JURNAL RISET INFORMATIKA **P-ISSN: 2656-1743 |E-ISSN: 2656-1735**
Vol. 5, No. 2 March 2023 DOI: https://doi.org/10.34288/jri.v5i2.205
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Evaluation of Machine Learning Using the K-NN Algorithm To Determine the Quality of Meat Before Consumption

Feronika[-1*)], Masrizal[-2], Ibnu Rasyid Munthe[-3]

[1,3] Department of Information system, [2]Department of Informatics Management
[1,2,3] Universitas Labuhanbatu
Sumatera Utara, Indonesia
[1*)] feronika@gmail.com, [2]masrizal120405@gmail.com, [3]ibnurasyidmunthe@gmail.com

(*) Corresponding Author

Abstract

Meat is one of the sources of animal protein for humans, and one of the requirements that must be met so that the human body does not lack protein, especially animal; this protein can be obtained from beef, chicken, and other meats, but the most important thing here is the content contained in meat, whether it has been contaminated with chemicals, e.g., chicken that has been injected with chemicals that cause the chicken to look fat, or beef whose flexibility has decreased and the pH is getting more acidic. This research tries to predict meat quality by looking at two parameters: flexibility and acidity. The programming language used is R Language, using the k-NN method or Algorithm to determine the meat's condition suitable for consumption. In detail, it will be processed in Machine Learning using the k-NN Algorithm; there are two criteria for consumption of meat, namely good or not good for consumption; in detail, the output will be explained using a specific graph using a plot function, and array data will be specifically classified to represent values. The value of 2 variables, namely feasible or not suitable for consumption.

Keywords: Machine Learning, K-NN Algorithm, R Language, Meat, Acidity Prediction, Flexibility Prediction

Abstrak

Daging adalah salah satu sumber protein hewani bagi manusia, dan salah satu syarat yang harus dipenuhi agar tubuh manusia tidak kekurangan protein khususnya hewani, protein ini bisa didapatkan dari sapi, ayam, dan daging lainnya, namun yang terpenting disini adalah kandungan yang terdapat pada daging, apakah sudah terkontaminasi dengan zat kimia, misalnya ayam potong yang telah disuntik dengan zat kimia yang menyebabkan ayam kelihatan gemuk, atau daging sapi yang tingkat kelenturannya sudah berkurang dan pH yang semakin asam. Riset ini mencoba untuk melakukan prediksi kualitas daging dengan melihat dua parameter yaitu kelenturan dan keasaman. Bahasa pemrograman yang digunakan adalah R Language, menggunakan metode atau algoritma k-NN yang dapat menentukan kondisi daging layak untuk dikonsumsi. Secara detail akan diolah pada Machine Learning menggunakan algoritma k-NN ini, terdapat dua kriteria daging konsumsi yaitu bagus atau kurang bagus untuk dikonsumsi, secara detail output akan dijelaskan menggunakan specific graph menggunakan plot function, dan data array akan secara spesifik diklasifikasikan untuk merepresentasikan nilai-nilai dari 2 variable yaitu layak atau tidak layak konsumsi.

Kata kunci: Machine Learning, K-NN Algorithm, R Language, Daging, Prediksi Keasaman, Prediksi Kelenturan

INTRODUCTION

The economic factor is an essential and decisive factor to be able to escape the *recession* in Indonesia in 2023, so what needs to be done by farmers is to utilize the existing land for several types of livestock, including broiler breeders. Beef or chicken in the freezer for a long time causes the level of flexibility in the beef to be reduced or hard. Likewise, with cattle farms, broiler breeders, and other breeders such as rabbits, goats, and others whose type of cuisine tends to be "*Sate*." Researchers need to write and share in detail new methods such as AIoT & LoRaWAN (Adi & Kitagawa, 2020), (Liani et al., 2021), (Mukti et al., 2021). Machine Learning and k-NN Algorithm (Lv et al., 2021), (Du & Li, 2019), ((Jia, 2022) in solving quality problems of beef, chicken, and other types of meat consumption. One of the Indonesian people's favorite foods is 'Meatballs', like a resident of Malang, East Java, Indonesia, is one of the fans of meatballs with many variants of meatballs.

Meatballs are also made from meat, e.g., beef. However, if when it is processed into meatballs, you see the quality of the beef being grilled, is it healthy, fresh, or the condition of the meat is no longer red or bluish.

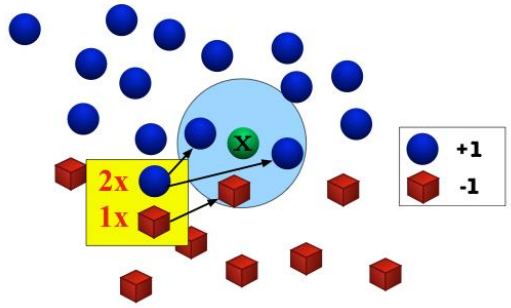


Figure 1. k-NN Algorithm overview

Figure 1 can be seen as an example of a binary classification with k=3. The green dot in the middle is the test sample x. The labels of the three neighbors are 2×(+1), and 1×(-1) yields the majority prediction (+1). K-NN Algorithm ((Pawlovsky & Matsuhashi, 2017),(Palacios & Suzuki, 2019) uses a distance matrix, and it concludes that the closest parameter is the best conclusion. Equation 1 shows the formula for this K-NN Algorithm (Chang & Liu, 2011), (Yao & Cao, 2022), (Salim et al., 2020), (Song et al., 2020), (Li et al., 2020).

$$\text{dist}(x,x') \geq \max \text{dist}(x,x''), (x'',y'') \in S_x \quad \text{.....(1)}$$

$$\text{dist}(x, z)=\left(\sum_{r=1}^{d} |X_r - Z_r|^p\right)^{1/p} \quad \text{.....(2)}$$

$$h(x)=\text{mode}\ (\{y''(\ x'',y'') \in S_x\ \}) \quad \text{.....(3)}$$

$$\sqrt{\sum_{i=1}^{k} x_i - y_i{}^2} \quad \text{.....(4)}$$

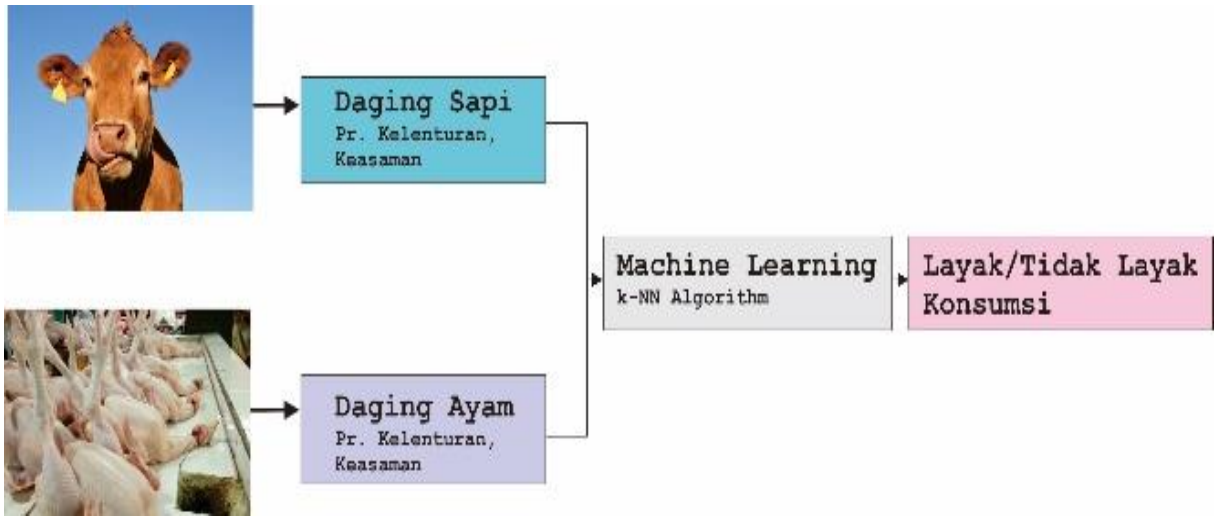


Figure 2. k-NN Algorithm for the feasibility of meat consumption

Furthermore, Figure 2 explains the method used in this manuscript where the k-NN method (Yao & Cao, 2022), (Song et al., 2020), (Sun et al., 2020), (Sushmitha & Jagadeesh, 2022) can solve the problem of Eligible or Not Like Meat consumed by the community, especially in Indonesia.

RESEARCH METHODS

This research will focus on using the k-NN Algorithm (Zhai, 2022), (Li et al., 2020), (Jia, 2022), (Sushmitha & Jagadeesh, 2022), (Sun et al., 2020), (Yunneng, 2020), (Palacios & Suzuki, 2019) and (Setia & Garg, 2021). This Algorithm is used to classify new objects based on attributes and samples from training data, and this Algorithm uses the predicted value of the new instance value.

Step-by-step to running of k-NN Algorithm in general:

1. Choose k nearest neighbors randomly
2. Map dataset to vector space
3. Separate the dataset into training data and test data
4. Calculate the distance, d, between the test data and the training data
5. Sorting d from smallest to largest
6. Take and separate k-sorting data
7. Observe the majority class
8. Classify test data by majority

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v5i2.205
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Step-by-step to running of k-NN Algorithm on this project:

1. Group product data into product array objects according to their respective categories
2. Group the product brand data (label) into the brand array object according to their respective categories
3. Plot an Array of products using different symbols
4. Plot array brands
5. Plot legend

Furthermore, it will be explained in full in the following flowchart, where the value of k must be determined first, then calculate the distance, sort it, and classified the test data. Moreover, Pseudo code is made to understand the flow of the program so that readers do not experience difficulties in understanding (Setia & Garg, 2021).

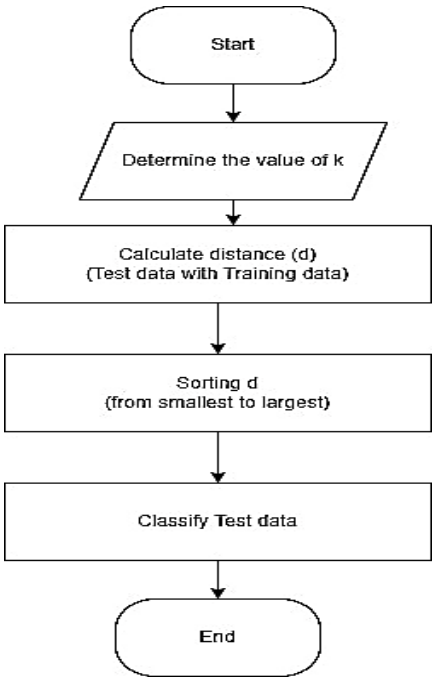


Figure 3. k-NN Algorithm flowchart

Pseudocode_1

// X: training data
// Y: class labels of X
//x: unknown sample
Classify (X,Y,x)
For i = 1 to m do
Compute distance d(X,..,x)
end for
Compute set 1 containing indicates for the k smallest distances d (X,...x).
End for
Compute set 1 containing indicates for the k smallest distances d (X,..x).
Return majority label for {Yi where I e l}

RESULTS AND DISCUSSION

The following table data determines the value to look for or predictions that will give the best quality value on meat quality. Moreover, Table 1 is a Classification of Meat Quality.

Table 1. Classification of Meat Quality

Brand	acidity value	Flexibility value	Category
A	7	1.2	Good
B	6	1.7	Not good
C	8	1.5	Good
D	5	1.3	Not good
E	9	1.0	Good
F	9.5	1.4	Good
G	8.3	0.8	Not good
H	7.5	1.1	?

Furthermore, the next step is to run the program using R Language using K-NN Algorithm (Salim et al., 2020), (Riquelme et al., 2020); the display on the Website page is data taken; this is stage 1. next is The relationship between flexibility and uniformity of product data as specifically shown in figure 4.

Code_1_running

```
x<-c(7,6,8,5,9,9.5,8.3,7.5)
> y<-
c(1.2,1.7,1.5,1.3,1.0,1.4,0.8,1.1)
> bagus_x<-c(7,8,9,9.5)
> bagus_y<-c(1.2,1.5,1.0,1.4)
> kurang_x<-c(6,5,8.3)
> kurang_y<-c(1.7,1.3,0.8)
> test_x<-(7.5)
> test_y<-(1.1)
> merk_bagus<-c('A','C','E','F')
> merk_kurang<-c('B','D','G')
> merk_test<-c('H')
>plot(x,y,col="blue",main="dataproduk"
,cex=1.3,pch=16,xlab="keasamaan",ylab=
"kelenturan")
```

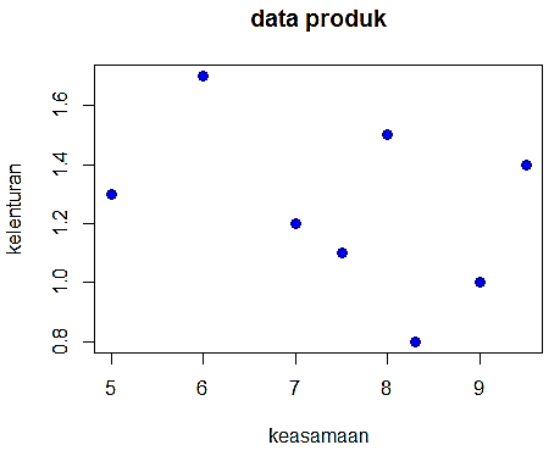
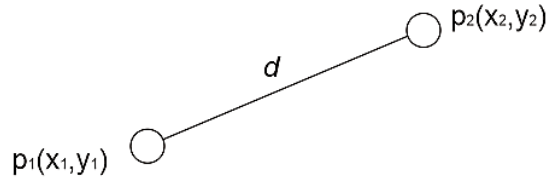


Figure 4. The relationship between flexibility and uniformity of product data

Step 2 is to determine the value of k; for example, if k is 3, then k=3, we can choose k odd, 1, 3, or 5. Then step 3 is to calculate the distance (d) between the test data (H) and neighbors; based on the results of observations of the type of data and the shape of the graph plot, then we can use the Euclidean distance formula to determine the distance d as the Figure 5.



Euclidean distance (d) = $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

Figure 5. The Euclidean distance (d)

Then we will determine the coordinate value, the coordinate value is seen in Table 2.

Table 2. Coordinate value

Number	Coordinate
1	A(7, 1.2)
2	B (6, 1.7)
3	C(8,1.5)
4	D(5,1.3)
5	E(9,1.0)
6	F (9.5,1.4)
7	G(8.3,0,8)
8	H(7.5,1.1)

$d_{HA} = \sqrt{7.5 - 7^2 + 1.1 - 1.2^2}$ = 0.509902
$d_{HB} = \sqrt{7.5 - 6^2 + 1.1 - 1.7^2}$ = 1.615549
$d_{HC} = \sqrt{7.5 - 8^2 + 1.1 - 1.5^2}$ = 0.640312
$d_{HD} = \sqrt{7.5 - 5^2 + 1.1 - 1.3^2}$ = 2.507987
$d_{HE} = \sqrt{7.5 - 9^2 + 1.1 - 1.0^2}$ = 1.503333
$d_{HF} = \sqrt{7.5 - 9.5^2 + 1.1 - 1.4^2}$ = 2.02237
$d_{HG} = \sqrt{7.5 - 8.3^2 + 1.1 - 0.8^2}$ = 0.85440

Stage 4 is sorting the results of the d calculations from the smallest to the most significant d. Then choose d as much as the value of k, namely k = 3 pieces, before sorting the results of the d calculation as following Table 3.

Table 3. Distance calculation

d	d value	Classification
HA	0.509902	A (Good)
HB	1.615549	B (Good)
HC	0.640312	C (Not good)
HD	2.507987	D (Not good)
HE	1.503333	E (Not good)
HF	2.022375	F (Good)
HG	0.854400	G (Good)

The next step is to determine a good product with a cross and determine the test data. The plot results can be seen in Figure 3. In Figure 3, those marked with a cross are of good quality, while those not marked with a cross are data that are not of good quality. The box is a sample of test data. Moreover, The relationship between flexibility and uniformity of product data is presented explicitly in Figure 6.

```
Code_2_running
>points(bagus_x,bagus_y,col="red",pch=
4,lwd=2,cex=2)
>points(test_x,test_y,col="black",pch=
22,lwd=2,cex=2)
```

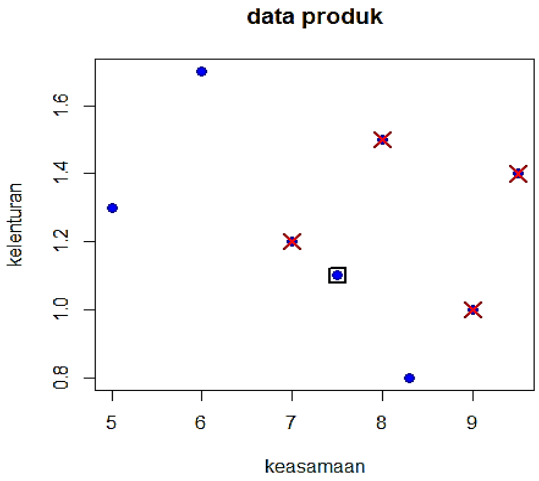


Figure 6. The relationship between flexibility and uniformity of product data

CONCLUSIONS AND SUGGESTIONS

Conclusion

Using the K-NN Algorithm using R Language, the meat quality can be determined by several parameters, namely the value of flexibility and the value of acidity, which is inputted for processing or the data to be trained by the K-NN Algorithm. Good quality of the meat was found at acidity seven and flexibility 1.2, 8 and 1.5, 9 and 1.0, and 9.5 and 1.4

Suggestion

The new method can be compared to produce an error comparison value (%) of the quality of processed meat or meat consumed by the

community if it is seen not only as two parameters but can be more.

REFERENCES

Adi, P. D. P., & Kitagawa, A. (2020). Performance evaluation of low power wide area (LPWA) LoRa 920 MHz sensor node to medical monitoring IoT based. *EECCIS 2020 - 2020 10th Electrical Power, Electronics, Communications, Controls, and Informatics Seminar*, 278–283. https://doi.org/10.1109/EECCIS49483.2020.9263418

Chang, Y., & Liu, H. (2011). Semi-supervised classification algorithm based on the KNN. *2011 IEEE 3rd International Conference on Communication Software and Networks, ICCSN 2011*, 9–12. https://doi.org/10.1109/ICCSN.2011.6014376

Du, S., & Li, J. (2019). Parallel Processing of Improved KNN Text Classification Algorithm Based on Hadoop. *2019 7th International Conference on Information, Communication and Networks (ICICN)*, 167–170. https://doi.org/10.1109/ICICN.2019.8834973

Jia, X. (2022). Detecting Water Quality Using KNN, Bayesian and Decision Tree. *Proceedings - 2022 Asia Conference on Algorithms, Computing and Machine Learning, CACML 2022*, 323–327. https://doi.org/10.1109/CACML55074.2022.00061

Li, B., Wan, S., Xia, H., & Qian, F. (2020). The Research for Recommendation System Based on Improved KNN Algorithm. *Proceedings of 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications, AEECA 2020*, 796–798. https://doi.org/10.1109/AEECA49918.2020.9213566

Liani, Y. A., Munthe, I. R., Irmayani, D., Broto, B. E., Yanris, G. J., Prasetya, D. A., Haryanto, R., Adi, P. D. P., Muslikh, A. R., & Arifuddin, R. (2021). The Broiler Chicken Coop Temperature Monitoring Use Fuzzy Logic and LoRAWAN. *Proceeding - ICERA 2021: 2021 3rd International Conference on Electronics Representation and Algorithm*, 161–166. https://doi.org/10.1109/ICERA53111.2021.9538771

Lv, W., Huang, H., Tang, W., & Chen, T. (2021). Research and Application of Intersection Similarity Algorithm Based on KNN Classification Model. *2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA)*, 141–144. https://doi.org/10.1109/CAIBDA53561.2021.00037

Mukti, F. S., Adi, P. D. P., Prasetya, D. A., Sihombing, V., Rahanra, N., Yuliawan, K., & Simatupang, J. (2021). Integrating Cost-231 Multiwall Propagation and Adaptive Data Rate Method for Access Point Placement Recommendation. *International Journal of Advanced Computer Science and Applications*, *12*(4), 772–777. https://doi.org/10.14569/IJACSA.2021.0120494

Palacios, P. A., & Suzuki, Y. (2019). An Immune Algorithm that Uses a Master Cell for Component Selection for the kNN Method. *Pan American Health Care Exchanges, PAHCE*, *2019-March*. https://doi.org/10.1109/GMEPE-PAHCE.2019.8717318

Pawlovsky, A. P., & Matsuhashi, H. (2017). The use of a novel genetic algorithm in component selection for a kNN method for breast cancer prognosis. *2017 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, 1–5. https://doi.org/10.1109/GMEPE-PAHCE.2017.7972084

Riquelme, J. A., Barrientos, R. J., Hernandez-Garcia, R., & Navarro, C. A. (2020). An exhaustive algorithm based on GPU to process a kNN query. *2020 39th International Conference of the Chilean Computer Science Society (SCCC)*, 1–8. https://doi.org/10.1109/SCCC51225.2020.9281231

Salim, A. P., Laksitowening, K. A., & Asror, I. (2020). Time Series Prediction on College Graduation Using KNN Algorithm. *2020 8th International Conference on Information and Communication Technology, ICoICT 2020*. https://doi.org/10.1109/ICOICT49345.2020.9166238

Setia, V., & Garg, S. (2021). Psychometric Assessment of Teacher Mindset using hybrid KNN-SVM Algorithm. *Proceedings of the 5th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2021*, 951–955. https://doi.org/10.1109/I-SMAC52330.2021.9640923

Song, H., Zhu, E., & Ma, L. (2020). *Design of embedded real-time target tracking system based on KNN algorithm*. 443–446. https://doi.org/10.1109/ITAIC49862.2020.9339101

Sun, Y., Chen, M., & Lin, G. (2020). Design of Intelligent Paper Counting System Based on KNN Algorithm. *Proceedings of the 32nd Chinese Control and Decision Conference, CCDC 2020*, 2910–2914. https://doi.org/10.1109/CCDC49329.2020.9164216

Sushmitha, K., & Jagadeesh, P. (2022). Feature Extraction and Classification of Bone Tumor using CNN Classifier with KNN Classifier. *2022 International Conference on Business Analytics for Technology and Security, ICBATS 2022*. https://doi.org/10.1109/ICBATS54253.2022.9759035

Yao, L., & Cao, T. (2022). Research on Talent Classification Based on Improved KNN Algorithm. *Proceedings - 2022 International Symposium on Control Engineering and Robotics, ISCER 2022*, 252–256. https://doi.org/10.1109/ISCER55570.2022.00050

Yunneng, Q. (2020). A new stock price prediction model based on improved KNN. *Proceedings - 2020 7th International Conference on Information Science and Control Engineering, ICISCE 2020*, 77–80. https://doi.org/10.1109/ICISCE50968.2020.00026

Zhai, H. (2022). Improving KNN Algorithm Efficiency Based on PCA and KD-tree. *Proceedings - 2022 International Conference on Machine Learning and Knowledge Engineering, MLKE 2022*, 83–87. https://doi.org/10.1109/MLKE55170.2022.00021