

Traffic Prediction System Using Machine Learning (ML)

Priyadharshini S
Department of Computer Science & Engineering,
SRMIST, Vadapalani,
Chennai, India
ps4267@srmist.edu.in

Ramita K
Department of Computer Science & Engineering,
SRMIST, Vadapalani,
Chennai, India
rk4337@srmist.edu.in

Sanjaya D T
Department of Computer Science & Engineering,
SRMIST, Vadapalani,
Chennai, India
ds7057@srmist.edu.in

Mr. Rajavel M
Department of Computer Science & Engineering,
SRMIST, Vadapalani
Chennai, India
rajavelm@srmist.edu.in

Abstract--- Traffic has its own consequences and the congestion from the same leads to consumption of time, energy, frustration and gasoline almost every other day as a routine. On one side, the government is also affected by congestion from traffic as goods transportation should be done with easy flow, reduction of pollution should be ensured in certain places along with the safety measures. Globally congestion is considered a major problem that has the capability to impact every level persisting in the society. Every driver would know about the causes of traffic jams that happen usually in a journey. Unexpected delays happen due to road constructions and works, accidents. The speed of traffic flow goes down due to bad conditions of weather. When the timing of the traffic signal below par, the capacity is limited in smaller roads. A network of roads where the capacity is also seem to be limited is what usually leads to increasing congestion in traffic globally. The main objective is to classify the traffic state. In the first step we the data is analyzed which then is later pre-processed. Application of a machine learning algorithm is the third step here over dataset of traffic prediction and the performance and accuracy is calculated from the result after applying the suitable algorithm.

Machine learning helps a computer to make decisions from previously acquired data and training over them. This concept comes under artificial intelligence where the necessity to program the computer for a purpose each time especially is not required. Programs are developed in computer and presented new data over which the machine learning algorithms are applied. The programs are usually written in python for machine learning. Then the prediction is achieved by training the presented data and testing them using the suitable machine learning algorithm over the dataset. There are three types of machine learning. The first one being supervised learning deals with inputs labelled prior feeding to the machine. Whereas in unsupervised learning there is no labelling of data and the machine learns by the algorithm given after the input is clustered. The third one which is the reinforcement learning receives feedback after interacting with the environment that could be positive or negative and based on the feedback received, makes decision for the performance to be better.

Different algorithms in machine learning are used to perceive patterns and information for decision making by data scientists. These are done in python generally. Depending upon the learning method and predicting technique the algorithms are categorized and classified into two. The first classification is supervised learning whereas the second classification is unsupervised learning. The data are given as inputs in supervised learning. Prediction of the class of provided input that are data points is known as classification. Classes are also known as targets or labels for the input data. The approximation of variables of input(x) to variables of output(y) over a mapping function for the predictive modelling performed for the task of classification. Classification comes under the approach of supervised learning in statistics and machine learning where the computer learns from classifying the observations presented made newly after learning from the input values fed. Dataset can be a multi class or bi class that is used to determine an output only between two options. Handwriting recognition, speech recognition are some of the examples under problems of classification.

1. INTRODUCTION

The aim of the paper is to gather the information of traffic flow over time and accuracy by a tool that will be used for prediction after developing. Anything that can lead to a traffic jam such as signals in traffic, rallies roads getting repaired, accidents and so on are all a part of traffic environment as it affects the traffic state at any given time. To make a proper decision, a driver requires the information of the above mentioned factors and when these are gathered prior with approximation closely correct it takes the driver to easily decide and move on without delays. Autonomous automobiles will also be benefited by the same in future cases. When it comes to the analysis of data, deep learning and genetic algorithm are found to be the one main problem though the machine learning community does not deal much with them. The issues that did rise from the dataset complexity are dealt. The improvement in accuracy is also achieved after using the proposed algorithm over the existing algorithm. The integration of the application and web server is also made here. To increase the accuracy to a higher range the algorithm will be improvised.



Process Of Machine Learning

machine learning algorithm over the given mapping function which is $y=f(x)$ where x is the input value and y is the output value. The approximation should be performed over the mapping function that when a new input data is given to the machine it correctly predicts the output. Logistic regression, decision trees are some techniques that come under algorithms of supervised learning. Algorithm from supervised training trains the model with labelled inputs to give the particular output for the same. Further-more they can be put as classification problem. From the variables of attributes in a succinct model the dependent attribute value could also be predicted for the construction. Under the classification category the attribute that is dependent is found out to be numerical. After the values getting observed the model of classification concludes the result. There can be more than one input and more than one output. The outcome is of two choices in classification model.

B. Dataset preparation:

There are over 15000 data in the dataset. Further there were 5 classes after classification

Normal
MoveOn
Busy
Heavy
Worst

C. PROPOSED SYSTEM

To determine the state of traffic of a particular place a model is proposed that uses classification. The previous condition data is gathered and stored in a dataset that is utilized by the machine to understand the problem and learn. First the analysis is performed over the dataset. For the values missing specific functions are performed including outliers and such. After this process the model for classification is developed and the dataset is split into training and testing sets. The model understands and learns from the data fed as input values under the training set and upon application of the suitable algorithm from machine learning and uses this learning to classify the test data to predict traffic state. The performance is calculated after applying the algorithm.

2. LITERATURE REVIEW

Title: "Short term traffic flow prediction for a non-urban highway using Artificial Neural Network"

Author: Kranti Kumara, M. Paridab, V.K. Katiyar

Year: 2013

The historical data and raw data are collected from the past conditions of traffic flow and artificial neural network application is done to predict the current status over some processes. The volume of the traffic, speed and time and other factors like density and day of the week are taken as input variables. Each vehicle is taken into consideration as the speed varies for each and those are taken as input variables in a separate manner whereas the previous reports had taken the average or combined speed of the vehicles for the prediction of flow of traffic. Even though the interval for time were increasing somewhere between 5 to 15 minutes and the speed as input was taken separately, the performance was found to be consistent after applying the artificial neural network over it.

Title: "A Traffic Prediction Using Machine Learning: Literature Survey"

Author: Ji Yoon Kim

Year: 2021

It is difficult for the government to make a decision between providing subsidy to the workers not benefited or find some other way for the reduction of cost of commute by offering some percentage of transmit to them. To perform the same it is crucial to calculate accuracy. Usually there is a pattern followed in the traffic flow data even if it is a large set, so a model that is suitable for the handling is required in these cases. Thus, this study uses "LSTM AND GRU" for the traffic status prediction.

Title: "Short-Term Traffic Flow Prediction Based on CNN-BILSTM with Multicomponent Information"

Author: Weiqing Zhuang and Yongbo Cao

Year: 2022

Over the last few years the need for traffic information has become high and prediction of traffic status with accuracy in the field of transportation also became a requirement everywhere. For the construction in the area of smart cities a system such as "Intelligent transportation system" plays a major role. "Convolutional neural network and bi-directional long short term memory" model is used in order to build a model for prediction with multiple level of steps. The input variables were nothing but special aspects in the model for prediction and those were used for the extraction of series of time. The "BILSTM" model proved to be increasing the level of accuracy in prediction and after comparison with others models in mean absolute error and percentage and root mean square error. They were found to be reduced to an extent of 30.4%, 32.2% and 39.6% with the mentioned respective factors.

Title: "A Research of Traffic Prediction using Deep Learning Techniques"

Author: B.Karthika, N.UmaMaheswari, R.Venkatesh

Year: 2019

The study in the field of traffic flow prediction is essential and there are various technologies such as "Intelligent transport systems" in this modern world that can be used for the same. The construction of smart cities require traffic data too. Data mining, deep learning algorithm and techniques are used to and implemented to forecast traffic related data. This can be used to eliminate the information that is not of much requirement from the real time level. There are possibilities of these data to be "Big data", thus some methods are followed to acquire the accuracy subsequently.

Title: "Internet Traffic Prediction by W-Boost: Classification and Regression"

Author: Hanghang Tong, Chongrong Li, Jingrui He, and Yang Chen

Year: 2005

Management and control, network design, optimization are done with the help of traffic prediction via internet. There are certain issues in network traffic like the non-linear nature of it and "self-similar" that makes it difficult to perform and calculate with accuracy. This study brought up a proposal of a scheme that is named "W-Boost" that perceives and solves in two ways that either using classification method or regression method to predict traffic. Upon bringing up complexity that is low to the algorithm this scheme performs in a way that the non-linearity in the flow of traffic is captured. Adoption of piece-wise constant for regression model due to weak learns can be seen. For the benefit of the two models with traffic information correlation and the proposal of an update in new weight is made. The "W-Boost" shows more efficiency and the result after demonstrating the scheme on traffic network that shows both "self-similarity" and "non-linearity".

3. SYSTEM ARCHITECTURE AND DESIGN

The aim is to gather the information of traffic flow over time and accuracy by a tool that will be used for prediction after developing using a model from machine learning. Upon comparing algorithms under supervising method the accuracy is predicted with the best result after updating the classification model.

A. FEASIBILITY STUDY

WRANGLISH OF DATASET

Data is loaded and cleanliness is also seem to be checked. To analyze the dataset it is trimmed and then later on cleaned. The steps should be ensured and the decisions should be justified when it comes to cleaning of the data

B. COLLECTION OF DATA

The splitting of the given dataset into testing and training set is made for the prediction from the set of data. Usually, 70% of data is assigned to the training set and 30% is assigned to the testing set.

C. PREPROCESSING

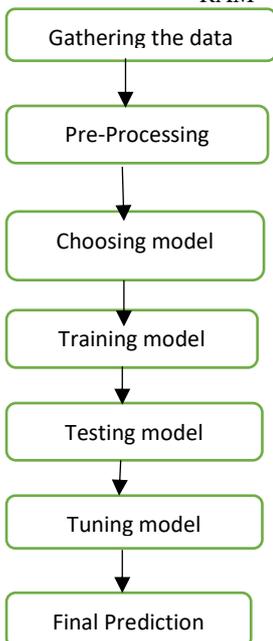
To avoid inconsistency from missing values from the collected data, data are preprocessed for efficiency improvement over the algorithm and get the best result. Removal of outliers and conversions of variable are also required to be completed.

D. PREDICTIVE MODEL CONSTRUCTION

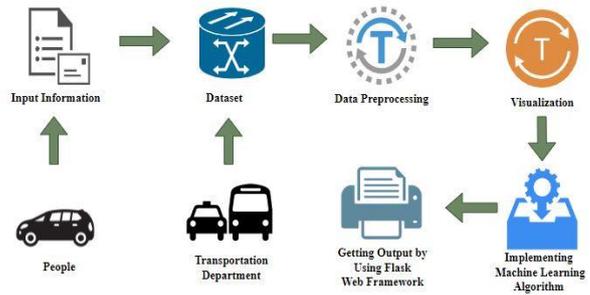
Historical and raw data are gathered first when it comes to the process of machine learning. Then the data should be pre-processed as the raw data cannot be used directly without pre processing as input values. Thirdly the suitable algorithm for the particular problem in machine learning is figured out and applied to the model made. After splitting the model into testing and training the output is predicted and compared with the actual output and errors are also minimized along the process. Tuned model is made and brought into application along with tuned time in order to make sure the accuracy is better and high. Thus, upon following all these steps the prediction is made by the machine.

E. ENVIRONMENTAL REQUIREMENTS

1. Software Requirements:
 - Operating System : Windows
 - Tool : Anaconda Navigator with Jupyter Notebook
2. Hardware requirements:
 - Processor : Pentium IV or Pentium III
 - Hard disk : minimum of 80 GB
 - RAM : minimum of 2 GB



F. SYSTEM ARCHITECTURE



4. METHODOLOGY

A. DATA PRE-PROCESSING

The error rate resulted from a machine learning model can be figured out using some techniques called validation techniques that are almost closer to the actual error rate occurred. Validation technique need not be used for a set of data that has records as huge as say the population represented. It is used in scenarios where the dataset is not as huge as the representation of population in real time and when there is a necessity to find the values that are identified to be missing, to identify the type of data and its description and duplicate values existing in the dataset. An evaluation is made that is completely declared to be unbiased from the data sample used for the fitting of the machine learning model. Along with the “hyper parameters” getting tuned in the model the evaluation is performed on the set of data that is split for training. The configuration of model is implemented with the skill set from the data set used for validation purpose. This results in the evaluation biasing. In order for the hyper parameters to be tuned in the machine learning model by a fine handling, engineers working for machine learning use the data and for evaluation to run frequently, the model uses validation set. This process also deals with some time consumption due to the tasks performed. These are collection and analysis of data, content and quality along with the structure of data getting approved and addressed altogether. The selection of the machine learning algorithm that should be incorporated into the model is based on the knowledge acquired from the understanding and learning of the characteristics of data upon identifying them. The cleaning process of data is the first step and some methods from the “pandas” library in python is used for dealing with the values missing in the set. Although cleaning is the primary task here, modelling with exploration required the most time to be given. Data can be found missing due to various reasonable actions like the source being a mistake or others and this can be learnt from viewing it statistically. To handle filling, identify and learn about 27 the approaches statistically to perform computation of the values that are found to be missing, the type of them plays a key role. Thus the source of the values missing needs to be learnt before adding them with the code. Data could be found missing due to various reasons like during the transmission of data from a database the data was found to be lost, an error occurred during coding, the user forgot to fill or denied to fill a field due to personal reasons and interpretations perceived from it.

MODULE DIAGRAM



B. EXPLORATION DATA ANALYSIS OF VISUALIZATION

Descriptions that can be counted as quantitative and data estimations are the majorly concentrated in statistics and subsequently these applied statistics use the visualization of data to perform the necessary tasks. Machine learning also use visualization of data after pre-processing usually. This is used to understand the dataset and acts as a tool for learning the same. Patterns, outliers, corrupted data and others can be figured out with the help of data visualization. Plotting and charts can be done to showcase the relationships just with some amount of understanding in the domain. The relationships are usually for stakeholders and thus visceral. Visualizing data helps in better understanding rather than being presented over a sheet of dataset with tabulation as samples. This is incorporated in machine learning for efficiency for work and to make it easy for dealing with. Representation can be done by using python in machine learning for this purpose. Histograms, bar charts, box plots for distributions, line plots for time series can be represented for data visualization of a dataset.

MODULE DIAGRAM



False positives(FP)- A payer is predicted as defaulter. Here the predicted class and actual class gives a yes and no respectively. Example: When actual class tells a person is dead whereas the predictive calls the person to be alive.

False negatives(FN)-A defaulter is predicted to be a payer. Here the predicted class and actual class gives a no and yes respectively. Example: When the actual class tells a person is found to be alive whereas the predictive class calls the person to be dead.

True positives(TP)- A person that paid is called as payer. Here the prediction is made right and the actual and predictive class both give the same positive outcome that is the predictive class and actual class gives a yes and yes respectively. Example: When the actual class signifies a person to be alive and the predictive class also approves and figures out the person to be alive.

True negatives(TN)- A person that did pay is said to be a defaulter. Here the prediction is made and the actual and predictive class both give the same negative outcome that is the predictive class and actual class giving a no and no respectively. Example: When the actual class shows the person is dead and the predictive class also identifies the person to be dead.

C. PREDICTION RESULT BY ACCURACY

To predict a value some predictors are used. These predictors are independent and used with equations that are linear in nature for an algorithm of “Logistic regression”. The range of prediction of values lies between a negative and positive infinity. This model gives the best accuracy as a result on comparing with other models. We calculate the “true positive rate” with the formula – $TPR=TP/(TP+FN)$. The false positive rate is calculated similarly with the formula- $FPR=FP/(FP+TN)$.

Accuracy- The correctness in the predictive model is calculated with proportion along with the frequency of correctness that are default or non-default in the model and this is known as accuracy.

Accuracy calculation: The formula for accuracy is $(TP+TN)/(TP+TN+FP+FN)$. The proportion of predictions from observations made correctly to the total number of observations in the model is the accuracy that is one of a performance measure. Just having a high accuracy does not define the model to be working at its best. Even though the accuracy is said to be a crucial measure for

performance, it gives the best result only upon using it over a set of data that is symmetric with the positives and negative values being the same in terms of numbers

Precision: The formula for precision is $TP/(TP+FP)$. The proportion of predictions made from the observations correctly to the total number of observations in the model gives the precision value. When the rate of precision is found high, it means the number of false positives are low in such cases. In this model we have acquired 0.788 which is comparatively said to be better.

Recall: The formula for recall is $TP/(FN+TP)$ and recall is the proportion of the correctly predicted values from the actual defaulters. That is the ratio between positive observations 30 predicted correctly to the overall observations in the actual class.

F1 Score - Precision and recall is taken to calculate the weighted average for F1 score which means the score accepts the false positive as well as false negative into consideration. This is said to be more useful than accuracy though it is difficult to perceive more importantly in a distribution with asymmetric classes. When the false positives and negatives have a similar count, accuracy is more efficiently calculated. On a contrary, where the count is not similar, precision and recall could be used.

General Formula:

$F\text{- Measure} = 2TP / (2TP + FP + FN)$

F1-Score Formula:

$F1\text{ Score} = 2*(Recall * Precision) / (Recall + Precision)$

D. ALGORITHM EXPLANATION

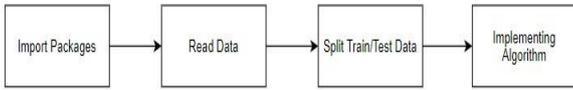
When a machine is fed with data to learn by giving training using a computer program and using the knowledge acquired it given output a machine learning model by classifying, it is known as classification in supervised machine learning. There are two kind of categories of data. The first one is bi-class where the identification or choice is made between two options only. One out of two options is the output always. Example for this can the gender of a person. On the contrary a multi class has so many outcome possibilities. An example for this can be speech recognition. Supervised machine learning labels the data with the existing and knowledge that is acquired in the run. After observing the pattern in data that are not labelled yet, the algorithm assigns the suitable label to them.

E. RANDOM FOREST CLASSIFIER

Random forest is nothing but a decision tree which is constructed at the time of implementation the random forest algorithm technique in the code. It follows the method of ensemble learning. Various tasks like regression, classification etc., can be performed efficiently with it. It is also known as “Random decision trees”. When there is a dataset given that is huge, over fitting and has high complexity on splitting the sets for training, random forest algorithm can be implemented in those cases. It is also comes under supervised learning in machine learning algorithms. In order to develop a model that is powerful for the purpose of prediction, “ensemble learning” can be applied to it as it combines algorithms of various types or algorithms of the same type over multiple times. “Multiple decision trees” are resulted from combining many decision trees, ultimately leading to a “forest of trees”. This is done by the random forest algorithm. Both the tasks under machine learning like regression and classification can be performed using this algorithm. There are some steps to follow when performing the “random forest algorithm”. The first one is to pick and collect some random number of records say N from the given set of data and build a decision tree for the records. Then we have to select the trees count required for the algorithm and follow the previous statement over and over again. Each tree tries to find an output value in the forest when it comes to a regression problem. The average value of all the trees are calculated and then the final value is decided upon the prediction. On the other hand, a

classification problem the prediction of class or category to which the newly acquired data belongs is searched and matched by the trees in the forests. The class or category with the highest number of pick is chosen in the record.

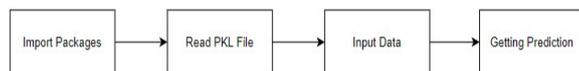
MODULE DIAGRAM



F. DEPLOYMENT

A web framework is something that does not have the need for any libraries or tools to perform a task and “flask” is one such web framework that is also known as micro framework. It is known to be written in python. It does not have the components where common features will be available in already existing libraries that can belong to some third parties like form validation, database abstraction layer and so on. There are some features in applications that can be added to the extensions supported by “Flask” like it was implemented by the same. Open authentication technologies and form validation and other tools related to common frameworks are used with the help of extensions. This was an idea that popped up from an April fool’s joke which was later thought about, considered and made into a software application according to “Ronacher”. A group of international Python programmers came up with this idea in the year of 2004. The “Pocoo” and “jinja” projects were developed after a system that was bulletin board was made by “Ronacher and Georg Brand”. A new project was formed in April, 2016 known as “Pallets project” that was given the Flask development authorities of libraries related. This was done as the “Pocoo” band was disbanded. Flask stood second after Django in the voting process of frameworks in github as the most widely used and popular framework with the approval of python developers by October 2020. It can be identified from “BSD” license. Compared to other frameworks in python, “flask” has gained more popularity and usage.

MODULE DIAGRAM



GIVEN INPUT EXPECTED OUTPUT

Input: data values

Output: predicting output

5. ANALYSIS

When data is collected from various sources, it is collected in raw format, which makes analysis impossible. To achieve better results from the applied model in the Machine Learning method, the data must be properly organized. Some Machine Learning models require information in a specific format; for example, the Random Forest algorithm does not accept null values. As a result, in order to run the random forest algorithm, null values must be managed from the original raw data set. Another consideration is that the data set be formatted in such a way that multiple Machine Learning and Deep Learning algorithms can be run on the same dataset. Python is being used in this machine learning process because it allows us to access various datasets.

Python packages used:

1. Sklearn- Scikit learn is package that is used for machine learning purpose as it contains many algorithms for the same in

python. Modules such as “Logistic Regression, accuracy_score, DecisionTreeClassifier and train_test_split” are utilized in the modules done.

2. NumPy- It is a python module that deals with numerical calculations and mathematical functions. Reading data and manipulating them in arrays can be done using Numpy.

3. Pandas- Files can be written and read using this module and with data frames we can manipulate data effectively.

4. Matplotlib- With the help of data frames we can manipulate data effectively. Visualization of data can be done to find out the patterns existing from the set of data we are using.

A. RANDOM FOREST ALGORITHM

Random forest is nothing but a decision tree which is constructed at the time of implemented the random forest algorithm technique in the code. It follows the method of ensemble learning. Various tasks like regression, classification etc., can be performed efficiently with it. It is also known as “Random decision trees”. When there is a dataset given that is huge, over fitting and has high complexity on splitting the sets for training, random forest algorithm can be implemented in those cases. It is also comes under supervised learning in machine learning algorithms. In order to develop a model that is powerful for the purpose of prediction, “ensemble learning” can be applied to it as it combines algorithms of various types or algorithms of the same type over multiple times. “Multiple decision trees” are resulted from combining many decision trees, ultimately leading to a “forest of trees”. This is done by the random forest algorithm. Both the tasks under machine learning like regression and classification can be performed using this algorithm. There are some steps to follow when performing the “random forest algorithm”. The first one is to pick and collect some random number of records say N from the given set of data and build a decision tree for the records. Then we have to select the trees count required for the algorithm and follow the previous statement over and over again. Each tree tries to find an output value in the forest when it comes to a regression problem. The average value of all the trees are calculated and then the final value is decided upon the prediction. On the other hand, a classification problem the prediction of class or category to which the newly acquired data belongs is searched and matched by the trees in the forests. The class or category with the highest number of pick is chosen in the record.

6. DISCUSSIONS AND CONCLUSION

Data cleaning and processing came first. It was then followed by analyzing missing values, exploratory analysis and construction and evaluation of model. On the public test set, the best accuracy will be identified by an accuracy score that is found to be higher. This application can assist in determining the traffic state prediction. It provides better results in classification problem. It is strong in preprocessing outliers, irrelevant variables, and a mix of continuous, categorical and discrete variables. It produces out of bag estimate error which has proven to be unbiased in many tests and it is relatively easy to tune with.

FUTURE RESEARCH DIRECTIONS

We can compare more than one machine learning algorithms. We can combine this process to IOT based system. We can deploy this model from any cloud based system.

7. REFERENCES

- [1] Pucher, J., Korattyswaropam, N., Mittal, N., & Ittyerah, N. (2005). Urban transport crisis in India. *Transport Policy*, 12, 185-198
- [2] Zhang, G., Patuwo, B.E., & Hu, M.Y. (1998). Forecasting with artificial neural networks: the state of art. *International Journal of Forecasting*, 14, 35-62.
- [3] Yasdi, R. (1999). Prediction of road traffic using a neural network approach. *Neural Computing and Applications*, 8, 135-142
- [4] Smith, B.L., & Demetsky, M.J. (1994). Short-term Traffic Flow Prediction: Neural Network Approach. *Transportation Research Board*, 1453, 98-104.
- [5] Pamula, T. (2011). Road traffic parameters prediction in urban traffic management systems using neural networks. *Transport Problems*, 6, 23-28.
- [6] K. Cho, B. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, "Learning Phrase Representations using RNN
- [7] Encoder-Decoder for Statistical Machine Translation", in *EMNLP 2014*, Doha, Qatar, 2014
- [8] M. Li et al., "Scaling Distributed Machine Learning with the Parameter Server," *Proc. USENIX OSDI*, 2014, pp. 583-98
- [9] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-Term Traffic Forecasting: Where We Are and Where We're Going," *Transportation Research Part C: Emerging Technologies*, vol. 43, part 1, 2014, pp. 3-19.
- [10] Schapire, R.E.: *The Boosting Approach to Machine Learning: an Overview*. MSRI Workshop on Nonlinear Estimation and Classification (2002)
- [11] Khotanzad, P., Sadek, N.: *Multi-Scale High-Speed Network Traffic Prediction Using Combination of Neural Network*. *IJCNN* (2003) 1071-1075
- [12] Hall, J., Mars, P.: *The Limitations of Artificial Neural Networks for Traffic Prediction*. *IEE Proceedings on Communications* (2000) 114-118
- [13] Vlahogianni, E.I., Golias, J.C., & Karlaftis, M.G. (2004). Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, 24, 533-557.
- [14] Kranti Kumara, M. Paridab, V.K. Katiyar (2013). Short term traffic flow prediction for a non-urban highway using Artificial Neural Network
- [15] Ji Yoon Kim (2021). *A Traffic Prediction Using Machine Learning: Literature Survey*
- [16] Weiqing Zhuang and Yongbo Cao (2022). *Short-Term Traffic Flow Prediction Based on CNN-BILSTM with Multicomponent Information*
- [17] B.Karthika, N.UmaMaheswari, R.Venkatesh (2019). *A Research of Traffic Prediction using Deep Learning Techniques*
- [18] Hanghang Tong, Chongrong Li, Jingrui He, and Yang Chen (2005). *Internet Traffic Prediction by W-Boost: Classification and Regression*.