iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43305

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Efficient Model to Predict the Traffic Congestion through Deep Learning

Hrushikesh Shelke[1], Vivek Lokhande[2], Akshay Pawar[3], Mahesh Solanke[4], Prof. Tanvi Ghodke[5]
[1, 2, 3, 4, 4]*Computer Dept, KJCOEMR, Pune*

Abstract: ***Traffic in urban centers has considered a worsening issue across the world, impacting development and people's daily lives. It also occurs in large and medium regions of India, posing a serious threat to the country's progress. Humans could gain knowledge from the administration and management of urban vehicular networks that road congestion can be ameliorated or diminished if we can anticipate road congestion which will occur within few moments or has already occurred in a few seconds and implement prompt, adequate traffic abatement methodologies. As a consequence, traffic congestion prediction is crucial for improving the transportation system's energy consumption and dependability. A variety of strategies to this problem have been investigated and discussed in this research article for this goal. This has enabled the successful implementation of our solution for a traffic congestion prediction system based on Artificial Neural Networks and Decision Making. The result analysis has been crucial in the realization of the improved accuracy of the prescribed approach.***
Keywords: Artificial Neural Network, Deep learning, Traffic Congestion prediction ,Preprocessing , Segarigation.

I. INTRODUCTION

Congestion is one of the most pressing concerns we face on a daily basis. The current design of signalized crossings is a conceptual framework that cannot adjust to traffic changes. The stationary program uses a pre timer based system, which may cause in lengthier signal wait periods. During peak hours, certain sections are more congested, while others are less congested. Drivers on the packed road must wait longer than is necessary since the stationary timer system does not adapt for rising congestion.

With the rapid development in population and economic growth, there are a rising number of urban vehicles on the streets. Cities' accelerated expansion is intensifying. Suburban transportation is becoming progressively complicated, and traffic congestion in metropolitan areas is becoming especially noticeable. One of them is traffic congestion on the roadways. If traffic difficulties in urban areas are not solved quickly, they will culminate in increasingly crowded areas and, in severe circumstances, traffic immobilization. The very first approach in resolving traffic congestion is to prevent it from arising in the initial instance. As a consequence, generating an accurate current traffic estimate allows for the formulation of targeted preventative measures and early alerts. Time is the utmost important commodity in today's hyper society. The ability to organize your time well is vital for accomplishment. A traffic junction is one of the locations that will be several times throughout the day yet will be unable to handle. A traffic signal with a defined frequency in a junction might provide a few vacant lanes while also packed sections. International and domestic experts have made significant headway in conducting prediction studies in order to anticipate the challenge of traffic congestion. Time series approach, parametric regression model, and artificial neural approach have all been deployed to roadway traffic congestion forecasting by academics from many professions and domains. Over the last several decades, there has been a lot of scientific curiosity in traffic congestion operation and supervision. This is due to the increasing requirement for movement, which is a straightforward result of technological development and urbanization. Intelligent Transportation Systems have showed promise in delivering efficient and appropriate urban traffic administration and monitoring throughout the years. The reliability of traffic information supplied to vehicular traffic, regulatory entities, or enterprises is a significant factor in the achievement gained via the use of these systems for urban transportation planning. As a result, short-term traffic estimation is an important study problem that has been investigated in multiple academic domains, with involvement from statistics, computer engineering, and supply chain management. The great proportion of short-term traffic forecasting techniques, however, concentrates on linkage or junction-level traffic flow prediction. Although this strategy has resulted in significant benefits in urban road congestion control, it is rejected when used to multi-modal, comprehensive urban network-wide transportation planning. Large-scale traffic administration is advantageous because it also assists traffic controllers in efficiently managing the traffic network and allocating services. Recognizing how traffic changes inside a network over time is crucial for properly managing road congestion inside an urban

©IJRASET: All Rights are Reserved

network model. This research study devotes section 3 to an overview of previous work as a literature survey; section 4 elaborates on the suggested methodology; section 5 analyzes the system's performance; and lastly, section 6 finishes the article with hints for further refinement.

II. LITERATURE SURVEY

To create the MFTCPV traffic congestion framework, which includes a raw data layer, a data processing layer, and a data presentation layer, L. Li et al. [1] proposed the LSTM-SPRVM deep model for the prediction of traffic congestion characteristics. The system was tested using real data from Whitemud Drive in Canada. Most heuristic techniques for parameter optimization of RVM have the drawback that the convergence speed is excessively quick in the early iterations, resulting in a loss in population diversity in the latter rounds.To overcome the problem, the authors aim to integrate RVM with complicated heuristic algorithms like the Adaptive Genetic Algorithm (AGA).Experiments reveal that LSTM-SPRVM outperforms other approaches in terms of accuracy.

TrafficNet is a proposal by P. Wang et al. to bridge the gap between current sophisticated deep learning algorithms and practical implementation.Three deep TrafficNetsdepend on residual learning are constructed and pre-trained using the ILSVRC-2012 dataset after rounds of testing and training [2]. The traffic application receives both the network structure and the weight. A traffic picture collection was created from 14470 active security cameras in Shaanxi Province, China, to fine-tune the TrafficNet. The original collection is made up of images that represent a wide range of lighting, weather conditions, and scenarios. The training picture set for the crowded and uncongested roads is kept in a balanced data size after pre-processing and augmentation. To pre-train transferred TrafficNet, this relatively little dataset is utilized. The detection is shown in the validation dataset.

The optimized GRU technique is suggested by J. Zhao et al. to estimate truck travel speed on metropolitan express routes in non-recurrent crowded circumstances. This document intends to offer truck drivers a useful information service, which depends on data generated by trucks traveling on Beijing's Sixth Ring Road [3].The original data for this research comes from Beijing truck driving. This work provides a GPS map-matching method that can fulfill the accuracy and efficiency of matching while addressing the problem of a huge volume of original data. Then, from the original data, the truck data moving on Beijing's Sixth Ring Road is removed.The GRU model is fed the derived historical time series of average trip speed. Adam, Adaddala, and Rmsprop are used to improve the weights in the GRU model, addressing the problem that the frequently used weight optimization method SGD cannot adaptively modify the learning rate.

J. S. Angarita-Zapata et al. created a taxonomy to classify families of TF problems using supervised regression learning. This is how the taxonomy is constructed. The challenges are classified using two sorts of criteria: traffic and modelling requirements. The first handles transportation-related concerns like data source type, contextual, and forecast geographical coverage. the input and output parameters, as well as the projected time horizons and data time step Unexpectedly, the second is data should be employed based on a set of criteria [4]. Output values, as well as prediction algorithms, may be described using DPP approaches. The most common forms of TF problems include the use of point detectors as the main source of information. To overcome the problems described above, traditional machine learning methodologies like as NNs, regression, and instance-based techniques are widely utilized.

Z. Bartlett et al. used distinct temporal data segments to study different magnitudes of temporal patterns to see how contextual temporal data affect prediction accuracy. In addition, unlike prior prediction models, the authors have suggested a dynamic temporal contextual framework that can dynamically include both short and long-term temporal trends [5]. This is accomplished by using several temporal data segments as input features, and the model may dynamically determine which is significant for regression to offer an accurate forecast in real-time via online learning. Using an existing real dataset, the distinct temporal data segments and suggested framework were examined and compared to a comparable prediction model.

Under typical conditions, the proposed LSTM network by J. Zheng et al. [6] can reliably anticipate traffic flow depending on reasonably steady time series. The traffic system on roadways, on the other hand, is stochastic and complicated, and it is frequently influenced by unusual circumstances such as severe weather, traffic accidents, and huge events. The efficiency of traffic flow forecasting is examined using three sample time-series analytic methods: ARIMA model, backpropagation neural network (BPNN), and LSTM. Because of its tremendous potential for online deployment, the ARIMA model has been widely adopted to anticipate traffic flow. Without a precise mathematical model, the BPNN is the most prevalent approach for processing traffic data. However, BPNN's shallow structure makes it inefficient in handling large amounts of data.

S. Feng et al. developed a traffic forecast hybrid model that incorporates both regional and global spatial correlations (T-LGGCN). On the one hand, in the setting of a global road infrastructure, the suggested model may take into consideration the spatial correlation among distant sensors [7].

The authors, on the other hand, may improve the capability of local spatial correlation analysis by separating node feature analysis and GCN local aggregation. When contrasted to the baselines, the outcomes show the importance of fully incorporating global and local spatial connection. To summarize, the T-LGGCN model accurately incorporates traffic flow's spatial-temporal correlation and may be utilized to forecast traffic flow.

J. Mena-Oreja et al. published the most thorough assessment of deep learning-based traffic forecast algorithms to date. The research examines current traffic prediction approaches and suggests some novel models for predicting traffic flow and speed. For a shared benchmark, the comparison is done using the same datasets [8]. The accuracy of the predictions is compared under general traffic conditions as well as for the onset and dissolution of traffic jams. This is a significant addition since the authors showed that effectively forecasting traffic, in general, does not always indicate that a deep learning approach can accurately predict traffic in congested settings when traffic factors vary more over time.In addition, the research found that error recurrent models outperform other deep learning prediction approaches when it comes to predicting traffic speed and flow in general and in congested areas.

Z. Bartlett et al. used distinct temporal data segments to study different magnitudes of temporal patterns to see how contextual temporal data affect prediction accuracy. In addition, unlike prior prediction models, the authors have suggested a dynamic temporal contextual framework that can dynamically include both short and long-term temporal trends. This is accomplished by employing several temporal data segments as input features, and the model may dynamically determine which is significant for regression to deliver an accurate forecast in real-time through online learning [9]. Using an existing real dataset, the distinct temporal data segments and suggested framework were examined and compared to a comparable prediction model.

K. N. Bui proposes the RSDC issue, a new methodology for traffic density classification using road sound datasets inspired by recent breakthroughs in urban sound classification using CNN-based approaches.For the categorization of traffic conditions, traffic sound data is collected on major thoroughfares at various time intervals. For the feature extraction procedure, a novel CNN architecture with five convolutional layers is presented [10]. Then, to enhance accuracy, numerous well-known machine learning models are used. The experiment shows that the proposed technique for traffic condition classification using road sound datasets has promising results.

III. PROPOSED METHODOLOGY

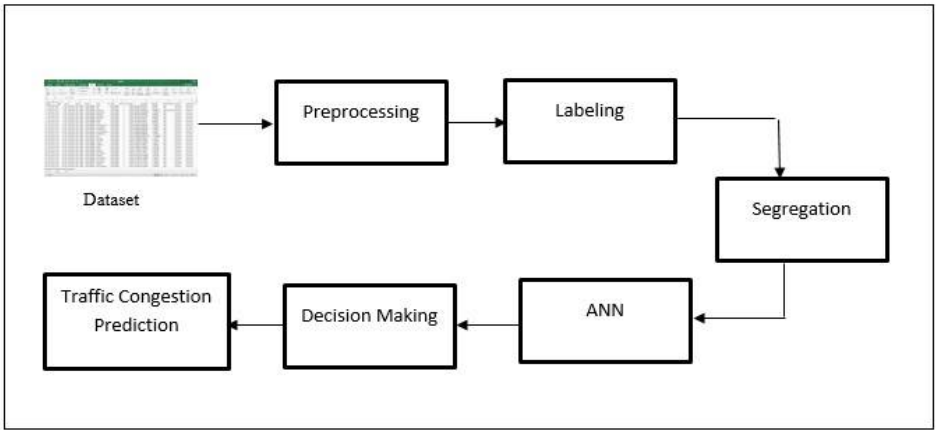


Figure 1: Traffic Congestion Prediction System Overview

The presented system for traffic congestion prediction is depicted in the above figure 1. The steps that are involved in the development of the model is narrated with the below mentioned steps.

1) *Step 1: Data collection and Data feeding* –The proposed model downloaded the traffic dataset from the URL : https://www.kaggle.com/fedesoriano/traffic-prediction-dataset. This dataset is stored in the form of aworkbook in .csv format. The dataset contains attributes like Datetime, Juntion, Vehicles and Sensor ID. An interactive GUI is developed to browse and feed this dataset to the proposed system. Once it is fed to the system, then, by using the pandas library of python the proposed model reads the dataset into its object.

2) *Step 2: Preprocessing and Labeling*- Here in this step the number of rows and columns is identified in the dataset. The data set contains 48120 Rows and 4 columns as attributes as it is mentioned in the previous step. Then the missing values are checked in every column and then the Senor ID column is dropped as it not having any role in the prediction of the traffic. The Datetime Column contains the concatenated date and time, which is then split into individual attributes like year, month,Date,hour and Day. These five attributes are added toJunction number, vehicles and Datetime to form 8 attributes. Which are labeled with a list to use in the next step of Data segregation.
3) *Step 3: Data Segregation* –Here in this step of data segregation the data are divided into two lists like X and Y. X contains some attributes like 'Year', 'Month' ,'Date', 'Hour', 'Day' and Y contains Vehicles as attributes. Then these lists are split using the split function for X and Y attributes with test size of 2% with random state on. Then four lists are created like X_train, X_test, Y_train and Y_test. Now a minMax scalar is used to create its object. Then this object is used to normalize the data for all X_train, X_test, Y_train and Y_test lists respectively to feed to the Artificialneural network of the next step.
4) *Step 4: Artificial Neural Network*- This is the core step of the proposed model which uses keras and tensorflow libraries for implementation of the ANN. Where a sequential neural network object is createdin the beginning. A dense layer of 200 kernels with uniform initializer kernels are added with activation function " Relu" is added with input dimension of 5 as mentioned in the previous step. This is followed by a Dense layer with a 20% dropout rate. In the second layer of the ANN a dense layer of 150 kernels with uniform initializer kernels are added with an activation function " Relu" is added with input dimension of 5 as mentioned in the previous step. This is followed by a Dense layer with a 20% dropout rate. This same thing is repeated for 3rd and 4th layer of ANN with 100 and 50 kernel size respectively with dropout percentage of 20%.

In the end all the trained tensors are collected in a single unit with sigmoid Activation function to enrich the training model. Finally an "Adam" Optimizer is used to compile and fit the model along with the 150 batch size for 100 Epochs.

The used Sigmoid and Relu activation function are depicted in equation 1 and 2.

$$Sigmoid = \frac{1}{(1+e^{-x})} \quad \text{_____(1)}$$

$$Relu=max(0,x) \quad \text{_____(2)}$$

Where x is the input attributes values

The architecture of the artificial neural network is depicted in the figure 2.

ANN Architecture	
Layer	Activation
Sequential	
Dense , Kernel=200, Dim=5	relu
Dropout 20%	
Dense , Kernel=150, Dim=5	relu
Dropout 20%	
Dense , Kernel=100, Dim=5	relu
Dropout 20%	
Dense , Kernel=50, Dim=5	relu
Dropout 20%	
Dense , Kernel=1	Sigmoid
Adam Optimizer	
Batch size 150	
Epochs 100	

Figure 2: Architecture of the artificial Neural network

5) *Step 5: Decision Making for Traffic Congestion Prediction:* Here in this process the X_test and Y_Test lists with months and dates are fed for the estimation of the prediction of flood and landslide. Then a decision is taken to perform the inverse transformation and reshape the list to predict the traffic congestion in the different junctions.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. RESULTS AND DISCUSSIONS

The proposed model for the traffic congestion prediction is deployed in window machine configured with Core i5 processor and with 8 GB primary memory. To implement the model Python programming language is used with Spyder IDE. Keras and Tensorflow deep learning libraries are imported for the purpose of coding artificial neural network.

The proposed model is evaluated for the estimation of the RMSE(Root mean square Error).This is a continuous evolution of the error rates between the actual values of Traffic Congestion and predicted values of the Traffic Congestion. This is evaluated based on the below equation 3.

$$\mathbf{RMSE}_{fo} = [\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$$

Where,

∑ - Summation

(Zfi - Zoi) ^{2} - Differences Squared for the expected and predictedTraffic Congestiondata.

N - Number of predicted data.

The obtained results are depicted in the below two graphs in the figure 3 and 4.

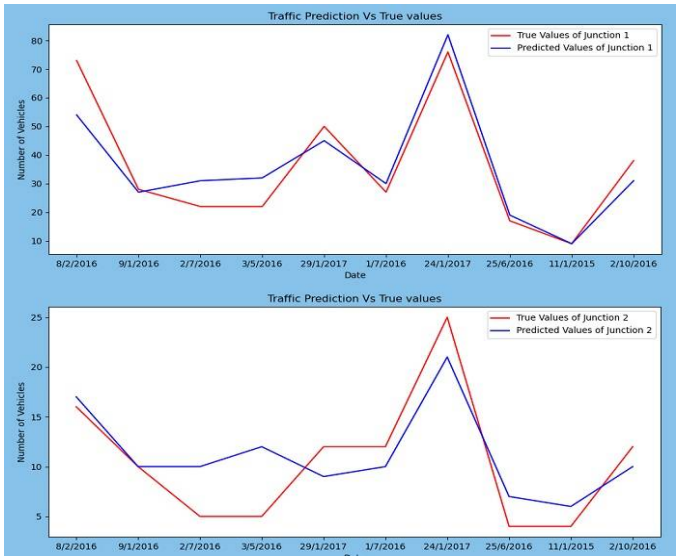


Figure 3: RMSE Estimation for Congestion Prediction ofJunction 1 and 2

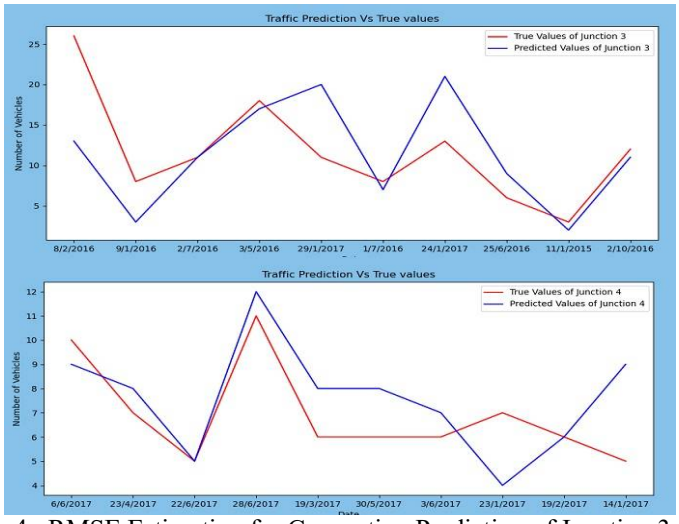


Figure 4: RMSE Estimation for Congestion Prediction of Junction 3 and 4.

On observing the graphs it is evident that X- Axis of the line indicates the Date Attributes and Y- axis represents the number of the vehicles that are passed through that junction. Here there are two line RED and BLUE. A Red colored line indicates the true values and Blue colored indicates the predicted number of vehicles that are passed to produce congestion. On close observations the RMSE for the 4 junctions were obtained as Root Mean Square Error for Junction 1 is: 6.6867 , Root Mean Square Error for Junction 2 is: 2.8016, Root Mean Square Error for Junction 3 is: 8.0978 , and Root Mean Square Error for Junction 4 is: 2.4259.
These values eventually indicate the best implementation of the model for prediction of the traffic congestion in the different junctions of the city.

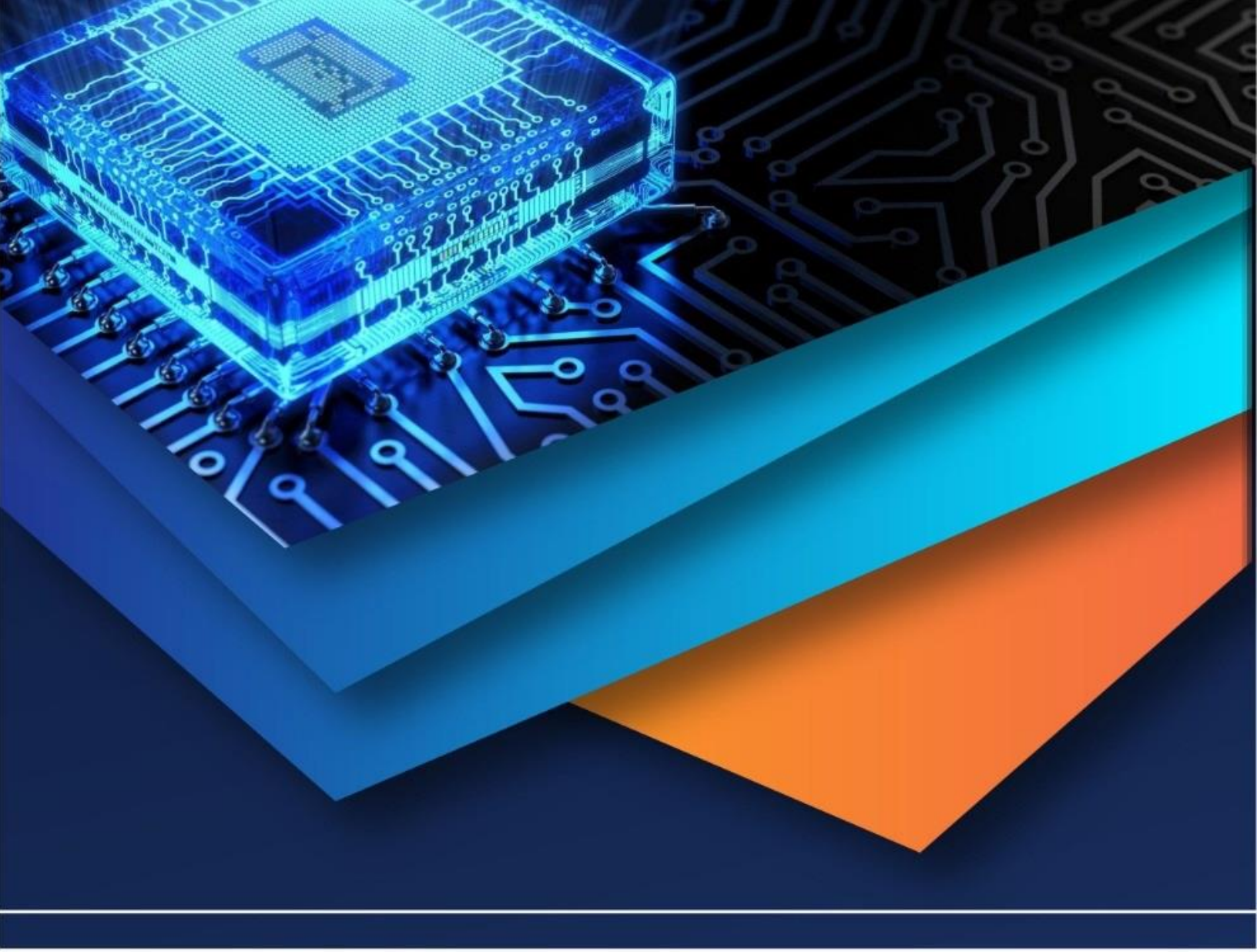
V. CONCLUSION AND FUTURE SCOPE

The presented approach for the purpose of achieving an effective Traffic congestion prediction through the use of deep learning methodologies has been elaborated in this research article. One of the most problematic occurrences in the urban life is prevalence of the traffic congestion that is rampant and present all across the city on certain hours. This traffic congestion is seemingly random and can be extremely complex to determine. Therefore, this approach utilizes a traffic dataset that is provided to the system as an input. This dataset is initially preprocessed to remove the redundant data and incomplete data. The preprocessing further improves the execution accuracy as it eliminates the unnecessary elements form the data. The Preprocessed data is then subjected to the labeling task where the attributes are converted into their respective integer forms. The preprocessed and labeled data is then provided for the segregation procedure. The preprocessed, labeled and segregated data is provided to the Artificial Neural Networks approach which performs hidden layer and output layer estimation which results in the probability scores. These probability scores are effectively classified using the if-then rules of the Decision making approach. This results in highly accurate traffic congestion prediction. The approach has been effectively evaluated for its performance which has achieved satisfactory results.
For the purpose of future work the approach can be implemented on the live data and a suitable alert can be generated to the users through a mobile application.

REFERENCES

[1] L. Li, H. Lin, J. Wan, Z. Ma, and H. Wang, "MF-TCPV: A Machine Learning and Fuzzy Comprehensive Evaluation-Based Framework for Traffic Congestion Prediction and Visualization," in IEEE Access, vol. 8, pp. 227113-227125, 2020, DOI: 10.1109/ACCESS.2020.3043582.
[2] P. Wang et al., "Regional Detection of Traffic Congestion Using in a Large-Scale Surveillance System via Deep Residual TrafficNet," in IEEE Access, vol. 6, pp. 68910-68919, 2018, DOI: 10.1109/ACCESS.2018.2879809.
[3] J. Zhao et al., "Truck Traffic Speed Prediction Under Non-Recurrent Congestion: Based on Optimized Deep Learning Algorithms and GPS Data," in IEEE Access, vol. 7, pp. 9116-9127, 2019, DOI: 10.1109/ACCESS.2018.2890414.
[4] J. S. Angarita-Zapata, A. D. Masegosa, and I. Triguero, "A Taxonomy of Traffic Forecasting Regression Problems From a Supervised Learning Perspective," in IEEE Access, vol. 7, pp. 68185-68205, 2019, DOI: 10.1109/ACCESS.2019.2917228.
[5] Z. Bartlett, L. Han, T. T. Nguyen, and P. Johnson, "A Novel Online Dynamic Temporal Context Neural Network Framework for the Prediction of Road Traffic Flow," in IEEE Access, vol. 7, pp. 153533-153541, 2019, DOI: 10.1109/ACCESS.2019.2943028.
[6] J. Zheng and M. Huang, "Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning," in IEEE Access, vol. 8, pp. 82562-82570, 2020, DOI: 10.1109/ACCESS.2020.2990738.
[7] S. Feng, J. Huang, Q. Shen, Q. Shi, and Z. Shi, "A Hybrid Model Integrating Local and Global Spatial Correlation for Traffic Prediction," in IEEE Access, vol. 10, pp. 2170-2181, 2022, DOI: 10.1109/ACCESS.2021.3139934.
[8] J. Mena-Oreja and J. Gozalvez, "A Comprehensive Evaluation of Deep Learning-Based Techniques for Traffic Prediction," in IEEE Access, vol. 8, pp. 91188-91212, 2020, DOI: 10.1109/ACCESS.2020.2994415.
[9] Z. Bartlett, L. Han, T. T. Nguyen, and P. Johnson, "A Novel Online Dynamic Temporal Context Neural Network Framework for the Prediction of Road Traffic Flow," in IEEE Access, vol. 7, pp. 153533-153541, 2019, DOI: 10.1109/ACCESS.2019.2943028.
[10] K. N. Bui, H. Oh and H. Yi, "Traffic Density Classification Using Sound Datasets: An Empirical Study on Traffic Flow at Asymmetric Roads," in IEEE Access, vol. 8, pp. 125671-125679, 2020, DOI: 10.1109/ACCESS.2020.3007917.

*****



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)