Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 215 - 220
DOI: https://doi.org/10.33330/icossit.v1i1.776
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT

ISSN 2723-4509 (Online)

DATA MINING IMPLEMENTATION FOR PRINTER SALES PREDICTION USING NAIVE BAYES METHOD

Rahayu Mayang Sari[1], Yori Apridonal M[2*]
[1]Computer Systems, Universitas Pembangunan Panca Budi Medan, Indonesia
[2]Information Systems, Sekolah Tinggi Manajemen Informatika dan Komputer Royal, Indonesia

Corresponding author:
yori@royal.ac.id

Keywords:
data mining
naïve bayes
prediction
printer sales

ABSTRACT

Sales is an important element in a business, for the sake of the smooth running of the business is expected to smooth sales and get large profits. The RCA Computer store is one of the computer shops engaged in the sale of printers in the city of Kisaran. Because of the current co-19 pandemic that hit the RCA Computer Range Store can not predict printer sales in the future. Data mining is one of the many fields of science that is growing rapidly, where data mining is able to process large data into information. Naïve Bayes predicts future opportunities based on past experience. The results of this study the Naive Bayes method can make predictions from printer sales data and provide the highest number of printer type predictions from the highest value of the type of printer based on brand attributes, type attributes, and payment attributes.

INTRODUCTION

In line with the development of information technology at the moment, data mining is one of the many fields of science that is growing rapidly, where data mining is able to process large data into information [1]. Another definition says that data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and related knowledge from various large databases [2].

Naïve Bayes is one of the most effective and efficient inductive learning tools for machine learning and data mining [3]. Another thing says Naive Bayes is a classification with probability and statistical methods raised by the British scientist Thomas Bayes, which predicts future opportunities based on previous experience so that it is known as Bayes theorem. The theorem is combined with naive where it is assumed that conditions between attributes are independent. The Naive Bayes classification is assumed that the presence or absence of certain characteristics of a class has nothing to do with the characteristics of other classes [4].

Prediction or forecasting is a technique for estimating expected demand for a product or activity for several periods in the future [5]. Forecasting can also be interpreted as a method used to predict future uncertainty in an effort to make better decisions [6].

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 215 - 220
DOI: https://doi.org/10.33330/icossit.v1i1.776
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

Matrix Laboratory (Matlab) is software that uses a matrix basis in its utilization. Matrix used in Matlab is fairly simple so it can be easily used. General Matlab is for: a). Mathematics and computing; b). Development and algorithm; c). Modeling, simulation and prototype making; d). Data analysis, exploration and visualization; e). Making applications including creating GUI (Graphical User Interface) [7].

Matlab-based programming for predicting printer sales makes users easier because it only includes a few variables from the sample to calculate parameters. Matlab programming used is created using the GUI (Graphic User Interface) application [8].

Sales is an important element in a business, for the sake of the smooth running of the business is expected to smooth sales and get large profits. The RCA Computer store is one of the computer shops engaged in the sale of printers in the city of Kisaran. Because of the current co-19 pandemic that hit the RCA Computer Range Store can not predict printer sales in the future. If you order a large number of printers and only a few printers are sold, this will cause printer stock to accumulate. With data mining using the naïve bayes method it is hoped that it can help the RCA Computer Store to predict printer sales in the future so that there are no printer stock that has accumulated and can maximize printer sales.

METHOD

Naive Bayes is a simple probabilistic classification that calculates a set of probabilities by adding up the frequency and combination of values from a given dataset. The algorithm uses the Bayes theorem and assumes all the independent or non-interdependent attributes given by values to class variables.

Naive Bayes is based on the simplification assumption that attribute values are conditionally mutually independent if output values are given. In other words, given the value of output, the probability of observing together is a product of individual probabilities. The advantage of using Naive Bayes is that this method only requires a small amount of training data to determine the estimated parameters needed in the classification process. Naive Bayes often works far better in most complex real-world situations than expected [9].

Input data variable is a set of data from an activity process that can be normalized manually [10]. In this printer sales forecasting system, input data will be processed by the printer brand, printer type, classification and payment method.

The naïve bayes scheme in Matlab is:

a. Upload data
b. Read the training data
c. Show prediction results

The Naïve Bayes Method Formula in Matlab

```
NB= NaiveBayes.fit(X,Y,'dist',{'kernel','kernel','kernel'});

% the final result of the calculation:

prediction = NB.predict(X);
```

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 215 - 220
DOI: https://doi.org/10.33330/icossit.v1i1.776
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

RESULT AND DISCUSSION

Table 1. Data Training

BRAND	TYPE	CLASIFICATION	PAYMENT
Canon	Ip 2770	Inkjet	Cash
Epson	Lx 310	Dot Matrik	Cash
Canon	Mp 287	Inkjet	Cash
Canon	Mp 287	Inkjet	Cash
Canon	Ip 2770	Inkjet	Cash
Hp	Dj 2545	Inkjet	Cash
Epson	L 220	Inkjet	Cash
Canon	Ip 2770	Inkjet	Bank Transfer
Epson	L 220	Inkjet	Bank Transfer
Brother	J 100	Inkjet	Bank Transfer
Epson	Lx 310	Dot Matrik	Bank Transfer
Epson	Lx 310	Dot Matrik	Bank Transfer
Hp	P 1102	Laserjet	Bank Transfer
Epson	L 120	Inkjet	Bank Transfer
Epson	L 350	Inkjet	Bank Transfer
Brother	J 100	Inkjet	Bank Transfer
Epson	Lx 310	Dot Matrik	Bank Transfer
Brother	J 200	Inkjet	Bank Transfer
Epson	Lx 310	Dot Matrik	Cash
Epson	Lx 310	Dot Matrik	Cash
Canon	Ip 2770	Inkjet	Cash
Canon	Ip 2770	Inkjet	Cash
Canon	Ip 2770	Inkjet	Cash
Canon	Mp 287	Inkjet	Bank Transfer
Canon	Lbp 6030	Laserjet	Bank Transfer
Hp	P 1102	Laserjet	Cash
Canon	Mp 287	Inkjet	Bank Transfer
Brother	J 200	Inkjet	Bank Transfer
Epson	Lx 310	Dot Matrik	Bank Transfer
Epson	L 120	Inkjet	Cash
Canon	Ip 2770	Inkjet	Bank Transfer
Canon	Mp 287	Inkjet	Bank Transfer
Hp	P 1102	Laserjet	Cash
Brother	J 200	Inkjet	Cash
Epson	L 220	Inkjet	Bank Transfer
Canon	Ip 2770	Inkjet	Bank Transfer

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 215 - 220
DOI: https://doi.org/10.33330/icossit.v1i1.776
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

Table 1. Data Training

BRAND	TYPE	CLASIFICATION	PAYMENT
Epson	L 120	Inkjet	Cash
Brother	J 200	Inkjet	Cash

Implementation

a. Upload data training to matlab

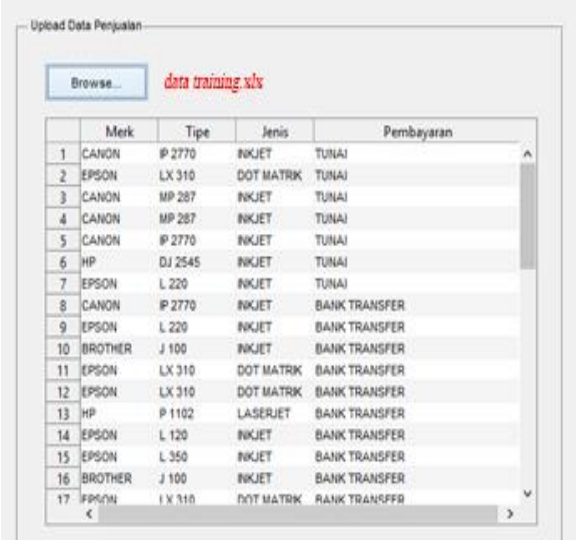

Figure 1. Interface Sales Data Upload form

b. Process of Calculation of Naïve Bayes on Matlab

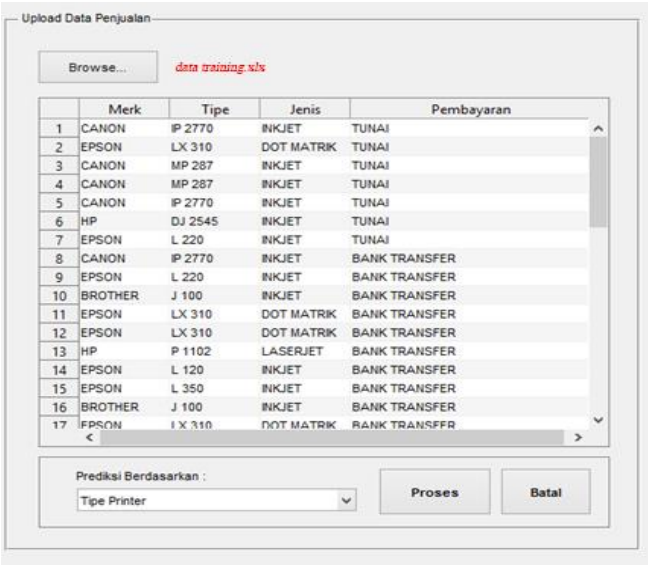

Figure 2. Predicted Output Interface

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 215 - 220
DOI: https://doi.org/10.33330/icossit.v1i1.776
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

c. Output of Matlab Prediction Results

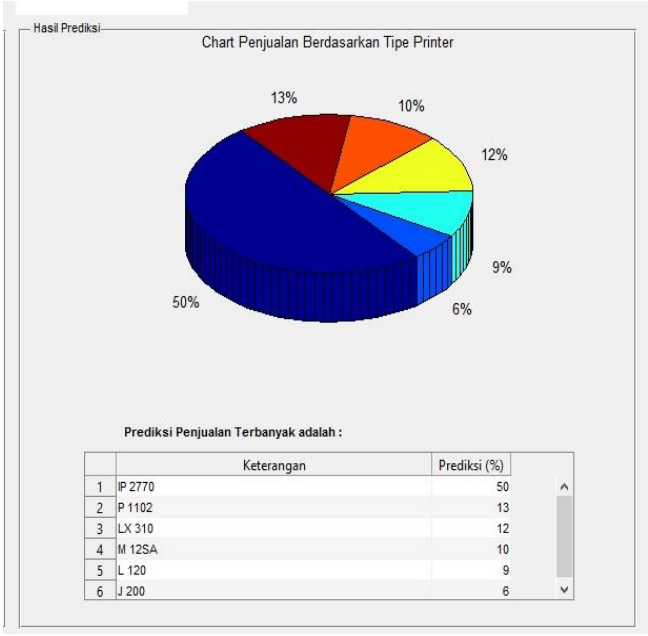

Figure 3. Output of Matlab Prediction Results

CONCLUSION

From the results of this study, by using the Naive Bayes method to predict printer sales types based on brand, type, and payment from the RCA Computer Shop, the following conclusions can be drawn.The Naive Bayes method can make predictions from printer sales data. The Naive Bayes method only provides the highest number of printer type predictions from the highest value for the type of printer based on brand attributes, type attributes, and payment attributes. The results of the printer type prediction process and the measurement of the level of data accuracy are very dependent on the training data to be processed in the system.

BIBLIOGRAPHY

[1] R. M. Sari, V. Tasril, and Y. A. M, "Prediksi Jumlah APBD Kota Payakumbuh dengan Metode K-Means," *IPTEKS Terap.*, vol. 14, pp. 45–50, 2020.
[2] Y. Apridonal M, "PENERAPAN DATA MINING MENGGUNAKAN METODE ASSICIATION RULE DENGAN ALGORITMA APRIORI UNTUK ANALISA POLA PENJUALAN BARANG," *JURTEKSI*, vol. V, no. 2, pp. 193–198, 2019.
[3] S. Syarli and A. Muin, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi

ISSN 2723-4509 (Online)
Kisaran, August 19th, 2020, page. 215 - 220
DOI: https://doi.org/10.33330/icossit.v1i1.776
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT

Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2016.

[4] M. Sabransyah, Y. N. Nasution, and F. D. T. Amijaya, "Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung," *J. EKSPONENSIAL*, vol. 8, no. 2, pp. 111–118, 2017.

[5] B. M. Atmega, "Jurnal Momentum ISSN : 1693-752X PERANCANGAN TRAFFIC LIGHT Jurnal Momentum ISSN : 1693-752X," vol. 17, no. 1, pp. 1–7, 2015.

[6] N. Fauziah, S. Wahyuningsih, and Y. N. Nasution, "Peramalan Mengunakan Fuzzy Time Series Chen (Studi Kasus : Curah Hujan Kota Samarinda)," *Statistika*, vol. 4, no. 2, 2016.

[7] A. Atina, "Aplikasi Matlab pada Teknologi Pencitraan Medis," *J. Penelit. Fis. dan Ter.*, vol. 1, no. 1, p. 28, 2019.

[8] I. Yuliastuti, L. Rahmasari, and R. -, "Pembuatan Aplikasi Program Matlab untuk Menganalisa Sifat Lasing Kaca TZBN Yang Didadah Ion Nd3+ sebagai Bahan Material Host Laser (Halaman 97 s.d. 102)," *J. Fis. Indones.*, vol. 18, no. 54, pp. 97–102, 2015.

[9] Y. H. Hui *et al.*, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI JUMLAH PRODUKSI BARANG BERDASARKAN DATA PERSEDIAAN DAN JUMLAH PEMESANAN PADA CV. PAPADAN MAMA PASTRIES. Volume 1.," *J. Mantik Penusa*, vol. 1, no. 2, pp. 16–21, 2017,

[10] R. B. Purnama, "Perancangan Prediksi Untuk Menentukan Indeks Harga Saham Menggunakan Jaringan Syaraf Tiruan," *Kinetik*, vol. 2, no. 2, p. 125, 2017.