Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di Kedai Kopi Krintji

TUGAS AKHIR

Ario Bimo Kuntjoro Vincent
41518010151

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di kedai Kopi Krintji

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
Ario Bimo Kuntjoro Vincent
41518010151

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

Nama Mahasiswa	:	Ario Bimo Kuntjoro Vincent
NIM	:	41518010151
Judul Tugas Akhir	:	Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di Kedai Kopi Krintji

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 03 Agustus 2022

Ario Bimo Kuntjoro Vincent

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Ario Bimo Kuntjoro Vincent
NIM	:	41518010151
Judul Tugas Akhir	:	Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di Kedai Kopi Krintji

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 03 Agustus 2022

Ario Bimo Kuntjoro Vincent

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Ario Bimo Kuntjoro Vincent
NIM	:	41518010151
Judul Tugas Akhir	:	Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di Kedai Kopi Krintji

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	✓
		Jurnal Nasional Terakreditasi	✓		
		Jurnal International Tidak Bereputasi		Diterima	
		Jurnal International Bereputasi			
	Disubmit/dipublikasikan di :	Nama Jurnal	: EDAS Conference and Journal Management System		
		ISSN	:		
		Link Jurnal	: https://www.edas.info/		
		Link File Jurnal Jika Sudah di Publish	:		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 03 Agustus 2022

Mengetahui
Dosen Pembimbing TA

Rahmat Budianto, Dr. Prof

Ario Bimo Kuntjoro Vincent

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518010151
Nama	:	Ario Bimo Kuntjoro Vincent
Judul Tugas Akhir	:	Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di kedai Kopi Krintji

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2022

(Saruni Dwiasnati, ST, MM, M.Kom)

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518010151
Nama	:	Ario Bimo Kuntjoro Vincent
Judul Tugas Akhir	:	Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di kedai Kopi Krintji

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2022

(Achmad Kodar, Drs. MT)

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518010151
Nama	:	Ario Bimo Kuntjoro Vincent
Judul Tugas Akhir	:	Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di kedai Kopi Krintji

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 3 Agustus 2022

(Wawan Gunawan, S.Kom, MT)

LEMBAR PENGESAHAN

NIM	: 41518010151
Nama	: Ario Bimo Kuntjoro Vincent
Judul Tugas Akhir	: Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di kedai Kopi Krintji

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 22 Juli 2022

Menyetujui,

(Rahmat Budiarto, Dr. Prof)
Dosen Pembimbing

Mengetahui,

(Wawan Gunawan, S.Kom, MT)	(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Koord. Tugas Akhir Teknik Informatika	Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan kehadirat Allah SWT atas segala rahmat dan karunia-Nya, dikarenakan Tugas Akhir yang berjudul "Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di kedai Kopi Krintji" dapat diselesaikan dalam jangka waktu yang ditentukan. Laporan Tugas Akhir ini dibuat sebagai syarat untuk LULUS sebagai sarjana Ilmu Komputer dari Universitas Mercu Buana.

Penulis menyadari bahwa pembuatan Tugas Akhir ini tidak terlepas dari bantuan dan bimbingan berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Kedua orang tua saya yang telah membesarkan saya, selalu mendoakan yang terbaik untuk saya agar dapat menyelesaikan tugas akhir dan juga tak luput terus mendukung saya karena bisa menyelesaikan kuliah dengan baik.
2. Ibu Saruni Dwiasnati, ST, MM, M.Kom selaku Dosen Pembimbing Akademik.
3. Bapak Rahmat Budiarto, Dr. Prof selaku Dosen Pembimbing Tugas Akhir yang telah membimbing serta memberikan nasihat dalam penyusunan Tugas Akhir sampai selesai.
4. Bapak Emil R. Kaburuan, Ph.D selaku Kepala Program Studi Informatika yang telah memberikan arahan selama perkulihan.
5. Bapak Wawan Gunawan, S.Kom, MT selaku Koordinator Tugas Akhir Prodi Informatika yang telah memberikan arahan selama periode tugas akhir.
6. Teman – teman saya di group panik yang telah membantu dan mensupport saya dalam menyelesaikan Tugas Akhir.
7. Semua pihak dan personal yang tidak dapat disebutkan satu per satu yang terlibat dalam pembuatan Tugas Akhir ini sehingga dapat selesai dengan baik.

Semoga perjuangan saya dalam mengusul Tugas Akhir ini akan menjadi perjalanan saya yang berharga, dan semoga saya berserta teman – teman yang menjalani Tugas Akhir diberikan Kesehatan dan kelulusannya.

Jakarta, 22 Juli 2022
Penulis

Ario Bimo Kuntjoro Vincent

DAFTAR ISI

HALAMAN SAMPUL.....i
HALAMAN JUDUL.....i
LEMBAR PERNYATAAN ORISINALITAS.....ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR.....iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....iv
LEMBAR PERSETUJUAN PENGUJI v
LEMBAR PENGESAHAN viii
ABSTRAK ix
ABSTRACT.....x
KATA PENGANTAR xi
DAFTAR ISI xiii
NASKAH JURNAL 1
KERTAS KERJA 7
BAB 1. LITERATUR REVIEW.....8
BAB 2. ANALISIS DAN PERANCANGAN 19
BAB 3. SOURCE CODE.....21
BAB 4. DATASET 22
BAB 5. TAHAPAN EKSPERIMEN 25
BAB 6. HASIL SEMUA EKSPERIMEN 28
DAFTAR PUSTAKA.....34
LAMPIRAN DOKUMEN HAKI 36
LAMPIRAN KORESPONDENSI.....38
LAMPIRAN SOURCE CODE 41

NASKAH JURNAL

Menu Classification using KNN and Naïve Bayes Algorithm at Kopi Krintji Café

Ario Bimo Kuntjoro Vincent
Informatics Dept.
Faculty of Computer Science
Universitas Mercu Buana
Jakarta, Indonesia
41518010151@student.mercubuana.ac.id

Anis Cherid
Informatics Dept.
Faculty of Computer Science
Universitas Mercu Buana
Jakarta, Indonesia
anis.cherid@mercubuana.ac.id

Deris Stiawan
Computer Engineering Department
Faculty of Computer Science
Universitas Sriwijaya
Palembang, Indonesia
deris@unsri.ac.id

Rahmat Budiarto
Informatics Dept.
Faculty of Computer Science
Universitas Mercu Buana
Jakarta, Indonesia
Rahmat.budiarto@mercubuana.ac.id

***Abstract*— nowadays there are a lot of cafes in Indonesia, especially in Jakarta, which is growing very rapidly. Youngsters and adults come just for hang out with their friends or families. Of course customers have favorite and non-favorite foods and drinks in the menu. To increase and maintain the customer numbers, a café must keep being creative and monitoring their menu. This research helps the Krintji Coffee Café innovating new variants of food and drink to be added to their menus to comfort customers. This research work utilizes the Naïve Bayes and K– Nearest Neighbor (KNN) algorithms to classify the Kopi Krintji Café 's menu. The experimental results show that KNN algorithm provides an accuracy value of 98%, precision and recall, each of 98% with 10-fold cross validation test of 98% accuracy. On the other hand, Naïve Bayes algorithm provides 97% accuracy, 95% precision and 93% recall while accuracy during the validation using k-fold cross validation with values of k are 10, 15, 20 and 25 is 96%.**

Keywords*— *Café menu, Classification, Naïve Bayes, K-Nearest Neighbor

I. INTRODUCTION

Nowadays there are a lot of cafes in Indonesia, especially in Jakarta, which is growing very rapidly. Youngsters and adults come just for hang out with their friends or families. Even many of them have reached the point of success and one example is the Kopi Krintji cafe located in Kebayoran Baru area, Southern Jakarta. Kop Krintji is one of the hits and famous cafes in the Southern Jakarta area, because the place is comfortable and the price is reasonable for the customers. To increase and maintain the customer numbers, a café must keep being creative and monitoring their menu according to the favorite menus

The main benefit of data science is gaining insight from the daily running data of a business. In the case of Kopi Kerinci café, sales transaction data is one of insights sources that can support business decision making, however, most sales transaction data is not utilized. It is only stored as an archive and only being used for generating sales report [1].

Large number of customers who visiting Kopi Krintji, have food and drinks of their choice or their favorites and at the same time they have foods and drinks that they do not like. This research helps the Krintji Coffee Café innovating new variants of food and drink to be added to their menus to comfort customers. This research work utilizes intelligent classifiers to classify whether an item in the Kopi Krintji Café's menu is popular or not. The authors choose KNN algorithm as classifier and Naïve Bayes algorithm as comparison.

This study uses the KNN algorithm and Naïve Bayes to assist Kopi Krintji cafe in determining new menu that will being liked by the customer and become favorite menu. The KNN algorithm is a method for classifying objects based on learning

data that is closest to the object. Learning data is projected into multi-dimensional spaces, where each dimension represents a feature of the data [2]. The Naïve Bayes algorithm is a Simple probabilistic-based prediction technique using Bayes's theorem (Bayes' rule) assuming strong (naïve) independence or non-dependence [3]. In other words, in Naïve Bayes the model used is an independent feature model [4].

II. RESEARCH METHOD

Figure 1 illustrates the research method carried on in this paper.

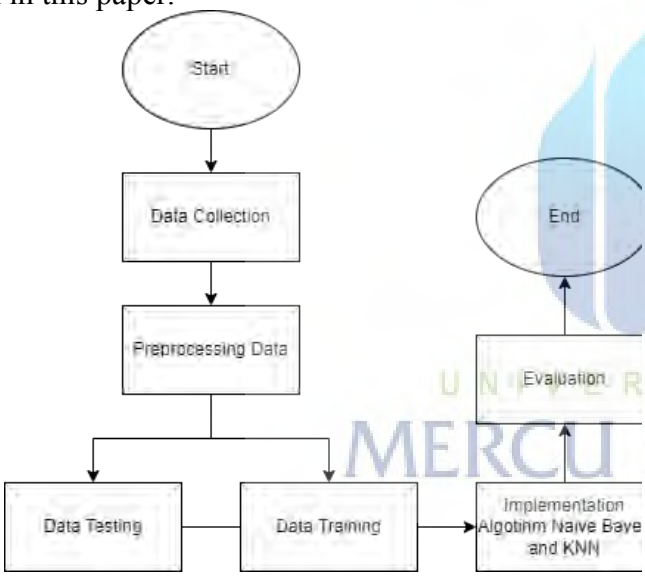

Figure 1. Research Method

A. Data Collection

The dataset is collected from transaction during 5-month period from January to May 2021. The dataset has a total of 2000 data. Table 1 shows the attributes of the dataset

Table 1. Dataset's Attributes

No	Attribute	Type	Information
1	No	*Int*	Numbers on food and beverage sales at Krintji Coffee
2	*Date*	*Date-time*	Customer Order date
3	*Item Name*	*Object*	The name of the menu in Krintji Coffee
4	*Item Variant Name*	*Object*	Krintji Coffee menu variants
5	*Category Name*	*Object*	Categories of Krintji Coffee menu
6	*Price*	*Int*	Prices from the Krintji Coffee menu
7	*Item Sold*	*Int*	Menus that have been sold
8	*Item Refunded*	*Int*	Cancelled items
9	*Gross Sales*	*Int*	Revenue from the menu sold

B. Data Preprocessing

Data preprocessing is a process that transforms raw data into a more understandable form [5]. Based on this research, the data pre-processing stage was carried out [6], including:

- Data Cleaning: checking blank values, duplicate data, and missing values [7].
- Data Reduction: this study does not use all attributes. The No, SKU, and Refund Item columns were removed, because it would interfere with the performance of the algorithm.
- Data transformation: change the data type on the *date* column from the original *string (object)* to the *datetime* data type; change the data type in the menu status column from the original string (object) to an integer. Next, change the data type on the column. Change the format and data type of the *Item Name* and *Category Name* columns from strings (objects) to integers. The results of preprocessing are presented in Table 2.

Table 2. Results of Data Preprocessing

No	Attribute	Data type	Data example
1	*Date*	Datetime	2021-01-01
2	*Item Name*	Int	31
3	*Item varian name*	Int	0
4	*Category name*	Int	22
5	*Price*	Int	200000
6	*Item sold*	Int	17
7	*Gross Sales*	Int	18000

C. Data Labeling

The result of the data labeling process on the Kopi Krintji cafe menu is presented in The Figure 2.

No	Date	Item Name	Item Variant Name	Category Name	SKU	price	Item Sold	Item Refunded	Gross Sales	Status Menu
1	2021-01-01 00:00:00	Mandailing	0	kopi single	NaN	20000	17	0	340000	Tidak Laris
2	2021-01-01 00:00:00	Teh panas	0	bukan kopi	NaN	12000	9	0	108000	Tidak Laris
3	2021-01-01 00:00:00	Bali Kintamani	0	kopi single	NaN	20000	10	0	200000	Tidak Laris
4	2021-01-01 00:00:00	Cheese cream drinks	Thai cream cheese	kopi di krintji	NaN	20000	8	0	160000	Tidak Laris
5	2021-01-01 00:00:00	Pisang original	Goreng	Cemilan	NaN	15000	11	0	165000	Tidak Laris

Figure 2. Data Grouping

D. Implementation of the Naïve Bayes and KNN Algorithms

After getting the desired data through the process of data pre-processing and data labeling, the two classifiers, Naïve Bayes and KNN are implemented. The value of independent variable x consists of *Item Name, Category Name, Price, Item Sold and Gross Sale,* and dependent value variable y as a classification label, which consists of menu status.

- Naïve Bayes Algorithm
  Bayes' method is a statistical approach to inducting inference on classification problems. It uses the basic concepts and definitions of Bayes' Theorem, then uses them to perform classifications in Data Mining [8]. Bayesian classification has similar classification capabilities to decision trees and neural networks [9] and is proven to have high accuracy and detection speed when applied to databases with large data [10].
- K–Nearest Neighbor (KNN) Algorithm
  K-Nearest Neighbor (KNN) is one of the algorithms used in classifying problems [11]. The working principle of KNN is to find the closest distance between the data to be evaluated with the nearest neighbor in the training data [12]. The KNN algorithm is one of the simplest algorithms for solving classification problems and often produces competitive and significant results [13].

E. Evaluation

In evaluating a classifier, it is necessary to analyze and measure the extent of the accuracy of the results that have been achieved by the classifier using the Confusion Matrix [14]. Then the evaluation results will be given in term of a classification report consisting of Accuracy, Precision and Recall values [15].

- $accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
- $precision = \frac{TP}{TP+FP}$
- $recall = \frac{TP}{TP+FN}$

Where TP is true positive, TN is false negative, FP is true positive, and FN is false negative.

III. EXPERIMENTAL RESULTS AND DISCCUSION

The experiments are carried out on Google Collaboratory platform using Python programming language and utilizing Pandas, Numpy, Seaborn, Matplotlib, and Sklearn libraries. The results of the experiment were obtained by dividing the percentage of *split validation* of the initial dataset into *training* data and *testing* data. In this division, *testing* data has a percentage of 20%, and *training* data is 80%. 5 (five) experiments using *k-fold cross validation* with the values of k are 5, 10, 15, 20 and 25.

A. Algoritma K–Nearest Neighbor

The following is the result of the visualization of the confusion matrix presented in Figure 3.

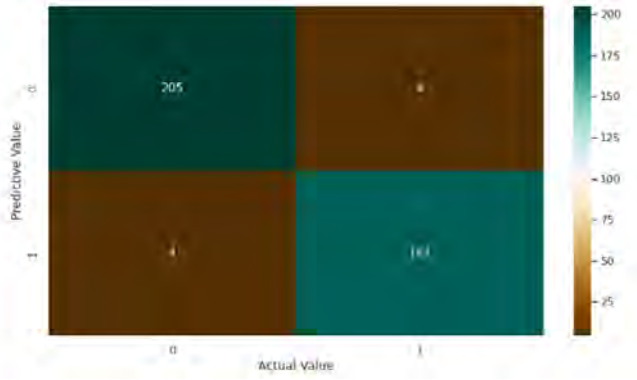

Figure 3. Confusion Matrix of the KNN

The following is an explanation of the *confusion matrix*.

1) *True positive* is positive data that has been classified correctly. The number of *true positive* in the experiment is 205.
2) *True negative* is negative data that successfully classified correctly. The number of *true* negatives that were successfully classified correctly in the KNN model is 187.
3) *False positive* is negative data but is classified as positive data. The number of *false positive* values in the KNN model reaches 4.
4) *False negative* is positive data but is classified as negative data. The number of *false negatives* is 4.

The results of the KNN algorithm are presented in Table 3.

Table 3. Experimental Results on KNN

Label	Metric			
	K	Accuracy	Precision	Recall
0	5	97%	98%	98%
1	5	97%	98%	98%

The results of the experiment using *the k-fold* values are presented in Table 4.

Table 4. KNN Accuracy Result of k-Fold cross validation

Experiment	Naïve Bayes	
	K-Fold	Accuracy
1	5 k-fold	93%
2	10 k-fold	98%
3	15 k-fold	93%
4	20 k-fold	94%
5	25 k fold	94%

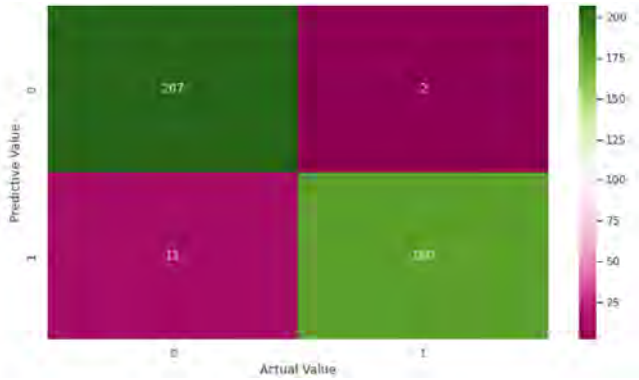

Figure 4. Confusion Matrix of Naïve Bayes

B. Naïve Bayes Algorithm

The following is the result of the visualization of the *confusion matrix* presented in Figure 4.

The confusion matrix of Naïve Bayes algorithm produces the following metric performance.

1) The number of *true positives* of the experiment is 207.
2) The number of *true* negatives that were successfully classified correctly in the KNN model is 180.
3) The number of *false positive* values in the KNN model reaches 2.
4) The number of *false negative* is 11.

The following are the results of the Naïve Bayes algorithm presented in Table 5. Naïve Bayes Algorithm results.

Table 5. Naïve Bayes Algorithm Results

Label	Metric		
	Accuracy	Precision	Recall
0	97%	95%	99%
1	97%	99%	93%

The results of cross validation experiment using Naïve Bayes algorithm are presented in Table 6.

Table 6. Result K – Fold Naïve Bayes

Experiment	K-NN K-Fold	Accuracy
1	5 k-fold	95%
2	10 k-fold	96%
3	15 k-fold	96%
4	20 k-fold	96%
5	25 k-fold	96%

C. Discussion

Based on the observation of the experimental results of the two algorithms on the testing data, some facts are revealed as follows.

- Overall, the accuracy value in the testing data for the K-Nearest Neighbors (KNN) Algorithm is 98%, which makes KNN algorithm a more appropriate algorithm in classifying the best-selling and not in-demand menus in Kopi Krintji cafe compared to Naïve Bayes algorithm which has a lower accuracy value, which only reaches 97%.

From the bar chart in Figure 5 we can find out the menu categories that are often ordered by Kopi Krintji café customers and also menu categories that are rarely or even never ordered by Krintji Coffee customers. It is worth to notice that the menu categories in the chart are not correlated with the statement whether a menu in Kopi Krintji sells (0) or Not Sells (1) but is just a category of several menus (precisely from the "Item Name" column) that are successfully sold every day.

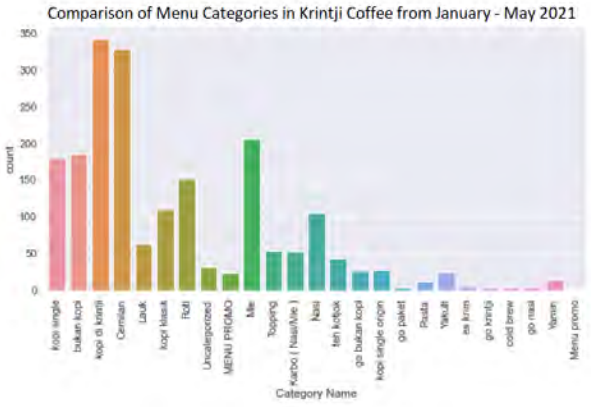

Figure 5. Kopi Krintji Menu Sales Category for the Period of January – May 2021

The following are the detailed insights obtained from the chart in Figure 8.

- The menu category that is always sold every day is the coffee menu category which has a total value of more than 300 (quite thin ahead of the menu category, namely Snacks), precisely coffee category has a value of 342.
- In addition to the coffee menu category, there are also other categories that are quite often ordered by the customers, i.e.: the Snack menu category which has a total value of 329, Noodles with a total value of 206, and a non-coffee category that has a total value of 186.
- On the other hand, the menu categories that are rarely or even never ordered by the customers are menu categories such as go package (4), cold brew (4), go Krintji (4), go rice (3), and Promo Menu (1).
- Finally, it can be seen that the menu at the Kopi Krintji cafe that are not in the interests of customers are go package, pasta, ice cream, go Krintji, cold brew and go rice.

IV. CONCLUSION

This study has shown the benefit of data science techniques in gaining insights from business data. Predicting a menu that sell well and is not in demand at the Kopi Krintji cafe helps the cafe owner in determining which menu should be kept and which menu should be dropped and replace them with more attractive ones. The insights revealed from the data analysis including the menu that always sold every day, i.e.: the coffee menu category, which has a total value of 342. On the other hand, the menu categories: go package, cold brew, go Krintji are unpopular menus. Thus, may be considered to be removed from the menu.

In addition, the KNN algorithm provides better accuracy result compared to the Naïve Bayes algorithm with an accuracy level of 98%, precision and recall of 98%. The cross validation provides the best value of accuracy with 10-fold at 98%. As for the Naïve Bayes algorithm, it has an accuracy

of 97%, precision of 95% and recall of 93%. Cross validation resulted in the best k-fold values of k-10, 15, 20 and 25 with an accuracy value of 96%. Thus, it can be seen that KNN and Naïve Bayes algorithms are equally good in classifying the best-selling and non-selling menus at Kopi Krintji café menus.

As for future work, authors plan to implement other classifiers to improve the accuracy of the classification, such as, support vector machine as well as artificial neural network.

References

[1] Y. N. Rin Hirakawa, "Study on Door to Door Sales.pdf." .

[2] X. Song, T. Xie, and S. Fischer, "A memory-access-efficient adaptive implementation of kNN on FPGA through HLS," *Proc. - 2019 IEEE Int. Conf. Comput. Des. ICCD 2019*, no. Iccd, pp. 177–180, 2019, doi: 10.1109/ICCD46524.2019.00030.

[3] S. Wang, J. Ren, and R. Bai, "A Regularized Attribute Weighting Framework for Naive Bayes," *IEEE Access*, vol. 8, pp. 225639–225649, 2020, doi: 10.1109/ACCESS.2020.3044946.

[4] E. S. Palupi, "Android Sales Prediction During Pandemic Using Naïve Bayes and K-Nn Methods Based on Particle Swarm Optimization," *J. Ris. Inform.* , vol. 4, no. 1, pp. 23–28, 2021, doi: 10.34288/jri.v4i1.279.

[5] J. Vieira, R. U. I. P. Duarte, and H. C. Neto, "kNN-STUFF : kNN STreaming Unit for Fpgas," vol. 7, 2019, doi: 10.1109/ACCESS.2019.2955864.

[6] A. Phinyomark, E. Ibanez-Marcelo, and G. Petri, "Resting-State fMRI Functional Connectivity: Big Data Preprocessing Pipelines and Topological Data Analysis," *IEEE Trans. Big Data*, vol. 3, no. 4, pp. 415–428, 2017, doi: 10.1109/tbdata.2017.2734883.

[7] L. Ma, Q. Pei, L. Zhou, H. Zhu, L. Wang, and Y. Ji, "Federated Data Cleaning: Collaborative and Privacy-Preserving Data Cleaning for Edge Intelligence," *IEEE Internet Things J.* , vol. 8, no. 8, pp. 6757–6770, 2021, doi: 10.1109/JIOT.2020.3027980.

[8] L. Jiang, H. Zhang, and Z. Cai, "A novel bayes model: Hidden naive bayes," *IEEE Trans. Knowl. Data Eng.* , vol. 21, no. 10, pp. 1361–1371, 2009, doi: 10.1109/TKDE.2008.234.

[9] D. Wang, D. Yuan, and C. Miao, "Sparse Naïve Bayes Base on Entropy Correlation for GPR Image Denoising," *2020 IEEE 3rd Int. Conf. Electron. Commun. Eng. ICECE 2020*, pp. 167–171, 2020, doi: 10.1109/ICECE51594.2020.9353029.

[10] Venkatesh and K. V. Ranjitha, "Classification and Optimization Scheme for Text Data using Machine Learning Naïve Bayes Classifier," *2018 IEEE World Symp. Commun. Eng. WSCE 2018*, pp. 33–36, 2019, doi: 10.1109/WSCE.2018.8690536.

[11] M. N. Ab Wahab, A. Nazir, A. T. Z. Ren, M. H. M. Noor, M. F. Akbar, and A. S. A. Mohamed, "Efficientnet-Lite and Hybrid CNN-KNN Implementation for Facial Expression Recognition on Raspberry Pi," *IEEE Access*, vol. 9, pp. 134065–134080, 2021, doi: 10.1109/ACCESS.2021.3113337.

[12] Y. Tan, "An Improved KNN Text Classification Algorithm Based on K-Medoids and Rough Set," *Proc. - 2018 10th Int. Conf. Intell. Human-Machine Syst. Cybern. IHMSC 2018*, vol. 1, pp. 109–113, 2018, doi: 10.1109/IHMSC.2018.00032.

[13] M. S. Sarma, Y. Srinivas, M. Abhiram, L. Ullala, M. S. Prasanthi, and J. R. Rao, "Insider threat detection with face recognition and KNN user classification," *Proc. - 2017 IEEE Int. Conf. Cloud Comput. Emerg. Mark. CCEM 2017*, vol. 2018-Janua, pp. 39–44, 2018, doi: 10.1109/CCEM.2017.16.

[14] Y. K. Gu, B. Xu, H. Huang, and G. Qiu, "A Fuzzy Performance Evaluation Model for a Gearbox System Using Hidden Markov Model," *IEEE Access*, vol. 8, pp. 30400–30409, 2020, doi: 10.1109/ACCESS.2020.2972810.

[15] Q. Xue, Y. Zhu, and J. Wang, "Joint Distribution Estimation and Naïve Bayes Classification under Local Differential Privacy," *IEEE Trans. Emerg. Top. Comput.* , vol. 9, no. 4, pp. 2053–2063, 2021, doi: 10.1109/TETC.2019.2959581.

KERTAS KERJA

Ringkasan

Kertas kerja ini berisi tentang kelengkapan material dari artikel jurnal dengan judul "Komparasi Algoritma K-Nearest Neighbor dan Naïve Bayes untuk Klasifikasi Menu Laris dan Tidak Laris di Kedai Kopi Krintji". Seluruh hasil penelitian Tugas Akhir yang tidak dimasukkan kedalam artikel jurnal. Pada kertas kerja ini disajikan terdiri dari literatur review, analisis perancangan, source code, dataset yang digunakan, tahapan eksperimen dan seluruh hasil eksperimen.

- Bagian 1 : Literature Review menjabarkan jurnal yang terkait dengan penelitian.
- Bagian 2 : Analisis dan Perancangan menjelaskan masalah terhadap penelitian.
- Bagian 3 : Source Code menyajikan kumpulan kode disetiap proses yang dilakukan peneliti terdiri dari membaca data, preprocessing data, data visualisasi, data modelling, implementasi algoritma.
- Bagian 4 : Pengambilan data atau *data collection* untuk penelitian.
- Bagian 5 : Tahapan Eksperimen berisi pengumpulan data, preprocessing data, data training data testing, impelementasi algoritma, dan evaluasi.
- Bagian 6 : Hasil semua eksperimen yang dilakukan pada penelitian secara keseluruhan yang mencakup skenario pengujian.