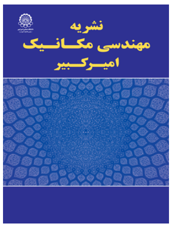

Amirkabir Journal of Mechanical Engineering

Amirkabir J. Mech. Eng., 53(9) (2021) 1177-1180
DOI: 10.22060/mej.2021.19222.6980

Comparison of Support Vector Machine and K-Means Algorithms Performance in Extracting the Real Driving Cycle of Combined Tehran-Amol

T. Qaraati[1], A. Momenimovahed[1], M. Azadi[2]*, S. A. Moosavian[3]

[1] Department of Mechanical Engineering, Imam Khomeini International University, Tehran, Iran
[2] Department of Mechanical Engineering, Semnan University, Semnan, Iran
[3] Department of Agriculture Engineering, Shahriyar Technical and Vocational University, Tehran, Iran

ABSTRACT: Driving cycles represent the vehicle speed as a function of time and are used in vehicle design, fuel management, and the improvement of standard indicators. In this study, four combined driving cycles were extracted using real data. The data was collected from a passenger car with a gasoline engine under real driving conditions while driven from Tehran to Amol based on the car chasing method. A code was generated in MATLAB software to create the desired cycle using support vector machine and K-means algorithms considering mid-range and mean values as group centers. The characteristic parameters of the cycles such as the average speed and the percentage of the car travel time at idle, cruise, accelerating, and decelerating conditions were also calculated. These cycles were compared based on the mean relative error, the root-mean-square error, and the Chi-square test. The results showed that the cycles extracted by the support vector machine were closer to the allowable time interval (less than 1800 seconds); however, the cycle extracted by the K-means algorithm with the mean value as the centers of the generated categories, recorded the least errors. This cycle, in addition to spending most of its time in accelerated motion, represented a greater amplitude of acceleration and velocity fluctuations than other cycles.

Review History:

Received: Nov. 06, 2020
Revised: May, 13, 2021
Accepted: May, 14, 2021
Available Online: May, 19, 2021

Keywords:

Driving cycle
Support vector machine
K-means algorithm
Combined cycle
Classification and clustering

1- Introduction

Driving cycles are data of the speed versus the time that are normally used to assess the performance of the vehicles. There are two common perspectives on categorizing driving cycles.

In the first view, the classification is based on the purpose, which the cycle was proposed. In this way, the cycle falls into one of the legislative or non-legislative groups. The first group is created to determine or modify criteria and is often commissioned by governments. The second group, however, is created for research purposes usually by researchers. The second view classifies cycles into real and synthetic groups according to the data collection method. Synthetic data collection is performed in the laboratory by placing the vehicle in acceleration, deceleration, constant speed, and stopping conditions. The diagram of these cycles shows small fractures and many continuous lines, which makes them non-representative of the real driving conditions. In contrast, in the real cycle data recording, the vehicle is driven on the road under real-world driving conditions and records information, needed to generate the cycle.

Zhao et. [1] proposed the Xi'an driving cycle by taking data from an electric vehicle, using the car chasing method. They grouped the micro-trips with K-means and support vector machine algorithms. They chose the cycle generated by the support vector machine as the final cycle. Their results showed that this cycle spent a short time in the cruise state. Fotouhi and Montazeri [2] employed an advanced vehicle location recorder and followed the car chasing method to collect the data required for the production of the Tehran cycle from a passenger vehicle. Comparing the created cycle with standard cycles, it was found that the Tehran cycle had a higher average speed than others. Anida and Salisa [3] collected the data needed to generate the real cycle of Kuala Terengganu from three electric vehicles with different battery panels including parallel, series, and single-plate. Data were then turned into micro-trips and categorized by K-means clustering. This cycle was given as an input to AUTONOMIE software to estimate the pollution and the fuel consumption of the vehicles. The results indicated that this cycle was more in the acceleration and deceleration states. The amount of fuel consumption and the carbon dioxide production in the series battery panels was estimated to be more than others.

2- Methodology

The route had a length of 194 km, 79% was on highways, and the rest was on major and minor city streets. Another important feature of this route was its wide range of altitude changes so that it covers a range of about 97 to about 2700 meters above sea level.

*Corresponding author's email: m_azadi@semnan.ac.ir

Table 1. Turbine characteristics

Traffic condition	Light	Semi-light	Semi-heavy	Heavy
Mean speed (km/h)	> 65	40-65	20-40	< 20
Driving duration (s)	> 900	300-900	120-300	< 120

Table 2. Turbine characteristics

Parameters	Value
Duration (s)	1868
Percent of driving time	82.80
Percent of idle time	17.20
Percent of cruise time	4.31
Percentage of accelerated time	38.04
Percentage of deceleration time	40.45
Average driving speed (km/h)	30.26
Average travel speed (km/h)	25.05
Maximum speed (km/h)	123.23
Average positive acceleration (m/s^2)	1.96
Average negative acceleration (m/s^2)	-1.79
Deviation from the average travel speed	37.72
Deviation from the average acceleration	2.36

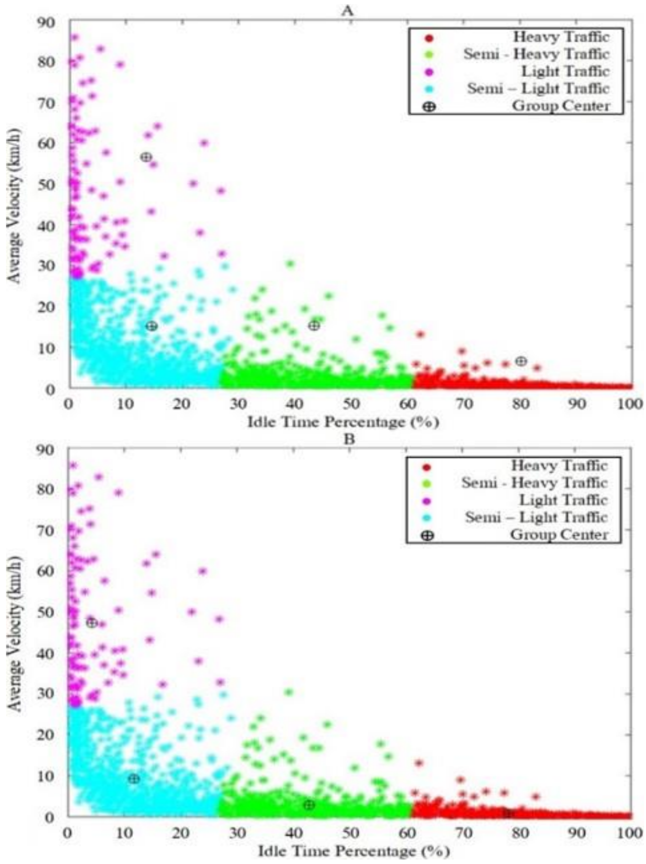

Fig. 1. Classification results of micro-trips using the K-means algorithm (Picture A: Mid-range and Picture B: Mean)

In this research, the K-means method was compared with the support vector machine. Note that the former is unsupervised while the latter is a supervised grouping algorithm. The advantages of the K-means method include its good accuracy and low computational volume [2]. The support vector machine algorithm, however, is one of the most widely used algorithms in solving nonlinear problems and identifying high-dimensional patterns. The mean speed and the driving duration of labeled data, which are used to train the support vector machine, are presented in Table 1.

3- Results

During 55.3 hours of data collection, 100,303 data points were recorded, of which 1658 micro-trips were obtained. The distance between the two stops, as stated in the literature [2], was considered as a micro-trip. It should be noted that the vehicle was considered as a stop if the vehicle speed and the acceleration were less than 0.1 km/h and 0.1 m/s2, respectively. The results of clustering from support vector machine and K-means algorithms are illustrated in Figs. 1 and 2. To create the cycle, mean and mid-range values were used. These two points are usually the best points, representative of the behavior of a bunch of data [5]. Therefore, if the micro-trips are relatively closer to these points in the production of cycles, the cycle will be more similar to the characteristics of total data. The similarity of the cycles with total data was also compared based on the mean relative error, the root-mean-square error, and the Chi-square test. By calculating the errors for the cycles, the one clustered by K-means, based on the mean value, illustrated the most similarity to total data. The parameters of this cycle are shown in Table 2.

4- Conclusions

Several driving cycles were estimated from speed-time data collected from a passenger vehicle in the Tehran-Amol route. The results revealed that,

• The K-means algorithm recorded better results than the support vector machine. This may be due to the effects that were defined in the training process.

• The supervised algorithms due to the training process, could classify data more accurately. But the generated driving cycle was not necessarily the best one.

• The cycle with the most similarity to total data was a very low speed cycle, which was due to the roughness of the path.

• The acceleration changes of this cycle (Tehran-Amol) were higher than other previously studied cycles that could be attributed to the Iranian driving culture.

• Investigating the average positive and negative acceleration values, as well as the deviation from the average travel speed and the acceleration, it is clear that the cycle was significantly different than other cycles.

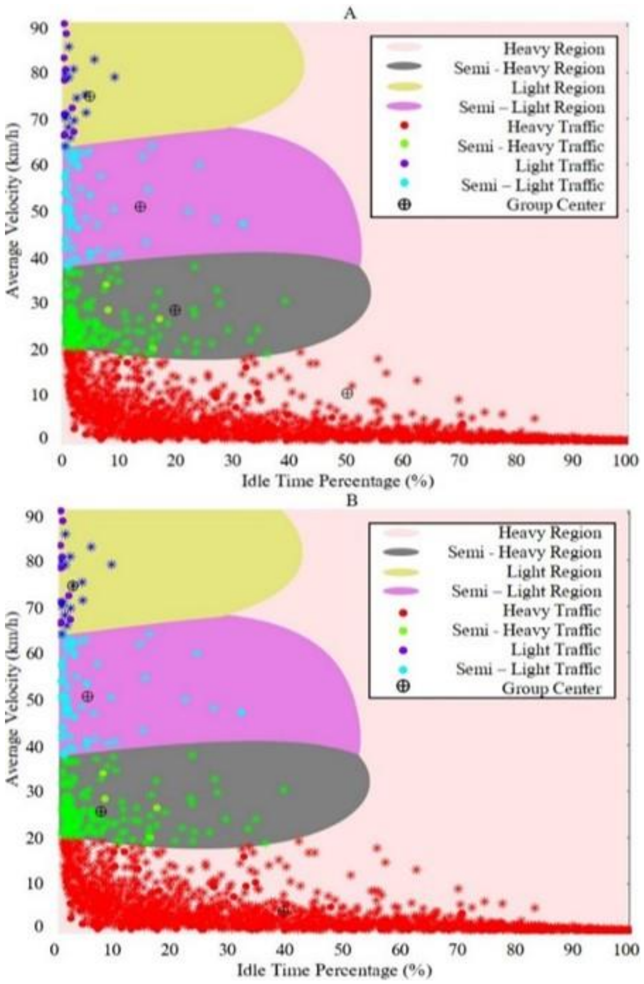

Fig. 2. Classification results of micro-trips using the support vector machine (Picture A: Mid-range and Picture B: Mean)

References

[1] X. Zhao, Q. Yu, J. Ma, Y. Wu, M. Yu, Y. Ye, Development of a representative EV urban driving cycle based on a K-means and SVM hybrid clustering algorithm, Journal of Advanced Transportation, 2018 (2018) 1890753.

[2] A. Fotouhi, M. Montazeri-Gh, Tehran driving cycle development using the K-means clustering method, Scientia Iranica, 20 (2013) 286-293.

[3] I.N. Anida, A.R. Salisa Driving cycle development for Kuala Terengganu city using K-means method, International Journal of Electrical and Computer Engineering, 9 (2019) 1780-1787.

[4] Y. Nguyen, N.D. Bui, T.-D. Nghiem, A. Le, GPS data processing for driving cycle development in Hanoi, Vietnam, Journal of Engineering Science and Technology, 15 (2020) 1429-1440.

[5] Median vs. average to describe normal, Natural Resources Conservation Service, United States Department of Agriculture, 2019, https://www.wcc.nrcs.usda.gov/.

HOW TO CITE THIS ARTICLE
T. Qaraati, A. Momenimovahed, M. Azadi, S. A. Moosavian, Comparison of Support Vector Machine and K-Means Algorithms Performance in Extracting the Real Driving Cycle of Combined Tehran-Amol, Amirkabir J. Mech Eng., 53(9) (2021) 1177-1180.

DOI: 10.22060/mej.2021.19222.6980