

RESEARCH ARTICLE

Fed-NTP: A Federated Learning Algorithm for Network Traffic Prediction in VANET

SANAZ SHAKER SEPASGOZAR  **AND SAMUEL PIERRE, (Senior Member, IEEE)**

Mobile Computing and Networking Research Laboratory (LARIM), Department of Computer and Software Engineering, Polytechnique Montreal, Montreal, QC H3T 1J4, Canada

Corresponding author: Sanaz Shaker Sepasgozar (sanaz-2.shaker-sepasgozar@polymtl.ca)

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 $R^2 - SCORE$ (0.975).

INDEX TERMS Vehicular ad-hoc network, federated learning, network traffic prediction, deep learning, artificial intelligence, intelligent transportation system.

I. 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) [1], LSTM [2] and Gated Recurrent Unit (GRU) [3] 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 [4], [5], [6]. In this way, the distributed data will keep in devices and the local data does not transfer between different users and servers. FL is

The associate editor coordinating the review of this manuscript and approving it for publication was Omer Chughtai.

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 [7]. 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 [8]. 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 [9]. 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.

II. 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 decentralized algorithm to cope with privacy-preserving issues. For instance, transfer learning has been implemented in [10] for traffic prediction, the requirements for implementing FL with the aim of traffic estimation were proposed in [11], a selective model aggregation based on FL for classification problem was proposed in [12], Privacy-preserving blockchain-based FL was proposed in [13] for predicting traffic flow and other hybrid methods based on FL were proposed in [14] and [15].

A novel wireless traffic prediction framework based on FL was proposed in [16] 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 [17], [18]. They compared their results with five other baselines such as Lasso: A linear model for regression, Support Vector Regression (SVR) [19], LSTM [20], FedAvg [5] and FedAtt [21] 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 [22] 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 [23] and compared their result with centralized algorithms including GRU [24], SAE [25], LSTM [26], and SVM [27]. 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 [28] 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 spatiotemporal 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. [29] 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. [30] 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.

TABLE 1. Summary of related work in traffic prediction.

Proposed method	Contribution	Reference
FedDA	Wireless Traffic Prediction based on FL	[16]
FedGRU	Privacy preserving algorithm to traffic flow prediction	[22]
FeDSTN	Traffic flow prediction based on historical traffic data	[28]
FedTSE	Vehicular speed and traffic flow prediction in FL way based on LSTM	[29]
RF-GRU-NTP	Network traffic flow prediction based on road and algorithms network parameters utilizing AI	[30]
ARIMA-SA-BPNN	Network traffic prediction based on historical data	[31]

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. [31] 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. 1 represents a summary of the related work regarding the existing methods for traffic prediction.

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.

III. 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 [30]. However, there are still factors

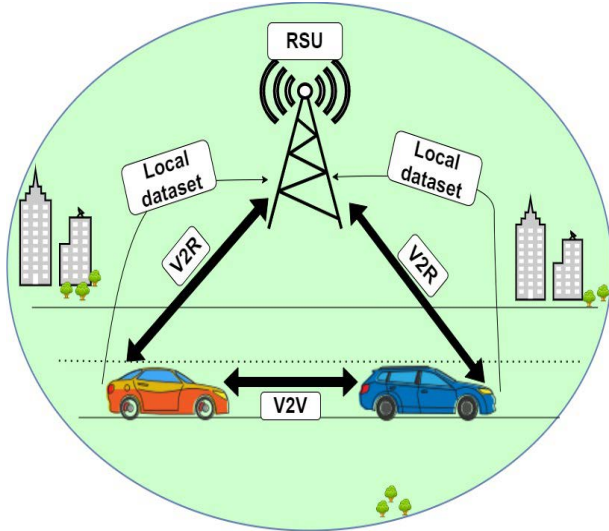


FIGURE 1. VANET architecture and training in centralized learning.

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. Figure 1 shows the architecture of the VANET environment and model training in centralized learning where the local dataset will be sent to the server.

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.

A. 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. Figure 2 shows the fundamental structure of the LSTM algorithm.

The equations of the gates are as follows:

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

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

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

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

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

$$h_t = o_t \tanh(C_t) \quad (6)$$

Deciding about which information is needed to be kept or ignored, is under forget gate responsibility and has been calculated in (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 (2,3) and is responsible to determine what information is relevant

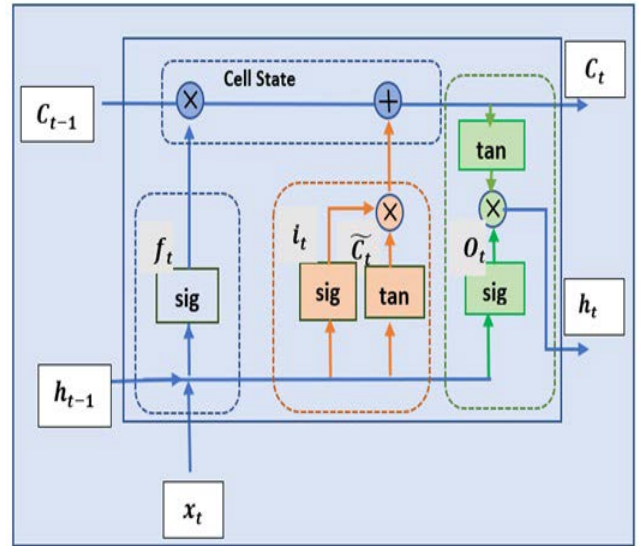


FIGURE 2. Structure of LSTM algorithm.

to add from the current step. In (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_c , 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 (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 (5,6), 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 [32].

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.

B. 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 [33]. Figure 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

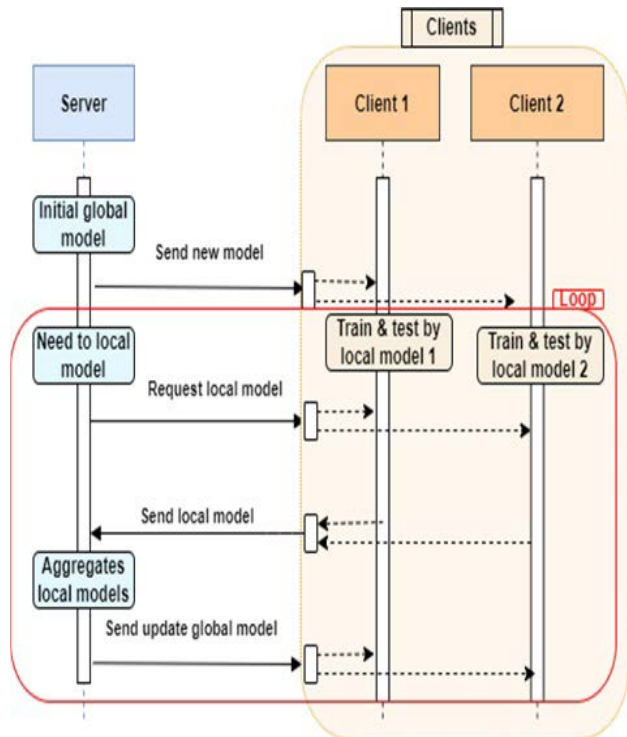


FIGURE 3. Sequence diagram for federated learning.

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 [34].

C. 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 [9] 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 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 Figure 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 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.

IV. DATA PREPARATION AND PERFORMANCE EVALUATION

A. 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 [9]. 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 [35].

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.

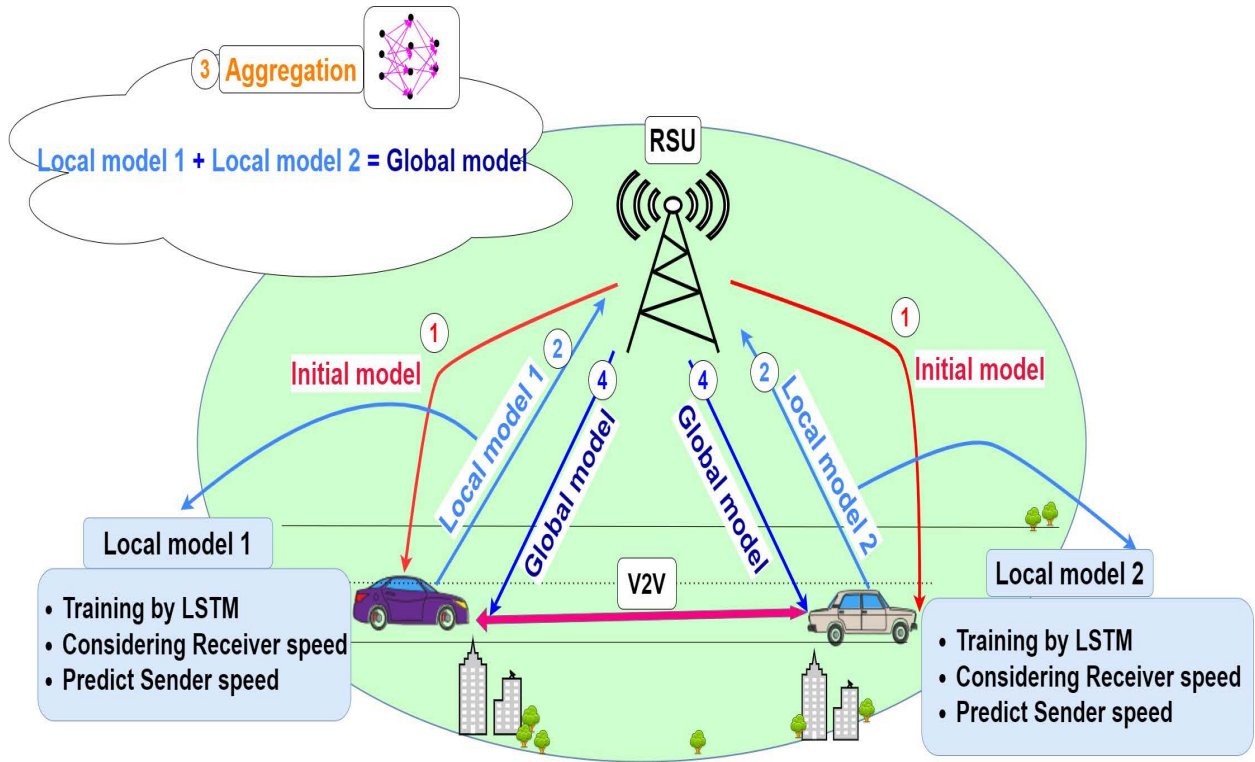


FIGURE 4. The architecture of proposed Fed-NTP model.

B. IMPLEMENTATION

We used python version 3.7 [36] and PyTorch [37] to implement the DL algorithm, for local training in the client. To design the FL algorithm, one extension library of Python called PySyft [38] has been used. PySyft can provide the requirements of FL algorithms and can be used by some main DL frameworks namely PyTorch and TensorFlow [39] that can be used for private and secure DL algorithms.

The implementation occurred in the Google Colab platform [40] 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 StandardScaler, 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 [41] 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.

C. 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 [42]:

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |x_i - \hat{x}_i| \tag{7}$$

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (x_i - \hat{x}_i)^2 \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \tag{9}$$

$$R^2 = 1 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - x_{avg})^2} \tag{10}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right|, \tag{11}$$

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 sets. The proposed Fed-NTP model can achieve the best results in terms of used evaluation metrics mentioned above.

D. EXPERIMENTAL RESULTS

We compared the performance of the proposed Fed-NTP model with three centralized and one decentralized

TABLE 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

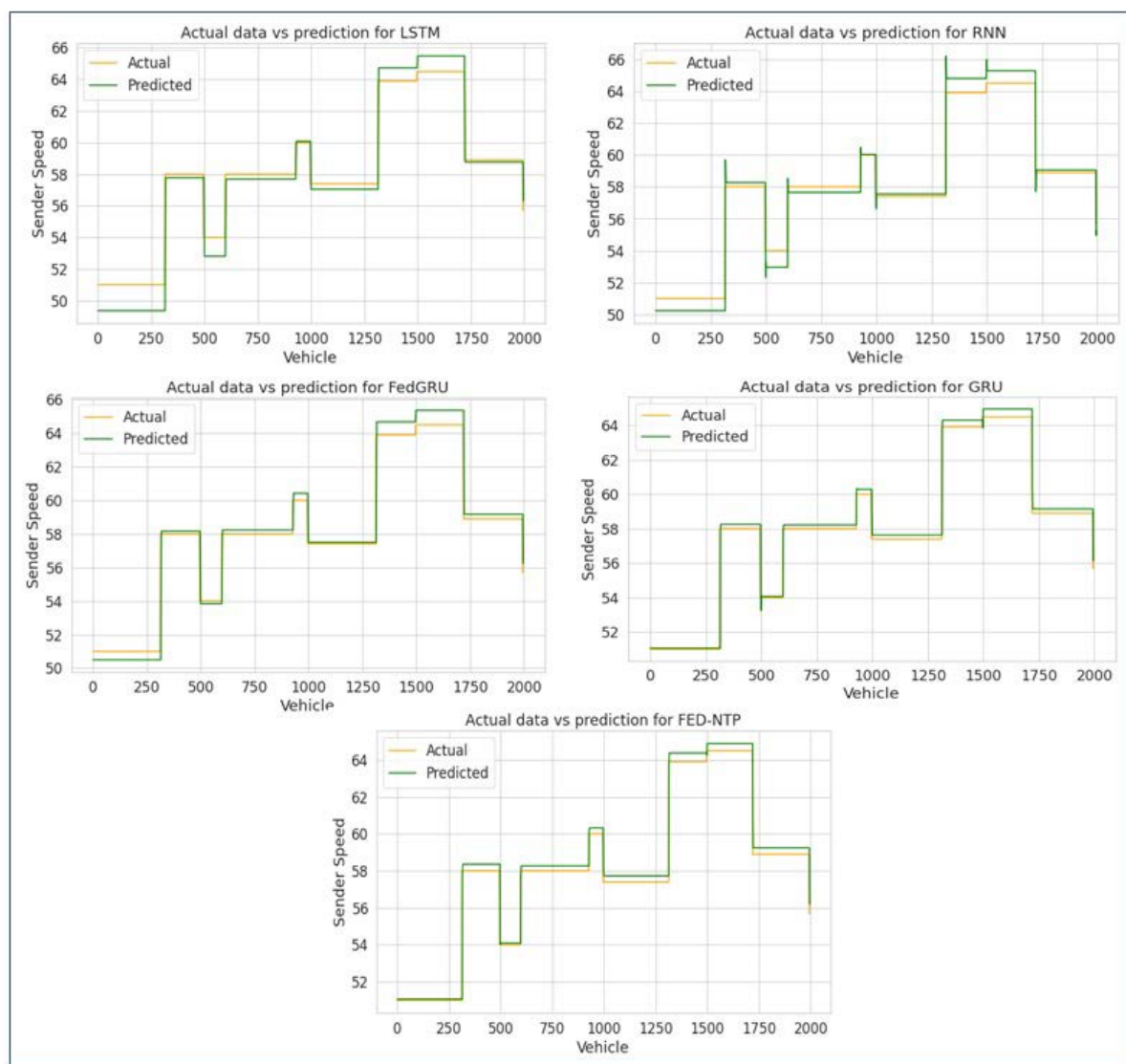


FIGURE 5. Network traffic prediction results for the LSTM, GRU, FedGRU and Fed-NTP.

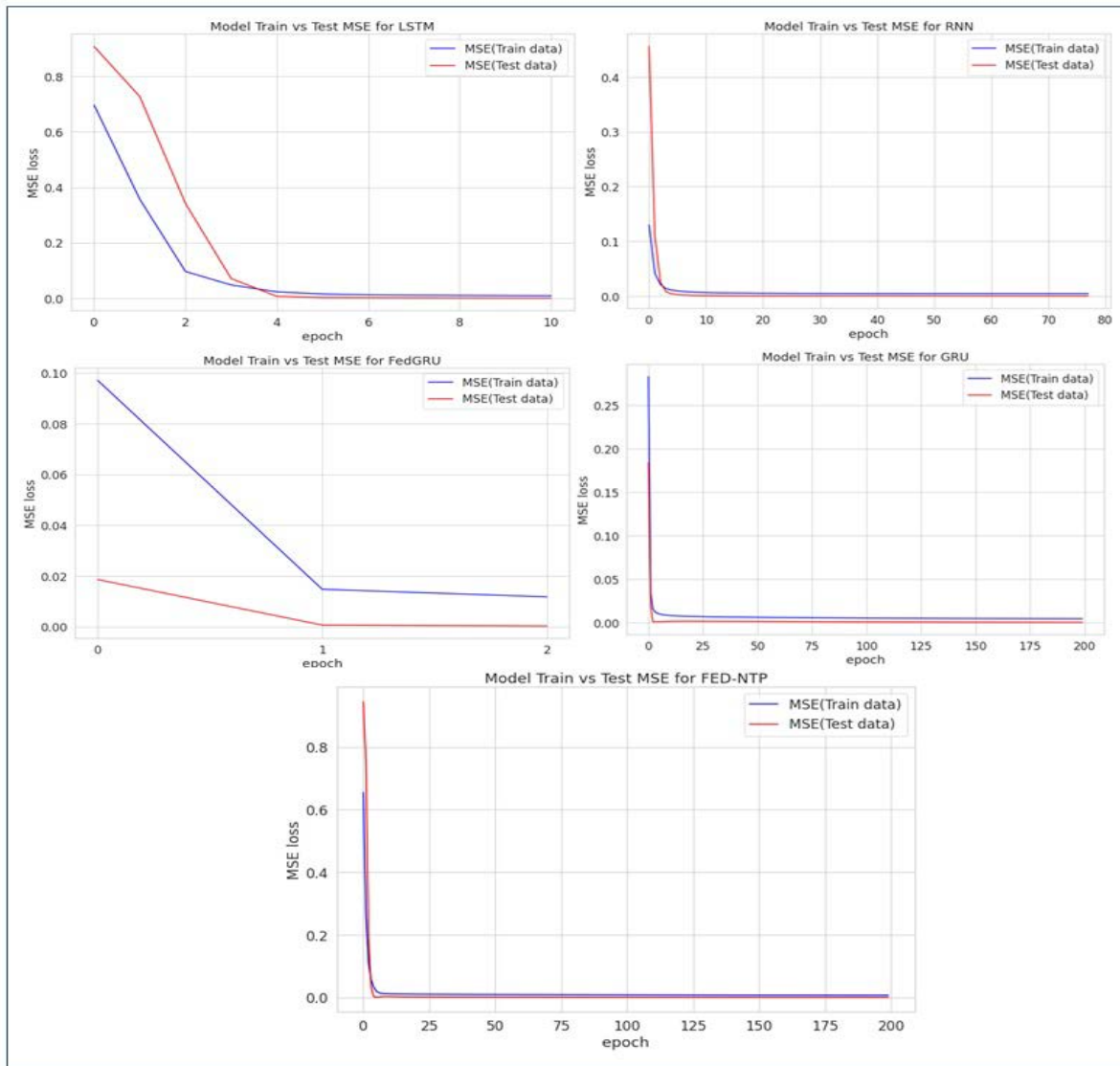


FIGURE 6. The MSE loss for the LSTM, GRU, RNN, FedGRU and Fed-NTP algorithms.

algorithm including LSTM [26], GRU [24], RNN [43] and FedGRU [22] 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. 2 shows the evaluation performance of different algorithms based on various metrics.

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 Figure 5. The actual data is illustrated in the orange line and the predicted data is depicted in the green line. As shown in Figure 5, the proposed Fed-NTP model performs more precisely than other algorithms with the lowest difference between the actual and predicted data.

Figure 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 Figure 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 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.

V. 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 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.

REFERENCES

- [1] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” 2014, *arXiv:1412.3555*.
- [2] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, “LSTM: A search space Odyssey,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2016.
- [3] R. Jozefowicz, W. Zaremba, and I. Sutskever, “An empirical exploration of recurrent network architectures,” in *Proc. Int. Conf. Mach. Learn.*, Oct. 2015, pp. 2342–2350.
- [4] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, and J. Roselander, “Towards federated learning at scale: System design,” *Proc. Mach. Learn. Syst.*, vol. 1, pp. 374–388, Apr. 2019.
- [5] B. McMahan, E. Moore, D. Ramage, S. Hampson, and Y. Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Proc. Artif. Intell. Statist.*, Apr. 2017, pp. 1273–1282.
- [6] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, “Federated optimization: Distributed machine learning for on-device intelligence,” 2016, *arXiv:1610.02527*.
- [7] S. Saha and T. Ahmad, “Federated transfer learning: Concept and applications,” *Intelligenza Artificiale*, vol. 15, no. 1, pp. 35–44, Jul. 2021.
- [8] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020.
- [9] R. M. Fujimoto, R. Guensler, M. P. Hunter, H. Wu, M. Palekar, J. Lee, and J. Ko. (Mar. 15, 2006). *CRAWDAD Dataset Gatech/Vehicular*. Accessed: Mar. 2006. [Online]. Available: <https://crawdad.org/gatech/vehicular/20060315>
- [10] 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 *Proc. 4th Int. Conf. Intell. Transp. Eng. (ICITE)*, Sep. 2019, pp. 115–120.
- [11] M. V. S. D. Silva, L. F. Bittencourt, and A. R. Rivera, “Towards federated learning in edge computing for real-time traffic estimation in smart cities,” in *Proc. Anais do Workshop de Computação Urbana (CoUrb)*, Dec. 2020, pp. 166–177.
- [12] 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. 23920–23935, 2020.
- [13] Y. Qi, M. S. Hossain, J. Nie, and X. Li, “Privacy-preserving blockchain-based federated learning for traffic flow prediction,” *Future Gener. Comput. Syst.*, vol. 117, pp. 328–337, Apr. 2021.
- [14] J. Pei, K. Zhong, M. A. Jan, and J. Li, “Personalized federated learning framework for network traffic anomaly detection,” *Comput. Netw.*, vol. 209, May 2022, Art. no. 108906.
- [15] C. Zhang, S. Zhang, J. J. Q. Yu, and S. Yu, “FASTGNN: A topological information protected federated learning approach for traffic speed forecasting,” *IEEE Trans. Ind. Informat.*, vol. 17, no. 12, pp. 8464–8474, Dec. 2021.
- [16] C. Zhang, S. Dang, B. Shihada, and M.-S. Alouini, “Dual attention-based federated learning for wireless traffic prediction,” in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, May 2021, pp. 1–10.
- [17] T. Italia, “Telecommunications—SMS, call, internet—TN,” Harvard Dataverse, May 2015, doi: [10.7910/DVN/QLCABU](https://doi.org/10.7910/DVN/QLCABU).
- [18] T. Italia, “Telecommunications—SMS, call, internet—MI,” Harvard Dataverse, 2015, doi: [10.7910/DVN/EGZHFV](https://doi.org/10.7910/DVN/EGZHFV).
- [19] H. Feng, Y. Shu, S. Wang, and M. Ma, “SVM-based models for predicting WLAN traffic,” in *Proc. IEEE Int. Conf. Commun.*, Jun. 2006, pp. 597–602.
- [20] C. Qiu, Y. Zhang, Z. Feng, P. Zhang, and S. Cui, “Spatio-temporal wireless traffic prediction with recurrent neural network,” *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 554–557, Aug. 2018.
- [21] S. Ji, S. Pan, G. Long, X. Li, J. Jiang, and Z. Huang, “Learning private neural language modeling with attentive aggregation,” in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8.
- [22] Y. Liu, J. J. Q. Yu, J. Kang, D. Niyato, and S. Zhang, “Privacy-preserving traffic flow prediction: A federated learning approach,” *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7751–7763, Aug. 2020.
- [23] C. Chen, *Freeway Performance Measurement System (PeMS)*. Berkeley, CA, USA: Univ. California, 2002.
- [24] R. Fu, Z. Zhang, and L. Li, “Using LSTM and GRU neural network methods for traffic flow prediction,” in *Proc. 31st Youth Acad. Annu. Conf. Chin. Assoc. Automat. (YAC)*, Nov. 2016, pp. 324–328.

- [25] Y. S. Lv, Y. J. Duan, W. W. Kang, Z. X. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Jan. 2014.
- [26] 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," *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.
- [27] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renew. Energy*, vol. 29, no. 6, pp. 939–947, 2004.
- [28] 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 Trans. Intell. Transp. Syst.*, early access, Mar. 17, 2022, doi: [10.1109/TITS.2022.3157056](https://doi.org/10.1109/TITS.2022.3157056).
- [29] 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 *Proc. IEEE INFOCOM 2022-IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, vol. 29, May 2022, pp. 1–6.
- [30] 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, 2022.
- [31] 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," *Comput. Netw.*, vol. 193, Jul. 2021, Art. no. 108102.
- [32] K. Smagulova and A. P. James, "Overview of long short-term memory neural networks," in *Deep Learning Classifiers With Memristive Networks*. Cham, Switzerland: Springer, Apr. 2019, pp. 139–153.
- [33] 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, Jun. 2022.
- [34] N. L. Giménez, M. M. Grau, R. P. 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.
- [35] 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 *Proc. 2nd ACM Int. Workshop Veh. Ad Hoc Netw. (VANET)*, 2005, pp. 83–84.
- [36] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual CreateSpace*. Scotts Valley, CA, USA: CWI, Jan. 1995.
- [37] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–12.
- [38] T. Ryffel, A. Trask, M. Dahl, B. Wagner, J. Mancuso, D. Rueckert, and J. Passerat-Palmbach, "A generic framework for privacy preserving deep learning," 2018, [arXiv:1811.04017](https://arxiv.org/abs/1811.04017).
- [39] A. Géron, *Hands-on Machine Learning With Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, CA, USA: O'Reilly Media, Sep. 2019.
- [40] *Google Colab*. Accessed: Jun. 12, 2022. [Online]. Available: <https://colab.research.google.com/>
- [41] Q. V. Le, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, and A. Y. Ng, "On optimization methods for deep learning," in *Proc. ICML*, Jan. 2011, pp. 265–272.
- [42] A. Botchkarev, "Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology," 2018, [arXiv:1809.03006](https://arxiv.org/abs/1809.03006).
- [43] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Applying convolutional neural network for network intrusion detection," in *Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICACCI)*, Sep. 2017, pp. 1222–1228.



SANAZ SHAKER SEPASGOZAR received the B.Sc. degree in computer software engineering from Islamic Azad University Central Tehran Branch, Tehran, Iran, in 2010, and the M.Sc. degree in IT management—information resource from Islamic Azad University Science and Research Branch, Tehran, in 2017. She is currently pursuing the Ph.D. degree in computer engineering with Polytechnique Montreal, QC, Canada. Her research interests include traffic flow prediction

on the road and vehicular *ad-hoc* networks using deep learning and machine learning algorithms.



SAMUEL PIERRE (Senior Member, IEEE) is currently a Full Professor with the Department of Computer and Software Engineering, Polytechnique Montreal, and the Director of the Mobile Computing and Networking Research Laboratory (LARIM). His research interests include wired and wireless communications, mobile computing and networking, cloud computing, and e-learning. He has authored or coauthored more than 600 technical publications, including papers in refereed

archival journals, textbooks, patents, and book chapters. He has received several awards, including the Prix Poly 1873 for excellence in teaching and training, in 2001 and 2005, and the Knight of the National Order of Quebec, in 2009. He is a fellow of the Engineering Institute of Canada, in 2003, and Canadian Academy of Engineering, in 2008. In December 2011, he was appointed as a member of the Order of Canada. In May 2014, he received an Honorary Doctorate from the University of Quebec at Trois-Rivieres (UQTR), and the University of Quebec in Outaouais (UQO), in November 2016. In 2017, he obtained the El Fassi Prize from the Agence universitaire de la Francophonie (AUF) to highlight the action of a person, who has exerted a significant influence through the quality of his expertise and the innovative nature of his achievements at the international level in the fields of research, training, development, and international cooperation, governance, and/or transfer of knowledge or skills. In 2020, he received the Grand Prize for Professional Excellence from the Order of Engineers of Quebec (OIQ). In 2021, he received the Gold Medal from Engineers Canada.

• • •