Jurnal Sosial dan Teknologi (SOSTECH)
Volume 4, Number 7, *Juli* 2024
p-ISSN **2774-5147** ; e-ISSN **2774-5155**

Optimizing KNN Algorithm Using Elbow Method for Predicting Voter Participation Using Fixed Voter List Data (DPT)

Ikmal Maulana[1*], Rusdianto Roestam[2]
[1,2] President University, Indonesia
Email: ikmalmaulana@gmail.com

Abstract

The purpose of this study is to produce maximum predictions for election participation rates. KNN (K-Nearest Neighbors) is one of the machine learning algorithms used to classify or regress data. The KNN algorithm works by finding the closest K training data from test data to be classified. Although the KNN (K-Nearest Neighbors) algorithm has advantages such as being easy to implement and being able to handle non-linear data, this algorithm also has several weaknesses, one of which is the determination of the value of K which is very ordinary and subjective. Therefore in this study optimization of the value of K on KNN using the Elbow method. The dataset used is the Fixed Voters List (DPT) in the 2019 General Elections in Karawang Regency. The final results of the experiments in this study, the highest achievement was obtained with a Mean Squared Error (MSE) value of 0.0018, a Root Mean Squared Error (RMSE) value of 0.0422, and a Mean Absolute Percent Error (MAPE) value of 6.36%. The highest accuracy produced in this study was 95.63% and the lowest was 93.64% with an average accuracy of 95.02%.

***Keywords:** K-Nearest Neighbor (KNN), Elbow, General Election, Voter List*

INTRODUCTION

General Elections are held nationally throughout Indonesia at the same time. PelectionIn general, the idea of democracy was born referring to John Locke and Rousseau, where the purpose of holding elections is to guarantee justice, equality, and freedom for every citizen in a country (Rosidin, Huda, and Burhanuddin 2021); (Syakur et al. 2018) One of the stages in the election is the compilation and updating of voter data. This stage is very crucial because it relates to the rights of citizens that are protected by law. The source of voter data in the General Election comes from the Population and Potential Voter Data (DP4) which comes from the Minister of Home Affairs (Kharima and Ihsan 2022); (Talakua 2023). Regency/Municipal KPU makes the DP4 synchronized with the Election Final Voters List (DPT) which is updated on an ongoing basis from the previous election.

The Fixed Voters List (DPT) is compiled based on the distribution of voters at the Village/Kelurahan level which is distributed to each Polling Place (TPS). In the Final Voter List (DPT) there are 9 data elements, namely Name, NIK, NKK, Gender, Marital Status, Address, Type of Disability, Place of Birth, and Date of Birth.

To measure the level of people's participation in the General Election, we compare the total number of voters in the Permanent Voter List (DPT) with the number of voters present in each TPS obtained from the DAA1 form. The objective of this research is to create a predictive model for public participation in each TPS concerning the Final Voter List (DPT) that will be determined so that before the DPT is determined, an evaluation of the preparation of the DPT which is oriented toward increasing community participation can be carried out. The results of predictions on the model can be used as a reference in preparing programs to increase community participation. The purpose of this study is to

e-ISSN 2774-5155
p-ISSN 2774-5147

produce maximum predictions for election participation rates.

In increasing the level of accuracy in an algorithm, an optimization technique is used (Riyan and Nendi 2024). One of the optimization techniques that can be used is the Elbow Method. Some previous studies that used the Elbow Method as an optimization method are as follows (Dewi and Pramita 2019) stated that the results of k-medoid clustering with silhouette coefficient produce better cluster quality because it has a lower DBI value than k-medoid clustering with the elbow method. Winarta and Kurniawan (2021) The test results show that the Elbow method works very well in producing an optimal cluster, which is found at k = 3 with an SSE difference value of 1257, 862 with k test = 5. Furthermore, research from Hartanti (2020) stated that the application of the Elbow method in accordance with this study resulted in the best number of clusters was 3, then by applying the K-Means method data was produced that the number of students was 66 students, with a calculation process of 9 iterations resulting in 3 clusters, namely the "Ready", "Quite Ready", and "Not Ready" categories, with the results of the "Ready" cluster as many as 7 students, the "Quite Ready" cluster as many as 30 students, and the "Not Ready" cluster as many as 29 students. From the three studies, the Elbow Method is used for the optimization of the K-Means algorithm by producing the optimal number of clusters (Umargono, Suseno, and Gunawan 2019).

RESEARCH METHODS

.

In this study, the model created by the author uses one of the models in machine learning, namely K-Nearest Neighbor (KNN). To perfect the model in the KNN algorithm, improvisation was carried out on the K parameter using the elbow method.

K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a machine learning algorithm used in classification and regression. The main principle behind KNN is that similar data tends to reside in the same environment. Therefore, in KNN, the class or regression value of a data is determined by the majority of the class or regression values of its k nearest neighbors. One of the advantages of KNN is that it is strong in an effective and simplest training process compared to other machine learning algorithms (Jadhav and Channe 2016); (Gorade, Deo, and Purohit 2017). One of the weaknesses of KNN is the number of nearest neighbors (parameter K) which is biased and subjective. The K value describes several nearest neighbors that are used for the classification process (Wahyudi, Sulthan, and Suhartini 2021); (Hidayat, Purnomo, and Yudhanto 2022). The results of the resulting classification accuracy and quality are strongly influenced by the determination of the K value parameter (Simanjuntak, Sutrisno, and Mahmudy 2014); (Alaydrus and Soebroto 2023).

Here's how the steps in KNN work:

1. Determine the parameter K (number of nearest neighbors)
2. Calculate the distance between test data and training data.
3. Determine the size of the short distance at the value of K
4. Concatenates several appropriate class labels
5. Determine the class according to the majority criteria

In the KNN (K-Nearest Neighbors) algorithm, the distance between the new data and the training data is used to determine the nearest neighbor of the new data. There are several methods that can be used to calculate the distance between new data and training data, namely Euclidean distance, Manhattan distance, and Minkowski distance. In this research, the writer will use the Euclidean distance

e-ISSN 2774-5155
p-ISSN 2774-5147

to calculate the distance. This method calculates the distance between two points in space using the following formula:

$$dxy = \sqrt{\sum (x_i - y_i n\ i = 1)^2}$$

Information:

d	=	Distance value
X	=	Test data
Y	=	Training data

Elbows

To solve the problem of ambiguity in determining the value of K in KNN, the authors use the Elbow method to determine the optimal K parameter. The elbow method is a technique in cluster analysis that is used to determine the optimal number of clusters in data. This technique is based on an elbow curve which shows the relationship between the number of clusters and the inertia value (sum squared distance) of the data in these clusters. The results of these calculations will be visualized in the form of an elbow curve graph, where the x-axis shows the number of clusters and the y-axis shows the value of inertia. In the elbow curve graph, it can be seen that the value of inertia tends to decrease when the number of clusters is increased, but at a certain point, the decrease in inertia value begins to slow down and forms an elbow on the curve (elbow).

The Elbow method helps interpret and validate consistency in cluster analysis and select the optimal number of clusters by adjusting the model to a range of K values (Khalid and Wade 2020). The way to determine the comparison is by knowing the SSE (Sum of Squared Error) value of each class/group (Syakur et al. 2018). The more dominant the number of K in the class automatically the lower the SSE value (Irwanto, Purwananto, & Soelaiman, 2012); (Umargono et al. 2019).

The following is the equation to find the SSE value :

$$SSE = \Sigma\ (xi - ci)2$$

Information:

xi	=	The data point in a cluster
xi	=	cluster center for the cluster

Here are the steps on Elbow [14]

1) Determine the initial parameters K
2) Increase K parameters
3) Calculating the sum of square error (SSE)
4) Analyze SSE which has decreased
5) Determine the parameter (value) K formed by the right angle.

e-ISSN 2774-5155
p-ISSN 2774-5147

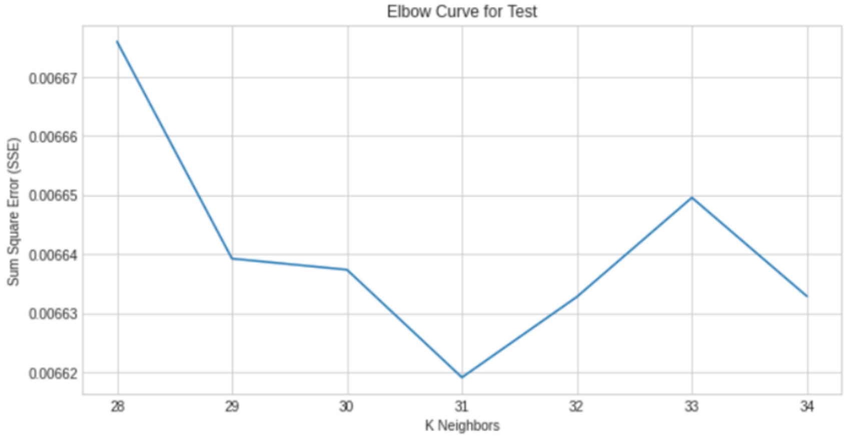

Figure 1. Elbow Curve

Evaluation of Model Testing

Evaluation of model testing in this study uses Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage error (MAPE).

Mean Squared Error (MSE)is a measure of the deviation or error between the actual value and the predicted value of data. MSE is calculated by calculating the difference between the predicted value and the actual value, then increasing the difference twice, adding up all these values, then dividing by the total amount of data (Montgomery, Peck, and Vining 2021). Here's the MSE equation:

$$MSE = \frac{\sum_{t=1}^{n}(At - Ft)^2}{n}$$

At	=	Actual value
Ft	=	Predictive Value
n	=	The amount of data

Root Mean Squared Error (MSE)is a measure of the deviation or error between the actual value and the predicted value of data. RMSE is the squared form of MSE. RMSE is calculated by calculating the difference between the predicted value and the actual value, then raising the difference twice, adding up all these values, dividing by the total amount of data, and taking the square root of the result (Montgomery et al. 2021); (Gunst and Mason 2018). Here's the RMSE equation:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}}$$

At	=	Actual value
Ft	=	Predictive Value
n	=	The amount of data

Mean Absolute Percentage error (MAPE)is a measure of the deviation or error between the actual value and the predicted value in the form of a percentage. MAPE is calculated by calculating the absolute difference between the predicted value and the actual value, dividing it by the actual value, and then taking the average of all these values. Then, the result is multiplied by 100% to get the result in percentage form (Montgomery et al. 2021). The following is the MAPE equation:

$$MAPE = \frac{\sum_{t=1}^{n}\left|\left(\frac{A_t - F_t}{A_t}\right)100\right|}{n}$$

e-ISSN 2774-5155
p-ISSN 2774-5147

At	=	Actual value
Ft	=	Predictive Value
n	=	The amount of data

Of the three test methods, in principle the lower the error value obtained, the better the quality of the model built. After MSE, RMSE, and MAPE are calculated, the quality of the accuracy of the predictions produced can be seen. The accuracy value is obtained from a 100% reduction minus the MAPE value.

RESULTS AND DISCUSSION

The first step in this research is to prepare the dataset. The dataset is obtained from the results of synchronizing the Final Voter List (DPT) and the DAA1 form. The number of datasets used is 2,741 with 14 attributes.

Table 1. Dataset

No	Attribute Name	Information	Information
1	Number of Voters	Number of voters in TPS	Independent
2	Male Voter (L)	Number of male voters	Independent
3	Female voter (F)	Number of female voters	Independent
4	Marital status (SM)	Number of voters who are married	Independent
5	Marital status (BM)	Number of unmarried voters	Independent
6	Marital status (PM)	Number of voters who have ever been married (widows/widowers)	Independent
7	Late adolescent age (UR)	Number of voters aged 17-25 years	Independent
8	Early adulthood (UD1)	Number of voters aged 26-35 years	Independent
9	Late adulthood (UD2)	Number of voters aged 36-45 years	Independent
10	Early elderly (UL1)	Number of voters in the late elderly age group (46-55 years)	Independent
11	Late elderly (UL2)	Number of voters in the late elderly age group (55-65 years)	Independent
12	Seniors (UM)	Number of voters in the senior age group (> 55 years)	Independent
13	Disabled voters (DIS)	Number of disabled voters	Independent
14	Attendance presentation (PARMAS)	Voter turnout presentation	dependent

From these two data sources, then the normalization process is carried out. This is to make the range of values the same without breaking the headline on the value of each attribute. Following are the results of the final dataset used.

e-ISSN 2774-5155
p-ISSN 2774-5147

```
row_data
```

	JP	L	P	SM	BM	PM	UR	UD1	UD2	UL1	UL2	UM	DIS	PARMAS
0	0.89	0.81	0.71	0.67	0.68	0.43	0.56	0.36	0.41	0.54	0.53	0.38	0.0	0.65
1	0.73	0.65	0.61	0.63	0.29	0.56	0.40	0.28	0.42	0.40	0.41	0.44	0.0	0.74
2	0.62	0.58	0.49	0.57	0.28	0.24	0.38	0.28	0.25	0.38	0.32	0.35	0.0	0.67
3	0.81	0.79	0.61	0.73	0.34	0.40	0.46	0.33	0.40	0.49	0.38	0.52	0.0	0.76
4	0.96	0.92	0.74	0.78	0.59	0.56	0.63	0.41	0.37	0.56	0.67	0.43	0.0	0.80
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2736	0.92	0.90	0.70	0.82	0.46	0.32	0.62	0.39	0.51	0.44	0.45	0.38	0.0	0.72
2737	0.87	0.82	0.69	0.66	0.59	0.59	0.61	0.33	0.53	0.42	0.40	0.37	0.0	0.74
2738	0.80	0.80	0.60	0.69	0.46	0.33	0.49	0.32	0.38	0.44	0.51	0.44	0.0	0.65
2739	0.42	0.42	0.32	0.38	0.23	0.13	0.33	0.12	0.22	0.23	0.24	0.25	0.0	0.76
2740	0.03	0.03	0.02	0.02	0.01	0.05	0.01	0.00	0.01	0.01	0.04	0.05	0.0	0.64

2741 rows × 14 columns

Figure 2. Display of the research dataset

Process of Training and Testing

The first step is to determine the ratio of test and training data with the following comparison:

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size = 0.20, random_state=0)
```

Figure 3 Comparison of the ratio of test and training data

From the picture above, in this experiment, the ratio used was 80% for training data and 20% for test data of the total dataset used. While the random_state function is to make the test data and training data consistent using the same data, in this test it is made 0 so that the resulting training data and test data values remain consistent.

Next, determine the optimal K parameter using the elbow. Due to the large number of datasets, the K range is made longer, between 1 and 100. So the graph appears below

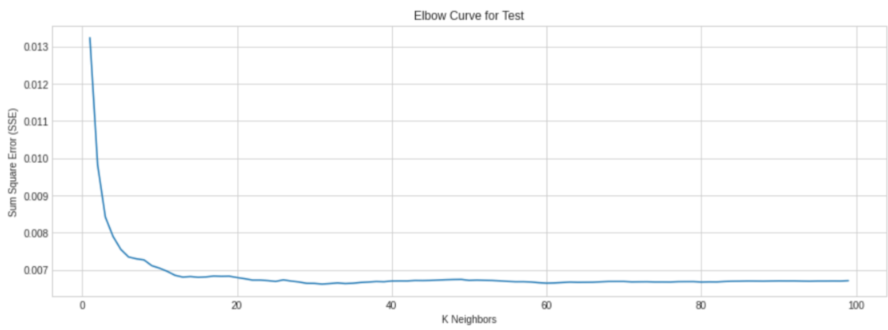

Figure 4. Curve with K range elbows (1-100)

From the resulting curve, it can be identified that there is a decrease in the error value between points 20 to 60. Therefore, reloading is carried out in a smaller range, namely 20 to 40 which is produced in the graph below.

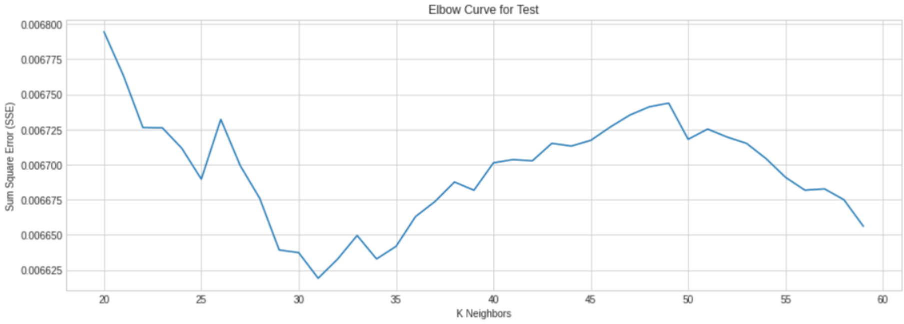

Figure 5. Curve with elbow range K (20-40)

From the second curve with a smaller range, it can be seen that there is a sharp decrease that forms an elbow, namely at the value of K = 31. Thus, the value of K=31 is predicted as optimal K.

```
[30] prediksi = y_predict[0:25]
     aktual = y_test [0:25]

     print("Nilai aktual :", aktual)
     print("")
     print("Nilai prediksi :", prediksi)

     Nilai aktual : [0.77 0.71 0.79 0.78 0.78 0.77 0.84 0.82 0.75 0.72 0.76 0.76 0.74 0.76
      0.76 0.67 0.72 0.87 0.72 0.73 0.78 0.75 0.78 0.82 0.79]

     Nilai prediksi : [0.75806452 0.75870968 0.78774194 0.7716129  0.79419355 0.78870968
      0.78516129 0.78516129 0.7583871  0.76645161 0.76677419 0.78419355
      0.79677419 0.76225806 0.77548387 0.75612903 0.78709677 0.78870968
      0.79483871 0.79645161 0.76709677 0.77516129 0.80612903 0.79290323
      0.79129032]
```

Figure 6. The Results of The Comparison of The Predicted Value and The Actual Value

From the figure, the prediction results are displayed with the y_predict variable while the actual data is described with the y_test variable

Model Testing

From the prediction results obtained, the model is tested and evaluated using the MSE (Mean Squared Error), (RMSE) Root Mean Square Error, and MAPE (Mean Squared Error) parameters. *Percentage error).* These three measurement methods in principle measure the amount of error in each prediction result. The smaller the resulting value, the better the accuracy obtained in the model. The following are the results of testing on the first try.

Table 2. Calculation of MSE, RMSE and MAPE

No	actual	predicted	Error	|Errors|	Error square	|%Errors|
1	0.77	0.75806452	0.01	0.01	0.00	1.55
2	0.71	0.75870968	-0.05	0.05	0.00	6.86
3	0.79	0.78774194	0.00	0.00	0.00	0.29
4	0.78	0.7716129	0.01	0.01	0.00	1.08
5	0.78	0.79419355	-0.01	0.01	0.00	1.82
6	0.77	0.78870968	-0.02	0.02	0.00	2.43
7	0.84	0.78516129	0.05	0.05	0.00	6.53
8	0.82	0.78516129	0.03	0.03	0.00	4.25
9	0.75	0.7583871	-0.01	0.01	0.00	1.12

e-ISSN 2774-5155
p-ISSN 2774-5147

No	actual	predicted	Error	|Errors|	Error square	|%Errors|
10	0.72	0.76645161	-0.05	0.05	0.00	6.45
11	0.76	0.76677419	-0.01	0.01	0.00	0.89
12	0.76	0.78419355	-0.02	0.02	0.00	3.18
13	0.74	0.79677419	-0.06	0.06	0.00	7.67
14	0.76	0.76225806	0.00	0.00	0.00	0.30
15	0.76	0.77548387	-0.02	0.02	0.00	2.04
16	0.67	0.75612903	-0.09	0.09	0.01	12.86
17	0.72	0.78709677	-0.07	0.07	0.00	9.32
18	0.87	0.78870968	0.08	0.08	0.01	9.34
19	0.72	0.79483871	-0.07	0.07	0.01	10.39
20	0.73	0.79645161	-0.07	0.07	0.00	9.10
21	0.78	0.76709677	0.01	0.01	0.00	1.65
22	0.75	0.77516129	-0.03	0.03	0.00	3.35
23	0.78	0.80612903	-0.03	0.03	0.00	3.35
24	0.82	0.79290323	0.03	0.03	0.00	3.30
25	0.79	0.79129032	0.00	0.00	0.00	0.16
		Total		**0.82**	**0.04**	**109**
			Mean Squared Error	**MAP**	**0.0018**	
			Mean Squared Error	**RMSE**	**0.0422**	
			Mean Absolute Percent Error	**MAPE**	**4.37**	%
			accuracy	**acc**	**95.63**	%

To confirm and compare the results of other tests, it is necessary to carry out more varied experiments in this study. In another experiment, adjustments were made to the split data ratio and the random_state value with the same dataset amounting to 2,741.

This follow-up experiment was carried out 12 times, plus the following is a recap of the test results from the 12 trials that were carried out.

Table 3. Comparison of experimental results

EXP	DATA	TEST	RAND	K	MAP	RMSE	MAPE	acc	RANK
1	50	0.2	0	2	0.0028	0.0524	6.36	93.64	12
2	2741	0.2	0	31	0.0018	0.0422	4.37	95.63	1
3	2741	0.1	0	30	0.0018	0.0424	4.42	95.58	3
4	2741	0.1	0	17	0.0021	0.0461	4.94	95.06	9
5	2741	0.1	0	142	0.0020	0.0446	4.62	95.38	7
6	2741	0.1	5	47	0.0019	0.0439	4.62	95.38	6
7	2741	0.1	5	28	0.0017	0.0418	4.38	95.62	2
8	2741	0.1	5	264	0.0020	0.0452	4.68	95.32	8
9	2741	0.2	0	54	0.0019	0.0436	4.61	95.39	5
10	2741	0.2	10	40	0.0037	0.0609	5.85	94.15	10
11	2741	0.3	0	31	0.0020	0.0444	4.61	95.39	4
12	2741	0.3	20	45	0.0038	0.0617	6.20	93.80	11
	AVERAGE				0.0023	0.0474	4.97	95.02	

e-ISSN 2774-5155
p-ISSN 2774-5147

To compare the accuracy of models without using elbows in determining K parameters, it is necessary to carry out comparative tests. Tests were carried out in experiments 2, 6, and 9 with K values that were not too far away from the previous tests

Table 4. K Value

EXP	DATA	TEST	RAND	K	MAP	RMSE	MAPE	acc
2	2741	0.2	0	30	0.0066	0.0814	7.4261	92.57
6	2741	0.1	5	46	0.0046	0.0677	6.5239	93.48
9	2741	0.2	0	55	0.0067	0.0818	7.5158	92.48

From the table above, it can be concluded that the accuracy obtained by optimizing the K value with elbows is higher than the experiment without elbows. Although the difference in K values is not too great, this impacts the accuracy of the model.

CONCLUSION

Based on the results of accuracy testing on 12 trials, it can be concluded that the highest accuracy value was obtained in the second experiment, which was 95.63% with MAPE 4.37%, MSE 0.0018, and RMSE 0.0422. The lowest accuracy was obtained in the 1st experiment with an accuracy of 93.64% with a MAPE of 6.36%, MSE of 0.0028, and RMSE of 0.00524. By comparing several changes in the training dataset variables in the split test and random_state, it can be concluded that K-NN works very well with different variations of the split test and random. This can be seen from the average accuracy value of 95.02%. To determine the K value parameter using the elbow method, the optimal K value can be determined. With the elbow method, you can compare the results obtained from several choices of K values displayed on the elbow chart. As in experiments 3-5 and 6-9, with the same split test and random_state with varying K values, the accuracy values are also different. In carrying out the training model using the K-NN method in this study, only one distance measurement approach was used, namely Euclidean Distance. In future research, it may be necessary to try other distance measurement approaches such as Manhattan distance, *Minkowski Distance*, and Chebychev Distance so that the accuracy obtained in the training model can be compared

BIBLIOGRAPHY

Alaydrus, Zaien Bin Umar, and Arief Andy Soebroto. 2023. "Klasifikasi Masa Panen Varietas Unggul Kedelai Menggunakan K-Nearest Neighbor." *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer* 7(5):2101–6.

Dewi, Dewa Ayu Indah Cahya, and Dewa Ayu Kadek Pramita. 2019. "Analisis Perbandingan Metode Elbow Dan Silhouette Pada Algoritma Clustering K-Medoids Dalam Pengelompokan Produksi Kerajinan Bali." *Matrix: Jurnal Manajemen Teknologi Dan Informatika* 9(3):102–9.

Gorade, Sudhir M., Ankit Deo, and Preetesh Purohit. 2017. "A Study of Some Data Mining Classification Techniques." *International Research Journal of Engineering and Technology* 4(4):3112–15.

Gunst, Richard F., and Robert L. Mason. 2018. *Regression Analysis and Its Application: A Data-Oriented Approach*. CRC Press.

Hartanti, Ninik Tri. 2020. "Metode Elbow Dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional." *J. Nas. Teknol. Dan Sist. Inf* 6(2):82–89.

Hidayat, Taufiqurrakhman Nur, Fendi Aji Purnomo, and Yudho Yudhanto. 2022. "PDAM Performance Clustering Using K-Means." Pp. 148–52 in *2022 1st International Conference on Smart Technology, Applied Informatics, and Engineering (APICS)*. IEEE.

Jadhav, Sayali D., and H. P. Channe. 2016. "Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques." *International Journal of Science and Research (IJSR)* 5(1):1842–45.

Khalid, Haider, and Vincent Wade. 2020. "Topic Detection from Conversational Dialogue Corpus with Parallel Dirichlet Allocation Model and Elbow Method." *ArXiv Preprint ArXiv:2006.03353*.

Kharima, Nadya, and Mohamad Ihsan. 2022. "Pemenuhan Hak Penyandang Disabilitas Dalam Penyusunan Daftar Pemilih Tetap (DPT) Pada Pemilihan Umum (Pemilu) Serentak Tahun 2024." *Jurnal Bawaslu DKI* 7(3):99–118.

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. 2021. *Introduction to Linear Regression Analysis*. John Wiley & Sons.

Riyan, Ade Bani, and Ikhsan Nendi Ikhsan Nendi. 2024. "K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data." *Journal Transnational Universal Studies* 2(2):60–66.

Rosidin, Utang, Uu Nurul Huda, and Burhanuddin Burhanuddin. 2021. "Penanganan Tindak Lanjut Temuan Dan Laporan Dugaan Pelanggaran Pemilu Oleh Badan Pengawas Pemilihan Umum Dalam Mewujudkan Pemilu Yang Berintegritas."

Simanjuntak, Tri Halomoan, Ir Sutrisno, and Wayan Firdaus Mahmudy. 2014. "Implementasi Modified K-Nearest Neighbor Dengan Otomatisasi Nilai K Pada Pengklasifikasian Penyakit Tanaman Kedelai."

Syakur, Muhammad Ali, B. Khusnul Khotimah, E. M. S. Rochman, and Budi Dwi Satoto. 2018. "Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster." P. 12017 in *IOP conference series: materials science and engineering*. Vol. 336. IOP Publishing.

Talakua, Jeane. 2023. "Fulfillment Of Political Rights Of Vulnerable Groups With Disabilities And Inclusive Election Steps In 2024 In Leihitu District, Central Maluku Regency." *Journal Of Humanities And Social Studies* 1(03):946–57.

Umargono, Edy, Jadmiko Endro Suseno, and S. Gunawan. 2019. "K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based-on Mean and Median." Pp. 234–40 in *Proceedings of the International Conferences on Information System and Technology*. SCITEPRESS—Science and Technology Publications Setubal, Portugal.

Wahyudi, Imam, M. Burhanis Sulthan, and Luluk Suhartini. 2021. "Analisa Penentuan Cluster Terbaik

Pada Metode K-Means Menggunakan Elbow Terhadap Sentra Industri Produksi Di Pamekasan." *Jurnal Aplikasi Teknologi Informasi Dan Manajemen (JATIM)* 2(2):72–81.

Winarta, Andre, and Wahyu Joni Kurniawan. 2021. "Optimasi Cluster K-Means Menggunakan Metode Elbow Pada Data Pengguna Narkoba Dengan Pemrograman Python." *JTIK (Jurnal Teknik Informatika Kaputama)* 5(1):113–19.