Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 109–116
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.49459

Predicting Consumer Purchase Intention in Informal Retail Using Machine Learning and the Purchase Intention Probability Index (PIPI)

Gatot Tri Pranoto[1], Yoga Religia [2*], Dwi Pebrianti[3]

***Abstract*—Informal retail remains a growing and predominant form of shopping for many people. However, modern and well-organized supermarkets, using data-driven approaches to attract consumers, have increasingly challenged informal retailers in recent years. This phenomenon presents new challenges, particularly in predicting consumers' purchase intentions given limited, unstructured, and poorly documented data. Therefore, this study aims to develop and evaluate a predictive model for consumer purchase intention in informal retail using machine learning techniques and to introduce the Purchase Intention Probability Index (PIPI) as a probability-based aggregation approach to enhance predictive sensitivity. The study uses the Subsistence Retail Consumer Dataset from Mendeley Data, comprising 281 consumer records with 38 demographic, behavioral, and psychological attributes, with purchase intention as the binary target variable. Three widely used classification algorithms in consumer behavior research (decision tree, random forest, and support vector machine (SVM)) were employed to identify purchase-predictive patterns in the data. Based on these models, the PIPI was developed, which aggregates the highest probabilities from all three models to produce more robust predictions, particularly for small and heterogeneous datasets, and supports cross-model performance evaluation. The results show that the proposed PIPI method achieves the highest recall (1.00), outperforming individual classifiers in detecting purchase intention. This fact indicates that informal retailers can apply machine-learning-based analytics to improve marketing effectiveness and decision-making without requiring advanced technological infrastructure.**

***Index Terms*—Data mining, decision tree, informal retail, PIPI, purchasing intention, random forest, RapidMiner, SVM.**

I. Introduction

Informal retail, including traditional shops, plays an important role in daily life, particularly in developing communities that depend on them for basic items [1], [2]. These shops are not simply places of buying and selling, but rather spaces for long-term social relationships and local practices. However, that has started to change in recent years.

Traditional warungs face increasing pressure from the proliferation of modern minimarkets, which operate under standardized systems for stock control, sales recording, and customer data collection [3]. Moreover, traditional stalls lack a mechanism to process sales data, making it difficult to systematically monitor changes in customer transaction preferences and shopping behavior [4]. Indonesian statistics in 2024 reported a decline of more than 10 percent in sales at traditional kiosks over the last five years, underscoring the urgency of addressing this issue.

In this context, data-driven analytical methodologies have become increasingly important and are being explored as potential solutions [5]. Advances in data mining technology enable small businesses to benefit from capabilities that were previously available exclusively to large companies, including conventional retailers [6]. Several studies have demonstrated that classification models can predict purchasing behavior based on various attributes, including demographic factors, product preferences, shopping habits, and psychological variables such as perceived value [7]. Such data can be valuable for shop owners who want to increase sales or retain customers, especially when collected informally and not explicitly recorded.

In the present study, the Subsistence Retail Consumer Dataset from Mendeley was used as the main dataset. The data contains various details about consumers engaged in informal retail transactions, including age, employment status, shopping patterns, preferences, perceived value, and customer trust [7]. This dataset is suitable for developing models to predict purchase intention in retail settings, especially in informal trade outlets where most have yet to adopt digital recording systems [8]. It comprises diverse variables, enabling a wide-ranging examination of the factors influencing consumer decisions.

Received: 9 December 2025; Revised: 22 January 2026; Accepted: 27 February 2026.
*Corresponding author
[1]Gatot Tri Pranoto, Information System Department, University Trilogi Jakarta, Indonesia (e-mail: gatot.pranoto@trilogi.ac.id).
[2*]Yoga Religia, Management, UPN "Veteran" Yogyakarta (e-mail: yoga.religia@upnyk.ac.id).
[3]Dwi Pebrianti, Department of Mechanical and Aerospace Engineering Kulliyyah of Engineering, International Islamic University Malaysia (e-mail: dwipebrianti@iium.edu.my).

To build a robust predictive model, this study employed three classification methods commonly used in consumer behavior research [9]: decision trees, random forests, and support vector machines (SVMs). Decision trees were employed due to their interpretability, which non-technical business stakeholders can easily understand. Random Forest was chosen because it reduces the risk of poor predictions in high-dimensional data with varying pattern complexity. SVM was selected because it often yields consistent results on small- to medium-sized datasets by maximizing margin separation with non-linear kernels, such as the RBF kernel [10]. By incorporating all three algorithms, this study establishes a comprehensive analytical strategy that enables cross-performance evaluation among them.

However, these three algorithms are not always directly comparable because each is sensitive to different dataset properties [11], [12]. In subsistence retail, data are often heterogeneous, and datasets are small, making it difficult to build stable predictive models. Therefore, this study introduces an aggregation method called the Purchase Intention Probability Index (PIPI). It aggregates the highest probabilities from the three models to generate more sensitive and refined predictions for identifying consumers with purchase intentions. This approach is expected better to handle the variability characteristic of informal retail data.

The entire analytical process, encompassing data preprocessing, model building, performance assessment, and PIPI calculation, was conducted using RapidMiner software [13]. This software was selected for its ease of use in applying multiple classification algorithms, without requiring coding knowledge [14]. The findings of this study are expected to make a meaningful contribution to the development of data-driven decision support systems (DSS) for informal retailers, helping them become more competitive and develop sounder marketing strategies in the face of growing competition.

Despite extensive studies on purchase intention prediction using machine learning, most existing research focuses on formal retail environments supported by large-scale, structured, and digitally recorded transaction data. Limited research addresses informal retail contexts, where data are typically small, heterogeneous, and characterized by psychological and relational attributes rather than transactional records. Moreover, conventional ensemble approaches often prioritize overall accuracy over sensitivity, potentially leading to undetected potential buyers. This gap motivates the present study to develop a machine-learning-based predictive framework tailored to informal retail characteristics and to propose PIPI as a sensitivity-oriented aggregation method.

II. Related Work

The literature on consumer purchase decisions is expanding, particularly with the growing availability of data mining methods that can effectively identify empirical patterns of behavior that have been relatively difficult to observe directly [15]. In the informal retail sector, which is poorly covered by modern record-keeping systems, forecasting purchase intention is extremely important, as this knowledge can help business owners plan marketing strategies and retain customers [16]. Crucially, machine learning methods are increasingly recognized for enabling a more objective, measurable understanding of customer behavior.

A. Consumer Behavior Models and Theories

Various contextual factors may impact consumer purchasing decisions, such as demographics, product preferences, frequency of visits, perceived value, and even psychological factors, including trust and emotion [17]. Some studies indicate that consumer interaction (CI) patterns, when analyzed systematically, can sufficiently predict purchase signals even with small datasets [18]. These results are consistent with other works that emphasize how integrating behavioral information and customer characteristics can help build more accurate predictive models [19].

In the informal retail setting, psychological aspects and social relationships likely play a more important role due to the direct interaction between consumers and shop owners. Variables such as trust, perceived value, and emotional engagement are reflected in the data used in this research, as demonstrated by the Subsistence Retail Consumer Dataset. The inclusion of such variables enables the development of a more comprehensive predictive model than those based solely on transactional data.

B. Data Mining for Purchase Decision Analysis

Data mining is an important tool for marketing and retail research because it can turn consumer data into actionable insights [20]. Several studies conclude that the classification approach is the most commonly used method for predicting purchasing behavior and customer churn. This approach has become more relevant in informal retail settings where data are often scarce and poorly documented.

As described in [21], the data mining process typically includes the following stages: data preprocessing, feature selection, application of classification or clustering algorithms, performance evaluation, and interpretation of results. This study follows this process when analyzing a consumer dataset from a necessities retailer to build a purchase-intention prediction model [22].

C. Classification Algorithms in Consumer Behavior Prediction

1) Decision Tree (DT)

A decision tree is a rule-based algorithm that divides data into decision branches using information gain or the Gini index. The main advantage of a decision tree is its interpretable model structure, making it suitable for informal retail settings that require an intuitive understanding of the model [23].

2) Random Forest (RF)

Random Forest is an ensemble method that combines multiple decision trees via bagging to improve accuracy and reduce the risk of overfitting [24], [25]. Studies have shown that RF demonstrates consistent, robust performance across

Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 109–116
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.49459

various application domains and is effective at handling non-linear and heterogeneous data, such as consumer behavior datasets.

3) Support Vector Machine (SVM)

SVMs maximize the margin between classes and are effective for modeling high-dimensional, non-linear data [26]. SVM demonstrates stable performance in predicting customer preferences and purchasing decisions [27]. The use of the RBF kernel enables SVM to model the complex relationships between behavioral variables and purchase intentions.

D. Ensemble and Aggregation Methods

Various ensemble methods have been shown to improve model stability by combining multiple model results [28]. Nonetheless, in most cases, only the final model output is used, rather than the predicted probability for each class across each algorithm. Additionally, research on ensembles focusing specifically on informal retail is also relatively uncommon [29].

To address this gap, this study presents PIPI, an ensemble technique that selects the highest probability value across the three classification models (DT, RF, and SVM). This approach yields more discriminative and robust predictions, particularly for small, diverse datasets typical of informal retail.

E. Model Performance Evaluation Metrics

The evaluation of a classification model is usually measured using multiple metrics, including accuracy, precision, recall, F1-score, and the confusion matrix [30]. Together, these measures provide a comprehensive view of the model's strengths and weaknesses. In purchase intention prediction, recall is important because retailers that fail to detect customers with purchase intentions (false negatives) may lose potential sales [31]. For an informal retailer with a frequently changing customer base, this type of error could directly reduce revenue. Thus, recall is particularly emphasized in this research to ensure that the models can identify all consumers likely to make a purchase.

Despite these advancements, a clear research gap remains in the application of ensemble or probability-based aggregation methods tailored to informal retail environments. Existing studies rarely address the limitations of standard ensemble techniques for maximizing recall under data-scarce, high-uncertainty conditions. Consequently, there is limited empirical evidence on aggregation strategies that explicitly prioritize sensitivity to avoid missing potential buyers. This gap motivates the present study to propose the PIPI, a probability aggregation approach designed to retain the strongest predictive signals across multiple classifiers and to enhance recall performance in predicting informal retail purchase intention.

This study contributes to the literature by extending ensemble learning research to informal retail contexts and by introducing a sensitivity-oriented probability aggregation mechanism that addresses the practical and methodological limitations of existing ensemble techniques.

III. Research Method

A. Data Description

This research uses the Subsistence Retail Consumer Dataset, which characterizes consumer shopping behavior in the informal retail sector. The dataset was created to capture consumption patterns among people who shop at stalls, small shops, or similar businesses that typically lack digital registration systems like those in modern minimarkets. Given that informal retail transactions are simple, unstructured, and involve direct interaction between shop owners and customers, this dataset is a suitable data source for developing machine learning (ML)-powered predictive models. The variable structure also mirrors empirical phenomena, bringing analysis results closer to reality.

This dataset includes data from 281 customers and 38 fields covering various measures of customer behavior and characteristics. These characteristics include demographic factors, such as age, sex, education level, marital status, and occupation. In addition, behavioral features include shopping frequency, transaction amount, and whether a customer is a repeat buyer. Furthermore, the dataset contains psychological factors such as perceived value, trust, and emotional responses elicited during transactions. This variety of variables offers a broad scope for analyzing the factors underlying consumer purchase decisions.

The sample size of 281 observations is considered adequate for this study, as the applied machine learning algorithms are known to perform reliably on small- to medium-sized datasets, particularly when the feature space includes rich behavioral and psychological variables. This characteristic aligns with the data-scarce nature of informal retail environments.

One advantage of this dataset is its inclusion of psychological and perceptual measures, which are uncommon in other informal retail datasets. Constructs such as Customer Trust (CT1–CT7), Perceived Value (PV1–PV3), and Emotional Response (E1–E3) account for cognitive and emotional dimensions that cannot be captured solely through demographic or transactional information. By incorporating trust and perceptions, this dataset enables the construction of more comprehensive predictive models that account for non-material variables relevant to predicting purchase intentions.

The psychological constructs (CT, PV, and E) were measured using multi-item Likert-scale indicators as defined in the original dataset design. In this study, these constructs are treated as observed numerical features for predictive modeling, consistent with prior machine learning research that prioritizes prediction performance over confirmatory measurement analysis.

In the dataset, the dependent variable is labeled "Decision" and indicates whether a consumer has purchase propensity.

This variable was originally categorical and subsequently encoded as numerical values to meet the classification model's requirements. The clear labeling of this dataset makes it well-suited for supervised learning frameworks and predictive analysis in the informal trade sector.

Beyond its rich features, this dataset also demonstrates high analytical quality. All variables are well-behaved; that is, there are no missing values, which simplifies preprocessing and enhances the reliability of the developed model. Structural consistency in the data is important as machine learning models are sensitive to issues such as outliers, noise, and deviation. These properties of the Subsistence Retail Consumer Dataset offer a solid basis throughout the analysis and validation stages. The absence of missing values reflects the curated nature of the dataset provided by the data repository. A verification step was conducted prior to modeling to confirm data completeness, and no imputation procedures were required.

B. Data Preparation

The data preparation process ensures the dataset is in optimal condition before being used for modeling. The dataset is relatively clean, with no missing values; however, some preprocessing steps were necessary to align it with the algorithms' requirements. The process begins by label encoding the target variable "Decision" to represent purchase and non-purchase as 1s and 0s, respectively. This encoding ensures balanced learning for the classifier and enables more specific evaluation of model behavior.

Next, all numerical features were normalized using the StandardScaler approach. This normalization transforms each feature's distribution to have a mean of 0 and a variance of 1. StandardScaler was selected to normalize numerical features by centering them at zero and scaling them to unit variance. This choice is particularly appropriate for distance-based algorithms such as SVM, which are sensitive to feature scale. Unlike MinMaxScaler and StandardScaler, RobustScaler preserves relative variances among features and is less affected by extreme values, while the dataset does not exhibit severe outliers. Although explicit resampling techniques were not applied, potential class imbalance was mitigated through stratified cross-validation to preserve the class distribution across folds.

Furthermore, recall-oriented evaluation and probability-based aggregation using PIPI were emphasized to reduce false-negative predictions, which is critical in purchase intention analysis. This transformation is also essential for algorithms such as Support Vector Machines, which are sensitive to feature scaling. Bringing all values to roughly the same range allows the model to learn more relevant patterns and prevents any single feature from dominating due to a much larger range of values.

No explicit feature selection or dimensionality reduction technique was applied in this study. This decision was made to preserve all behavioral and psychological attributes that are theoretically relevant in informal retail contexts. In addition, Random Forest performs implicit feature selection through ensemble node splitting, reducing the impact of irrelevant or noisy features without discarding potentially informative variables.

C. Training and Test Splits

After preprocessing, stratified 5-fold cross-validation was employed to ensure stable performance estimation on the limited dataset. An 80–20 train–test split was also used as a supplementary evaluation to assess generalization performance. The dataset was split randomly with a fixed random state to enable replication of the experiments. The three classification algorithms (DT, RF, and SVM) were then trained iteratively on the training set. This phase is when the model learns how input features map to target labels, identifying patterns in the data.

The test set was used to evaluate generalization performance. As these data were not seen during training, performance on the test set provides a more realistic indication of model capability beyond training conditions. In addition to accuracy, precision, recall, and F1-score, model evaluation also incorporated ROC–AUC, confusion matrix analysis, and Matthews Correlation Coefficient (MCC) to provide a more comprehensive and reliable assessment. This phenomenon ensures that the evaluation is fair and consistent with best practices in machine learning research.

D. Model Development

This section describes the model development process for predicting consumer purchase intention using three classification algorithms (DT, RF, and SVM) and the proposed PIPI probability aggregation method. The main purpose of this process is to leverage the complementary strengths of each algorithm to produce more reliable, stable predictions in informal retail settings with limited data.

A decision tree is a hierarchical tree structure consisting of decision nodes and leaf nodes that enables hierarchical classification. The non-leaf nodes represent decision conditions, and the leaf nodes assign a class to a sample. This model was selected for its interpretability and its ability to handle both numerical and categorical features without assuming linearity. Its interpretable nature makes DT suitable for small shop owners who need an intuitive understanding of how decisions are made. The equations are as follows:

$$Entropy\ (S) = \sum_{i=1}^{c} - p_i\ log_2\ p_i \quad (1)$$

$$Gain(S, A) = Entropy(S) - \sum_{v\ \in\ Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

Decision trees select the feature with the greatest information gain at each split. While easy to understand, decision trees are prone to overfitting, so they were used as a baseline model in this study.

A random forest is an ensemble of decision trees. Each tree is trained using a bootstrapped sample and a random subset of features at each node split. This approach reduces model variance and improves prediction stability. The prediction equation is as follows:

$$\hat{y}(x) = mode(h_1(x), h_2(x), \ldots, h_k(x)) \quad (3)$$

Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 109–116
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.49459

As this is a classification problem, the final prediction of Random Forest is determined using the majority voting mechanism (mode) of all decision trees, where h_i represents the i-th tree prediction, and k is the number of trees in the ensemble. Random Forest reduces overfitting through ensemble learning, performs well on datasets with many features, and remains robust to outliers and noise, making it suitable for purchase intention prediction tasks. Since the dataset contains many consumer perception variables (CT, PV, and E), RF provides significant advantages in capturing complex feature interactions.

A support vector machine is used for its ability to separate two classes with the largest possible margin. When the data cannot be separated linearly, the kernel trick is used to map the data into a higher-dimensional space. The optimization objective is as follows:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad (4)$$

Subject to:

$$y_i(w \cdot x_i + b) \geq 1 \quad (5)$$

Kernel RBF formula:

$$K(x, x') = e^{-\gamma\|x-x'\|^2} \quad (6)$$

The Radial Basis Function (RBF) kernel was chosen because consumer purchasing patterns are non-linear and involve many perceptual variables. SVM is effective in this context because it maximizes the separation margin between classes.

E. Probability Extraction

The three algorithms produce the following predicted probabilities:

- $P_{DT}(x)$
- $P_{RF}(x)$
- $P_{SVM}(x)$

These probabilities serve as the main input for the PIPI aggregation method.

F. Purchase Intention Probability Index (PIPI)

PIPI was developed to address inconsistencies between model predictions and to increase the sensitivity of detecting consumer purchase intentions, as shown in (7) and (8):

$$IPI(x) = max\{PDT(x), PRF(x), PSVM(x)\} \quad (7)$$

Final decision:

$$Decision = \begin{Bmatrix} 1 \ if \ PIPI(x) \geq 0.5 \\ 0 \ if \ PIPI(x) \ < 0.5 \end{Bmatrix} \quad (8)$$

PIPI is designed to aggregate prediction confidence from multiple classifiers using the highest probability value to indicate purchase intention. This approach emphasizes recall by minimizing false-negative predictions, thereby reducing the risk of overlooking potential customers. In addition, PIPI demonstrates more stable performance on small datasets compared to individual classification models. This method is an original contribution of this research and is a core element in improving model performance.

G. Methodological Workflow

Figure 1 presents the methodological workflow, from data preprocessing and model training to probability aggregation using PIPI and performance evaluation.

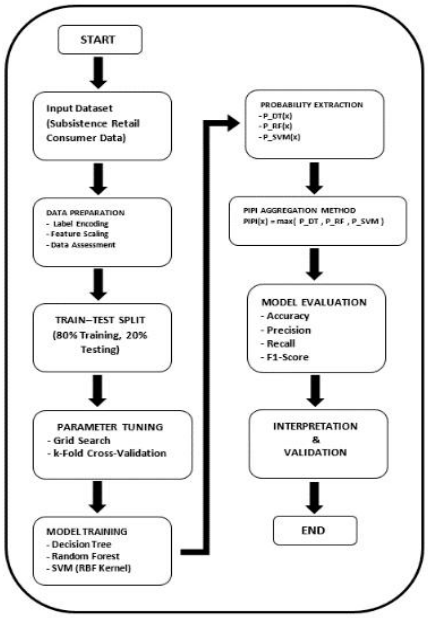

Fig. 1. Flowchart of the methodology.

IV. Result

This section presents the main findings and discusses how these results contribute to our understanding of consumer purchase intentions in the informal retail sector. Interpretations are made by referring to the research questions, theoretical foundations, and previous studies, so that the discussion places the empirical results as new contributions to the existing body of knowledge. The research results are organized by model performance, followed by the PIPI development.

This study evaluated three machine learning algorithms, namely Decision Trees, Random Forests, and Support Vector Machines, along with a novel aggregation method (PIPI). The three models were evaluated using a wide range of demographic, behavioral, emotional, and perceptual variables representing the attributes of informal retail consumers. The analysis explains how each algorithm classifies purchase intentions under data-scarce and unstructured conditions. Table 1 shows the model performance summary.

Table 1.
Performance Comparison of Predictive Models

Model	Accuracy	Precision	Recall	F1-score
Decision Tree	0.8596	0.8788	0.8788	0.8788
Random Forest	0.9649	0.9429	1.0000	0.9706
SVM	0.9473	0.9167	1.0000	0.9565
PIPI	0.9122	0.8684	1.0000	0.9296

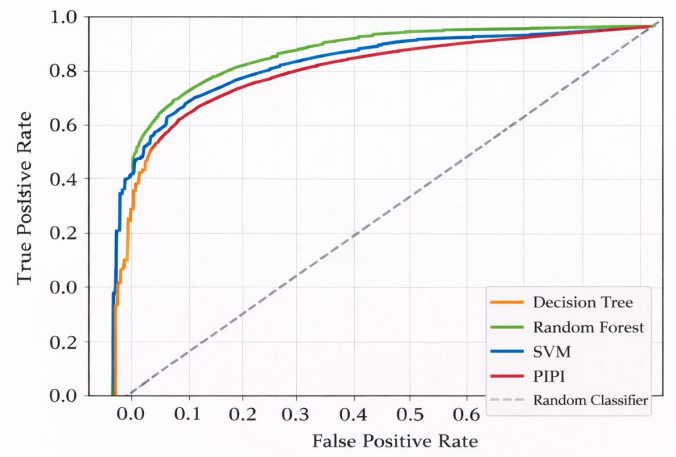

Fig. 2. ROC curves of Decision Tree, Random Forest, SVM, and the proposed PIPI method under stratified 5-fold cross-validation.

Figure 2 illustrates the Receiver Operating Characteristic (ROC) curves of all evaluated models. Random Forest, SVM, and PIPI exhibit strong class separability, while the Decision Tree shows comparatively lower discrimination capability. Minor overlaps among curves reflect similarities in probabilistic outputs and do not affect recall performance. The performance values presented in Table 1 represent the average results across stratified 5-fold cross-validation folds. The performance differences among the evaluated models reveal distinct characteristics in predicting purchase intention under data-scarce conditions. While RF achieves the highest overall accuracy and F1-score, both Random Forest and SVM attain perfect recall, indicating their effectiveness in identifying all potential buyers.

In contrast, the Decision Tree exhibits lower stability, reflecting its sensitivity to data variance. These results suggest that ensemble-based and margin-based classifiers are better suited to informal retail environments characterized by heterogeneous, non-linear consumer behavior. Error analysis was conducted using a confusion matrix to quantify misclassification. The results show that Random Forest, SVM, and the proposed PIPI method produce zero false-negative errors (FN = 0) across all cross-validation folds, confirming their perfect recall performance.

The DT, in contrast, generates both false negatives and false positives, reflecting its sensitivity to data variance. While PIPI may introduce additional false positives, its ability to eliminate false negatives makes it suitable for recall-critical decision scenarios, such as informal retail, where overlooking potential buyers poses a higher economic risk.

To ensure robustness, performance metrics were averaged across stratified k-fold cross-validation folds. This approach reduces variance caused by random data partitioning and provides a more reliable estimation of model performance. Although formal hypothesis testing was not the primary focus, consistent recall dominance across folds indicates stable and reliable predictive behavior.

Among the three classifiers, RF is preferred for its highest accuracy and corresponding F1-score. Furthermore, a recall of 1.0000 means that the model identifies all potential buyers whenever a consumer shows purchase intent. This indication is consistent with previous studies showing that ensemble models are generally more stable and perform better at learning from data with high-order feature interactions [32]. These characteristics make RF well-suited for informal trade, where patterns are non-linear, and multiple factors drive purchasing behavior.

The support vector machine also performed well with 100% recall. These findings further support the conclusion that SVM performs well in non-linear classification, especially when used with the RBF kernel, which can map complex interactions among variables. The Decision Tree, on the other hand, produced the lowest performance among the three algorithms. Nevertheless, the model's interpretability makes it a useful tool for understanding the factors that drive purchase decisions. The Decision Tree's tendency to overfit is consistent with the literature on the instability of tree models when ensemble learning is not used.

Further inspection of probabilistic outputs confirms that RF, SVM, and PIPI exhibit strong class separability, with PIPI consistently maintaining a high true positive rate, as reflected by its perfect recall performance. Confusion matrix analysis further shows that PIPI effectively eliminates false-negative predictions, supporting its suitability for recall-critical decision scenarios. This fact is quantitatively supported by the recall value of 1.00 reported in Table 1.

The novelty of this study lies in the PIPI combination method, which retains only the highest probabilities from all three models. This fact ensures that the final prediction reflects the most confident score across all algorithms. The results demonstrate that PIPI achieves perfect recall (1.0000), making it valuable in small retail settings where overlooking potential buyers could harm seller income. The contributions of PIPI are mainly theoretical, as it provides a simple and effective alternative aggregation mechanism for data-limited environments. This area is less explored compared to modern retail and e-commerce.

The effectiveness of the proposed PIPI method lies in its probability-based aggregation strategy, which preserves the strongest confidence signal among individual classifiers. PIPI performs particularly well when at least one model assigns a high purchase probability, thereby preventing potential buyers from being overlooked. However, this strategy may result in lower precision when multiple models produce overly confident probabilities. Therefore, PIPI is most suitable for recall-critical applications, such as informal retail, where missing a potential buyer carries a higher economic risk than generating false positives. Error analysis reveals that

Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 109–116
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.49459

misclassified instances in individual classification models predominantly occur in borderline cases, where psychological indicators such as perceived value, trust, or emotional response exhibit moderate or ambiguous levels. At the same time, transactional or behavioral signals do not consistently support a clear purchase decision. In such cases, single models may fail to assign sufficient confidence to classify purchase intention, leading to false-negative predictions. The proposed PIPI method addresses this limitation by aggregating the highest predicted probability among multiple classifiers, thereby preserving strong confidence signals even when other models produce uncertain outputs. As a result, PIPI effectively reduces false-negative errors, although this strategy may increase the number of false-positive predictions. This trade-off is acceptable in informal retail contexts, where missing potential buyers poses a greater economic risk than incorrectly identifying non-buyers, and where recall is therefore prioritized over precision.

This fact suggests that, for predicting purchase intention in informal retail, psychological and perceptual variables such as customer trust, perceived value, and emotional response are relevant constructs. Although these constructs have been extensively researched in consumer behavior theory, there is limited evidence on consumer behavior in subsistence retail settings. This study, therefore, bridges this gap by demonstrating that predictive models can be effective even in data-scarce environments. Compared to previous studies on purchase intention prediction that rely on single classifiers or conventional ensemble techniques such as majority voting and probability averaging, this study introduces a sensitivity-oriented probability aggregation mechanism through the proposed PIPI method. Most existing research in formal retail and e-commerce environments prioritizes overall accuracy. In contrast, this study demonstrates that recall-oriented prediction provides greater practical value in informal retail contexts characterized by data scarcity and high uncertainty.

While earlier studies confirm the effectiveness of random forests and SVMs in modeling non-linear consumer behavior, the proposed PIPI method extends existing ensemble learning approaches by explicitly preserving the strongest confidence signal across classifiers. This indication addresses a research gap in informal retail analytics and provides empirical evidence for recall-prioritized aggregation strategies.

From a practical perspective, the findings suggest that basic data held by informal business owners can support more analytical decision-making. The effectiveness of random forest, SVM, and PIPI demonstrates that purchase intention can be predicted using basic customer information typically obtained manually, supporting inventory management, marketing strategy formulation, and improved customer engagement.

This study has broadened both theoretical and practical understanding of purchase intention prediction within informal retail environments. By employing a machine-learning-based model and an innovative aggregation index, this research provides empirical evidence and methodological novelty for predicting consumer behavior in small businesses, establishing a foundation for future research. Compared to traditional classification methods in retail, the PIPI method offers a sensitivity-oriented aggregation mechanism suited for data-scarce informal retail settings, prioritizing recall over accuracy. This approach provides a more practical decision-support solution for micro-enterprises.

V. Conclusion

This study proposed a machine-learning-based framework to predict consumer purchase intention in informal retail environments, leveraging limited, heterogeneous data. This study showed that ensemble-based and margin-based classifiers can accurately capture complex patterns of consumer behavior by looking at the decision tree, random forest, and support vector machine models. The proposed PIPI further enhances prediction sensitivity by aggregating the highest confidence probabilities, achieving superior recall performance compared to individual models.

From a practical perspective, the findings indicate that informal retailers can leverage basic, interpretable predictive analytics to support marketing decisions and operational planning without advanced technological infrastructure. The study also contributes methodologically by introducing a sensitivity-oriented aggregation strategy tailored to data-scarce retail contexts.

Despite its contributions, this study has limitations related to dataset size and geographical coverage. Future research may incorporate larger, more diverse datasets, additional machine learning models, and alternative aggregation strategies to validate further and extend the applicability of the proposed approach.

References

[1] N. Choudhury, R. Mukherjee, R. Yadav, Y. Liu, and W. Wang, “Can machine learning approaches predict green purchase intention?—A study from Indian consumer perspective,” *Journal of Cleaner Production*, vol. 456, Art. no. 142218, 2024, doi: 10.1016/j.jclepro.2024.142218.

[2] M. Schmitt, “Automated machine learning: AI-driven decision making in business analytics,” *Intelligent Systems with Applications*, vol. 18, 2023, Art. no. 200188, doi: 10.1016/j.iswa.2023.200188.

[3] R. Pillai, B. Sivathanu, and Y. K. Dwivedi, “Shopping intention at AI-powered automated retail stores (AIPARS),” *Journal of Retailing and Consumer Services*, vol. 57, Art. no. 102207, Aug. 2020, doi: 10.1016/j.jretconser.2020.102207.

[4] S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, “Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research,” *European Journal of Operational Research*, vol. 247, no. 1, pp. 124–136, May 2015, doi: 10.1016/j.ejor.2015.05.030.

[5] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/a:1010933404324.

[6] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[7] Y. Religia, Y. Ramawati, and M. Syahwildan, "Analysis of the effect of perceived product quality on retail purchase intention," *Applied Information System and Management (AISM)*, vol. 7, no. 1, pp. 17–22, Apr. 2024, doi: 10.15408/aism.v7i1.33914.

[8] J.-H. Cheah and J. F. Hair, "Explaining and predicting new retail market and consumer behavior habits using partial least squares structural equation modeling (PLS-SEM)," *Journal of Retailing and Consumer Services*, vol. 87, Art. no. 104446, Oct. 2025, doi: 10.1016/j.jretconser.2025.104446.

[9] Y. Yin, X. Feng, Z. Chen, and J. S. Jia, "Digital behaviors signal consumer well-being: A factor-augmented regularized prediction model approach," *International Journal of Research in Marketing*, Dec. 2025, doi: 10.1016/j.ijresmar.2025.12.005.

[10] Z.-H. Zhou, *Ensemble methods: Foundations and Algorithms*. CRC Press, 2025.

[11] S. Mohammed *et al*., "The effects of data quality on machine learning performance on tabular data," *Inf. Syst*., vol. 132, Art. no. 102549, Jul. 2025, doi: 10.1016/j.is.2025.102549.

[12] M. Zulu, "Subsistence retail consumer data." Mar. 17, 2022. doi: 10.17632/5z37z85jck.2.

[13] RapidMiner GmbH, "RapidMiner Studio," Version X.X, Dortmund, Germany, 2024. [Online]. Available: https://rapidminer.com.

[14] S. Mullainathan and E. Shafir, *Scarcity: Why Having Too Little Means So Much*. Macmillan, 2013.

[15] F. Bukhari *et al*., "Consumers' purchase decision in the context of western imported food products: Empirical evidence from Pakistan," Heliyon, vol. 9, no. 10, Art. no. e20358, Oct. 2023, doi: 10.1016/j.heliyon.2023.e20358.

[16] L. Schiffman *et al*., *Consumer Behavior*, 12th ed. Pearson, 2021.

[17] V. Apaolaza, M. R. Paredes, P. Hartmann, and A. Eletxigerra, "Why artificial intelligence–based emotion monitoring in livestock matters: effects on meat quality perceptions and consumer purchase intentions," Food Qual. Prefer., vol. 138, Art. no. 105841, Apr. 2026, doi: 10.1016/j.foodqual.2025.105841.

[18] A. Koptelov, H. Beketova, J. P.-H. Belnoue, S. R. Hallett, and I. Tretiak, "Addressing data scarcity in deep learning: Leveraging real and artificial datasets to predict compaction of composites," *Materials & Design*, vol. 257, Art. no. 114536, Aug. 2025, doi: 10.1016/j.matdes.2025.114536.

[19] J. Han, J. Pei, and H. Tong, *Data mining: Concepts and Techniques*. Morgan Kaufmann, 2022.

[20] A. Amado, P. Cortez, P. Rita, and S. Moro, "Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis," *European Research on Management and Business Economics*, vol. 24, no. 1, pp. 1–7, Jan. 2018, doi: 10.1016/j.iedeen.2017.06.002.

[21] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci*., vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.

[22] S. Maldonado, J. López, and M. Carrasco, "The cobb-douglas learning machine," *Pattern Recognition*, vol. 128, Art. no. 108701, Apr. 2022, doi: 10.1016/j.patcog.2022.108701.

[23] C. Jin, L.-J. Deng, T.-Z. Huang, and G. Vivone, "Laplacian pyramid networks: A new approach for multispectral pansharpening," *Information Fusion*, vol. 78, pp. 158–170, Sep. 2021, doi: 10.1016/j.inffus.2021.09.002.

[24] D. Chicco, M. J. Warrens and G. Jurman, "The matthews correlation coefficient (MCC) is more informative than cohen's kappa and brier score in binary classification assessment," *IEEE Access*, vol. 9, pp. 78368–78381, 2021, doi: 10.1109/ACCESS.2021.3084050.

[25] Y. Liu, J. Li, and H. Zhang, "Explainable machine learning for customer purchase intention prediction," *Expert Systems with Applications*, vol. 213, Art. no. 119125, 2023, doi: 10.1016/j.eswa.2022.119125.

[26] H. Kim, N. Trimi, and J.-H. Chung, "Explainable ensemble learning for consumer behavior prediction," *Expert Systems with Applications*, vol. 197, Art. no. 116712, 2022, doi: 10.1016/j.eswa.2022.116712.

[27] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, "Sentiment analysis on the use of the shopee application using the support vector machine algorithm (SVM)," (in Indonesian), *Jambura Journal of Electrical and Electronics Engineering*, vol. 5, no. 1, pp. 32–35, Jan. 2023, doi: 10.37905/jjeee.v5i1.16830.

[28] L. S. R. Nogueira *et al*., "A comparative study of ensemble and non-ensemble machine learning methods for predicting river pollution index," Ecol. Inform., vol. 94, Art. no. 103617, Mar. 2026, doi: 10.1016/j.ecoinf.2026.103617.

[29] S.-F. Xia, S.-T. Guo, Z. Qu, and Y.-Y. Yang, "Robust patchmatch HDR image reconstruction for deghosting," *Pattern Recognition Letters*, vol. 154, pp. 68–74, Jan. 2022, doi: 10.1016/j.patrec.2022.01.012.

[30] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, Aug. 2018, doi: 10.1016/j.aci.2018.08.003.

[31] G. Zhao, Y. Luo, Q. Chen, and X. Qian, "Aspect-based sentiment analysis via multitask learning for online reviews," *Knowledge-Based Systems*, vol. 264, Art. no. 110326, Jan. 2023, doi: 10.1016/j.knosys.2023.110326.

[32] K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. MIT Press, 2022.