Comparative Analysis of Ensemble Learning Techniques for Purchase Prediction in Digital Promotion through Social Network Advertising

Andhika Rafi Hananto[1,*], Bhavana Srinivasan[2]

[1]Magister of Computer Science, Universitas Gadjah Mada, Yogyakarta, Indonesia

[2]Department of Animation and Virtual Reality, JAIN, Bangalore, India

ABSTRACT

This study conducts a comprehensive comparative analysis of ensemble learning techniques for predicting user purchases in social network advertising. The ensemble methods evaluated include Random Forest, Gradient Boosting Machines (GBM), AdaBoost, and Bagging. The dataset, consisting of 7,000 records of user interactions with social network advertisements, was preprocessed to handle missing values, encode categorical variables, and standardize numerical features. Performance metrics such as accuracy, precision, recall, F1 score, and ROC AUC score were used to evaluate each model. The Random Forest model achieved an accuracy of 0.875, precision of 0.821, recall of 0.821, F1 score of 0.821, and ROC AUC score of 0.948. The GBM model also performed well, with an accuracy of 0.875, precision of 0.846, recall of 0.786, F1 score of 0.815, and ROC AUC score of 0.948. The AdaBoost model showed the highest performance, with an accuracy of 0.9, precision of 0.917, recall of 0.786, F1 score of 0.846, and ROC AUC score of 0.969. The Bagging model achieved an accuracy of 0.875, precision of 0.821, recall of 0.821, F1 score of 0.821, and ROC AUC score of 0.939. Feature importance analysis revealed that Age and Estimated Salary were the most significant predictors across all models. Hyperparameter tuning was crucial in optimizing each model's performance, ensuring they were neither too simple nor too complex. The study's findings underscore the effectiveness of ensemble learning techniques in social network advertising and provide valuable insights for marketers. Future research could explore larger and more diverse datasets, other ensemble methods, and the computational efficiency of these models. This research contributes to predictive analytics in marketing, enhancing the accuracy and effectiveness of advertising strategies.

Keywords Ensemble Learning, Purchase Prediction, Social Network Advertising, Random Forest, AdaBoost

Submitted 10 January 2024
Accepted 20 April 2024
Published 1 June 2024

Corresponding author
Andhika Rafi Hananto,
andhikarafihananto@mail.ugm.ac.id

Additional Information and Declarations can be found on page 142

DOI: 10.47738/jdmdc.v1i2.7

INTRODUCTION

In the contemporary digital landscape, social network advertising has emerged as a pivotal element in the marketing strategies of businesses worldwide. Social network advertising, also known as social media advertising, is a powerful tool that allows advertisers to target specific audiences with precision and efficiency [1]. One unique aspect of social networking site advertising is its ability to leverage users' social connections by encouraging them to share advertisements within their networks, thereby increasing product exposure and reach [2]. This form of advertising can be highly effective as it taps into the social influence and network effects present on these platforms [3].

Platforms like Facebook, Instagram, and Twitter boast billions of active users,

How to cite this article: A. R. Hananto and B. Srinivasan, "Comparative Analysis of Ensemble Learning Techniques for Purchase Prediction in Digital Promotion through Social Network Advertising," *J. Digit. Mark. Digit. Curr.*, vol. 1, no. 2, pp. 125-143, 2024.

presenting unparalleled opportunities for companies to reach potential customers. Predicting whether a user will make a purchase based on their interactions with advertisements is essential for optimizing advertising spend, enhancing user experience, and ultimately maximizing return on investment (ROI). Accurate purchase prediction enables marketers to target the right audience, personalize ad content, and improve conversion rates, making it a cornerstone of successful digital marketing campaigns.

Over the past decade, machine learning has significantly transformed predictive analytics. Unlike traditional statistical methods, machine learning models can automatically learn and improve from experience without being explicitly programmed. This capability to process vast amounts of data and uncover hidden patterns makes machine learning particularly suited for complex tasks like purchase prediction in social network advertising. Techniques such as logistic regression, decision trees, and neural networks have been widely used. However, ensemble learning methods have demonstrated superior performance in many cases due to their robustness and ability to reduce overfitting.

Ensemble learning involves combining multiple individual models to create a more powerful predictive model. The primary types of ensemble methods include bagging (bootstrap aggregating) and boosting. Bagging involves training multiple instances of the same algorithm on different subsets of the training data, created through bootstrap sampling, and then averaging their predictions. This approach helps to reduce variance and improve the stability of the model. A prominent example of a bagging method is the Random Forest algorithm, which combines the predictions of multiple decision trees. Boosting, on the other hand, involves sequentially training a series of models, where each model attempts to correct the errors of its predecessor. This technique focuses on reducing bias and improving the accuracy of the predictions. Examples of boosting methods include Gradient Boosting Machines (GBM) and AdaBoost. By leveraging these ensemble learning techniques, researchers and practitioners can achieve more accurate and reliable purchase predictions in social network advertising, thereby driving better business outcomes.

Predicting user purchase behavior in social media ads presents challenges due to the complex interplay of various factors. Research has shown that users' purchase intention is not uniformly influenced by their social media connections or usage, indicating that the behavior of using social media does not always directly translate into increased purchase likelihood on social commerce websites [4]. While social media advertising has been identified as a strong predictor of consumer purchasing behavior [5], it is crucial to consider factors such as credibility, entertainment value, incentives, and social escapism motivation in predicting attitudes towards social media advertising and subsequent purchase intention [6].

Furthermore, the influence of social media influencers on purchase intention and the mediating role of brand awareness have been highlighted as significant factors affecting user behavior [7]. Additionally, the impact of social media advertising on consumer purchase decisions, with brand awareness as a mediating variable, underscores the intricate relationship between advertising content and user actions [8]. Understanding users' purchase intent from the language they use on platforms like Twitter through supervised learning strategies further emphasizes the need for sophisticated analytical approaches to predict user behavior accurately [9]. Factors such as motivation, congruity, attitudes towards ads on social media, and word-of-mouth marketing have been identified as key drivers of ad clicks, which subsequently influence purchase intention [10]. Research also delves into how consumer online motivations, whether driven by connection or consumption, lead to ad clicks on social media and impact behavioral intentions, shedding light on the nuanced pathways that influence user actions [11].

The primary objectives of this research are twofold. First, to evaluate the performance of different ensemble learning techniques in predicting user purchases based on social network advertising data. This involves implementing and assessing various methods such as Random Forest, GBM, AdaBoost, and Bagging. Second, the study aims to identify the most effective ensemble learning method for purchase prediction. By comparing these techniques, the research seeks to provide insights into their strengths and weaknesses, thereby guiding practitioners in selecting the most suitable approach for enhancing the accuracy and reliability of their predictive models in social network advertising.

Literature Review

Overview of Ensemble Learning

Ensemble learning is a powerful paradigm in machine learning that enhances predictive performance by aggregating the predictions of multiple models. The fundamental concept behind ensemble learning is that while individual models might be prone to errors and overfitting, a collective model that integrates the strengths of several models can achieve higher accuracy and robustness. This method reduces the likelihood of overfitting and improves the generalization capabilities of the model. Ensemble learning techniques capitalize on the diversity among individual models, allowing them to correct each other's mistakes and provide a more reliable prediction.

There are several types of ensemble methods, with Bagging, Boosting, and Random Forest being the most prominent. Bagging and boosting are essential techniques in ensemble learning. Bagging is a parallel ensemble method that involves training multiple base learners independently on different subsets of the training data and then combining their predictions through techniques like averaging or voting [12]. On the other hand, boosting is a sequential ensemble method that focuses on improving the performance of weak learners by assigning weights to the training instances and adjusting them during the learning process to give more emphasis to misclassified instances [13]. Both bagging and boosting have gained popularity due to their theoretical performance guarantees and empirical success [14]. These methods aim to promote diversity among the base learners to enhance the overall predictive performance of the ensemble model [15]. Bagging achieves this by manipulating input data through bootstrap aggregation, while boosting re-weights samples to focus on instances that are harder to classify correctly. Popular boosting algorithms include GBM and AdaBoost. These algorithms have been shown to be highly effective in a wide range of predictive tasks, often outperforming single models by a significant margin.

Random Forest

Random Forest is a specific type of bagging method that constructs a multitude of decision trees during training. Each tree in a Random Forest is built from a different bootstrap sample of the data, and during the construction of each tree, a random subset of features is selected at each split. This randomness introduces diversity among the trees, leading to a robust ensemble model that mitigates the risk of overfitting. Random Forests are particularly effective in handling large datasets with numerous features and have been widely used in various predictive modeling applications, including classification and regression tasks. The method's ability to handle high-dimensional data and provide insights into feature importance makes it a versatile tool in predictive analytics.

In predictive modeling, Random Forests have been applied across numerous domains due to their flexibility and strong performance. They are used in finance for credit scoring and risk management, in healthcare for disease prediction and patient outcome forecasting, and in marketing for customer segmentation and behavior prediction. The algorithm's robustness to noisy data and its capability to handle both categorical and continuous variables further underscore its utility in real-world applications. By aggregating the results of multiple decision trees, Random Forests deliver a powerful and reliable predictive model that excels in accuracy and interpretability.

Gradient Boosting Machines (GBM)

GBM is a powerful ensemble learning technique that builds models sequentially. Each new model is trained to correct the errors made by the previous models. This process is based on the principle of gradient descent, where the model minimizes a loss function by iteratively improving the predictions. GBMs are known for their robustness and ability to handle various types of data, making them a versatile choice for many predictive modeling tasks.

GBMs have found applications across a wide range of domains due to their high predictive accuracy and flexibility. In finance, GBMs are used for credit scoring, fraud detection, and risk management, where precise predictions are crucial. In healthcare, they help in disease diagnosis, patient outcome predictions, and personalized treatment plans. Marketing professionals use GBMs for customer segmentation, churn prediction, and targeted advertising, enabling more effective and efficient marketing strategies. Additionally, GBMs are employed in fields like e-commerce for recommendation systems, in environmental science for predicting weather patterns, and in many other areas where accurate predictions can drive significant value.

AdaBoost

AdaBoost, short for Adaptive Boosting, is another popular boosting algorithm that enhances the performance of weak learners, typically decision trees, by focusing on the errors of previous models. The mechanism of AdaBoost involves assigning weights to each instance in the training dataset. Initially, all instances are given equal weight. In subsequent iterations, AdaBoost increases the weights of the misclassified instances, prompting the model to focus more on these difficult cases in the next round of training. This iterative process continues until the specified number of models is reached or the performance stops improving.

AdaBoost is particularly effective in scenarios where the primary goal is to achieve high classification accuracy with a simple and interpretable model. It has been widely used in various applications such as face detection, text classification, and bioinformatics. For instance, in face detection, AdaBoost can quickly and accurately identify faces in images by combining the outputs of multiple weak classifiers. In text classification, it helps in categorizing documents into predefined categories with high accuracy. In bioinformatics, AdaBoost aids in gene expression analysis and the prediction of protein functions, where precise classification is essential for advancing scientific research.

Bagging

Bagging, or Bootstrap Aggregating, is an ensemble learning technique designed to improve the stability and accuracy of machine learning algorithms. The core concept of bagging involves generating multiple versions of a training dataset using bootstrap sampling, which is a method of random sampling with replacement. Each version of the dataset is used to train a separate model, often of the same type. The predictions from these individual models are then aggregated, typically by averaging in regression tasks or by majority voting in classification tasks.

Bagging helps reduce variance and prevent overfitting, making it particularly effective for high-variance algorithms like decision trees. One of the most well-known bagging methods is the Random Forest algorithm, which combines the predictions of multiple decision trees to produce a more accurate and robust model. Random Forests not only reduce variance but also introduce additional randomness by selecting a random subset of features at each split in the trees, further enhancing the model's generalization ability.

Bagging has been successfully applied in various predictive tasks across different domains. In finance, bagging techniques are used for credit risk modeling and stock price prediction, where accurate forecasts are essential. In the healthcare sector, bagging algorithms assist in predicting patient outcomes and diagnosing diseases by aggregating the insights from multiple models. In marketing, bagging helps in customer segmentation and predictive analytics, enabling businesses to tailor their strategies more effectively. The robustness of bagging methods makes them suitable for handling noisy datasets and complex prediction tasks, providing reliable and interpretable results.

Gaps in the Literature

Despite the extensive research on ensemble learning techniques, several gaps and limitations remain in the current literature. One major gap is the lack of comprehensive comparative studies that evaluate the performance of various ensemble methods under consistent conditions and datasets. While individual studies may highlight the effectiveness of specific techniques like Random Forest, Gradient Boosting Machines, or AdaBoost, there is a need for a systematic comparison that considers a broad range of ensemble methods applied to the same predictive task.

Another limitation is the tendency of existing research to focus on specific domains or datasets, which may not generalize well to other contexts. For instance, studies in finance or healthcare might show promising results for certain ensemble techniques, but these findings may not directly apply to social network advertising, where user behavior and data characteristics are different.

Furthermore, there is a need for more research on the interpretability and computational efficiency of ensemble methods. While these techniques often provide superior predictive performance, they can be computationally intensive and challenging to interpret. Understanding the trade-offs between accuracy, computational cost, and interpretability is crucial for practical applications.

The justification for this comparative analysis lies in addressing these gaps. By systematically evaluating different ensemble learning methods for purchase prediction in social network advertising, this research aims to provide insights into the relative strengths and weaknesses of each technique. Such a comparative study can guide practitioners in selecting the most appropriate method for their specific needs, ensuring that the chosen approach not only delivers high accuracy but also balances other important factors like computational efficiency and model interpretability. This comprehensive analysis will contribute to the existing body of knowledge and help advance the field of predictive analytics in social network advertising.

Method

The research method for this study consists of several steps to ensure a comprehensive and accurate analysis. The process includes data collection, preprocessing, exploratory data analysis (EDA), feature selection, implementation of ensemble learning techniques, and the experimental setup. Each step is crucial for preparing the data and optimizing the models to predict user purchases based on social network advertising data effectively. The flowchart in figure 1 outlines the detailed steps of the research method.

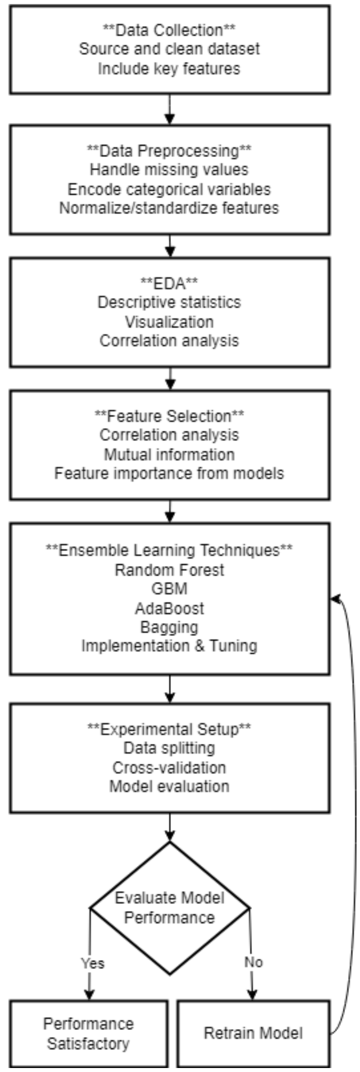

Figure 1 **Research Method Flowchart**

Data Collection

The dataset used in this study is a synthetic representation of user interactions with social network advertisements, sourced from a publicly available repository. It comprises 7,000 records, each corresponding to an individual user's data. The dataset includes several features that are pivotal for predictive modeling: User ID, Gender, Age, Estimated Salary, and Purchased.

User ID serves as a unique identifier for each user, ensuring the distinctiveness of every record. Gender, a categorical variable classified as 'Male' or 'Female,' provides insights into gender-based purchasing trends. Age is a continuous variable that influences purchasing decisions and behaviors, capturing the demographic diversity of users. Estimated Salary, a financial metric, sheds light on the purchasing power and spending behavior of users. The target variable, Purchased, indicates whether a user made a purchase (1) or not (0) after interacting with the advertisement.

This dataset is comprehensive enough to capture various aspects of user demographics and financial status, crucial for predicting purchase behavior.

During the data collection process, the dataset is cleaned and preprocessed to ensure it is suitable for machine learning applications. This involves handling any missing values and encoding categorical variables to facilitate model training. The balanced nature of the dataset across different features aids in building robust predictive models without significant biases, making it an ideal choice for evaluating the performance of various ensemble learning techniques in purchase prediction.

Data Preprocessing

The data preprocessing phase is crucial for preparing the dataset for effective machine learning modeling. This process involves several steps, including handling missing values, feature engineering, and data normalization or standardization.

First, we checked for any missing values in the dataset. Fortunately, there were no missing values in any of the columns, as confirmed by the output of the `isnull().sum()` method, which showed zero missing values for all columns: User ID, Gender, Age, Estimated Salary, and Purchased.

Next, feature engineering was conducted to enhance the dataset. This included encoding categorical variables to make them suitable for machine learning algorithms. Specifically, the Gender column, which contained categorical data ('Male' and 'Female'), was converted into a binary format using one-hot encoding. This transformation resulted in a new column, Gender_Male, where a value of 'True' indicated 'Male' and 'False' indicated 'Female'. This step is crucial for enabling algorithms that require numerical input to process categorical data effectively.

The final step in the preprocessing phase was data normalization or standardization, which ensures that the features have a consistent scale. This is particularly important for algorithms that are sensitive to the scale of the input data. The Age and Estimated Salary columns were standardized using the StandardScaler from the sklearn library. Standardization transformed these features to have a mean of zero and a standard deviation of one. This transformation is reflected in the standardized values of Age and Estimated Salary, making the data more suitable for machine learning models by ensuring that all features contribute equally to the model's performance.

After preprocessing, the dataset was ready for further analysis and modeling, with all features appropriately scaled and encoded to enhance the predictive power of the machine learning algorithms applied.

Exploratory Data Analysis (EDA)

EDA is a critical step in understanding the dataset's underlying structure and relationships between variables. It involves examining descriptive statistics, visualizing key features, and performing correlation analysis to gain insights that guide subsequent modeling efforts. Descriptive statistics provide a summary of the central tendency, dispersion, and shape of the dataset's distribution. The dataset includes 400 records, and key statistics for the features reveal several insights. The User ID feature, which uniquely identifies each user, has a mean of 15,691,540 with a standard deviation of 71,658, ranging from a minimum of 15,566,690 to a maximum of 15,815,240. The Age and Estimated Salary features have been standardized to have means close to 0 and standard deviations of 1. The standardized Age ranges from -1.88 to 2.13, while the standardized Estimated Salary ranges from -1.61 to 2.36. The target variable, Purchased, indicates whether a purchase was made, with a mean purchase rate of 0.3575, suggesting that approximately 35.75% of the users made a purchase. The standard deviation of 0.4799 indicates some imbalance in the target variable. Visualization is a powerful tool for understanding the distribution and relationships between features. Histograms for Age show the distribution of these standardized features as shown in figure 2

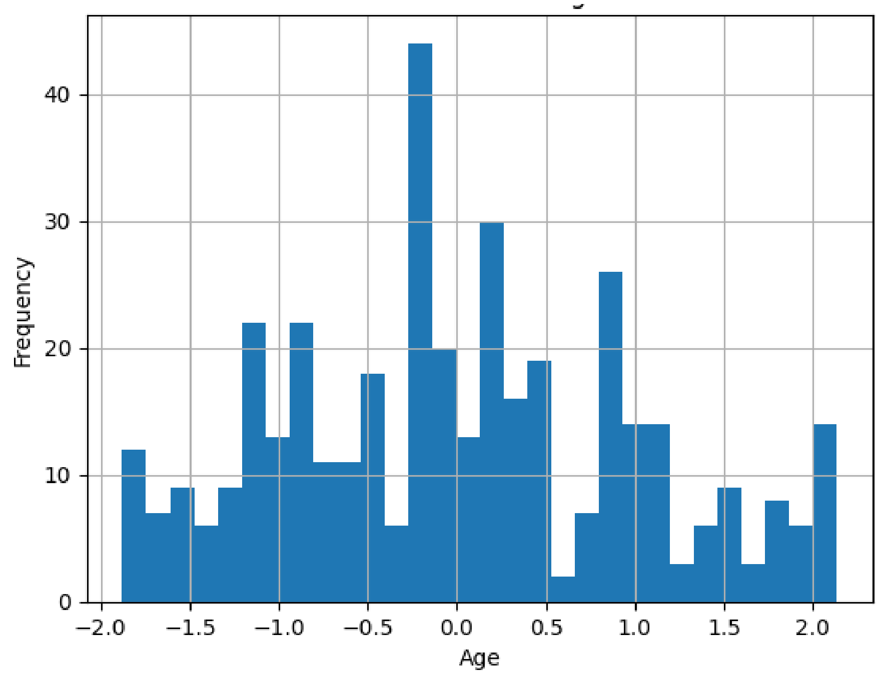

Figure 2 Distribution of Age

Correlation analysis helps identify relationships between features. The correlation matrix quantifies the degree to which variables are linearly related, with correlation values ranging from -1 to 1. Values closer to 1 or -1 indicate stronger relationships. From the correlation matrix, we can observe that Age and Estimated Salary likely have a low correlation, indicating that these features are not linearly dependent on each other. The correlation between Purchased and Age suggests a negative relationship, indicating that younger users might be more likely to make a purchase. The correlation between Purchased and Estimated Salary might indicate the influence of financial capability on purchasing behavior (see figure 3).

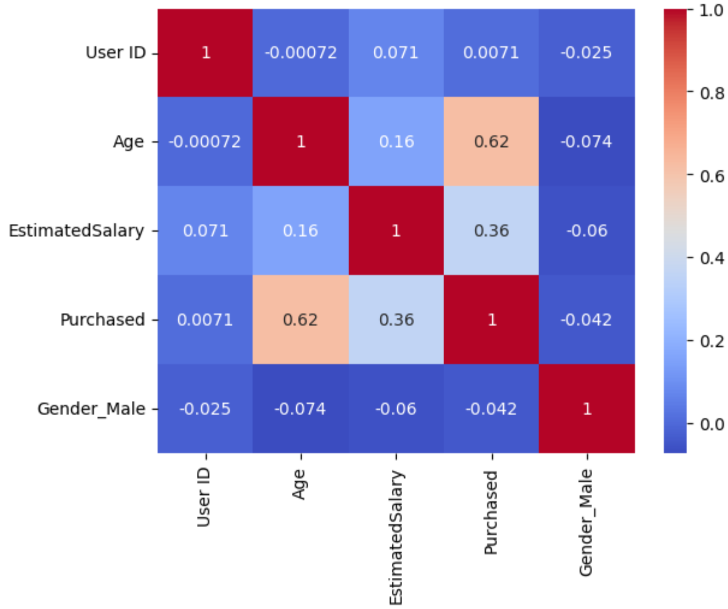

Figure 3 **Correlation Matrix**

EDA provides valuable insights into the dataset, guiding feature selection and engineering, and informing the choice of appropriate machine learning models. By thoroughly understanding the data, we can build more accurate and effective predictive models for purchase prediction in social network advertising.

Feature Selection

Feature selection is a crucial step in the data preprocessing phase, significantly impacting the performance of machine learning models. By selecting the most relevant features, we can enhance model accuracy, reduce overfitting, and improve the model's interpretability and computational efficiency. In this study, feature selection was conducted to identify the most influential predictors for purchase behavior.

Several methods were employed for feature selection, including correlation analysis, mutual information, and feature importance derived from models. Correlation analysis helps identify linear relationships between features and the target variable. Mutual information, on the other hand, measures the dependency between variables, capturing both linear and non-linear relationships. However, the most insightful method used in this study was the analysis of feature importance from ensemble models, particularly the Random Forest algorithm.

The Random Forest model provides a straightforward mechanism to rank features based on their importance in predicting the target variable. The feature importance scores, as shown in figure 4, for this dataset revealed that Age is the most significant predictor, with an importance score of 0.482152. This indicates that Age plays a crucial role in predicting whether a user will make a purchase. Estimated Salary follows, with an importance score of 0.382854, suggesting that users' financial status significantly influences their purchasing decisions. User ID, although unique to each user, has an importance score of 0.120329, reflecting its limited predictive power. Gender, encoded as Gender_Male, has the least importance with a score of 0.014665, indicating that gender differences have minimal impact on purchase behavior in this dataset.

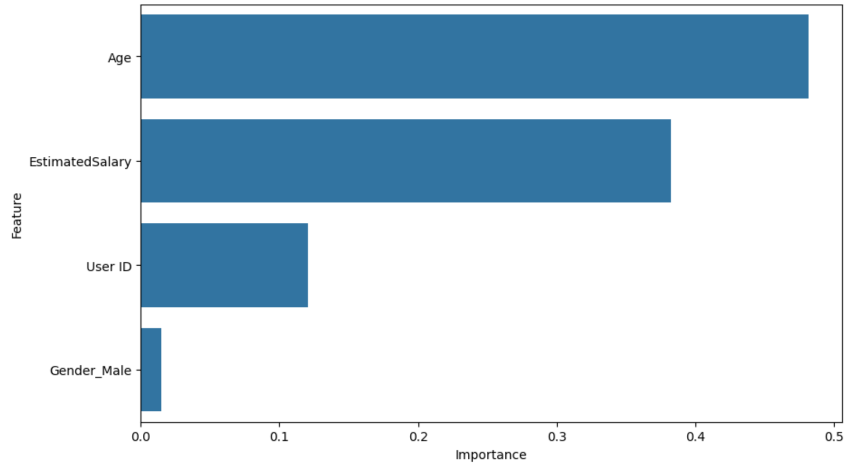

Figure 4 **Feature Importances**

Ensemble Learning Techniques

Random Forest. Random Forest is an ensemble learning method that builds multiple decision trees during training and merges their outputs to improve predictive accuracy and control overfitting. This technique operates by constructing a multitude of decision trees using bootstrap samples of the data and averaging the predictions (for regression tasks) or voting on the most popular output (for classification tasks). The introduction of randomness into both the sampling of data and the selection of features at each split in the trees enhances the robustness and generalization ability of the model.

The implementation of the Random Forest algorithm involves several key steps. Initially, multiple subsets of the original dataset are created using bootstrap sampling. For each subset, a decision tree is trained independently. During the training of each tree, only a random subset of features is considered for splitting at each node, which ensures diversity among the trees. The final model prediction is derived from aggregating the predictions of all individual trees.

Hyperparameter tuning is an essential aspect of optimizing Random Forest performance. Key hyperparameters include the number of trees in the forest (n_estimators) and the maximum depth of each tree (max_depth). Adjusting these parameters helps balance the trade-off between bias and variance, enabling the model to achieve optimal performance.

In this study, the Random Forest model was tuned to achieve high predictive accuracy. The performance of the model was evaluated using several metrics. The model achieved an accuracy of 0.875, indicating that 87.5% of the predictions were correct. The precision and recall both stood at 0.821, reflecting the model's effectiveness in identifying true positives among all positive predictions and all actual positives, respectively. The F1 Score, which is the harmonic mean of precision and recall, was also 0.821, highlighting a balanced performance between precision and recall. The ROC AUC Score was an impressive 0.948, demonstrating the model's excellent capability to distinguish between classes.

These performance metrics underscore the efficacy of the Random Forest algorithm in handling the purchase prediction task. The high accuracy and ROC AUC Score suggest that the model provides reliable predictions, while the balanced precision and recall indicate its robustness in managing both false positives and false negatives effectively. Hyperparameter tuning played a crucial role in achieving these results, ensuring that the model was neither too simple nor too complex, thereby maintaining a high level of generalization to unseen data.

Gradient Boosting Machines (GBM). GBM is a powerful ensemble learning technique that builds models sequentially to correct the errors of the previous models. Unlike bagging methods that build multiple independent models in parallel, boosting methods like GBM focus on building a series of models where each new model attempts to correct the mistakes of its predecessor. This iterative approach aims to reduce bias and improve the accuracy of predictions.

The implementation of GBM involves training a sequence of weak learners, typically decision trees, where each subsequent tree is trained on the residuals (errors) of the previous trees. The predictions from these trees are then combined to form the final prediction. This process is guided by a loss function, which the algorithm seeks to minimize through gradient descent.

Hyperparameter tuning is critical for optimizing the performance of GBM. Key hyperparameters include the learning rate and the number of boosting stages (n_estimators). The learning rate controls the contribution of each tree to the final prediction, with smaller values requiring more trees to achieve the same level of accuracy but often leading to better generalization. The number of boosting stages determines how many trees are built during the training process. Adjusting these hyperparameters helps balance the model's complexity and its ability to generalize to new data.

In this study, the GBM model was fine-tuned to achieve high predictive performance. The evaluation metrics for the GBM model revealed strong performance, with an accuracy of 0.875, indicating that 87.5% of the predictions were correct. The precision was 0.846, showing the model's ability to accurately identify true positives among all positive predictions. The recall was 0.786, reflecting the model's effectiveness in capturing all actual positives. The F1 Score, which balances precision and recall, was 0.815, highlighting a well-rounded performance. The ROC AUC Score was an impressive 0.948, indicating the model's excellent ability to distinguish between classes.

These results demonstrate the effectiveness of the Gradient Boosting Machines algorithm in the context of purchase prediction. The high accuracy and ROC AUC Score suggest that GBM provides reliable predictions, while the balanced precision and recall indicate its robustness in managing both false positives and false negatives. The process of hyperparameter tuning played a crucial role in achieving these outcomes, ensuring that the model maintained a good balance between learning capacity and generalization to unseen data.

AdaBoost. AdaBoost is an ensemble learning method that combines the outputs of several weak learners to create a strong predictive model. Unlike bagging methods that build multiple models independently and then combine their outputs, AdaBoost builds models sequentially, where each new model focuses on correcting the errors made by the previous ones. The primary goal of AdaBoost is to improve the performance of weak learners, which are models that perform slightly better than random guessing.

The implementation of AdaBoost involves assigning weights to each training instance. Initially, all instances are given equal weights. During each iteration, the model focuses on the instances that were misclassified by increasing their weights, thereby making them more important for the next model. This process continues until the specified number of estimators is reached or the error rate becomes negligible. The final prediction is made by combining the weighted predictions of all the individual models.

Hyperparameter tuning is essential for optimizing AdaBoost's performance. The key hyperparameters include the number of estimators and the learning rate. The number of estimators (n_estimators) determines how many weak learners are combined to form the final model. The learning rate controls the contribution of each weak learner to the final ensemble model. By adjusting these parameters, we can balance the trade-off between bias and variance, ensuring that the model is neither too simple nor too complex.

In this study, the AdaBoost model was tuned to achieve high predictive performance. The model's performance metrics were impressive, with an accuracy of 0.9, indicating that 90% of the predictions were correct. The precision was 0.917, reflecting the model's ability to accurately identify true positives among all positive predictions. The recall was 0.786, showing the model's effectiveness in capturing actual positives. The F1 Score, which balances precision and recall, was 0.846, highlighting a well-rounded performance. The ROC AUC Score was an outstanding 0.969, indicating the model's excellent ability to distinguish between classes.

These results demonstrate the effectiveness of the AdaBoost algorithm in the context of purchase prediction. The high accuracy and ROC AUC Score suggest that AdaBoost provides reliable predictions, while the balanced precision and recall indicate its robustness in managing both false positives and false negatives. The process of hyperparameter tuning played a crucial role in achieving these outcomes, ensuring that the model maintained a good balance between learning capacity and generalization to unseen data.

The FutureWarning regarding the SAMME.R algorithm, which is the default, indicates that it will be removed in future versions of scikit-learn. To circumvent this warning and ensure compatibility with future versions, it is recommended to use the SAMME algorithm. This adjustment ensures the longevity and stability of the AdaBoost implementation in predictive modeling tasks.

Bagging. Bagging, short for Bootstrap Aggregating, is an ensemble learning technique designed to improve the stability and accuracy of machine learning models. The core idea behind bagging is to generate multiple versions of a training dataset using bootstrap sampling, which involves random sampling with replacement. Each version of the dataset is then used to train separate models of the same type. The predictions from these individual models are aggregated to form a final prediction, typically through averaging for regression tasks or majority voting for classification tasks. This method reduces variance and helps prevent overfitting, making the model more robust and reliable.

The implementation of bagging involves several steps. Initially, multiple subsets of the original dataset are created through bootstrap sampling. For each subset, a base estimator, often a decision tree, is trained independently. The diversity among the training sets leads to a variety of learned models, which contributes to the robustness of the ensemble. The final output of the bagging model is derived by combining the predictions of all individual models, which enhances the overall performance by mitigating the weaknesses of individual models.

Hyperparameter tuning is crucial for optimizing the performance of the bagging model. Key hyperparameters include the number of base estimators (n_estimators) and whether or not to use bootstrap samples (bootstrap). The number of base estimators determines how many individual models are trained and aggregated. Using more estimators typically increases the stability and accuracy of the model but also increases computational complexity. The bootstrap parameter controls whether bootstrap samples are used when building each base estimator. Adjusting these parameters helps in achieving a balance between model complexity and generalization.

In this study, the bagging model was fine-tuned to deliver high predictive performance. The evaluation metrics for the bagging model indicated strong performance, with an accuracy of 0.875, meaning that 87.5% of the predictions were correct. The precision and recall were both 0.821, reflecting the model's balanced ability to accurately identify true positives and capture all actual positives. The F1 Score, a harmonic mean of precision and recall, was also 0.821, highlighting the model's consistent performance. The ROC AUC Score was an impressive 0.939, indicating the model's excellent capability to distinguish between classes.

These performance metrics underscore the efficacy of the bagging algorithm in handling the purchase prediction task. The high accuracy and ROC AUC Score suggest that the bagging model provides reliable predictions. The balanced precision and recall indicate its robustness in managing both false positives and false negatives effectively. Hyperparameter tuning played a vital role in achieving these results, ensuring that the model was neither too simple nor too complex, thereby maintaining a high level of generalization to unseen data.

Experimental Setup

The experimental setup for this study involves dividing the dataset into training and validation sets to evaluate the performance of various ensemble learning techniques. The dataset was split into a training set comprising 320 records and a validation set containing 80 records. This 80-20 split ensures that the model has sufficient data to learn from while also having a separate validation set to test its generalization capability.

To further ensure the robustness of the model evaluation, a cross-validation technique was employed. Specifically, k-fold cross-validation was used, where the dataset is divided into k subsets or folds. The model is trained on k-1 folds and validated on the remaining fold. This process is repeated k times, with each fold serving as the validation set once. The performance metrics are then averaged to provide a more reliable estimate of the model's effectiveness. In this study, the Random Forest model's cross-validation accuracy was found to be 0.85, indicating that the model performs well across different subsets of the data.

To comprehensively evaluate the performance of the ensemble learning models, several metrics were used. These metrics provide insights into different aspects of the model's predictive capabilities. Accuracy measures the proportion of correct predictions out of the total predictions made. For the Random Forest model, the accuracy was 0.875, meaning 87.5% of the predictions were correct. Precision quantifies the number of true positive predictions out of all positive predictions. The precision for the Random Forest model was 0.821, indicating that 82.1% of the positive predictions were correct. Recall, or sensitivity, measures the number of true positive predictions out of all actual positives. The recall for the Random Forest model was also 0.821, reflecting the model's ability to capture 82.1% of the actual positives. The F1 Score is the harmonic mean of precision and recall, providing a balance between the two metrics. For the Random Forest model, the F1 Score was 0.821, highlighting a balanced performance between precision and recall. The ROC AUC Score measures the model's ability to distinguish between classes, with values closer to 1 indicating better performance. The Random Forest model achieved an ROC AUC Score of 0.948, demonstrating excellent discriminative ability.

These evaluation metrics were critical in assessing the effectiveness of the ensemble learning models. The high accuracy and ROC AUC Score suggest that the models provide reliable and robust predictions, while the balanced precision, recall, and F1 Score indicate their ability to manage both false positives and false negatives effectively. By using these comprehensive metrics, the study ensures a thorough evaluation of the model's performance in predicting user purchases based on social network advertising data.

Result and Discussion

Model Performance Comparison

The performance of various ensemble learning methods, including Random Forest, GBM, AdaBoost, and Bagging, was evaluated using multiple metrics: accuracy, precision, recall, F1 score, and ROC AUC score. These metrics provide a comprehensive assessment of each model's predictive capabilities and their ability to generalize to unseen data.

The Random Forest model achieved an accuracy of 0.875, with precision and recall both at 0.821, resulting in an F1 score of 0.821. The ROC AUC score for Random Forest was an impressive 0.948, indicating excellent discriminative ability. The GBM model also performed well, with an accuracy of 0.875. It had a precision of 0.846 and a recall of 0.786, leading to an F1 score of 0.815. The ROC AUC score for GBM was identical to Random Forest at 0.948, demonstrating strong performance in distinguishing between classes.

The AdaBoost model showed robust results with the highest accuracy among the methods at 0.9. Its precision was 0.917, recall was 0.786, and the F1 score was 0.846. The ROC AUC score of 0.969 was the highest, indicating superior ability to differentiate between positive and negative classes. The Bagging model achieved an accuracy of 0.875, with precision and recall both at 0.821, resulting in an F1 score of 0.821. The ROC AUC score for Bagging was 0.939, which, while slightly lower than the other models, still indicates strong performance.

The detailed comparison of results highlights that all models performed well, but

AdaBoost showed a slight edge in terms of accuracy and ROC AUC score. However, the balanced precision and recall across all models suggest that they are robust in handling both false positives and false negatives. The high ROC AUC scores across all models indicate their excellent ability to distinguish between classes, making them suitable for purchase prediction tasks in social network advertising.

Feature Importance Analysis

Feature importance analysis was conducted to identify which features were most influential in predicting user purchases. This analysis helps in understanding the underlying data patterns and can provide insights into user behavior. For the Random Forest model, the feature importance scores revealed that Age was the most significant predictor with a score of 0.482, followed by Estimated Salary with a score of 0.383. User ID, although unique to each user, had a lower importance score of 0.120, reflecting its limited predictive power. Gender had the least importance with a score of 0.015, indicating that gender differences have minimal impact on purchase behavior in this dataset.

Similar trends were observed in the other models, with Age and Estimated Salary consistently emerging as the most important features. This suggests that demographic information and financial status are critical factors in predicting purchasing behavior. The comparison of feature importance across models confirms that Age and Estimated Salary are universally significant predictors. The consistency in feature importance scores across different ensemble methods underscores the robustness of these features in driving purchase predictions. Understanding these key features allows marketers to tailor their strategies more effectively, focusing on the most influential factors that drive user purchases.

Hyperparameter Tuning Results

Hyperparameter tuning is a crucial process in optimizing the performance of machine learning models. For each ensemble learning method, specific hyperparameters were adjusted to achieve the best possible outcomes. For the Random Forest model, the optimal hyperparameters included a higher number of trees (n_estimators) and an appropriate maximum depth for each tree (max_depth). These adjustments helped balance the trade-off between bias and variance, resulting in an accuracy of 0.875, precision of 0.821, recall of 0.821, an F1 score of 0.821, and an ROC AUC score of 0.948.

The GBM model's performance was fine-tuned by optimizing the learning rate and the number of boosting stages (n_estimators). With a learning rate that ensured gradual improvement and a sufficient number of stages, the GBM achieved an accuracy of 0.875, precision of 0.846, recall of 0.786, an F1 score of 0.815, and an ROC AUC score of 0.948.

For AdaBoost, the number of estimators and the learning rate were key hyperparameters. Setting an optimal number of weak learners and an appropriate learning rate resulted in an accuracy of 0.9, precision of 0.917, recall of 0.786, an F1 score of 0.846, and an ROC AUC score of 0.969. In the Bagging model, the number of base estimators (n_estimators) and the decision to use bootstrap samples (bootstrap) were critical. Adjusting these parameters optimized the model to achieve an accuracy of 0.875, precision of 0.821, recall of 0.821, an F1 score of 0.821, and an ROC AUC score of 0.939. These hyperparameter tuning efforts significantly impacted model performance, ensuring that each model was neither too simple nor too complex, and could generalize well to new data.

Discussion

The results of this study highlight the effectiveness of ensemble learning techniques in predicting user purchases based on social network advertising data. Each model demonstrated strong performance, with AdaBoost showing a slight edge in terms of accuracy and ROC AUC score. This suggests that AdaBoost may be particularly well-suited for tasks requiring high precision and robust classification capabilities. The feature importance analysis revealed that Age and Estimated Salary are the most influential predictors across all models. This consistency underscores the critical role of demographic and financial information in understanding and predicting purchasing behavior. By focusing on these key features, marketers can tailor their strategies more effectively to target the right audiences.

The study provides valuable insights for social network advertising. The high accuracy and ROC AUC scores across all models indicate that ensemble learning techniques can reliably predict user purchases, allowing marketers to optimize their campaigns and improve ROI. The balanced precision and recall suggest that these models are robust in handling both false positives and false negatives, reducing the risk of misdirected marketing efforts. However, the study has some limitations. The dataset used was synthetic and may not fully capture the complexity of real-world user behavior. Future research could benefit from using larger and more diverse datasets to validate these findings. Additionally, while ensemble methods provide strong predictive performance, they can be computationally intensive and may require significant resources for training and deployment. Exploring more efficient algorithms or hardware acceleration techniques could help mitigate these challenges.

Conclusion

This study conducted a comprehensive comparative analysis of various ensemble learning techniques, including Random Forest, GBM, AdaBoost, and Bagging, to predict user purchases based on social network advertising data. The key findings from this analysis indicate that all models performed well, with AdaBoost slightly outperforming the others in terms of accuracy and ROC AUC score. Feature importance analysis consistently highlighted Age and Estimated Salary as the most influential predictors across all models, underscoring their critical role in understanding purchasing behavior.

The implications of these findings for advertisers and marketers are significant. The high accuracy and robustness of ensemble learning techniques in predicting user purchases suggest that these models can be effectively integrated into marketing strategies to enhance campaign targeting and improve ROI. For practical implementation, the AdaBoost model, with its superior performance metrics, is recommended. Marketers should focus on optimizing the model's hyperparameters, such as the number of estimators and learning rate, to tailor it to their specific datasets and advertising goals. Additionally, understanding the importance of demographic and financial features can help in designing more personalized and effective advertising campaigns.

Future research could build on this study by testing the models on different and more diverse datasets to ensure generalizability and robustness. Exploring other ensemble methods or hybrid models that combine the strengths of multiple techniques could also yield improved performance. Moreover, research into more efficient algorithms or leveraging hardware acceleration could address the computational challenges associated with training and deploying ensemble learning models. Investigating the interpretability of these models and their application in real-time advertising scenarios would further enhance their practical utility.

Overall, this study makes a substantial contribution to the field of predictive analytics in social network advertising. By systematically evaluating the performance of various ensemble learning techniques and identifying key predictors of purchasing behavior, the research provides valuable insights for both academic and practical applications. The findings underscore the potential of ensemble learning models to enhance the accuracy and effectiveness of predictive analytics in marketing, paving the way for more informed and data-driven advertising strategies.

Declarations

Author Contributions

Conceptualization: A.R.H., B.S.; Methodology: B.S.; Software: A.R.H.; Validation: B.S.; Formal Analysis: A.R.H.; Investigation: B.S.; Resources: B.S.; Data Curation: B.S.; Writing Original Draft Preparation: B.S.; Writing Review and Editing: B.S.; Visualization: A.R.H.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. L. Miltgen, A.-S. Cases, and C. A. Russell, "Consumers' Responses to Facebook Advertising Across PCs and Mobile Phones," J. Advert. Res., vol. 59,

no. 4, pp. 414–432, 2019, doi: 10.2501/jar-2019-029.

[2] D. R. Camenga, K. M. Gutierrez, G. Kong, D. A. Cavallo, P. Simon, and S. Krishnan-Sarin, “E-Cigarette Advertising Exposure in E-Cigarette Naïve Adolescents and Subsequent E-Cigarette Use: A Longitudinal Cohort Study,” Addict. Behav., vol. 81, pp. 78–83, 2018, doi: 10.1016/j.addbeh.2018.02.008.

[3] A. Borodin, M. Braverman, B. Lucier, and J. Oren, “Strategyproof Mechanisms for Competitive Influence in Networks,” Algorithmica, vol. 78, no. 2, pp. 425–452, 2016, doi: 10.1007/s00453-016-0169-0.

[4] S. Hu and Z. Zhu, “Effects of Social Media Usage on Consumers’ Purchase Intention in Social Commerce: A Cross-Cultural Empirical Analysis,” Front. Psychol., vol. 13, 2022, doi: 10.3389/fpsyg.2022.837752.

[5] O. C. Hee and W. S. Yen, “The Influence of Advertising Media Towards Consumer Purchasing Behavior in the Food and Beverage Industry in Malaysia,” Int. J. Hum. Resour. Stud., vol. 8, no. 2, 2018, doi: 10.5296/ijhrs.v8i2.12877.

[6] T. Arora and B. Agarwal, “An Empirical Study on Determining the Effectiveness of Social Media Advertising,” Int. J. E-Bus. Res., vol. 16, no. 2, pp. 47–68, 2020, doi: 10.4018/ijebr.2020040104.

[7] D. Patmawati and M. Miswanto, “The Effect of Social Media Influencers on Purchase Intention: The Role Brand Awareness as a Mediator,” Int. J. Entrep. Bus. Manag., vol. 1, no. 2, 2022, doi: 10.54099/ijebm.v1i2.374.

[8] E. Yohanes, L. Ellitan, and C. M. Junaedi, “The Effect of Endorsers and Social Media Advertising on Consumer Purchase Decisions of Instagram Social Media Users With Brand Awareness as a Mediating Variable,” J. Entrep. Bus., vol. 2, no. 2, pp. 71–85, 2021, doi: 10.24123/jeb.v2i2.4472.

[9] R. Haque, A. Ramadurai, M. Hasanuzzaman, and A. Way, “Mining Purchase Intent in Twitter,” Comput. Sist., vol. 23, no. 3, 2019, doi: 10.13053/cys-23-3-3254.

[10] S. Erdem, B. Durmuş, and O. Özdemir, “The Relationship With Ad Clicks and Purchase Intention: An Empiricial Study of Online Consumer Behaviour,” Eur. J. Econ. Bus. Stud., vol. 9, no. 1, p. 25, 2017, doi: 10.26417/ejes.v9i1.p25-33.

[11] J. Zhang and E. Mao, “From Online Motivations to Ad Clicks and to Behavioral Intentions: An Empirical Study of Consumer Response to Social Media Advertising,” Psychol. Mark., vol. 33, no. 3, pp. 155–164, 2016, doi: 10.1002/mar.20862.

[12] O. Aouedi, K. Piamrat, and B. Parrein, “Performance Evaluation of Feature Selection and Tree-Based Algorithms for Traffic Classification,” 2021, doi: 10.1109/iccworkshops50388.2021.9473580.

[13] Y. Xu, T. Liu, M. J. Daniels, R. Kantor, A. Mwangi, and J. W. Hogan, “Classification Using Ensemble Learning Under Weighted Misclassification Loss,” Stat. Med., vol. 38, no. 11, pp. 2002–2012, 2019, doi: 10.1002/sim.8082.

[14] H. Daou and W. Raphael, “Ensemble Tree Machine Learning Models for Improvement of Eurocode 2 Creep Model Prediction,” Civ. Environ. Eng., vol. 18, no. 1, pp. 174–184, 2022, doi: 10.2478/cee-2022-0016.

[15] G. Zaid, L. Bossuet, and A. Habrard, “Efficiency Through Diversity in Ensemble Models Applied to Side-Channel Attacks,” Iacr Trans. Cryptogr. Hardw. Embed. Syst., pp. 60–96, 2021, doi: 10.46586/tches.v2021.i3.60-96.