IJLSSM
INTERNATIONAL JOURNAL OF LAW SOCIAL SCIENCES AND MANAGEMENT
E-ISSN: 3048 - 1317
Volume 1, No. 3
July 2024

Application of the K-Nearest Neighbor Method in Determining Laptop for Classes

***[1] Rofika Qolbi, [2] Tundo, [3] Salsabila Putri Wibowo, [4] Yuma Akbar**

[1,2,3,4] Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika

*Corresponding Author:
rofikaqolbi45@gmail.com

Abstract

Laptops are one of the basic needs in today's modern life. Laptops are used in a wide variety of activities such as work, study, and entertainment. This research aims to be able to predict the class of laptops in the Ilda Computer store. In this process, the K-Nearest Neighbor Algorithm (KNN) method will be applied. There are 2 types of data that will be used in this study, namely training data totaling 80 data and test data as many as 6 data. In the data, there are 7 criteria that will be used, namely Price, Screen Size, Resolution, OS, RAM, Processor Type, and Laptop Class. In this study, it was obtained that the application of the KNN Algorithm can help in determining the prediction of the Laptop Class. And also the application of the KNN algorithm with K=3 obtained the best performance results with an accuracy value of 50%, a presicion of 50%, and a recall of 66%. Meanwhile, with K=4, the best performance results were obtained with an accuracy value of 50%, presicion of 66%, and recall of 50%. Finally, the K=5 obtained the best performance with an accuracy value of 66%, a presicion of 33%, and a recall of 100%.

Keywords: *Algorithm; Data Mining; KNN; Prediction; Laptop*

1. INTRODUCTION

One of the very rapid developments of the times is the development of the technological age where technology is now increasingly developing and provides many conveniences for human life, technological developments from time to time occur as a result of the increasing development of science owned by humans. This makes laptop manufacturers must be able to answer consumer needs according to their market share. Producers are also required to continue to develop long-term strategies to deal with changing consumer tastes in a short period of time. This increase in technology has given birth to products that can facilitate or help human work; one form of this technological innovation is laptops. Laptops have become one of the basic necessities in today's modern life. Laptops are used in various activities such as work, study, and

entertainment. Because of this, the need for laptops is increasing, and the laptop market is growing rapidly with a variety of different brands, models, and specifications on the market at varying prices. In essence, laptops have undergone significant development so that they have many specifications with advanced technology from various well-known brands such as Acer, Asus, Lenovo, HP, and others, but each company produces laptop units are not many or limited according to the generation created, the use of laptops for everyone is also increasing and plays a very important role in carrying out certain activities in their fields, However, this use requires an adequate specification to make performance easier, besides that generally people look from the point of view of big brands to determine the quality obtained and the high price.(Hardiyanto & Rozi, 2020)(Evaristus Nardo et al., 2023)(Alfani P R et al., 2021)(Nur Amalia et al., 2024)

Based on previous research titled "Implementation of Decision Support Systems in Selecting the Best Laptop in the Computer Science Department Using the *Simple Multi Attribute Rating Technique*"has results in the form of a ranking of each alternative analyzed so that it provides a clear and objective picture based on the specified criteria. Based on the percentage of correctness, a result of 100% was obtained where the ranking sequence generated between the manual calculation and the system obtained the same result. Another research with the title "Supporting System for the Convenience of Buying Used Laptops by Applying Methods (Warjaya et al., 2024)*Preference Selection Index* (PSI)" results that research with the *Preference Selection Index* can be petrified in the process of buying a used laptop with the criteria set and weight, namely: Price, Year of Release, RAM, *Hard Disk*, Battery Endurance. *Preference selection index method* It is helpful in making a decision in buying a used laptop in a number of alternatives that exist. Finally, based on previous research with the title "Application of Methods (Hutagaol & Hutahaean, 2019)*Semi Average* In Forecasting Laptop Sales in Stores *Big Computer*"The forecast results in 2023 are obtained with an average value of 3.70833 if rounded to 4 sales numbers every month in 2023.(Wiranto et al., 2023)

This study aims to determine the Laptop Class based on several criteria, namely Price, Screen Size, Resolution, OS, RAM, and Processor Type *The* research will use

the *K-Nearest Neighbor* (KNN) Algorithm method, and it is hoped that the results of this research can determine the Laptop Class that will help in the selection of laptops.

2. METHOD

Data is collected from the ILDA *Computer* Store. The next step involves the data conversion process, where the data obtained is initially discrete and then converted into continuous data. Once the data has been successfully converted, the next step is to carry out the analysis process. After the analysis process, it will be continued with the stage of implementing the KNN Algorithm. Here is the entire process seen in Figure 1

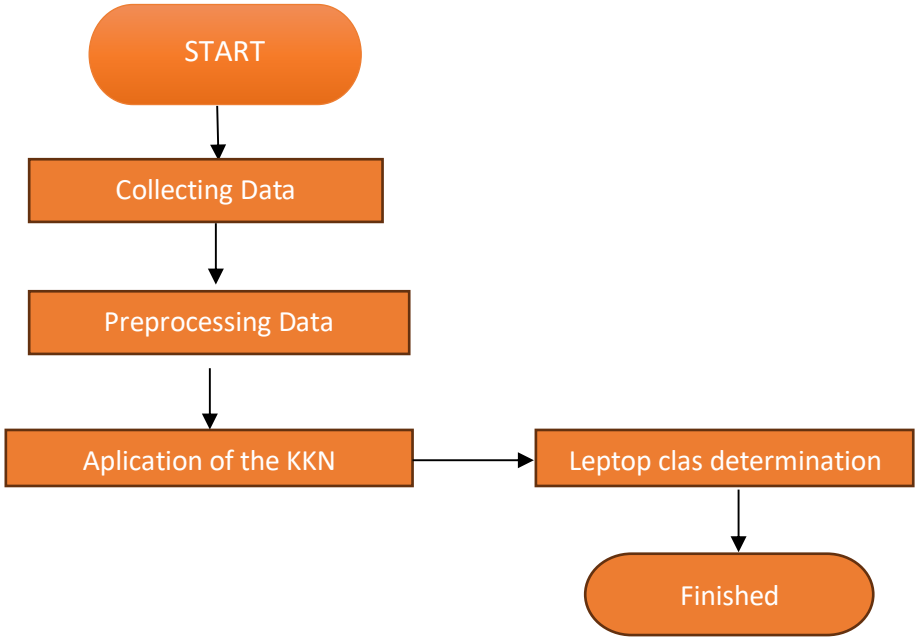

Figure 1. Research Stages

Collecting Data

Data Data collection is the process of obtaining relevant information or facts with a specific purpose. At this stage, the data collected came from the ILDA Computer take. The data used in this study are divided into 2 categories (Azis et al., 2024), namely:

a. Training Data

   The training data used in this research amounted to 80 data.

b. Test Data

   In this study, there are a total of 6 test data to be used.

Preprocessing Data

Data preprocessing is a step for data cleaning, for example removing *noise* and inconsistent data, data integration where multiple data sources can be combined into a single part, data *transformation* where data is changed and consolidated into a form appropriate for data aggregation and reduction operations, including feature selection and extraction. In this study, the method used consists of 2 methods, namely:(Simbolon, 2024)

a. Cleaning Data

   Data Cleaning is a cleaning technique in data that is usually used to handle missing values in data.(Nugroho & Wijana, 2015)

b. Transformation Data

   Data Transformation involves the process of converting raw data into a format that is more suitable for analysis or modeling by modifying the scale, shape, or distribution of the data. The goal is to meet the assumptions of statistical models, reduce outlier influence, or improve understanding of data. In this study, (Ali & Rinaldi, 2023) discrete data will be changed to continue.

Application of the KNN Algorithm

K-Nearest Neighbor, also known by the abbreviation K-NN, is an algorithm that aims to identify similarities between new and old data. This algorithm works by putting new data into categories that are most similar to pre-existing categories. So, the (Wafiqi et al., 2021)*K-Nearest Neighbor* algorithm stores all the old data and classifies the new data based on the degree of similarity. *K-Nearest Neighbor* (K-NN) is one of the algorithms commonly used for data classification. This algorithm makes use of the closest distance on each object to classify the data. In the calculation, two types of data are involved, namely test data and data sets. The next process involves calculating the

distance between the test data and (Dewi et al., 2022)*the dataset* using *the Euclid* (*Query Distance*) method. Euclid's equation 1 (*Query Distance*) can be seen below:

$$Euclidean\ Distance = \sqrt{\sum_{k}^{d} = (XData\ Uji^2 - XData\ Latih^2)} \quad 1)$$

Based on the above formula, the K-NN algorithm can determine the distance between the test data and the training data, aiding in classifying the new data into appropriate categories. This approach provides a foundation for research that aims to predict the class level of laptops in ilda computer stores.

Confusioen Matrix

The Confussion Matrix is used for testing in estimating objects that have True and False (Pare & Marini, 2022)values. The test sequence is carried out in the form of confusion *matrix* tabulation where *the predicted class* will be displayed at the top of the matrix and the observed class on the left. Each numerical data indicates how many cases of the observed class are actually to be predicted. Here is the (Sabda, 2023) *confusion matrix* formula seen in equation 2, equation 3 and equation 4.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} x100\% \quad (2)$$

Accuracy is used to measure the performance of the method.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Precision is used to measure the level of accuracy between the information requested by the user and the answer provided by the system.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

Recall is used to measure the success rate of the system in retrieving information.

3. RESULTS AND DISCUSSION

Collecting Data

At this stage, 80 training data and 6 test data were used. The following is the training data used in this study which is shown in the table 1.

Table 1. Training Data

Laptop Name	Price	Screen Size	Resolution	OS	RAM	Processor Type	Laptop Class
Apple Macbook Pro MJLT2	27500000,00	15,4	2880x1800 *Pixel*	Mac OS X	16	*Core* i7	*High end*
Apple MacBook Air MJVE2ID / A	3000000,00	14,0	1366x768 *Pixel*	Mac OS X	16	*Core* i7	*Low end*
Dell Inspiron 14-3458 | Core i3-4005U	7526300,00	14,0	1366x768 *Pixel*	DOS	4	*Core* i3	*High end*
Lenovo IdeaPad 110-14ISK	1499999,00	14,0	1366x768 *Pixel*	Windows 10	4	*Core* i3	*Low end*
Lenovo IdeaPad S215-6495	3300000,00	11,6	1366x768 *Pixel*	Windows 8	2	E *Series*	*Low end*
Axioo Neon TKMC125	2699999,00	14,0	1366x768 *Pixel*	DOS	4	*Core* i3	*Low end*
Lenovo IdeaPad S215-6495	2699999,00	14,0	1366x768 *Pixel*	Windows 8	2	*Core* i3	*Low end*
Asus Zenbook 14X OLED Space Edition	2750000,00	14,0	2880x1800 *Pixel*	Mac OS X	16	*Core* i7	*Low end*
...	...	...	...	...	...	...	...
Asus TUF Gaming A15 FA506	2699999,00	15,6	2880x1800 *Pixel*	Windows 10	8	*Core* i5	*Low end*

Data Transformation

At this stage, the existing training data is changed, which was initially *discrete* data is changed to *continue*. First, the Price (H) criterion is divided into 3 parts, namely, Good = 3, Not Bad = 2, Poor = 1. Second, the Screen Size (UK) is divided into 3 i.e., Good = 3, Not Bad = 2, Less = 1. Third, the Resolution (R) criteria are divided into 2 parts, namely, Good = 2 and Not Bad = 1. Fourth, the OS criteria are divided into 3 parts, namely, Good = 30, Not Bad = 20, Poor = 10. Fifth, the RAM criteria are divided into 3 parts, namely, Good = 3, Not Bad = 2, Poor = 1. Sixth, the Processor (P) criteria are divided into 3 parts, namely, Good = 3, Not Bad = 2, Poor = 1. Finally, the Laptop Class (KL) criteria are divided into 2 parts, namely, *Highend* = 2 and *Lowend* = 1. The following is table 2 which contains the training data that has been changed.

Table 2. Training Data Transformation

Laptop Name	H	UK	R	OS	RAM	P	KL
Apple Macbook Pro MJLT2	3	3	2	10	3	3	1
Apple MacBook Air MJVE2ID / A	1	2	1	10	3	3	2
Dell Inspiron 14-3458 | Core i3-4005U	2	2	1	30	1	1	1
Lenovo IdeaPad 110-14ISK	1	2	1	20	1	1	2
Lenovo IdeaPad S215-6495	2	1	1	20	1	1	2
Axioo Neon TKMC125	1	2	1	30	1	1	2
Lenovo IdeaPad S215-6495	1	2	1	20	1	1	2
Asus Zenbook 14X OLED Space Edition	1	3	2	10	3	3	2
…	…	…	…	…	…	…	…
Asus TUF Gaming A15 FA506	1	3	2	20	2	2	2

Application of the KNN Algorithm

The steps to apply the KNN algorithm to the determination of the Laptop Class are:

a. The K values used are 3, 4, and 5,

b. Calculate *the Euclid* distance (*Query Distance*) of each training data with the test data using equation 1. The following are the results of the calculation of the test data shown in table 3:

Table 3. Results of the Implementation of the KNN Algorithm

Data Test 1	Data Test 2	Data Test 3	Data Test 4	Data Test 5	Data Test 6
10,14889	10,04988	1,414214	1,414214	3,162278	10,14889
10,24695	10,14889	2,828427	2,828427	2,44949	10,24695
10,58301	10,48809	20,12461	20,12461	20,07486	10,19804
3,605551	3,316625	10,3923	10,3923	10,0995	2,236068
3,605551	3,605551	10,3923	10,3923	10,19804	2,645751
10,63015	10,53565	20,19901	20,19901	20,04994	10,24695
3,605551	3,316625	10,3923	10,3923	10,0995	2,236068
10,14889	10,04988	2,44949	2,44949	2,44949	10,14889
2,236068	1,732051	10,19804	10,19804	10,0995	1
…	…	…	…	…	…
3,162278	2,828427	10,34408	10,34408	10,04988	2

From the results contained in table 3, the K-value will then be determined.

Table 4. K=3 Value

	High end	*Low end*
Test Data 1	3	1
Test Data 2	4	3
Test Data 3	4	4
Test Data 4	4	4
Test Data 5	1	2
Test Data 6	2	2

Table 5. K=4 Values

	High end	*Low end*

Test Data 1	3	1
Test Data 2	4	3
Test Data 3	4	4
Test Data 4	4	4
Test Data 5	5	4
Test Data 6	2	2

Table 6. K=5 Value

	High end	*Low end*
Test Data 1	4	2
Test Data 2	4	3
Test Data 3	4	4
Test Data 4	4	4
Test Data 5	5	4
Test Data 6	3	6

Determination of Laptop Class

At this stage, the results of determining the Laptop Class for the existing test data will be obtained. The following are the results of determining the Laptop Class shown in table 7.

Table 7. Results of Determination of Test Data for Laptop Class Criteria

Laptop Name	KL (K=3)	KL (K=4)	KL (K=5)
Apple Macbook Pro MJLT2	High end	High end	Low end
Apple MacBook Air MJVE2ID / A	Low end	Low end	Low end
Dell Inspiron 14-3458 | Core i3-4005U	High end	High end	Low end
Lenovo IdeaPad 110-14ISK	Low end	Low end	Low end
Lenovo IdeaPad S215-6495	Low end	Low end	Low end
Axioo Neon TKMC125	Low end	Low end	Low end

It can be seen in table 7 that the results of the test data determination for the Laptop Class with K=3, K=4, and K=5 have the same value. Furthermore, a confusion matrix calculation will be carried out based on the results of determining the Laptop Class test data using equations 2, 3, and 4. The following is table 8 for the K=3 confusion matrix, table 9 for the K=4 confussion matrix, and table 10 for the K=5 confusion matrix.

Table 8. Confusion Matrix (K=3)

Actual	Predicted		Sum
	Highend	Lowend	
Highend	TP = 2	FP = 2	4
Lowend	FN = 1	TN = 1	2
Sum	3	3	6

After the value of TP, FP, FN, TN is known, then the performance value can be calculated as follows.

$$\text{Accuracy} = \frac{2+1}{6} = \frac{3}{6} \text{ x } 100\% = 50\%$$

$$\text{Precision} = \frac{2}{2+2} = \frac{2}{4} = 50\%$$

$$\text{Recall} = \frac{2}{2+1} = \frac{2}{3} = 66\%$$

Table 9. Confusion Matrix (K=4)

Actual	Predicted		Sum
	Highend	Lowend	
Highend	TP = 2	FP = 1	3
Lowend	FN= 2	TN =1	3
Sum	4	2	6

After the value of TP, FP, FN, TN is known, then the performance value can be calculated as follows.

$$Accuracy = \frac{2+1}{6} = \frac{3}{6}\ x\ 100\%$$
$$= 50\%$$

$$Precision = \frac{2}{2+1} = \frac{2}{3} = 66\%$$

$$Recall = \frac{2}{2+2} = \frac{2}{4} = 50\%$$

Table 10. Confusion Matrix (K=5)

Actual	Predicted		Sum
	Highend	Lowend	
Highend	TP = 1	FP = 2	3
Lowend	FN= 0	TN =3	3
Sum	0	6	6

After the value of TP, FP, FN, TN is known, then the performance value can be calculated as follows.

$$Accuracy = \frac{1+3}{6} = \frac{4}{6}\ x\ 100\% = 66\%$$

$$Precision = \frac{1}{1+2} = \frac{1}{3} = 33\%$$

$$Recall = \frac{1}{1+0} = \frac{1}{1} = 100\%$$

Based on the results of the calculation above, it can be concluded that from testing the prediction data and actual data using the KNN algorithm with K=3, the best performance results with an accuracy value of 50%, a probability of 50%, and a recall of 66% can be obtained. Meanwhile, with K=4, the best performance results were obtained with an accuracy value of 50%, presicion of 66%, and recall of 50%. Finally, the K=5 obtained the best performance with an accuracy value of 66%, a presicion of 33%, and a recall of 100%.

4. CONCLUSION

Based on the results of the study, several conclusions were obtained, namely:

a. The application of the K-Nearest Neighbor algorithm method to determine the Laptop Class based on data taken from the Ilda Computer Store can help in determining the laptop class
b. Based on the results of the study using the KNN algorithm with K=3, the best performance results were obtained with an accuracy value of 50%, a presicion of 50%, and a recall of 66%. Meanwhile, with K=4, the best performance results were obtained with an accuracy value of 50%, presicion of 66%, and recall of 50%. Finally, the K=5 obtained the best performance with an accuracy value of 66%, a presicion of 33%, and a recall of 100%.
c. Based on the results, the KNN algorithm used in determining the most appropriate laptop class is K=5 with the greatest accuracy value compared to K=4 and K=3.
d. It is hoped that this research can help subsequent research and the results of this research can be developed again by future researchers.

References

Alfani P R, A. W., Rozi, F., & Sukmana, F. (2021). PREDIKSI PENJUALAN PRODUK UNILEVER MENGGUNAKAN METODE K-NEAREST NEIGHBOR. JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika), 06, 155–160.

Ali, I., & Rizki Rinaldi, A. (2023). PENERAPAN METODE K-NEAREST NEIGHBOR UNTUK PREDIKSI PENJUALAN SEPEDA MOTOR TERLARIS. Jurnal Mahasiswa Teknik Informatika, 7(1).

Angga Sabda, M. (2023). Implementasi Data Mining Dalam Memprediksi Penjualan Parfum Terlaris Menggunakan Metode K-Nearest Neighbor. Jurnal Sistem Komputer Dan Informatika (JSON) Hal: 415−, 422(2). https://doi.org/10.30865/json.v5i2.7194

Azis, A., Zy, A. T., & Sunge, A. S. (2024). Prediksi Penjualan Obat Dan Alat Kesehatan Terlaris Menggunakan Algoritma K-Nearest Neighbor. Jurnal Teknologi Dan Sistem Informasi Bisnis, 6(1), 117–124. https://doi.org/10.47233/jteksis.v6i1.1078

Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. Building of

Informatics, Technology and Science (BITS), 3(4), 639–648. https://doi.org/10.47065/bits.v3i4.1408

Evaristus Nardo, Asrul Sani, Agusta Pratama Wibawa, & Nur Nawaningtyas Pusparini. (2023). Penentuan Rekomendasi Laptop Terbaik Bagi Customer Dalam Sistem Pendukung Keputusan Dengan Metode SAW Di CV. Jeflin Laptop. Jurnal Publikasi Sistem Informasi Dan Manajemen Bisnis, 3(1), 113–124. https://doi.org/10.55606/jupsim.v3i1.2367

Hardiyanto, B., & Rozi, F. (2020). PREDIKSI PENJUALAN SEPATU MENGGUNAKAN METODE K-NEAREST NEIGHBOR. JOEICT(Jurnal of Education and Information Communication Technology), 04, 13–18.

Hutagaol, J., & Hutahaean, K. M. (2019). Seminar Nasional Sains & Teknologi Informasi (SENSASI) Sistem Pendukung Kepetusan Pembelian Laptop Bekas dengan Menerapkan Metode Preference Selection Index (PSI). Seminar Nasional Sains & Teknologi Informasi (SENSASI), 446–451. http://prosiding.seminar-id.com/index.php/sensasi/issue/archivePage|446

Nugroho, R. S., & Wijana, K. (2015). PROGRAM BANTU PREDIKSI PENJUALAN BARANG MENGGUNAKAN METODE KNN Studi Kasus: U.D. ANANG. Jurnal EKSIS, 08, 83–93.

Nur Amalia, O., Suarna, N., & Prihartono, W. (2024). MENENTUKAN KEPUTUSAN KONSUMEN UNTUK MEMBELI LAPTOP MENGGUNAKAN METODE DECISION TREE ID3 (STUDI KASUS MAHASISWA DI CIREBON). Jurnal Mahasiswa Teknik Informatika, 8(2).

Rosalina Pare, F., & Ferdinand Marini, L. (2022). Prediksi Harga Jual Beras Eceran Menggunakan Metode K-Nearst Neighbor Di Kabupaten Manokwari. Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi, 357–360.

Simbolon, I. N. (2024). PREDIKSI KUALITAS AIR SUNGAI DI JAKARTA MENGGUNAKAN KNN YANG DIOPTIMALISASI DENGAN PSO. Jurnal Informatika Dan Teknik Elektro Terapan, 12(2). https://doi.org/10.23960/jitet.v12i2.4191

Ulul Azmi Wafiqi, A., Arvian James, B., Huga Ramadhan, A., & Nizar, A. (2021). Prediksi Tingkat Stres Pada Mahasiswa UNUGHA Cilacap Menggunakan Algoritma K-Nearest Neighbor. Jurnal TEKNO KOMPAK, 18(2), 331–343.

Warjaya, A., Meirza, A., Rafaela Puteri, N., & Yandra Niska, D. (2024). Implementasi Sistem Pendukung Keputusan Dalam Pemilihan Laptop Terbaik dalam Jurusan Ilmu Komputer Menggunakan Metode Simple Multi Attribute Rating Technique. Jurnal Teknologi Informatika Dan Komunikasi (JICT), 6(1), 198–208. https://doi.org/10.52661

Wiranto, F., Latipa Sari, H., & Elfianty, L. (2023). PENERAPAN METODE SEMI AVERAGE DALAM PERAMALAN PENJUALAN LAPTOP DI TOKO BIG

COMPUTER. Journal of Science and Social Research, 3, 796–806. http://jurnal.goretanpena.com/index.php/JSSR