ICISC

PROCEEDINGS
of the
5th International Conference on Inventive Systems and Control (ICISC 2021)

7-8 January 2021

Technical Sponsors

Springer

JCT
COLLEGE OF ENGINEERING AND TECHNOLOGY

UNSUPERVISED DEEP LEARNING ON SPATIAL-TEMPORAL TRAFFIC DATA USING AGGLOMERATIVE CLUSTERING

S. SENTHILARASI[1] AND Dr. S. KAMALAKKANNAN[2]
[1]Research Scholar, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Department of Computer Science, Chennai.
[2]Associate Professor and Research Supervisor, Department of Computer Science, School of Computing Sciences
Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai.
[1]CORRESPONDING AUTHOR: senarasi80@gmail.com
[2] kannan.scs@velsuniv.ac.in

ABSTRACT:

Now-a-days, road traffic arises of the increased population because of employment in urban areas of all places. This creates to be a challenging aspect , uses certain measures to solve then and there by the authorized people in situation. The varying types of services related to traffic streams are pedestrian census, turning movement census and others. The highlighting features are tabulated on the basis of traffic management techniques. The proposed study on Agglomerative Clustering on traffic data have been satisfied with its enhanced features. An agglomerative clustering is an extensive model of hierarchial clustering , bottom-up approach that combines the similarities of samples in clusters. The quite number of methods to find optimal clusters in different ways are briefly discussed to support the clustering. Experiments studied on California-Traffic solution-Data from SWITRS conducted with traffic data proves the optimal number of clusters formed, its validity using method of clustering accuracy and picturization of time series on traffic data are shown in different categories.

KEYWORDS: Traffic Services, Traffic Management Techniques, Agglomerative Clustering, Optimal Clusters, Time Series Visualization

1. INTRODUCTION:

At present, the growth of human population is excessive in urban areas which states to be the top most challenges for worldwide countries. Such rapid growth in city populations, mode of transport, movement of vehicles, causes of air pollution affects both the environment and human life[1]. The excessive movement of vehicles leads to road traffic congestion that creates critical issue in urban places. An article's report shows the information nearly 600 billion rupees are spent each year because of road traffic congestion (or) its dally frequently occurs on high extent roads, highways added with wastage of fuel[8].

The prediction of traffic is considered as an crucial part by providing correct and valid traffic details for both voyagers and traffic agencies in a modern Intelligent Transportation Systems(ITS) and Advanced Traveller Information System (ATIS). The attributes of traffic details are such as traffic blockage positions, traffic size and mass flows in earlier, can support the officials in better way to manage the traffic strategies and voyagers to plan alternate routing plan of actions[2].

An assessment of traffic contains the different types of related collection of data methods examined on different view of traffic streams . These views comprehends the subsequent services[3].

Different types of Services

1. Pedestrian Census
2. Turning movement census
3. Bicycle census
4. Link flow traffic census
5. Rail and bus punctuality census
6. Travel Time
7. Round about census
8. Parking census technique
9. Queue length census
10. Recognition of number plates
11. Illegal movement census
12. Traffic Saturation flows

1.1 Higlighting features in traffic management techniques:-

Due to the excessive growth of population in place of urban areas shows the most biggest strengths for worldwide administrations. The increased use of motor vehicle on roadways not only affects the surrounding environment also harms the quality of human life and their wealth[4] .

Table 1: Different Types of Traffic Management

TYPES OF TRAFFIC MANAGEMENT	BASIC FEATURES	ADDITIONAL FEATURES	PROS	CONS
Manual Traffic Control	- It is the simplest way of managing the traffic - It involves a policeman standing on every interconnection of roads to control the regular traffic flow	- A traffic police controls the regular traffic flow by standing on traffic poll by showing the sign board - In case of large volume of traffic, he/she signals the vehicle drivers to drive (or) not to drive	The people on roadways can understand about the traffic in view of emergencies , traffic police can prioritize the movement of vehicles	The system involves human as main role in controlling traffic is inefficient
Automatic Traffic Control	It is an automatic traffic control system	-The system adds three colored signals red, yellow and green -Each signal light is set with timings in terms of seconds to glow sequentially to alert the	It overcomes the difficulties of manual traffic	It cannot differentiate the movement of vehicles with emergencies and without emergencies.

		travellers.		
Image Processing Traffic Control	It involves cameras to record images of large volume of traffic lying on roads.	-It is normally fixed on high pillars to cover the traffic to certain distance from its positioned view. -The recorded images are validated by a chip in computer to find the number of vehicles on road to compute the timings for signal lights especially red and green to manage the traffic density	It dynamically changes the recorded timings of signal lights	It cannot capture the images clearly in case of heavy rain and heavy traffic jam
Wireless Traffic Control	Emergency Vehicle is facilitated with RF-Transmitter and Receiver placed on the signal pole	-It alerts the RF_Receiver to send data to main control system. -The system computes the timing approximately for the green signal to glow during the movement of emergency vehicles and other vehicles to stop by signaling the red light	It provides a solution in case of emergency vehicles passing on the road	It's maintenance cost is high due to sensors, load cells in accumulating traffic data
IRIS	It is an original source software project to observe and control the lay of traffic lying on	-It gives the real-time information existing on highways to find traffic incidents, control the traffic	It is better than wireless traffic control	It is expensive to buy and its recovery maintenance costs also expensive

	connecting roadways such as interstate and highways	flow and relay information of travellers. -It is using GPL license, ATMS Software tool .		

2.RELATED WORK:

Recently, a quite number of papers have been published by various authors in Spatial – Temporal combination on road Traffic data by using different Clustering techniques. The following tabular form depicts the study of different papers related to road traffic data [5],[6],[7],[8],[9],[10],[11],[12],[13],[14],[15],[16],[17],[18],[19].

Table 2: Study of Spatial-Temporal Traffic data

S.NO.	AUTHOR	TITLE OF THE PAPER	CLUSTERING TECHNIQUE	ACHIEVEMENTS	YEAR	EXPERIMENT STUDY
1	Wendy Weijermans and Eric Van Berkum	Analyzing highway flow patterns using cluster analysis	Ward clustering	Maintains traffic profile daily of individual clusters, Integration of clusters	2005	Dutch highway location
2	Weichiet Ku, George R. Jagadeesh	A clustering-based approach for Data-Driven imputation of Missing traffic data	KMeans Clustering	A SDAE model is constructed to structure the Spatial-temporal relationships for every group of road divisions	2016	Jurong West area of Singapore collected in Quantum Inventions
3	Mei Yeen Choong, Lorita Angeline, Reneee Ka Yinchin, Kiam Beng Yeao, Kenneth Tze Kin Teo	Vehicle Trajectory Clustering for traffic intersection surveillance	KMeans and Fuzzy Cmeans	Implements KMeans and Fuzzy Cmeans on trajectories, Usage of Longest Common Subsequences with different trajectories lengths to pinpoint noise and outliers.	2016	Computer Vision Robotics Research Laboratory(CVRR) of 100 vehicle trajectories
4	Vishwajeet Pattanaik, P.K. Gupta, S.K. Singh and Mayank Singh	Smart real-time traffic congestion estimation and clustering technique for urban vehicular roads	KMeans Clustering	Establishes traffic estimation system on Global Positioning System(GPS) to activate the availability of mobile phones for	2016	Several roadmaps of NewDelhi

				development. Discovers shortest path from Source to Destination using Dijkstra's algorithm		
5	Maithew Aven	Daily flow Pattern Recognition by Spectral Clustering	KMeans Algorithm, Spectral Clustering	Proves the accurate relationships arises between the dataset values	2017	New York City Intersection case study
6	Senyan Yang, Jianping Wu, Geqi Qi and Kun Tian	Analysis of traffic state variation patterns for urban road network based on Spectral Clustering	Spectral Clustering	Extracts the traffic features variances for the road network , transforms traditional clustering to graph partitions	2017	North East region of Beijing
7	Mohammad Billah, Arash Masooki, Farzana Rahman and Jay A. Farrell	Roadway feature mapping from point cloud data: A Graph-based clustering approach	Graph based clustering		2017	Advanced Highway Maintenance and Construction technology Research Center (AHMCT) – 3D point clouddata
8	Weirong Liu, Member IEEE Gaorang Qin, Yun He and Fei Jiang	Distributed Cooperative Reinforcement learning –based traffic signal control that integrates V2X network's Dynamic Clustering	Dynamic Clustering	Initializes the partitions of clusters to retain in correct size in addition with lane and destination to maximize the stability of clusters. Balancing of traffic load using Reinforcement learning	2017	V2X Network's data collection
9	Guojieng Shen, Chaohuan Chen et.al.	Research on Traffic Speed Prediction by Temporal Clustering nalysis and Convolutional Neural Network with Deformable Kernels	Hierarchial Clustering	Initiation of deformable convolution layer into DCNN to resize the kernels, extends the strengths of DCNN to structure the Spatial-Temporal data	2018	Xiasohan district of Hangzhou traffic police Brigade , Zhejiang, China
10	Sakhawat Hosain Sumit and Shamin	C-means Clustering and deep-neuro-fuzzy classification	Cmeans, Fuzzy clustering	Handles Overlapping uncertainities which	2018	Environmental data(750 days 12/1/2006 -

	Akhter	for road weight measurement in traffic		provides the ability to perform task with multidimensional featured data		12/10/2008)collected from AccuWeather
11	Yuntao Chang and Bin su	Construction of the Driving cycle of vehicles Queuing at toll station	Fuzzy Cmeans clustering, Markov Method	A model on space-based vehicles driving constructed in relation to queuing vehicles, Comparison of effect between manual and ETC Collection System	2018	Jiangqiao toll station on G2, Beijing-Shanghai Expressway
12	Weixiang Xu and Jiaojiao Li	An improved algorithm for clustering uncertain traffic data streams based on Hadoop Platform	AF- DBSCAN clustering	An UTD-SPCQ algorithm is used to receive data and its buffering are realized ,MapReduce Framework of distributed structure in hadoop platform	2019	Berijing's taxi dataset
13	Tharingu Bandaragoda, Darwin De Silva,	Trajectory Clustering of road traffic in Urban environment using incremental machine learning in combination with hyperdimensional computing	Hyperdimensional computing	An unsupervised incremental learning approach addresses two key problems in traffic profile. Enables the Variable-length trajectories of traffic trips.	2019	Bluetooth Traffic Monitoring System(BTMS), Arterial roads in Victoria, Australia
14	Hiroki Watanobe, Tomas Maly, Johannes Wallner	Cluster-Linkage analysis in Traffic Data Clustering for Development of Advanced Driver Assistance Systems	Hybrid Clustering(on the basis of K-Covers and KMeans Algorithms)	Discovery of optimal number of clusters, Cluster models using average Silhouette Width, Reduced test effort in the development of ADAS	2020	Vehicle-Pedestrian near crashes in SHRP2 dataset
15	Weijing Qi, Bjorn LandFeldt, Qingyang Song, Lei Guo, Abbas	Traffic Differentiated Clustering Routing in DSRC and C-V2X Hybrid Vehicular Networks	One-hop clustering	A Traffic Differentiated Clustering Routing(TDCR) mechanism in a Software Defined	2020	Data collection from CHs to data center

	Jamliopour			Network (SDN) enables hybrid vehicular network. Includes centralized one-hop clustering and dispatch of data		

3. PROPOSED STUDY IN TRAFFIC DATA:

The Proposed study on traffic data utilizes the bottom-up approach of hierarchial clustering- Agglomerative Clusters, Which merges the samples of similar features closer recursively till the single cluster is formed based on samples outcomes.

The proposed model of traffic data (Figure 1) shows the input of various features , has to be found and clustered using Agglomerative clustering approach . The similarities of clusters are given importance and notified rather than dissimilar features considered as outliers. The outcome of traffic data contains the number of clusters formed are evaluated using measures of aggregation, and the validity of clusters are found using the method of Silhouette Coefficient.

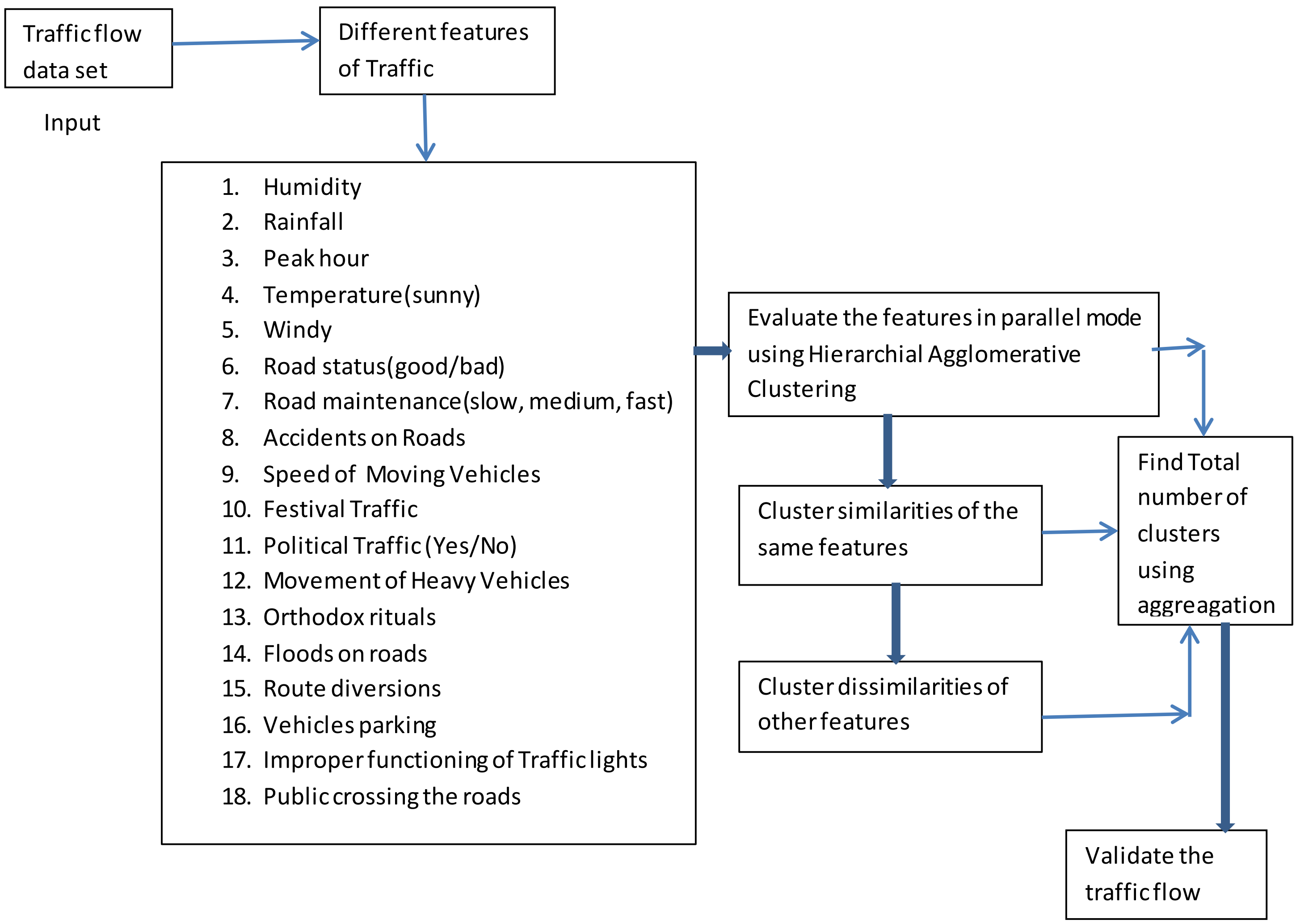


Figure 1: TRAFFIC FLOW USING AGGLOMERATIVE CLUSTERING MODEL

Output

3.1 Preprocessing the Traffic data:

The traffic data are of different types mainly depends on analysis and visualization related methods. The actual data are generally raw, data with errors, uncertainities , number of missing values, outliers (or) items mismatching. In particular raw data is considered for analysis and visualization.

3.1.1 Data Collection:

The number 'n' of datasets can be created and accumulated by sensing devices fixed in traffic signals, vehicles (or) CCTV's fixed on traffic accumulating places on roadways.

Examples related to Traffic data are

- Vehicles GPS data
- Locations of GSM
- Recording Video / Image of surveillance devices.

The sensors mode of process can be classified into three groups namely

3.1.1.1 Based on Location – The object's location is recorded by getting into the sensor range. For example, at a road's criss-cross places, a video recording device catches the location and wandering of foot travelers iff crossed across the monitor.

3.1.1.2 Based on Activity – Whenever a device reacts to certain activities, its related information is noted. For example, the user's GSM location is noted in automatic process during a call.

3.1.1.3 Based on Device – An object's device track record and directs the location and other related concepts back actively. Example- for every 20 seconds, a car sends a signal to the data center with the use of GPS-device.

3.1.2 Properties of Traffic data:-

In many number of instances, traffic data holds different attributes of spatial and temporal additional information. These properties can be classified into three classes namely

- Numbers
- Category
- Text

3.1.2.1 Numbers:

It represents the continuous valued variables which means the data object's quantitative values. Every numbered value explains any one data object's aspect particularly speed, weight etc. Most of the attributes are time-variant , time-oriented techniques in visualization.

For instance, A better choice for visualization is histogram.

3.1.2.2 Category:

It represents the discrete set of valued variables that explains the data objects status. Type of vehicles, directions and incidents are instances of categorical properties. This property is simplified in visual form as color mapping, assigning a specified color to represent a value.

3.1.2.3 Text:

It represents the words, lexical information that illustrates additional information related to traffic such as vehicles included in an incident, intersection points and others. The attributes containing semantic information are necessary to analyze and describe traffic situations[20].

4. EXPERIMENTAL STUDY:

The Traffic data is collected from California-Traffic solution-Data from SWITRS[26]. The data includes the collisions from January 1st 2001 til the mid of October 2020. The dataset contains three main tables, consisting of collisions, parties and victims. The collision table contains 74 field names related to traffic by describing with 9.17m rows. The parties table holds the information about the people's age, gender and serious behavior included in collision. The victims table consists of the details related to the people got injury in the collision.

The proposed algorithm on Spatial-Temporal data accepts the 'n' number of attributes on traffic data as input which has to be read and normalized using *scaler* functions. The normalized data are then transform into *data frames* to predict the number of clusters. The dimensionality of the data has been reduced. Finally, the outcome of clusters are displayed in advanced feature of hieararchial clustering *dendrogram* , clusters in circular form with k values (2,3,4,5,6) . The clusters validity have been computed using the method of Silhouette Coefficient , which proves the Strong Clustering .

4.1 Algorithm: Spatial – Temporal Traffic data using Agglomerative Clustering

Input: Dataset consisting of attributes(1,2,….n)
Ouput: Provides optimal number of clusters

1. Read the data(1,2,…n)attributes
2. Repeat
   (i) Normalize the data
   (ii) Fit the normalized data in DataFrame
   (iii) Transform the normalized data to predict clusters
   (iv) Add dendrogram to display the output of hierarchial clustering
   (v) Cluster the data for K=2,3,4,5and 6, where k represents the number of clusters
   (vi) Compute the clustering validity using Silhouette coefficient
   Until the stopping condition is met
3. Output the optimal number of clusters

4.2 Agglomerative Clustering:

Agglomerative Clustering is a plan of hierarchial clustering. It is otherwise known as Connectivity based clustering, which is a behavior of cluster analysis that searches to construct a hierarchy of clusters. Hierarchial clustering is established on the central plan of objects being more closed to neighboring objects compared to distant objects. A cluster can be outlined in wider by the extreme value required in connecting parts of clusters.

With varying distances, clusters in difference can be stated by means of "Dendrogram"(Figure 2) which issues an increased hierarchy of clusters that combines with one another lying at fixed distances.

In a dendrogram,

X axis – Objects are placed, Clusters do not mix.

Y axis – Distance is marked to merge the clusters[21].

4.2.1 Factors of Agglomerative Clustering:

1) Initially, it looks more trustworthy due to over clustering. Representation of clustering from a Convolutional Neural Network are initiated with random weights, which are not efficient, whereas the closest neighbors and over clustering are generally acceptable.
2) The over clustering can be integrated to make it as a better representation of learning.
3) It is a recurrent type of process that occurs in a recurrent framework[22].

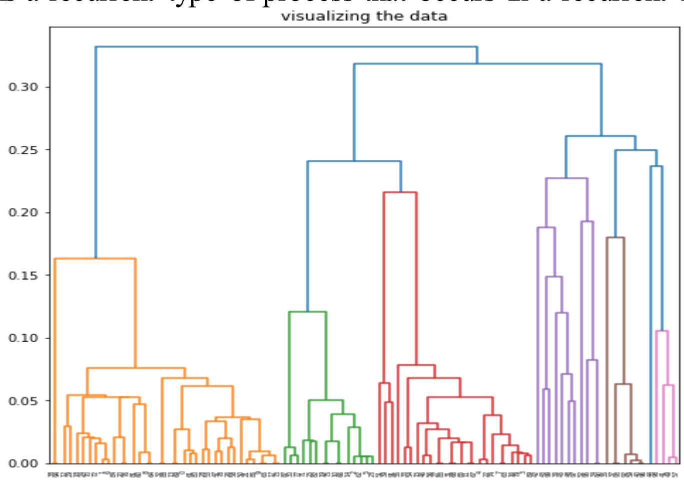


Figure 2: Dendrogram of Traffic data

4.2.2 Methods to find Optimal Clusters:

The validity of clustering indexes are generally defined by integrating both the compactness and divisibility of the clusters. The closeness of elements in clusters are evaluated by compactness where the difference between two clusters are represented using the divisibility measure.

The common methods to yield the clustering quality are

- Davies -Bouldin Index
- Dunn Index
- Elbow Method
- Silhouette Method

4.2.2.1 Clustering Method- Davies - Bouldin Index:

The average similarity exists among every cluster with its similar are measured using DBIndex. The minimum values of DBIndex states that clusters are compactly tight and distinguishes the cluster in a better form. The aim of this index is to fulfill “minimum within-cluster variance and maximum between cluster seperations”[23].

The Davies – Bouldin Index can be computed as follows[25]:

$$DB = \frac{1}{n}\sum_{i=1}^{n} max_{j\neq i}\left(\frac{\sigma_i + \sigma_j}{d\left(c_i, c_j\right)}\right)$$

4.2.2.2 Clustering Method – Dunn Index:

The Dunn Index(DI) measures the dataset clusters, is expected to be maximum and the clusters are separated. The clusters are so compact and distinguished by increasing the intercluster distance, whereas, the intra-cluster distance is in decreasing[23].

The Dunn Index is computed using the form[25]:

$$D = \frac{\min_{i\leq a\leq b\leq n} X(\mathrm{a}, \mathrm{b})}{\max_{i\leq a\leq c\leq n} Y(\mathrm{c})}$$

Where i,j,k are clusters indices

X = distance variable of inter-clusters

Y = distance variable of intra-clusters

4.2.2.3 Clustering Method – Elbow:

It is an analytical method of expansion and verification of standards within cluster analysis planned to assist in finding dataset’s suitable number of clusters[24].

4.2.2.4 Clustering Method – Silhouette Coefficient:

The Silhouette Coefficient depicts the fitness of objects within the cluster (Figure 3). The clustering quality is measured between the range -1 and +1.

+1 score represents the samples are distanced from its nearest clusters.

0 score represents the sample is too close to the decision limits which divides the two nearest clusters.

-1 score represents the samples are assigned to the wrong cluster.

The Silhouette score can be computed as follows

Silhouette Score = (a-b)/max(a-b)

Where ‘a’ represents the distance to the points in the closest cluster

‘b’ represents intra-cluster distance of all the available points[25].

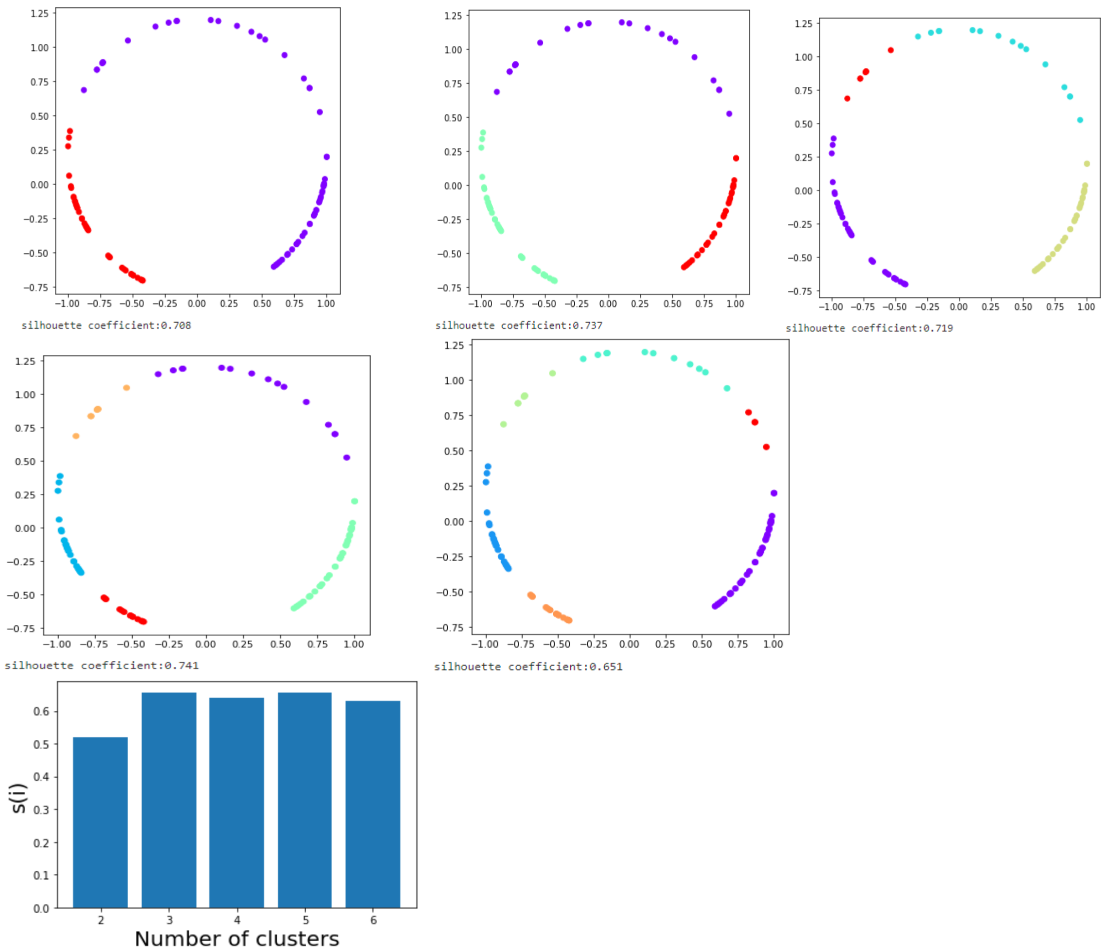


Figure 3: Optimal number of clusters and its validity

4.3 Time Series Visualization in Traffic data:

Visualization plays a major task in analysis of time series and its forecasting. The plotting of data issues the most valuable features to find temporal models such as (trends, cycles and seasonality) to enhance the type of model.

There are six varying types of visualization applied on time series data (univariate time series).

1. Line plots
2. Histograms and Density Plots
3. Box and Whisker Plots
4. Heat Maps
5. Lag Plots (or) Scatter Plots
6. Autocorrelation Plots

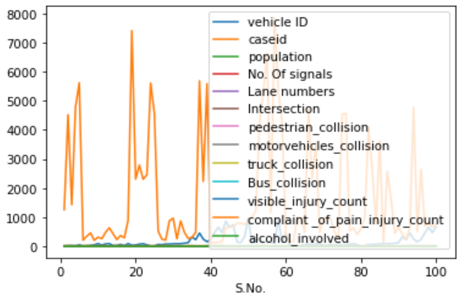


Figure 4: Line Plot

Figure 5: Histograms and Density Plots

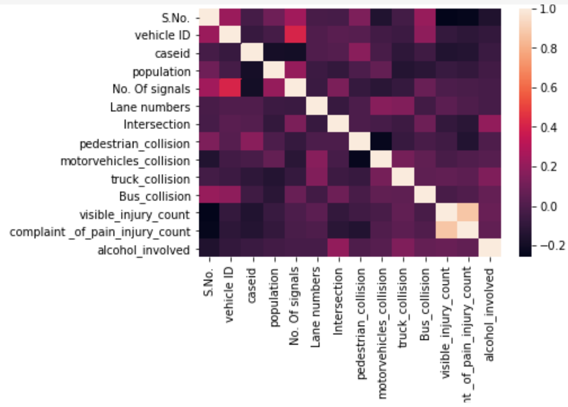


Figure 7: Heat Map

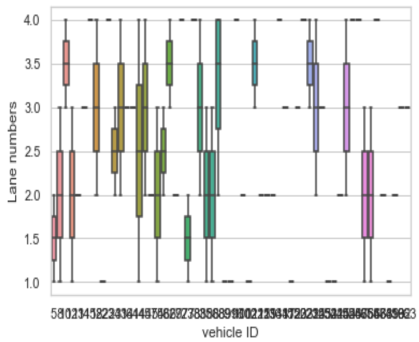
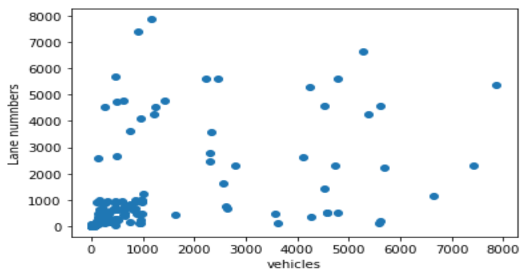


Figure 6: Box and Whisker Plot



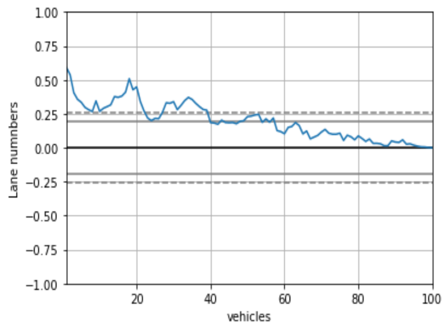


Figure 8: Lag Plot (or) Scatter Plot

Figure 9: Autocorrelation Plot

Time Series Using Line Plots – The traffic data are enumerated and its recordings for each date stored as columns using a new data frame.

Time Series Using Histogram and Density Plots – A histograms combines the values into bins , number of results available in each bin provides intuition into the basic diffusion of the results.

Time Series Using Box and Whisker Plots – This type of plot is created for each day of traffic and placed side-by-side towards direct match-up.

Time Series Using Heat Map –This concept plots the matrix numbers as an aspect, where the every cell values related to the matrix are given a unique color.

Time Series Using Lag plot – It is a convenient type of plot which exhibits the relation between every result and its lag value called as Scatter Plot[27].

Time Series Using Autocorrelation Plot – It is also known as "Correlation Coefficients", which can be computed for every outcome and varying lag values.

Values near to 0 represents a weak correlation

Values near to -1 (or) 1 represents a strong correlation.

5. DISCUSSION:

The agglomerative clustering on traffic data was experimented on California-Traffic solution-Data from SWITRS showed in familiar concept of advanced hierarchial clustering *dendrogram* , also validity of clusters formed has been with the help of clustering accuracy method. The concept of dendrogram is unique in hierarchial clustering , where as it not applicable in other types of clustering. The visualization of traffic data in terms of time series has been depicted in different categories . The proposed study on this paper is just a step regarding the better representation of clustering, where the concept has to be further enhanced by the comparative study with other clustering models to prove the increased level of accuracy. In addition to, the agglomerative clustering approach has to be portrayed in a most better form compared to other clustering models.

6. CONCLUSION:

This paper is focused on the road traffic data occurring in high density population areas . the similarities of collision happens due the movement of vehicles on road has been classified and studied on the California-Traffic solution-Data from SWITRS. The model of Agglomerative Clustering had been applied to analyze and evaluate the optimal number of clusters using dendrogram. The validity of clusters have been proven with Strong relationship with clusters using the measure of Silhouette Coefficient ,which is one of a method of cluster accuracy. The view of traffic data occurs in time series have been shown in varying methods. The accuracy and representation of clusters are valid compared to KMeans Clustering , Spectral Clustering, OPTICS clustering and BIRCH clustering. Further , the clustering on agglomerative approach has to be enhanced in a better representation added with accuracy, execution time, speed and performance.

REFERENCES:

[1] *Nestor Cardenas Benitez, Raul Aquino-Santos, Pedro Magna-Espinoza et.al*, Traffic Congestion System through connected and vehicles and Big data, Sensors 2016

[2] *Shengnan Guo, Youfang Lin, Shijie Li, Zhaoming Chen and Huaiyu Wan*, Deep Spatial-Temporal 3D Convolutional Neural Networks for Traffic Data Forecasting, IEEE Transactions on Intelligent Transportation Systems, 2019.

[3] VTRAC- Intelligent Traffic Solutions

[4] *Siddamma, G. Pahupatimath, Prof. Shipa Madhavanavar*, Different techniques used in Traffic Control System: An introduction, IJERT

[5] *Wendy Weijermars and Eric Van Berkum*, Analyzing highway flow patterns using cluster analysis, IEEE,2005,pp(831-836)
[6] *Wei Chiet Ku, George R. Jagadeesh, Alo Prakash, Thambipillai Srikanthan*, A clustering – Based Approach for Data-Driven Imputation of Missing Traffic data, IEEE, 2016
[7] *Mei Yeen Choong, Lorita angeline, Renee Ka Yin Chin, Kiam Beng Yeo, Kenneth Tze Kin Teo*, Vehicle Trajectory Clusteriung for Traffic Intersection Surveillance, IEEE, 2016, International Conference on Consumer Electronics-Asia
[8] *Vishwajeet Pattanaik, Mayank Singh, P.K. Gupta and S.K.Singh*, Smart real-time Traffic congestion estimation and clustering technique for urban vehicular roads, IEEE, 2016
[9] *Maithew Aven*, Daily Traffic flow pattern recognition by spectral clustering, CMC Senior Theses, 2017, pp(1-16)
[10] *Senyan Yang, Jianping Wu, Geqi Qi and Kun Tian*, Analysis of Traffic state variation patterns for urban road network based on spectral clustering, Advances in Mechanical Engineering, 2017
[11] *Mohammad Billah, Arash Maskooki, Farzana Rahman and Jay A. Farrell*, roadway feature Mapping from point cloud data: A Graph-based Clustering Approach, IEEE, 2017, Intelligent Vehicles Symposium
[12] *Weirong Liu, Member, IEEE Gaorang Qin, Yun He and Fei Jiang*, Distributed Cooperative Reinforcement Learning – based Traffic Signal control that integrates V2X Networks Dynamic Clustering, IEEE, 2017, Transactions on Vehicular Technology
[13] *Guojiang Shen, Chaohuan Chen, Qihong Pan, Si Shen and Zhi liu*, Research on Traffic speed prediction by Temporal Clustering Analysis and Convolutional Neural Network with Deformable Kernels, IEEE Access, 2018, pp(51756-51765)
[14] *Sakhawat Hosain Sumit, Shamim Akhter*, C-means Clustering and deep-neuro-fuzzy classification for road weight measurement in traffic management system, Springer, 2018
[15] *Yuntao Chang, Bin Su*, Construction of the Driving Cycle of Vehicles Queuing at Toll Station, IEEE, 2018, International Conference on Intelligent Transportation Systems(ITSC)
[16] *Weixiang Xu and Jiaoliao Li*, An improved algorithm for clustering uncertain traffic data srtreams based on Hadoop Platform, World Scientific, 2019,pp(1950203-1 to 1950203-19)
[17] *Tharingu Bandaragoda, Daswin De Silva et.al*, Trajectory Clustering of road traffic in urban environments using incremental machine learning in combination with hyper dimensional computing, IEEE, 2019, International Transportation Systems Conference
[18] *Hiroki Watanobe, Tomas Maly, Johannes Wallner, Gunther Prokop*, Cluster-Linkage Analysis in Traffic Data Clustering for Development of Advanced Driver Assistance Systems, ICICT, 2020, pp(54-59)
[19] *Weijing Qi, Bjorn Lanfeldt, Qingyang Song, Lei Guo, Abbar Jamlaipour*, Traffic Differentiated Clustering Routing in DSRC and C-V2X Hybrid Vehicular Networks, IEEE,2020, Transaction on Vehicular Technology
[20] *Wei Chen, Fang Zhou Guo and Fei-Yue Wang*, A survey of Traffic Data Visualization, IEEE, 2015, Transactions on Intelligent Transportation Systems, pp(1-15)
[21]docs.rapidminer.com/latest/studio/operators/modeling/segmentation/agglomerative-clustering.html
[22] *Jianwei Yang, Devi Parikh, Dhruv Batra*, Joint Unsupervised learning of Deep Representations and Image Clusters, 2016.
[23] *Sadia Nawrin, Md. Rahatur Rahman, Sharmin Akhter*, Exploring K-Means with Interval Validity Indexes for Data Clustering in Traffic Management System, IJACSA 2017

[24]medium.com/@masraudheena/4-best-ways-to-find-optimal-number-of-clusters-for-clustering-with-python-code-706199f9957C
[25] tutorialspoint.com/machine-learning-with-python/machine-learning-with-python-analysis-of-silhouette-score.htm
[26] kaggle.com/alexgude/California-traffic-collision-data-from-switrs
[27] machinelearningmastery.com/time-series-data-visualization-with-python