International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

Machine Learning Based Traffic Volume Count Prediction

Nishant Patil[1], Megha Natekar[2], Ratan Gore[3], Chandrashekhar Raut [4]

[1,2,3]Student, Computer Engineering Department, Mumbai University, Datta Meghe College of Engineering, Navi Mumbai, Maharashtra, India
[4]Professor, Computer Engineering Department, Mumbai University, Datta Meghe College of Engineering, Navi Mumbai, Maharashtra, India

---

Abstract - The transportation industry was accountable for 28% of worldwide CO2 emissions in 2014. the number of traffic-related deaths in 2013 was 1.25 million. Additionally, holdup at peak hours reaches unacceptable levels in many parts of the earth these are all serious issues caused by current transportation systems, and optimization through the usage of recent technologies is vital for the required improvements. Tons of the innovation that are a part of the solution already exists. Various Business sectors and government agencies and individual travelers require precise and appropriate traffic flow information. It helps the riders and drivers to make better travel judgments to alleviate traffic congestion, improve traffic operation efficiency, and reduce carbon emissions. Machine learning provides better accuracy for Traffic volume flow prediction. It's addressed as a major element for the success of advanced traffic volume management systems, advanced public transportation systems, and traveler information systems. The rationale of this extension is to develop a prescient demonstration utilizing different machine learning calculations and to record the end-to-end steps. The Metro Interstate Activity Volume dataset could also be a relapse circumstance where we are trying to anticipate the esteem of a ceaseless variable. We'll be analyzing how the drift of month-to-month interstate activity volume changes over an extended time between 2012 and 2018. Concurring to the discoveries, the month-to-month activity volume remains an equivalent indeed even though the knowledge appears a somewhat upward drift a short time recently applying time arrangement strategies.

Key Words: **Traffic Volume, Random Forest, Machine Learning, Webapp, prediction, RSME, MAE**

1. INTRODUCTION

Traffic jams on urban Network are increasing day by day, because the traffic demand increases, and the speed of the vehicles is drastically reduced thus causing longer vehicular queuing and more such cases substantially hamper the traffic flow by giving rise to holdup.

Such situations highlight towards the drawback such as

- Increase in pollution
- Wear and tear of vehicles
- Delays may result in late arrival etc.

1.1 Motivation

With the progress of urbanization and therefore the recognition of automobiles, transportation problems are becoming more and more challenging: the traffic volume flow is congested, wear n tear of vehicles, delays end in the late time of arrival at the meeting, accidents are frequent, and wastage of fuel while waiting in traffic, the traffic environment is becoming worse, to unravel this problem and to assist society, we've chosen our topic as traffic volume prediction.

1.2 Problem Definition

Now? The question arises of how to improve the capacitor y of the road network. To solve this problem the first solution that occurs to most of us is to build more highways, expanding the number of lanes on the road. However, according to the study done by scholars, expanding the road capacity will cause more serious traffic conditions. Therefore, traffic volume prediction is one of the most famous.

1.3 Objective

The objective of this study is to seek out a traffic volume predictor suitable for real implications. This predictor must be accurate in terms of computation cost and power consumption. Within the go after such a predictor, we've included the subsequent contributions: We compare existing schemes to seek out their effectiveness for real-time applications

2. Literature Review

Traffic volume prediction is integrated by a selection of technologies. Machine learning is one of the foremost famous of those systems. It can improve traffic efficiency, ease congestion, increase road capacity, and reduce traffic accidents and environmental pollution. Road sources are mainly gathered from the high mobility vehicles on the highway or on urban roads, which makes it so important to figure out what percentage of vehicles are progressing to be on the given road segment within the longer term. To affect this, the traffic volume prediction system will provide highly reliable future traffic. according to the historical traffic pattern and thus the position over the entire road network.

which will cause simulating the long run workload and possible computing capacity. Most of the traffic data reports are actual time, but sometimes it is not so favourable because we use this report once we plan which route, we must always go. Assume that we are getting to the office during working hours and that we at traffic information and choose the simplest or shorter route to reach our destination but holdup occurred, the problem is to urge actual-time information about traffic comes whence to resolve this issue by using forecasting? It's aiming to be great, but what causes can impact traffic conditions? We'd prefer to research it. Many causes can affect traffic conditions. This and ancient traffic conditions are often considered predicting, these suggestions are very simple, if traffic is so heavy immediately, also acceptable is that after ten or twenty minutes the traffic situation would be same ancient traffic situation, Different weekdays and weekends may behave in several traffic situations, and perhaps they will also alter traffic conditions. With the increasing cost of gasoline, the demand for an efficient routing system to scale back traffic jams is extremely necessary.

3. Proposed Methodology

how to make this existing system more efficient and enforce traffic environment for efficient and accurate transportation, which may help us better arrange transportation resources, disperse the traffic flow before it's overloaded, and even provide more abundant on-road entertainment.

Where one such need arises towards the prediction of traffic volume count.

Importance of traffic volume:

- Better implies for advancement of infrastructures.
- Provides way better implies to utilize streets
- Accurate activity volume forecast can help course arranging, and relieve activity congestion.
- All of these planning will also help the government and rest of bodies

System Flow:

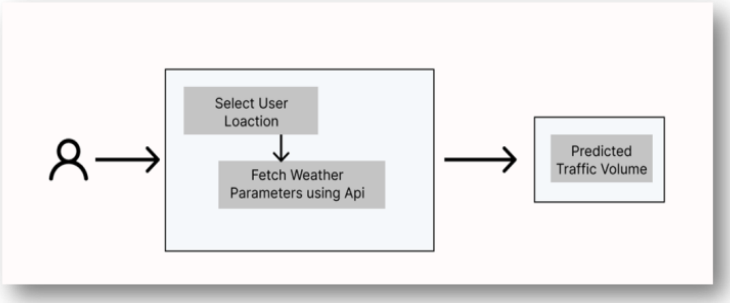


Fig 3.1 System flow

4. Dataset and technology used

Our dataset contains 10 column and 48205 rows. First 5 records shown below

	holiday	temp	rain_1h	snow_1h	clouds_all	weather_main	weather_description	date_time	location	traffic_volume	Location
0	None	288.28	0.0	0	40	Clouds	scattered clouds	02-10-2012 09:00	Rabale	5545.0	Rabale
1	None	289.36	0.0	0	75	Clouds	broken clouds	02-10-2012 10:00	Gansoli	4516.0	Rabale
2	None	289.58	0.0	0	90	Clouds	overcast clouds	02-10-2012 11:00	Thane	4767.0	Rabale
3	None	290.13	0.0	0	90	Clouds	overcast clouds	02-10-2012 12:00	Koparkhairane	5026.0	Rabale
4	None	291.14	0.0	0	75	Clouds	broken clouds	02-10-2012 13:00	nerul	4918.0	Rabale

Machine learning: could also be a technique of knowledge analysis that automates analytical model building. it is a branch of AI-supported the thought that systems can learn from data, identify patterns and make decisions with minimal human intervention.

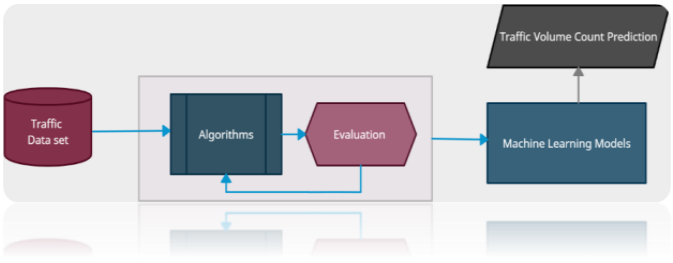
Python: is an interpreted high-level general-purpose programming language. Its design philosophy emphasizes code readability with its use of great indentation. Its language constructs also as its object-oriented approach aim to assist programmers to write clear, logical code for small and large-scale projects.

NumPy: NumPy is one of the foremost commonly used packages for scientific computing in Python. It provides a multidimensional array object, also as variations like masks and matrices, which can be used for various math operations.

Pandas: pandas are a fast, powerful, flexible and easy to use open-source data analysis and manipulation tool, built on top of the Python programming language.

Lasso regression: Lasso regression could also be a kind of linear regression that uses shrinkage. Shrinkage is where data values are shrunk towards a central point, a bit like the mean. The lasso procedure encourages simple, sparse models.

Ridge regression: Ridge regression could also be a model tuning method that's used to analyses any data that suffer from multicollinearity. This method performs L2 regularization. When the issue of multicollinearity occurs, least-squares are unbiased, and variances are large, this leads to predicted values being distant from the particular values.



Random-forest: A random forest could also be a machine learning technique that's used to solve regression and

classification problems. It utilizes ensemble learning, which can be a method that mixes many classifiers to provide solutions to complex problems. A random forest algorithm consists of numerous decision trees.

5. Result

According to our work we get following RSME values and MAE values

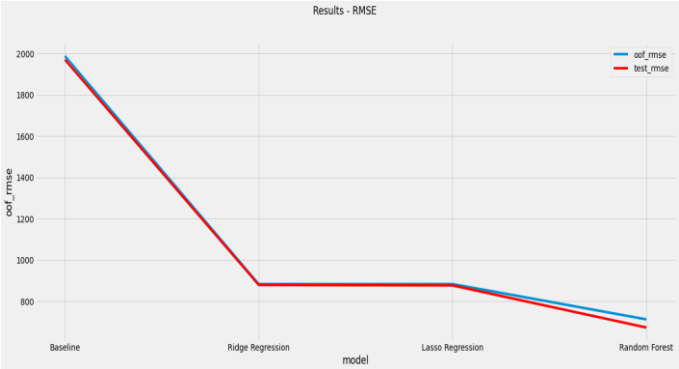


Fig 5.1 RSME values of All algorithm

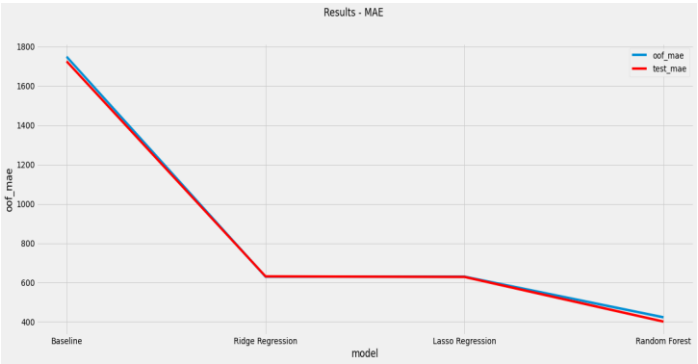


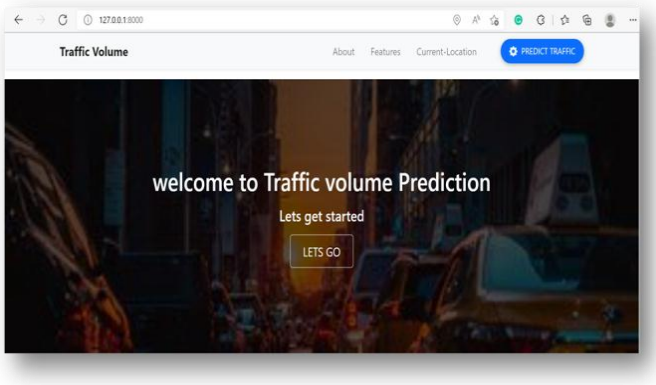
Fig 5.2 MAE values of All algorithm

Table -1: comparison of all algorithm with their RSME and MAE value.

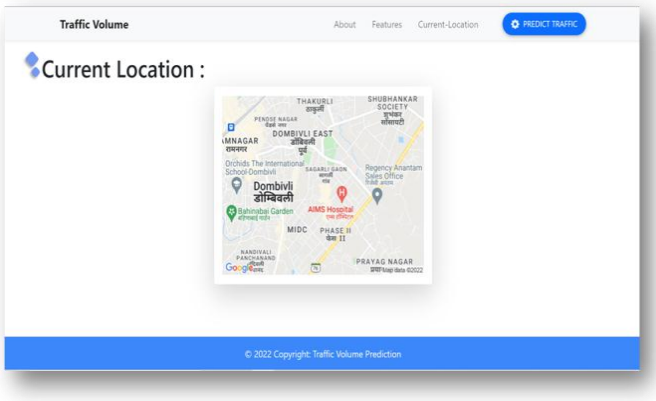
ALGORITHM	RMSE	MAE
Ridge Regression	878.76	631.14
Lasso Regression	876.49	628.38
Random Forest	672.31	401.08

Random forest is giving the lowest RSME & MAE score as compared to others, so going with this model would be a good idea.

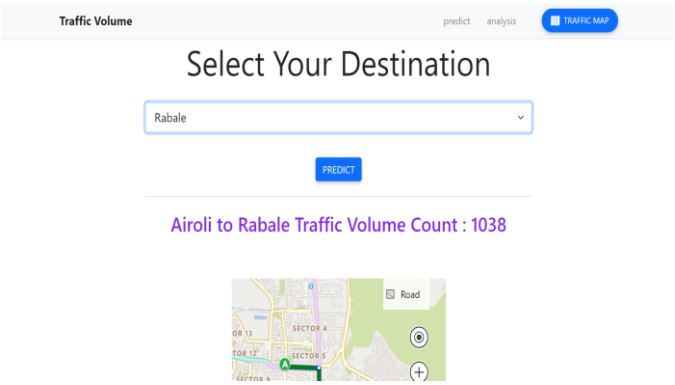
6. Implementation



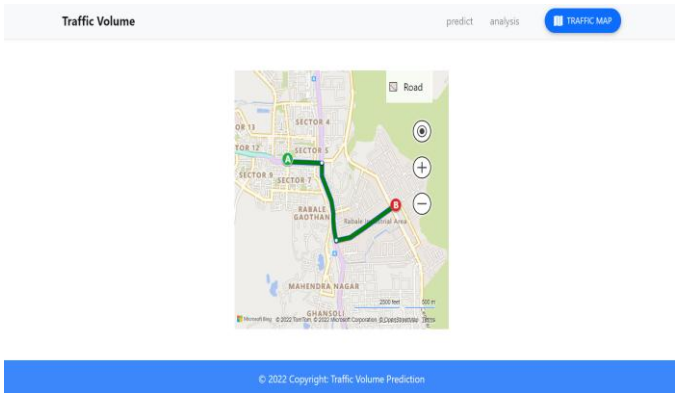
The landing page of the traffic volume provides an interface for the user to access the website and to predict the traffic volume count along with the entire route from source to destination. This page includes the navigation links i.e. About, Features and Current-location and predict traffic and also includes let's Go button to navigate to the predict page.



Current location represents the starting location or source location.



Here the user can enter the destination, selecting from the dropdown the required destination and then using the predict button to predict the traffic volume. Thus, the traffic volume will be predicted giving us the desired result.



The predicted result shows the traffic volume count and thus the complete route from source to destination on the live map for user to get a lively experience.

6. CONCLUSIONS

Random forest is giving the least RSME & MAE score as compared to others, so going with this demonstrate would be a great thought. us extend Machine Learning based traffic volume prediction model frameworks bargains with data innovation, machine learning. In our venture, we focused on the ML models utilized in activity expectation errands. In spite of the fact that profound learning and hereditary calculation is an vital issue in information examination, it has not been managed with broadly by the ML community. The proposed calculation gives higher precision than the existing calculations too, it moves forward the complexity issues all through the dataset. The framework can be examined and a parcel work can be done. The calculations will be advance moved forward to much higher exactness. Ready to moreover coordinated the net server and the application

REFERENCES

[1] 'Vehicular traffic analysis and prediction using machine learning algorithms' by A Moses, R Parvathi - 2020 International Conference

[2] https://www.kaggle.com/code/meemr5/traffic-volume-prediction-time-series-starter

[3] https://www.researchgate.net/publication/351295007_A_Traffic_Prediction_for_Intelligent_Transportation_System_using_Machine_Learning