Different types of CNN on real data, to find parameters of a network, which are optimal for network traffic prediction implemented in [81]. They run the algorithm in 1000 iterations with a [0.01-0.5] learning rate and tried different algorithms like multi-layer perceptron (MLP), CNN-RNN, CNN, CNN-LSTM and CNN-GRU. The results showed that the CNN and its different types of it are able to overcome other classical machine learning algorithms. The authors in [2], tried to develop and optimize the traffic learning by deploying the Taguchi method via layer-by-layer features categorized with the unsupervised algorithm. They proposed an optimized structure for traffic flow prediction trained with stacked autoencoder (SAE) and Levenberg–Marquardt (LM) algorithm to implement a short-term traffic prediction. Their approach aims to increase the precision of traffic flow prediction. The proposed model is a kind of deep architecture of the neural network using the Taguchi method to learn the features of traffic flow. To implement and evaluate their method, they used real collected data and implemented some constraints to the hidden layers, like considering five hidden layers. They assumed that the neurons were inactive most of the time and used the LM algorithm for training the input that the last autoencoder generated. They compared

their result with different algorithms like a radial basis function NN (RBFNN), hybrid exponential smoothing and the LM algorithm with NNs (EXP-LM), particle swarm optimization algorithm with NNs (PSOINN) and they found that the accuracy rate of prediction in the proposed model is almost 90% which had the best performance. The traffic flow data might be irregular because of unpredictable situations, and the proposed model was practical in dealing with this problem.

A new method for LTE network traffic prediction is presented in [82]. They used three different machine learning algorithms, including the RF, Bagging, and SVM on public cellular traffic datasets with the aim of network traffic prediction. They evaluated their performances using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination R^2 . The results showed that the Bagging has good results in a combination of numerical and categorical features. The Bagging and RF got the lowest and highest RMSE, 2.59 and 3.38, respectively. Moreover, the Bagging and RF got the lowest and highest MAE, 1.60 and 2.19 as well as the Bagging got the highest R^2 with 50.8%. However, the Bagging took 116 (s) and RF took 112 (s) time for learning, and the SVM took 6 (s) that is less time in comparison with them.

3.2.2 Road Traffic Prediction

Deep Neural Network (DNNs) can forecast the traffic flow with big data. However, there are some challenges about spatial-temporal issues. In [1], the authors proposed a forecast model to cover these problems and to analyze the internal mechanism of DNN based traffic flow prediction (DNN-BTF) on traffic flow data with reliable precision, in order to enhance the exactness of prediction. For validation of their model, they used data from the open-access database PeMS¹. Moreover, they used the RNN and CNN (with an entire convolutional network) algorithms for temporal and spatial features, respectively.

They adopted three different indexes for performance evaluation: the Mean Relative Error (MRE), the MAE, and the RMSE. Furthermore, they compared their proposed model with the adaptive absolute shrinkage and selection operator (LASSO) [83], traditional shallow back-propagation neural network (BPNN), stacked autoencoder (SAE) [84], DeepST [85], and Sequence to Sequence learning (StoS) methods. The results showed that their model outperformed them. In [86], the authors proposed a spatial-temporal feature selection algorithm using GRU (GRU+STFSA) for prediction of short-term traffic flow, which was a combination of spatial and temporal analysis using the GRU. They compared the results with CNN and simple GRU algorithms. They evaluated their model by using the MAE, the

¹<http://pems.dot.ca.gov/>.

Mean Absolute Percentage Error (MAPE), and the RMSE metrics. The proposed model showed better performance in stability and accuracy with a minor error in prediction.

The hybrid CNN-LSTM model proposed in [87] for traffic flow prediction of GPS data. They adopted a greedy policy approach to train the CNN-LSTM model to solve hybrid models' complexity and time-consuming training process. The experimental results showed that the deep hybrid proposed model could predict the traffic more accurately in less run time than the Linear(a model with dense layer using softmax), CNN and CNN-LSTM model, considering both temporal and spatial features. The authors in [88], presented a hybrid framework for predicting short-term traffic flow based on Support Vector Regression (SVR). They applied the RF for finding the compelling features and improved the Genetic Algorithm (GA) to find optimal features. They collected data from real-world datasets to implement the proposed RF-CGASVR model, and after selecting features by the RF, the training phase was done. Intending to evaluate the proposed model, they tried it into two layouts. The first one was straight, which was designed with the aim of performance evaluation of light scenarios. The second one was the crossroad layout, which was about intersections and different roads. They used the RMSE and MAPE as metrics for performance evaluation. The results showed that the proposed model performed well while they did not consider long-term traffic. They also compared the evaluation results of proposed model with different algorithms like ARIMA and SVR-based models. However, there are some studies about traffic prediction considering weather conditions using different deep learning algorithms [89–93].

Based on the large amount of data produced in smart cities by different intelligent vehicles, road parameters that can cause road traffic could affect network traffic. The Artificial Intelligence (AI) approach can be an optimal solution for network traffic prediction. However, the previous works focused on network traffic prediction or road traffic prediction, so the lack of study on considering these two issues together motivated us to propose a novel network traffic prediction approach based on road traffic flow.

Table 3.1 shows a summary of previous studies in road traffic prediction, network traffic prediction and their proposed methods.

3.3 Methodology

In this section, we propose a novel RF-GRU-NTP model using machine learning and deep learning algorithms with the aim of network traffic flow prediction in VANET. Fig. 3.1 shows the architecture of VANET environment.

VANET is a dynamic environment due to the presence of moving characters (like vehicles) in

Table 3.1 Summary of related work in road and network traffic prediction

Proposed method	Category	Comparison algorithms	Approach
LSTM	Network traffic prediction	vlstm, mlstm, dlstm, arima, darima and asl	[15]
EAM	Network traffic prediction	GRU	[16]
ANFIS	Network traffic prediction	LSTM	[17]
CNN	Network traffic prediction	MLP, CNN-RNN, CNN-LSTM and CNN-GRU	[18]
SAE-LM	Network traffic prediction	RBFNN, EXP-LM and PSONN	[3]
LTE network prediction method	Network traffic prediction	RF, Bagging and SVM	[19]
DNN-BTF	Road traffic prediction	RNN and CNN	[2]
GRU+STFSA	Road traffic prediction	GRU and CNN	[23]
Hybrid CNN-LSTM	Road traffic prediction	Linear and CNN	[24]
RF-CGASVR	Road traffic prediction	ARIMA and SVR-based models	[25]

the network that make it complicated in prediction problems. By applying machine learning algorithms on appropriate datasets, we can increase prediction accuracy in traffic problems [94]. Furthermore, deep learning algorithms are employed to predict complex patterns in faster and more accurate ways.

Many parameters can affect the network traffic flow, so the proposed model is divided into three phases with the objective of network traffic prediction considering different effective parameters. Fig. 3.2 shows the architecture of the proposed model, including three phases.

3.3.1 Phase 1 - Network Traffic Flow

In the first phase, we just focused on V2R communications, and we predicted the network traffic flow based on the packets sent by vehicles to the RSUs. In this step, we implemented classification methods, then we tried different machine learning algorithms like the KNN [95],

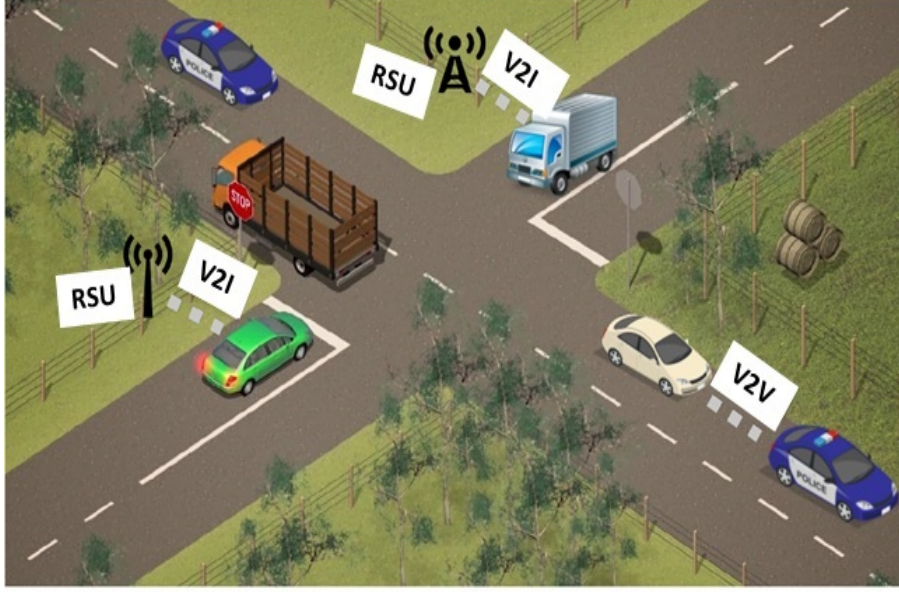


Figure 3.1 The architecture of VANET.

RF [96], NB [97], and SVM [98], intending to have the best performance and accuracy in prediction results with regards to precision, recall, F1-score, and accuracy. Moreover, we used the area under the Receiver Operating Curve (ROC) metric, confusion matrix, and computation time to evaluate the algorithms considering all aspects of the fittest algorithm aiming to network traffic prediction.

Random Forest Algorithm

The RF is a supervised learning algorithm composed of several decision trees that are combined together to get more accurate prediction results. Therefore, we can use the RF algorithm in regression and classification problems, as well as it can overcome the main problem of machine learning algorithms that is over fitting [99]. In our research work, we deployed four different machine learning algorithms on a dataset that got by GPS [100], and we got the best results with the RF algorithm for network traffic prediction. The vehicular network dataset (the network type was 802.11 adhoc) that we used was including the measure of short-range communications performance between vehicles together and vehicle and RSUs. Some parameters like latitude, longitude, speed of sender, speed of the receiver, packet receiving are reported in the dataset. We selected 10000 records of one day to predict network traffic flow based on packets that are sent by vehicles to the RSUs. To evaluate all deployed classification methods, we used a confusion matrix intending to find out the number of correct classifier's predictions, the incorrect ones, and even in the situation that they got confused.

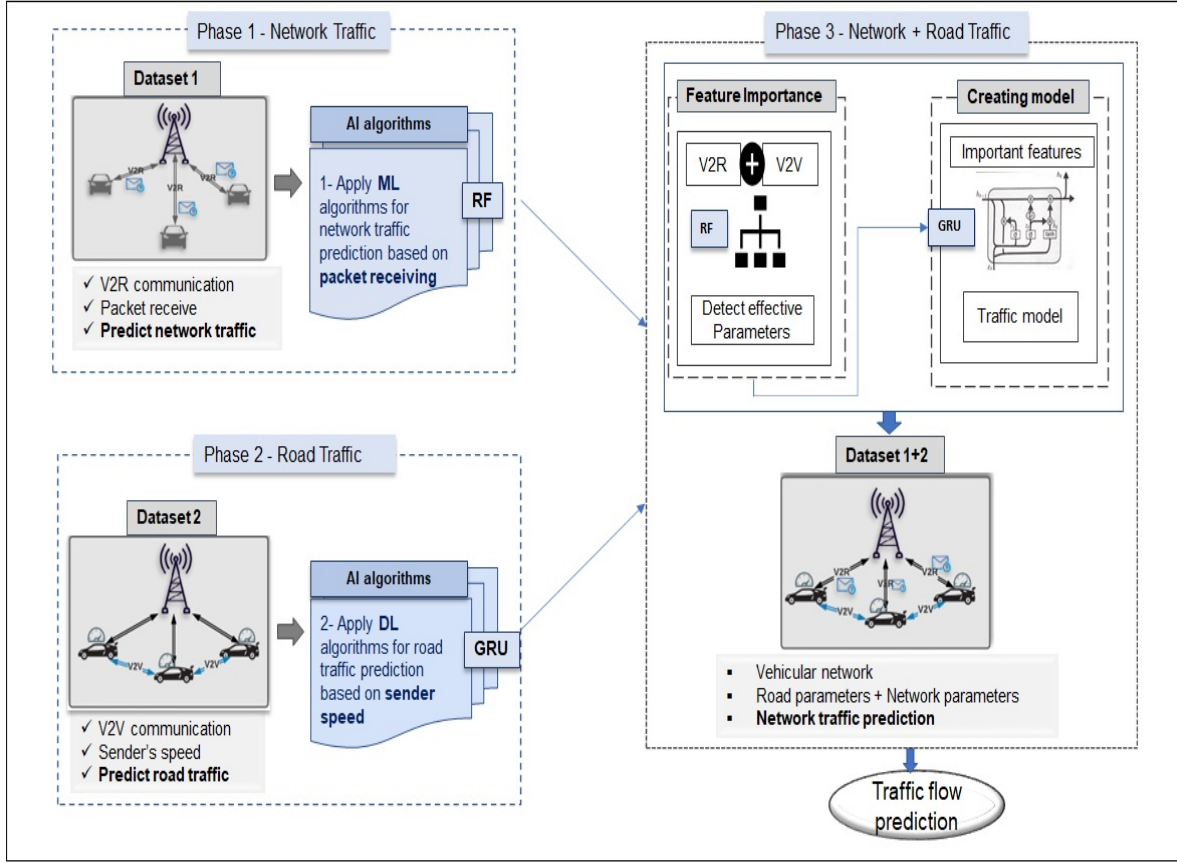


Figure 3.2 The architecture of VANET.

Classification Metrics

We used four metrics to evaluate our classification model [101]: the precision, the recall, the F1-score, and the accuracy.

1. *Precision* shows the number of true positive predicted observations over all the positive predicted including false and true, which is denoted by:

$$Precision = \frac{TP}{TP + FP} \quad (3.1)$$

2. *Recall* represents the ratio of true positive predicted observations over all the observa-

tions in a real class, which is designated by:

$$Recall = \frac{TP}{TP + FN} \quad (3.2)$$

3. *F1 – score* is affected by Precision and Recall, and it is denoted by:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.3)$$

4. *Accuracy* shows the number of true positive and true negative observations over all the number of samples, which is designated by:

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \times 100 \quad (3.4)$$

where,

TP: True Positive;

FP: False Positive;

TN: True Negative;

FN: False Negative.

Fig. 3.3 shows the first phase workflow of the proposed model including the algorithms, libraries, tools, and metrics that we used. Our vehicular network dataset contains V2R communications that got by GPS receiver.

We labeled data into two classes: class 1 for a case that packet is received, and we assumed it in a non-traffic situation; and class 0 for not receiving a packet that we presumed as a traffic situation.

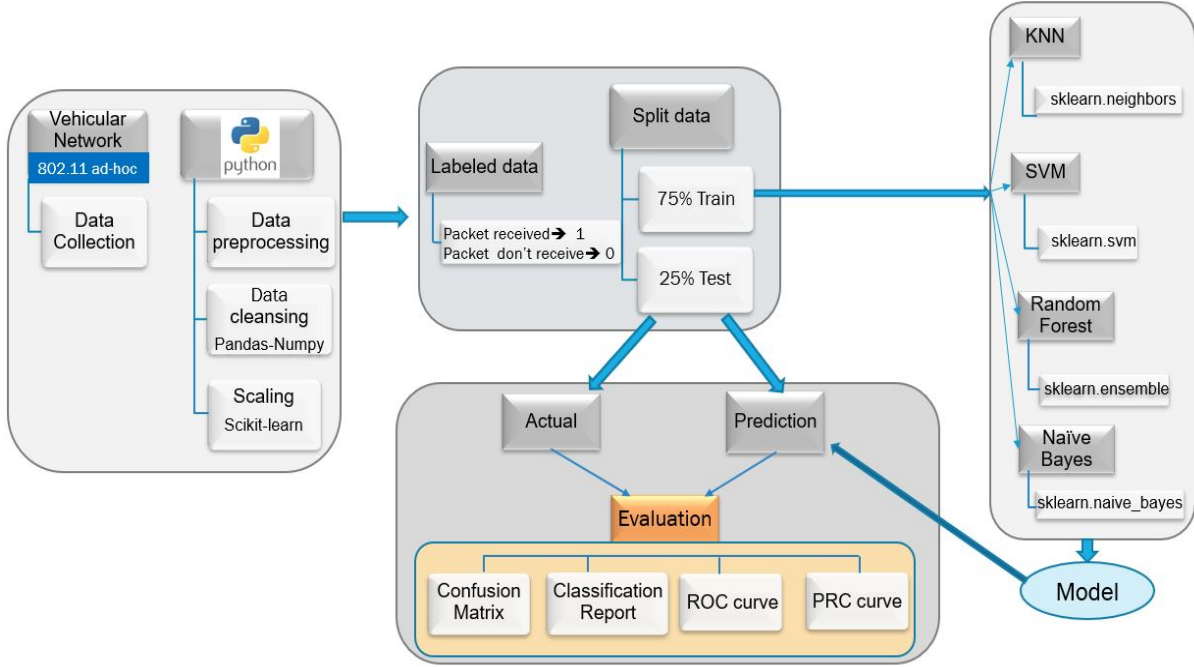


Figure 3.3 The workflow of network traffic prediction (phase1).

$$\text{Traffic flow classes} = \begin{cases} \text{In non-traffic situation;} \\ \text{Packet received: class1} \\ \text{In traffic situation;} \\ \text{Packet does not receive: class 0} \end{cases} \quad (3.5)$$

After that, we split the dataset based on different tries and considering the best computational cost into 75% for train and 25% for test. At the next level, we implemented four different machine learning algorithms on the train set to predict network traffic and create our model with the highest performance. Based on various metrics that we used to evaluate the models, we had got the best results when we deployed the RF algorithm for network traffic prediction considering packet receiving as a network parameter.

3.3.2 Phase 2 - Road Traffic Prediction

The second phase focuses on V2V communication considering the speed of vehicles (“sender speed”), which are sending packets as a road parameter to predict the road traffic flow. At this phase, we tried regression methods and deployed three different deep learning algorithms, including the GRU, Bi-LSTM, and LSTM. We evaluated their performances using some

evaluation metrics for deep regression algorithms, and we got the best result with the GRU algorithm in road traffic prediction.

GRU Algorithm

The RNN algorithm has different types, and the GRU is one of the efficient algorithms derived from the LSTM. The GRU inherits the essential advantage of RNN, which is learning features automatically. It can act like the LSTM and memorize the performance of prediction for the long term in a faster way. The LSTM has three gates containing input gate, forget gate, and output gate while the GRU algorithm has just two gates including update gate and reset gate which makes the GRU less complicated and more efficient [86]. The role of the reset gate is determining how much of the previous information needs to be forgotten, and the update gate defines the amount of memory needed to keep around and pass through to the future information [75]. Fig. 3.4 shows the internal structure of the GRU algorithm, and its equations are as follows:

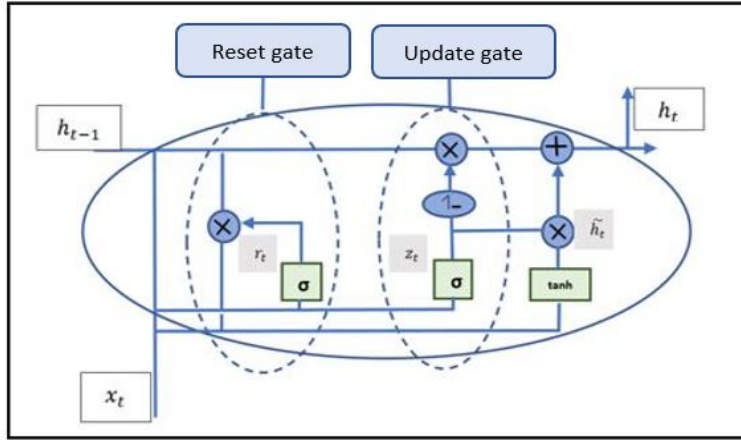


Figure 3.4 The structure of the GRU.

$$\text{Update gate:} \quad Z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (3.6)$$

$$\text{Reset gate:} \quad r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (3.7)$$

$$\text{Process input:} \quad \tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (3.8)$$

$$\text{Hidden state update: } h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (3.9)$$

The update gate has been calculated in (3.6) for time step t , where x_t is the input of update gate and W_z represents its weight that their multiplication would be accumulate with the sum of h_{t-1} that shows the past information in time $t - 1$ and its weight that shows with U_z . Then the result will be squashed between 0 and 1 by applying the sigmoid activation function (σ) [102]. Activation functions can normalize the output of neural network models. This gate determines how much of the previous information is necessary for passing to the future. The reset gate is calculated in (3.7) that is similar to the update gate. However, the difference is that the reset gate will decide which past information is needed to forget. The effect of gates on the final output will be calculated in (3.8). W shows the weight of the update gate, r_t represents the input that got by the reset gate, the output of the previous neuron showed by h_{t-1} and the input of the current neuron is presented by x_t , which they are multiply on $\tanh$ as a nonlinear activation function. This procedure shows the process for a new memory content. The output of the current neuron is represented by h_t in (3.9), z_t represents the update gate, h_{t-1} is showing the output of the previous unit and $\tilde{h}_t$ represents the output value that is pending in the current unit [86], [102].

We considered a situation for traffic flow in the road, which is when the speed of senders drops off under 60 Km/h, we would have traffic in the road considering the speed of vehicles in our dataset are in the range of (0, 104 Km/h) and there was not any speed limit.

$$\text{Road traffic happening} = \begin{cases} \text{In non-traffic situation;} \\ \text{Sender speed} > 60 \text{ Km/h} \\ \text{In traffic situation;} \\ \text{Sender speed} < 60 \text{ Km/h} \end{cases} \quad (3.10)$$

After data collection, we did pre-process, data cleansing, and scaling using the MinMax method [103]. We split the dataset into 75% for train and 25% for a test. Based on our dataset and parameters, we implemented three different regression algorithms, including the LSTM, GRU and Bi-LSTM, to speed prediction and created our model with the highest performance. To implement the regression algorithms, the Keras library alongside TensorFlow was used [104]. The data has been passed 120 times for training; thus, the model was fit to run on 120 epochs. Different optimizers have been tried like Stochastic Gradient Descent methods

(SGDs) [105], Adaptive Gradient Algorithm (AdaGrad) [106], Adaptive Moment Estimation (Adam) [107], and Root Mean Square Propagation (RMSProp) [108]. We got the best result from the Adam optimizer for stochastic optimization, then to avoid overfitting and improve the performance, we added a dropout layer into our algorithms, then we evaluated our model using different evaluation metrics. Fig. 3.5 shows the workflow and structure of phase two.

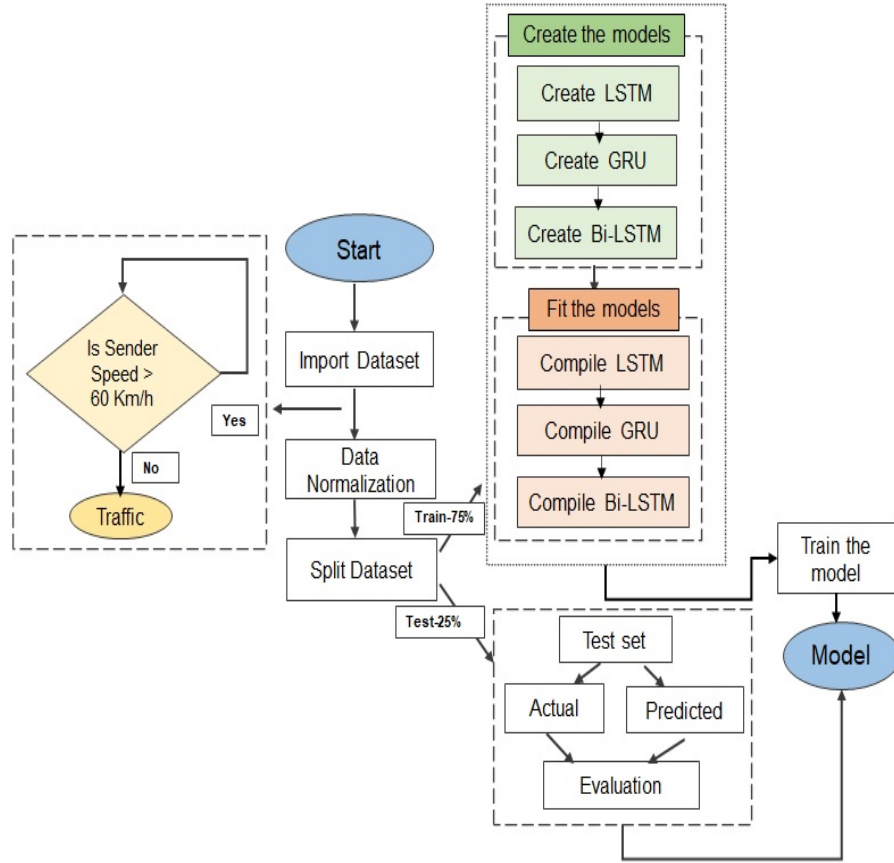


Figure 3.5 The workflow of road traffic prediction (phase 2).

Regression Evaluation Metrics

We used four evaluation metrics including the Mean absolute error (MAE), the Mean Squared Error (MSE), the Root Mean Squared Error (RMSE) and the $R^2 - SCORE$, to evaluate our regression models [109].

1. MAE (Mean absolute error), shows the sum of the total difference between the actual

values and the predicted values, which is denoted by:

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |x_i - \hat{x}_i| \quad (3.11)$$

2. MSE (Mean Squared Error), illustrates the difference between the actual and the predicted value by squared, which is designated by:

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (3.12)$$

3. RMSE (Root Mean Squared Error), can show a more accurate error rate by squared MSE metric, which is denoted by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (3.13)$$

4. $R^2 - SCORE$, represents the ratio and the efficiency of the predicted model and how to fit it in comparison with the actual value, as well as it ranges from 0 to 1 where the closest value to 1 shows the better model. It is denoted by:

$$R^2 = 1 - \frac{\sum(x_i - \hat{x}_i)}{\sum(x_i - x_{avg})} \quad (3.14)$$

where,

x_i = Real Value;

$\hat{x}_i$ = Predicted Value;

x_{avg} = Scores of all outputs are averaged with uniform weight;

n = The number of examples of test or verification set.

Based on evaluation results, we got the best performance by deploying the GRU algorithm to predict “sender speed” as a road parameter to predict road traffic.

3.3.3 Phase 3 - Network Traffic Prediction Considering Road Parameters

Finally, in the third phase, we intend to predict the network traffic flow considering the parameters that affect road traffic. We propose an RF-GRU-NTP model to predict the network traffic. To the best of our knowledge, it is the first time that network traffic flow

is predicted based on effective parameters in road traffic flow. Furthermore, we propose a model that combines deep learning and machine learning algorithms to predict the network traffic flow. The structure of the proposed model in Phase 3 is presented in Fig. 3.6.

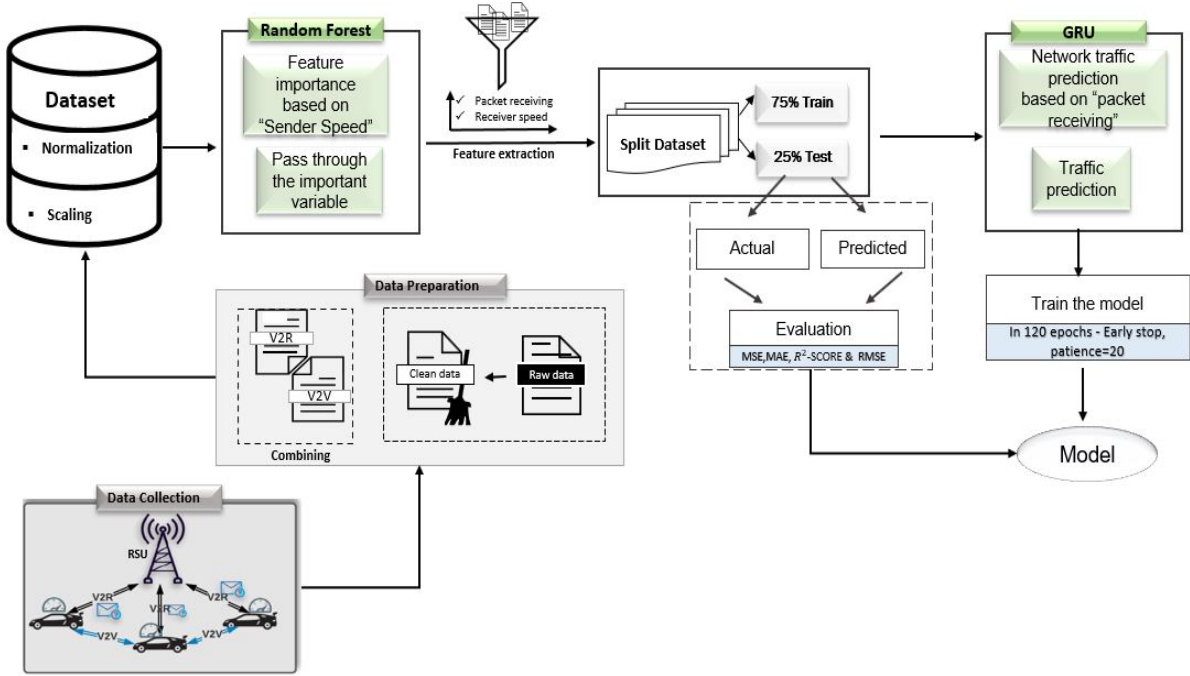


Figure 3.6 The structure of the proposed RF-GRU-NTP model (phase 3).

This phase includes a machine learning algorithm as a feature selection and a deep learning algorithm in terms of network traffic prediction considering road network parameters. In this way, at the first step, we combined V2V and V2R datasets. We investigated on parameters including “receiver speed”, “packet receive”, “time”, “signal strength”, and “noise strength”, then by deploying the RF algorithm, the features that are effective on “sender speed” as a road parameter in road traffic happening have been found. Afterward, we passed the important features through the GRU as inputs to predict the network traffic flow. Also, we considered 75% of data for training and 25% of that for testing.

Effective Parameters Detection by The Random Forest

The RF algorithm plays the role of detecting effective parameters on the network traffic flow. The RF algorithm is based on selecting a vector (A, B) irregularly and including trees that are growing to the decision tree. Then, by voting through the n trees and considering the training data (A, B), which provides k samples for building a regression tree, it extracts the c subset and puts them back randomly. The $i(i \in c)$ subspace has been extracted,

and the other samples that did not extract in the subspace are out-of-bag (OOB) data. To construct the regression tree's feature space, a fixed dimension vector would be chosen from an M dimensional vector, and it can be considered an input variable [110]. Then considering minimum variance, which is computed during the growth process, the splitting can be calculated as follows:

$$I = \min_s \frac{\sum_{s=1}^k (y_s - \hat{y}_s)^2}{k} \quad (3.15)$$

where,

I = Optimal splitting variable;

S = Embedded sample dimension;

y_s = Value of variable;

$\hat{y}_s$ = Average value of variable.

The RF algorithm is made by the growth of trees, and we can compute the effect of out of bag data on the model as follows:

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (3.16)$$

where,

x_i = Real Value;

$\hat{x}_i$ = Predicted Value;

n = The number of samples out of bag.

The RF algorithm, in terms of forming the out of bag data matrix, will generate n new trees and will get the different ranks of $MSE[MSE_1, MSE_2, \dots, MSE_n]$. Moreover, measuring the importance of the input variable is calculated as follows [110]:

$$Vim = \frac{\sum_{j=1}^n (MSE_j - MSE_{nj})^2}{n \cdot S_E} \quad (3.17)$$

where,

Vim = Importance score;

$n(j \in n)$ = Number of decision tree;

S_E = Standard error of n decision tree.

Network Traffic Prediction by GRU

After finding out which features are more important to be as an input variable for GRU, we made a new filtered dataset for more accurate prediction using effective parameters. At this level, we filtered our combined dataset (V2V+V2R) and defined traffic based on “sender speed” and assume if the sender speed is less than 60 Km/h, then we will have traffic. Then using RF, we figured that the most effective parameters on “sender speed” were “receiver speed” and “packet receiving”. Then by applying the GRU algorithm, we predicted the network traffic flow based on the most important parameters in the road and network. Based on evaluation results, the proposed model can predict the network traffic flow more accurately than other algorithms like LSTM and Bi-LSTM, which have been tried.

3.4 Data Preparation and Performance Evaluation

3.4.1 Dataset

We used a vehicular network dataset (the network type was 802.11 ad-hoc) [100] to measure short-range communications performance based on V2V and V2R communications in a highway. For data gathering, they put an external antenna on the roof of vehicles. The vehicle’s longitude, latitude, speed, and heading were reported by GPS every two seconds. The data got from the highway in Atlanta has five usual lanes and one High Occupancy Vehicle (HOV) lane, which has been monitored between 2 pm and 5 pm. It reports the location’s information with an accuracy of five to seven meters, and they got information of location via interpolation. In V2R communication, 1470 bytes of packets were broadcasted by the senders at a mean rate of about 150 packets/s. In V2V communication, the sender and receiver were installed on vehicles [111]. Both datasets based on V2V and V2R communication were used in our implementation. For the first phase, our focus was on network traffic prediction. We used a dataset based on V2R communication, and we predicted the network traffic flow based on receiving packets by the RSUs. Then, for the second phase, we used the dataset of V2V communication; and we targeted the “sender speed” to predict road traffic, while both sender and receiver vehicles were in the same lane (the second one from the right side of the road) but they had different distances. Finally, for the third phase, since we aimed to predict the network traffic flow considering the effect of road traffic on it, we combined V2V and V2R datasets. After data collection, we did pre-processing, data cleansing and scaling using

StandardScaler method which applied in order to scale and normalize the data. Then, using a machine learning algorithm, we extracted the effective parameters in a combined dataset of network traffic flow.

To implement the algorithms, we used Python version 3.6 [112], and for each phase, we used different libraries. For the first phase, different machine learning algorithms had been implemented. We used Scikit-learn, Pandas, NumPy, Matplotlib, Mlxtend, and some more libraries intending to deploy machine learning algorithms to find out the fittest algorithm for network traffic prediction. In consequence, we find that the RF had better performance. For phases two and three, which implemented deep learning algorithms, we used Keras libraries and TensorFlow [104] to implement the LSTM, Bi-LSTM and RF-GRU-NTP. After data collection, we did pre-process, data cleansing, and scaling to have more accurate prediction results.

3.4.2 Performance Evaluation of The Network Traffic Prediction

As mentioned previously, we used different metrics, such as confusion matrix, precision, recall, F1-score, accuracy, ROC curves and Precision-Recall curve as classification metrics to evaluate the machine learning algorithms. For the deep learning part (phase two and three) we used the MSE, MAE, RMSE, and $R^2 - SCORE$ to evaluate the deep learning algorithms. Moreover, for all the phases, we computed execution time to find the most suitable algorithm in our model.

We performed a confusion matrix as an evaluation metric in classification problems, which shows the different situations between true labels and the predicted ones [113], as shown in Fig. 3.7.

		Prediction	
		True	False
Actual	True	TP	FP
	False	FN	TN

Figure 3.7 The structure of the confusion matrix.

where,

TP: When (Actual = 1, Predicted = 1)

FP: When (Actual = 0, Predicted = 1)

TN: When (Actual = 0, Predicted = 0)

FN: When (Actual = 1, Predicted = 0)

The true labels are presented in the rows representing the correct label and the predicted labels are placed in the column to determine the value that the classifiers have predicted. The number of times that true labels and predicted ones are matched would be presented in the diagonal.

In the first phase, our target was network traffic prediction based on packet receiving and we considered two classes for packet receiving in our dataset. Then, we labeled them: class 0 for NO (not receiving the packet) and class 1 for Yes (receiving the packet) as shown in (3.5). After that, we split the dataset into 75% for train and 25% for test and implemented four different classification algorithms. Then, we evaluated the algorithms by a confusion matrix. Fig. 3.8 shows the confusion matrix of the RF, NB, KNN and SVM algorithms.

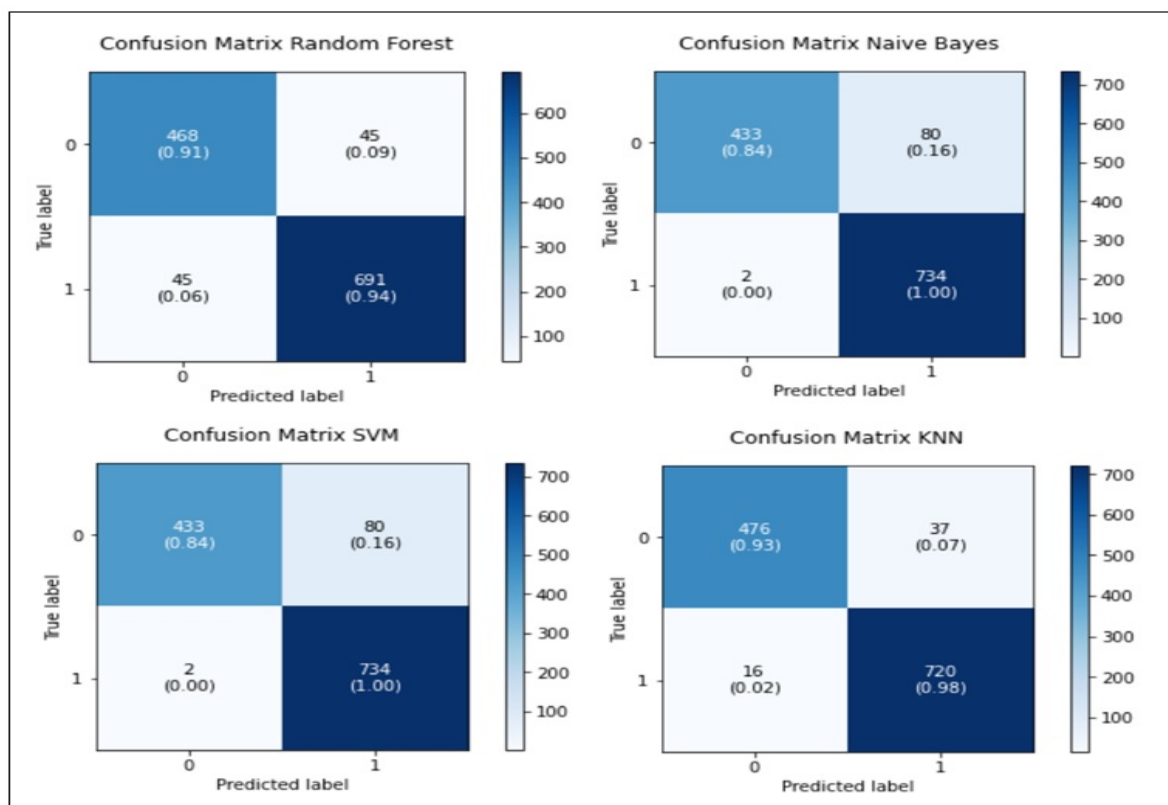


Figure 3.8 Confusion matrices of the RF, NB, SVM and KNN.

The confusion matrix of the RF indicates in 468 cases that the true label was 0 and it has been predicted correctly; and in 45 cases, the true label was 0, and it has been predicted wrong. Moreover, the prediction results for label 1 are shown in the next row. The NB's confusion matrix indicates 433 correct prediction cases for label 0 and 734 true prediction cases for label 1. The confusion matrix of the SVM and KNN algorithms shows that the SVM has better prediction in label 1 with 734 true cases, and the KNN has better performance at predicting label 0 with 476 correct cases.

One of the common graphical metrics to present the result of classification problems is the Receiver Operator Characteristic (ROC) curve considering the Area Under the Curve (AUC). Moreover, another metric called the Precision-Recall (PR) curve can give us more valuable graphs to evaluate the performance of the algorithms [111]. By using the ROC curve, we clarify the accuracy of classifiers in prediction results.

The ROC curve will calculate the probability of predictions, and it shows the False Positive Rate (FPR) on the X-axis and True Positive Rate (TPR) on the Y-axis. The TPR can describe the model's predictive performance in positive class in the situation that the true label is positive as well. Fig. 3.9 depicts the results of four different algorithms that we used for network traffic prediction. However, describing the prediction of positive class in the condition, where the true label is negative, is the responsibility of FPR. Moreover, there is a classifier with a random acting level shown in a red dotted line. It would separate the area into two parts that can improve the performance estimation in a way, where the curves that have better performance would be located above this line and close to the top-left corner, and the poor algorithms would be placed under the red dotted line.

Furthermore, the larger AUC indicates that the algorithms have better performance. As shown in Fig. 3.9, the blue curve, which belongs to the KNN, and the orange curve that indicates the RF algorithm, are closer to the top left corner and they have larger AUC, which means they have higher performance, respectively. In addition, the green curve that represents the SVM and the red one that belongs to the NB algorithm, show the lower performance with 0.94 and 0.95 AUC consecutively.

However, the other evaluation metric for machine learning algorithms is the PR curve that presents a balance between precision indicated at the Y-axis and recall pointed out at the X-axis. Fig. 3.10 illustrates the PR curve for different machine learning algorithms that we implemented. The blue dotted horizontal line demonstrates a "baseline" classifier, where the ideal classifier is above this line and the closest to the right top corner. Thus, the lowest performance would belong to the curves that are closer to the baseline. As depicted in Fig. 3.10, the KNN and RF algorithms, shown in orange and green, respectively, have the highest

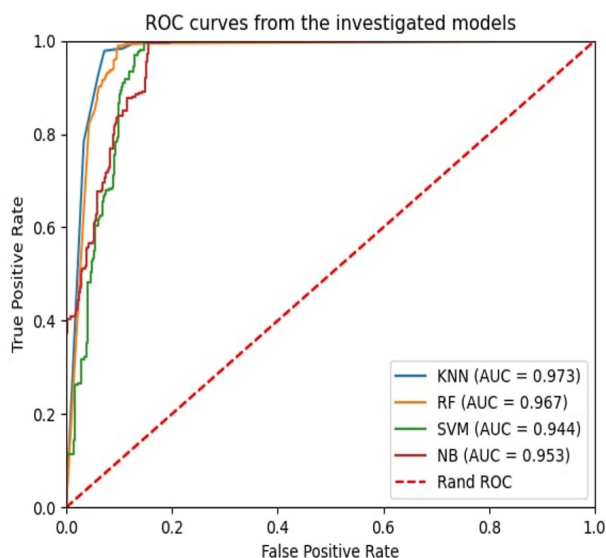


Figure 3.9 ROC curve for KNN, RF, SVM and NB.

performance; and the NB and SVM, shown in purple and red, have the lowest performance accordingly.

We tried classification reports and other metrics like precision, recall and F1-score to get more insight into our model performance. Table 3.2 shows the evaluation results of each algorithm. Therefore, the KNN algorithm has the highest accuracy with 96%, and the SVM has the lowest one with 91% in our prediction model, while the RF and NB give the same percentage of accuracy with 93%.

Table 3.2 Classification reports

Algorithm	Accuracy (%)	Label (packet receiving)	Precision	Recall	F1-score
KNN	96	No	0.97	0.91	0.94
		Yes	0.94	0.98	0.96
RF	93	No	0.94	0.91	0.92
		Yes	0.93	0.96	0.95
SVM	91	No	0.99	0.85	0.91
		Yes	0.90	0.99	0.95
NB	93	No	0.99	0.85	0.91
		Yes	0.90	0.99	0.95

After the performance evaluation that we have done, the results of the KNN and RF was almost the best. However, the other factor that should be considered in performance evalua-

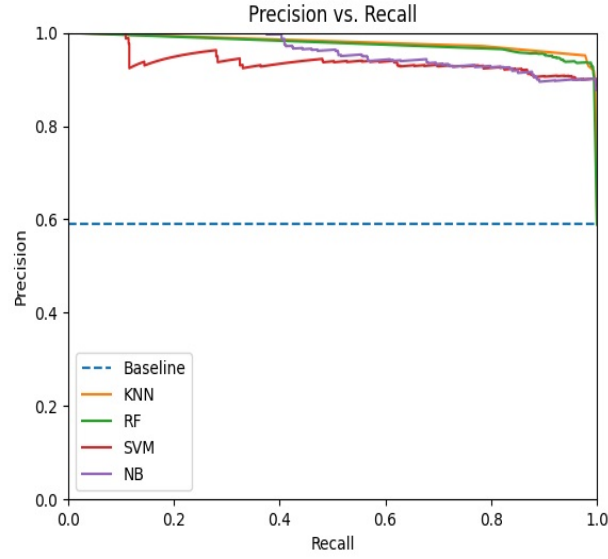


Figure 3.10 PR curve for the KNN, RF, SVM and NB in network traffic prediction.

tion is the execution time to find the best algorithm for network traffic prediction, specifically when the volume of data is large. Table 3.3 present the execution time for each algorithm based on the results. The KNN is the most time-consuming algorithm, and the NB took less time. Therefore, based on all the results that we have got, we can conclude that the RF is the perfect match for our model with network traffic prediction, because it is not as time consuming as KNN. On the other hand, its performance is good enough in other evaluations.

Table 3.3 Execution time for the RF, NB, SVM and KNN

Algorithm	Execution time(s)
RF	0.1403975
NB	0.0819984
SVM	0.1957004
KNN	0.7682998

For the second phase, our target was road traffic prediction based on “sender speed”. We deployed three different deep learning algorithms; then, we evaluated their performances to find the fittest algorithm for our model. The MAE and the RMSE are two standard evaluation metrics to represent the average performance of the model [114].

However, we calculated the MSE as a metric for evaluating the prediction results and the

$R^2 - SCORE$ that is the most important metric for our model, because it can indicate how the predicted model is good in future observation which the higher value of it, indicates a better prediction model, and it means the difference between the actual and predicted value is insignificant. Finally, we considered the time that each of the deep learning algorithms took for fitting the model. The results of calculating the MAE for all algorithms are illustrated in Fig. 3.11, where it shows that the LSTM experienced the highest error at the beginning. On the other hand, the GRU has the lowest one, and it has less iteration to get the best result, however, the LSTM and Bi-LSTM have more epochs, respectively.

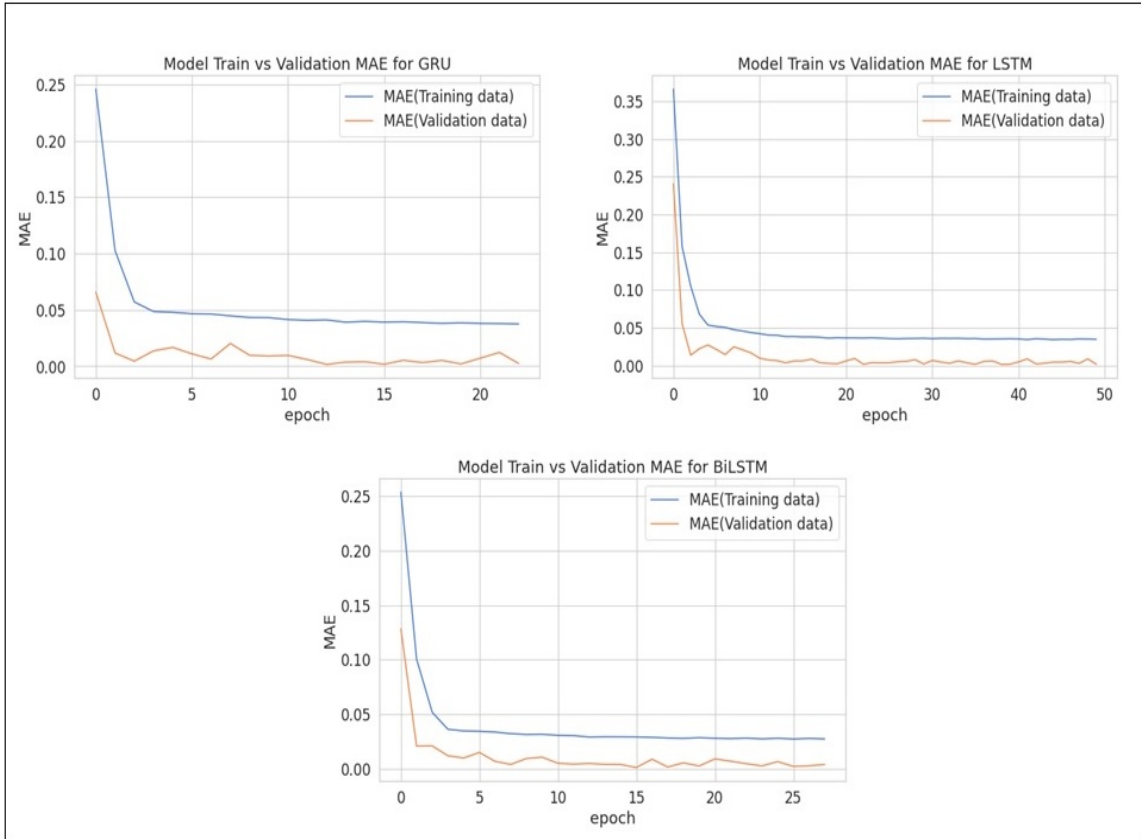


Figure 3.11 The MAE for GRU, LSTM and Bi-LSTM (phase2).

By considering the growing number of vehicles and facing a large amount of data, execution time is a vital parameter for choosing our model's fittest algorithm. Thus, we calculated the time that each algorithm took for fitting the model, and as illustrated in Table 4.1, the GRU is the fastest one with 1 minute and 3 seconds in model fitting, and the Bi-LSTM and LSTM are the slowest, respectively.

Then, we predicted the sender speed with the aim of road traffic prediction by implementing

Table 3.4 Duration time for GRU, Bi-LSTM and LSTM

Algorithm	Duration time (m)	Epochs
GRU	1:03.246713	Epoch 49/120
Bi-LSTM	3:24.082677	Epoch 56/120
LSTM	5:27.551359	Epoch 53/120

the GRU, LSTM, and Bi-LSTM algorithms. As shown in Fig. 3.12, the blue line indicates the test data and the prediction that different algorithms have done is in the orange line. The results show that the GRU has performed better than the LSTM and Bi-LSTM. Thus, we can see that the difference between test and predicted data is insignificant.

Moreover, we calculated some evaluation metrics like the MSE, MAE, RMSE and $R^2 - SCORE$ for all three algorithms for the final evaluation step. As demonstrated in Table 3.5, we got the best results in all metrics with the GRU algorithm. Furthermore, the highest $R^2 - SCORE$, the most crucial factor for evaluating our model, was 0.995, which belonged to the GRU algorithm.

Table 3.5 Evaluation metrics for Bi-LSTM, GRU and LSTM

Algorithm	MSE	MAE	RMSE	$R^2 - SCORE$
Bi-LSTM	2.062	0.98	1.436	0.993
GRU	1.402	0.416	1.184	0.995
LSTM	1.975	0.984	1.405	0.993

Based on all evaluation ways done, the GRU is the best algorithm that can fit our model in the lowest time and error, and the highest $R^2 - SCORE$.

Finally, we predicted network traffic flow considering road traffic flow by implementing machine learning and deep learning algorithms for the third phase. By deploying the RF algorithm, we detected the most effective features on “sender speed”. Fig. 3.13 shows that the “receiver speed” and “packet receiving” are the most effective parameters on “sender speed”, which we consider as a network traffic parameter.

The next step is passing the most influential parameters on network traffic as input variables into the GRU algorithm with the aim of network traffic prediction and implementing the proposed RF-GRU-NTP model. Fig. 3.14 shows the MAE for the LSTM, RF-GRU-NTP and Bi-LSTM algorithms. As depicted, the LSTM experienced the highest error at the

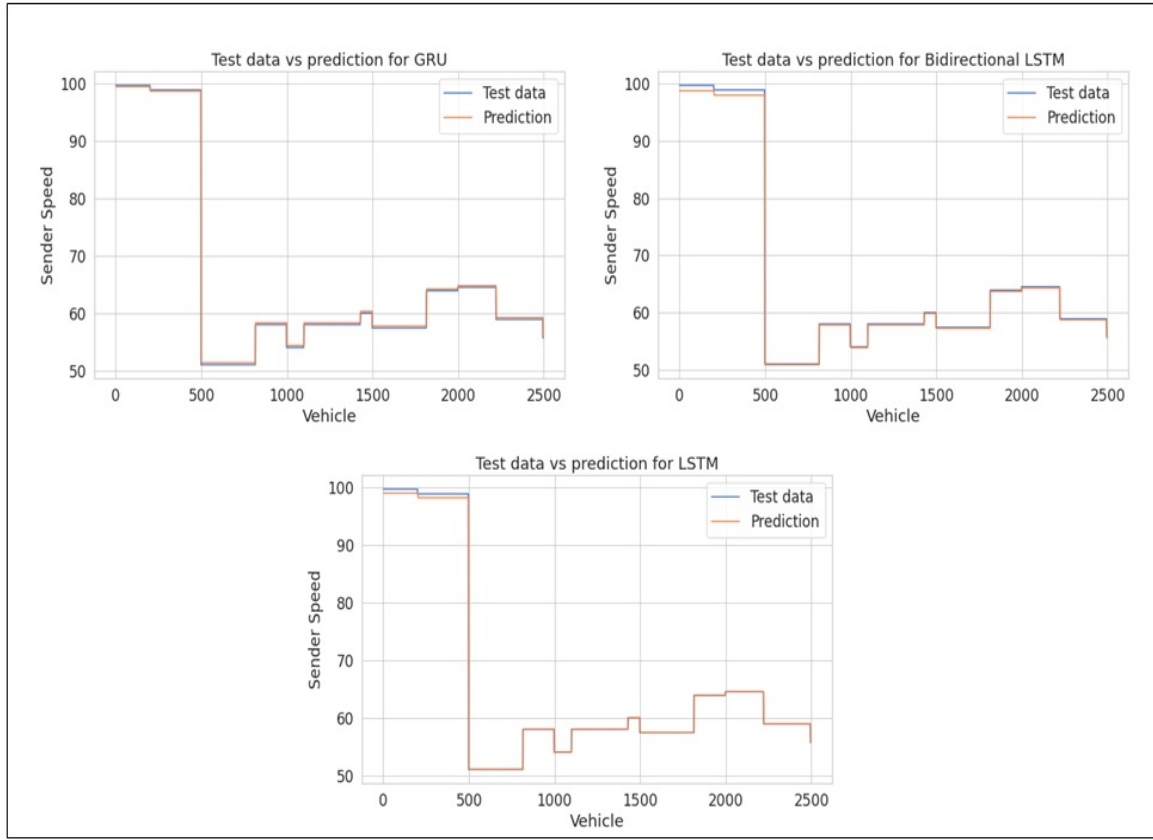


Figure 3.12 Test data vs prediction for the LSTM, GRU and Bi-LSTM (phase2).

beginning and Bi-LSTM had the most iteration while the proposed RF-GRU-NTP had better performance in less than 25 iterations.

We predicted network traffic flow considering road traffic parameters and evaluated the performance of our model using four different evaluation metrics as the MAE, MSE, RMSE, and $R^2 - SCORE$. Table 3.6 shows the value of these evaluation metrics for each algorithm, demonstrating that the proposed RF-GRU-NTP model has the lowest error and the highest $R^2 - SCORE$, which means it has better performance in comparison with the LSTM and Bi-LSTM.

Fig. 3.15 depicts the prediction results for sender speed considering effective parameters on it in order to network traffic prediction. The blue line indicates the test data and the prediction that has been done by models is in the orange line. As shown, the proposed RF-GRU-NTP model performed better than other pure algorithm that we implemented, and the difference between test and predicted data is insignificant while the highest difference between predicted

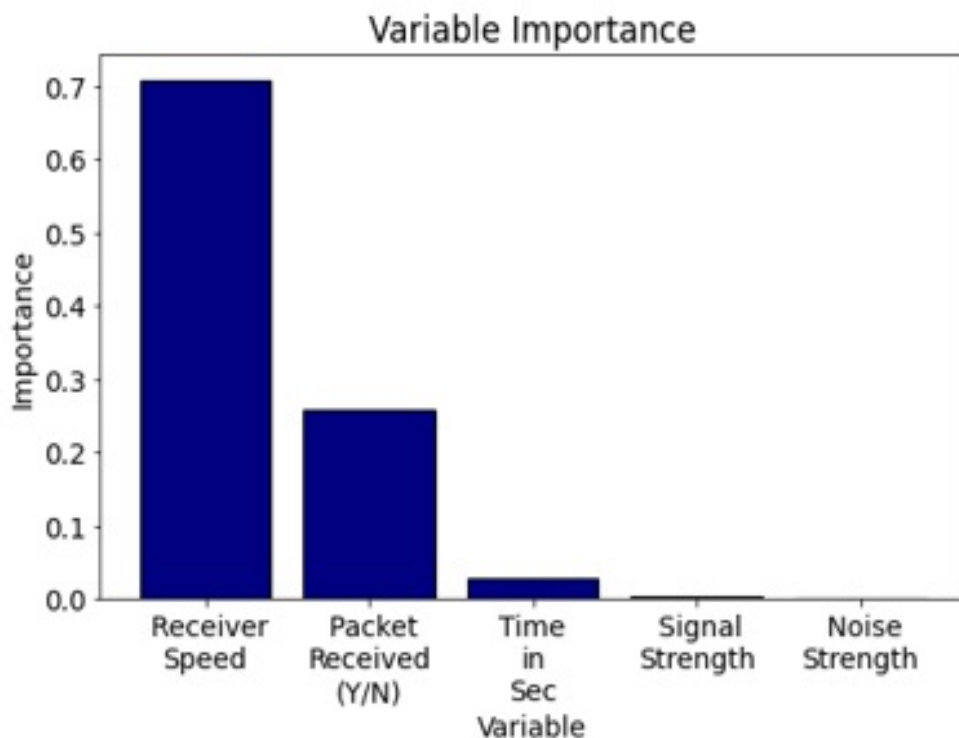


Figure 3.13 Effective parameters in network traffic.

Table 3.6 Evaluation metrics for Bi-LSTM, RF-GRU-NTP and LSTM

Algorithm	MSE	MAE	RMSE	$R^2 - SCORE$
Bi-LSTM	1.911	0.711	1.382	0.993
RF-GRU-NTP	1.437	0.372	1.199	0.995
LSTM	1.677	0.427	1.295	0.994

and actual value belongs to the Bi-LSTM algorithm.

After calculating the evaluation metrics, we computed the time that each algorithm took for fitting the model. Moreover, we set an early stop method to stop training when the algorithm has not improved, considering 20 for patience's value. As shown in Table 3.7, whereas the RF-GRU-NTP algorithm could fit the model in less than 34 seconds, the Bi-LSTM and LSTM needed about 2 minutes and 4 minutes, respectively.

Fig. 3.16 shows the difference in fitting time between the proposed RF-GRU-NTP model and the pure algorithms that we implement them. The proposed RF-GRU-NTP model has significant different fitting times from other algorithms.

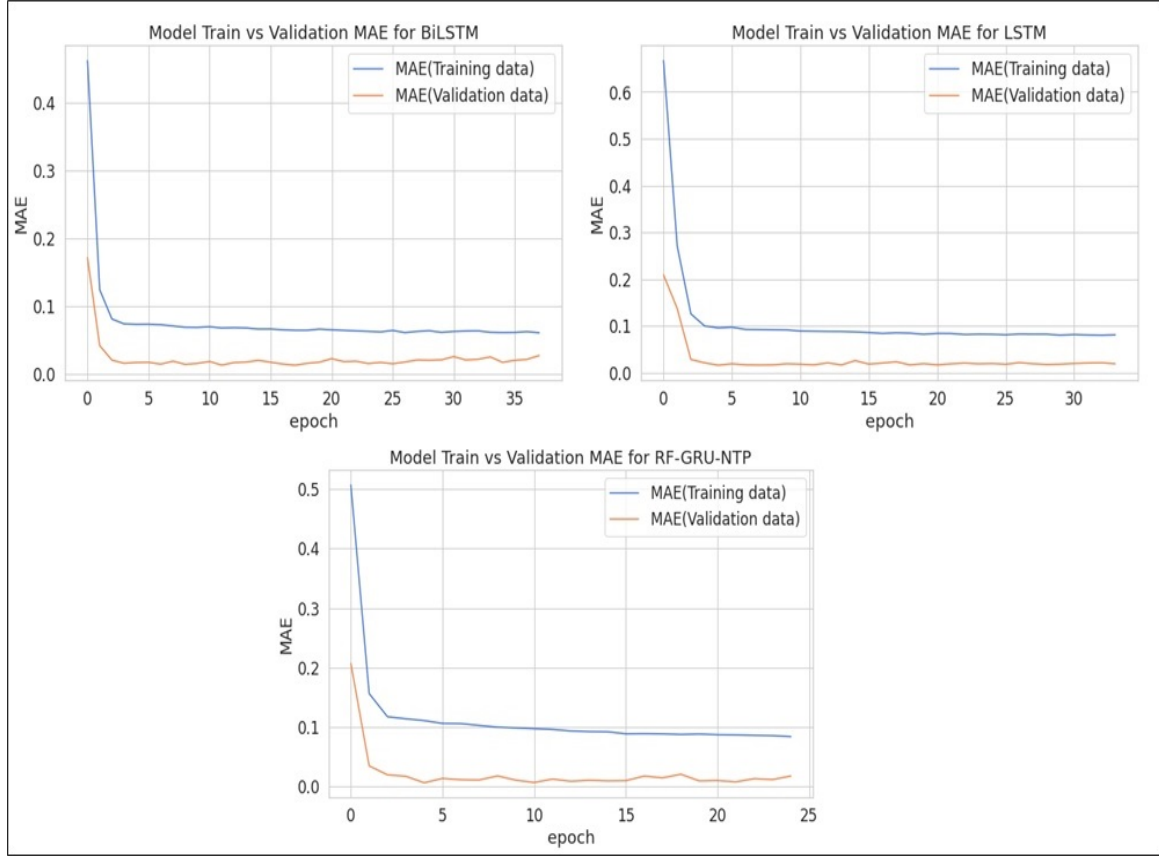


Figure 3.14 The MAE for GRU, LSTM and Bi-LSTM (phase3).

Consequently, based on all results that we got, the proposed RF-GRU-NTP model has the best performance in network traffic prediction considering road traffic parameters.

3.5 Conclusion

In this paper, we proposed an RF-GRU-NTP model with the aim of network traffic flow prediction based on the traffic in the road and network simultaneously. We divided our research into three phases. In the first phase, we focused on network traffic prediction. We used the V2R dataset and considered the receiving packets sent by vehicles to the RSUs as a network parameter to predict network traffic flow. Then, we tried different machine learning algorithms like the RF, NB, KNN, and SVM algorithms, and we evaluated them using some classification metrics. After all evaluations, the RF has the better performance to predict network traffic flow while our target was “packet receiving”.

In the second phase, we tried to predict the road traffic flow using the V2V dataset while

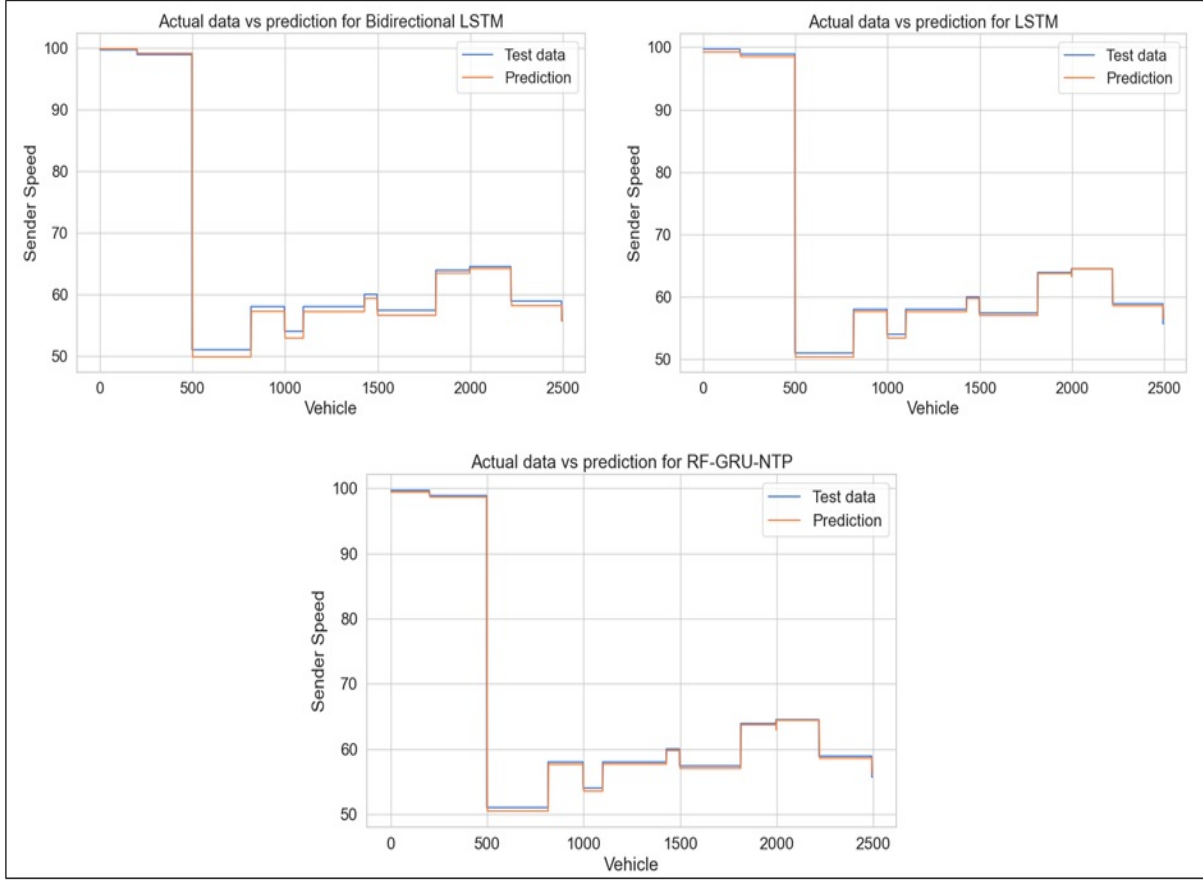


Figure 3.15 Test data vs prediction for the LSTM, RF-GRU-NTP and Bi-LSTM (phase 3).

our target was “sender speed” to define the road traffic. We assumed that the traffic would happen on the road if the senders’ speed were less than 60 Km/h. Therefore, we implemented different deep learning algorithms, including the LSTM, GRU, and Bi-LSTM. Finally, we evaluated the results using some regression evaluation metrics, which, based on the results we got, the GRU was the fittest algorithm for road traffic prediction.

Then in the third phase, we implemented our target, which is network traffic flow considering road traffic flow, by combining machine learning and deep learning algorithms. For this purpose, we combined V2V and V2R datasets and used the RF algorithm for feature selection. We found the most important features, which were “packet receive” and “receiver speed” that can affect “sender speed” and the network traffic flow. Then by implementing the proposed RF-GRU-NTP model, we predicted network traffic flow.

Therefore, we compared our results with a pure algorithm like LSTM and Bi-LSTM to make sure that the proposed model has good results in network traffic flow prediction.

Table 3.7 Duration time for RF-GRU-NTP, Bi-LSTM and LSTM

Algorithm	Duration time (m)
RF-GRU-NTP	0:33.302000
Bi-LSTM	02:06.236182
LSTM	04:09.028719

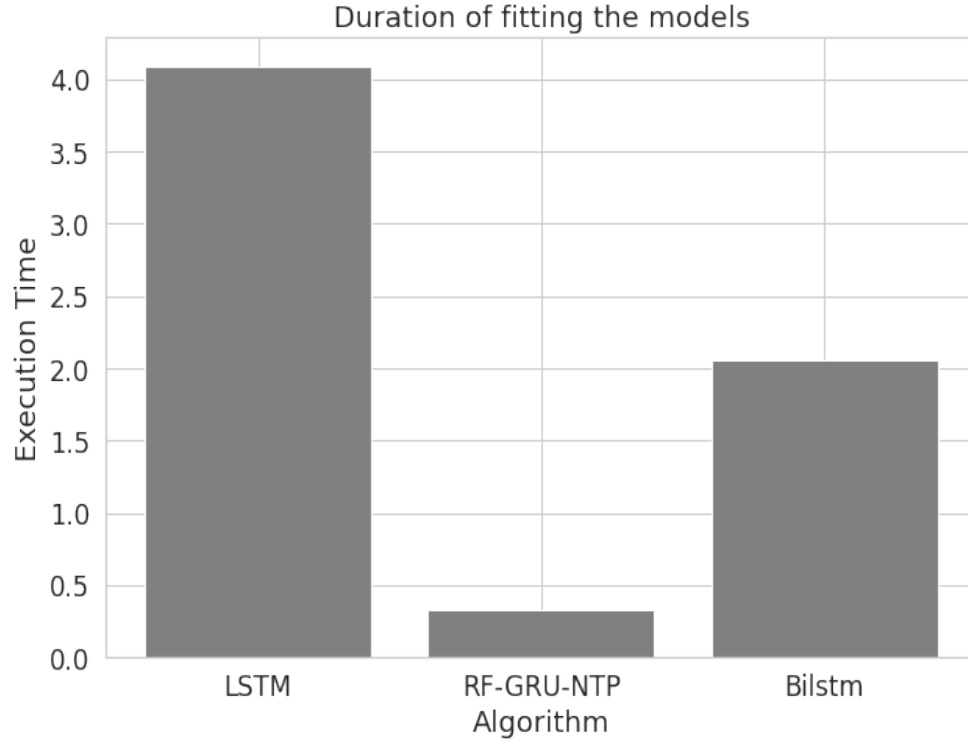


Figure 3.16 Time for fitting the LSTM, RF-GRU-NTP and Bi-LSTM (phase 3).

The main complexity of the proposed model was combining two datasets in order to implement machine learning and deep learning algorithms with the aim of network traffic prediction considering different types of parameters. To the best of our knowledge, this is the first research that predicts the network traffic flow based on road traffic flow.

However, by growing up the number of vehicles, the volume of produced data by them would take shape of big data which in our future work we will implement our proposed model in big data.

CHAPTER 4 ARTICLE 2: A COMPARATIVE STUDY OF ARTIFICIAL INTELLIGENCE ALGORITHMS FOR NETWORK TRAFFIC PREDICTION IN VANET

Sanaz Shakersepasgozar, Samuel Pierre

Mobile Computing and Network Research Laboratory (LARIM)
Department of Computer and Software Engineering, Polytechnique de Montréal
Montreal, QC H3T 1J4, Canada
Email : {sanaz.shaker-sepasgozar, samuel.pierre}@polymtl.ca.

Status: published in *18th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 10 October 2022.

Abstract

Increasing the number of vehicles and their communications in smart cities is a critical issue that leads to road and network traffic. Traffic prediction with high accuracy and less complexity is a challenge in Intelligent Transportation System (ITS). In Artificial Intelligence (AI), Machine learning (ML) algorithms are promising solutions to prediction problems, and Deep Learning (DL) algorithms are used for more complicated issues. In this paper, we propose a comparative analysis of the prediction performance of the most five common AI algorithms used to solve classification problems. Different evaluation metrics are employed to analyze algorithms and get the most accurate one for selecting such problems. Simulation results on the Vehicular Ad-Hoc Network (VANET) dataset revealed that Random Forest (RF) as a traditional ML algorithm performed better than other algorithms in terms of accuracy (96%) and execution time (0.73 minute) for traffic prediction.

Keywords: Vehicular Ad-Hoc Network, network traffic prediction, Artificial Intelligence, machine learning method, neural network, Intelligent Transportation System

4.1 Introduction

One of the most common technologies deployed in smart cities is ITS, which VANET is a crucial part. This latter consists of wireless communication between Vehicle-to-Road Side Unit (V2R) and Vehicle-to-Vehicle (V2V). Different types of AI algorithms are applied for discovering the pattern and extracting valuable knowledge for prediction [115]. Several studies

address network traffic prediction but finding the best algorithm with the highest prediction accuracy and minimum error is still a challenge. Consequently, this challenge can affect the network performance in terms of delay and accuracy. Regarding the importance of accuracy of prediction and execution time based on increasing the volume of data, we need to implement the most precise algorithm considering its execution time.

In this paper, we implement four different machine learning algorithms including RF [96], K-Nearest Neighbors (KNN) [95], Support Vector Machine (SVM) [98] and Naive Bayes (NB) [97]. Moreover, we try a Multi-Layer Perceptron (MLP) [116], which is a class of feed-forward Artificial Neural Network (ANN). MLP algorithm is used for solving classification problems, and it relies on a neural network. It uses the back-propagation technique for training, which can produce a single output from a set of inputs.

Moreover, we use a real dataset collected from Global Positioning System (GPS) and is taken from V2R communication based on packets that are sent by vehicles to the Road Side Units (RSUs). Our main target is finding the best classification algorithm to predict the network traffic based on the “packet receiving” parameter. We define traffic situations as when the packet has not been received and in non-traffic situations, the vehicles will receive the packet. We compare the prediction results of the five above listed AI algorithms for classification problems. The main contribution of this paper is to show that RF is the best AI algorithm in terms of accuracy, execution time and error.

The rest of this paper is organized as follows. Section II presents related work on different machine learning and deep learning algorithms for traffic prediction. Section III explains the proposed method for implementing KNN, RF, SVM, NB and MLP algorithms. The evaluation results are illustrated in Section IV and the conclusion is drawn in Section V.

4.2 Related Work

With the growth of the number of vehicles and their communications in smart cities, network traffic prediction becomes a major issue. A wide range of related work has focused on deploying various AI algorithms, in order to propose an accurate and accelerated traffic prediction model. A large number of these works adapted ML algorithms [117], [118], plenty of them presented DL algorithms [81], [119] and some works combined DL and ML algorithms together to propose a prediction model [120], [121].

A novel prediction model was proposed in [122] with the aim of predicting network traffic accurately. They combined three different algorithms: (1) Local Mean Decomposition (LMD) for decomposing the network traffic time series to get components of product func-

tion, (2) Bidirectional Long Short-Term Memory (BiLSTM) algorithm for predicting the network traffic data and (3) the Bayesian optimization algorithm to optimize the hyperparameters in BiLSTM. They implemented their proposed model on two actual datasets and used some evaluation metrics including Root Mean Square Error (RMSE), Relative Root Mean Square Error (RRMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Square Sum Error (SSE), Theil Inequality Coefficient (TIC), Index of Agreement (IA), and Coefficient of Determination (R2). The prediction results showed that their model outperformed Auto Regressive Integrated Moving Average (ARIMA), Empirical Mode Decomposition (EMD), BiLSTM, SVM and MLP algorithms.

Lim *et al.* [123] performed network traffic classification by implementing Convolutional Neural Network (CNN) and Residual Network (ResNet) for classifying network traffic. Their main objective was to fit unseen packets by enabling a deep learning model. They implemented the proposed model on the generated packet-based dataset and analyzed the performance of the network traffic classification model using the $F1 - score$.

A framework on a Long Short Term (LSTM) neural network algorithm to perform the network traffic prediction at a short time scale (i.e., <30 seconds) was used in [124]. The authors compared their results with RNN and ARIMA algorithms by using Normalized Root Mean Square Error (NRMSE) as an evaluation metric. Therefore, the results showed that the proposed model performed better than compared algorithms.

An improved Support Vector Machine algorithm called Cost-Sensitive SVM (CMSVM) has been proposed in [125]. This proposed algorithm is based on multi-class SVM algorithm. The authors used active learning for every single class of traffic and considered weight values, in order to solve the imbalance problem in network traffic recognition. They did the implementation by using two imbalanced datasets. The experimental results showed that the proposed CMSVM algorithm was able to overcome the imbalance problem, decrease the computation cost and enhance the accuracy of classification.

A method for predicting the Long-Term Evaluation (LTE) network edge traffic was proposed in [82]. They implemented three algorithms including the SVM, RF and Bagging on the public cellular traffic dataset. They used the RMSE, MAE and R2 as evaluation metrics. Based on these metrics, the Bagging algorithm has the lowest error, but its training time is 116s, while the SVM takes 6s for training. Therefore, SVM is significantly better than the Bagging and the RF which have both the worst error and training time. Thus, the algorithm that has the best performance in case of error evaluation is the most training time-consuming.

In [126], the authors attempted to solve the network traffic prediction problem by implementing various Recurrent Neural Network (RNN) architectures like simple RNN, LSTM

and Gated Recurrent Unit (GRU) algorithms. They focused on three problems including packet distribution prediction, volume prediction and packet protocol prediction. To evaluate the performance, they used the Mean Squared Error (MSE), metric and compared the various architectures of the RNN algorithm with ARIMA, Naive Method (NM) and Moving Average (MA). The experiment results showed that the simple RNN, LSTM and GRU outperformed other algorithms in terms of accuracy error evaluation.

4.3 Proposed Model for Network Traffic Prediction

In this section, we present a comparison between various traditional ML algorithms (i.e., KNN, RF, SVM and NB) and one of the most common AI algorithms based on the neural network called MLP for network traffic prediction. Fig. 4.1 depicts the architecture of our proposed prediction model to compare the five selected algorithms.

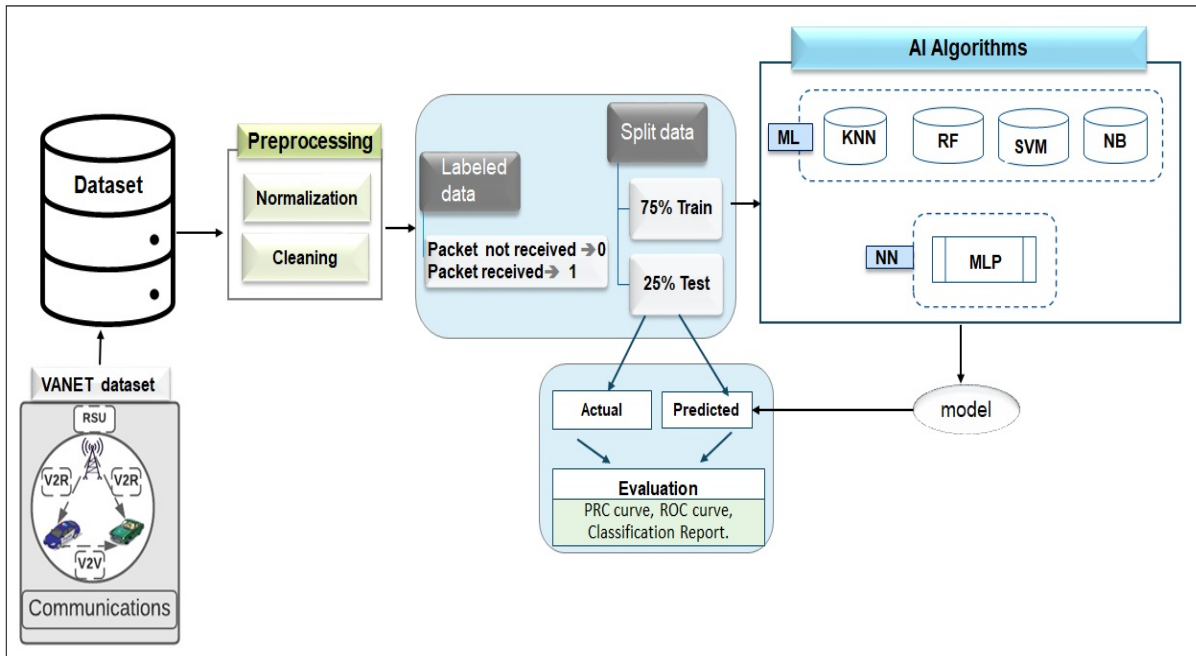


Figure 4.1 The architecture of proposed comparison model.

4.3.1 Dataset

We used a real dataset based on VANET communications [100], where the type of network is 802.11 Ad-Hoc. It contains two kinds of communication data based on V2V and V2R for measuring short-range communications. The authors installed a GPS antenna on each

vehicle and collected the longitude, latitude, speed and heading every two seconds. In our comparison work, we focused on V2R communication where 1470 bytes of packets had been broadcasted via the senders at an average rate of around 150 packets/s.

In the preprocessing step, we did data cleaning and scaling using the MinMaxScaler method for normalizing and scaling the data.

Moreover, for network traffic prediction based on V2R communication, we labeled data into class 0 for not receiving packets in a traffic situation and class 1 for receiving packets in a non-traffic situation. Then, we split data into 75% for training and 25% for tests. We implemented five different AI algorithms (i.e., KNN, RF, SVM, NB and MLP). Finally, we evaluated their performance based on various evaluation metrics, in order to find the most suitable algorithm for network traffic prediction in VANET.

4.3.2 Implementation

To implement the KNN, RF, SVM, NB and MLP algorithms, we used Python version 3.7 [112] and some libraries including Sklearn, Pandas, NumPy, Matplotlib and Mlxtend. For implementing the MLP algorithm, Adam was utilized as an optimizer and we chose ReLU as an activation function including 300 iterations. All the implementations took place in Google Colab using GPU accelerator to have better performance and improve the processing time.

4.3.3 Selected AI Algorithms

This section presents the five selected AI algorithms for network traffic prediction: KNN, RF, SVM, NB, and MLP.

- KNN: one of the straightest algorithms in ML is the nearest neighbor used for both regression and classification problems. The KNN classification has two stages: (1) searching the nearest neighbors and (2) defining the classes that used those neighbors. Selecting the k nearest neighbor is done by distance metric [95]. Although the KNN is simple and easy to implement, finding sample take a long time to perform.
- RF: includes a set of decision trees selected randomly from the training set, then for finding the final class, the votes will aggregate to make the final decision. The RF can decrease the overfitting in decision tree algorithms and enhance the accuracy [96].
- SVM: can be used for both binary and multiclass classifications or regression problems with high accuracy. It is based on separating data by a hyperplane in an optimal way to make classes by considering training samples, which exist at the edge of each class [98].

- NB: is a powerful algorithm for binary and multiclass classification problems and works based on Bayes theory that connects the different conditional probabilities of random events. The NB is a scalable and fast algorithm but it cannot learn the feature's relation.
- MLP: is an appropriate prediction algorithm in labeled data and regression problems. The MLP is a feed forward class of ANN algorithm that adjusts suitable connection between input and output data. It consists of an input layer, a hidden layer including a certain number of neurons called perceptron and an output layer, where each of these layers has a full connection with the weighted connection to the units in the previous layer and the output will go through a nonlinear activation function designated by (4.1) [116]:

$$y = \varphi\left(\sum_{i=1}^n w_i x_i + b\right) = \varphi(W^T X + b) \quad (4.1)$$

where W denotes the weights, X represents the vector of inputs, b is the bias and φ is the activation function. The training part of the MLP tries to adapt the weights and biases to their ideal value for minimizing the error, which is denoted by [116]:

$$E = \frac{1}{l} \sum_{i=1}^l (T_i - Y_i)^2 \quad (4.2)$$

where T_i denotes the predicted value, Y_i is the actual value and l is the training set size. The architecture of MLP network is illustrated in Fig. 4.2 [116].

The MLP algorithm is an adaptive neural network algorithm that can learn to do the different tasks based on the training dataset. Moreover, it is a fully connected algorithm, which has too many parameters that might cause inefficiency.

In this paper, we implemented four traditional ML algorithms and one simple neural network algorithm for network traffic prediction to solve classification problems. We used some evaluation metrics, in order to evaluate the performance of these algorithms. The RF algorithm gives the best result in network traffic prediction in terms of accuracy and execution time.

After implementing these five algorithms considering traffic happening in the network based on packet receiving, we need to evaluate them using various evaluation metrics.

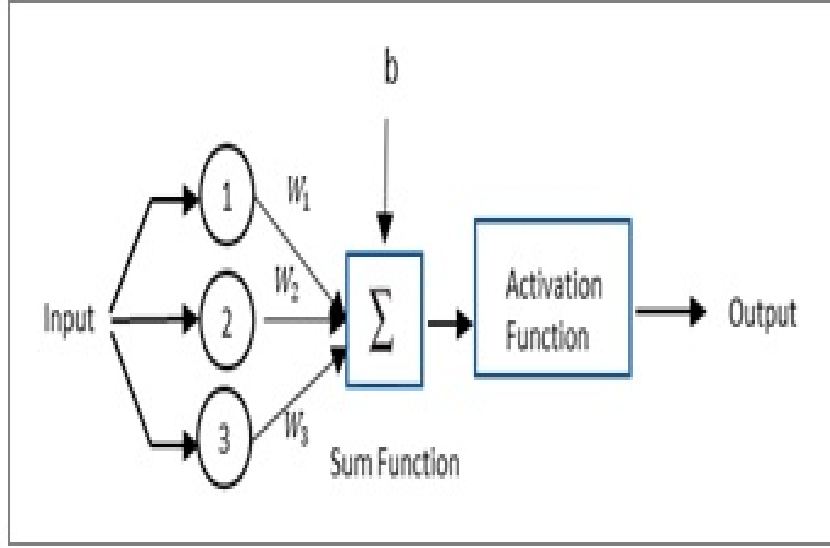


Figure 4.2 The architecture of MLP network.

4.3.4 Evaluation Metrics

Intending to compare the performance of different algorithms, we used various evaluation metrics (e.g., Receiver Operating Curve (ROC) curve, Precision-Recall (PR) curve, calculating the execution time, etc.) [111]. Moreover, we tried four other classification metrics including *Precision*, *Recall*, *F1 – score* and *Accuracy* which are calculated, respectively, by (4.3), (4.4), (4.5) and (4.6) [101].

$$Precision = \frac{T_P}{T_P + F_P} \quad (4.3)$$

$$Recall = \frac{T_P}{T_P + F_N} \quad (4.4)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.5)$$

$$Accuracy = \frac{T_P + T_N}{(T_P + T_N + F_P + F_N)} \times 100, \quad (4.6)$$

4.4 Simulation Results

In the VANET network, vehicles have different wireless communications like V2V and V2R, where we consider the *sending packets* as the parameter for these communications. We assume traffic and non-traffic situations in the network based on packet receiving and compare the results of the five selected algorithms (i.e., KNN, RF, SVM, NB and MLP) for predicting it in V2R communication. Then, by investigating the performance of each algorithm, we find the most performed algorithm for network traffic prediction in classification problems.

To measure the performance of algorithms in classification problems, we used one of the most prevalent graphical tools called the Receiver Operating Characteristics (ROC) curve. Then, we calculated the Area Under the Curve (AUC), where the large size of this area is a good indication of the performance of each algorithm. Different probabilities of prediction are shown by the ROC curve, where the False Positive Rate (FPR) indicates on the X-axis and the True Positive Rate (TPR) on the Y-axis. The nearest curve to the top left corner, related to the biggest AUC, has the best performance in prediction. Fig. 4.3 depicts the performance of the five different algorithms based on the ROC curve and the calculated AUC.

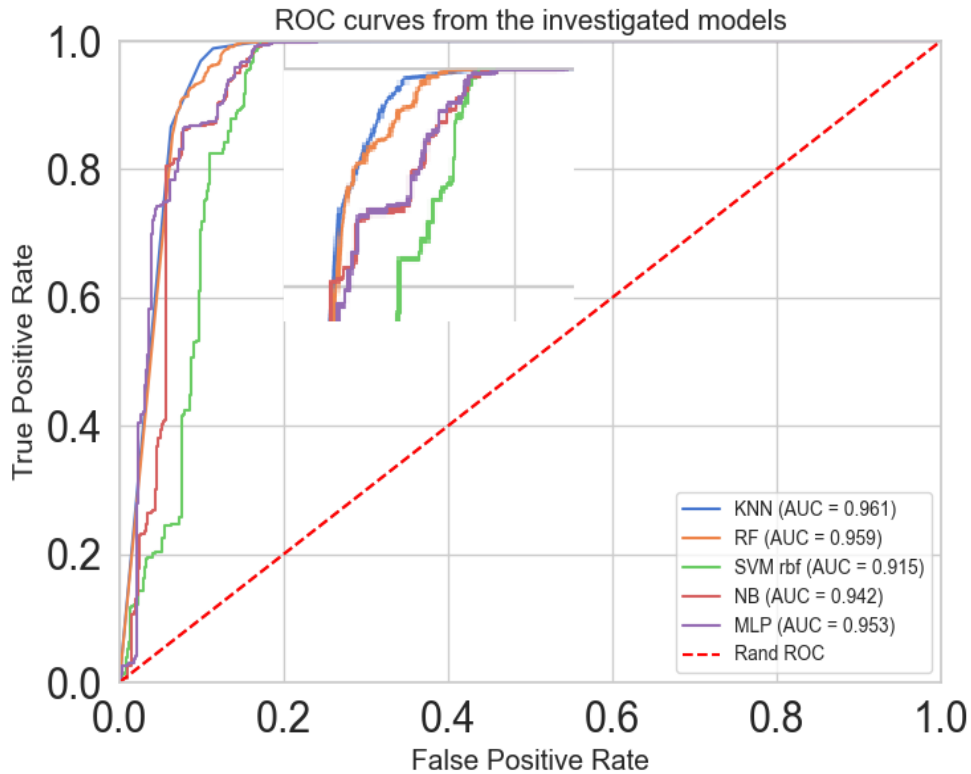


Figure 4.3 ROC curve for KNN, RF, SVM, NB and MLP.

The red dotted line, which separates the whole area into two parts, is a classifier that acts on a random level. The curves located above it have better performance than the ones that are placed under it. As shown in Fig. 3, the blue and orange curves that belong to KNN and RF, respectively, are the nearest curves to the top left corner and have the biggest AUC that shows their satisfying performance. The SVM, which is represented by the green curve, performs worse than all other algorithms. The NB and MLP, which are illustrated by the red and purple curves, respectively, have almost the same performance, specifically when TPR is ranging from 0.8 to 1.0.

There are other metrics like *accuracy*, *precision*, *recall* and *F1 – score* that can get more insight into the performance of the different algorithms for network traffic prediction and we calculated them, as illustrated in Table 4.1.

Table 4.1 Classification reports

Algorithm	Accuracy (%)	Label (packet receiving)	Precision	Recall	F1-score
KNN	97	No	0.95	0.89	0.92
		Yes	0.97	0.99	0.98
RF	96	No	0.99	0.81	0.89
		Yes	0.95	0.99	0.97
MLP	96	No	0.99	0.81	0.89
		Yes	0.95	0.99	0.97
SVM	95	No	0.87	0.88	0.87
		Yes	0.97	0.96	0.97
NB	89	No	0.68	0.88	0.77
		Yes	0.97	0.89	0.93

The KNN with 97% has the highest accuracy, the RF and MLP have the same accuracy with 96%, and the SVM and NB with 95% and 89% have the lowest accuracy accordingly.

Furthermore, we used another evaluation metric name precision-recall (PR) curve to show the balance between precision and recall, which are pointed out at the Y-axis and X-axis, respectively. We compared the PR curves of the five selected algorithms, as illustrated in Fig. 4.4.

The red dotted line, which is displayed horizontally, is a baseline classifier for more insight into the performance of algorithms in a way that the algorithms placed above this line and near the right top corner have better performance.

The comparison results present that the KNN (i.e., the blue curve) and RF (i.e., the orange curve) algorithms have better performance. The NB algorithm, which is related to the red curve that even crosses the baseline, has the worst performance. The MLP algorithm (i.e.,

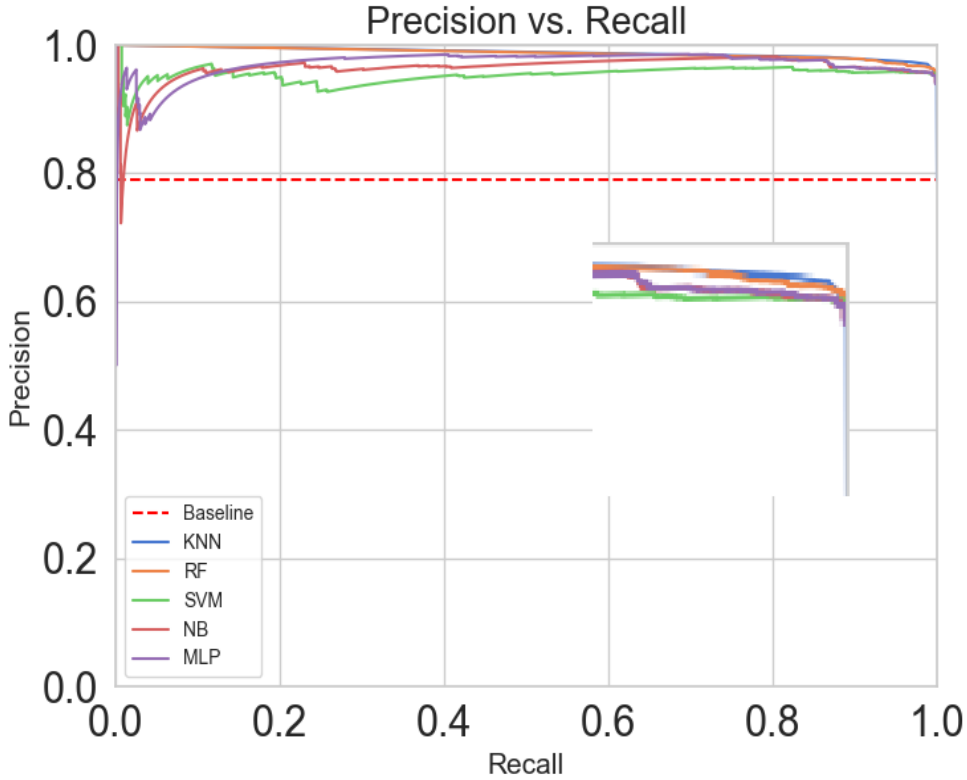


Figure 4.4 PR curve for KNN, RF, SVM, NB and MLP.

the purple curve) acts better than the NB and the SVM (i.e., the green curve) algorithms.

We tried different evaluation metrics for network traffic prediction in classification problems. However, there is another important parameter to find out the most performed algorithm which is the execution time. Based on the growth of the volume of data in wireless communications, the time that each algorithm takes to be executed could be a challenge. Though, we calculated the execution time for each of the five selected algorithms, as shown in Fig. 4.5.

The KNN algorithm, which has the best performance in other evaluation metrics, is considered the most time-consuming algorithm for network traffic prediction. As depicted in Fig. 5, the KNN takes 6.47 minutes, which is the highest time compared with other algorithms. Therefore, it could not be the most performed algorithm for our prediction. The SVM and MLP take 3.95 and 3.2 minutes, respectively. The NB is the fastest algorithm with 0.07 minute, but in other evaluation metrics, it has poor results. Finally, the RF has satisfying results as the KNN in other evaluation metrics. It takes 0.73 minute as execution time, which is an acceptable result for prediction.

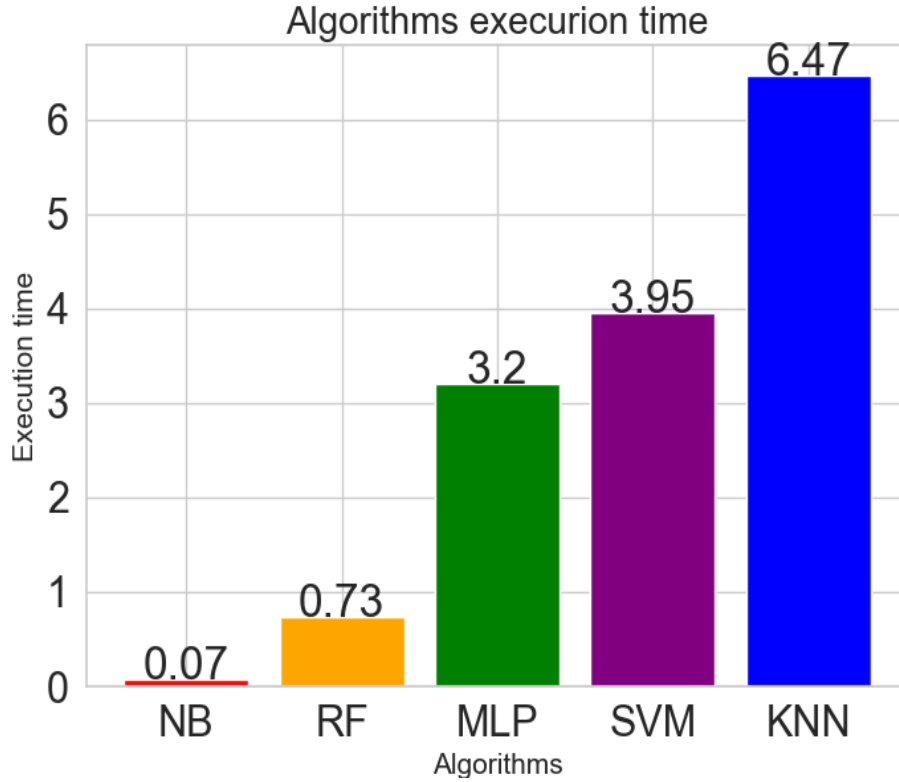


Figure 4.5 Execution time for network traffic prediction using KNN, RF, SVM, NB and MLP.

Therefore, based on all the results obtained from different evaluation metrics to compare the performance of different traditional machine learning algorithms and one simple neural network algorithm for network traffic prediction, the RF is the most performed algorithm in terms of accuracy, execution time and prediction error.

4.5 Conclusion

In this paper, we investigated the performance of four traditional machine learning algorithms (i.e., KNN, RF, SVM and NB) and tried MLP as a neural network algorithm for network traffic prediction in VANET. We used a real dataset based on V2R communication and we implemented these five algorithms to find the most performed one for classification problems. We considered packet receiving as a network traffic parameter and we assumed in a non-receiving packet situation that traffic is happening, and when the packet is received, there is no traffic in the network. We analyzed the performance of various evaluation metrics, such as ROC curve, PR curve, classification report and execution time. The evaluation results show that the KNN with a slight difference from the RF, performs better than other

algorithms, but its execution time for network traffic prediction remarkably takes higher time (6.47 minutes). It is enough that we do not consider the KNN as the best algorithm. The fastest algorithm is NB with 0.07 minute of execution time, but it performs weakly in other evaluation metrics. The MLP algorithm gives better results in evaluation metrics than SVM, even in execution time they take 3.2 minutes and 3.95 minutes, respectively. Based on all evaluation metrics, we can conclude that the RF is the most performed algorithm for network traffic prediction in VANET with a good execution time (0.73 minute) and reasonable results in other evaluation metrics.

CHAPTER 5 ARTICLE 3: FED-NTP: A FEDERATED LEARNING ALGORITHM FOR NETWORK TRAFFIC PREDICTION IN VANET

Sanaz Shakersepasgozar, Samuel Pierre

Mobile Computing and Network Research Laboratory (LARIM)

Department of Computer and Software Engineering, Polytechnique de Montréal

Montreal, QC H3T 1J4, Canada

Email: {sanaz.shaker-sepasgozar, samuel.pierre}@polymtl.ca.

Status: published in *IEEE Access*, 14 November 2022.

Abstract

During the last years, the volume of data produced in smart cities has been growing up, which can cause network traffic. Some of the challenges in an Intelligent Transportation System (ITS) are predicting the network traffic with the highest accuracy, keeping the security of data and being less complex. Artificial Intelligence (AI) algorithms are advantageous solutions to predict, control and avoid network traffic. However, such algorithms brought some costs to the privacy field. Accordingly, besides having an accurate prediction, preserving the privacy of data is an important challenge that should be considered. To cope with this problem, we propose a Federated learning algorithm for Network Traffic Prediction (Fed-NTP) based on Long Short-Term Memory (LSTM) algorithm to train the model locally, which can predict the network traffic flow accurately while preserving privacy. We implement the LSTM algorithm in a decentralized way by using the federate learning (FL) algorithm on the Vehicular Ad-Hoc Network (VANET) dataset and predict network traffic based on the most influential features of network traffic flow in the road and network. Simulation results reveal that the proposed model besides preserving the privacy of data, takes an obvious advantage over other well-known AI algorithms in terms of errors in prediction and the highest $R^2 - SCORE$ (0.993).

Keywords: Vehicular Ad-Hoc Network, federated learning, network traffic prediction, deep learning, Artificial Intelligence, Intelligent Transportation System

5.1 Introduction

The wireless communication method has swift development in smart cities and Vehicular Ad-Hoc Network (VANET) is one of the major parts of this kind of communication in Intelligent

Transportation System (ITS). The basic structure of VANET is made of vehicles, Road-Side Units (RSUs) and their communication, which are Vehicle-to-Vehicle (V2V) and Vehicle-to-Road Side Units (V2R). Since the number of vehicles and smart devices is growing up, predicting the network traffic is always a critical issue including challenges, such as accuracy in prediction, the volume of data and data privacy.

Accordingly, AI algorithms are the appropriate solution with high accuracy for predicting problems while we are facing a big and complicated dataset. In the past few years, various types of deep learning (DL) algorithms have been proposed and some of them include Recurrent Neural Network (RNN) [69], LSTM [72] and Gated Recurrent Unit (GRU) [73] found more suitable for time-series problems. These kinds of algorithms are able to learn long dependencies for a long period of time to predict traffic flow. However, the main challenge is that they should be able to preserve the privacy of data in transferring process between users and servers.

The implementation of AI algorithms relies on accessibility to the data for creating models with the aim of predicting, controlling or avoiding traffic in smart cities. The availability of data in a centralized way for different servers can lead to data leakage. Consequently, the AI algorithms should be implemented in a refined way where there is no need to transfer the local data in the network to train the model. For this purpose, the federated learning (FL) algorithm was proposed for the first time in 2016 by Google [127], [128] [129]. In this way, the distributed data will keep in devices and the local data does not transfer between different users and servers. FL is able to use data from various organizations to train the model by using DL or machine learning (ML) algorithms in a way that just the model will be transferred [130]. This leads to keeping the privacy of data and preventing leakage. In fact, instead of sending data to the centralized location, the AI models are transmitted to the locations where the local data is placed.

Nowadays, there are various types of modern distributed devices (e.g., wearable devices, mobile phones, autonomous vehicles, etc.) that are generating valuable data. However, storing data in a local place and keeping the training process at the edge must become a vital mission [131]. FL algorithm is proposed as a developing technology that is distributed with the concept of privacy-preserving. This algorithm makes us unnecessary to share private data over the network. In this process, the ML models can create connections to the data without access to their location or the local data. Therefore, they can overcome the disadvantages of centralized models, which require training models locally and need the availability of data.

In this paper, we propose the Fed-NTP algorithm with the aim of privacy-preserving of data for network traffic prediction considering road and network parameters. We implement the

LSTM algorithm for training the local models then apply the FL algorithm to keep the data in virtual clients and transfer the model to the server for network traffic prediction. The proposed Fed-NTP model has been implemented on a real dataset got by Global Positioning System (GPS) based on V2V and V2R communications in the VANET network [100]. Our main goal is to implement an algorithm that can keep the privacy of data along with high prediction accuracy. The major contributions of this work are resumed as follows:

- We propose a decentralized deep learning Fed-NTP model to predict network traffic flow. The main point of the proposed model is that there is no need to send data to the server and just the trained model will be communicated. Implementing the LSTM algorithm in a federated way can ease the way for training the models in local locations without moving data to a central server. In fact, we do not bring the data to the model although, the model will be sent to the location where the data is placed.
- We focus on predicting network traffic flow in VANET and we assume “sender speed” as a traffic parameter in V2R communication and consider a threshold to detect the network traffic. We go beyond the network parameters for traffic prediction in the network and consider the road parameters that can affect network traffic flow. Therefore, we consider the “receiver speed” as a road parameter in V2V communication and take advantage of it to predict the “sender speed” for network traffic flow. Then, we compare the prediction results with different centralized and decentralized algorithms besides the proposed model Fed-NTP.
- Finally, a real VANET dataset is tested to analyze the outcome of the proposed model. Based on calculating different evaluation metrics in network traffic prediction and comparing the results with other baseline algorithms, the experimental results revealed the accurateness of the proposed Fed-NTP model with the lowest error in prediction while data privacy is preserved.

The innovation point of this work lies in network traffic prediction considering road and network conditions in a secure way to keep the privacy of data. By implementing the LSTM algorithm which is one of the most accurate DL algorithms in regression problems in a decentralized way by employing FL learning, we can predict network traffic flow based on road and network parameters securely and privately.

The rest of the paper is organized as follows. Section II provides a review of previous research that implemented FL and AI algorithms for traffic prediction. Section III introduces the proposed model and describes the implementation. The evaluation results and experimental

validation are shown in Section IV. Finally, in Section V, we conclude this paper and address future research.

5.2 Background and Related Work

Traffic prediction is a critical issue that so many researchers have implemented different AI algorithms including DL and ML algorithms to get efficient results. In the past few years, FL was deployed in prediction problems in different fields such as healthcare and transportation. Also, so many researchers implemented the FL algorithm to cope with privacy-preserving issues. For instance, transfer learning has been implemented in [132] for traffic prediction, the requirements for implementing FL with the aim of traffic estimation were proposed in [133], a selective model aggregation based on FL for classification problem was proposed in [134], Privacy-preserving blockchain-based FL was proposed in [135] for predicting traffic flow and other hybrid methods based on FL were proposed in [136, 137].

A novel wireless traffic prediction framework based on FL was proposed in [138] named Dual Attention-Based Federated Learning (FedDA) to train a prediction model by multiple edge clients. They proposed a data-sharing strategy due to the heterogeneous nature of traffic data. The augmented traffic data has been transferred to the central server and then a quasi-global model is got and shared between all base stations (BSs). Then they clustered the BSs into different groups based on their traffic pattern and information's geolocation. They used two real-world big datasets [139], [140]. They compared their results with five other baselines such as Lasso: A linear model for regression, Support Vector Regression (SVR) [141], LSTM [142], FedAvg [128] and FedAtt [143] which the first three of them were based on centralized training and the last two one was trained based on FL which is a decentralized algorithm. The experimental results based on mean squared error (MSE) and mean absolute error (MAE) as two regression evaluation metrics, showed that the proposed FedDA overcame the other baselines.

The authors in [144] proposed an FL-based gated recurrent unit neural network algorithm (FedGRU) for road traffic prediction and applied FL with the aim of privacy-preserving. A secure parameter aggregation mechanism has been employed to train the global model whose main duty was aggregating the local trained models in the cloud for road traffic prediction. Then, to improve the performance of the network traffic prediction model, they applied an ensemble clustering-based FedGRU. They implemented the proposed model in a dataset derived from PeMS [145] and compared their result with centralized algorithms including GRU [146], SAE [84], LSTM [147], and SVM [148]. The experimental results showed that their proposed decentralized model performed better than the centralized baselines.

A Federated Deep Learning based on the Spatial-Temporal Long and Short-Term Networks (FedSTN) algorithm was proposed in [149] to predict the traffic flow based on historical traffic data. The proposed algorithm has three components: 1) Recurrent Longterm Capture Network (RLCN) module whose task is capturing short and long spatiotemporal features, 2) Attentive Mechanism Federated Network (AMFN) module that is responsible for sharing the short-term spatio-temporal hidden information and trained based on Vertical Federated Learning (VFL) and 3) Semantic Capture Network (SCN) module that is employed to take some features like non-Euclidean connections and Point of Interest (POI). They compared their results with different centralized and decentralized proposed approaches, and the experimental results showed that the proposed algorithm can achieve higher prediction accuracy.

Yuan et al. [150] proposed an FL framework for traffic state estimation (TSE) called FedTSE. They designed their proposed model based on the LSTM algorithm for training the local models to predict vehicular speed. Then, to deal with resource limitations, the authors proposed deep reinforcement learning (DRL) to download/upload the model parameters. They consider three communication modes for TSE including FedTSE-Syn with identical numbers for training epochs, FedTSE-Asyn with a different number of training epochs for each RSU, and FedTSE-Asyn (Weight) that consider the penetration of training epoch. The authors set eight RSUs to compare the prediction performance of the proposed model. The MSE metrics have been used to evaluate the FedTSE and the results proved that the proposed model can reduce the error in comparison with other models.

Moreover, there is some other research for traffic prediction that implemented different AI algorithms with high accuracy in prediction. Sepasgozar et al. [120] proposed a model named Random Forest-Gated Recurrent Unit-Network Traffic Prediction algorithm (RF-GRU-NTP), which is able to predict the network traffic flow considering the parameters that can affect road traffic as well. They implemented their model in the VANET network and divided their work into three phases. In the first phase, they predicted the network traffic based on V2R communication. In the second phase, they predicted road traffic based on V2V communication. In the third phase, they found the most effective parameters in traffic by using random forest (RF). Then, by implementing the GRU algorithm, the authors predicted the network traffic flow based on road and network parameters. Moreover, they compared their proposed model with LSTM and Bidirectional-LSTM (Bi-LSTM) algorithms. The experimental results proved that the proposed model gives the minimum prediction error and execution time.

Yang et al. [151] proposed a new method based on SA (Simulated Annealing) optimized ARIMA-BPNN (Autoregressive Integrated Moving Average model-Back Propagation Neural

Network) for traffic prediction. The authors used historical network traffic data to evaluate the proposed prediction method by using some metrics including MAE, Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE). The obtained results proved that the proposed method overcomes the traditional network traffic prediction.

Table 5.1 represents a summary of the related work regarding the existing methods for traffic prediction.

Table 5.1 Summary of Related Work in Traffic Prediction

Proposed method	Contribution	Reference
FedDA	Wireless Traffic Prediction based on FL	[138]
FedGRU	Privacy preserving algorithm to traffic flow prediction	[144]
FeDSTN	Traffic flow prediction based on historical traffic data	[149]
FedTSE	Vehicular speed and traffic flow prediction in FL way based on LSTM	[150]
RF-GRU-NTP	Network traffic flow prediction based on road and algorithms network parameters utilizing AI	[120]
ARIMA-SA-BPNN	Network traffic prediction based on historical data	[151]

Despite all the research done, predicting network traffic flow considering two types of communications (i.e., V2V and V2R) and considering road and network parameters while keeping the privacy of data, motivated us to propose an accurate algorithm that can preserve the privacy of data for network traffic prediction.

5.3 Methodology

In this section, we explain the proposed Fed-NTP model in the VANET environment in detail. The existence of dynamic characters (vehicles) in the VANET is the main reason for making it complicated for prediction problems. On the other hand, DL algorithms are promising solutions in terms of predicting complex patterns [120]. However, there are still factors such as effective parameters in network traffic and privacy of data that can help us make an accurate prediction while the privacy of data is kept. Fig. 5.1 shows the architecture

of the VANET environment and model training in centralized learning where the local dataset will be sent to the server.

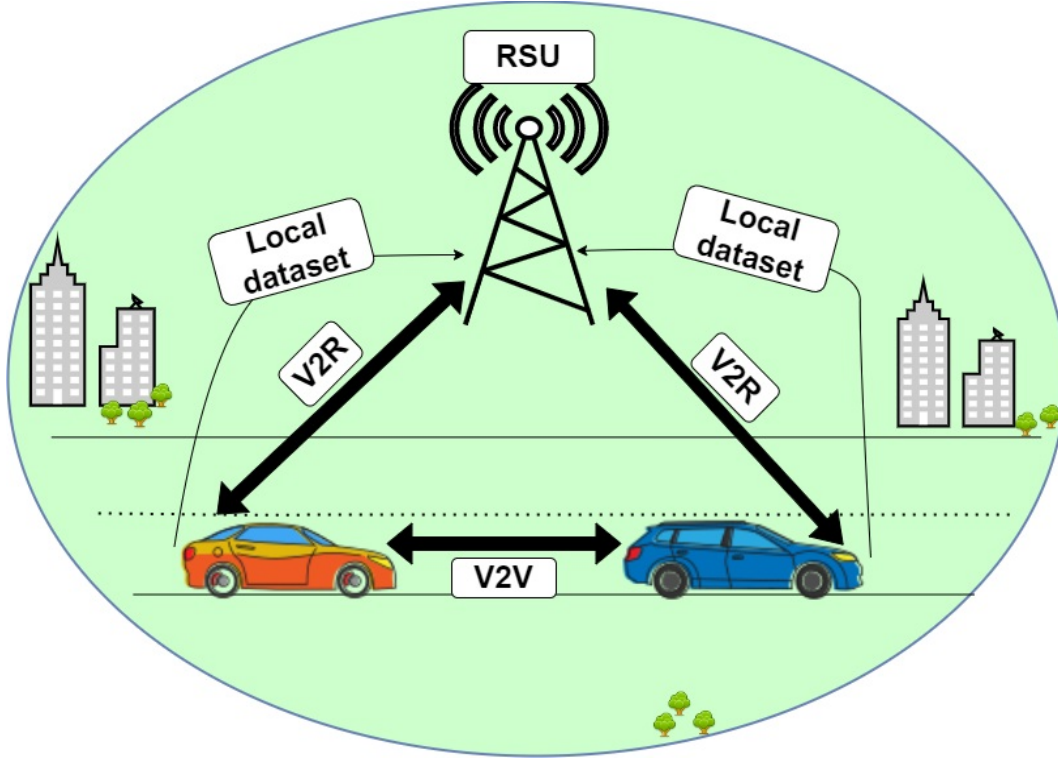


Figure 5.1 VANET architecture and training in centralized learning.

Sending the local data to the server might lead to a risk of information leakage. While decentralized algorithms (i.e., FL) do not need the local data for the training process, the model will send the local data to the server. Therefore, we implement the LSTM algorithm with the aim of having high accuracy in a federated way in terms of keeping data locally and just sending the model to the server.

5.3.1 LSTM Algorithm

The LSTM is one type of RNN algorithm which can learn dependencies in prediction problems. It consists of three gates including input, output and forgets gate. Fig. 5.2 shows the fundamental structure of the LSTM algorithm.

The equations of the gates are as follows:

$$\text{Forget gate:} \quad f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5.1)$$

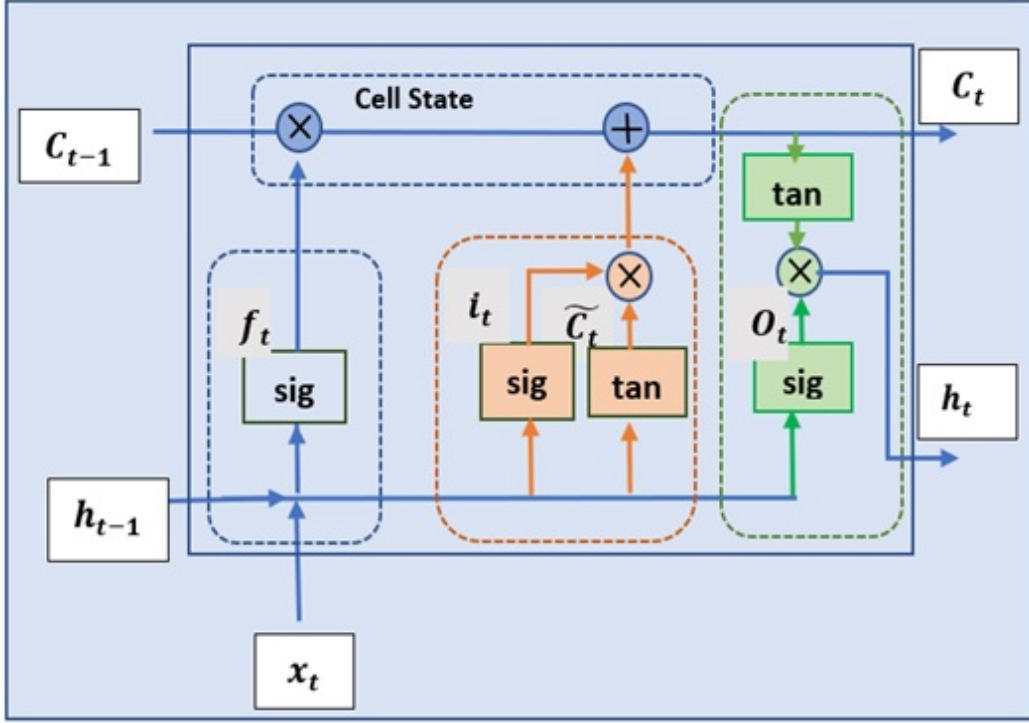


Figure 5.2 Structure of LSTM algorithm.

$$\text{Input gate:} \quad i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (5.2)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (5.3)$$

$$\text{Cell state:} \quad C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5.4)$$

$$\text{Output gate:} \quad o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5.5)$$

$$h_t = o_t \tanh(C_t) \quad (5.6)$$

Deciding about which information is needed to be kept or ignored, is under forget gate responsibility and has been calculated in (5.1). For timestep t , x_t is the input and h_{t-1} represents the hidden state that will be passed through the sigmoid function. The weight matrix between forget gate and the input gate is represented by W_f and the connection bias at t is represented by b_f . The input gate is calculated in (5.2, 5.3) and is responsible to determine what information is relevant to add from the current step. In (5.2), i_t represents

the input gate at t , W_i represents the weight of sigmoid between input and output gate, $\tilde{C}_t$ represents the weight matrix of $\tanh$ function between information in cell state, the output part is represented by W_o , and the b_c represents the bias vector. Storing information from the new state will be done in the cell state, which is calculated in (5.4) where C_{t-1} represents the previous time step. Finally, the value of the next hidden state will be chosen by the output gate that is calculated in (5.5, 5.6), where o_t represents the output gate at t , W_o represents the weight matrix of output gated, b_o represents the bias vector and h_t represents the LSTM output [152].

In this paper, we implemented the LSTM algorithm for a local training process to keep the data locally and create a model in the client (i.e., vehicles) and sending the model to the server.

5.3.2 Federated Learning

Traditional machine learning algorithms or other centralized algorithms need to transfer local data to a central server for implementation. However, FL algorithm performs the training of the models locally without transferring data over the network, and it can be implemented on a server in a distributed way. This ability allows FL to overcome the drawbacks of other centralized algorithms. Moreover, transferring the trained model instead of local data reduces the usage of the bandwidth, decreases energy consumption and increases privacy [153]. Fig. 5.3 shows the sequence diagram for the FL workflow.

In the first step, the server by using random parameters, will create a neural network model and initialize it. The new model will be sent to all clients, and when the clients receive it, they will start the training and testing process on the local data. After that, the server will send a request to the clients for getting their trained local model and when it is done, the first round of FL has been completed. Then, the server will aggregate all the trained local models, create a new global model and send it back to the clients [154].

5.3.3 Proposed FeD-NTP

In this section, we introduce the detail of the proposed Fed-NTP model, which is planned to predict network traffic flow considering road and network parameters, while the privacy of data is preserved. For this purpose, we used the VANET dataset [100] which includes V2V and V2R communications. To define the network traffic, the “sender speed” as a network parameter in V2R communication and the “receiver speed” as a road parameter in V2V communication, are considered. We assumed when the “sender speed” drops below

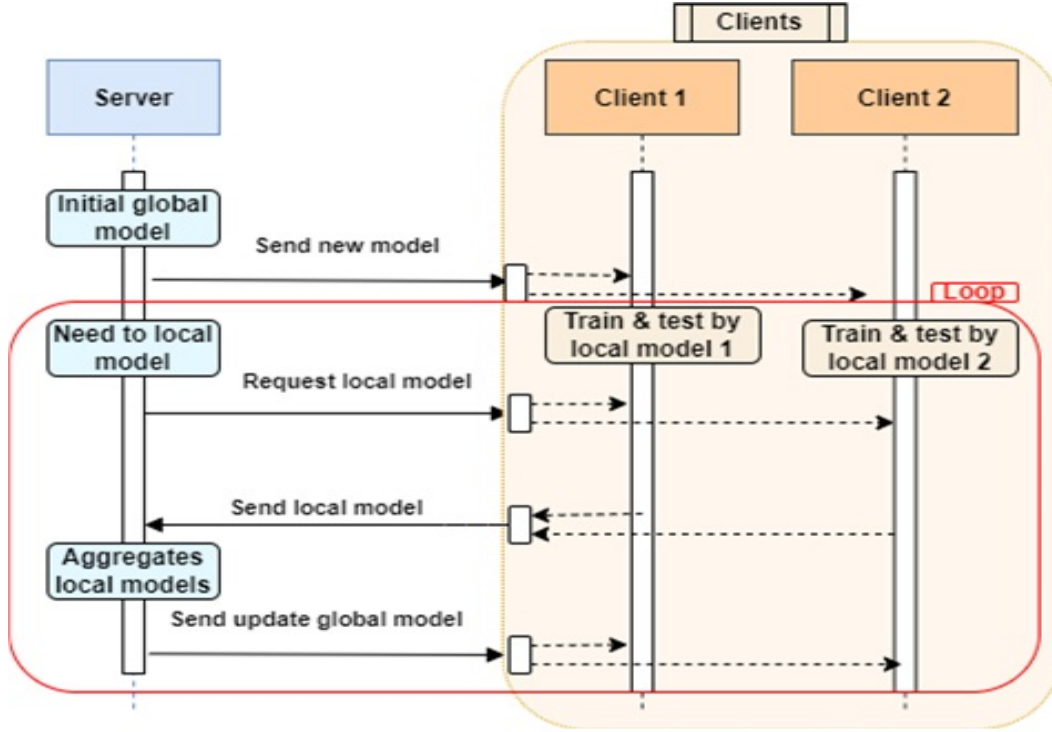


Figure 5.3 Sequence diagram for federated learning.

60 Km/h, then the traffic will occur in the network considering that the speed range of all vehicles in the dataset is between (0, 104 Km/h) without any speed limitation. Considering this assumption, we predict “sender speed” in V2R communication while we consider the “receiver speed” in V2V communication in our prediction.

However, besides having accurate network traffic prediction, we intend to keep the data locally and just the model will transfer in the network and send to the server. The architecture of the proposed Fed-NTP model is shown in Fig. 5.4.

In the first place, the initial model based on random parameters which are created by the server or RSU in our work will send to the clients (step 1). Then the RSU needs the trained local model and not the local data which the FL algorithm will do the process. At this step, the clients that are vehicles in our case, start to do the train and test process. The past three values of “receiver speed” as road parameter and three past values of “sender speed” as network parameter have been taken for predicting the next value of “sender speed” and predicting the network traffic by implementing the LSTM algorithm. Implementing the LSTM as a powerful DL algorithm with high prediction accuracy can lead to improving the prediction results in the federated way as well. After the local model has been trained by the LSTM, it will be sent to the server and the local data will keep in the local place where is

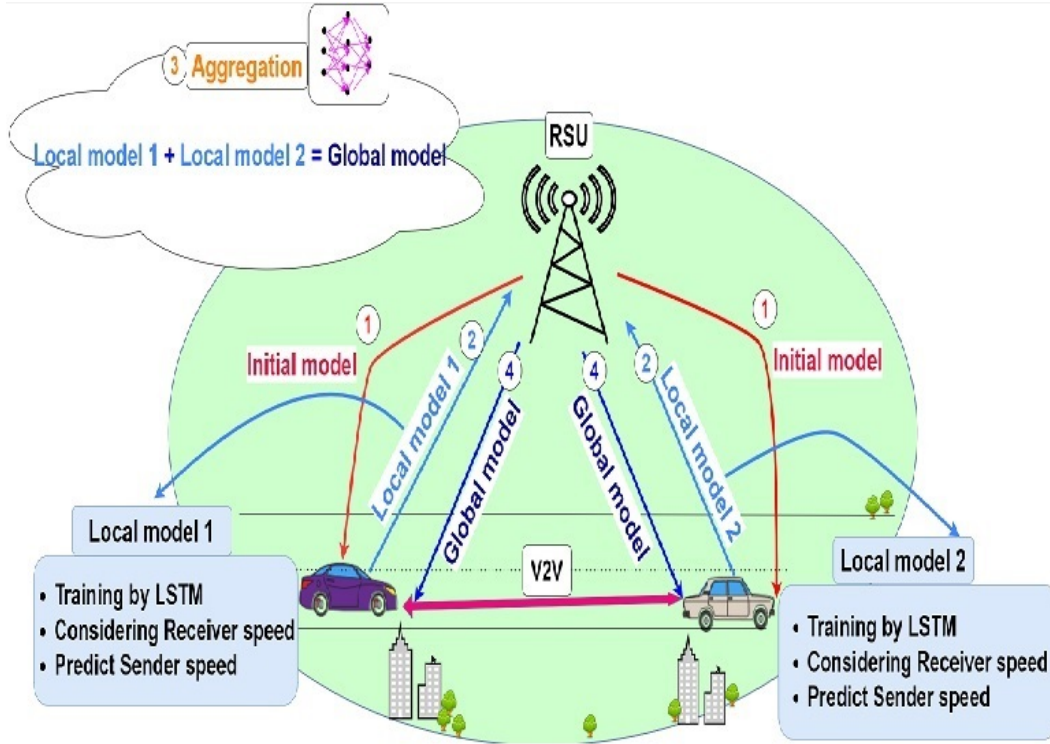


Figure 5.4 The architecture of proposed Fed-NTP model.

in the vehicles (step 2). After that, the RSU will aggregate all the local models which have been sent by vehicles and make a global model (step 3). Finally, the global model will send back to the vehicles to be trained again (step 4).

During this process in the VANET environment, each vehicle plays the role of client or worker node, which by implementing DL can train the dataset locally. The local dataset is split into some pieces and each client will receive one piece of it. Moreover, the clients do not transfer their datasets to each other. The proposed decentralized model has the ability to maintain the privacy of data by not distributing data over clients and across the network.

The experimental results show the superiority of the proposed Fed-NTP model in terms of the different evaluation metrics that we apply, and we go beyond getting high prediction accuracy by keeping the privacy of data.

5.4 Data Preparation and Performance Evaluation

5.4.1 Dataset

We adopted a real VANET dataset (802.11 ad-hoc network type) including two types of communication (i.e., V2V and V2R) to measure short-range communications on the highway [100]. The data is collected from the external antennas that were installed on the roof of vehicles. Some features, such as the longitude, latitude, speed and heading were reported by GPS every two seconds. The location of data gathering was from a highway in Atlanta with six lanes including five usual lanes and one High Occupancy Vehicle (HOV) lane that has been monitored during the day between 2 pm and 5 pm. The accuracy of the location information that got by GPS, was reported around five to seven meters that were conducted by interpolation.

The V2V communication was measured based on the following vehicles and the V2R experiments, got from moving vehicles and an RSU located on an elevated bridge. The number of packets in V2R communication was 1470 bytes, which were broadcasted by the senders at an approximate rate of 150 packets/s [111].

We used the “sender speed” as a network traffic parameter from V2R communication and we consider the “receiver speed” from V2V communication as a road traffic parameter, which can affect traffic in the network and predict network traffic flow with high accuracy while data privacy is preserved too.

5.4.2 Implementation

We used python version 3.7 [112] and PyTorch [155] to implement the DL algorithm, for local training in the client. To design the FL algorithm, one extension library of Python called PySyft [156] has been used. PySyft can provide the requirements of FL algorithms and can be used by some main DL frameworks namely PyTorch and TensorFlow [104] that can be used for private and secure DL algorithms.

The implementation occurred in the Google Colab platform [157] while the GPU has been used as a hardware accelerator to improve the processing. For the preprocessing step, after data cleaning, we normalized the data using StandardScalar, then we split the dataset based on the segmentation method, which is commonly used, into 80% for the train set and 20% for the test set. After that, we defined two virtual clients for implementing the FL algorithm and the LSTM has been implemented for the local training process in the clients to create local models for sending to the RSU.

For the model evaluation, we choose MSE loss that shows the deviation between predicted values and the actual one in the regression problem. The SGD [105] has been used as an optimizer in our work, which is an adaptive optimizer for FL algorithms. Moreover, during the implementation for finding the most efficient hyperparameters, we used a checkpoint that made us set the number of epochs = 200, learning rate $\alpha = 0.1$, batch size = 128, for the proposed Fed-NTP and baseline algorithms, and eight hidden layers considered in LSTM implementation.

5.4.3 Evaluation Metrics

We adapted five different regression evaluation metrics including MAE, MSE, RMSE, R^2 – *SCORE* and MAPE to evaluate the prediction accuracy as follows [109]:

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |x_i - \hat{x}_i| \quad (5.7)$$

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (5.8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (5.9)$$

$$R^2 = 1 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - x_{avg})^2} \quad (5.10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right|, \quad (5.11)$$

where x_i presents the real value, $\hat{x}_i$ indicates the predicted value, x_{avg} presents the score of averaged weighted and n indicates the number of tests or verification set. The proposed Fed-NTP model can achieve the best results in terms of used evaluation metrics mentioned above.

5.4.4 Experimental Results

We compared the performance of the proposed Fed-NTP model with three centralized and one decentralized algorithm including LSTM [147], GRU [146], RNN [81] and FedGRU [144] respectively. The performance evaluation was based on five different evaluation metrics

with the aim of network traffic flow prediction influenced by “sender speed” as a network traffic parameter and “receiver speed” as a road traffic parameter. Among these competing methods, the experimental results revealed that the proposed Fed-NTP can overcome the baseline models in terms of accuracy in prediction and preserving privacy.

Table 5.2 shows the evaluation performance of different algorithms based on various metrics.

Table 5.2 Models Evaluation

Type	Algorithm	Best epoch	RMSE	MSE	MAE	MAPE	$R^2 - SCORE$
Centralized	LSTM	10/200	0.938	0.879	0.687	0.012	0.948
	RNN	77/200	0.645	0.427	0.489	0.009	0.975
	GRU	199/200	0.408	0.166	0.264	0.004	0.99
Decentralized	FedGRU	10/200	0.938	0.879	0.687	0.012	0.948
	Fed-NTP (Proposed)	199/200	0.342	0.117	0.161	0.003	0.993

All the algorithms were executed in 200 epochs, and each one got the best results in various epochs. The evaluation results revealed that the proposed Fed-NTP model takes a definite advantage over centralized and decentralized algorithms with the lowest error and the highest $R^2 - SCORE$. The LSTM algorithm has the worst results with the highest error in prediction and the GRU algorithm has lower error but not better than the proposed model.

The prediction results of the “sender speed” as a parameter in network traffic flow are shown in Fig. 5.5. The actual data is illustrated in the orange line and the predicted data is depicted in the green line. As shown in Fig. 5.5, the proposed Fed-NTP model performs more precisely than other algorithms with the lowest difference between the actual and predicted data.

Fig. 5.6 depicts the MSE loss for the LSTM, GRU, RNN, FedGRU and Fed-NTP algorithms. The MSE for the training data is represented in the blue curves and the test data is illustrated in the red curves. As shown in Fig. 5.6, the FedGRU experienced the less iteration and got the best results in the second epoch which is the lowest one, and the proposed model got the best results in the last epoch while it has less fluctuation during the process and better performance in comparison with other algorithms.

The results that we got from all the evaluation metrics show that the proposed Fed-NTP algorithm can achieve the lowest error in terms of network traffic prediction considering road

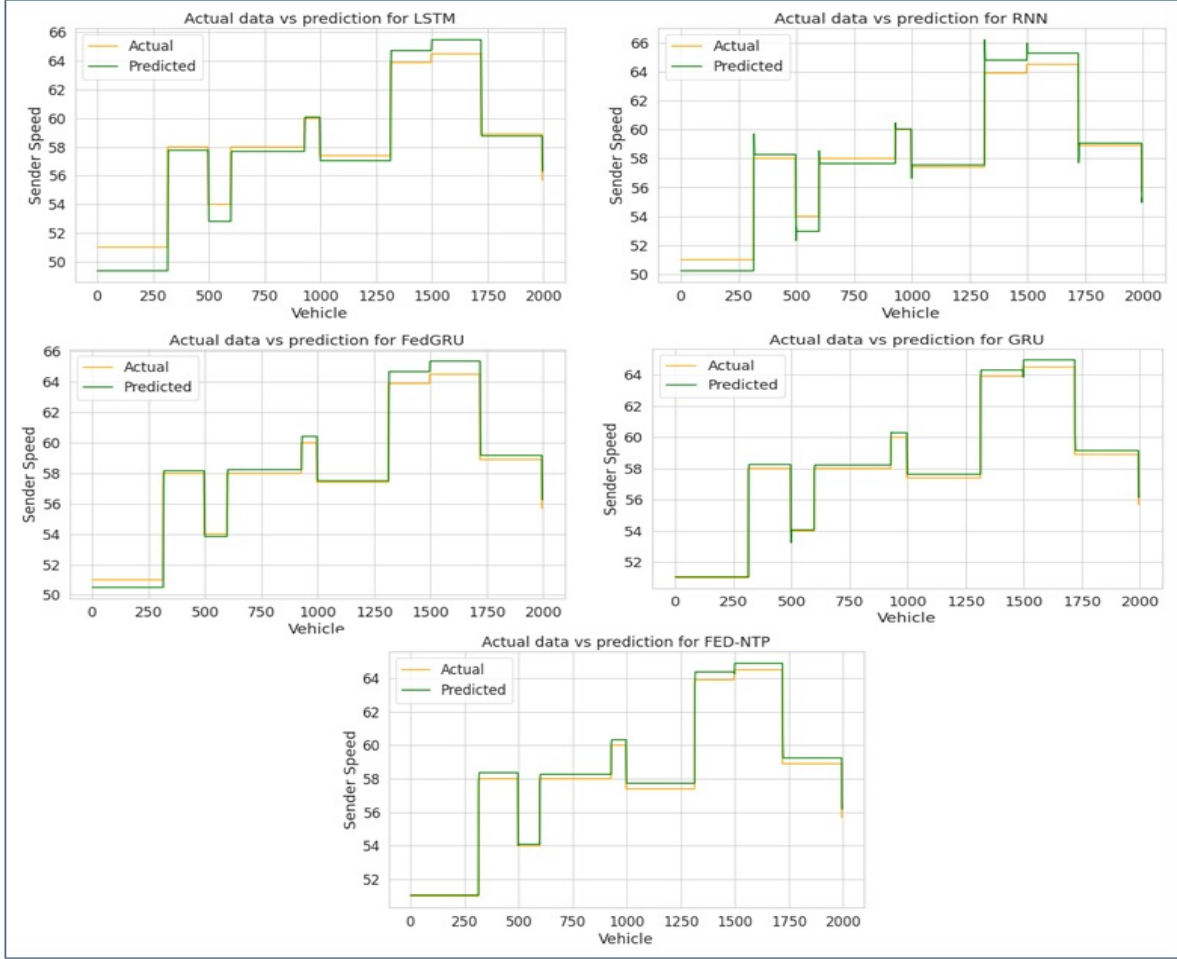


Figure 5.5 Network traffic prediction results for the LSTM, GRU, FedGRU and Fed-NTP.

and network parameters. Also, the simulation results revealed that the proposed algorithm is able to get the highest $R^2 - SCORE$ in comparison with the baselines.

5.5 Conclusion

In this paper, our primary purpose was to achieve the trade-off between the security of data and the accuracy of network traffic flow prediction considering road and network parameters altogether. We proposed the FED-NTP model to predict network traffic for the local training in the clients (vehicles). The LSTM algorithm has been implemented with the aim of a local training process and for keeping the data locally, we utilized the FL algorithm. Traditional machine learning algorithms need to transfer local data and a central server for implementation whereas in decentralized ML such as FL, there is no need to transfer the local data and just the trained model would be transferred, and it can be implemented on a distributed

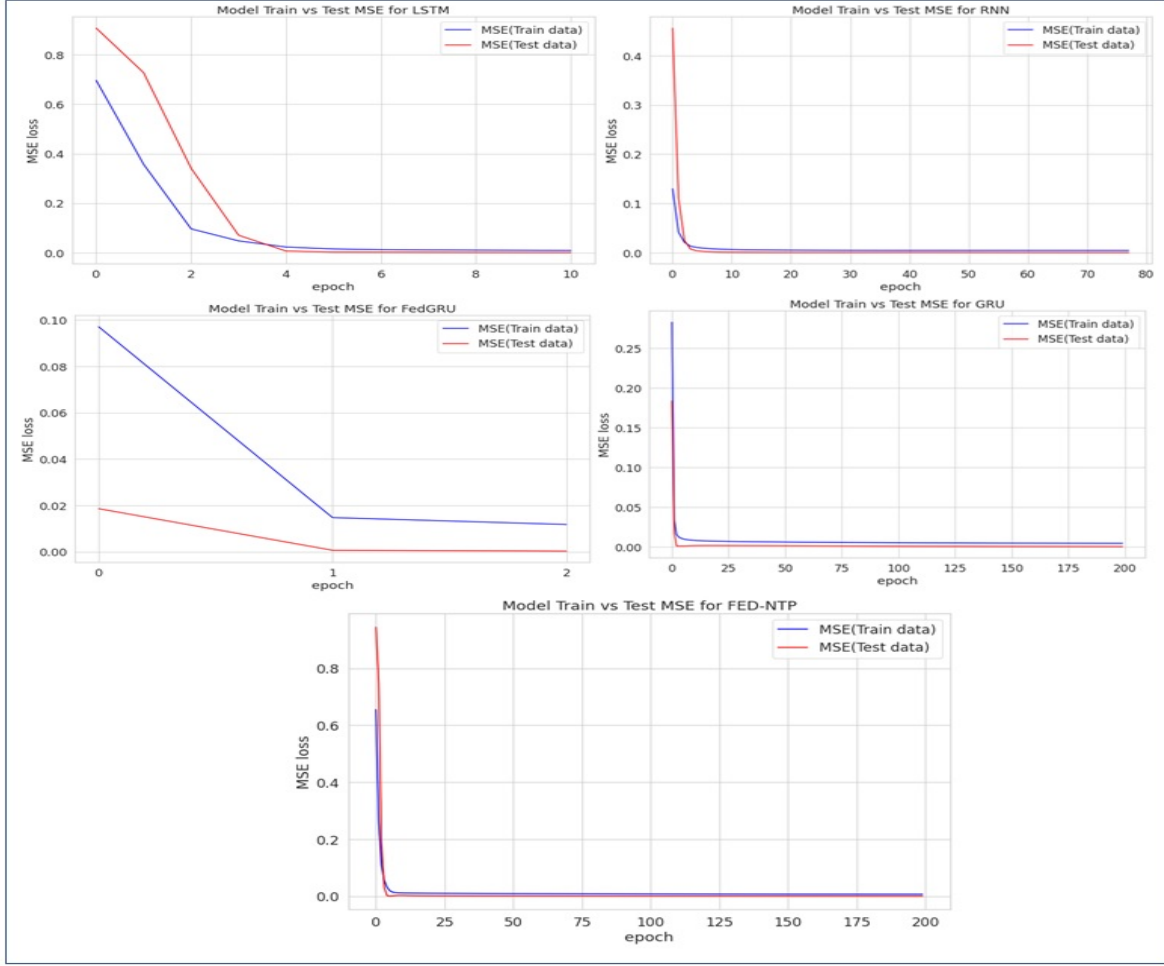


Figure 5.6 The MSE loss for the LSTM, GRU, RNN, FedGRU and Fed-NTP algorithms.

server. A real VANET dataset based on V2V and V2R communication has been used for network traffic prediction. The “sender speed” is considered as a network traffic parameter and we assumed that traffic has happened at a speed lower than 60 Km/h, and the “receiver speed” is considered as a road parameter and we monitored its effect on network traffic flow as well.

Besides the proposed Fed-NTP model, we implemented three centralized and one decentralized deep learning algorithm including the LSTM, GRU, RNN, and FedGRU respectively. The simulation results have been analyzed in terms of different evaluation metrics such as MAE, MSE, RMSE, MAPE and $R^2 - SCORE$. The results revealed that the proposed model performed more precisely and took explicit benefit over the baseline algorithms.

The PySyf library and its PyTorch extension have some complications to implement deep learning algorithms which were the main complexity of our proposed model. Implementing

deep learning algorithms in a federated way besides keeping their accuracy in prediction was another challenge in our work.

However, there are more opportunities to take advantage of implementing AI algorithms, in terms of network traffic prediction in different wireless networks such as the fifth generation (5G) and the sixth generation (6G). The new Generations of networks are able to improve the speed of communication and reliability, along with decreasing the delay in the network which in our future work we will consider 5G and 6G networks.

Acknowledgment

The authors would like to thank Dr. Franjeh El Khoury for her valuable comments and proofreading of this paper.

CHAPTER 6 GENERAL DISCUSSION

The proposed methods are discussed in this chapter. Traffic flow in the vehicular network can appear in two types including road traffic and network traffic. Road traffic can occur due to various environmental parameters such as weather conditions, the number of vehicles, accidents, road closures, the number of intersections and traffic lights and many more. The mentioned road traffic parameters which cause road traffic happening can lead to greenhouse emissions, wasting time and fuel, increasing injury and death and etc. One of the most important research topics is road traffic prediction with the aim of controlling and avoiding traffic on the road. VANET as improved technology is able to assist ITS in order having wireless communications between vehicles and installed infrastructures on the roads.

The other type of traffic flow in a vehicular network is network traffic happening which has different effective network parameters. Network traffic detection is based on the performance of the network that some factors such as sender speed, throughput, packet loss, and delay can affect the quality of service in the network and decrease its performance. So much research and valuable works have been done to predict, control and avoid road traffic flow [158–160] and network traffic flow [81, 161, 162] separately. Various machine learning and deep learning methods have been proposed in these two fields independently. However, the effect of these two types of traffic flow on each other is an important issue that is not taken to account.

When road traffic happens due to the mentioned reasons, road and network traffic will occur in a chain. In that case, the speed of vehicles that are playing the role of packet senders will drop and the number of their communication in each unit of the region will increase. Then the network load will increase and all these happening can lead to delays in the network, packet loss, decrease in the QoS and so on. Although, considering road traffic and network traffic together, is a critical issue that is hidden in most views.

Therefore, implementing AI algorithms for traffic flow prediction in both types of road traffic and network traffic is a promising solution. Due to the growth of smart vehicles in ITS and the volume of data that is produced by them, analyzing the data is a big challenge considering that its security is also maintained. Most of the AI algorithms need the data to do the training process which means the data will transfer over the network which might increase the risk of data leakage. For solving this problem, some decentralized algorithms have been proposed that make it possible to transfer the trained model instead of the raw data. Due to the volume of produced data in smart cities, different aspects of security become important and the most important one is preserving the privacy and security of data.

In this dissertation, both traffic types including road traffic flow and network traffic flow are considered. Various network and environmental factors that can affect traffic flow have been investigated. Also, different AI algorithms have been implemented for traffic flow prediction. An intelligent Network traffic prediction model considering road and network parameters propose in Chapter 3. Several machine learning classification algorithms including random forest, K-Nearest Neighbors, Support Vector Machines, and Naive Bayes algorithms are implemented to predict network traffic flow based on packet receiving. The random forest got the best results in terms of prediction error. Then three different deep learning regression algorithms including Long Short-Term Memory, Gated Recurrent Unit and Bi-directional Long Short-Term Memory are implemented with the aim of road traffic flow prediction based on sender speed. The Gated Recurrent Unit algorithm shows the best performance in road traffic prediction based on regression evaluation metrics. Finally in the last step, By implementing the random forest, the most effective parameters in the road and network on network traffic flow are identified and after implementing LSTM, Bi-LSTM and GRU, the GRU is able to overcome other algorithms in terms of error prediction and execution time.

Investigation about the best AI algorithm in terms of accuracy and execution time for network traffic prediction is proposed in Chapter 4. Four various machine learning algorithms including random forest, K-Nearest Neighbors, Support Vector Machine, Naive Bayes and one class of feed-forward Artificial Neural Networks named Multi-Layer Perceptron are implemented. A real VANET dataset based on Vehicle to-roadside unit communication is used. Two types of traffic situations based on Packet receiving as a network traffic parameter are assumed. By utilizing different classification evaluation metrics, the random forest gets the best result in network traffic prediction in terms of accuracy and execution time.

Besides the accuracy of network traffic prediction, another important point is maintaining data security. A Federated learning algorithm for Network Traffic Prediction (Fed-NTP) proposes in Chapter 5. The LSTM algorithm is implemented to train the model locally and send it to the server instead of sending the local data. Network traffic flow prediction based on “sender speed” as a network parameter while the privacy of data is preserved, is the main target of the proposed Fed-NTP model. Utilizing a federated learning algorithm leads to keeping the data with the client which are vehicles in this work. Then the trained model will transfer over the network to network traffic prediction. Various centralized algorithms such as LSTM, RNN, GRU and a decentralized one named FedGRU [144] are implemented. the simulation results revealed that the proposed Fed-NTP model overcomes other baselines.

Data mining and analyzing it is a key part of prediction problems. In this dissertation, a real VANET dataset [100] is adapted. The data including V2V and V2R communications

collected from the installed antenna on the vehicles. Different parameters such as longitude, latitude, speed and heading were reported by Global Positioning System. In each chapter, different parameters and communications are considered for network traffic prediction.

CHAPTER 7 CONCLUSION

7.1 Summary of Contributions

Road traffic measuring is based on the number of vehicles on the road, the travel time, the speed of vehicles and many more environmental factors. From another standpoint, network traffic computing is based on the rate of packet loss, throughput, connectivity, sender speed, delay and more network parameters. However, the noteworthy point is that these two types of traffic can affect each other.

Artificial intelligence has made a great contribution to science. AI algorithms act as powerful tools which make machines learn like humans and provide the ability to predict through what they have learned. In the past few years, traffic flow in vehicular networks become a critical issue that AI has provided efficient algorithms for them. There are three mechanisms To improve the performance of the vehicular network including control, prediction and avoidance of traffic flow. The main body of this dissertation is based on predicting network traffic flow to prevent that from happening using AI algorithms. Since AI has so many branches, finding the fittest algorithm for different problems is an important issue that is discussed in this dissertation. Due to the increasing growth of smart devices in smart cities and the large amount of data that they produce, maintaining the security of data is a critical issue. The data generated by the smart devices, need to be transferred over the network to implement AI algorithms, which can increase the risk of data leakage. Security and privacy of data in network traffic flow forecasting that employs the proposed method is another contribution.

Network traffic prediction is a crucial issue in VANET which can be affected by various network parameters. On the other hand, the parameters which can lead to road traffic can also be effective factors in network traffic. From another point of view, the traffic prediction problem is proposed in both kinds of regression and classification types. According to both effective road and network parameters and both types of traffic prediction problems (regression and classification) an AI-based model by implementing machine learning and deep learning algorithm is proposed in this dissertation. The proposed RF-GRU-NTP model is including three phases which in the first phase different ML algorithms such as KNN, RF, NB and SVM are implemented for network traffic prediction. The network traffic flow is predicted as a classification problem in VANET. Various classification evaluation metrics including precision, recall, F1-score, and accuracy used. Additionally, other metrics such as ROC curve, confusion matrix and computation time have been used to evaluate the performance of algorithms. Based on all evaluation metrics the RF obtain an acceptable result in all

evaluation metrics. In the second phase, the road traffic flow by implementing three different DL algorithms such as GRU, Bi-LSTM and LSTM is predicted. The road traffic flow is predicted as a regression problem and various regression evaluation metrics such as MAE, MSE, RMSE and the $R^2 - SCORE$ have been used to evaluate the performance of implemented algorithms. The evaluation results showed that the GRU with the highest $R^2 - SCORE$ (0.995) and lowest prediction error (MSE= 1.402, MAE= 0.416 and RMSE= 1.184) has the best performance. Finally, in the third phase the proposed RF-GRU-NTP model, predict the network traffic flow considering the most effective parameters in both road and network. The prediction results were evaluated by regression metrics and compared with LSTM and Bi-LSTM algorithms. The results revealed that the proposed model has better performance in terms of execution time (0.33 minute) and prediction error.

Since the AI method has so many branches, finding the fittest algorithm, which can meet all the requirements for prediction is not an easy task. In other contributions of this dissertation, the performance of five common AI algorithms for network traffic prediction as a classification problem has been analyzed. Various evaluation metrics such as ROC and PR curves, Precision, Recall, F1-score, Accuracy and calculation execution time are used to evaluate the performance of implemented AI algorithms. The comparison result between RF, KNN, SVM, NB and MLP shows that the RF and KNN with a small difference, have the best performance. However, the KNN took the longest time to execute which was 6.47 minutes which can make it far from being the best algorithm. While the RF that got acceptable results in other evaluation metrics took 0.73 minutes for executing which can prove that it has the best performance in network traffic prediction.

Finally, implementing a deep learning algorithm in a secure way for network traffic prediction in VANET is considered in this dissertation. Regarding keeping the privacy of data, proposing a Fed-NTP model to predict network traffic flow by keeping the local data in the client places; then, implementing the LSTM algorithm in a federated way and sending the trained model over the network and to the server are other contributions in this dissertation. The proposed model is compared with three different centralized algorithms including LSTM, RNN, GRU and one decentralized algorithm called FedGRU and the result showed the proposed Fed-NTP algorithm overcomes other baselines.

7.2 Limitations

The greatest limitation of this dissertation is the lack of a new real VANET dataset. Since one of the major contributions of this work is network traffic prediction considering road traffic, a real VANET dataset including road and network information in the form of one

dataset could be very useful. However, an optimum solution was dividing the whole process into different phases based on road and network communication separately; then, by finding the most effective parameters, predict network traffic flow considering road factors.

Moreover, in order to network traffic prediction based on road traffic parameters, we had to divide the dataset into two parts including V2R for the network traffic part and V2V for the road traffic part. Using one dataset including two types of communications (V2V and V2R) for two different traffic flow predictions, limited the number of features. While the contributions met by implementing AI algorithms, having a large number of features such as weather conditions could increase its functionality. However, to cover this crisis, we implement an ML algorithm (RF) to find out the importance of features on network traffic flow with the aim of increasing the accuracy of prediction.

A real VANET dataset including V2V and V2R communication that reported the data by installing antennae on the roof of vehicles and gathered data by GPS was the best choice at the time. The vehicles were communicated with DSRC technology and some valuable features such as “packet receiving”, “sender speed” and “receiver speed” were playing a decisive role in achieving the goal. However, another type of communication in VANET such as V2X could improve the implementation.

Due to the large amount of data that is produced by smart devices in smart cities the security of data is a real challenge. Implementing decentralized algorithms such as federated learning helps to transfer the trained model instead of data over the network.

In Chapter 3 the “sender speed” was taken as a network traffic flow parameter and the effect of other factors on it has been evaluated. Moreover, having more detail about the network such as throughput and delay could improve the network traffic prediction. So, we utilized other network parameters like “packet receiving” with the aim of identifying the traffic situation in the network.

In Chapter 5 implementing a DL algorithm in a federated way in order to network traffic prediction while road traffic is considered, while the federated learning libraries have some limitations was one of our biggest challenges in this dissertation. The PySyft library for implementing FL had some complications with compatibility for implementing DL algorithms which forced us to make a handcraft of DL algorithms; then, implement it in a federated way.

7.3 Future Work

In the context of this dissertation, several future works are listed as follows:

- The future work can be predicting the network traffic flow in 5G or 6G of vehicular networks based on artificial intelligence methods.
- Since real VANET data is used in this dissertation, implementing the proposed models in a heterogeneous Vehicular Network (HetVNET) environment which integrates various wireless access networks such as DSRC and LTE can be considered as future work.
- Considering the security of data and implementing the FL which the main process will employ on the edge, using storage services and cloud computing (e.g., Google Cloud Platform (GCP), Amazon Web Services (AWS) and Microsoft Azure) with the aim of analyzing data for network traffic prediction.
- Proposing network traffic prediction and considering more network factors such as delay and throughput while the environmental parameters like weather conditions are considered.
- Based on the growth of generated data by smart devices, implementing the proposed model in Chapter 3 on big data can be an adaptive achievement for smart cities.
- With regards to the proposed approach in Chapter 5 of this dissertation, proposing a secure model that not only keeps the security of data but also considers possible attacks such as penetration attacks, and backdoor attacks on the network.

REFERENCES

- [1] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 166–180, May 2018.
- [2] H.-F. Yang, T. S. Dillon, and Y.-P. P. Chen, "Optimized structure of the traffic flow forecasting model with a deep learning approach," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp. 2371–2381, July 2016.
- [3] X. Wang, Z. Ning, X. Hu, E. C.-H. Ngai, L. Wang, B. Hu, and R. Y. Kwok, "A city-wide real-time traffic management system: Enabling crowdsensing in social internet of vehicles," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 19–25, Sept. 2018.
- [4] Y. Xing, X. Ban, X. Liu, and Q. Shen, "Large-scale traffic congestion prediction based on the symmetric extreme learning machine cluster fast learning method," *Symmetry*, vol. 11, no. 6, p. 730, Mar. 2019.
- [5] N. Taherkhani and S. Pierre, "Improving dynamic and distributed congestion control in vehicular ad hoc networks," *Ad Hoc Networks*, vol. 33, pp. 112–125, Oct. 2015.
- [6] Y. Liu, X. Chen, C. Chen, and X. Guan, "Traffic big data analysis supporting vehicular network access recommendation," in *2016 IEEE international conference on communications (ICC)*. IEEE, May 2016, pp. 1–6.
- [7] R. Gasmi and M. Aliouat, "Vehicular ad hoc networks versus internet of vehicles-a comparative view," in *2019 international conference on networking and advanced systems (ICNAS)*. IEEE, June 2019, pp. 1–6.
- [8] J. Jakubiak and Y. Koucheryavy, "State of the art and research challenges for vanets," in *2008 5th IEEE Consumer Communications and Networking Conference*. IEEE, Jan. 2008, pp. 912–916.
- [9] W. Tong, A. Hussain, W. X. Bo, and S. Maharjan, "Artificial intelligence for vehicle-to-everything: A survey," *IEEE Access*, vol. 7, pp. 10 823–10 843, Jan. 2019.
- [10] B. Haider, S. Henna, A. Gul, and F. Aadil, "A survey on mobility management techniques in vanets," in *2016 IEEE International Conference on Computer and Information Technology (CIT)*. IEEE, Dec. 2016, pp. 125–133.

- [11] O. AlFarraj, A. Tolba, S. Alkhalaf, and A. AlZubi, "Neighbor predictive adaptive handoff algorithm for improving mobility management in vanets," *Computer Networks*, vol. 151, pp. 224–231, Mar. 2019.
- [12] H. Özköse, E. S. Arı, and C. Gencer, "Yesterday, today and tomorrow of big data," *Procedia-Social and Behavioral Sciences*, vol. 195, pp. 1042–1050, July 2015.
- [13] N. Taherkhani and S. Pierre, "Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 11, pp. 3275–3285, Nov. 2016.
- [14] M. S. Bouassida and M. Shawky, "A cooperative congestion control approach within vanets: formal verification and performance evaluation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, pp. 1–13, 2010.
- [15] M. Sepulcre, J. Mittag, P. Santi, H. Hartenstein, and J. Gozalvez, "Congestion and awareness control in cooperative vehicular systems," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1260–1279, June 2011.
- [16] C.-W. Hsu, C.-H. Hsu, and H.-R. Tseng, "Mac channel congestion control mechanism in iee 802.11 p/wave vehicle networks," in *2011 IEEE Vehicular Technology Conference (VTC Fall)*. IEEE, Sept. 2011, pp. 1–5.
- [17] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, and S. S. Iyengar, "A survey on deep learning: Algorithms, techniques, and applications," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–36, Sept. 2018.
- [18] H. Al Najada and I. Mahgoub, "Big vehicular traffic data mining: Towards accident and congestion prevention," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, Sept. 2016, pp. 256–261.
- [19] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, pp. 1–16, Dec. 2016.
- [20] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of big data*, vol. 2, no. 1, pp. 1–21, Dec. 2015.

- [21] R. Jia, P. Jiang, L. Liu, L. Cui, and Y. Shi, "Data driven congestion trends prediction of urban transportation," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 581–591, June 2017.
- [22] O. El Joubari, "Mobility and traffic models for vanets," Ph.D. dissertation, Université Paris-Saclay, Mar. 2022.
- [23] L. M. Giripunje, A. Vidyarthi, and S. K. Shandilya, "Adaptive congestion prediction in vehicular ad-hoc networks (vanet) using type-2 fuzzy model to establish reliable routes," *Wireless Personal Communications*, pp. 1–22, May 2022.
- [24] M. M. Kunt, I. Aghayan, and N. Noii, "Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods," *Transport*, vol. 26, no. 4, pp. 353–366, Dec. 2011.
- [25] K. S. Jadaan, M. Al-Fayyad, and H. F. Gammoh, "Prediction of road traffic accidents in jordan using artificial neural network (ann)," *Journal of Traffic and Logistics Engineering*, vol. 2, no. 2, June 2014.
- [26] A. Darccedil; in and A. Buuml; lent, "A neural network (nn) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics," *Scientific Research and Essays*, vol. 5, no. 19, pp. 2832–2836, Oct. 2010.
- [27] A. A. Kather and P. Anitha, "Employing a uav unit for the data dissemination protocol in vanet," Aug.
- [28] H. Touluni, M. Miyara, B. Nsiri, and M. Boudhane, "Traffic congestion prevention system," *International Journal of Communication Networks and Information Security*, vol. 11, no. 1, pp. 140–147, Apr. 2019.
- [29] F. Zhou, Q. Yang, T. Zhong, D. Chen, and N. Zhang, "Variational graph neural networks for road traffic prediction in intelligent transportation systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2802–2812, July 2020.
- [30] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, Dec. 2013.
- [31] D. Pescaru, "Urban traffic congestion prediction based on routes information," in *2013 IEEE 8th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. IEEE, May 2013, pp. 121–126.

- [32] C. Menelaou, S. Timotheou, P. Kolios, C. G. Panayiotou, and M. M. Polycarpou, "Minimizing traffic congestion through continuous-time route reservations with travel time predictions," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 1, pp. 141–153, Dec. 2018.
- [33] B. Sony, A. Chakravarti, and M. M. Reddy, "Traffic congestion detection using whale optimization algorithm and multi-support vector machine," in *International Conference on Advances in Civil Engineering (ICACE-2019)*, vol. 21, Mar. 2019, p. 23.
- [34] X. Zheng, D. Zhang, H. Gao, Z. Zhao, H. Huang, and J. Wang, "A novel framework for road traffic risk assessment with hmm-based prediction model," *Sensors*, vol. 18, no. 12, p. 4313, Dec. 2018.
- [35] N. S. Nafi, R. H. Khan, J. Y. Khan, and M. Gregory, "A predictive road traffic management system based on vehicular ad-hoc network," in *2014 Australasian Telecommunication Networks and Applications Conference (ATNAC)*. IEEE, Nov. 2014, pp. 135–140.
- [36] A. Mejdoubi, H. Fouchal, O. Zytoune, and M. Ouadou, "A distributed predictive road traffic management system in urban vanets," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, June 2019, pp. 37–42.
- [37] I. Loumiotis, K. Demestichas, E. Adamopoulou, P. Kosmides, V. Asthenopoulos, and E. Sykas, "Road traffic prediction using artificial neural networks," in *2018 South-Eastern European Design Automation, Computer Engineering, Computer Networks and Society Media Conference (SEEDA_CECNSM)*. IEEE, Sept. 2018, pp. 1–5.
- [38] D. F. Specht *et al.*, "A general regression neural network," *IEEE transactions on neural networks*, vol. 2, no. 6, pp. 568–576, Nov. 1991.
- [39] S. Arshad, Q. Raza, and M. Hussain, "Vanets embedded novel traffic management paradigm for smart city milieu," in *Proceedings of the 3rd International Conference on Future Networks and Distributed Systems*, July 2019, pp. 1–10.
- [40] F. J. Braz, J. Ferreira, F. Gonçalves, K. Weege, J. Almeida, F. Baldo, and P. Gonçalves, "Road traffic forecast based on meteorological information through deep learning methods," *Sensors*, vol. 22, no. 12, p. 4485, June 2022.
- [41] R. Sathiyaraj, A. Bharathi, S. Khan, T. Kiren, I. U. Khan, and M. Fayaz, "A genetic predictive model approach for smart traffic prediction and congestion avoidance for

- urban transportation,” *Wireless Communications and Mobile Computing*, vol. 2022, Apr. 2022.
- [42] A. M. Nagy and V. Simon, “Improving traffic prediction using congestion propagation patterns in smart cities,” *Advanced Engineering Informatics*, vol. 50, p. 101343, Oct. 2021.
 - [43] C. C. Ihueze and U. O. Onwurah, “Road traffic accidents prediction modelling: An analysis of anambra state, nigeria,” *Accident Analysis & Prevention*, vol. 112, pp. 21–29, Mar. 2018.
 - [44] A. Rasaizadi and S. Seyedabrishami, “Stacking ensemble learning process to predict rural road traffic flow,” *Journal of Advanced Transportation*, vol. 2022, June 2022.
 - [45] J. Liu, J. Wan, D. Jia, B. Zeng, D. Li, C.-H. Hsu, and H. Chen, “High-efficiency urban traffic management in context-aware computing and 5g communication,” *IEEE Communications Magazine*, vol. 55, no. 1, pp. 34–40, Jan. 2017.
 - [46] W. Wang, Y. Bai, C. Yu, Y. Gu, P. Feng, X. Wang, and R. Wang, “A network traffic flow prediction with deep learning approach for large-scale metropolitan area network,” in *NOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium*. IEEE, Apr. 2018, pp. 1–9.
 - [47] H. Shi, C. Pan, L. Yang, and X. Gu, “Agg: A novel intelligent network traffic prediction method based on joint attention and gcn-gru,” *Security and Communication Networks*, vol. 2021, Sept. 2021.
 - [48] P.-L. Gregoire, C. Desjardins, J. Laumonier, and B. Chaib-draa, “Urban traffic control based on learning agents,” in *2007 IEEE Intelligent Transportation Systems Conference*. IEEE, Sept. 2007, pp. 916–921.
 - [49] M. Ali Hassoune, Z. Mekakia Maaza, and S.-M. Senouci, “Vertical handover decision algorithm for multimedia streaming in vanet,” *Wireless Personal Communications*, vol. 95, no. 4, pp. 4281–4299, Aug. 2017.
 - [50] C.-M. Huang, D.-T. Dao, and M.-S. Chiang, “A bursty multi-node handover scheme for mobile internet using the partially distributed mobility management (bmh-dmm) architecture,” *Telecommunication Systems*, vol. 69, no. 1, pp. 113–130, Sept. 2018.
 - [51] S. Mahajan, R. HariKrishnan, and K. Kotecha, “Prediction of network traffic in wireless mesh networks using hybrid deep learning model,” *IEEE Access*, vol. 10, pp. 7003–7015, Jan. 2022.

- [52] P. Lopez-Garcia, E. Onieva, E. Osaba, A. D. Masegosa, and A. Perallos, "A hybrid method for short-term traffic congestion forecasting using genetic algorithms and cross entropy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 2, pp. 557–569, Nov. 2015.
- [53] Y. Liu, "Intelligent processing in vehicular ad hoc networks: a survey," *arXiv preprint arXiv:1903.11916*, Mar. 2019.
- [54] R. Wang, Y. Zhang, L. Peng, G. Fortino, and P.-H. Ho, "Time-varying-aware network traffic prediction via deep learning in iiot," *IEEE Transactions on Industrial Informatics*, Mar. 2022.
- [55] A. Mehmood, A. Khanan, A. H. H. Mohamed, S. Mahfooz, H. Song, and S. Abdullah, "Antsc: An intelligent naïve bayesian probabilistic estimation practice for traffic flow to form stable clustering in vanet," *IEEE Access*, vol. 6, pp. 4452–4461, July 2017.
- [56] R. More, A. Mugal, S. Rajgure, R. B. Adhao, and V. K. Pachghare, "Road traffic prediction and congestion control using artificial neural networks," in *2016 International Conference on Computing, Analytics and Security Trends (CAST)*. IEEE, Dec. 2016, pp. 52–57.
- [57] N. M. Balamurugan, M. Adimoolam, M. H. Alsharif, and P. Uthansakul, "A novel method for improved network traffic prediction using enhanced deep reinforcement learning algorithm," *Sensors*, vol. 22, no. 13, p. 5006, July 2022.
- [58] M. Liu, M. Zhang, W. Zhao, C. Song, D. Wang, Q. Li, and Z. Wang, "Prediction of congestion degree for optical networks based on bp artificial neural network," in *2017 16th International Conference on Optical Communications and Networks (ICOON)*. IEEE, Aug. 2017, pp. 1–3.
- [59] Z. Tian, S. Li, Y. Wang, and X. Wang, "A network traffic hybrid prediction model optimized by improved harmony search algorithm," *Neural Network World*, vol. 25, no. 6, p. 669, Nov. 2015.
- [60] D. Wei, "Network traffic prediction based on rbf neural network optimized by improved gravitation search algorithm," *Neural Computing and Applications*, vol. 28, no. 8, pp. 2303–2312, Aug. 2017.
- [61] S. Wang, L. Nie, G. Li, Y. Wu, and Z. Ning, "A multi-task learning-based network traffic prediction approach for sdn-enabled industrial internet of things," *IEEE Transactions on Industrial Informatics*, Jan. 2022.

- [62] R. L. Abduljabbar, H. Dia, and P.-W. Tsai, “Unidirectional and bidirectional lstm models for short-term traffic prediction,” *Journal of Advanced Transportation*, vol. 2021, Mar. 2021.
- [63] G. Xiong, Z. Li, H. Wu, S. Chen, X. Dong, F. Zhu, and Y. Lv, “Building urban public traffic dynamic network based on cps: an integrated approach of big data and ai,” *Applied Sciences*, vol. 11, no. 3, p. 1109, Jan. 2021.
- [64] S. Kolekar, S. Gite, B. Pradhan, and K. Kotecha, “Behavior prediction of traffic actors for intelligent vehicle using artificial intelligence techniques: A review,” *IEEE Access*, vol. 9, pp. 135 034–135 058, Sept. 2021.
- [65] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020.
- [66] D.-H. Le, H.-A. Tran, S. Souihi, and A. Mellouk, “An ai-based traffic matrix prediction solution for software-defined network,” in *ICC 2021-IEEE International Conference on Communications*. IEEE, June 2021, pp. 1–6.
- [67] D. Santos, J. Saias, P. Quaresma, and V. B. Nogueira, “Machine learning approaches to traffic accident analysis and hotspot prediction,” *Computers*, vol. 10, no. 12, p. 157, Nov. 2021.
- [68] A. Ali, Y. Zhu, and M. Zakarya, “Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction,” *Neural networks*, vol. 145, pp. 233–247, Jan. 2022.
- [69] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *arXiv preprint arXiv:1412.3555*, Dec. 2014.
- [70] A. Karpathy, J. Johnson, and L. Fei-Fei, “Visualizing and understanding recurrent networks,” *arXiv preprint arXiv:1506.02078*, June 2015.
- [71] M. Coşkun, Ö. YILDIRIM, U. Ayşegül, and Y. Demir, “An overview of popular deep learning methods,” *European Journal of Technique (EJT)*, vol. 7, no. 2, pp. 165–176, Dec. 2017.
- [72] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, “Lstm: A search space odyssey,” *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp. 2222–2232, July 2016.

- [73] R. Jozefowicz, W. Zaremba, and I. Sutskever, “An empirical exploration of recurrent network architectures,” in *International conference on machine learning*. PMLR, June 2015, pp. 2342–2350.
- [74] P. Sun, A. Boukerche, and Y. Tao, “Ssgru: A novel hybrid stacked gru-based traffic volume prediction approach in a road network,” *Computer Communications*, vol. 160, pp. 502–511, July 2020.
- [75] P. T. Yamak, L. Yujian, and P. K. Gadosey, “A comparison between arima, lstm, and gru for time series forecasting,” in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, Dec. 2019, pp. 49–55.
- [76] Z. Cui, R. Ke, Z. Pu, and Y. Wang, “Deep bidirectional and unidirectional lstm recurrent neural network for network-wide traffic speed prediction,” *arXiv preprint arXiv:1801.02143*, Jan. 2018.
- [77] F. Falahatraftar, S. Pierre, and S. Chamberland, “A centralized and dynamic network congestion classification approach for heterogeneous vehicular networks,” *IEEE Access*, vol. 9, pp. 122 284–122 298, Aug. 2021.
- [78] A. Lazaris and V. K. Prasanna, “Deep learning models for aggregated network traffic prediction,” in *2019 15th International Conference on Network and Service Management (CNSM)*. IEEE, Oct. 2019, pp. 1–5.
- [79] S. Troia, R. Alvizu, Y. Zhou, G. Maier, and A. Pattavina, “Deep learning-based traffic prediction for network optimization,” in *2018 20th International Conference on Transparent Optical Networks (ICTON)*. IEEE, July 2018, pp. 1–4.
- [80] T. H. Aldhyani, M. Alrasheedi, A. A. Alqarni, M. Y. Alzahrani, and A. M. Bamhdi, “Intelligent hybrid model to enhance time series models for predicting network traffic,” *IEEE Access*, vol. 8, pp. 130 431–130 451, July 2020.
- [81] R. Vinayakumar, K. Soman, and P. Poornachandran, “Applying convolutional neural network for network intrusion detection,” in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, Sept. 2017, pp. 1222–1228.
- [82] N. Stepanov, D. Alekseeva, A. Ometov, and E. S. Lohan, “Applying machine learning to lte traffic prediction: comparison of bagging, random forest, and svm,” in *2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, Oct. 2020, pp. 119–123.

- [83] Y. Kamarianakis, W. Shen, and L. Wynter, "Real-time road traffic forecasting using regime-switching space-time models and adaptive lasso," *Applied stochastic models in business and industry*, vol. 28, no. 4, pp. 297–315, July 2012.
- [84] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, Sept. 2014.
- [85] J. Zhang, Y. Zheng, D. Qi, R. Li, and X. Yi, "Dnn-based prediction model for spatio-temporal data," in *Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems*, Oct. 2016, pp. 1–4.
- [86] G. Dai, C. Ma, and X. Xu, "Short-term traffic flow prediction method for urban road sections based on space-time analysis and gru," *IEEE Access*, vol. 7, pp. 143 025–143 035, Sept. 2019.
- [87] Z. Duan, Y. Yang, K. Zhang, Y. Ni, and S. Bajgain, "Improved deep hybrid networks for urban traffic flow prediction using trajectory data," *Ieee Access*, vol. 6, pp. 31 820–31 827, June 2018.
- [88] L. Zhang, N. R. Alharbe, G. Luo, Z. Yao, and Y. Li, "A hybrid forecasting framework based on support vector regression with a modified genetic algorithm and a random forest for traffic flow prediction," *Tsinghua Science and Technology*, vol. 23, no. 4, pp. 479–492, Aug. 2018.
- [89] J.-Y. Yang, L.-D. Chou, Y.-C. Li, Y.-H. Lin, S.-M. Huang, G. Tseng, T.-W. Wang, and S.-P. Lu, "Prediction of short-term average vehicular velocity considering weather factors in urban vanet environments," in *2010 International Conference on Machine Learning and Cybernetics*, vol. 6. IEEE, July 2010, pp. 3039–3043.
- [90] F. Li, W. Chen, and Y. Shui, "Study on connectivity probability of vanets under adverse weather conditions at 5.9 ghz," *IEEE Access*, vol. 8, pp. 547–555, Dec. 2019.
- [91] X. Bao, D. Jiang, X. Yang, and H. Wang, "An improved deep belief network for traffic prediction considering weather factors," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 413–420, Feb. 2021.
- [92] F. I. Rahman, "Short term traffic flow prediction using machine learning-knn, svm and ann with weather information." *International Journal for Traffic & Transport Engineering*, vol. 10, no. 3, June 2020.

- [93] Y. Bie, T. Z. Qiu, C. Zhang, and C. Zhang, "Introducing weather factor modelling into macro traffic state prediction," *Journal of advanced transportation*, vol. 2017, Feb. 2017.
- [94] N. Julio, R. Giesen, and P. Lizana, "Real-time prediction of bus travel speeds using traffic shockwaves and machine learning algorithms," *Research in Transportation Economics*, vol. 59, pp. 250–257, Nov. 2016.
- [95] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers-a tutorial," *ACM Computing Surveys (CSUR)*, vol. 54, no. 6, pp. 1–25, July 2021.
- [96] M. Pal, "Random forest classifier for remote sensing classification," *International journal of remote sensing*, vol. 26, no. 1, pp. 217–222, Jan. 2005.
- [97] J. Ren, S. D. Lee, X. Chen, B. Kao, R. Cheng, and D. Cheung, "Naive bayes classification of uncertain data," in *2009 Ninth IEEE international conference on data mining*. IEEE, Dec. 2009, pp. 944–949.
- [98] G. M. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: directing training data acquisition for svm classification," *Remote Sensing of Environment*, vol. 93, no. 1-2, pp. 107–117, Oct. 2004.
- [99] N. Donges, "A complete guide to the random forest algorithm," Sept. 2021. [Online]. Available: <https://builtin.com/data-science/random-forest-algorithm>
- [100] R. M. Fujimoto, R. Guensler, M. P. Hunter, H. Wu, M. Palekar, J. Lee, and J. Ko, "CRAWDAD dataset gatech/vehicular (v. 2006-03-15)," Downloaded from <https://crawdad.org/gatech/vehicular/20060315>, Mar. 2006.
- [101] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, Nov. 2011.
- [102] S. Kostadinov, "Understanding gru networks." Dec. 2017. [Online]. Available: <https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be>
- [103] S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *arXiv preprint arXiv:1503.06462*, Mar. 2015.
- [104] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.", Sept. 2019.

- [105] Q. V. Le, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, and A. Y. Ng, “On optimization methods for deep learning,” in *ICML*, Jan. 2011.
- [106] J. Duchi, E. Hazan, and Y. Singer, “Adaptive subgradient methods for online learning and stochastic optimization.” *Journal of machine learning research*, vol. 12, no. 7, July 2011.
- [107] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, Dec. 2014.
- [108] T. Tieleman, G. Hinton, *et al.*, “Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude,” *COURSERA: Neural networks for machine learning*, vol. 4, no. 2, pp. 26–31, Oct. 2012.
- [109] A. Botchkarev, “Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology,” *arXiv preprint arXiv:1809.03006*, Sept. 2018.
- [110] Y. Xuan, W. Si, J. Zhu, Z. Sun, J. Zhao, M. Xu, and S. Xu, “Multi-model fusion short-term load forecasting based on random forest feature selection and hybrid neural network,” *Ieee Access*, vol. 9, pp. 69 002–69 009, Sept. 2021.
- [111] H. Wu, M. Palekar, R. Fujimoto, R. Guensler, M. Hunter, J. Lee, and J. Ko, “An empirical study of short range communications for vehicles,” in *Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks*, Sept. 2005, pp. 83–84.
- [112] G. vanRossum, “Python reference manual,” *Department of Computer Science [CS]*, no. R 9525, Jan. 1995.
- [113] J. Davis and M. Goadrich, “The relationship between precision-recall and roc curves,” in *Proceedings of the 23rd international conference on Machine learning*, June 2006, pp. 233–240.
- [114] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance,” *Climate research*, vol. 30, no. 1, pp. 79–82, Dec. 2005.
- [115] Y.-J. Lee and O. Min, “Comparative analysis of machine learning algorithms to urban traffic prediction,” in *2017 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, Oct. 2017, pp. 1034–1036.

- [116] A. Y. Nikraves, S. A. Ajila, C.-H. Lung, and W. Ding, "Mobile network traffic prediction using mlp, mlpwd, and svm," in *2016 IEEE International Congress on Big Data (BigData Congress)*. IEEE, June 2016, pp. 402–409.
- [117] D. Xu, Y. Wang, P. Peng, S. Beilun, Z. Deng, and H. Guo, "Real-time road traffic state prediction based on kernel-knn," *Transportmetrica A: Transport Science*, vol. 16, no. 1, pp. 104–118, Dec. 2020.
- [118] W. Chen, Z. Shang, and Y. Chen, "A novel hybrid network traffic prediction approach based on support vector machines," *Journal of Computer Networks and Communications*, vol. 2019, Feb. 2019.
- [119] L. Nie, X. Wang, L. Wan, S. Yu, H. Song, and D. Jiang, "Network traffic prediction based on deep belief network and spatiotemporal compressive sensing in wireless mesh backbone networks," *Wireless Communications and Mobile Computing*, vol. 2018, Jan. 2018.
- [120] S. S. Sepasgozar and S. Pierre, "Network traffic prediction model considering road traffic parameters using artificial intelligence methods in vanet," *IEEE Access*, vol. 10, pp. 8227–8242, Jan. 2022.
- [121] A. R. Mohammed, S. A. Mohammed, and S. Shirmohammadi, "Machine learning and deep learning based traffic classification and prediction in software defined networking," in *2019 IEEE International Symposium on Measurements & Networking (M&N)*. IEEE, July 2019, pp. 1–6.
- [122] Z. Tian and P. Song, "A novel network traffic combination prediction model," *International Journal of Communication Systems*, vol. 35, no. 7, p. e5097, May 2022.
- [123] H.-K. Lim, J.-B. Kim, J.-S. Heo, K. Kim, Y.-G. Hong, and Y.-H. Han, "Packet-based network traffic classification using deep learning," in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*. IEEE, Feb. 2019, pp. 046–051.
- [124] S. Nihale, S. Sharma, L. Parashar, and U. Singh, "Network traffic prediction using long short-term memory," in *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, July 2020, pp. 338–343.
- [125] S. Dong, "Multi class svm algorithm with active learning for network traffic classification," *Expert Systems with Applications*, vol. 176, p. 114885, Aug. 2021.

- [126] N. Ramakrishnan and T. Soni, “Network traffic prediction using recurrent neural networks,” in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, Dec. 2018, pp. 187–193.
- [127] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konečný, S. Mazzocchi, B. McMahan, *et al.*, “Towards federated learning at scale: System design,” *Proceedings of Machine Learning and Systems*, vol. 1, pp. 374–388, Apr. 2019.
- [128] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Artificial intelligence and statistics*. PMLR, Apr. 2017, pp. 1273–1282.
- [129] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, “Federated optimization: Distributed machine learning for on-device intelligence,” *arXiv preprint arXiv:1610.02527*, Oct. 2016.
- [130] S. Saha and T. Ahmad, “Federated transfer learning: Concept and applications,” *Intelligenza Artificiale*, vol. 15, no. 1, pp. 35–44, Jan. 2021.
- [131] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, May 2020.
- [132] Y. Ren, X. Chen, S. Wan, K. Xie, and K. Bian, “Passenger flow prediction in traffic system based on deep neural networks and transfer learning method,” in *2019 4th International Conference on Intelligent Transportation Engineering (ICITE)*. IEEE, Sept. 2019, pp. 115–120.
- [133] M. V. S. da Silva, L. F. Bittencourt, and A. R. Rivera, “Towards federated learning in edge computing for real-time traffic estimation in smart cities,” in *Anais do IV Workshop de Computação Urbana*. SBC, Dec. 2020, pp. 166–177.
- [134] D. Ye, R. Yu, M. Pan, and Z. Han, “Federated learning in vehicular edge computing: A selective model aggregation approach,” *IEEE Access*, vol. 8, pp. 23 920–23 935, Jan. 2020.
- [135] Y. Qi, M. S. Hossain, J. Nie, and X. Li, “Privacy-preserving blockchain-based federated learning for traffic flow prediction,” *Future Generation Computer Systems*, vol. 117, pp. 328–337, Apr. 2021.

- [136] J. Pei, K. Zhong, M. A. Jan, and J. Li, “Personalized federated learning framework for network traffic anomaly detection,” *Computer Networks*, vol. 209, p. 108906, May 2022.
- [137] C. Zhang, S. Zhang, J. James, and S. Yu, “Fastgnn: A topological information protected federated learning approach for traffic speed forecasting,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8464–8474, Jan. 2021.
- [138] C. Zhang, S. Dang, B. Shihada, and M.-S. Alouini, “Dual attention-based federated learning for wireless traffic prediction,” in *IEEE INFOCOM 2021-IEEE conference on computer communications*. IEEE, May 2021, pp. 1–10.
- [139] T. Italia, “Telecommunications-sms, call, internet-mi,” *Hardvard Dataverse*, May 2015.
- [140] —, “Telecommunications-sms, call, internet-mi.(2015),” *Hardvard Dataverse*. [Online]. Available: <https://doi.org/10.7910/DVN/EGZHFV>
- [141] H. Feng, Y. Shu, S. Wang, and M. Ma, “Svm-based models for predicting wlan traffic,” in *2006 IEEE international conference on communications*, vol. 2. IEEE, June 2006, pp. 597–602.
- [142] C. Qiu, Y. Zhang, Z. Feng, P. Zhang, and S. Cui, “Spatio-temporal wireless traffic prediction with recurrent neural network,” *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 554–557, Jan. 2018.
- [143] S. Ji, S. Pan, G. Long, X. Li, J. Jiang, and Z. Huang, “Learning private neural language modeling with attentive aggregation,” in *2019 International joint conference on neural networks (IJCNN)*. IEEE, July 2019, pp. 1–8.
- [144] Y. Liu, J. James, J. Kang, D. Niyato, and S. Zhang, “Privacy-preserving traffic flow prediction: A federated learning approach,” *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7751–7763, Apr. 2020.
- [145] C. Chen, *Freeway performance measurement system (PeMS)*. University of California, Berkeley, 2002.
- [146] R. Fu, Z. Zhang, and L. Li, “Using lstm and gru neural network methods for traffic flow prediction,” in *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE, Nov. 2016, pp. 324–328.

- [147] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, “Long short-term memory neural network for traffic speed prediction using remote microwave sensor data,” *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, May 2015.
- [148] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, “Support vector machines for wind speed prediction,” *Renewable energy*, vol. 29, no. 6, pp. 939–947, May 2004.
- [149] X. Yuan, J. Chen, J. Yang, N. Zhang, T. Yang, T. Han, and A. Taherkordi, “Fedstn: Graph representation driven federated learning for edge computing enabled urban traffic flow prediction,” *IEEE Transactions on Intelligent Transportation Systems*, Mar. 2022.
- [150] X. Yuan, J. Chen, N. Zhang, C. Zhu, Q. Ye, and X. S. Shen, “Fedtse: Low-cost federated learning for privacy-preserved traffic state estimation in iov,” in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, May 2022, pp. 1–6.
- [151] H. Yang, X. Li, W. Qiang, Y. Zhao, W. Zhang, and C. Tang, “A network traffic forecasting method based on sa optimized arima–bp neural network,” *Computer Networks*, vol. 193, p. 108102, July 2021.
- [152] K. Smagulova and A. P. James, “Overview of long short-term memory neural networks,” in *Deep Learning Classifiers with Memristive Networks*. Springer, 2020, pp. 139–153.
- [153] A. R. Javed, M. A. Hassan, F. Shahzad, W. Ahmed, S. Singh, T. Baker, and T. R. Gadekallu, “Integration of blockchain technology and federated learning in vehicular (iot) networks: A comprehensive survey,” *Sensors*, vol. 22, no. 12, p. 4394, June 2022.
- [154] N. Llisterri Giménez, M. Monfort Grau, R. Pueyo Centelles, and F. Freitag, “On-device training of machine learning models on microcontrollers with federated learning,” *Electronics*, vol. 11, no. 4, p. 573, Feb. 2022.
- [155] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, *et al.*, “Pytorch: An imperative style, high-performance deep learning library,” *Advances in neural information processing systems*, vol. 32, 2019.
- [156] T. Ryffel, A. Trask, M. Dahl, B. Wagner, J. Mancuso, D. Rueckert, and J. Passerat-Palmbach, “A generic framework for privacy preserving deep learning,” *arXiv preprint arXiv:1811.04017*, Nov. 2018.

- [157] “Google colab,” 2022. [Online]. Available: <https://colab.research.google.com>
- [158] N. S. Nafi and J. Y. Khan, “A vanet based intelligent road traffic signalling system,” in *Australasian Telecommunication Networks and Applications Conference (ATNAC) 2012*. IEEE, Nov. 2012, pp. 1–6.
- [159] B. Cherkaoui, A. Beni-Hssane, M. El Fissaoui, and M. Erritali, “Road traffic congestion detection in vanet networks,” *Procedia Computer Science*, vol. 151, pp. 1158–1163, Jan. 2019.
- [160] C. Jayapal and S. S. Roy, “Road traffic congestion management using vanet,” in *2016 International conference on advances in human machine interaction (HMI)*. IEEE, Mar. 2016, pp. 1–7.
- [161] F. Falahatraftar, S. Pierre, and S. Chamberland, “An intelligent congestion avoidance mechanism based on generalized regression neural network for heterogeneous vehicular networks,” *IEEE Transactions on Intelligent Vehicles*, June 2022.
- [162] M. Taruk, E. Budiman, H. J. Setyadi, *et al.*, “Comparison of tcp variants in long term evolution (lte),” in *2017 5th international conference on electrical, electronics and information engineering (ICEEIE)*. IEEE, Oct. 2017, pp. 131–134.

APPENDIX A CLASSIFICATION PERFORMANCE METRICS

$$Precision = \frac{TP}{TP + FP} \quad (\text{A.1})$$

$$Recall = \frac{TP}{TP + FN} \quad (\text{A.2})$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (\text{A.3})$$

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \times 100 \quad (\text{A.4})$$

APPENDIX B GRU ALGORITHM EQUATIONS

$$\text{Update gate:} \quad Z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (\text{B.1})$$

$$\text{Reset gate:} \quad r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (\text{B.2})$$

$$\text{Process input:} \quad \tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (\text{B.3})$$

$$\text{Hidden state update:} \quad h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (\text{B.4})$$

APPENDIX C REGRESSION PERFORMANCE METRICS

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |x_i - \hat{x}_i| \quad (C.1)$$

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (C.2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (C.3)$$

$$R^2 = 1 - \frac{\sum (x_i - \hat{x}_i)}{\sum (x_i - x_{avg})} \quad (C.4)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (C.5)$$

APPENDIX D MINIMUM VARIANCE

$$I = \min_s \frac{\sum_{s=1}^k (y_s - \hat{y}_s)^2}{k} \quad (\text{D.1})$$

APPENDIX E IMPORTANCE SCORE

$$Vim = \frac{\sum_{j=1}^n (MSE_j - MSE_{nj})^2}{n \cdot S_E} \quad (\text{E.1})$$

APPENDIX F MLP ALGORITHM EQUATIONS

$$y = \varphi\left(\sum_{i=1}^n w_i x_i + b\right) = \varphi(W^T X + b) \quad (\text{F.1})$$

$$E = \frac{1}{l} \sum_{i=1}^l (T_i - Y_i)^2 \quad (\text{F.2})$$

APPENDIX G LSTM ALGORITHM EQUATIONS

$$\text{Forget gate:} \quad f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (\text{G.1})$$

$$\text{Input gate:} \quad i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (\text{G.2})$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (\text{G.3})$$

$$\text{Cell state:} \quad C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (\text{G.4})$$

$$\text{Output gate:} \quad o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (\text{G.5})$$

$$h_t = o_t \tanh(C_t) \quad (\text{G.6})$$