ISSN 2085-4552

Sales Prediction at PT.World Infinite Network: A Comparative Study Of Naïve Bayes and Adaptive Neuro-Fuzzy Inference System

Jenie Sundari[1], Aden Irman[2]
[1] Universitas Bina Sarana Informatika: Faculty of Technology and Informatics, Jakarta, Indonesia
[2] Universitas Nusa Mandiri Faculty of Technology and Informatics, Jakarta, Indonesia
[1] Jenie.jni@bsi.ac.id, [2]12190102@nusamandiri.ac.id

Accepted 08 January 2026
Approved 23 January 2026

***Abstract*— This study analyzes historical sales transaction data from PT. World Infinite Network to support sales prediction and pattern discovery in the IT product domain. The dataset consists of structured transaction records containing product categories, sales volume, and time-based attributes, which are processed using data mining techniques. Two predictive methods—Naïve Bayes and Adaptive Neuro-Fuzzy Inference System (ANFIS)—are applied to model sales trends and classify purchasing behavior. Naïve Bayes is utilized for probabilistic classification due to its computational efficiency, while ANFIS is employed to capture nonlinear relationships through fuzzy inference and neural learning. Model performance is evaluated using classification accuracy. Experimental results show that the Naïve Bayes algorithm achieves an accuracy of 19.05%, indicating limited predictive capability on the given dataset and suggesting that sales patterns exhibit high variability and weak conditional independence among features. Despite the low accuracy, the findings highlight the comparative behavior of probabilistic and neuro-fuzzy approaches and provide insights into the suitability of different data mining methods for sales prediction in small or noisy transactional datasets. This study contributes empirical evidence for method selection in sales analytics and offers a baseline for improving predictive performance through feature engineering and data enrichment.**

***Index Terms*— sales, adaptive, neuro, data mining.**

I. INTRODUCTION

The development of digital technology has transformed communication patterns through the emergence of new media that accelerate information dissemination. This shift promotes a digital-oriented lifestyle and marks the rise of the information society, supported by Information and Communication Technology (ICT) infrastructure. Broader access to information enhances public participation through more open two-way communication and feedback mechanisms. [1].

One of the techniques for data processing is **data mining**, which aims to discover connections between data that are not yet known to users and to make them easily accessible. Based on these informational relationships, decisions can be derived. Description, estimation, prediction, classification, clustering, and association are among the main categories that constitute data mining. In short, classification refers to assigning an individual or entity into a specific category, generally referred to as a "class." According to previous studies, the **Naïve Bayes** method possesses several advantages: it performs rapid calculations, uses a simple algorithm, and yields highly accurate results. These features make it one of the most widely used methods in classification. To estimate the parameters required for classification, the Naïve Bayes method needs only a small amount of training data [2].

The **Apriori algorithm** is used to calculate association rules between objects, which indicate how two or more items interact with one another. In other words, association rules describe how two or more products are related—for example, evaluating whether customers who buy product A are also likely to buy product B. This algorithm, based on association rules, was developed by R. Agrawal and R. Srikant in 1994, specifically for analyzing shopping cart data and identifying products that are frequently purchased together. In the healthcare field, Apriori can also be applied to determine patient responses to medications.

Data mining using the Apriori algorithm is designed to find data that most frequently appear in a database. Event information within the database is recorded as items [3].

PT. World Infinite Network is a system integrator company that provides hardware, software, and maintenance services for information technology (IT) devices. Since its establishment in 2011, PT. World Infinite Network has collaborated with various

governmental and non-governmental institutions. Some of its government partners include the Department of Communication, Informatics, and Statistics (DISKOMINFOTIK) of the DKI Jakarta Provincial Government, the Human Resources Division of the Indonesian National Police (SSDM Polri), the Central Statistics Agency (BPS), Bank Negara Indonesia (BNI), Angkasa Pura, POS Indonesia, and the National Traffic Management Center (NTMC) of the Indonesian National Police. The company's product and service offerings include servers, personal computers (PCs), notebooks, laptops, storage devices, IT peripherals, and maintenance services [4]

II. THEORY

Prediction is a systematic method used to estimate future events based on past and present information [5]. Prediction is employed to obtain insights about future changes that may affect policy implementation and its consequences. For example, governments may use prediction to estimate the impact of economic policies on economic growth, inflation, and unemployment rates [6].

Sales refer to an integrated effort in designing strategies aimed at fulfilling the needs and desires of buyers, with the goal of achieving profitable transactions. Sales serve as the primary source of revenue for a company, as profits are generated from these activities [7].

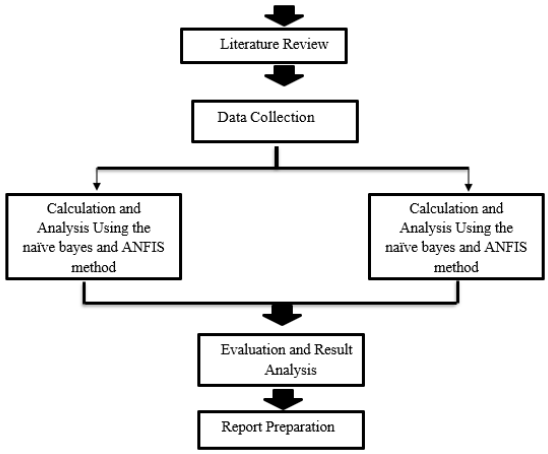
Data mining is an instrument that enables users to quickly access large volumes of data by extracting information from massive datasets through statistical, mathematical, and artificial intelligence approaches [8]. It is a method in the field of computer science used to discover knowledge from data, transforming it into useful and meaningful information.

In pattern recognition, the Bayesian Decision Theorem is an important statistical technique. Based on a simplifying assumption, the **Naïve Bayes** method assumes that if the output value is given to the attributes, these attributes are conditionally independent of one another. For categorical data features, Naïve Bayes is easy to apply—for example, when calculating the "gender" feature with values such as "male" and "female." However, additional methods are required before applying Naïve Bayes to handle numerical features. Consequently, the calculation of probability (likelihood) values for numerical (continuous) features differs from that for categorical (discrete) features[9].

The **Adaptive Neuro-Fuzzy Inference System (ANFIS)** is a hybrid approach that combines a fuzzy inference system with the adaptive learning capabilities of an artificial neural network. Its purpose is to create a model capable of capturing the relationship between the input and output of a complex system, even when such relationships are difficult to describe mathematically[10]

III. METHOD

The research process serves as an essential stage to facilitate the study, as it follows a structured, standardized, sequential, systematic, and logical framework. A systematic research procedure helps researchers conduct their studies effectively and in a well-directed manner. The following are the methods employed by the authors.



IV. RESULT AND DISCUSSION

In conducting this research, the sample data in the **Sales Data Table of PT. World Infinite Network** consisted of 22 product items, along with sales data covering a two-year period (2021–2022). The use of the **ANFIS** aimed to fulfill all the rules defined within the algorithm by meeting the minimum **support** and combining each item within the database, as well as the minimum **confidence** requirement, which represents the strength of the relationship between association rules and items.

Data processing was also carried out using the **Naïve Bayes Algorithm** in accordance with the standard implementation procedures of the algorithm. This approach enables the company to determine appropriate strategies based on the calculation results derived from the database that has been implemented using both the ANFIS and Naïve Bayes algorithms.

The following is a list of product items from the sales data at **PT. World Infinite Network:**

Table 1. List of Product Sales Items

No	Nama Item	Kode
1	Dell Latitude 3420	DL3
2	Dell Latitude 5420	DL5

3	Dell Latitude 3590	DL35
4	Acer Travelmate P214	ATP214
5	HP Envy X360	HPEX3
6	HP Probook 430	HPP43
7	HP Pav Plus 14	HPPP14
8	Lenovo V14	LV14
9	Lenovo Thinkpad X1 Carbon	LTPX1C
10	Lenovo Ideapad S530	LIDS530
11	Acer VM4960	AVM4960
12	Acer Veriton	AV
13	Dell Optiplex 3000 SFF	DO3000
14	Dell Optiplex 5260	DO5260
15	HP Pro A G3	HPPAG3
16	Lenovo Thinksystem ST250	LTST250
17	Dell PowerEdge R740	DPR740
18	Dell PowerEdge R540	DPER540
19	Microsoft Office Home And Business 2019	MOHB2019
20	Microsoft Office 365	MO365
21	Backpack	BPCK

Source: Research Results, 2023

Table 2. data training

Item	Brand	Type	Sales
DL3	Dell	Notebook	38
DL5	Dell	Notebook	22
DL35	Dell	Notebook	16
ATP214	Acer	Notebook	9
HPEX3	HP	Notebook	13
HPP43	HP	Notebook	18
HPPP14	HP	Notebook	14
LV14	Lenovo	Notebook	21
LTPX1C	Lenovo	Notebook	6
LIDS530	Lenovo	Notebook	30
AVM4960	Acer	PC	112
AV	Acer	PC	1
DO3000	Dell	PC	6
DO5260	Dell	PC	2
HPPAG3	HP	Server	2
LTST250	Lenovo	Server	6
DPR740	Dell	Server	0
DPER540	Dell	Server	13

Table 3. data testing

Item	Brand	Type	Sales
DL3	Dell	Notebook	8
DL5	Dell	Notebook	4
DL35	Dell	Notebook	5
ATP214	Acer	Notebook	9
HPEX3	HP	Notebook	2
HPP43	HP	Notebook	1
HPPP14	HP	Notebook	1
LV14	Lenovo	Notebook	0
LTPX1C	Lenovo	Notebook	0
LIDS530	Lenovo	Notebook	90
AVM4960	Acer	PC	10
AV	Acer	PC	4
DO3000	Dell	PC	3
DO5260	Dell	PC	0
HPPAG3	HP	Server	2
LTST250	Lenovo	Server	4
DPR740	Dell	Server	0
DPER540	Dell	Server	1
MOHB2019	Microsoft	Software	4
MO365	Microsoft	Software	1
BPCK	Unbrand	Accessories	120

In the data grouping process, the dataset was divided into a ratio of **80% for training data** and **20% for testing data**.

Table 4 naïve bayes calculation

	true_DL3	true_DL5	true_ATP	true_HPP	true_HPE	true_HPPP	true_LV14	true_LTP	true_LIDS530
pred_DL3	1	0	0	0	0	0	0	0	0
pred_DL5	0	1	0	0	0	0	0	0	0
pred_DL35	0	0	1	0	0	0	0	0	0
pred_ATP	0	0	0	1	0	0	0	0	0
pred_HPEX3	0	0	0	0	1	0	0	0	0
pred_HPP43	0	0	0	0	0	1	0	0	0
pred_HPPP14	0	0	0	0	0	0	1	0	0
pred_LV14	0	0	0	0	0	0	0	1	0
pred_LTPX1C	0	0	0	0	0	0	0	0	1
pred_LIDS530	0	0	0	0	0	0	0	0	0

Based on the image above, the results of the Naïve Bayes Algorithm calculation using RapidMiner Studio show an outcome of **19.05%**.

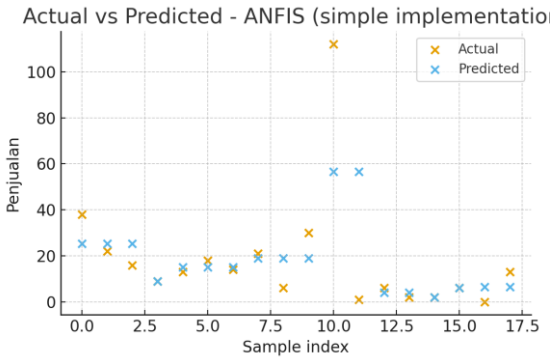


figure 1. ANFIS prediction

Model: Simple ANFIS (2 inputs: Brand & Item Type; 2 Gaussian MFs per input → 4 rules; first-order consequent).

Training method: Alternating Least Squares for consequent parameters combined with numerical gradient descent for premise parameters (mean & sigma).

Final RMSE (Root Mean Square Error): ≈ **19.4646** (sales units).

V. CONCLUSION

Based on the experimental results, the Naïve Bayes algorithm achieved an accuracy of **19.05%**, indicating limited predictive capability for this dataset. In contrast, the **Simple ANFIS model** with two input variables (Brand and Item Type), four fuzzy rules, and first-order consequents produced a **Final RMSE of approximately 19.4646 sales units**. This suggests that the ANFIS model was able to represent the relationship between inputs and sales output with a moderate level of error. Overall, the ANFIS approach provides a more flexible and adaptive modeling framework compared to the Naïve Bayes classifier for this type of sales prediction problem.

REFERENCES

[1] I. Astrid Faidlatul Habibah, "The Information Society Era as an Impact of New Media," *J. Teknol. dan Inf. Bisnis*, vol. 3, no. 2, pp. 350–363, 2021, doi: https://doi.org/10.47233/jteksis.v3i2.255.

[2] S. J. P.Bhargavi, "Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils," *Int. J. Comput. Sci. Netw. Secur.*, vol. 9, no. 8, pp. 117–122, 2009.

[3] R. A. Mahessya and S. Indrawati, "Implementasi Metode Anfis Data Mining Dalam Menyeleksi Beasiswa Di Smpn 7 Sorolangun," *J. Process.*, vol. 12, no. 1, pp. 904–915, 2017, [Online]. Available: http://ejournal.stikom-db.ac.id/index.php/processor/article/view/379

[4] L. Genetic and A. Viewer, "Artificial life Genetic Algorithm Viewer 1 . 0," pp. 11–15.

[5] S. Adiguno, Y. Syahra, and M. Yetri, "Implementasi Data Mining C.45 LINEAR REGRESI DAN KMEANS DENGAN MENGGUNAKAN FRAMEWORK DJANGO PYTHON," *J. Sist. Inf. TGD*, vol. 1, no. 4, pp. 275–281, 2022.

[6] G. Galih, "Data Mining di Bidang Pendidikan untuk Analisa Prediksi Kinerja Mahasiswa dengan Komparasi 2 Model Klasifikasi pada STMIK Jabar," *J. Teknol. Sist. Inf. dan Apl.*, vol. 2, no. 1, p. 23, 2019, doi: 10.32493/jtsi.v2i1.2643.

[7] A. Safitri Sembiring, N. Laila, and A. Wahyuni Lubis, "Analisis Harga Pokok Penjualan dan Laba Kontribusi terhadap Volume Penjualan pada Perum Bulog Divre Sumut," *ILTIZAM J. Shariah Econ. Res.*, vol. 7, no. 1, pp. 109–123, 2023, doi: 10.30631/iltizam.v7i1.1841.

[8] M. S. Mochammad Faid, Ahmad Supri, "Implementasi Data Mining C.45 LINEAR REGRESI DAN KMEANS DENGAN MENGGUNAKAN FRAMEWORK DJANGO PYTHON," pp. 2–3, 2023.

[9] A. Virrayyani and S. Sutikno, "Prediksi Penjualan Barang Menggunakan Metode Adaptive Neuro-Fuzzy Inference System (ANFIS)," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 2, no. 2, pp. 57–63, 2016, doi: 10.23917/khif.v2i2.2554.

[10] Fahrillah Fahrillah and Zaehol Fatah, "Pengelompokan Data Nilai Siswa Madrasah Ta'Hiliyah Menggunakan Metode K-Means Clustering," *J. Ris. Sist. Inf.*, vol. 2, no. 1, pp. 53–59, 2025, doi: 10.69714/0v1pkz05.