Jurnal Riset Multidisiplin Edukasi
https://journal.hasbaedukasi.co.id/index.php/jurmie
Halaman : 860-871

DESIGNING A PRODUCT CLASSIFICATION DASHBOARD FOR MARKETING STRATEGY USING K-NEAREST NEIGHBOR

Kadek Intan Cahya Putri[1], Anak Agung Ngurah Hary Susila[2], Ni Putu Sutramiani[3]
Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana, Bali[1,2,3]
Email: intancahya1202@gmail.com[1], harysusila@unud.ac.id[2], sutramiani@unud.ac.id[3]

Informasi

Volume : 2
Nomor : 7
Bulan : Juli
Tahun : 2025
E-ISSN : 3062-9624

Abstract

The development of information technology has driven the use of sales data to support data-driven business decision-making. This study aims to design a dashboard to classify Orlenalycious Padangsambian's products using the K-Nearest Neighbor (K-NN) algorithm to determine more accurate marketing strategies. The methods used include collecting sales data from the Moka POS system, data preprocessing, classification using the K-NN algorithm with K=5, and visualizing the classification results in a Streamlit-based dashboard. The classification results divide the products into three categories: Highly Popular, Popular, and Fairly Popular. The proposed marketing strategy refers to the 4P Marketing Mix, where highly popular products are promoted intensively, popular products are pushed through advertising, and fairly popular products are evaluated or promoted through bundling. The resulting dashboard displays informative visualizations such as pie charts and bar charts to facilitate the analysis of sales trends and product performance. This study provides a solution for Orlenalycious to design more efficient and effective data-driven marketing strategies, as well as offering an easier way to monitor and evaluate product performance in real-time.

Keywords: *Dashboard, K-Nearest Neighbor, Data Visualization, Marketing Strategy, Marketing Mix 4P.*

Perkembangan teknologi informasi telah mendorong penggunaan data penjualan untuk mendukung pengambilan keputusan bisnis yang berbasiskan data. Penelitian ini bertujuan merancang sebuah dashboard untuk mengklasifikasikan produk Orlenalycious Padangsambian dengan menggunakan algoritma K-Nearest Neighbor (K-NN), guna menentukan strategi pemasaran yang lebih tepat. Metode yang digunakan mencakup pengumpulan data penjualan dari sistem Moka POS, preprocessing data, klasifikasi menggunakan algoritma K-NN dengan parameter K=5, dan visualisasi hasil klasifikasi dalam bentuk dashboard berbasis Streamlit. Hasil klasifikasi produk dibagi menjadi tiga kategori: Sangat Diminati, Diminati, dan Cukup Diminati. Strategi pemasaran yang diusulkan mengacu pada Marketing Mix 4P, di mana produk yang sangat diminati dipromosikan secara intensif, produk yang diminati didorong melalui iklan, dan produk yang cukup diminati dievaluasi atau dipromosikan melalui bundling. Dashboard yang dihasilkan menampilkan visualisasi data yang informatif, seperti pie chart dan bar chart, untuk memudahkan analisis tren penjualan dan performa produk. Penelitian ini memberikan solusi bagi Orlenalycious untuk merancang strategi pemasaran berbasis data yang lebih efisien dan efektif, serta menawarkan cara yang lebih mudah dalam memonitor dan mengevaluasi performa produk secara real-

time.

Kata Kunci: *Dashboard, K-Nearest Neighbor, Visualisasi Data, Strategi Pemasaran, Marketing Mix 4P.*

A. INTRODUCTION

In the current era of digital business transformation, data plays a central role in shaping strategic decisions. The abundance of product sales data provides valuable opportunities for businesses to extract insights that can improve marketing performance. However, without appropriate tools and techniques, these data often remain underutilized.

Product classification is a crucial step in identifying trends, customer preferences, and market segmentation. By categorizing products based on performance or features, businesses can design more targeted and efficient marketing strategies. A commonly used algorithm for classification tasks is the K-Nearest Neighbor (KNN), known for its simplicity and effectiveness in pattern recognition based on proximity between data points.

KNN does not require model training in advance and makes classification decisions based on the majority vote among the closest neighbors. This approach is particularly beneficial when working with relatively small or structured datasets. Nevertheless, the output of the KNN algorithm must be presented in a way that is comprehensible to decision-makers.

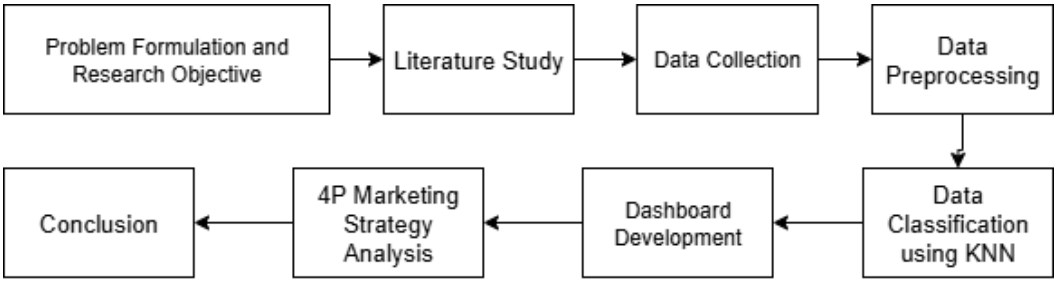
To bridge this need, dashboard technology offers a visual solution that integrates data processing with user-friendly presentation. Dashboards can display classification results through dynamic graphs, filters, and key performance indicators (KPIs) that enable quick understanding and action.

This study focuses on the development of an interactive dashboard to present product classification results generated by the K-Nearest Neighbor algorithm. The system is designed to assist decision-makers in determining appropriate marketing strategies based on data-driven classification. The dashboard aims to provide real-time insights into product positioning and category performance. The research also evaluates the accuracy of the classification model and the usability of the dashboard interface.

B. RESEARCH METHOD

This study was conducted at Orlenalycious Store, located in Perumahan Jl. Purba Indah Raya Gg. II No.9, Padangsambian, West Denpasar, Bali, from September 2024 until completion. The research methodology serves as a structured guide to ensure the project is

completed systematically, clearly, and efficiently. The research process involves several key stages, including problem formulation, literature review, data collection, data preprocessing, product classification using the K-Nearest Neighbor (K-NN) algorithm, data visualization through dashboard design, and marketing strategy analysis.



Problem Formulation and Research Objective

The study begins by identifying issues related to the lack of effective decision-making tools in managing product sales performance at Orlenalycious, a local SME. The core problem lies in the inability to classify products effectively based on their sales trends, which limits the accuracy of marketing strategies. Thus, the objective of this research is to design and implement a dashboard system that visualizes product classification using the KNN algorithm, enabling the business to make data-driven marketing decisions based on real-time analytics.

Literature Study

A comprehensive literature review is conducted to understand current approaches in data classification, dashboard development, and marketing strategy frameworks. Particular attention is paid to previous works on the implementation of the KNN algorithm for product classification due to its simplicity, high accuracy with labeled data, and robustness with numerical features. Studies on dashboard visualization tools such as Streamlit, Power BI, and Google Data Studio are also explored. Moreover, research on the 4P Marketing Strategy (Product, Price, Place, Promotion) is reviewed to guide the final analysis.

Data Collection

The data for this research was collected from Orlenalycious' point-of-sale system (MOKA POS), spanning from March 2022 to December 2024. The dataset includes: Product name and category, Quantity sold, Date and time of transaction, Sales channel (e.g., Walk-in, GoFood, ShopeeFood), Gross and net sales values, Payment method. In addition to transactional data, primary data was collected via interviews with the store owner and questionnaires to assess pre- and post-dashboard implementation perceptions.

Data Preprocessing

To prepare the dataset for machine learning, several preprocessing steps were conducted:

- Data Cleaning: Removal of null values, correction of invalid dates, and elimination of duplicate records.
- Data Transformation: Conversion of date formats and extraction of new time-based features (e.g., month, year).
- Feature Engineering: Computation of new variables such as average selling price, product frequency, and dominant sales channel.
- Data Aggregation: Transaction-level data was grouped by product for classification purposes.
- Data Labeling: Products were labeled into three classes based on cumulative sales contribution thresholds: Highly Demanded (<85%), Moderately Demanded (85–95%), Low Demand (>95%)

This labeled dataset was used as the target variable for classification.

Data Classification using KNN

The core classification process was executed using the K-Nearest Neighbor (KNN) algorithm with the following stages:

- Data Splitting: The dataset was divided into training (70%) and testing (30%) sets.
- K-Value Optimization: Several values of k were tested (e.g., k=3, k=5, k=7) to identify the most accurate configuration.
- Distance Calculation: Euclidean Distance was used to measure the similarity between data points.
- Model Evaluation: Performance metrics such as accuracy, precision, and recall were calculated to assess the model. The best model was selected for deployment in the dashboard.

Dashboard Development

A dynamic, user-friendly dashboard was developed using Streamlit, a Python-based data visualization framework. The dashboard includes: Overview statistics (total sales, transaction count, quantity sold), Line charts for historical sales trends, Bar charts for top-selling product categories, Pie charts for sales channel distribution, Scatter and bubble plots for product lead time analysis, Tables summarizing product classification results.

This dashboard is designed to assist business owners in quickly identifying key

performance areas and making strategic decisions.

4P Marketing Strategy Analysis

Based on the classification results, marketing strategies were formulated using the 4P framework:

- Product: Identify which products require promotion, modification, or discontinuation based on performance.
- Price: Determine optimal pricing strategies for each product category based on sales contribution.
- Place: Evaluate the effectiveness of each sales channel and focus efforts on high-contribution platforms.
- Promotion: Schedule campaigns based on historical high-performance periods (e.g., holidays, seasonal trends).

The outcome of this analysis supports Orlenalycious in optimizing their product positioning and improving sales performance.

Conclusion

The final stage includes summarizing the findings of the research and evaluating the impact of the dashboard on business decisions. A post-implementation questionnaire was distributed to assess user satisfaction, dashboard usability, and perceived decision-making improvements. Recommendations for future research include integrating predictive forecasting features and expanding to multichannel data sources.

Literature Study

Business Intelligence

Business Intelligence (BI) is a tool used to collect, manage, and analyze large amounts of business data to generate better decisions and appropriate actions for improving company performance [1]. BI enables organizations to leverage technologies such as data mining, analytics, and data visualization to identify patterns and business trends, thereby crafting more effective strategies for improving productivity and operational efficiency.

Classification

Classification is a technique used to build models from unlabeled data, which are then used to categorize new data based on similarity of characteristics to predefined categories [2]. Classification is widely applied across various fields including marketing, supporting decision-making by segmenting data. This enables businesses to enhance strategies such as more targeted marketing or forecasting product demand and consumer behavior.

K-Nearest Neighbor (K-NN)

K-Nearest Neighbor (K-NN) is an algorithm that classifies data by grouping objects based on the closest distance between them [3]. It is widely used due to its simplicity and effectiveness in clustering products or customers. K-NN can improve prediction accuracy regarding consumer preferences or product performance. The K-NN algorithm maps data into multidimensional space and calculates distances using methods such as Euclidean Distance:

$d(x,y) = \sqrt{\sum^{i)8} (x_a - y_a)^2}$

Where x and y are feature vectors representing training and test data, respectively, and i represents the number of dimensions [4].

Dashboard

A dashboard is a visualization tool that displays key information needed to achieve one or several objectives, simplifying the monitoring and analysis of data [5]. Dashboards provide an overview of business performance, allowing users to understand trends and metrics in real-time.

Streamlit

Streamlit is an open-source Python library for building web apps for data science and machine learning [6]. Built on NumPy, Pandas, and Matplotlib, Streamlit enables developers to quickly create dashboards and visualizations using an intuitive API. Visualizations in this research include line charts, bar charts, scatter plots, bubble charts, and pie charts presented in a real-time dashboard.

Marketing Strategy

Marketing strategy refers to decision-making processes concerning the allocation of marketing budgets in consideration of environmental conditions and competition [7]. It plays a key role in identifying the right market segments, positioning products, and tailoring promotional tactics. The objectives of a strong marketing strategy include improving team coordination, enhancing adaptability, supporting logical decision-making, and serving as a performance benchmark.

Soft Selling

Soft selling is a sales technique that uses storytelling to offer products subtly rather than through direct promotion [8]. It emphasizes the customer experience and emotions associated with the product, and is particularly effective for building long-term customer trust and loyalty.

Hard Selling

Hard selling delivers a direct and clear promotional message, encouraging immediate customer responses [9]. Techniques like discounts, packages, and limited-time offers are typically used to boost short-term sales.

Python and Libraries

Python is a popular programming language known for its readable syntax and flexibility. Libraries such as Pandas, NumPy, and Matplotlib make it ideal for data analysis, web app development, and AI [10].

- Scikit-learn: Used for implementing machine learning algorithms like classification, regression, and clustering [11].
- NumPy: Facilitates fast numerical computation using multidimensional arrays [12].
- Pandas: Supports data manipulation and analysis using its DataFrame structure [13].
- Matplotlib: Enables creation of a wide variety of static, animated, and interactive visualizations [14].

Google Colaboratory

Google Colab is a cloud-based platform that allows users to write and run Python code in-browser without setup, making it highly collaborative and accessible [15].

Machine Learning Concepts

Machine Learning (ML) allows systems to learn and make decisions based on data without explicit programming [16]. There are three learning paradigms:

- Supervised Learning: Trains models on labeled datasets to predict outputs [16].
- Unsupervised Learning: Identifies patterns in unlabeled data [16].
- Reinforcement Learning: Combines elements of supervised and unsupervised learning using reward feedback from the environment [16].

Overfitting and Underfitting

Overfitting occurs when a model learns the training data too well, failing to generalize to new data [17]. Underfitting happens when a model is too simplistic, unable to learn data patterns effectively [18].

Marketing Mix 4P

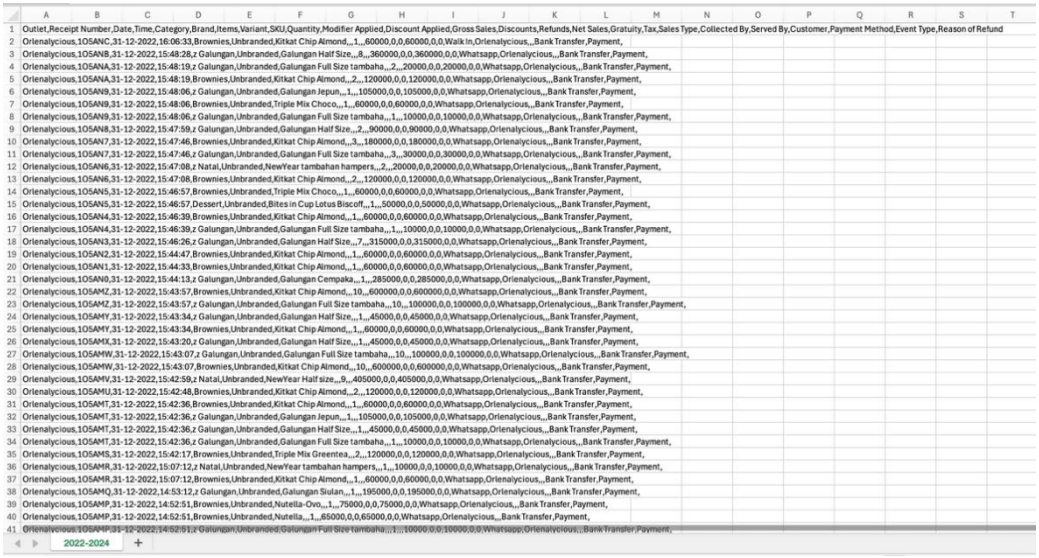
The Marketing Mix 4P (Product, Price, Place, Promotion) is a foundational framework for developing marketing strategies [19]. Each element contributes to delivering value and achieving business objectives.

C. RESULTS AND DISCUSSION

This section presents the results and discussion of the research, covering the process of data preparation, classification using the K-Nearest Neighbor (K-NN) algorithm, development of a product classification dashboard, and marketing strategy analysis based on the classified results. Each sub-section includes detailed explanations supported by figures, code outputs, and relevant interpretations.

Data Collection and Preprocessing

The research began with data collection from Orlenalycious' digital cashier system, MOKA POS. The data comprised 126,479 transaction records collected over a 34-month period from March 2022 to December 2024. The dataset included attributes such as receipt number, date, time, product category, item name, quantity, gross sales, net sales, sales type, and payment method.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Outlet,Receipt Number,Date,Time,Category,Brand,Items,Variant,SKU,Quantity,Modifier Applied,Discount Applied,Gross Sales,Discounts,Refunds,Net Sales,Gratuity,Tax,Sales Type,Collected By,Served By,Customer,Payment Method,Event Type,Reason of Refund																			
2	Orlenalycious,1OSANC,31-12-2022,16:06:33,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Walk In,Orlenalycious,,,Bank Transfer,Payment,																			
3	Orlenalycious,1OSANB,31-12-2022,15:48:28,z Galungan,Unbranded,Galungan Half Size,,,8,,,360000,0,0,360000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
4	Orlenalycious,1OSANA,31-12-2022,15:48:19,z Galungan,Unbranded,Galungan Full Size tambaha,,,2,,,20000,0,0,20000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
5	Orlenalycious,1OSANA,31-12-2022,15:48:19,Brownies,Unbranded,Kitkat Chip Almond,,,2,,,120000,0,0,120000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
6	Orlenalycious,1OSAN9,31-12-2022,15:48:06,z Galungan,Unbranded,Galungan Jepun,,,1,,,105000,0,0,105000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
7	Orlenalycious,1OSAN9,31-12-2022,15:48:06,Brownies,Unbranded,Triple Mix Choco,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
8	Orlenalycious,1OSAN9,31-12-2022,15:48:06,z Galungan,Unbranded,Galungan Full Size tambaha,,,1,,,10000,0,0,10000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
9	Orlenalycious,1OSAN8,31-12-2022,15:47:59,z Galungan,Unbranded,Galungan Half Size,,,2,,,90000,0,0,90000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
10	Orlenalycious,1OSAN7,31-12-2022,15:47:46,Brownies,Unbranded,Kitkat Chip Almond,,,3,,,180000,0,0,180000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
11	Orlenalycious,1OSAN7,31-12-2022,15:47:46,z Galungan,Unbranded,Galungan Full Size tambaha,,,3,,,30000,0,0,30000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
12	Orlenalycious,1OSAN6,31-12-2022,15:47:08,z Natal,Unbranded,NewYear tambahan hampers,,,2,,,20000,0,0,20000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
13	Orlenalycious,1OSAN6,31-12-2022,15:47:08,Brownies,Unbranded,Kitkat Chip Almond,,,2,,,120000,0,0,120000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
14	Orlenalycious,1OSAN5,31-12-2022,15:46:57,Brownies,Unbranded,Triple Mix Choco,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
15	Orlenalycious,1OSAN5,31-12-2022,15:46:57,Dessert,Unbranded,Bites in Cup Lotus Biscoff,,,1,,,50000,0,0,50000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
16	Orlenalycious,1OSAN4,31-12-2022,15:46:39,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
17	Orlenalycious,1OSAN4,31-12-2022,15:46:39,z Galungan,Unbranded,Galungan Full Size tambaha,,,1,,,10000,0,0,10000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
18	Orlenalycious,1OSAN3,31-12-2022,15:46:26,z Galungan,Unbranded,Galungan Half Size,,,7,,,315000,0,0,315000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
19	Orlenalycious,1OSAN2,31-12-2022,15:44:47,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
20	Orlenalycious,1OSAN1,31-12-2022,15:44:33,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
21	Orlenalycious,1OSAN0,31-12-2022,15:44:13,z Galungan,Unbranded,Galungan Cempaka,,,1,,,285000,0,0,285000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
22	Orlenalycious,1OSAMZ,31-12-2022,15:43:57,Brownies,Unbranded,Kitkat Chip Almond,,,10,,,600000,0,0,600000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
23	Orlenalycious,1OSAMZ,31-12-2022,15:43:57,z Galungan,Unbranded,Galungan Full Size tambaha,,,10,,,100000,0,0,100000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
24	Orlenalycious,1OSAMY,31-12-2022,15:43:34,z Galungan,Unbranded,Galungan Half Size,,,1,,,45000,0,0,45000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
25	Orlenalycious,1OSAMY,31-12-2022,15:43:34,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
26	Orlenalycious,1OSAMX,31-12-2022,15:43:20,z Galungan,Unbranded,Galungan Half Size,,,1,,,45000,0,0,45000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
27	Orlenalycious,1OSAMW,31-12-2022,15:43:07,z Galungan,Unbranded,Galungan Full Size tambaha,,,10,,,100000,0,0,100000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
28	Orlenalycious,1OSAMW,31-12-2022,15:43:07,Brownies,Unbranded,Kitkat Chip Almond,,,10,,,600000,0,0,600000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
29	Orlenalycious,1OSAMV,31-12-2022,15:42:59,z Natal,Unbranded,NewYear Half size,,,9,,,405000,0,0,405000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
30	Orlenalycious,1OSAMU,31-12-2022,15:42:48,Brownies,Unbranded,Kitkat Chip Almond,,,2,,,120000,0,0,120000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
31	Orlenalycious,1OSAMT,31-12-2022,15:42:36,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
32	Orlenalycious,1OSAMT,31-12-2022,15:42:36,z Galungan,Unbranded,Galungan Jepun,,,1,,,105000,0,0,105000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
33	Orlenalycious,1OSAMT,31-12-2022,15:42:36,z Galungan,Unbranded,Galungan Half Size,,,1,,,45000,0,0,45000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
34	Orlenalycious,1OSAMT,31-12-2022,15:42:36,z Galungan,Unbranded,Galungan Full Size tambaha,,,1,,,10000,0,0,10000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
35	Orlenalycious,1OSAMS,31-12-2022,15:42:17,Brownies,Unbranded,Triple Mix Greentea,,,2,,,120000,0,0,120000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
36	Orlenalycious,1OSAMR,31-12-2022,15:07:12,z Natal,Unbranded,NewYear tambahan hampers,,,1,,,10000,0,0,10000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
37	Orlenalycious,1OSAMR,31-12-2022,15:07:12,Brownies,Unbranded,Kitkat Chip Almond,,,1,,,60000,0,0,60000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
38	Orlenalycious,1OSAMQ,31-12-2022,14:53:12,z Galungan,Unbranded,Galungan Siulan,,,1,,,195000,0,0,195000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
39	Orlenalycious,1OSAMP,31-12-2022,14:52:51,Brownies,Unbranded,Nutella-Oreo,,,1,,,75000,0,0,75000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
40	Orlenalycious,1OSAMP,31-12-2022,14:52:51,Brownies,Unbranded,Nutella,,,1,,,65000,0,0,65000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			
41	Orlenalycious,1OSAMP,31-12-2022,14:52:51,z Galungan,Unbranded,Galungan Full Size tambaha,,,1,,,10000,0,0,10000,0,0,Whatsapp,Orlenalycious,,,Bank Transfer,Payment,																			

Preprocessing of the data involved multiple stages: data cleaning, formatting and standardization, aggregation, data matching and filtering, feature engineering, and data labeling. The cleaning process eliminated null values and corrected inconsistent formats, such as converting date fields to datetime format and standardizing item names. Feature engineering introduced calculated attributes like product frequency, average price, and lead time. The data was then labeled into three classes based on cumulative sales contribution: Highly Demanded (0-85%), Demanded (85-95%), and Less Demanded (>95%).

Classification Using K-NN

Following preprocessing, classification was performed using the K-Nearest Neighbor algorithm with a k-value of 5. The dataset was divided into training (70%) and testing (30%) subsets using stratified sampling to ensure class distribution balance. Categorical variables

such as Sales_Type and Payment_Method were transformed using one-hot encoding, while numerical features were standardized to improve model accuracy.

The model was trained and tested, yielding a testing accuracy of 51.19% and training accuracy of 73.85%. The classification report showed a precision macro average of 0.35, indicating moderate precision across all classes. A confusion matrix was generated to identify the performance for each class, where the model correctly classified "Highly Demanded" items 32 times, "Demanded" items 11 times, and "Less Demanded" items 8 times.

Finally, the trained model was used to predict the class of all products in the dataset, and the predicted results were compared with the threshold-based labels. This comparison highlighted variations between manual and model-based classifications, providing insights for strategic product assessment.

Dashboard Development and Visualization

A Streamlit-based dashboard was developed to visualize the classification results. The dashboard allows users to upload .CSV files and provides multiple tabs, including Overview, Product Analysis, and Prediction. In the Overview tab, users can view general sales metrics such as total sales, total quantity, and total transactions. Sales trends and top-selling categories are visualized using line charts and bar charts.

The Product Analysis tab provides detailed insights into product performance through tables, scatter plots, bubble charts, and pie charts. Key indicators such as lead time, product frequency, and average price are displayed to support inventory and pricing strategies. The tab also visualizes sales distribution across various channels and payment methods, helping identify the most effective distribution platforms.

Marketing Strategy Analysis

The classified product data was further analyzed using the 4P Marketing Mix framework. The Product aspect focused on grouping products based on demand to guide stock planning. The Price strategy included bundling less demanded items with popular ones and implementing premium pricing for high-performance products.

For the Place strategy, evaluation of distribution channels revealed that GrabFood contributed 30.1% to sales, followed by Walk-In (24.9%), indicating these as primary channels for targeted campaigns. In terms of Promotion, multiple approaches were proposed, including loyalty cards through WhatsApp Business, broadcast promotions, and chatbot implementation to enhance customer interaction and engagement.

Evaluation Through Questionnaire

Post-research evaluation was conducted using a Likert-scale questionnaire targeted at Orlenalycious stakeholders. The average score for the promotion aspect was 3.5, indicating agreement on the effectiveness of promotional strategies, especially through GrabFood and WhatsApp. The distribution aspect scored 3.9, suggesting that channel diversification, such as expansion to ShopeeFood and GoFood, was well received. For the proposed bundling packages, the average score was 4.1, indicating high acceptance and potential impact on increasing marketplace performance.

Overall, the research findings support the effectiveness of using Business Intelligence and machine learning for product classification and strategic marketing. The dashboard serves as a practical tool for SMEs to make data-driven decisions that enhance sales performance and operational efficiency.

D. CONCLUSION

This study successfully demonstrates the integration of Business Intelligence and machine learning—specifically the K-Nearest Neighbor (K-NN) algorithm—as a valuable approach to support data-driven marketing decisions for small and medium-sized enterprises (SMEs). By utilizing a structured dataset of 126,479 sales transactions collected over a 34-month period from Orlenalycious, the research was able to classify products into three categories of demand: Highly Demanded, Demanded, and Less Demanded. Although the classification model achieved a moderate level of accuracy during testing, it offered meaningful insights into product performance and sales contribution that could be translated into practical marketing actions.

The development and implementation of an interactive dashboard using Streamlit served as an effective visualization tool, enabling users—particularly business owners—to interpret complex product and sales data more intuitively. The dashboard incorporated several visual elements, including line charts, bar graphs, scatter plots, and pie charts, to provide a comprehensive overview of sales performance and classification results. Additionally, the dashboard supported strategic planning using the Marketing Mix 4P framework (Product, Price, Place, Promotion), offering business insights such as price optimization, bundling strategies, promotion planning, and prioritization of sales channels based on historical performance.

Evaluation through stakeholder questionnaires revealed that the proposed dashboard and recommendations were perceived as highly useful. Positive responses were recorded particularly for bundled package promotions, the use of WhatsApp Business for direct communication, and the proposed loyalty card system—demonstrating the practical implications and user satisfaction with the system.

In conclusion, the research has effectively addressed the problem statements outlined in the introduction by producing a functional classification model and visualization tool to enhance marketing decision-making. For future research, it is recommended to explore alternative classification algorithms to improve prediction accuracy and to incorporate real-time data integration into the dashboard to allow for more dynamic and responsive business analysis.

E. REFERENCES

[1] Fauzi Ramadhan, M., Wibawa, K. S., & Kurniawan, A. Business Intelligence System for Food Distribution Monitoring. Journal of Applied Information and Computing Technology, 2022; 11(2): 132–140

[2] Dalbergio, M., Situmorang, M., & Christian, A. Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Risiko. Jurnal Informatika dan Komputer, 2019; 4(1): 17–25

[3] Cholil, M., Rahmat, H., & Munir, A. Klasifikasi Tingkat Keaktifan Mahasiswa dengan Algoritma K-Nearest Neighbor. Jurnal Teknologi dan Sistem Komputer, 2021; 9(3): 184–192

[4] Miya Juwita, A., Alamsyah, Y., & Putra, D. Evaluasi Akurasi Algoritma K-NN pada Klasifikasi Minat Baca Mahasiswa. Jurnal Riset Teknologi dan Inovasi Pendidikan, 2022; 5(1): 22–28

[5] Faid, M., Sari, R., & Yuliana, N. Analisis Pengaruh Nilai K terhadap Akurasi K-NN dalam Klasifikasi Produk. Seminar Nasional Teknologi Informasi dan Aplikasi Komputer, Surabaya. 2019; 10: 45–50.

[6] Wahidillah, A., Rofik, R., & Prasetyo, T. Penerapan Euclidean Distance pada Sistem Klasifikasi Produk UMKM. Jurnal Teknologi dan Sains, 2023; 12(1): 39–47.

[7] Sihombing, A. L., & Manurung, S. Desain Dashboard Interaktif untuk Monitoring Penjualan. Jurnal Sistem Informasi, 2019; 7(2): 111–119.

[8] Tholib, A. Dasar Penggunaan Streamlit untuk Aplikasi Data Science. In: Ramadhan, I., Kusuma, A., Editors. Python Tools for Machine Learning. Yogyakarta: Abutholib Press;

2021: 33–54.

[9] Rambe, A., Hanafiah, M., & Anshari, N. Strategi Pemasaran Digital dalam Era Transformasi Bisnis. Jurnal Bisnis dan Inovasi, 2022; 8(1): 99–106.

[10] Haque-Fawzi, M. G., & Sutrisno, H. Strategi Pemasaran: Konsep, Teori dan Implementasi, Edisi Kedua. Jakarta: Rajawali Pers; 2022.

[11] Faizaty, E., & Lestari, W. Strategi Soft Selling Melalui Media Sosial untuk Meningkatkan Brand Awareness. Jurnal Ilmu Komunikasi dan Media, 2021; 6(2): 78–86.

[12] Musrifah, H. Teknik Hard Selling dalam Peningkatan Penjualan Produk UMKM. Jurnal Ekonomi dan Kewirausahaan, 2021; 5(1): 24–29.

[13] Ma'arif, M. Pemrograman Python untuk Analisis Data. Yogyakarta: Graha Ilmu; 2020.

[14] Raharjo, S. Pengenalan Pustaka Python: Pandas, NumPy, dan Matplotlib. Jurnal Teknologi Informatika, n.d.; 4(2): 90–97.

[15] Andarsyah, R., Susanto, A., & Wibowo, S. Pemanfaatan Google Colaboratory dalam Pendidikan Berbasis Data. Jurnal Teknologi Edukasi dan Pembelajaran, 2024; 3(1): 55–63.

[16] Wijoyo, R., Fitriani, D., & Hasan, M. Penerapan Algoritma Machine Learning dalam Sistem Prediksi. Jurnal Sistem dan Informatika, 2024; 6(2): 140–153.

[17] Anhar, R., & Nugraha, Y. Permasalahan Overfitting dalam Model Pembelajaran Mesin. Jurnal Penelitian dan Pengembangan Teknologi Informasi, 2023; 5(1): 70–77.

[18] Firmansyach, I., & Kurniawan, T. Analisis Underfitting pada Proses Pelatihan Model AI. Seminar Nasional Teknologi Komputer dan Sains, Bandung. 2023: 120–126.

[19] Karim, F., Wibowo, A., & Suryana, D. Strategi Marketing Mix 4P untuk Destinasi Pariwisata. Jurnal Manajemen Pariwisata, 2021; 9(3): 33–41.