Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 2 No. 1 (2023) pp: 12-17
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Driver Predicting Behavior Based on Accelerometer and Gyroscope Sensors with K-Means Algorithm Method

Michael Sitorus[1], Melisa[2], Desty Herninda[3]
[1,2,3]Information System Departement, BRI Institute
[1]michael.sitorus@bri-institute.ac.id, [2]melisa260202@gmail.com, [3]destyaherninda36908@gmail.com

Abstract

The transportation with an increasing population level in Indonesia will also affect the smoothness of traffic. However, with the existence of transportation that is owned by someone, it is often misused in driving. Aggressive driving behavior is a major factor in traffic accidents. As reported AAA Foundation for Traffic Safety, 106,727 dangerous accidents – 55.7 percent of the over past four year period involved drivers engaging in one or more acts of aggressive driving. Therefore, to predict driver behavior, the K-Means algorithm used the Orange Application. Using the driver behavior dataset that has been recorded with Accelerometer and Gyroscope Sensor-based applications, the results obtained where the k value is the number of cluster 1 there are 814 items, cluster 2 has 997 items, cluster 3 has 1273 items. From this test, it was found that proportion of predicted driver behavior with Aggressive qualifications was 96.4%, Normal was 92.2% and from driver behavior with Low qualifications was 74.3%. The accuracy rate of research using the K-Means algorithm to determine predictions is 0.861 or 86.1%. The results of predictions can help prevent accidents or other risks that will occur in the future.

Keywords: Accelerometer, Gyroscope, Sensor, K-Means Algorithm, Predictions

1. Introduction

In the 4.0 era, there have been several modern transportations that have sprung up in Indonesia. The more transportation with an increasing population level in Indonesia will also affect the smoothness of traffic. However, with the existence of transportation that is owned by someone, it is often misused in driving. As we see and know, everyone has a different driving behavior.

Driver behavior can be categorized according to several types of speed, namely aggressive, normal, slow. Where all three have their respective advantages and disadvantages. Many people may think that people who like to drive at high speed are dangerous. [1]. Meanwhile, driving at high speed also increases the risks that could happen to motorists on the road. Such as the risk of crime, accidents, robbery, to get a ticket. We can actually minimize this risk.

Aggressive driving behavior is a major factor in traffic accidents. As reported by the AAA Foundation for Traffic Safety, 106,727 fatal accidents – 55.7 percent of the total – over the past four year period involved drivers engaging in one or more acts of aggressive driving [2]. This study was conducted over four years, and the most common aggressive acts were:

1. Erratic or reckless driving
2. Not obeying traffic signs or signs
3. Illegal turns
4. Speed up
5. Failure to grant the right of way
6. Illegal passing
7. Tailgating

Aggressive driving also causes countless no dangerious accidents that often result in serious injury. It's best to stay calm behind the wheel.

The problems raised have been researched with the title “Driver Behavior Scoring Application Using Fusion Sensors” [3]. In this study. Researchers reiterate the problem of driver behavior that can increase mortality. By predicting the behavior of drivers using applications based on Accelerometer and Gyroscope sensors with the K-Means Algorithm Method.

Traffic issues is a vital thing that needs to be studied on it [4]. With the existence of sophisticated technology in the current era, it makes it easier for humans to be able to record driver behavior data while driving through an application based on Accelerometer and Gyroscope sensors. The data will later help in predicting driver behavior in driving. The results of these predictions can prevent accidents or other risks that will occur in the future.

Received: 14-12-2022 | Accepted: 03-05-2023 | Published: 20-05-2023

2. Research Methods

The characteristics of drivers are important in understanding their behavior on the road. These characteristics encompass various factors, including the driver's natural abilities, learning capacity, and motives. By examining a driver's motives and attitudes, it becomes possible to gain insights into why they behave the way they do behind the wheel.

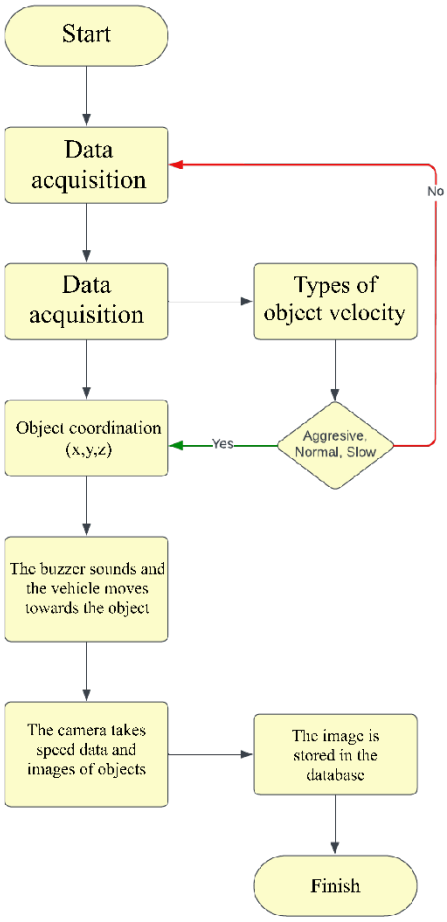


Figure 1. Driver Behavior Recorder Application System Work

For example, a driver who is motivated by fear of accidents or a sense of social responsibility may react differently in certain driving situations than one who is not. It is important to note, however, that driver characteristics can be impacted by a range of factors. The use of drugs or alcohol, as well as pain or fatigue, can significantly reduce a driver's effectiveness and alter their behavior on the road. Overall, understanding driver characteristics and how they influence behavior can help to improve road safety and reduce accidents.

The basis of the accelerometer is Newton's Second Law which says that the acceleration experienced by an object is directly proportional to the force and inversely proportional to the mass of the object acting on it. Accelerometer is a device used to measure acceleration and the effect of gravity on acceleration [5]. When the speed of an object changes, the object accelerates. Acceleration or what is often called acceleration is the rate of change of velocity with respect to time. Just like speed, acceleration is also divided into two, namely average acceleration and instantaneous. Acceleration is the result of the derivative of velocity with respect to time and velocity is obtained from the derivative of displacement with time.

Meanwhile, A gyroscope is a navigation tool that can used to measure the orientation of the device. Manner The work of the gyroscope is to measure angular velocity and how quickly it surrounds an axis [6]. Accelerometer and gyroscope sensors are sensors that are often used to determine the type of movement [7] and anggular position [8]. The workings of this system can be seen in Figure 1.

3. Results and Discussions

Research methods are basically a scientific way to obtain data with a specific purpose and usefulness. The scientific way means research activities that is based on scientific features, namely rational, empirical, and systematic. Rational means research activities carried out in a reasonable way, so that it is affordable by human reasoning. Empirical means the ways that done it can be observed by the human senses, so that others can observe and know the means used. Systematic means a process that used in research that uses certain steps that is logical [9].

Methods are steps or ways to achieve a certain goal. The research method can be explained as a way to achieve research objectives, in this case the steps are carried out with a scientific approach (rational, empirical and systematic). In general, research methods are classified into several types, including qualitative, quantitative, and combination. So based on the description of the theory above, it can be concluded that the research method is sharing ways, steps, and models for conducting a study. The method used by the author as a means and guide in this study is to use the K-Means Algorithm.

K-Means is one of the attempted non-hierarchical data clustering methods partition the existing data into the form of one or more clusters or groups so that the data which have equally detrimental characteristics in the same cluster and data which have different characteristics from other groups [10].

A well-known algorithm in the clustering method is K-means because its fairly simple algorithm and efficient[11]. The K-means algorithm uses a 14 iterative process to get a cluster database. It takes the desired number of initial clusters as input and produces the final centroid point as output.

DOI: https://doi.org/10.31004/riggs.v2i1.29

The K-means method will choose pattern k as the starting point of the centroid randomly. The number of iterations to reach the cluster centroid will be influenced by the initial cluster centroid candidates randomly. So that we get a way of developing the algorithm by determining the centroid cluster which is seen from the high initial data density in order to get higher performance.

K-Means Clustering is a method in data mining, which is used to perform grouping large amounts of data. Where from each data is in one The group, has characteristics that same or different characteristics with the data in other groups [12]. In its completion, the K-Means algorithm will produce a centroid point which is the goal of the K-Means algorithm. After the K-Means iteration stops, each object in the dataset becomes a member of a cluster. The cluster value is determined by searching all objects to find the cluster with the closest distance to the object. The K-means algorithm will group data items in a dataset into a cluster based on the shortest distance [13].

The initial centroid value chosen randomly which is the starting point, will be calculated the distance to all data using the Euclidean Distance formula. Data that has a short distance to the centroid will form a cluster. This process continues until there is no change in each of the 15 groups [14].

Steps in conducting clustering using the K means algorithm as follows [15]:

1. Specify the Number of clusters k.
2. Initialize the cluster center K. It can done in different ways. But the most numerous ways used is with determine randomly. Center of the cluster is assigned an initial value with random numbers.
3. Allocate all data/objects to in the nearest cluster. Closeness Two data/objects can be obtained based on the distance of the two objects aforementioned. To be able to calculate the distance of all data to the center point cluster then can use the Euclidean distance theory formulated below:
   D(ij)
   $= \sqrt{(X1i - Xij)^2 + (X2i - X2j)^2} + \cdots + (Xki - Xkj)^2$
   where:
   D(i,j) = Distance of data to i to center Cluster J
   Xki = Data to i attribute data to k
   Xkj = Center point to j on attribute to k
4. Recalculate cluster center with cluster membership that now. The cluster center is the average of all objects/data in specific clusters. If desired could also use the median of the cluster. So the average (mean) is not the only measure wearable.
   $Rk = \frac{1}{Nk}\ (X1k + X2k + \cdots + Xnk)$
   where:
   Rk = New average.
   Nk = Number of training patterns on the cluster (k).
   Xnk = Pattern to (n) that becomes cluster section (k).
5. Reassign each object using the center of the new cluster. If the center cluster does not change anymore then The clustering process is complete. Or Go back to step number 3.

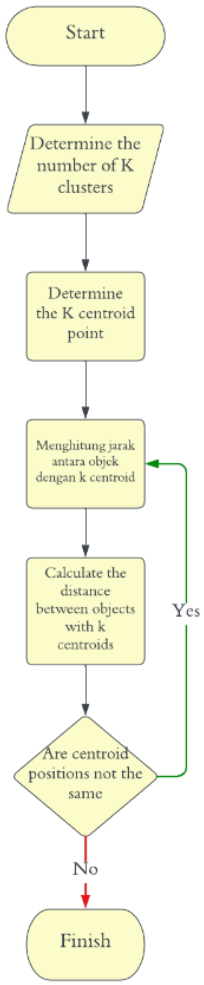


Figure 2. Flowchart K-Means Algorithm

The Research Process on Driver Behavior with the K-Means Algorithm is carried out using the Orange application. In Figure 3 is a Research Scheme using Orange.

DOI: https://doi.org/10.31004/riggs.v2i1.29

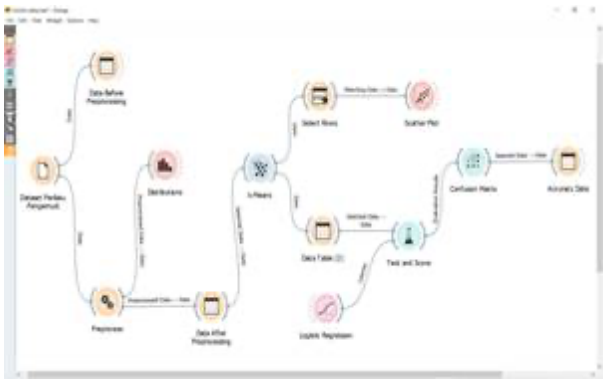
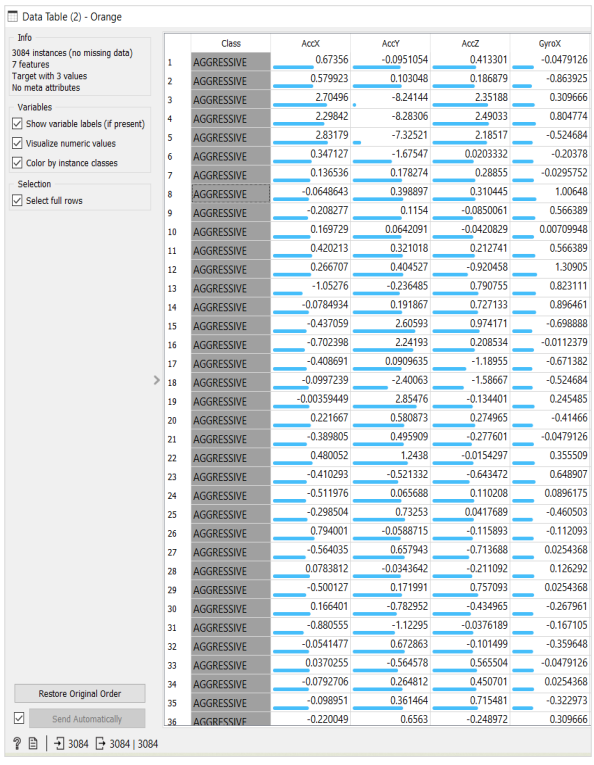


Figure 3. Driver Behavior Research Scheme Using the Orange Application

4. Conclusion

4.1. Research Dataset

This study uses a driver behavior dataset by taking 2 data samples per second. This data includes Speed (x, y, z axis in meters per second squared (m/s2), Rotation (X, Y, Z axis in degrees per second (°/s)), Classification label (Aggressive, Normal, Slow), Timestamp (time in seconds) All existing data is the result of an application recorder designed on the basis of accelerometer and gyroscope sensors.



Data Table (2) - Orange

Info
3084 instances (no missing data)
7 features
Target with 3 values
No meta attributes

Variables
Show variable labels (if present)
Visualize numeric values
Color by instance classes

Selection
Select full rows

Restore Original Order
Send Automatically

3084 | 3084 | 3084

	Class	AccX	AccY	AccZ	GyroX
1	AGGRESSIVE	0.67356	-0.0951054	0.413301	-0.0479126
2	AGGRESSIVE	0.579923	0.103048	0.186879	-0.863925
3	AGGRESSIVE	2.70496	-8.24144	2.35188	0.309666
4	AGGRESSIVE	2.29842	-8.28306	2.49033	0.804774
5	AGGRESSIVE	2.83179	-7.32521	2.18517	-0.524684
6	AGGRESSIVE	0.347127	-1.67547	0.0203332	-0.20378
7	AGGRESSIVE	0.136536	0.178274	0.28855	-0.0295752
8	AGGRESSIVE	-0.0648643	0.398897	0.310445	1.00648
9	AGGRESSIVE	-0.208277	0.1154	-0.0850061	0.566389
10	AGGRESSIVE	0.169729	0.0642091	-0.0420829	0.00709948
11	AGGRESSIVE	0.420213	0.321018	0.212741	0.566389
12	AGGRESSIVE	0.266707	0.404527	-0.920458	1.30905
13	AGGRESSIVE	-1.05276	-0.236485	0.790755	0.823111
14	AGGRESSIVE	-0.0784934	0.191867	0.727133	0.896461
15	AGGRESSIVE	-0.437059	2.60593	0.974171	-0.698888
16	AGGRESSIVE	-0.702398	2.24193	0.208534	-0.0112379
17	AGGRESSIVE	-0.408691	0.0909635	-1.18955	-0.671382
18	AGGRESSIVE	-0.0997239	-2.40063	-1.58667	-0.524684
19	AGGRESSIVE	-0.00359449	2.85476	-0.134401	0.245485
20	AGGRESSIVE	0.221667	0.580873	0.274965	-0.41466
21	AGGRESSIVE	-0.389805	0.495909	-0.277601	-0.0479126
22	AGGRESSIVE	0.480052	1.2438	-0.0154297	0.355509
23	AGGRESSIVE	-0.410293	-0.521332	-0.643472	0.648907
24	AGGRESSIVE	-0.511976	0.065688	0.110208	0.0896175
25	AGGRESSIVE	-0.298504	0.73253	0.0417689	-0.460503
26	AGGRESSIVE	0.794001	-0.0588715	-0.115893	-0.112093
27	AGGRESSIVE	-0.564035	0.657943	-0.713688	0.0254368
28	AGGRESSIVE	0.0783812	-0.0343642	-0.211092	0.126292
29	AGGRESSIVE	-0.500127	0.171991	0.757093	0.0254368
30	AGGRESSIVE	0.166401	-0.782952	-0.434965	-0.267961
31	AGGRESSIVE	-0.880555	-1.12295	-0.0376189	-0.167105
32	AGGRESSIVE	-0.0541477	0.672863	-0.101499	-0.359648
33	AGGRESSIVE	0.0370255	-0.564578	0.565504	-0.0479126
34	AGGRESSIVE	-0.0792706	0.264812	0.450701	0.0254368
35	AGGRESSIVE	-0.098951	0.361464	0.715481	-0.322973
36	AGGRESSIVE	-0.220049	0.6563	-0.248972	0.309666

Figure 3. Driver Behavior Dataset

4.2. Process

At this stage, the data that has been input is then clustered to determine the prediction results with the classification or category of Aggressive, Normal, Slow driver behavior. Before clustering the data will be entered into the preprocessing after getting the results of the data table that has been processed, then it can then be entered into a bar chart or distribution.

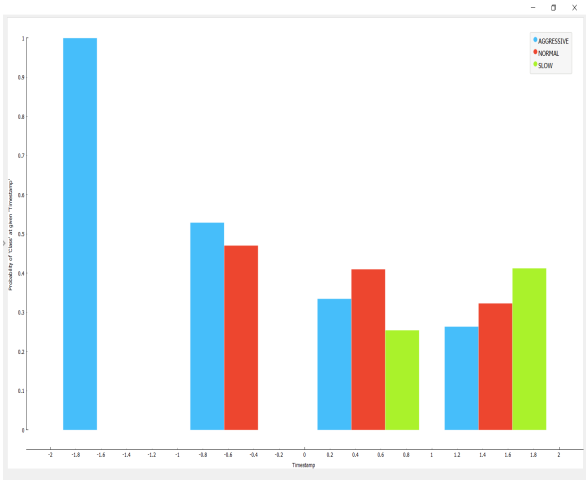


Figure 4. Bar Chart of Driver Behavior Classification

4.3. Results

The clustering results from the preprocessing and processing stages can then be processed with the K-Means Algorithm. At the processing stage of the K-Means Algorithm, 3 clusters are taken. To get the results of the Confusion Matrix, testing and scoring of the data that has been processed is first carried out. The test results and scoring use the results of the Logical Regression test which can be seen in Figure 7.

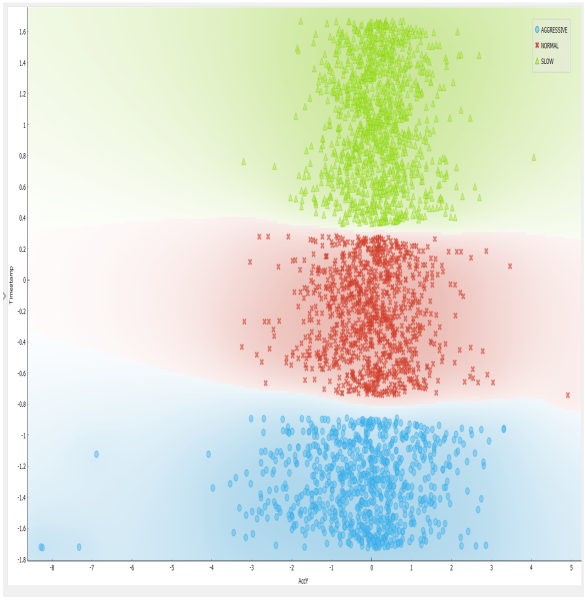


Figure 6. Scatter Plot Results Using the K-Means Algorithm

DOI: https://doi.org/10.31004/riggs.v2i1.29

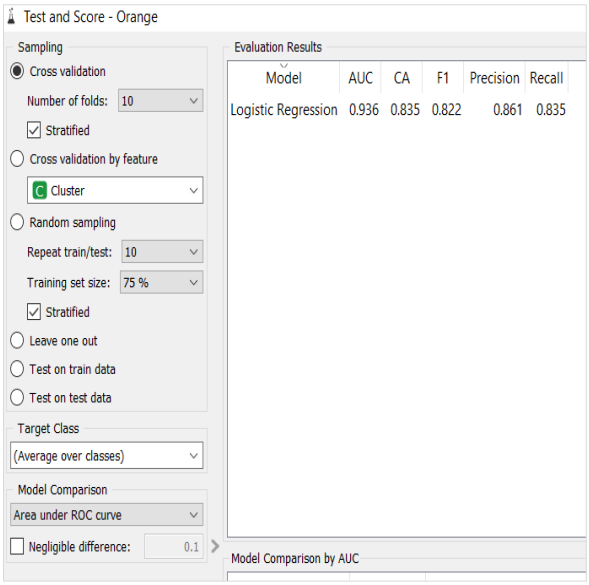


Figure 7. Test Results and Scoring

Where in Figure 8 and Figure 9, 3 clusters are taken according to the K-Means method where the k value is the number of cluster 1 there are 814 items, cluster 2 has 997 items, cluster 3 has 1273 items.

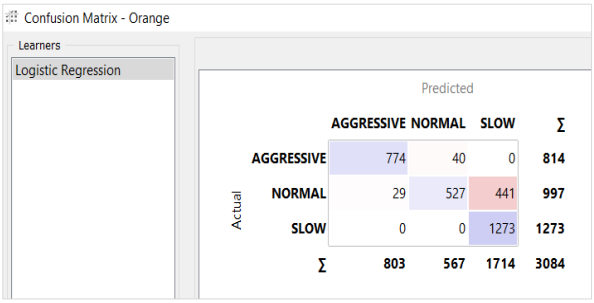


Figure 8. Confusion Matrix Results

From the figure below it can be seen that the total data is 3084 items. From this test, it was found that the proportion of predicted driver behavior with the Aggressive classification was 96.4%, Normal was 92.2% and from driver behavior with Low classification was 74.3%.

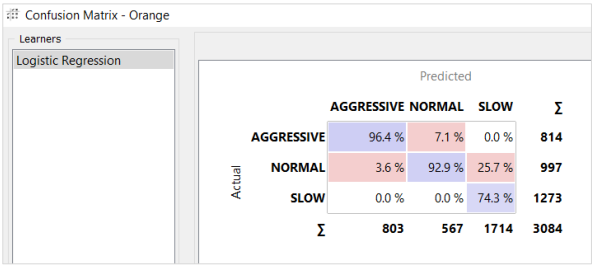


Figure 9. Percentage Rate Proportion of Predictions

The accuracy rate of research using the K-Means algorithm to determine predictions of driver behavior is 0.861 or 86.1%.

4. Conclusion

The conclusion, the research has produced several important findings. Firstly, the driver behavior data can be divided into three distinct clusters: Aggressive, Normal, and Low. Secondly, testing conducted using the Orange application resulted in an accuracy level of 86.1%, based on a recall value of 83.5% and a precision score of 86.1%. However, it is worth noting that while the system is capable of predicting driver behavior, it is not always precise. This lack of precision may be attributed to the values generated by the sensors in the data recording application. Overall, these findings suggest that technology can be used to gain insights into driver behavior, but further research is needed to improve the accuracy of these predictions

Reference

[1] R. Nugraha, "Menurut Studi, Orang yang Mengemudi Lambat Itu Juga Berbahaya dan Lebih Banyak Sebabkan Kematian," *Hai Grid*, Jan. 11, 2020. https://hai.grid.id/read/071983451/menurut-studi-orang-yang-mengemudi-lambat-itu-juga-berbahaya-dan-lebih-banyak-sebabkan-kematian?page=all (accessed Jan. 10, 2023).

[2] I. Cojocaru and S. Popescu, "Driving Behavior," Craiova, Dolj, Rumania, May 2022. Accessed: Jan. 10, 2023. [Online]. Available: https://www.kaggle.com/datasets/outofskills/driving-behavior

[3] A. Larasati, R. V. H. Ginardi, and Sarwosri, "Aplikasi Scoring Perilaku Pengemudi Menggunakan Sensor Fusi," *Jurnal Teknik ITS*, vol. 7, no. 2, pp. 1–7, Aug. 2018.

[4] Y. Oktopianto, S. Shofiah, F. A. Rokhman, K. P. Wijayanthi, and E. Krisdayanti, "Analisis DaerahRawan Kecelakaan(Black Site)Dan Titik Rawan Kecelakaan (Black Spot)Provinsi Lampung," *Borneo Engineering: Jurnal Teknik Sipil*, vol. 5, no. 1, pp. 40–51, Apr. 2021.

[5] M. Abadi and A. Saleh, "1Rancang Bangun Alat Pengukur Langkah Kaki dengan Sensor Accelerometer dan Fasilitas Komunikasi Wireless 2,4 GHz," *Center for Open Science*, pp. 1–6, Aug. 2021.

[6] Immersa Lab, "Pengertian Gyrscope Dan Cara Kerjanya," *Immersa Lab*, Feb. 07, 2018. https://www.immersalab.com/pengertian-gyroscope-dan-cara-kerjanya.htm/ (accessed Jan. 10, 2023).

DOI: https://doi.org/10.31004/riggs.v2i1.29

[7] H. Achmad and Wahyudi, "Rancang Bangun Inertial Measurement Unit Sebagai Sistem Monitoring Kendaraan Bergerak Berbasis Sensor Accelerometer dan Gyroscope," *Jurnal Rekayasa Elektrika*, vol. 9, no. 4, pp. 187–194, Oct. 2011.

[8] T. Perkasa and H. H. Rachmat, "Perancangan Alat Ukur Sudut Tekuk Lutut Wireless menggunakan Sensor Gyroscope berbasis ATMega 328 dan ATMega 2560," *Jurnal ELKOMIKA*, vol. 5, no. 1, pp. 30–47, Jun. 2017.

[9] Sugiyono, *Metode Penelitian Kualitatif, Kualitatif dan R&D*, 28th ed. Bandung: CV. Alfabeta, 2018.

[10] M. S. Zain, "Data Mining Dengan Metode K-Means Untuk Pengelompokan Mahasiswa Yang Mengunjungi Perpustakaan Berdasarkan Data Kunjungan Dan IPK," *UPI YPTK Jurnal EKOBISTEK*, pp. 1–16, Jan. 2018.

[11] T. H. Sardar and Z. Ansari, "An analysis of MapReduce efficiency in document clustering using parallel K-means algorithm," *Future Computing and Informatics Journal*, vol. 3, no. 2, pp. 200–209, May 2018, Accessed: Jan. 10, 2023. [Online]. Available: https://digitalcommons.aaru.edu.jo/fcij/

[12] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," *ReadyStar*, vol. 2, no. 1, pp. 394–403, Nov. 2019.

[13] B. Bangoria, N. Mankad, and V. Pambhar, "A Survey on Efficient Enhanced K-Means Clustering Algorithm," *International Journal for Scientific Research & Development (IJSRD)*, vol. 1, no. 9, Dec. 2013, Accessed: Dec. 21, 2022. [Online]. Available: https://ijsrd.com/Article.php?manuscript=IJSRDV1I9016

[14] A. Agrawal and H. Gupta, "Global K-Means (GKM) Clustering Algorithm: A Survey," *Int J Comput Appl*, vol. 79, no. 2, pp. 20–24, Oct. 2013.

[15] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan RapidMiner Dengan K-Means Clusteer Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi," *CESS (Journal of Computer Engineering System and Science)*, vol. 3, no. 2, pp. 173–178, Jul. 2018.

DOI: https://doi.org/10.31004/riggs.v2i1.29