



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 4, April 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

# Application of Machine learning in determination of LoS (Level of Service) of a Roadway from V/C (Volume to Capacity) Ratio and Speed Using K- Means

Sreeparvathy CM

Assistant Professor, Department of Civil Engineering, Mahatma Gandhi Institute of Technology, Hyderabad, India

**ABSTRACT:** A qualitative measure used to relate the quality of motor vehicle traffic service can be defined as Level of Service introduced by Human Capital Management (HCM). India, being a country with huge growth in private transportation in the recent times and does not have sufficient infrastructure to meet the demand resulted in congested roads and delay in traffic operations, thereby reducing the quality of service by these vehicles. Six LoS letters are defined by HCM, viz., A, B, C, D, E, and F, denoting A as the best quality of service and F as the worst helps in dealing with analysis of traffic data. The current availability of computational power and massive progress in the theory of artificial intelligence, machine learning models would be a suitable tool to overcome the mammoth tasks. In this paper, the application of K-Means Algorithm (Python) that is trained with a set of synthetic data set of more than five thousand data points for prediction of LoS using the input parameters: Volume and Speed of the vehicle are explored. Based on the learning ability of this algorithm, the data can be bucketed into six categories suitable for the prediction of LoS. For computation of LoS, volume to capacity ratio is to be estimated. K-means algorithm is an iterative algorithm that tries to partition the dataset into  $K$  pre-defined distinct non-overlapping clusters where each data point belongs to only one group. Based on input parameters the K-means algorithm helps to find to which cluster the new data point belongs and hence the output can be predicted based on the cluster. The program gives the category of LoS based on the input speed and range of volume on a roadway.

**KEYWORDS:** LoS, Human Capital Management, K- means, Volume, Speed, Python, Machine Learning

## I. INTRODUCTION

Quality of service is more important in today's networks. LOS is a commonly used qualitative measure in transportation engineering applicable to all services in management domain and analyzes roadways and intersections by categorizing traffic flow and assigning quality levels of traffic based on vehicle speed, density, congestion, etc.

In this paper, presentations of multivariate clustering methods in identifying the Level of service based on the defining parameters volume to capacity ratio and speed of flow is studied. As per the HCM, there are various methods for calculating the LoS for dissimilar types of facilities. Speed and V/C Ratio are dependable components in prediction of LoS. Introducing machine learning or clustering techniques, the investigatory analysis potential of the data is extracted accurately.

Examining data analysis involves prying data for unknown structures and examining for trends as well as incongruities. This may involve using prior knowledge or inflexible models in the exploration process. This analysis would help to reveals the structures in data. K-means can be an unsupervised learning technique, it is simple and has the capability of exploring the data and categorizing it.

The factors of LoS can be described as estimating the traffic flow characteristics along a road stretch which could be effectively collected using detectors. Such huge data could be explored and utilized for training the program. For a given roadway, the capacity is constant which is a quantitative measure. Traffic volume at any instant of time is the major factor that contributes to the quality experienced on the roadway.

### Need of Machine learning in Transportation Engineering

Increased globalization in the economy, population growth, development of industrial production, safety concerns, and environmental degradation are becoming a challenges in transportation. With the current availability of computational power and massive progress in the theory of artificial intelligence, machine learning models seem to be a suitable tool to overcome the challenges. Machine learning, as a subset of the tools of artificial intelligence, is a broad area of

computer science that provides machines and autonomous systems with the ability to learn and improve on previous experience. In other words, machine learning focuses on the development of computer programs that can autonomously access data and use it for learning and self-improvement. This core objective of machine learning can be achieved through many types of learning, including supervised learning, unsupervised learning, reinforcement learning, self-learning, feature learning, and association rules, among others.

Artificial neural networks, support vector machines, decision trees, evolutionary algorithms, Petri nets, and regression algorithms are the best models. It seems that machine learning is related and overlaps with other fields of information theory, like optimization, statistics, and data mining.

## II. LITERATURE REVIEW

[1] In this paper, performances of multivariate clustering methods in specifying flow pattern variations reconstructed by a macroscopic flow model are sought. The classification of flow states over the fundamental diagram are sought utilizing the methods of partitional cluster analyses by considering the stretch density. Non-hierarchical or partitional clustering analysis returned promising results on state classification which in turn helps to capture sudden changes on test stretch flow states. The measure of determination coefficient calculated by using the K-means method.

[2] Is application of driving condition recognition to the intelligent control of hybrid electric vehicles using K-means. Driving features are identified and used for driving segment clustering, using the k-means clustering algorithm.

[3] In this paper, Modified K-means Clustering (MKC) approach is used, a set of historical data is portioned into a group of meaningful sub-classes (also known as clusters) based on travel time, frequency of travel time and velocity for a specific road segment and time group. Clustering strategy can be used as a powerful tool of discovering hidden knowledge that can easily be applied on historical traffic data to predict accurate travel time.

### Methodology:

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. Clustering is one of the most common exploratory data analysis technique used to get perception about the structure of the data. It can be defined as the task of identifying clusters in the data such that data points in the same cluster are very similar while data points in different clusters are different. The decision of which similarity measure to use is application-specific.

Clustering is an unsupervised learning method since the ground truth to compare the output of the clustering algorithm to the true labels to evaluate its performance is absent. The only aim in this paper is to investigate the structure of the data by grouping the data points into distinct subgroups.

K-means algorithm is an iterative algorithm that tries to partition the dataset into K pre-defined distinct non-overlapping clusters where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum. The less variation we have within clusters, the more homogeneous the data points are within the same cluster. In this project K-Means is used to analyze traffic clusters of LoS based on two parameters V/C Ratio and Speed.

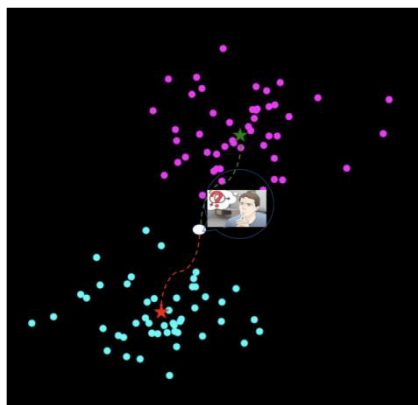


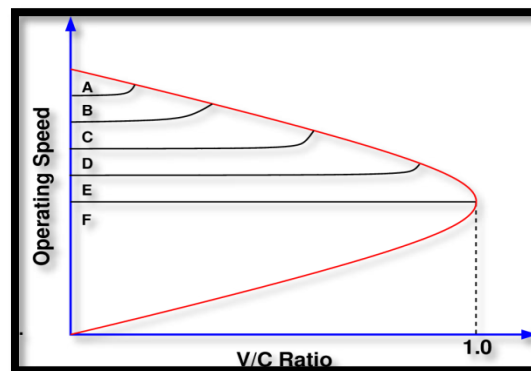
Figure 1 shows clustering technique using K-Means

The procedure is adopted to initiate the clustering models

1) Specify number of clusters as K.

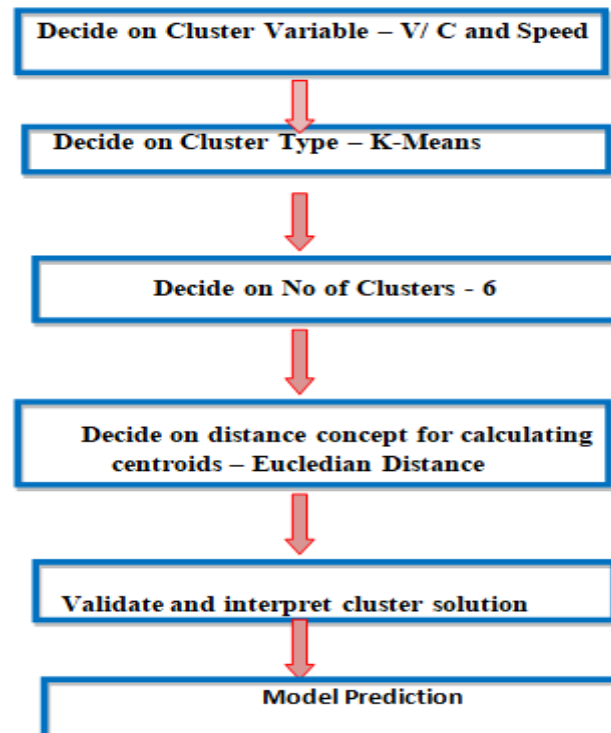
- 2) Initialize centroids by first shuffling the dataset and then randomly select K data points for the centroids without replacement.
- 3) Iterate until there is no change to the centroids.
- 4) Compute the sum of the squared distance between data points and all centroids.
- 5) Assign each data point to the closest cluster (centroid).
- 6) Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.

This approach of K - means follows to solve the problem can be called as Expectation- Maximization. The E-step is assigning the data points to the closest cluster. The M-step is computing the centroid of each cluster. The input parameters for the clustering of the traffic data into different Level of Service are Volume (No of vehicles on Roadway) and speed of the vehicle. The capacity of a given roadway facility is a constant. For computation of LOS, volume to capacity ratio is to be computed and this would be on the X axis or X vector component. The Y component is speed which can be represented on the Y-axis.



Level of Service with V/C ratio on X axis and Operating Speed on Y axis

An artificial data set of 5100 data points was created in Excel and was input into the program as CSV file format. The process includes stepwise as Data Collection → Pre Processing of Data → Input Parameters

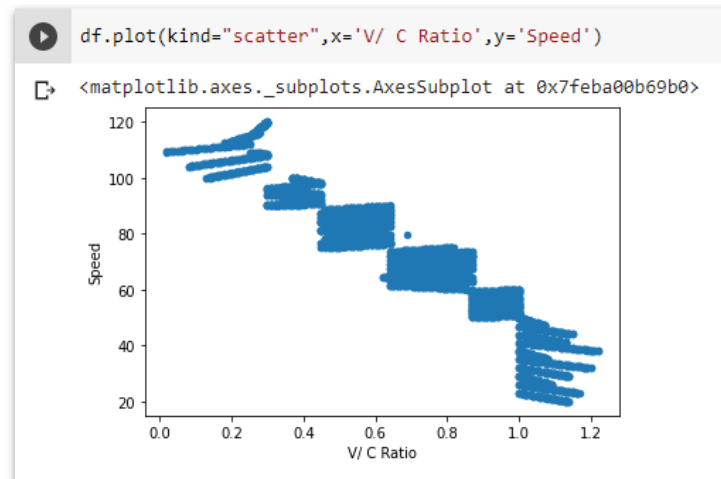


Flow Chart shows the Procedure in Clustering



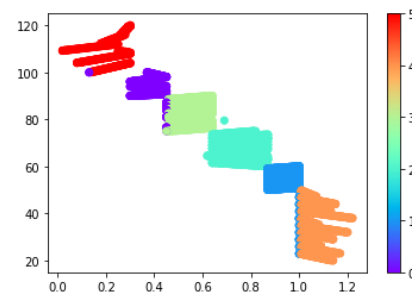
Python language is used in processing the data for the desired appropriate output as a part of Machine Learning. Input is given in the attributes as V/C Ratio and Speed.

The output can be visualized as follows



Graph shows the output as scattered

```
[ ] plt.scatter(df['V/ C Ratio'],df['Speed'],c=y_kmeans,cmap='rainbow')
plt.colorbar()
plt.show()
```



```
[ ] print(set(kmeans.labels_))
```

```
{0, 1, 2, 3, 4, 5}
```

Graph shows the output as scattered when applied colors

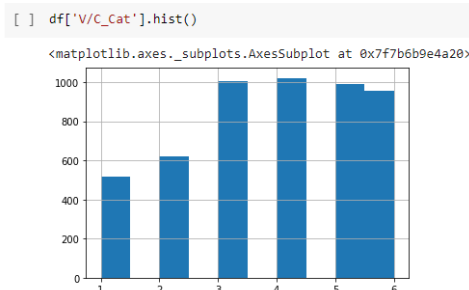


Figure shows Bucketing of Clusters

## Results

A case study is done with inputs as volume, capacity and speed and the model is effectively predicting the LoS of the roadway. The entire data set was used to train the model. The model effectively produced expected results. The various output as per different inputs are shown in figures below

#### Model Prediction

```
#@markdown Enter Data

Volume = 2850#@param {type: "number"}
Capacity = 2750#@param {type: "number"}
Speed = 48.40#@param {type: "number"}
```

Enter Data

Volume: 2850

Capacity: 2750

Speed: 48.40

#### Result Showing Worst Service Condition

#### Model Prediction

```
[66]
#@markdown Enter Data

Volume = 300#@param {type: "number"}
Capacity = 2750#@param {type: "number"}
Speed = 80#@param {type: "number"}
```

Enter Data

Volume: 300

Capacity: 2750

Speed: 80

#### Result Showing Good Service Condition

### III. CONCLUSION

Los being a qualitative measure used for the traffic condition was appropriately achieved using Python, a machine learning technique. In this paper, the given set has comparatively less data points and was synthetically developed. This being an unsupervised learning has less accuracy score when compared to manual testing which proves to give nearly exact output. The present work can be further extended with multivariate analysis. It is also possible conduct a supervised learning on the same problem can be considered as the scope of future work.

### REFERENCES

- [1] Mehmet Ali SİLGÜ1, Hilmi Berk ÇELİKOĞLU1, “ K-MEANS CLUSTERING METHOD TO CLASSIFY FREEWAY TRAFFIC FLOW PATTERNS” , 10.5505/pajes.2014.36449 , Research Article/Araştırma Makalesi , 10.01.2014
- [2] M. Montazeri-Gh, A. Fotouhi “Traffic condition recognition using the k-means clustering method” Scientia Iranica, www.sciencedirect.com, Scientia Iranica B (2011) 18 (4), 930–937, 7 June 2011
- [3] Rudra Pratap Deb Nath, Nihad Karim Chowdhury, Jae-Woo Chang “Modified K-Means Clustering for Travel Time Prediction Based on Historical Traffic Data”, <https://www.researchgate.net/publication/221019408> , 10.1007/978-3-642-15387-7\_55, September 2010
- [4] Building Materials In A Green Economy: Community-Based Strategies For Dematerialization Brian Milani Oise-Ut Aecdcp / U. Of T. Institute For Environmental Studies Supervisor: Prof. E.O’sullivan
- [5] Zhongheng Zhang Introduction to machine learning: k-nearest neighbors VOL 4 NO 11 JUNE 2016



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



**www.ijirset.com**

Scan to save the contact details