IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **Volume 11, Iss 4 , 2022**

VEHICLE COUNT PREDICTION APPROACH BY USING MACHINE LEARNING METHODOLOGY

Rama Devi Burri[1], V. B. V. N. Prasad[2]

[1]Department of Information Technology, Institute of Aeronautical Engineering, Dundigal-Telangana-500043.
[2]Department of Engineering Mathematics, Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, Green fields, Guntur, Andhra Pradesh, India -522302
E-mail: vbvnprasad@kluniversity.in
E-mail: ramaburri5@gmail.com.

DOI : 10.48047/IJFANS/11/ISS4/120

1. ABSTRACT

Rapidly growing cities with rising population mobility have results an exponential growth in the number of vehicles on the road. Historical data is used to count the vehicles for vehicle count prediction, previous information is gathered from a variety of sources, including daily newspapers, Internet sources, and other repositories. Artificial intelligence techniques are best for prediction giving accurate results. One of the subsets is Machine learning for artificial intelligence. This includes the random forest regression, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression, and Naive Bayes. These predictions are compared and assessed in order to find the best-performing optimized solutions.The aim of this paper is to utilize different Machine Learning Algorithms to predict the number of vehicles that will pass through a particular junction. By counting the number of vehicles which are moving in particular area, one can able to get information and may reduce the traffic.

Key words: Optimum solution, random forest regression, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Logistic Regression, Naive Bayes, Artificial intelligence.

2. INTRODUCTION

The number of vehicles in a city is an important parameter for urban development planning. And also use full resource for evaluation of economics as well as a useful source of data, post-evaluation of urban building projects. It can contribute the traffic control section of scientific management in ensuring transportation safety [6] by providing an accurate forecast of vehicle count. Short-term vehicle count forecasts are generally based on monthly and even daily, the traffic flow changes to evaluate variations in the movement of vehicles. Vehicle calculation is a non-stationary time series [2] with seasonal and periodic trends [1] and certain expansions. Due to these characteristics, six common time series prediction models [7] (Decision Tree, Random Forest, K-Nearest Neighbor, Logistic Regression, and Support Vector, Naive Bayes) have been developed. Traditional predictive methods necessitate large volumes of data collection, as well as the ability to synthesize the different elements that influence vehicle volume [3]. It's difficult to model this type of forecasting strategy. In other circumstances, the ideal can be effectively built, but the parameters are difficult to estimate because they lack the necessary data [4]. Because urban vehicles have two distinguishing characteristics: First, the data is seasonal and has a predictable pattern; instant, the sequences formed by counting number of urban vehicles [5] which is measured randomly. To

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 4 , 2022

approximate the series, one can use the Random Forest, Decision Tree, K-Nearest Neighbor, Support Vector Machine, Naive Bayes, and Logistic Regression models.

2.1 Vehicle Count Prediction:

Vehicle counting prediction, both supervised and unsupervised learning algorithms are used, which are subcategories of the machine learning. To properly forecast the outcome in "Vehicle Count Prediction,Decision Tree, Random Forest, and Support Vector Machine are the most commonly used machine learning algorithms. In this unsupervised learning algorithms are used to find patterns in data sets with data points that are neither classified nor labelled,and on the anther side one can use supervised learning algorithms. Supervised learning algorithms include regression, which produces a real value as an output, and classification, which is used to make yes or no decisions.

3. PROBLEM STATEMENT

Create classification methods to identify the count of vehicles passing through the junction by Using the attributes provided in the dataset, to address the shortcomings of the current system, one can present vehicle count prediction system based on a machine learning algorithms predict the number of vehicles count by using the dataset and some vehicle characteristics. Different methodologies and procedures are used to make the predictions.. These algorithms are more accurate than the existing algorithms.

4. EXISTING SYSTEM

In the existing system the vehicle count prediction is done by using algorithms, Naive Bayes and Support vector machine. These algorithms give less efficiency when compared to other algorithms which are developed in this paper.

5. PROPOSED SYSTEM

In order to solve the issues, which are present in the existing system and for improving the accuracy of vehicle count, the prediction [9] is done by using various techniques and different algorithms such as “Decision Tree, Random Forest, logistic Regression and K-Nearest Neighbor (KNN)”. These algorithms give more accuracy than the algorithms which are in the existing system.

6. VARIOUS ALGORITHMS

6.1. Decision Tree:

One of the supervised learning technique is the Decision tree, and it can be used to solve classification and regression problems. Mainly used to resolve classification problems. In this supervised learning technique, there are some internal nodes which consists of dataset attributes, branches which represent the decision rules and the leaf nodes represents the conclusion of the tree. Here two nodes in Decision Tree algorithm, they are Decision node and the Child node. Decision node represents the decisions of the tree and leaf node represents the output of the decisions and also the output of the tree. The decisions in this algorithm are taken based on the datasets and also based on problem decision. It was named as decision tree because it will represent tree, it starts with root node and followed by branches and at the last level there will be leaf node. Finally, trees are represented using the CART algorithm (Classification and Regression Tree algorithm).

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved,** Volume 11, Iss 4 , 2022

6.2Random Forest:

Random forest is one of the supervised learning technique, it is also a standard machine-learning algorithm and mainly to answer both classification, regression problems. Random Forest algorithm can be obtained by combining number of decision trees on different subsets of dataset. In this algorithm the results are be around to get more accuracy of the algorithm. Random Forest Regression is used to solve both classification and regression using multiple decision trees this process is called as bagging technique.

6.3 K-Nearest Neighbor (KNN):

One of the supervised learning techniques is K-Nearest Neighbor it is also a popular machine-learning algorithm and mainly this technique is used to solve both classification and regression problems[8]. This algorithm mainly to solve classification problems. This algorithm stores all the data and classify them into well suite category by using this algorithm. This algorithm will not take the data based on assumptions. It is also called as lazy learner algorithm because it stores the dataset without learning the data from the training data set. At training phase whenever the new data enters then it stores the dataset and then data is classified into a category same as new data.

6.4: Logistic Regression:

It is used to solve both classification and regression problems. This algorithm mainly used to solve classification problems. In this algorithm the output can be either Yes or no type,0 or 1, true or false type. This algorithm is used to predict dependent variables from the autonomous variables. The dependent variables output can be predicted by using this algorithm. This algorithm uses the concept of probability and the cost lies between 0 and 1.

6.5 Naive Bayes:

Naive Bayes is also one of the technique for supervised learning and is used for solving classification problems. This algorithm works based on Bayes theorem, it is used to make quick predictions and this algorithm mainly classified into two words they are Naive, Bayes. Specifies that occurrence of one feature is not depending on other features. Bayes will mainly deals with Bayes rule or Bayes law. This is one of the fast machine learning prediction model and provides better accuracy and also deals with some issues related to the Bayes rule.

Steps for the Algorithms:

1. Defining Problem
2. Collection of Data
3. Exploratory Data Analysis (EDA)
4. Feature Engineering
5. Train/Test Split
6. Evaluation of the model
7. Model Selection, Training
8. Prediction and Assessment
9. Model Improvement

7. METHODOLOGY

Step 1: Importing the panda's module for loading the data frame is the first step.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 4 , 2022

Step 2: Create functions that extract the month, day, and hours from a Timestamp (Date Time) and load them into various columns.

Step 3: Separate the class label from the rest of the data and save it in the target variable.

Step 4: Use Machine Learning methods to train the data and anticipate the results after testing.

Step 5: We present everything in a graph manner, with the number of vehicles passing by the signal on a daily, weekly, monthly, and annual average basis.

Step 6: We used several machine learning methods to determine the accuracy.

8. RESULTS

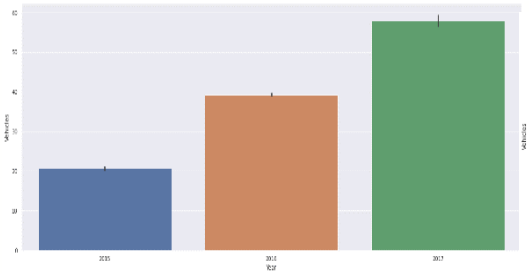

Figure 8.1: count of vehicles in a year

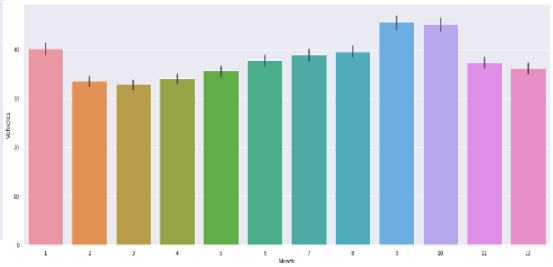

Figure 8.2: Count of vehicles in a month

From the **Figure 8.1**during the year 2015 the vehicle count is 20 and in the year 2016 the vehicle count has increased to nearly 40 and during the upcoming years the count of vehicles gradually increased to nearly 60.
As a result one can able to get knowledge, for every year due to increasing population the vehicle count also increasing, for this above **fig: 8.1** resenting the increasing number of vehicles.
The **figure 8.2** represents the vehicle in a month considering different day

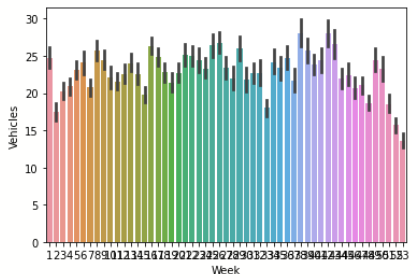

Figure 8.3: Count of vehicles in a week

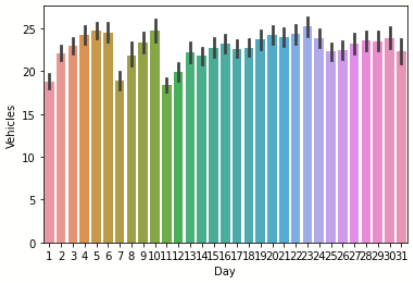

Figure 8.4: Count of vehicles in a day

From Figure 8.3: There are many variations of vehicle count during particular week. By Calculating the count one can easily recognize the traffic, in which time the traffic is less to travel.

The **Figure 8.4:** represents the count of vehicles in a day staring to ending, by having this one can set the as per their work. People can avoid traffic by noticing this.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 4 , 2022

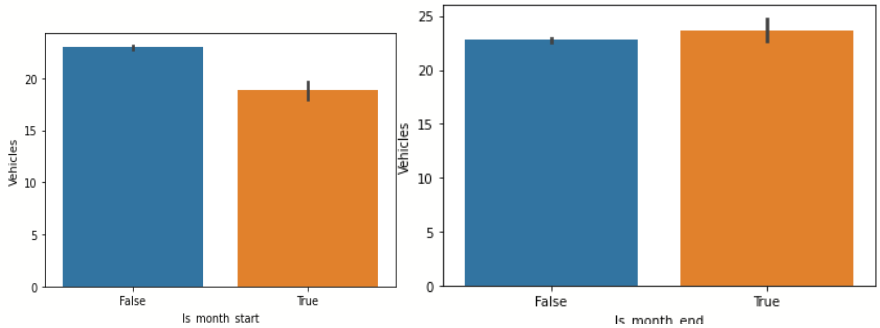

Figure 8.5: Count of vehicles in month start Figure: 8.6: Count of vehicles in month end

The **Figure 8.5** shows that during the month starting the count of vehicles is nearly 20 and if it is not month starting the count of vehicles is greater than 20.

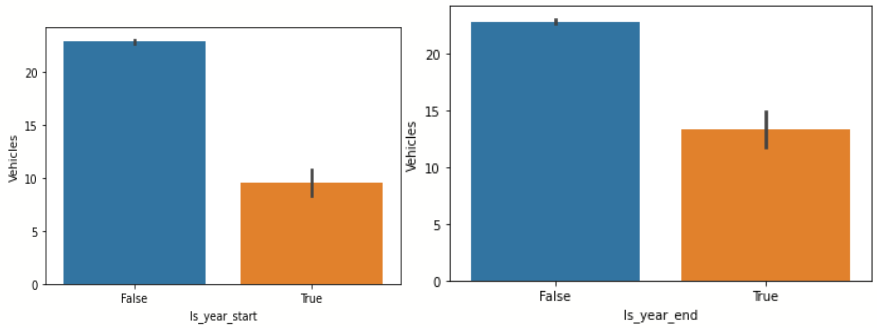

Figure8.7: Count of vehicles in Year start Figure8.8: Count of vehicles in year-end

Here we calculated the number of vehicles moving in year starting and ending

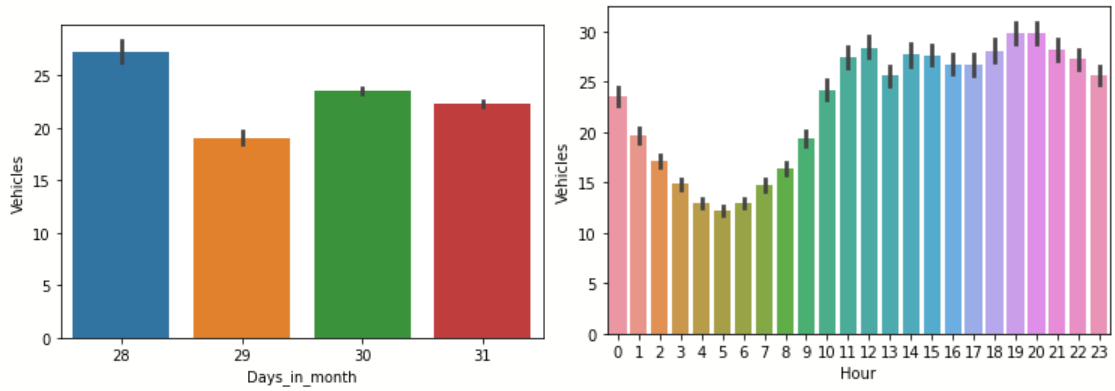

Figure8.9: Count of vehicles in day_in_monthFigure8.10: Count of vehicles in hour

In Figure 8.9 the count of vehicles in a particular day in month. Like public holidays, festivals the count is more when compared with the remaining days,

The figure 8.10 shows the count of vehicles in particular hour, here we counted the frequency of vehicles in particular time, like office, school going timings. And at starting of hour the vehicle count is low gradually the vehicle count increasing.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 4 , 2022

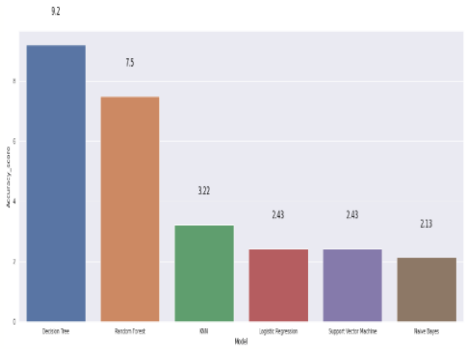

Figure8.11: Model and there accuracy score

This is the final Graph representing the accuracy of the score which is different for various algorithms

Name of the Algorithm	Decision Tree	Random Forest	KNN	Logistic Regression	Naïve Bayes
Accuracy	9.2	7.5	3.22	2.43	2.13

As per the above results Decision tree is Best score among all.

9. CONCLUSIONS

After observing different cases, like one particular day, in a week , month, and year the frequency of the vehicles can be calculated. From the above table Decision tree algorithm is best for vehicle count prediction.

for predicting the vehicle count in a particular junction. the traffic can be intimate to the people, and we can set them alternate for their routes. Government can take certain measures to reduce the traffic by building highways or creating alternating ways to reach the destination.

10. REFERENCES

1. HongMeng Ji. Analysis on Passenger Transport Requirment and Passenger Volume Forcast of High Speed Railway[D].BeijingJiaotong Universit,2015.
2. Diongue A K, Diop A, Ndongo M. Seasonal fractional ARIMA with stable innovations[J]Statistics & Probability Letters, 2008, 78(12):1404-1411.
3. Wang Y, Wang J, Zhao G, et al. Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China[J]. Energy Policy, 2012, 48(3):284-294.
4. ZhenLong Wang. Time series analysis[M]. BeiJing: China Statistics Press, 2000.
5. Kalekar P S. Time series Forecasting using Holt-Winters Exponential Smoothing[J]. Kanwal Rekhi School of Information Technology, 2004.
6. ChingFu Chen, YuHern Chang, YuWei Chang. Seasonal ARIMA forecasting of inbound air travel arrivals to Taiwan[J]. Transportmetrica, 2009, 5(2):125-140.
7. Montanari A, Rosso R, Taqqu M S. A seasonal fractional ARIMA Model applied to the Nile River monthly flows at Aswan[J]. Water Resources Research, 2000, 36(5):1249-1255.
8. Dr. B. Rama Devi, Sai Manvitha Enadula, Akshith Sriram Enadula "Recognition of Student Emotions in an Online Education System" 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT) 10.1109/ICECCT52121.2021.9616788.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 4 , 2022

9. Dr. B. Rama Devi, Ch. S. Sudha Sri ,B. Akhila Md. M. Asma"Prediction Of Missing Child Using Machine Learning" Vidyabharati International Interdisciplinary Research Journal), PP:388-394 ISSN: 2319-4979,
10. S. P. Mogal and D. I. Lalwani, "Experimental Investigations of Single and Multi-Fault in Rotor System," Applied Mechanics and Materials, vol. 541–542, pp. 628–634, Mar. 2014, doi: 10.4028/www.scientific.net/amm.541-542.628.
11. Haider, A., & Anwar, F. (2019). Design and analysis of low-power OTA for biosignal amplification. Journal of Medical Systems, 43(12), 355.
12. Zhao, S., & Li, Z. (2018). A low-power bulk-driven OTA with dynamic bias control for wearable healthcare devices. In Proceedings of the IEEE International Symposium on Low Power Electronics and Design (ISLPED) (pp. 225-230).