SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

Performance Comparison between K-Nearest Neighbor (KNN) Method and Naïve Bayes Method in Reward for Motorcycle Salesman

Muhammad Ayyasi Fawaz [1)*] , Khairul [2)] , Andysah Putera Utama Siahaan [3)]
[1,2,3)] Master of Information Technology Study Program, Panca Budi Development University, Indonesia
[1)] ayyasfawaz@gmail.com, [2]khairul@dosen.pancabudi.ac.id, [2] andiesiahaan@gmail.com

Submitted :July 1, 2024 | **Accepted** : July 28, 2024 | **Published** : July 31, 2024

Abstract: Motorcycles face challenges in boosting sales and maintaining employee loyalty. To address this, the company offers a reward salesman tour for employees who meet certain criteria. However, the current evaluation system is too simple and does not fully capture the quality of employees, especially their product knowledge and involvement in company campaigns. This study aims to solve these issues using data mining techniques, specifically the Naïve Bayes and K-Nearest Neighbors (KNN) methods. These methods were chosen for their accuracy and simplicity. The K-Nearest Neighbor method (K=11) showed an accuracy of 94.04%, a precision of 83.78%, and a recall of 96.87%, while the Naïve Bayes method showed an accuracy of 81.81%, a precision of 72.00%, and a recall of 81.25%.

Keywords: K Nearest Neighbor, Naïve Bayes, Machine Learning, Reward, Motorcycles

INTRODUCTION

PT. Honda Indako Trading Coy Krakatau is a business in the automotive field (cars, motorcycles, and power generators) and Johnson (speedboat engines) as well as their spare (https://indako.id/about/). Along with the company's current development, the company must continuously develop products and disseminate information to the public extensively.

The success and existence of the company can be seen through the number of products sold, as this is the main source of income that allows the company to grow, develop, and sustain its operations. With increasing sales, the company can increase its market share, expand product reach, strengthen customer relationships, build a strong brand, and create a fundamental competitive advantage. This can happen if the human resources (employees) possess the necessary qualities to sell a product. This quality becomes a factor in increasing the productivity of the company's performance. Therefore, to improve the quality and loyalty of employees annually rewards salesmen tours to employees who meet the criteria set by the management (Abdul Koda et al., 2022).

Rewarding employees is a form of appreciation from the company for their dedication to working with good quality. It also boosts employee morale and fosters to work even harder. Thus, the company creates more competent employees to carry out their duties (Faran & Triayudi, 2024). However, challenges such as overly simplistic assessment mechanisms have been found, meaning the winners do not fully represent truly high-quality employees. This relates to how well employees know the product and participate in helping the company's campaigns locally and nationally.

This issue can be addressed by observing patterns based on past processes or data. With rapid technological development, computers can process large data patterns using certain methods (Nasyuha, 2019). Data mining is a suitable way to solve this problem. Data mining is a process of data processing and extraction to retrieve information from collection data. In data mining, the process is conducted based on data collected in the past. This data is stored and organized in a data warehouse or database. The stored data is then reprocessed to obtain valuable and important new information (Faran & Triayudi, 2024)

Several methods, such as Naïve Bayes and K-Nearest Neighbors, can solve this problem. Naïve Bayes is a supervised learning method known for accuracy with simple calculations (Nasien et al., 2024). It performs statistical grouping that estimates membership likelihood using experience to predict what will happen in the future (Sholekhah et al., 2024). K-Nearest Neighbor functions to classify new subjects based on training sample data and attributes. It works based on the closest distance from training data to test data and then takes the majority to make predictions from the test sample. These methods are chosen for their high performance in classifying data, simplicity in computation, and good accuracy. The results of both methods will be compared to determine which method performs better in solving this problem.

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

LITERATURE REVIEW

Machine Learning

Machine learning is a branch of computer science that can operate without explicit programming. Many researchers think about ways to make AI perform at human levels. Machine learning is an aspect of artificial intelligence that studies how to process data. By definition, machine learning is the science or study of algorithms and statistical models used by computer systems to perform specific tasks without explicit instructions. It relies on patterns and inference.

The application of these techniques is related to machine learning and AI. The term 'machine' refers to algorithms or programs running on computers. Therefore, all knowledge inevitably involves data (Dinata & Hasdyna, 2020). The term "supervisor" here refers to labels on each piece of data. Labels are tags added to data in a model. Unsupervised learning does not use labels to predict target features/variables. Instead, it relies on the similarities of the attributes. If the attributes and properties of the extracted feature data have similarities, they will be grouped (clustering). Clustering is machine learning without regular supervision, where the data set must be automatically partitioned into clusters so that objects in the same cluster are more similar, while objects in different clusters are different (Nasution et al., 2022)

Data Mining

Data mining and knowledge discovery in databases are often used interchangeably to describe extracting hidden information from large databases. One of the stages in the overall KDD (Knowledge Discovery in Databases) process is data mining. Process of data mining, often referred to as knowledge discovery in databases, the application of scientific methods in data mining is crucial. Therefore, data mining can also be described as exploring and analyzing databases to find interesting patterns to accurately and potentially extract information and knowledge (Khairul et al., 2023).

K Nearest Neighbor

The K-Nearest Neighbor (KNN) method is a classification technique that determines the category based on the majority of nearby neighbors. KNN is performed by finding the group of k objects in the training data that looks similar to the object in the new or testing data. In this context, the optimal value of k for KNN depends on the data; generally, a higher value of k will reduce the effect of noise in classification but can also make the classification boundaries less distinct. An illustration of the use of the k value in the method is presented in Figure 1. The ideal k value is obtained through parameter optimization. The nearest neighbor algorithm is used in certain situations where classification predictions are based on the closest training data (in other words, k=1) (Ramadhan et al., n.d.)

1. Determine the parameter K to be used in the K-NN calculation.
2. Calculate the distance between the test data and the training data using a distance matrix, such as the Euclidean distance, as formulated in Equation 1
3. Sort the results based on the Euclidean distance values.
4. Determine the results from the K nearest neighbors.
5. The target output is the majority class.

(Nijunnihayah & Hilabi, 2024)

$$d(x,y) = \sqrt{\Sigma_{i=1}^{n}(x_i - y_i)^2}$$

d(x,y) = distance between data point x and data point y
xi = value of the feature in the training data
yi = yi value in testing data
n = value of the feature in the testing data
i = dimension of the data

Naïve Bayes

Naïve Bayes is a classification method that applies the principles of probability and statistics, introduced by the English scientist Thomas Bayes (Aulia et al., 2023). This method particularly uses probabilistic estimates based on prior experiences. Naïve Bayes utilizes statistical knowledge, including the application of probability theory, to handle supervised learning cases, where there is a dataset with attributes, classes, or labels as a basis for

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

information. One of the main advantages of using the Bayesian method is found in classical integral methods for obtaining a marginal model (Aini et al., 2021).

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Explanation as follows :

- P(B|A) = Conditional probability of B given A
- P(A|B) = Conditional probability of A given B
- P(A) = Probability of event A
- P(B) = Probability of event B (Khalaf et al., 2024)

Reward Salesman Tour

Employees are a crucial factor in any company, as having employees who meet the company's qualification standards ensures that the company's productivity is maintained and improved (Supiyandi et al., 2020). In addition, employees are a key asset for a company in its efforts to maintain operational continuity, growth, competitiveness, and achieve optimal profits. Consequently, the management of employees and the evaluation of employee performance are essential in determining the effectiveness of the HR management system in achieving the company's goals, vision, and mission (Nabilah et al., 2024)

METHOD

This study used quantitative research, which is a type of research that produces knowledge that can be obtained through the use of statistical techniques or other quantitative (measurement) means. Quantitative research can also be defined as an investigation into social issues based on the testing of a theory consisting of variables, measured in numbers, and analyzed using statistical procedures to determine whether the theory's predictive generalizations are accurate. According to Punch (1988), quantitative research is empirical research where the data are in the form of numbers that can be counted. This type of research focuses on collecting and analyzing numerical data (Ali et al., 2022). The form contains criteria aspects related to this case :

Table 1. Aspects of Salesman Reward Criteria

CODE	CRITERIA
A1	Matic Motorcycle
A2	Bebek Motorcycle
A3	Sport Motorcycle
A4	Premium Motorcycle
A5	New Product Knowledge
A6	Qualified Sales Personel
A7	My Hero App

Then, by the research objectives that have been set, the researcher wants to conduct a comparative and the conceptual framework above can be described as follows :

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

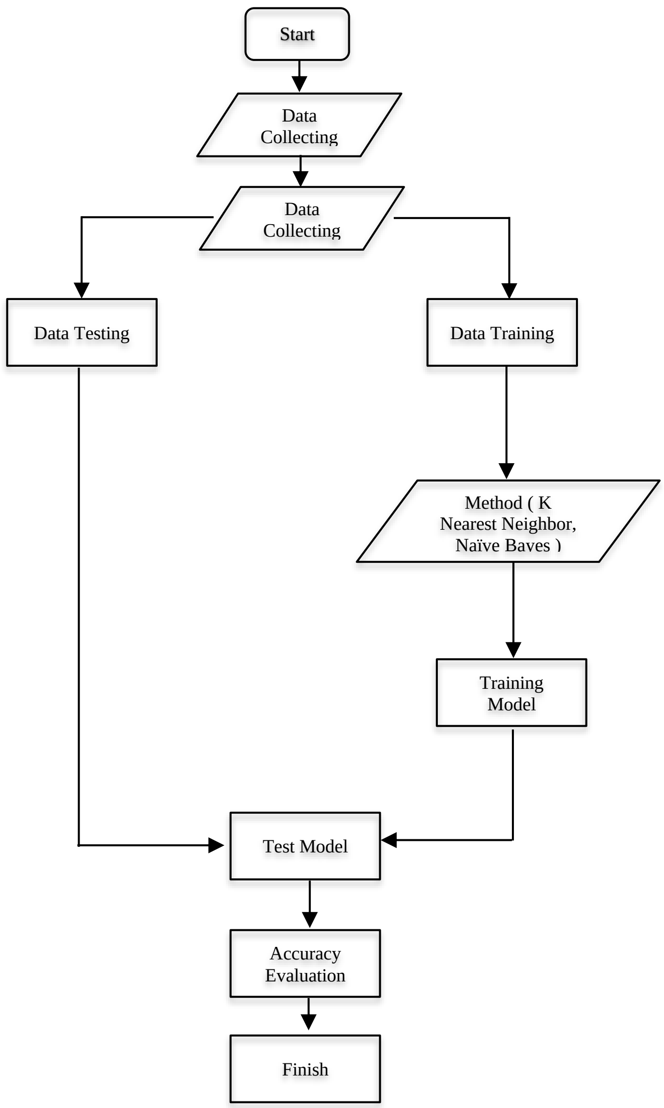

Figure 1. Research Conceptual Framework

The first step is to collect the data to be analyzed through observation until the data is successfully gathered. Then, the data will be split into training data, which will be the knowledge base for the system, and testing data, which will be used for testing purposes. The testing data functions to evaluate whether the system's predictions are accurate. Next, the K-Nearest Neighbor and Naïve Bayes methods are applied to the training data, generating models for each method. Once both models are created, they will be tested using the previously separated testing data. The results will produce performance values for each method, which will then be evaluated for accuracy and compared to determine which method performs better.

RESULT

K Nearest Neighbor Calculation Process

K- Steps to solve in determining the reward achievement for the salesman tour using the K-Nearest Neighbor method are as follows:

Table 2. Motorcycle Sales Data

Data	A1	A2	A3	A4	A5	A6	A7	STATUS
D1	72	24	6	0	71	96	10	Didn’t Receive Award
D2	108	96	12	4	84	77	30	Receive Award

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

D3	12	120	12	4	71	85	30	Didn't Receive Award
D4	144	96	9	8	81	68	10	Receive Award
D5	120	48	15	4	95	72	40	Didn't Receive Award
D6	60	120	3	8	69	65	40	Didn't Receive Award
D7	108	96	3	4	79	85	20	Didn't Receive Award
D8	108	144	3	0	90	77	10	Receive Award
D9	120	96	6	0	75	97	30	Receive Award
D10	120	72	0	0	97	92	20	Receive Award
…	…	…	…	…	…	…	…	…
D89	108	120	0	8	70	78	20	Receive Award

1. Determining Parameter K
The parameter is a variable that determines the number of nearest neighbors that will be used in the classification process. The parameter K can affect the accuracy of the classification. In this study, since the data is even in number, the parameter K will be an odd number. The values of parameter K in this study include 5, 8, and 11.

2. Calculating the distance between the training data and the testing data using the Euclidean Distance matrix
Training Data against Test Data taken from Training Data
D1 against D3 :

$$d(1,3) = \sqrt{(72-12)^2 + (24-120)^2 + (6-12)^2 + (0-4)^2 + (71-71)^2 + (96-85)^2 + (10-30)^2}$$
$$= 113{,}8815174$$

3. Sorting based on the Euclidean Distance Value

Table 3. Euclidean Distance Results

Data	Euclidean Distance	The Rank of proximity distance	Parameter (K=5)	Parameter (K=8)	Parameter (K=11)
1	113,8815174	74			
2	100,124922	55			
3	0	1			
4	135,6244816	86			
5	132,7102106	84			
6	53,05657358	13			
7	99,63433143	54			
8	101,4593515	58			
9	111,5885299	72			
10	121,8318513	76			
11	77,4209274	37			
12	123,1624943	79			
13	107,939798	64			
14	108,1850267	65			
15	69,20982589	28			
16	55,91064299	16			
17	39,67366885	6	Didn't get a reward	Didn't get a reward	Didn't get a reward

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

18	63,69458376	22			
19	78,12169993	40			
20	50,43808085	12			Didn't get a reward
21	56,02677931	17			
22	77,81388051	38			
23	111,3956911	71			
24	123,3450445	80			
25	78,17928114	41			
26	66,29479618	24			
27	27,83882181	2	Didn't get a reward	Didn't get a reward	Didn't get a reward
28	72,13875519	29			
29	136,0882067	87			
30	67,46851117	26			
31	103,1939921	62			
32	55,09083408	14			
33	111,2115102	69			
34	87,85783972	46			
35	66,03029608	23			
36	99,50879358	53			
37	69,15923655	27			
38	102,9708697	61			
39	75,94735018	34			
40	77,1362431	36			
41	47,8225888	10			Didn't get a reward
42	102,2594739	60			
43	73,32803011	30			
44	55,6596802	15			
45	66,57326791	25			
46	76,70071708	35			
47	62,87288764	21			
48	114,0657705	75			
49	36,0970913	4	Didn't get a reward	Didn't get a reward	Didn't get a reward
50	50,11985634	11			Get Reward
51	75,43208866	33			
52	78,55571271	42			
53	46,17358552	7		Get Reward	Get Reward
54	122,1106056	77			
55	86,75828491	45			

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024 e-ISSN: 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.13935 p-ISSN: 2541-044X

56	130,3341858	83			
57	142,6429108	88			
58	111,0045044	68			
59	111,2115102	69			
60	32,49615362	3	Didn't get a reward	Didn't get a reward	Didn't get a reward
61	88,10788841	47			
62	96,43650761	51			
63	104,0432602	63			
64	86,24384036	44			
65	133,7871444	85			
66	83,76156637	43			
67	62,09669879	20			
68	108,3697375	66			
69	100,8761617	56			
70	101,1879439	57			
71	91,5259526	50			
72	101,6366076	59			
73	77,9166221	39			
74	60,4814021	19			
75	124,8679302	82			
76	122,8006515	78			
77	59,5986577	18			
78	74,22937424	32			
79	123,7739876	81			
80	47,02127178	8		Didn't get a reward	Didn't get a reward
81	91,25239723	49			
82	113,4239834	73			
83	110,8873302	67			
84	47,70744177	9		Didn't get a reward	Didn't get a reward
85	89,98333179	48			
86	153,0816775	89			
87	37,08099244	5	Didn't get a reward	Didn't get a reward	Didn't get a reward
88	73,64781056	31			
89	97,03607577	52			

*name of corresponding author

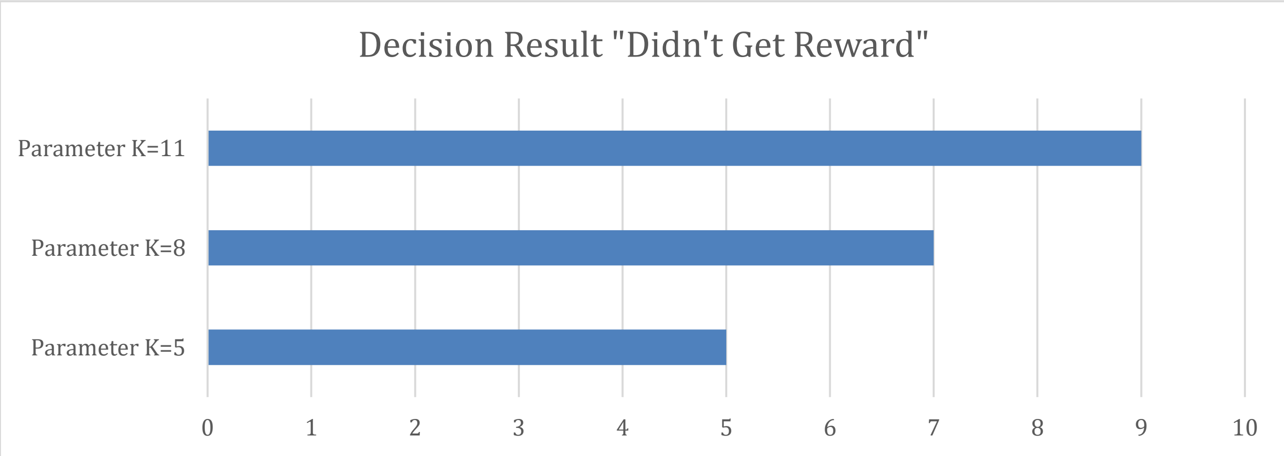

Figure 2. Decision Result "Didn't Get Reward"

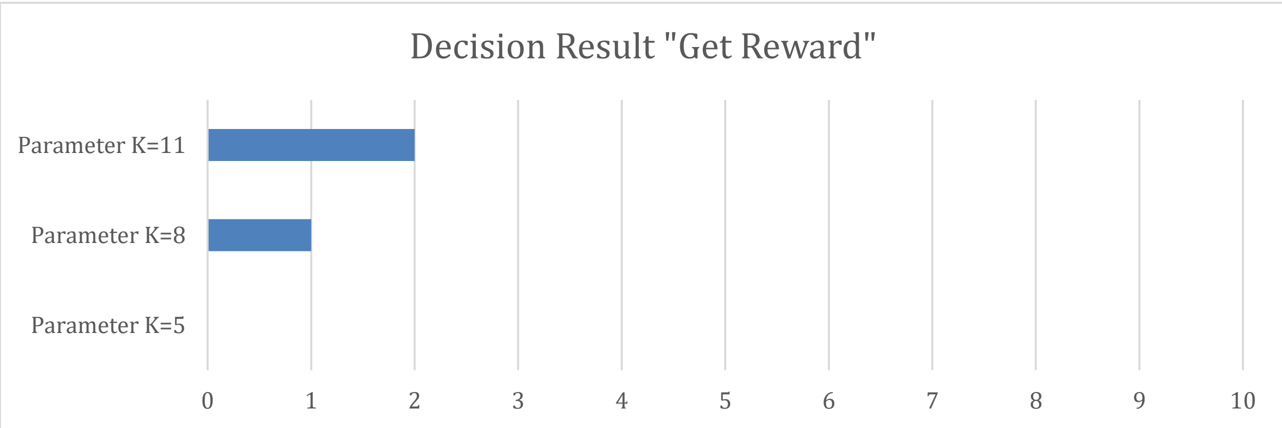

Figure 3. Decision Result " Get Reward"

4. Determining the result from the nearest K neighbors
The value of K is determined by the nearest distance, where the smaller the number, the closer the distance.

5. Target output is the majority class
Training Data against Test Data taken from Training Data
The value of K in this part is divided into 3 numbers: K=5, K=8, and K=11. For K=5, the ratio of those who received the reward to those who did not is 0:5; for K=8, the ratio is 1:7; and for K=11, the ratio is 2:9
6. Accuracy Evaluation for K-Nearest Neighbor
The following steps involve testing by examining the results from the Confusion Matrix:
Training Data against Test Data taken from Training Data

Table 4. Confusion Matrix K Nearest Neighbor (K=5)

	Class	
	Actual Yes	Actual No
Predicted Yes	31	6
Predicted No	1	50

Accuracy = (TP + TN) / (TP + TN + FP + FN) * 100 %
= (31 + 50) / (31 + 50 + 6 + 1) * 100 %
= (81 / 88) * 100 %
= 0,92045 * 100 %
= 94,04 %

Precision = (TP / (TP + FP)) * 100 %

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

= (31 / (31 + 6)) * 100 %
= (31 / 37) * 100 %
= 0,83783 * 100 %
= 83,78 %

Recall = (TP / (TP + FN)) * 100%
= (31 / (31 + 1)) * 100 %
= (31 / 32) * 100 %
= 0,96875 * 100 %
= 96,87 %

Table 5. Confusion Matrix K Nearest Neighbor (K=8)

	Original Class	
	Actual Yes	Actual No
Predicted Yes	31	5
Predicted No	1	51

Accuracy = (TP + TN) / (TP + TN + FP + FN) * 100 %
= (31 + 51) / (31 + 51 + 5 + 1) * 100 %
= (82 / 88) * 100 %
= 0,93181 * 100 %
= 93,181 %

Precision = (TP / (TP + FP)) * 100 %
= (31 / (31 + 5)) * 100 %
= (31 / 36) * 100 %
= 0,86111 * 100 %
= 86,111 %

Recall = (TP / (TP + FN)) * 100%
= (31 / (31 + 1)) * 100 %
= (31 / 32) * 100 %
= 0,96875 * 100 %
= 96,87 %

Table 6. Confusion Matrix K Nearest Neighbor (K=11)

	Class	
	Actual Yes	Actual No
Predicted Yes	31	6
Predicted No	1	50

Accuracy = (TP + TN) / (TP + TN + FP + FN) * 100 %
= (31 + 50) / (31 + 50 + 6 + 1) * 100 %
= (81 / 88) * 100 %
= 0,92045 * 100 %
= 94,04 %

Precision = (TP / (TP + FP)) * 100 %
= (31 / (31 + 6)) * 100 %
= (31 / 37) * 100 %
= 0,83783 * 100 %
= 83,78 %

Recall = (TP / (TP + FN)) * 100%
= (31 / (31 + 1)) * 100 %
= (31 / 32) * 100 %
= 0,96875 * 100 %
= 96,87 %

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

Naïve Bayes Calculation Process

Naïve Bayes Steps to solve in determining the reward achievement for the salesman method are as follows and we use the same dataset as K-Nearest Neighbor :

1. Probability Calculation for Each Class:
Training Data against Test Data taken from Training Data

Probability of the Class "Receiving Reward"
Total data receiving reward / Total number of all data
= 32 / 88
= 0.36364

Probability of the Class "Not Receiving Reward"
Total data not receiving reward / Total number of all data
= 56 / 88
= 0.63636

2. Determining the Mean Value of Each Class for Each Criterion:
Training Data against Test Data taken from Training Data

Determining the mean value of the class "Receiving Reward" for the criterion Matic
Total transaction data for the criterion Matic with the status of receiving reward / Total data receiving reward
= 3480 / 32
= 108.75

Determining the mean value of the class "Not Receiving Reward" for the criterion Matic
Total transaction data for the criterion Matic with the status of not receiving reward / Total data not receiving reward
= 3468 / 56
= 61.92857

3. Determining the Standard Deviation of Each Class for Each Criterion:
Training Data against Test Data taken from Training Data

Determining the standard deviation of the class "Receiving Reward" for the criterion Matic

$$= \frac{(108-108{,}75)^2+(144-108{,}75)^2+(108-108{,}75)^2+ \dots\dots\dots + (108-108{,}75)^2}{32-1}$$

= 25,444793

Determining the standard deviation of the class "Not Receiving Reward" for the criterion Matic

$$= \frac{(72-61{,}92857)^2+(12-61{,}92857)^2+(120-61{,}92857)^2+ \dots\dots\dots + (24-61{,}92857)^2}{56-1}$$

= 30,40668

Recapitulation of the Mean and Standard Deviation Values :

Table 7. Results Recapitulation of Mean and Standard Deviation Values

	Mean (Matic)	**Variance (Matic)**	**Mean(Bebek)**	**Variance(Bebek)**
Get an Award	108,75	25,44479318	120	24
Didn't get an award	61,92857143	30,40668227	61,92857143	33,96006058

Mean(Sport)	**Variance (Sport)**	**Mean (Premium)**	**Variance(Premium)**
6,5625	5,431030634	2,625	3,257203555
6,321428571	5,217098067	3,571428571	3,17837078

Mean (PK)	Variance(PK)	Mean (QS)	Variance(QS)	Mean (MH)	Variance(MH)
82,40625	10,31945788	82,875	10,17272702	20	3,535533906
83,91071429	9,998708463	80,35714286	8,665260221	21,96428571	6,985606724

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

4. Determining the Gaussian Distribution Value for Test Data:
Training Data against Test Data taken from Training Data

Determining the Gaussian distribution value of the class “Receiving Reward” for the criterion Matic

$$\frac{1}{\sqrt{2 * 3{,}14 * 25{,}44479}} - exp^{\frac{(12-108{,}75)^2}{2*25{,}44479^2}}$$

= 1,13746

Determining the Gaussian distribution value of the class “Not Receiving Reward” for the criterion Matic

$$\frac{1}{\sqrt{2 * 3{,}14 * 30{,}40668}} - exp^{\frac{(12-61{,}92857)^2}{2*30{,}40668^2}}$$

= 0,00340

The results obtained are as follows:

Table 8. Gaussian Distribution Value

	Matic	Motorbike	Sport	Premium	New Product Knowledge	Qualified Sales People	Apps My Hero	Status Award
	12	120	12	4	71	85	30	
Get Award	1,13746E-05	0,1662681	0,044511584	0,112067307	0,020992781	0,038380214	0,002067223	5,71385 E-16
Didn't Get Award	0,003408536	0,002723313	0,042298977	0,124413494	0,017339139	0,039893239	0,029476557	6,33828 E-13

5. Determining Reward Status:
By combining all of the criteria into one group, "receiving reward" classes are multiplied by their respective probabilities, the final step is to add the whole thing up

6. Evaluating Naïve Bayes Accuracy
The next step involves testing by looking at the results from the Confusion Matrix:

Training Data against Test Data taken from Training Data

Table 9. Confusion Matrix Naive Bayes

	Actual Yes	**Actual No**
Predicted Yes	26	10
Predicted No	6	46

Accuracy = (TP + TN) / (TP + TN + FP + FN) * 100 %
= (26 + 46) / (26 + 46 + 10 + 6) * 100 %
= (72 / 88) * 100 %
= 0,81818 * 100 %
= 81,81 %

Precision = (TP / (TP + FP)) * 100 %
= (26 / (26 + 10)) * 100 %
= (26 / 36) * 100 %
= 0,72222 * 100 %
= 72 %

Recall = (TP / (TP + FN)) * 100%

*name of corresponding author

SinkrOn: Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 3, July 2024
DOI : https://doi.org/10.33395/sinkron.v8i3.13935
e-ISSN: 2541-2019
p-ISSN: 2541-044X

= (26 / (26 + 6)) * 100 %

= (26 / 32) * 100 %

= 0,8125 * 100 %

= 81,25 %

DISCUSSION

This research focuses on the performance between the K-Nearest Neighbor (KNN) method and Naïve Bayes in predicting a salesman's reward in a sales tour. This study attempts to answer several questions about how to analyze the reward performance for a motorcycle sales tour by using the K-Nearest Neighbor (KNN) method, then using the Naïve Bayes method, and finally comparing the performance values between the K-Nearest Neighbor (KNN) method and Naïve Bayes in determining the reward performance for the motorcycle sales tour.

Table 10. *K-Nearest Neighbor* and *Naïve Bayes* Algorithms Comparison of Confusion Matrix

COMPARATIVE	Algorithms				
	K-NN (K=5)	K-NN (K=8)	K-NN (K=11)	Naïve Bayes	
Accuracy Value	94,04%	93,18%	94,04%	81,81%	
Recall Value	83,78%	86,11%	83,78%	72,00%	
Precision Value	96,87%	96,87%	96,87%	81,25%	

CONCLUSION

This research compares two algorithms, K-Nearest Neighbors and Naive Bayes, to evaluate the performance of a salesman motorcycle to get a reward. The results show that the K-Nearest Neighbor (K=11) showed an accuracy of 94.04%, a precision of 83.78%, and a recall of 96.87%, while the Naïve Bayes showed an accuracy of 81.81%, a precision of 72.00%, and a recall of 81.25%. This evaluation shows the high effectiveness of both algorithms in assessing performance, with K-Nearest Neighbor showing a higher result

REFERENCES

Abdul Koda, Rahayu, P., Pratama, A., Rafly, A., & Kaslani. (2022). Penentuan Bonus Karyawan Dengan Menggunakan Algoritma K-Nearest Neighbor. *KOPERTIP : Jurnal Ilmiah Manajemen Informatika Dan Komputer*, *4*(1), 14–20. https://doi.org/10.32485/kopertip.v4i1.115

Aini, N., Handoko, W., & Nurhaliza, R. (2021). *PREDIKSI PENERIMAAN BANTUAN PIP PADA SMKS AL-FURQON BATUBARA DENGAN METODE NAÏVE BAYES*. *1*(3), 219–226.

Ali, M. M., Hariyati, T., Pratiwi, M. Y., & Afifah, S. (2022). Metodologi Penelitian Kuantitatif dan Penerapannya dalam Penelitian. *Education Journal*, *2*(2), 1–6.

Aulia, A., Tanjung, M. R., Iqbal, M., & Wijaya, R. F. (2023). Implementasi Algoritma Naïve Bayes Dalam Menganalisis Jumlah Live Stream VTuber Skem. *Bulletin of Information Technology (BIT)*, *4*(4), 554–559. https://doi.org/10.47065/bit.v4i4.899

Dinata, R. K., & Hasdyna, N. (2020). *Machine Learning.pdf* (p. 23).

Faran, J., & Triayudi, A. (2024). Penerapan Algoritma K-Means Data Mining untuk Clustering Kinerja Karyawan Koperasi. *Penerapan Algoritma K-Means Data Mining Untuk Clustering Kinerja Karyawan Koperasi*, *4*(4), 2096–2108. https://doi.org/10.30865/klik.v4i4.1728

Khairul, K., Nasyuha, A. H., Ikhwan, A., H. Aly, M., & Ahyanuardi, A. (2023). Implementation of Multiple Linear Regression to Estimate Profit on Sales of Screen Printing Equipment. *Jurnal Infotel*, *15*(2), 55–61. https://doi.org/10.20895/infotel.v15i2.934

Khalaf, M. H., Sari, H. L., & Fredricka, J. (2024). *Sistem Pakar Mendiagnosis Penyakit Rhinosinusitis Dengan Menggunakan Metode Naïve Bayes*. *20*(1), 86–97.

Nabilah, J., Syahputra, A., & Arifitama, B. (2024). *Penilaian Kinerja Tenaga Pemasaran Untuk Menentukan Reward dan Benefit dengan Menggunakan Metode Weighted Product*. *4*(4), 2262–2271. https://doi.org/10.30865/klik.v4i4.1704

Nasien, D., Darwin, R., Cia, A., Leo Winata, A., Go, J., Charles Wijaya, R., & Charles Lo, K. (2024). *Perbandingan Implementasi Machine Learning Menggunakan Metode KNN, Naive Bayes, Dan Logistik Regression Untuk Mengklasifikasi Penyakit Diabetes*. *4*(1).

Nasution, D., Sirait, D. N., Wardani, I., & Dwiyanto. (2022). Optimasi Jumlah Cluster Metode K-Medoids Berdasarkan Nilai DBI Pada Pengelompokkan Data Luas Tanaman Dan Produksi Kelapa Sawit Di Sumatera

Utara. *Kumpulan JurnaL Ilmu Komputer (KLIK)*, *9*(2), 381.

Nasyuha, A. H. (2019). Sistem Pendukung Keputusan Menentukan Pemberian Pinjaman Modal dengan Metode Multi Attribute Utility Theory. *Jurnal Media Informatika Budidarma*, *3*(2), 117. https://doi.org/10.30865/mib.v3i2.1093

Nijunnihayah, U., & Hilabi, S. S. (2024). *Implementation of the K-Nearest Neighbor Algorithm to Predict Sales of Medical Devices in Medical Devices Implementasi Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Alat Kesehatan pada Media Alkes*. *4*(April), 695–701.

Ramadhan, I., Elmi, S., Efrizoni, L., Ramadhan, I., Elmi, S., & Efrizoni, L. (n.d.). *Implementasi algoritma k-nearest neighbor (knn) untuk prediksi bencana gunung berapi*. *01*, 58–65.

Sholekhah, F., Putri, A. D., Rahmaddeni, R., & Efrizoni, L. (2024). Perbandingan Algoritma Naïve Bayes dan K-Nearest Neighbors untuk Klasifikasi Metabolik Sindrom. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(2), 507–514. https://doi.org/10.57152/malcom.v4i2.1249

Supiyandi, S.-, Siahaan, A. P. U., & Alfiandi, A. (2020). Sistem Pendukung Keputusan Pemilihan Pegawai Honorer Kelurahan Babura dengan Metode MFEP. *Jurnal Media Informatika Budidarma*, *4*(3), 567. https://doi.org/10.30865/mib.v4i3.2107

*name of corresponding author