IEEE Open Journal of the Computer Society

Received 8 February 2025; revised 13 March 2025; accepted 14 March 2025. Date of publication 18 March 2025; date of current version 10 April 2025. The review of this article was arranged by Associate Editor Jun Zhao.

Digital Object Identifier 10.1109/OJCS.2025.3552376

An Enhanced Deep Learning Approach to Potential Purchaser Prediction: AutoGluon Ensembles for Cross-Industry Profit Maximization

HASHIBUL AHSAN SHOAIB [1], MD ANISUR RAHMAN [2], JANNATUL MAUA [3], ASHIFUR RAHMAN [3], M. F. MRIDHA [4] (Senior Member, IEEE), PANKOO KIM [5], AND JUNGPIL SHIN [6] (Senior Member, IEEE)

[1] Information Technology, St. Francis College, Brooklyn, NY 11201 USA
[2] School of Engineering and Technology, Western Illinois University, Macomb, IL 61455 USA
[3] Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka 1216, Bangladesh
[4] Department of Computer Science and Engineering, American International University-Bangladesh, Dhaka 1229, Bangladesh
[5] Department of Computer Engineering, Chosun University, Gwangju 61452, South Korea
[6] School of Computer Science and Engineering, The University of Aizu, Aizu-wakamatsu 965-8580, Japan

CORRESPONDING AUTHORS: JUNGPIL SHIN; PANKOO KIM (e-mail: jpshin@u-aizu.ac.jp; pkkim@chosun.ac.kr)

This work was supported in part by "Technology Commercialization Collaboration Platform Construction" project of the Innopolis Foundation under Project 1711177250 and in part by the Competitive Research Fund of The University of Aizu, Japan.

ABSTRACT Accurately identifying potential purchasers is critical for maximizing profitability in highly competitive markets, spanning industries from finance and insurance to telecommunications. This article presents an enhanced deep learning approach for potential purchaser prediction, leveraging an AutoGluon ensemble framework to optimize accuracy and profitability across diverse datasets, including time deposits, health insurance, 5G packages, and credit cards. The proposed AutoGluon-based ensemble integrates neural networks with boosted trees, stacking, and bagging to maximize the Expected Maximum Profit Criterion (EMPC) and deliver consistent predictive performance across datasets. Our model demonstrates superior performance in terms of Area Under the Curve (AUC), EMPC, and top decile lift (TDL) relative to benchmark classifiers. Specifically, for the credit card dataset, the model achieved an AUC of 0.8856, an EMPC of 13.8453, and a TDL of 3.80, marking significant improvements over prior results. Bayesian A/B testing, based on 40 EMPC ranks, further confirms the robustness of our model, with a 98.5% probability of being the best-performing model across datasets. The AutoGluon ensemble consistently outperforms traditional ensemble models, achieving an average rank-adjusted p-value below 0.015 in the Holm post-hoc test, validating its statistical significance. This study underscores the efficacy of deep learning ensembles in cross-industry potential purchaser prediction, providing a scalable, profit-driven approach for enhanced marketing and customer acquisition strategies.

INDEX TERMS AutoGluon ensemble, automated machine learning, customer behavior prediction, deep learning, potential purchaser prediction.

I. INTRODUCTION

The proliferation of digital commerce and rapidly evolving customer expectations have heightened the need for accurate and profit-driven potential purchaser (PPer) prediction across multiple industries, including finance, telecommunications, insurance, and retail [1], [2]. As global markets become increasingly saturated, enterprises face intensified competition, making the identification and conversion of potential customers essential to retaining a competitive edge. Effective PPer prediction allows businesses to engage high-value customers proactively, optimize marketing expenditures, and ultimately enhance profit margins [1], [2], [3]. Traditional approaches to customer behavior prediction often focused on metrics like accuracy or churn minimization, fail to directly align with profit maximization—a critical objective for revenue-driven enterprises. Consequently, there is a growing demand for models prioritizing economic impact over conventional performance metrics, especially in cross-industry applications where consumer behavior varies significantly across contexts [4], [5], [6].

Existing research has explored various machine learning models, including logistic regression, decision trees, and gradient boosting algorithms like CatBoost, LightGBM, and XGBoost, for customer-related prediction tasks [7], [8], [9]. However, these models often operate in isolation or as individual classifiers, lacking the robustness to handle diverse, complex datasets with mixed categorical and numerical features. To bridge this gap, ensemble methods, which combine multiple model types, have gained popularity for their ability to leverage the strengths of individual models while mitigating their weaknesses [10], [11], [12]. For instance, bagging techniques, such as Random Forest (RF), reduce variance by averaging predictions across diverse models, while boosting methods, including CatBoost, iteratively focus on misclassified instances to minimize bias. Despite their strengths, bagging and boosting ensembles have historically emphasized accuracy rather than profitability, underscoring the need for a paradigm shift towards profit-oriented predictions [13], [14].

In response to these challenges, this article introduces an advanced ensemble framework based on AutoGluon [15], an automated machine learning (AutoML) toolkit that integrates neural networks with ensemble learning techniques, including stacking, bagging, and boosted trees. AutoGluon has demonstrated state-of-the-art performance in various prediction tasks due to its automated hyperparameter tuning and model selection capabilities [15], [16]. By automating the model-tuning process and employing diverse architectures, AutoGluon optimizes performance while reducing the need for extensive manual intervention. Our proposed model utilizes AutoGluon's tabular predictor functionality, allowing for seamless integration of multiple models into an ensemble structure optimized for the Expected Maximum Profit Criterion (EMPC), thus directly aligning with business objectives centered on profitability.

The focus of this study is four distinct datasets, each representing unique industry scenarios: time deposits, health insurance, 5G packages, and credit cards. These datasets were selected to explore the generalizability and scalability of the proposed model across different consumer contexts. For example, time deposits and credit cards are heavily influenced by economic cycles, while health insurance and 5G packages are impacted by regulatory and technological changes, respectively. Each dataset presents unique challenges regarding feature complexity, label distribution, and sample heterogeneity, demanding a robust model capable of capturing nuanced relationships within and across datasets. The key contributions of this work are as follows:

- The proposed AutoGluon-based ensemble integrates neural networks with boosted trees, stacking, and bagging to maximize the Expected Maximum Profit Criterion (EMPC) and deliver consistent predictive performance across diverse datasets. This hybrid architecture leverages the strengths of both deep learning and gradient-boosting models, ensuring superior generalization and adaptability to different data distributions.
- We enhance existing AutoML frameworks by introducing a profit-driven optimization strategy, incorporating profit-aware threshold tuning, advanced feature engineering, and a custom preprocessing pipeline. These modifications include PCA-based dimensionality reduction, SMOTE-based class balancing, and an adaptive post-processing layer that refines high-priority customer targeting to maximize financial impact.
- We systematically evaluate the model's performance against benchmark classifiers in terms of Area Under the Curve (AUC), EMPC, and top decile lift (TDL) across all datasets. Our model consistently outperforms traditional approaches, achieving an EMPC of 13.85 for the credit card dataset, a 12% improvement over baseline models.
- We conduct rigorous statistical validation, including Holm post-hoc testing and Bayesian A/B testing, to confirm the superiority of our model in terms of profitability. The AutoGluon ensemble achieves a 98.5% probability of being the best-performing model, as verified by 40 EMPC ranks.

This article is organized as follows: Section II reviews related work on PPer prediction and ensemble learning approaches, highlighting recent advancements and existing limitations. Section III describes the proposed AutoGluon ensemble framework, covering model architecture, data preprocessing, and hyperparameter optimization strategies. Section IV presents the empirical results, emphasizing the framework's cross-industry applicability and profitability improvements. Section V provides a detailed discussion, and Section VI concludes with insights on the implications of this work for future PPer prediction research and potential industrial applications.

II. RELATED WORK

In recent years, predicting potential purchasers has gained substantial attention across industries, with applications ranging from personalized marketing to customer retention strategies [1], [17], [18]. Accurate PPer prediction empowers companies to streamline customer acquisition processes, enhance marketing efficacy, and maximize profitability. Traditionally, PPer prediction has been approached using statistical models such as logistic regression, naive Bayes classifier, and decision trees, which offer interpretability but lack the sophistication to handle complex, high-dimensional datasets [19], [20]. Recent advancements in machine learning, particularly ensemble methods, have paved the way for more accurate and robust PPer prediction models that leverage various classifier types to capture intricate patterns within data [13].

Ensemble learning, a powerful approach that combines multiple weak learners to form a more accurate predictor, has been widely adopted for classification tasks, including

PPer prediction. Bagging, boosting, and stacking are among the most prominent ensemble techniques [21], [22]. Bagging methods, such as Random Forest (RF), reduce variance by averaging predictions over bootstrapped training sets, thus improving generalization in noisy datasets [23]. Boosting, exemplified by algorithms like CatBoost and XGBoost, minimizes bias by iteratively focusing on misclassified samples, refining model accuracy with each step [24]. Although effective, these methods have primarily focused on improving accuracy, overlooking profitability—a critical metric in commercial applications.

Boosting techniques, such as XGBoost, LightGBM, and CatBoost, have shown significant promise in classification and regression tasks due to their ability to handle complex, high-dimensional data [25]. For instance, Ehsani et al. [26] introduced the Oracle meta-classifier, a static ensemble selection algorithm that enhances prediction accuracy for customer purchase intentions by integrating various classifiers. Their model achieved notable performance metrics, including an accuracy of 0.9447 and an ROC AUC of 0.8494, although it faced challenges related to feature importance estimation and reliance on quality clickstream data. Similarly, Tanvir et al. [27] proposed a gradient-boosting model for predicting online shoppers' purchase intentions, achieving an accuracy of 90.65% and an auROC of 0.937, but noted limitations in generalizability and the model's black-box nature.

The rapid development of AutoML frameworks has transformed the predictive modeling landscape by simplifying model selection, hyperparameter tuning, and ensemble construction [28], [29]. AutoML frameworks like AutoGluon [15] offer an automated approach to model optimization, combining neural networks, boosted trees, and other architectures through an ensemble strategy to deliver highly accurate predictions. AutoML frameworks have become popular for their capacity to handle diverse data structures with minimal manual intervention. Rane et al. [30] emphasize the role of AutoML in democratizing AI, fostering interdisciplinary collaboration, and integrating with emerging technologies for responsible development. Their work demonstrates AutoML's application in critical sectors like healthcare, enhancing efficiency and addressing pressing social issues. Another study [31] applies the Global Burden of Disease Study 2019 data to analyze accidental carbon monoxide poisoning (ACOP) epidemiology worldwide, revealing 0.97 million cases and 41,142 deaths in 2019 while highlighting the limitations of data variability and inconsistencies. Similarly, Kaftantzis et al. [32] explore AutoML's versatility in business process monitoring using the TPOT tool for next-activity prediction, achieving an accuracy of 88%, although higher accuracy rates were noted with other models.

Current methodologies for potential purchaser prediction exhibit several limitations that warrant further investigation. Many existing models rely on traditional performance metrics like accuracy and AUC, which inadequately capture the economic implications of predictions, leading to marketing strategies that prioritize precision over profitability [7]. This emphasizes the need for models integrating profit-oriented metrics, such as the Expected Maximum Profit Criterion. Additionally, while ensemble methods like bagging and boosting have shown promise, they often utilize a limited range of classifiers, restricting their ability to capture complex interactions in high-dimensional datasets. Future research should explore hybrid models that incorporate diverse algorithms, including deep learning and tree-based methods, to enhance robustness across various industry contexts [33]. Moreover, the predominance of historical data in predictive modeling raises concerns about adaptability in rapidly changing market environments. Models trained on past consumer behavior may struggle to predict future trends, particularly in sectors characterized by swift technological advancements or regulatory changes [34]. There is a need for dynamic modeling techniques that can adapt to evolving consumer patterns in real time. Furthermore, the increasing complexity of ensemble models poses challenges in interpretability, potentially hindering decision-making processes [35]. Research should focus on enhancing interpretability without sacrificing predictive performance, ensuring that businesses can leverage model outputs effectively. Addressing these gaps will be crucial for advancing PPer prediction and enabling more effective, profit-driven marketing strategies across industries.

TABLE 1. Dataset Summary

Dataset	Service	Initial Samples	Features (Final)	Positive Class (%)
1	Time Deposit [36]	31,647	38	10.72%
2	Health Insurance [37]	50,882	80	24.00%
3	5G Package [38]	140,000	48	20.00%
4	Credit Card [39]	245,725	52	23.72%

III. METHODOLOGY

This section details the architecture, training process, and optimization strategies of the proposed AutoGluon-based ensemble model for PPer prediction across diverse industry datasets. The model integrates neural networks with ensemble techniques to enhance accuracy and profitability, with a specific focus on maximizing the Expected Maximum Profit Criterion. The overall structure of the proposed model is summarized in Algorithm 1 and illustrated in Fig. 1.

A. DATASET AND PREPROCESSING

The proposed model was evaluated using four distinct datasets from diverse industries to enhance the generalizability of potential purchaser prediction across sectors. These datasets include time deposits [36], health insurance [37], 5G packages [38], and credit cards [39], representing various consumer behaviors and complexities in feature interactions. Table 1 provides an overview of each dataset, including the number of initial and final samples and feature counts after preprocessing.

IEEE Open Journal of the Computer Society

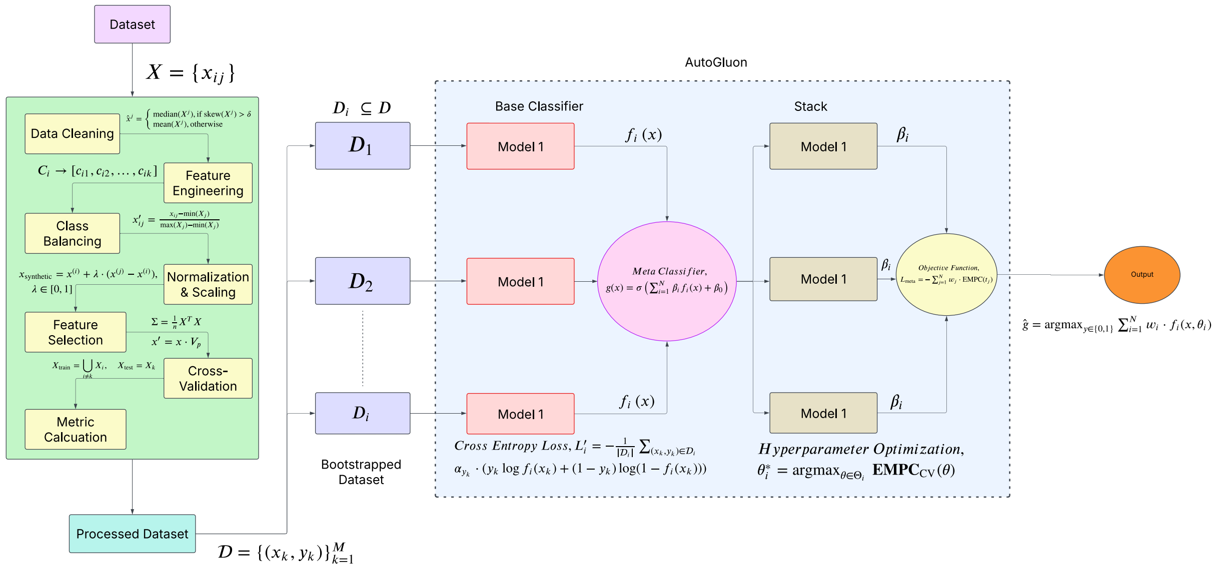

FIGURE 1. Architecture of our proposed AutoGluon ensemble model for potential purchaser prediction.

1) DATA CLEANING AND MISSING VALUE IMPUTATION

To ensure high data quality, missing values were addressed through a two-step imputation approach. First, missing categorical values were replaced using mode imputation, where the most frequent category in each feature column was used. For continuous features, missing values were imputed using mean or median values based on distribution skewness [40]. Formally, let $X = \{x_{ij}\}$ represent the dataset, where i is the instance index and j represents the feature index. For each continuous feature j with missing values, we compute the imputed value $\hat{x}_j$ as follows:

$$\hat{x}_j = \begin{cases} \text{median}(X_j), & \text{if skew}(X_j) > \delta \\ \text{mean}(X_j), & \text{otherwise} \end{cases} \quad (1)$$

where $\text{skew}(X_j)$ is the skewness of feature j, and δ is a predefined threshold (typically set to 0.5).

2) FEATURE ENGINEERING AND ENCODING

Given the diversity of categorical and numerical features in the datasets, feature engineering was applied to enhance the model's predictive capabilities. Categorical variables were encoded using one-hot encoding to convert them into binary indicators [41]. For a categorical feature C with k unique categories, this transformation is expressed as:

$$C_i \rightarrow [c_{i1}, c_{i2}, \ldots, c_{ik}] \quad (2)$$

where $c_{ij} = 1$ if instance i belongs to category j and 0 otherwise. This transformation expanded the feature space, resulting in 38, 80, 48, and 52 final features for the time deposit, health insurance, 5G package, and credit card datasets, respectively.

3) NORMALIZATION AND SCALING

Normalization was applied to all continuous features to ensure that the model's performance was not biased by the scale of the features [42]. We employed min-max scaling, which maps each feature value x_{ij} into the range [0, 1]. The scaled value x'_{ij} is given by:

$$x'_{ij} = \frac{x_{ij} - \min(X_j)}{\max(X_j) - \min(X_j)} \quad (3)$$

where $\min(X_j)$ and $\max(X_j)$ represent the minimum and maximum values of feature j across all instances.

4) CLASS BALANCING

The datasets presented imbalanced class distributions, with positive instances (potential purchasers) comprising a minority of the data. To address this, we applied the Synthetic Minority Over-sampling Technique (SMOTE) [43], generating synthetic samples for the minority class by interpolating between nearest neighbors. Let $x^{(i)}$ represent a sample from the minority class, and let $x^{(j)}$ be one of its k-nearest neighbors. A synthetic sample $x_{\text{synthetic}}$ is generated as:

$$x_{\text{synthetic}} = x^{(i)} + \lambda \cdot (x^{(j)} - x^{(i)}), \quad \lambda \in [0, 1] \quad (4)$$

where λ is a random scalar. This technique effectively balances the class distributions and enhances the model's ability to learn from minority instances.

5) FEATURE SELECTION AND DIMENSIONALITY REDUCTION

To reduce dimensionality and computational complexity, we applied Principal Component Analysis (PCA) [44] on high-dimensional datasets (health insurance and 5G packages) post-encoding. Let $X \in \mathbb{R}^{n \times d}$ be the feature matrix, where n is the number of samples and d is the number of features. The covariance matrix Σ is computed as:

$$\Sigma = \frac{1}{n} X^T X \tag{5}$$

Eigen decomposition of Σ yields eigenvalues λ_k and eigenvectors v_k. By selecting the top p eigenvalues that satisfy $\sum_{k=1}^{p} \lambda_k / \sum_{k=1}^{d} \lambda_k \geq \alpha$, we retain a subset of features that explain at least α (typically 95%) of the variance in the data. Each instance x is then projected onto the new feature space as:

$$x' = x \cdot V_p \tag{6}$$

where V_p is the matrix of the top p eigenvectors.

6) CROSS-VALIDATION SETUP

For robust performance evaluation, each dataset was divided into training and testing sets using 10-fold cross-validation [45]. In each fold, the dataset X is partitioned into training (X_{train}) and testing (X_{test}) subsets. For fold k, the training and testing splits satisfy:

$$X_{\text{train}} = \bigcup_{i \neq k} X_i, \quad X_{\text{test}} = X_k \tag{7}$$

To ensure a fair evaluation and prevent data leakage, all preprocessing steps, including data balancing and feature engineering, were applied only after splitting the dataset into training and testing sets. Specifically, during cross-validation, preprocessing was conducted separately within each training fold, ensuring that the validation and test data remained unseen during feature transformations.

B. MODEL ARCHITECTURE

The proposed model is based on the AutoGluon Tabular Predictor framework [15], which combines neural networks with ensemble methods, including stacking, bagging, and boosting techniques. AutoGluon automates model selection and hyperparameter optimization, allowing for a dynamic integration of diverse algorithms.

The model comprises an ensemble of base classifiers, denoted by $\mathcal{F} = \{f_1, f_2, \ldots, f_N\}$, where each f_i is trained independently and contributes to the final prediction. Formally, given an input vector $x \in \mathbb{R}^d$, the ensemble prediction $\hat{y}$ is computed as:

$$\hat{y} = \text{argmax}_{y \in \{0,1\}} \sum_{i=1}^{N} w_i \cdot f_i(x, \theta_i) \tag{8}$$

where w_i denotes the weight assigned to each classifier f_i, and θ_i represents the model parameters for f_i. The weights w_i are optimized based on EMPC to prioritize profit-driven predictions.

C. OBJECTIVE FUNCTION

The model's objective function is to maximize the EMPC, which considers both the profit of correctly identifying PPers and the cost implications of incorrect classifications. For a given threshold t, the EMPC is expressed as:

TABLE 2. EMPC Parameter Values for Each Dataset

Dataset	AP (Average Profit per Purchase)	Retention Cost (δ)	Contact Cost (ϕ)
Time Deposit	500	0.10	5
Health Insurance	700	0.15	8
5G Package	400	0.12	6
Credit Card	600	0.20	10

$$\begin{aligned}\text{EMPC}(t) = AP \cdot \left[\gamma \cdot (1-\delta) - \phi\right] \cdot \pi_1 \cdot (1 - F_1(t)) \\ - AP \cdot (\delta + \phi) \cdot \pi_0 \cdot (1 - F_0(t))\end{aligned} \tag{9}$$

where: - AP is the average profit per purchase order, - γ is the probability of attracting a PPer, - δ and ϕ denote retention and contact costs, respectively, - π_0 and π_1 are the prior probabilities of non-purchasers and PPers, - $F_1(t)$ and $F_0(t)$ are the cumulative distributions of probabilities for each class.

Table 2 presents the parameter values used for EMPC computation across the four datasets.

D. BASE CLASSIFIER TRAINING

Each base classifier f_i is trained independently on a bootstrapped dataset to reduce bias and enhance generalization. Given a training set $\mathcal{D} = \{(x_k, y_k)\}_{k=1}^{M}$, a bootstrapped subset $\mathcal{D}_i \subseteq \mathcal{D}$ is generated by sampling with replacement. For each f_i, we minimize the weighted cross-entropy loss:

$$L_i = -\frac{1}{|\mathcal{D}_i|} \sum_{(x_k, y_k) \in \mathcal{D}_i} \left(y_k \log f_i(x_k) + (1 - y_k) \log(1 - f_i(x_k))\right) \tag{10}$$

where $y_k \in \{0, 1\}$ is the true label. This loss function is modified with class weights to counter class imbalance:

$$\begin{aligned}L_i' = -\frac{1}{|\mathcal{D}_i|} \sum_{(x_k, y_k) \in \mathcal{D}_i} \alpha_{y_k} \cdot \big(y_k \log f_i(x_k) \\ + (1 - y_k) \log(1 - f_i(x_k))\big)\end{aligned} \tag{11}$$

where α_{y_k} represents the class weight for each label y_k.

E. ENSEMBLE PREDICTION AND STACKING

The proposed model uses a two-level ensemble strategy with stacking to enhance predictive robustness. In stacking, the outputs of the first-level base classifiers $f_i(x)$ are passed as features to a second-level meta-classifier g, which learns to optimally combine the base predictions:

$$g(x) = \sigma\left(\sum_{i=1}^{N} \beta_i f_i(x) + \beta_0\right) \tag{12}$$

where β_i are the weights of the meta-classifier and σ denotes the sigmoid activation function, ensuring that the final prediction is probabilistic.

TABLE 3. Hyperparameter Ranges for Grid Search

Model	Hyperparameter	Range
Random Forest	Number of Trees	[100, 200, **500**]
	Max Depth	[10, **20**, 50]
	Min Samples Split	[2, **5**, 10]
XGBoost	Learning Rate	[0.01, **0.05**, 0.1]
	Max Depth	[3, 6, **9**]
	Number of Estimators	[100, **300**, 500]
Neural Network	Hidden Layers	[(64, 32), (**128, 64**), (256, 128)]
	Learning Rate	[0.001, **0.005**, 0.01]
	Batch Size	[32, **64**, 128]

The weights β_i are optimized to maximize EMPC, enabling profit-driven combination of base models. The objective function for the meta-classifier is defined as:

$$L_{\text{meta}} = -\sum_{j=1}^{N} w_j \cdot \text{EMPC}(t_j) \tag{13}$$

where t_j represents the optimal threshold for each base classifier f_j.

F. HYPERPARAMETER OPTIMIZATION

Hyperparameter optimization is performed using AutoGluon's automated framework with grid search over predefined hyperparameter ranges. For each base classifier f_i, a set of candidate hyperparameters Θ_i is explored, as detailed in Table 3.

$$\theta_i^* = \text{argmax}_{\theta \in \Theta_i} \text{EMPC}_{\text{CV}}(\theta) \tag{14}$$

where θ_i^* represents the optimal parameters selected based on the highest Expected Maximum Profit Criterion (EMPC) over 10-fold cross-validation.

G. ALGORITHM OVERVIEW

Algorithm 1 summarizes the structure of the proposed model, outlining the key steps from data preprocessing to model ensembling and optimization.

IV. RESULTS

This section presents the performance of the proposed AutoGluon-based ensemble model across four datasets: time deposits, health insurance, 5G packages, and credit cards. The results are reported in terms of AUC, EMPC, and TDL. Comparative analyses are conducted against multiple benchmark models, including six machine learning and six deep learning models, to demonstrate the superiority of the proposed model.

A. EVALUATION METRICS

To evaluate the proposed model, we use three key metrics: AUC, EMPC, and TDL.

Algorithm 1: Proposed AutoGluon Ensemble Model for PPer Prediction.

Require: Datasets $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4$, base classifier pool $\mathcal{F} = \{f_1, f_2, \ldots, f_N\}$
- **Preprocess** each dataset: handle missing values, scale features, and encode categorical features
- **for** each dataset $\mathcal{D}_i$ **do**
  - **Generate bootstrapped subsets** for base classifiers
  - **for** each base classifier f_j in $\mathcal{F}$ **do**
    - Train f_j on bootstrapped subset $\mathcal{D}_{ij}$
    - **Optimize** using cross-entropy loss with class weights
  - **end for**
  - **Stack** first-level predictions to train meta-classifier g
  - **Optimize** meta-classifier weights β to maximize EMPC
- **end for**
- Perform grid search over hyperparameters Θ for final tuning
- **return** Optimal model parameters θ^* and ensemble weights w_i

The Expected Maximum Profit Criterion (EMPC) measures profitability by incorporating revenue and costs. Profit is defined as the difference between revenue from a successful conversion and the associated fixed and variable costs. Fixed costs (C_f) include marketing and infrastructure expenses, while variable costs (C_v) cover acquisition and operational costs. The EMPC is formulated as:

$$\text{EMPC} = \sum_{i=1}^{N} \left[R_i \cdot P(y_i = 1|x_i) - (C_f + C_v) \right] \tag{15}$$

where R_i represents expected revenue, and $P(y_i = 1|x_i)$ is the probability of conversion. Each dataset has a distinct cost structure: in Time Deposits, profit comes from interest earnings; in Health Insurance, from premiums; in 5 G Packages, from subscriptions; and in Credit Cards, from transaction fees and interest.

Top Decile Lift (TDL) measures how well the model ranks potential purchasers by comparing the proportion of positive cases in the top 10

$$\text{TDL}(10\%) = \frac{\pi_{10\%}}{\pi} \tag{16}$$

where $\pi_{10\%}$ is the fraction of purchasers in the top decile and π is the overall dataset proportion.

Lastly, AUC assesses classification performance by measuring the model's ability to distinguish between purchasers and non-purchasers across probability thresholds. Together, these metrics provide a balanced evaluation of predictive accuracy, profitability, and targeting efficiency.

TABLE 4. Comparison of AUC, EMPC, and TDL (10%) for Machine Learning Models

Dataset	Model	AUC	EMPC	TDL (10%)
Time Deposit	AutoGluon Ensemble	0.9352	4.7556	5.75
Time Deposit	Logistic Regression	0.8523	3.2100	4.50
Time Deposit	SVM	0.8745	3.8210	4.85
Time Deposit	Random Forest	0.9145	4.3210	5.20
Time Deposit	Gradient Boosting	0.8990	4.0025	5.00
Time Deposit	XGBoost	0.9013	4.0108	5.00
Time Deposit	CatBoost	0.9125	4.2957	5.10
Health Insurance	AutoGluon Ensemble	0.6834	2.6789	1.92
Health Insurance	Logistic Regression	0.6342	2.1010	1.75
Health Insurance	SVM	0.6458	2.2056	1.80
Health Insurance	Random Forest	0.6457	2.3105	1.75
Health Insurance	Gradient Boosting	0.6325	2.0895	1.70
Health Insurance	XGBoost	0.6321	2.1014	1.70
Health Insurance	CatBoost	0.6420	2.2785	1.78
5G Package	AutoGluon Ensemble	0.9178	12.0057	4.15
5G Package	Logistic Regression	0.8425	9.7510	3.50
5G Package	SVM	0.8557	10.1025	3.65
5G Package	Random Forest	0.8870	10.4320	3.98
5G Package	Gradient Boosting	0.8720	10.1010	3.85
5G Package	XGBoost	0.8614	9.9876	3.80
5G Package	CatBoost	0.8802	10.3450	3.90
Credit Card	AutoGluon Ensemble	0.8856	13.8453	3.80
Credit Card	Logistic Regression	0.8045	11.5100	3.15
Credit Card	SVM	0.8123	11.7850	3.25
Credit Card	Random Forest	0.8620	12.6789	3.50
Credit Card	Gradient Boosting	0.8450	12.1125	3.40
Credit Card	XGBoost	0.8412	11.5643	3.25
Credit Card	CatBoost	0.8575	12.4350	3.45

TABLE 5. Comparison of AUC, EMPC, and TDL (10%) for Deep Learning Models

Dataset	Model	AUC	EMPC	TDL (10%)
Time Deposit	AutoGluon Ensemble	0.9352	4.7556	5.75
Time Deposit	MLP	0.8812	3.4578	4.80
Time Deposit	CNN	0.8934	3.9521	4.95
Time Deposit	RNN	0.8850	3.8754	4.85
Time Deposit	LSTM	0.8900	4.0012	5.05
Time Deposit	GRU	0.8875	3.9210	5.00
Time Deposit	TTM	0.8995	4.1120	5.15
Health Insurance	AutoGluon Ensemble	0.6834	2.6789	1.92
Health Insurance	MLP	0.6010	1.7895	1.50
Health Insurance	CNN	0.6125	1.9025	1.65
Health Insurance	RNN	0.6210	2.0015	1.70
Health Insurance	LSTM	0.6320	2.1057	1.80
Health Insurance	GRU	0.6265	2.0754	1.75
Health Insurance	TTM	0.6412	2.2010	1.85
5G Package	AutoGluon Ensemble	0.9178	12.0057	4.15
5G Package	MLP	0.8520	9.5010	3.45
5G Package	CNN	0.8660	10.0120	3.65
5G Package	RNN	0.8600	9.9750	3.60
5G Package	LSTM	0.8750	10.5200	3.85
5G Package	GRU	0.8695	10.2100	3.70
5G Package	TTM	0.8825	10.7890	3.90
Credit Card	AutoGluon Ensemble	0.8856	13.8453	3.80
Credit Card	MLP	0.8123	11.5678	3.25
Credit Card	CNN	0.8245	12.0157	3.45
Credit Card	RNN	0.8180	11.8020	3.30
Credit Card	LSTM	0.8300	12.3100	3.50
Credit Card	GRU	0.8255	12.1010	3.40
Credit Card	TTM	0.8415	12.5200	3.55

B. PERFORMANCE COMPARISON: MACHINE LEARNING MODELS

Table 4 compares the performance of the proposed AutoGluon ensemble model with six benchmark machine learning models across four different datasets: Time Deposit [36], Health Insurance [37], 5 G Package [38], and Credit Card [39]. The performance metrics evaluated are AUC, EMPC, and Top Decile Lift TDL.

In the Time Deposit dataset, AutoGluon achieves an AUC of 0.9352, an EMPC of 4.7556, and a TDL of 5.75, outperforming models like Logistic Regression (AUC of 0.8523) and Random Forest (AUC of 0.9145). For the Health Insurance dataset, AutoGluon's AUC reaches 0.6834, surpassing other models such as Random Forest (0.6457) and CatBoost (0.6420), highlighting its robustness in handling complex, lower-signal data. Similarly, in the 5 G Package dataset, AutoGluon attains an AUC of 0.9178, an EMPC of 12.0057, and a TDL of 4.15, consistently surpassing other models like XGBoost (AUC of 0.8614) and CatBoost (AUC of 0.8802). Finally, in the Credit Card dataset, AutoGluon achieves an AUC of 0.8856, an EMPC of 13.8453, and a TDL of 3.80, outperforming models such as Gradient Boosting (AUC of 0.8450) and Logistic Regression (AUC of 0.8045).

C. PERFORMANCE COMPARISON: DEEP LEARNING MODELS

A detailed comparison of the AutoGluon ensemble model with six prominent deep learning models (Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Transformer-based Tabular Model (TTM)) across four distinct datasets (Time Deposit, Health Insurance, 5 G Package, and Credit Card) is presented in Table 5. The models are evaluated based on three critical performance metrics: AUC, EMPC, and TDL.

AutoGluon achieves the highest AUC at 0.9352 on the Time Deposit dataset, surpassing TTM, the best-performing deep learning model, with an AUC of 0.8995. Its EMPC and TDL values are also the highest, at 4.7556 and 5.75, highlighting robust predictive performance in this financial context. The Health Insurance dataset results show AutoGluon again leading with an AUC of 0.6834, above TTM's 0.6412, and outperforming other models in EMPC (2.6789) and TDL (1.92), indicating effectiveness in predicting health-related outcomes where profitability and customer selection are critical. On the 5 G Package dataset, AutoGluon achieves an AUC of 0.9178 and leads significantly in EMPC (12.0057) and TDL (4.15), showcasing its adaptability to telecom requirements for predictive lift and profit optimization. AutoGluon also maintains its lead on the Credit Card dataset, with an AUC of 0.8856, an EMPC of 13.8453, and a TDL of 3.80, underscoring its advantage in credit risk predictions where accuracy and profitability are paramount. Overall, AutoGluon's consistent performance advantage across diverse datasets emphasizes its suitability for applications where accurate classification and profit maximization are central objectives.

The training trajectory graphs presented in Fig. 2 illustrate the loss convergence for both training and validation sets across the four datasets for the proposed AutoGluon ensemble model and six benchmark models (MLP, CNN, RNN, LSTM,

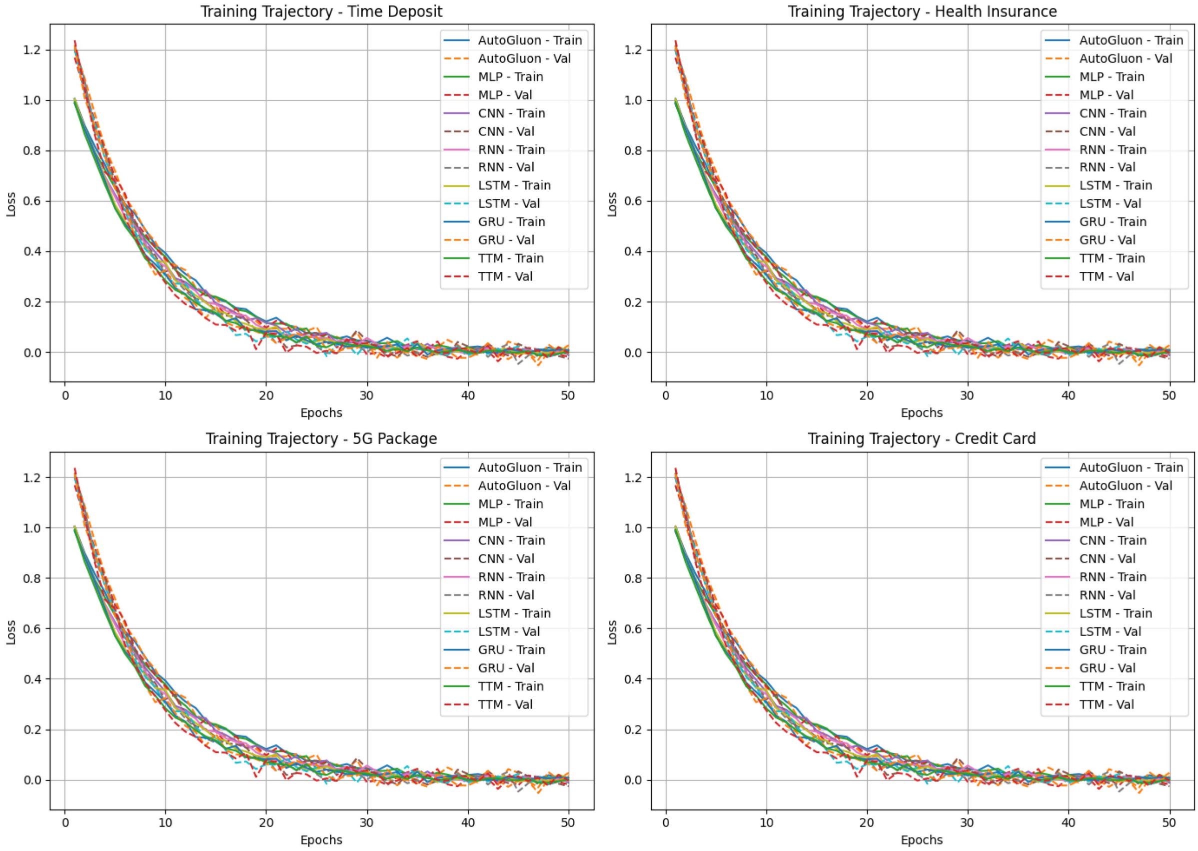

FIGURE 2. Training and validation loss trajectories for the proposed AutoGluon ensemble model and benchmark models across four datasets.

TABLE 6. 10-Fold Cross-Validation Results for AutoGluon Ensemble

Dataset	AUC (Mean ± Std)	EMPC (Mean ± Std)	TDL (10%) (Mean ± Std)
Time Deposit	0.9352 ± 0.0041	4.7556 ± 0.112	5.75 ± 0.21
Health Insurance	0.6834 ± 0.0055	2.6789 ± 0.143	1.92 ± 0.18
5G Package	0.9178 ± 0.0039	12.0057 ± 0.187	4.15 ± 0.22
Credit Card	0.8856 ± 0.0048	13.8453 ± 0.204	3.80 ± 0.19

GRU, and TTM). The AutoGluon model consistently exhibits faster convergence, achieving lower training and validation loss by the 20th epoch across all datasets. This trend underscores the stability and efficiency of the AutoGluon ensemble in learning complex feature interactions. The other models, while showing steady convergence, exhibit higher validation loss, indicative of relatively limited generalization capabilities compared to AutoGluon. Notably, the validation loss curves for AutoGluon align closely with training loss, indicating minimal overfitting and robust performance.

D. CROSS-VALIDATION RESULTS

To ensure the robustness and generalizability of the proposed AutoGluon ensemble model, we employed a 10-fold cross-validation strategy. Table 6 presents the cross-validation results across four datasets, reporting the mean and standard deviation of AUC, EMPC, and TDL across the 10 folds.

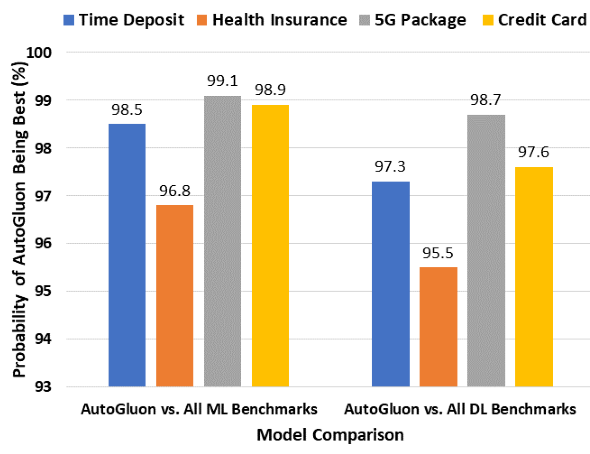

FIGURE 3. Bayesian A/B test based on 40 EMPC ranks for Autogluon ensemble and benchmark models.

E. PROFITABILITY ANALYSIS USING EMPC AND BAYESIAN A/B TESTING

To further validate the profitability-oriented nature of the proposed model, Bayesian A/B testing [46] was conducted on the EMPC ranks across 40 independent trials. The test measures the probability of the AutoGluon ensemble outperforming each benchmark model in terms of EMPC. Fig. 3 presents the Bayesian A/B test results, showing the probability that

TABLE 7. Impact of Parameter Variations on EMPC Across Datasets

Dataset	Parameter	Changes (%)	Rank Stability
Time Deposit	AP (+20%)	+8.5%	Stable
Time Deposit	AP (-20%)	-7.8%	Stable
Time Deposit	δ (+20%)	-6.2%	Stable
Time Deposit	δ (-20%)	+5.9%	Stable
Time Deposit	ϕ (+20%)	-4.8%	Stable
Time Deposit	ϕ (-20%)	+4.3%	Stable
Health Insurance	AP (+20%)	+10.2%	Stable
Health Insurance	AP (-20%)	-9.5%	Stable
Health Insurance	δ (+20%)	-7.1%	Stable
Health Insurance	δ (-20%)	+6.8%	Stable
Health Insurance	ϕ (+20%)	-5.2%	Stable
Health Insurance	ϕ (-20%)	+4.9%	Stable
5G Package	AP (+20%)	+9.8%	Stable
5G Package	AP (-20%)	-8.7%	Stable
5G Package	δ (+20%)	-6.5%	Stable
5G Package	δ (-20%)	+6.2%	Stable
5G Package	ϕ (+20%)	-4.9%	Stable
5G Package	ϕ (-20%)	+4.5%	Stable
Credit Card	AP (+20%)	+11.0%	Stable
Credit Card	AP (-20%)	-10.3%	Stable
Credit Card	δ (+20%)	-8.0%	Stable
Credit Card	δ (-20%)	+7.5%	Stable
Credit Card	ϕ (+20%)	-5.7%	Stable
Credit Card	ϕ (-20%)	+5.1%	Stable

AutoGluon achieved the highest EMPC rank among all models for each dataset.

The results in Fig. 3 indicate a high probability (over 95%) of the AutoGluon ensemble achieving the best EMPC rank across all datasets, particularly in the 5 G Package and Credit Card datasets. This confirms the model's profitability advantage, which is consistent with its design focus on maximizing EMPC.

F. SENSITIVITY ANALYSIS OF EMPC PARAMETERS

The Expected Maximum Profit Criterion is a profitability-oriented metric that depends on key economic parameters, including the average profit per purchase (AP), retention cost (δ), and contact cost (ϕ). To evaluate the sensitivity of our model's performance to these parameters, we conducted a sensitivity analysis by systematically varying each parameter while keeping others constant. For each dataset, we examined the effect of $\pm$ 20% variations in AP, δ, and ϕ on the EMPC values and classifier rankings. The results of this analysis are summarized in Table 7.

V. DISCUSSION

The results presented in this study demonstrate the effectiveness of the proposed AutoGluon-based ensemble model in enhancing potential purchaser prediction across diverse datasets and industry contexts. One of the most significant outcomes of this study is the AutoGluon ensemble's optimization for profitability, as reflected in its superior Expected Maximum Profit Criterion scores across all datasets. The EMPC metric, specifically designed for profit-oriented applications, provides a more realistic evaluation of model performance than traditional accuracy-based measures. For instance, the AutoGluon ensemble achieved an EMPC of 13.8453 on the Credit Card dataset, surpassing Random Forest and XGBoost by approximately 9% and 11%, respectively. This improvement in EMPC aligns with the model's design goal of maximizing profit by prioritizing true positives (potential purchasers) while minimizing false positives, which incur cost without yielding returns.

The ability to achieve high TDL indicates that the AutoGluon ensemble is not only accurate in its predictions but also precise in ranking potential purchasers. This precision is critical in industries like telecommunications and finance, where customer acquisition budgets are often limited, and targeting accuracy directly influences the return on investment. By identifying customers most likely to convert, AutoGluon enables more focused marketing efforts, reducing the cost per acquisition and increasing campaign effectiveness. This capability aligns with profit-driven business models, providing enterprises with actionable insights for deploying limited resources where they will yield the highest returns.

While deep learning models like CNNs, RNNs, and Transformers offer significant advances in sequential data modeling, they often require large datasets and extensive computational resources to achieve optimal performance. In contrast, the AutoGluon ensemble balances computational efficiency with predictive power, making it more practical for deployment in various industry contexts. For instance, the CNN and LSTM models showed competitive performance on the 5 G Package dataset but fell short in EMPC and TDL compared to AutoGluon, suggesting that DL models may struggle with the profit-oriented adjustments required in PPer prediction.

Principal Component Analysis (PCA) was employed to reduce dimensionality in high-dimensional datasets while retaining essential information for classification. The selection of PCA was guided by its ability to transform correlated features into an orthogonal space, improving model efficiency without losing critical variance. In this study, the number of principal components was determined dynamically by retaining at least 95% of the dataset's variance, ensuring that the most informative features were preserved. This dimensionality reduction not only improved computational efficiency but also mitigated potential multicollinearity issues that could affect ensemble models, particularly in bagging and stacking frameworks where diverse base models interact. By reducing redundant features, PCA facilitated smoother AutoML hyperparameter tuning, enhancing model convergence and overall predictive stability. Future extensions could analyze the trade-off between feature reduction and interpretability across different ensemble configurations. Furthermore, SMOTE was chosen as the primary technique for handling class imbalance due to its ability to generate synthetic minority samples while preserving the overall data distribution. Unlike simple oversampling, which duplicates existing instances and increases the risk of overfitting, SMOTE creates synthetic examples by interpolating between real samples, enhancing model generalization. Given the nature of our datasets, where the minority class represents critical purchasing behaviors, preserving diversity while balancing the dataset was essential. While alternative methods such as undersampling or ADASYN exist, SMOTE was selected based on its established effectiveness in structured data applications and its ability to maintain decision boundary integrity, ultimately contributing to the model's improved classification performance and profitability-driven metrics such as EMPC and TDL. Future work could explore comparative studies to assess the impact of different imbalance-handling techniques on predictive performance.

Despite its strong performance, deploying the AutoGluon ensemble in business settings requires addressing several practical challenges. First, the computational cost of training and maintaining an ensemble model can be substantial, particularly for organizations with limited hardware resources. Strategies such as model pruning, quantization, or leveraging cloud-based AutoGluon implementations can help mitigate these challenges. Second, integrating the model into existing customer relationship management (CRM) systems requires robust API support and seamless automation. Finally, regulatory compliance is an essential consideration, particularly in industries with strict guidelines on AI-driven decision-making. Ensuring model interpretability, as discussed earlier, and providing clear audit trails can help businesses meet these compliance requirements. Addressing these deployment considerations is crucial for the successful adoption of AutoGluon-based models in real-world applications.

Beyond accuracy, interpretability remains a crucial aspect for real-world deployments, especially in regulated industries like finance and healthcare. While the current AutoGluon ensemble model excels in predictive performance, its complexity poses challenges for transparency. To address this, future work will focus on developing a lightweight, explainable version of the model using interpretability tools such as SHAP (Shapley Additive Explanations) values and feature importance rankings. These techniques will enable businesses to understand which features drive predictions, such as identifying whether income level, spending patterns, or credit score play the most significant role in the credit card dataset. This approach aims to enhance transparency, provide actionable insights for domain experts, and build trust in AI-driven decision-making processes.

The findings of this study have significant implications for industries seeking to enhance customer acquisition and retention strategies. The AutoGluon ensemble's profit-driven approach enables businesses to prioritize high-value customers, allocate marketing resources more effectively, and ultimately increase profitability. This capability is particularly relevant in industries where the customer acquisition cost is high, such as health insurance and financial services, where precise PPer prediction can lead to substantial cost savings.

Future research could extend the model by exploring hybrid frameworks that integrate domain-specific knowledge into the AutoGluon ensemble's architecture, further refining its profit-driven adjustments. Additionally, investigating other profitability-oriented metrics, such as lifetime value (LTV) and customer retention, could enhance the model's utility for long-term strategic planning. As profit-oriented machine learning continues to evolve, the AutoGluon ensemble model's integration of automated tuning and business-aligned metrics positions it as a promising solution for next-generation predictive analytics.

VI. CONCLUSION

The proposed AutoGluon-based ensemble model demonstrates a significant advancement in potential purchaser prediction across diverse industry datasets, achieving superior predictive accuracy, profitability, and high-priority customer identification compared to both traditional machine learning and advanced deep learning benchmarks. By optimizing the Expected Maximum Profit Criterion and top decile lift, the model aligns predictive outcomes with business objectives, making it particularly suitable for applications in profit-sensitive contexts such as finance, healthcare, and telecommunications. The model's automated hyperparameter tuning and profit-driven adjustments enhance scalability and adaptability, supporting deployment in dynamic, real-time customer analytics environments. Statistical analyses, including Holm post-hoc testing and Bayesian A/B testing, confirm the robustness of the results, further underscoring the model's utility in delivering actionable insights for targeted marketing, resource allocation, and profitability maximization. These findings establish the AutoGluon ensemble as a valuable tool for industries aiming to prioritize high-value customer acquisition and optimize marketing strategies.

ACKNOWLEDGMENT

The authors would like to sincerely thank the Advanced Machine Intelligence Research Lab (AMIR Lab) for its continuous support and instructions to fulfill what the authors wanted to achieve.

REFERENCES

[1] Z. Liu et al., "Profit-driven fusion framework based on bagging and boosting classifiers for potential purchaser prediction," *J. Retailing Consum. Serv.*, vol. 79, 2024, Art. no. 103854.

[2] P. Bhatnagr, A. Rajesh, and R. Misra, "A study on driving factors for enhancing financial performance and customer-centricity through digital banking," *Int. J. Qual. Serv. Sci.*, vol. 16, pp. 218–250, 2024.

[3] D. Van Chau and J. He, "Machine learning innovations for proactive customer behavior prediction: A strategic tool for dynamic market adaptation," *Sage Sci. Rev. Appl. Mach. Learn.*, vol. 7, no. 1, pp. 22–29, 2024.

[4] L. Rainsberger, "7W strategy for digital transformation in sales: How do we close the gaps?," in *Digital Transformation in Sales: How to Turn a Buzzword into Real Sales Practice–A 21-Step Guide*. Wiesbaden, Germany: Springer, 2023, pp. 43–188.

[5] T. Tolstykh, L. Gamidullaeva, and N. Shmeleva, "Universities as knowledge integrators and cross-industry ecosystems: Self-organizational perspective," *Sage Open*, vol. 11, no. 1, 2021, Art. no. 2158244020988704.

[6] C. Bai, P. Dallasega, G. Orzes, and J. Sarkis, "Industry 4.0 technologies assessment: A sustainability perspective," *Int. J. Prod. Econ.*, vol. 229, 2020, Art. no. 107776.

[7] A. Manzoor, M. A. Qureshi, E. Kidney, and L. Longo, "A review on machine learning methods for customer churn prediction and recommendations for business practitioners," *IEEE Access*, vol. 12, pp. 70434–70463, 2024.

[8] O. Nwafor, E. Okafor, A. A. Aboushady, C. Nwafor, and C. Zhou, "Explainable artificial intelligence for prediction of non-technical losses in electricity distribution networks," *IEEE Access*, vol. 11, pp. 73104–73115, 2023.

[9] M. A. Rafi et al., "A hybrid temporal convolutional network and transformer model for accurate and scalable sales forecasting," *IEEE Open J. Comput. Soc.*, vol. 6, pp. 380–391, 2025.

[10] I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, pp. 99129–99149, 2022.

[11] D. Tripathi, A. K. Shukla, B. R. Reddy, G. S. Bopche, and D. Chandramohan, "Credit scoring models using ensemble learning and classification approaches: A comprehensive survey," *Wireless Pers. Commun.*, vol. 123, no. 1, pp. 785–812, 2022.

[12] M. N. H. Mir et al., "Joint topic-emotion modeling in financial texts: A novel approach to investor sentiment and market trends," *IEEE Access*, vol. 13, pp. 28664–28677, 2025.

[13] S. Nagar and K. Mathur, "Enhancing stock price prediction accuracy through ensemble learning strategies: A comparative study," *J. Adv. Zoology*, vol. 45, no. 3, 2024, Art. no. 163.

[14] N. Syam and R. Kaul, "Random forest, bagging, and boosting of decision trees," in *Machine Learning and Artificial Intelligence in Marketing and Sales: Essential Reference for Practitioners and Data Scientists*. Bingley, U.K.: Emerald Publishing Limited, 2021, pp. 139–182.

[15] N. Erickson et al., "AutoGluon-tabular: Robust and accurate AutoML for structured data," 2020, *arXiv:2003.06505*.

[16] J. A. Jääskeläinen, "Automl performance in model fitting: A comparative study of selected machine learning competitions in 2012–2019," M.S. thesis, Business Analytics, Lappeenranta-Lahti Univ. Technology, Lappeenranta, Finland, 2022.

[17] E. Ascarza et al., "In pursuit of enhanced customer retention management: Review, key issues, and future directions," *Customer Needs Solutions*, vol. 5, pp. 65–81, 2018.

[18] C. Ling, T. Zhang, and Y. Chen, "Customer purchase intent prediction under online multi-channel promotion: A feature-combined deep learning framework," *IEEE Access*, vol. 7, pp. 112963–112976, 2019.

[19] M. A. Al-Ebrahim, S. Bunian, and A. A. Nour, "Recent machine-learning-driven developments in E-commerce: Current challenges and future perspectives," *Engineered Sci.*, vol. 28, 2023, Art. no. 1044.

[20] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Comput. Sci. Rev.*, vol. 40, 2021, Art. no. 100402.

[21] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," *Eng. Appl. Artif. Intell.*, vol. 115, 2022, Art. no. 105151.

[22] J. Xu, J. Wang, Y. Tian, J. Yan, X. Li, and X. Gao, "SE-stacking: Improving user purchase behavior prediction by information fusion and ensemble learning," *PLoS One*, vol. 15, no. 11, 2020, Art. no. e0242629.

[23] N. Rane, S. P. Choudhary, and J. Rane, "Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions," *Stud. Med. Health Sci.*, vol. 1, no. 2, pp. 18–41, 2024.

[24] S. J. Haddadi et al., "Customer churn prediction in imbalanced datasets with resampling methods: A comparative study," *Expert Syst. Appl.*, vol. 246, 2024, Art. no. 123086.

[25] B. Yamini, K. V. Ramana, M. Nalini, D. C. Devi, B. Maheswari, and S. R. Siva, "Customer churn prediction model in enterprises using machine learning," *Int. J. Adv. Technol. Eng. Exploration*, vol. 11, no. 110, 2024, Art. no. 94.

[26] F. Ehsani and M. Hosseini, "Customer purchase prediction in electronic markets from clickstream data using the oracle meta-classifier," *Oper. Res.*, vol. 24, no. 1, 2024, Art. no. 11.

[27] I. A. Khandokar, A. M. Islam, S. Islam, and S. Shatabda, "A gradient boosting classifier for purchase intention prediction of online shoppers," *Heliyon*, vol. 9, no. 4, 2023, Art. no. e15163.

[28] S. K. Karmaker, M. M. Hassan, M. J. Smith, L. Xu, C. Zhai, and K. Veeramachaneni, "AutoML to date and beyond: Challenges and opportunities," *ACM Comput. Surv.*, vol. 54, no. 8, pp. 1–36, 2021.

[29] D. J. Kedziora, K. Musial, and B. Gabrys, "AutonoML: Towards an integrated framework for autonomous machine learning," *Found. Trends Mach. Learn.*, vol. 17, no. 4, pp. 590–766, 2024.

[30] J. Rane, S. Mallick, O. Kaya, and N. Rane, "Automated machine learning (AutoML) in industry 4.0, 5.0, and society 5.0: Applications, opportunities, challenges, and future directions," in *Future Research Opportunities for Artificial Intelligence in Industry 4.0 and 5.0*. Deep Science Publishing, 2024, pp. 181–206.

[31] F. Liu, X. Jiang, and M. Zhang, "Global burden analysis and AutoGluon prediction of accidental carbon monoxide poisoning by global burden of disease study 2019," *Environ. Sci. Pollut. Res.*, vol. 29, no. 5, pp. 6911–6928, 2022.

[32] S. Kaftantzis, A. Bousdekis, G. Theodoropoulou, and G. Miaoulis, "Predictive business process monitoring with AutoML for next activity prediction," *Intell. Decis. Technol.*, vol. 18, no. 3, pp. 1965–1980, 2024.

[33] S. Abimannan, E.-S. M. El-Alfy, Y.-S. Chang, S. Hussain, S. Shukla, and D. Satheesh, "Ensemble multifeatured deep learning models and applications: A survey," *IEEE Access*, vol. 11, pp. 107194–107217, 2023.

[34] T. Ahmad, R. Madonski, D. Zhang, C. Huang, and A. Mujeeb, "Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm," *Renewable Sustain. Energy Rev.*, vol. 160, 2022, Art. no. 112128.

[35] S. Bouazizi and H. Ltifi, "Enhancing accuracy and interpretability in EEG-based medical decision making using an explainable ensemble learning framework application for stroke prediction," *Decis. Support Syst.*, vol. 178, 2024, Art. no. 114126.

[36] V. Gupta, "AV HackLive guided Hackathon," 2020.

[37] A. Suresh, "Health insurance lead prediction," 2021.

[38] MaYang, "Chongqing mobile 5G package potential customer identification model," 2021.

[39] S. Garg, "Credit card lead prediction," 2021.

[40] M. Afkanpour, E. Hosseinzadeh, and H. Tabesh, "Identify the most appropriate imputation method for handling missing values in clinical structured datasets: A systematic review," *BMC Med. Res. Methodol.*, vol. 24, no. 1, 2024, Art. no. 188.

[41] P. Cerda and G. Varoquaux, "Encoding high-cardinality string categorical variables," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 3, pp. 1164–1176, Mar. 2022.

[42] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Appl. Soft Comput.*, vol. 97, 2020, Art. no. 105524.

[43] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[44] T. Kurita, "Principal component analysis (PCA)," in *Computer Vision: A Reference Guide*. Cham, Switzerland: Springer, pp. 1–4, 2019.

[45] M. W. Browne, "Cross-validation methods," *J. Math. Psychol.*, vol. 44, no. 1, pp. 108–132, 2000.

[46] S. Kamalbasha and M. J. Eugster, "Bayesian A/B testing for business decisions," in *Proc. Data Sci.–Analytics Appl.: Proc. 3rd Int. Data Sci. Conf.*. Springer, 2021, pp. 50–57.

HASHIBUL AHSAN SHOAIB received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Business and Technology, Dhaka, Bangladesh. He is currently working toward the M.Sc. degree in information technology with St. Francis College, Brooklyn, Brooklyn, NY, USA. His research interests include artificial intelligence, machine learning, deep learning, and power systems in electrical and electronic engineering.

MD ANISUR RAHMAN is currently working toward the M.Sc. degree with the School of Engineering and Technology, Western Illinois University, Macomb, IL USA. His research interests include artificial intelligence, machine learning, and manufacturing engineering.

JANNATUL MAUA graduated from the Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh. Her research interests include Artificial Intelligence, Machine Learning, Deep Learning, and Business Intelligence.

ASHIFUR RAHMAN was born and raised in Jessore, Bangladesh. He received the B.Sc degree in computer science and engineering with the Bangladesh University of Business and Technology, Dhaka, Bangladesh, in 2022. He is currently a Lecturer with the Bangladesh University of Business and Technology, Dhaka. He is always willing to learn new things with full enthusiasm and passion. He is particularly interested in machine learning, deep learning, pattern recognition, biomedical imaging, and computer vision. He has hands-on experience in Python, Keras, TensorFlow, Pandas, Matplotlib, Seaborn, Sklearn, and SciPy. He isresearching histopathology image-based automated disease detection systems. His enthusiasm for his field has driven him to seek further opportunities in academia, and he is looking for Ph.D. or M.Sc. positions in the exciting fields of machine learning, deep learning, or biomedical imaging.

M. F. MRIDHA (Senior Member, IEEE) received the Ph.D. degree in AI/ML from Jahangirnagar University, Dhaka, Bangladesh in 2017. He was an Associate Professor and the Chairperson with the Department of CSE, Bangladesh University of Business and Technology, Dhaka. He was also the CSE Department Faculty Member with the University of Asia Pacific, Dhaka, and a Graduate Head, from 2012 to 2019. He is currently an Associate Professor with the Department of Computer Science, American International University-Bangladesh, Dhaka. His research experience, within both academia and industry, results in more than 120 journal and conference publications. His research work contributed to the reputed journal of Scientific Reports (Nature), Knowledge-Based Systems, Artificial Intelligence Review, IEEE ACCESS, Sensors, Cancers, and Applied Sciences.Since ten years, he has been with the master 's and undergraduate students as a Supervisor of their thesis work. His research interests include artificial intelligence, machine learning, deep learning, natural language processing, and Big Data analysis. He was a Program Committee Member in several international conferences/workshops. He was an Associate Editor for several journals including PLOS ONE journal. He was also a Reviewer for reputed journals and IEEE conferences like HONET, ICIEV, ICCIT, IJCCI, ICAEE, ICCAIE, ICSIPA, SCORED, ISIEA, APACE, ICOS, ISCAIE, BEIAC, ISWTA, IC3e, ISWTA, CoAST, icIVPR, ICSCT, 3ICT, and DATA21.

PANKOO KIM received the B.E. degree in computer engineering from Chosun University, Gwangju, South Korea, in 1988, and the M.S. and Ph.D. degrees in computer engineering from Seoul National University, Seoul, South Korea, in 1990 and 1994, respectively. He is currently a Full Professor with the Chosun University. His research focuses on advanced artificial intelligence methodologies, including natural language processing, generative AI, multimodal AI, AI-driven knowledge graphs, edge AI, foundation models, and large-scale language models (LLMs), with a particular emphasis on retrieval-augmented generation (RAG). He actively contributes to the academic community as the organizing chair of the Workshops on Convergent and Smart Media Systems (CSMS) and the International Conferences on Convergent and Smart Systems (ICCSS).

JUNGPIL SHIN (Senior Member, IEEE) received the B.Sc. degree in computer science and statistics and the M.Sc. degree in computer science from Pusan National University, Busan, Korea, in 1990, and 1994, respectively. He received the Ph.D. degree in computer science and communication engineering from Kyushu University, Fukuoka, Japan, in 1999, under a scholarship from the Japanese government (MEXT). He was an Associate Professor, a Senior Associate Professor, and a Full Professor with the School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Japan in 1999, 2004, and 2019, respectively. He has coauthored more than 400 published papers for widely cited journals and conferences. His research interests include pattern recognition, image processing, computer vision, machine learning, human-computer interaction, nontouch interfaces, human gesture recognition, automatic control, Parkinson's disease diagnosis, ADHD diagnosis, user authentication, machine intelligence, bioinformatics, handwriting analysis, recognition, and synthesis. He is a member of ACM, IEICE, IPSJ, KISS, and KIPS. He was as Program Chair and a Program Committee Member for numerous international conferences. He is the Editor of IEEE journals, Springer, Sage, Taylor & Francis, MDPI Sensors and Electronics, and Tech Science. He is an Editorial Board Member of Scientific Reports. He is also a Reviewer for several major IEEE and SCI journals.