Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 4, August : 2023

SMART TRAFFIC SYSTEM WITH TRAFFICPREDICTION USING MACHINE LEARNING

Vishwasgowda T M, M.Tech Student, Department of InformationScience and Engineering, The National Institute of Engineering, Mysuru : 2020scn_vishwasgowdatm_a@nie.ac.in
Chinnaswamy C N, Associate Professor, Department of InformationScience and Engineering, The National Institute of Engineering, Mysuru : chinnaswamycn@nie.ac.in

Abstract
The traffic is one of the major issues faced by people in almost every city. Predicting the traffic has always been very challenging for transportation planning as well as the city manager. With the growth of population and usage of number of vehicles has increased the need for reliable traffic prediction system. The use of machine learning algorithms would be more suitable in predicting the traffic and help maintain the flow of vehicles. The paper focuses on formulating a method for better prediction of the traffic with the use of machine learning principles such as LSTM and Random Forest (RF) algorithm. Feature engineering is applied on the existing dataset for traffic forecasts and the impact of feature engineering is presented. The basic dataset had the following fields: Vehicle Id, Time, Date, Junction Id as the primary and significant parameters considered to anticipate the traffic at each zone. Along with these we have added the day and weekend features to enhance the results. The system is tested to have improved accuracy with the help of feature engineering.

Keywords: Traffic, Artificial Intelligence, Random Forest, LSTM, RMSE,Dataset, Testing and training.

I. Introduction

Our country is heading towards the advancement of urbanization with the increase in the usage of automobiles, transportation challenges are turning out to be progressively troublesome, also the accidents are common, and the traffic environment is degrading [1]. There are many elated works happening to boost the capacity of the road network. One of the alternatives that comes to our mind is to add an additional roadway that reduces the traffic flow, i.e., to increase the number of lanes on the road. However, according to a study conducted by academics, increasing road capacity will result in worse traffic circumstances [2][3]. Establishing an effective and accurate transportation system, which can help us better plan transportation resources, disperse traffic flow before it becomes overcrowded, and even provide more abundant on-road entertainment, is an efficient way to improve the traffic environment. The source of the traffic prediction could be taken from sensors adopted in roads, vehicle count, traffic cameras, parking system and many more. As part of Artificial Intelligence (AI), Machine Learning (ML) is one of the most important and popular growing fields these days [4][10]. Machine learning has recently emerged as an important and promising study field for transportation engineering, particularly in traffic prediction. Traffic congestion has an impact on the economy, either directly or indirectly. Every day, traffic congestion wastes people's valuable time and increasesthe expense of fuel.

Because traffic congestion is a big issue for individuals of all socioeconomic backgrounds, there must be small-scale traffic forecasting for people to live their lives without annoyance or strain. The comfort of road users is essential first and foremost to ensure the country's economic progress. This is only achievable if the traffic flow is smooth. To deal with this, we require traffic prediction so that we can estimate or predict future traffic to some degree. Figure 1 depicts the advantage of traffic prediction. The traffic flow prediction with the use of Big Data could be accurately measured, the traffic information will strongly help the improvement of business in the travels sector and other government agencies. The road user can make better travel decision, reduce traffic congestion and

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 4, August : 2023
could contribute to the carbon emission. The purpose of traffic flow prediction is to predict traffic as soon as feasible for users. Nowadays, traffic is extremely congested, and people on the road have no control over it. As a result, this study may be useful in predicting traffic. In summary, this paper's structures are divided into five primary components. Introduction, Related Work, Overview, Methodology, Results and Discussion, and Conclusion with Future Work are the sections.

Objective of this paper includes:

1. Analyzing the traffic pattern with time.
2. Feature engineering to enhance the existing dataset.
3. Increasing the accuracy and reducing the root mean square error with ML models.

The major benefits of implementing the traffic prediction system are identified and illustrated in the following fig 1.

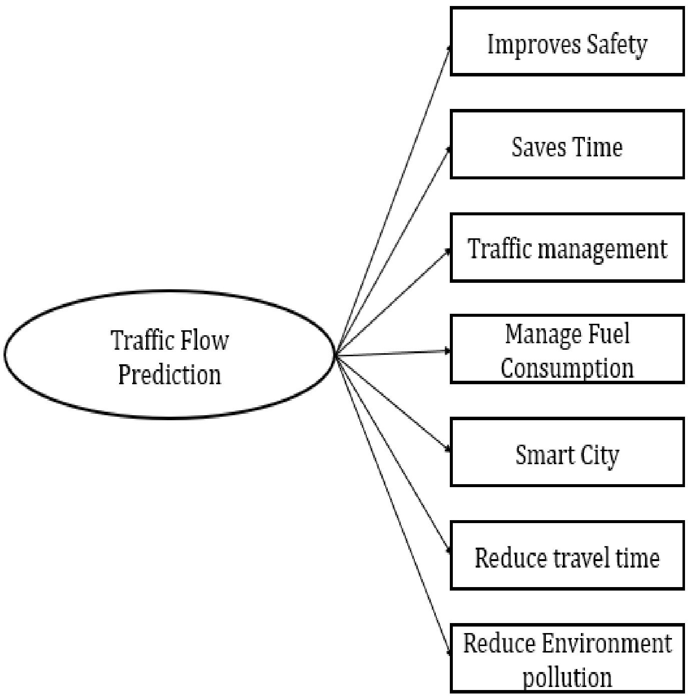


Fig 1: Benefits of implementing the traffic prediction system.

II. Literature Review:

In the last few years, transportation researchers have carried out a lot of research on the occurrence of traffic congestion in road transportation and the prediction of the traffic flow at various road networks. The traffic predication is usually based on real traffic flow, historical data collected from various sensors and captured by cameras and GPS. This helps in having a safe driving assistance for autonomous vehicles and reduced travel time [4]. Analysis using Random Forest algorithm, predicting the root mean square error, calculates the mean absolute error is described to predict the machine learning models. [5] has proposed a method called “scalable” that used to predict the traffic congestion of vehicles in a grid framework. Anwar and co-workers applied a spectral clustering-based method to supervise traffic congestion [6]. Considering the traffic flow density and different types of roads, Liang and his team developed a novel prediction model capable of estimating the next-time step traffic volume using a single road traffic segment to clarify traffic congestions using traffic flow variables such as the current inflow, outflow, and traffic volume, etc. [7]. The authors of [8] attempted to handle the traffic control problem using an ML system to deal with traffic difficulties. The authors used the Q- learning RL technique to manage traffic lights and created a simulation environment called Simulation of Urban Mobility (SUMO). SUMO allows you to watch cars in motion, control the vehicle's delay time, and alter the delay time. [9] developed a model

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 4, August : 2023
combining artificial neural networks and root mean squared error. Both were used as a metric by applying singular point probabilities. Traffic congestion has become a global pandemic that transportation researchers are racing against time to improve the effectiveness of intelligent transportation systems. Some researchers have been able to achieve good results when it comes to traffic flow prediction. A practical traffic flow parameter prediction model was created for traffic flow conditions estimations. An autoregressive model was combined with other predictive models [11][12] The authors of [13] developed a unique traffic prediction approach based on DL with the least prediction error and introduced the LSTM model. The dataset for this study was real-world traffic big data from performance measurement systems (PeMS). The number of optimized parameters used in this study should be increased. In addition, the model training time must bemanaged.

III. Methodology

a. **Dataset:** The collection of necessary data is identified; the following are those datasets that help us in the traffic prediction: This project's core dataset was a thorough traffic forecast record hosted on Kaggle [14]. This data set comprises information on 48120 traffic datasets collected between 2015 and 2017. For testing, 11808 records are used. It has five columns that collect information such as the date, time, vehicle ID, and junction ID. Through feature engineering, we generated an enhanced dataset with two additional attributes: day and weekend.

b. **Attributes of Dataset:** Few specific data attribute that contributes to the prediction are taken from the Kaggle dataset, they are as follows: Date, Time, Vehicle Id, Junction Id. These attributes are listed in our dataset and are the basic attributes used in analysis. We have used the featured engineering techniques to obtain accuracy. This helps in adding up few more attributes that contribute to the improvement in the results. The features that are added here are the Day and Weekend attribute, the day attribute specifies the day of week and weekend attribute is used to check if is a weekend, which is typically a holiday. Days are enumerated from 1 to 7. Weekend is a binary attribute.
If the day of week are Saturday and Sunday, then it is marked as holiday and is assigned with the value 1 and rest of the days are marked with 0. The traffic pattern is analyzed with these enhanced features.

c. **Proposed System:** The proposed model's fundamental premise has two main tasks. The primary job is to accurately anticipate traffic based on historical and advanced data for the specified location. The planned work architecture is depicted in Figure 2. The augmented feature set was created using feature engineering techniques.

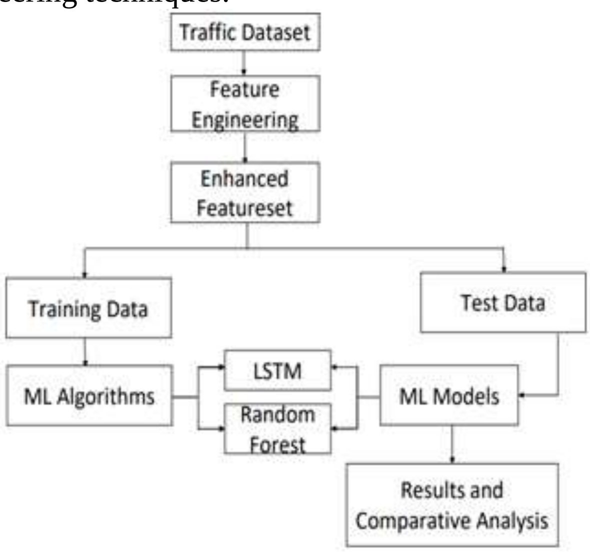


Fig 2: System Architecture

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 4, August : 2023

The process of changing raw data into features appropriate for machine learning models is known as feature engineering. It is the process of picking, extracting, and converting the most important features from accessible data to develop more accurate and efficient machine learningmodels.

IV. Machine Learning Methods:

Machine learning (ML) allows you to build prediction models that consider vast amounts of heterogeneous data from many sources. Numerous studies on the use of ML algorithms to forecast road traffic have been done.

a. **Random Forest:** To obtain accurate predictions, the random forest method generates numerous decision trees and integrates their data. Given enough training data, it is quite fast and can yield effective results. When applied to the problem of traffic congestion, this approach has an accuracy of 87.5 percent. Weather conditions, time period, specific road conditions, road quality, and vacations are employed as model input variables in this scenario.
b. **Long Short-Term Memory**: this algorithm basically works on a time series data, it considers the data generated in previous hour and the current hour data. With these data the analysis is performed. LSTM RNN variants that handle the vanishing gradient problem. According to a study that examined the performance of these models, the GRU model is more accurate in predicting traffic flow and is easier to train. A vast variety of papers recommend developing different types of NN models for traffic prediction, such as graph neural networks, fuzzy neural networks, Bayesian neural networks, and others, as well as hybrid techniques that integrate two or more algorithms. As of today, no single best technique has been discovered that can be utilized in all instances and produce the most accurate forecasts.

V. Results and Discussion

Traffic congestion is regarded as one of the most serious concerns confronting urban regions. As computing technologies progress, various advanced concepts are employed to estimate traffic congestion and flow. This study employs two machine learning concepts, LSTM and RF, to forecast traffic flow. The traffic data set was obtained from the internet and enhanced by adding two attributes. Apply machine learning concepts to forecast traffic flow after preprocessing.
The proposed system's performance is assessed using the metrics accuracy, precision, and RootMean Squared Error (RMSE), as given in Table 1.

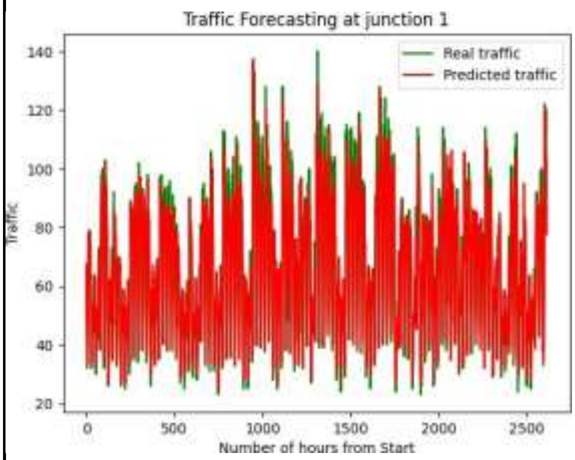
Models	**Basic Attributes**	**Enhanced Attribute**
LSTM	6.16 RMSE	4.57 RMSE
Random Forest	77.82 Accuracy	94.96 Accuracy

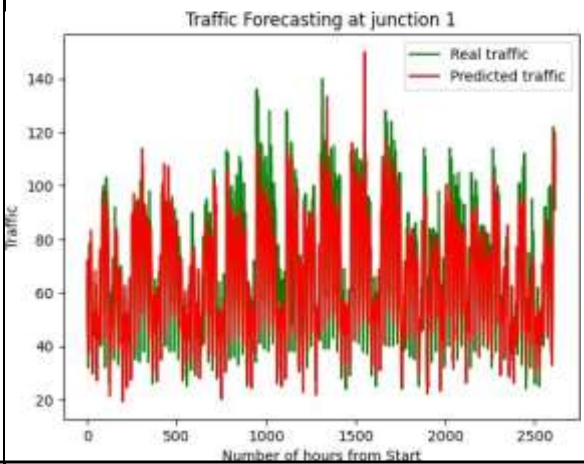
Table 1. Comparison of existing basic features approach and our enhanced features

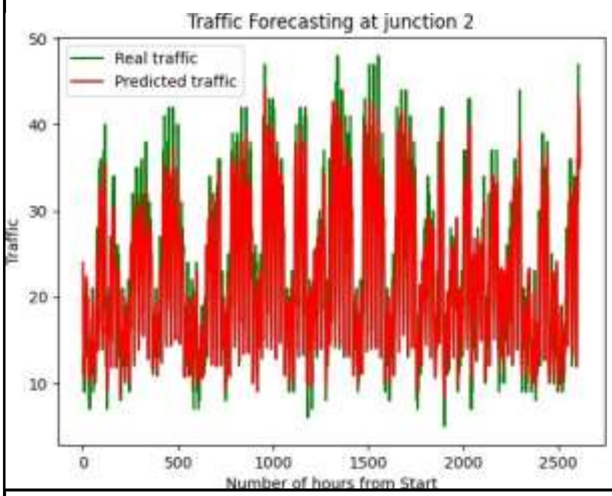
Models have been trained for both the basic and advanced datasets that we developed usingfeature engineering. We added two additional attributes to the advanced dataset. Both the basic and advanced datasets were subjected to a comparative study. We discovered that theaccuracy of the advanced dataset has improved.
Figure 3 shows the Traffic forecast i.e., real traffic v/s predicted traffic in all the junctions we have taken in our study. Using LSTM, we calculated Root mean square error for both the basic and advanced dataset. In the table we can observe that the error has been reduced for advanced dataset compared to basic dataset. For Random Forest model accuracy-score has beencomputed and the accuracy has been improved for the advanced dataset.

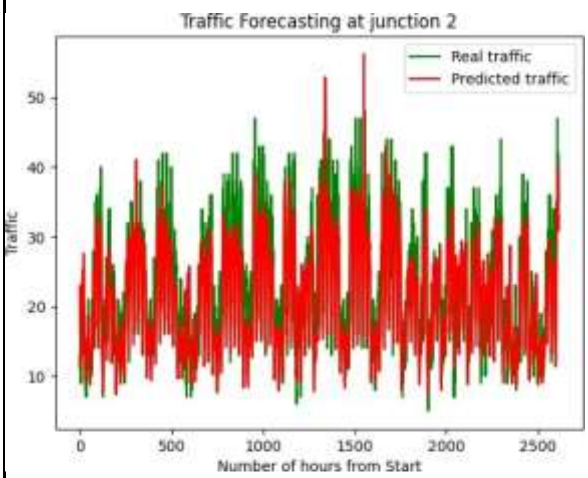
LSTM_ Basic	**LSTM_ Advanced**

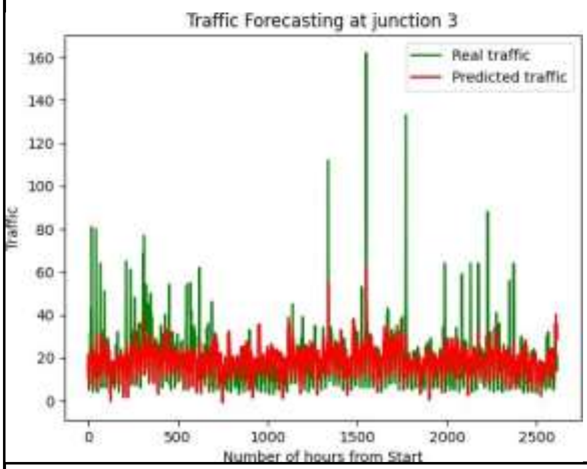
Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 4, August : 2023

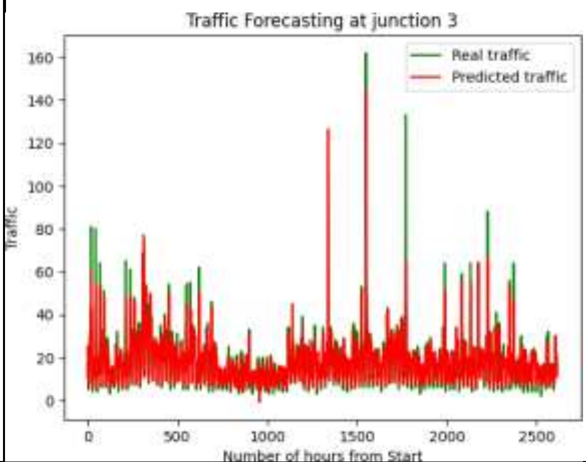












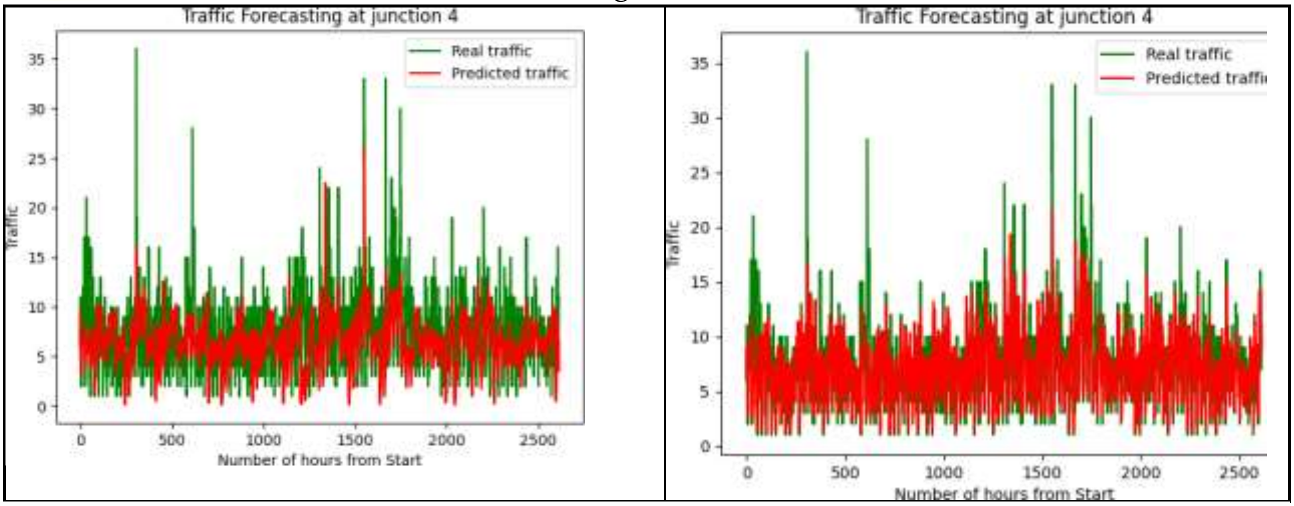


Figure 3: Traffic Forecasting at Junctions obtained by Basic LSTM and Advanced LSTMtechniques.

Conclusion

The traffic flow forecast system employs machine learning techniques. The proposed models were trained and tested using two datasets. One is a publicly available dataset, and we used feature engineering to generate a new dataset. Experiment results demonstrate that the models with enhanced dataset outperformed traditional dataset models.

The results of prediction obtained by applying LSTM are illustrated in Figure 3, the left-hand side graphs are the results with the original dataset and the right-hand side results are obtained with the enhanced dataset. It can be observed that the predicted values are closer to the actual values with feature engineered dataset. As observed in Table 1, the RMSE value with LSTM reduces with the enhanced dataset. A significant improvement in accuracy can also be observed with the RF method. The results illustrate the power of feature engineering. With a better understanding of the data, feature enhancements can be made and they can prove to be great boosters for the performance of ML models. The technique of feature engineering can be performed for the data pertaining to any domain and is very likely to yield better results.

Forecasting with such a great accuracy can assist individuals or users in gauging road traffic more easily in advance, and they can make a better-informed decisions in terms of navigation paths. With the ideas of Smart Cities gaining momentum, such systems will prove to be of great use in managing and monitoring city's traffic.

Future Scope:

- Build the model with neural network, by adding neurons and hidden layers to examine thecomplex traffic flow.
- The LSTM could be combined with CNN to learn the patterns of images to interpret thetraffic speed.

References

[1] Chen, C.; Liu, B.; Wan, S.; Qiao, P.; Pei, Q. An Edge Traffic Flow Detection Scheme Based on Deep Learning in an Intelligent Transportation System. IEEE Trans. Intell. Transp.Syst. 2021, 22, 1840–1852.

[2] Boukerche, A.; Wang, J. A performance modeling and analysis of a novel vehicular trafficflow prediction system using a hybrid machine learning-based model. Ad. Hoc. Netw. 2020, 106, 102224.

[3] Ferreira, Y.; Frank, L.; Julio, E.; Henrique, F.; Dembogurski, B.; Silva, E. Applying a Multilayer Perceptron for Traffic Flow Prediction to Empower a Smart Ecosystem. In Computational Science and Its Applications; Springer: Cham, Switzerland, 2019; pp. 633–648.

[4] Aditya Krishna K.V.S et al, Smart traffic analysis using Machine Learning, IJEAT, 2019.
[5] D'Andrea, E.; Marcelloni, F. Detection of traffic congestion and incidents from GPS traceanalysis. Expert Syst. Appl. 2017, 73, 43–56.
[6] Anwar, T.; Vu, H.L.; Liu, C.; Hoogendoorn, S.P. Temporal tracking of congested partitions in dynamic urban road networks. Transp. Res. Rec. 2016, 2595, 88–97.
[7] Liang, Z.; Wakahara, Y. Real-time urban traffic amount prediction models for dynamic route guidance systems. EURASIP J. Wirel. Commun. Netw. 2014, 2014, 1–13.
[8] Boukerche, A., & Wang, J. (2020). Machine learning-based traffic prediction models for intelligent transportation systems. Computer Networks, 181, 107530.
[9] Qian, Z.S.; Li, J.; Li, X.; Zhang, M.; Wang, H. Modeling heterogeneous traffic flow: A pragmatic approach. Transp. Res. Part B Methodol. 2017, 99, 183–204.
[10] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein generative adversarial networks, in: D.Precup, Y.W. Teh (Eds.), Proceedings of the 34th International Conference on Machine Learning, Vol. 70, PMLR, 2017, pp. 214–223.
[11] Kong, X.; Xu, Z.; Shen, G.; Wang, J.; Yang, Q.; Zhang, B. Urban traffic congestion estimation and prediction based on floating car trajectory data. Future Gener. Comput. Syst.2016, 61, 97–107.
[12] Davoodi, N.; Soheili, A.R.; Hashemi, S.M. A macro-model for traffic flow with consideration of driver's reaction time and distance. Nonlinear Dyn. 2016, 83, 1621–1628.
[13] 19. Xu, Y., Xu, W., Yin, F., Lin, J., & Cui, S. (2017, December). High-accuracy wireless traffic prediction: A GP-based machine learning approach. In GLOBECOM 2017- 2017 IEEE Global Communications Conference (pp. 1-6). IEEE.
[14] 20. Aibin, M., Chung, N., Gordon, T., Lyford, L., & Vinchoff, C. (2021, June). Onshort-and long-term traffic prediction in optical networks using machine learning. In 2021 International Conference on Optical Network Design and Modeling (ONDM) (pp. 1-6). IEEE.
[15] Y. Lv, Y. Duan, W. Kang, Z. Li, F.-Y. Wang, et al., Traffic flow prediction with big data: A deep learning approach, IEEE Trans. Intell. Transp. Syst. 16 (2) (2015) 865–873. [16]. Deng, D. Yu, et al., Deep learning: methods and applications, Found. Trends Signal Process. 7 (3–4) (2014) 197–387.
[17] Gadze, J.D.; Bamfo-Asante, A.A.; Agyemang, J.O.; Nunoo-Mensah, H.; Opare, K.A.B.An Investigation into the Application of Deep Learning in the Detection and Mitigation of DDOS Attack on SDN Controllers. Technologies 2021, 9, 14.