Novi Ekonomist
Vol 19(1), Year XIX, Issue 37, january – june 2025.
ISSN 1840-2313 (Print) 2566-333X (Online)
DOI: 10.69781/NOEK202537165

Submitted: 27.04.2025.
Accepted: 16.06.2025.
REVIEW ARTICLE
UDK: 316.77:004.738.5(4-672EU)

ENHANCING SOCIAL MEDIA AD CAMPAIGNS THROUGH ENSEMBLE-BASED OPTIMIZATION

Safiye Turgay
Sakarya University, Adapazarı, Turkey
sencer@sakarya.edu.tr
ORCID: 0000-0002-8706-0801

Mustafa Kavacık
Necmettin Erbakan University, Konya, Turkey
mustafa.kavacik@erbakan.edu.tr
ORCID: 0000-0003-3782-8154

Yunus Emre Torkul
Sakarya University, Adapazarı, Turkey
yunustorkul@gmail.com
ORCID: 0000-0001-9543-9363

Muhammed Sercan Şahin
Sakarya University, Adapazarı, Turkey
sercansahin2000@gmail.com
ORCID: 0009-0001-1037-9005

Rabia Ayşe Güzel
Sakarya University, Adapazarı, Turkey
r.ayse.guzel@gmail.com
ORCID: 0000-0002-0835-567X

Abstract: *The focus of this paper is to analyze the use of ensemble-based optimization for improvement of social media ad campaigns. The procedure of using more than one machine learning model yield very high possibilities in predictive accuracy, accuracy in targeting and most importantly in the issue of budgeting. Random Forest, Gradient Boosting & Stacking allow the advertisers to estimate the users' behavior better – CTR, & fight Ad Fraud in a more effective manner. The paper supports the finding by providing three real-life use cases which illustrate The study provides examples for real-time bidding, audience segmentation, as well as sentiment analysis where the proposed approach outcompeted the baseline in terms of performance. It turns out that the use of ensemble learning in social media marketing can be helpful in more effective ad spend distribution, increasing the rates of engagement, and, as a result, the probability of successful campaigns will be higher.*

Key words: *Ensemble Methods, Machine Learning, Social Media Advertising, Optimization, Random Forest, Gradient Boosting, Stacking, Click-Through Rate (CTR), Ad Targeting, Budget Allocation, Ad Fraud Detection, Audience Segmentation, Predictive Modeling*

JEL classification*: D83, M31, M37, C38*

1. INTRODUCTION

In the contemporary society dominated by social technologies, social media sites are the most strategic advertising spaces where the advertiser targets to capture the customer attention. Since billions of users are active daily and produce massive amounts of data, the difficulty is identifying specific audience segments and to campaign management fail to fit at the higher level and unable to handle the huge amount of data, thus resulting in ineffectual ad positioning and wasteful utilization of the budget.

Therefore, these problems can be effectively addressed by ensemble methods in machine learning where multiple models are combined to improve the accuracy and volatility of the predictions. In contrast to selecting only one type of model, Random Forest, Gradient Boosting and Stacking models use several models hence providing higher accuracy at different metrics. This kind of method can be especially helpful in solving some critical problems of social media advertising such as audience targeting, ctr-prediction, budget allocation, and fraud detection.

This paper focuses on the application of Ensemble based optimization for changing the social media ad campaigns. Thus, through machine learning methods, the targeting becomes much more accurate, advertising costs are minimized, and the effectiveness of the entire campaign is increased. In this paper, we also show by case studies and empirical analysis how advertisers can leverage on ensemble methods to provide better engagement rates and ROI for their advertising campaigns. How these techniques may apply to social media marketing highlights a critical fact: modern campaigns must generally use data-driven approaches supported by a variety of modelling techniques.

2. LITERATURE SURVEY

It is a rapidly growing field of interest in the last few years more specifically with contextual to social media advertising. Due to the increasing amount of data that is available in the market, researchers and practitioners have attempted to find new ways of improving the outcomes of advertisement initiatives. This literature review looks into relevant works of scholars who have analyzed ensemble methods as applied to enhancing social media advertising campaigns with a focus on the existing and emerging methods in this field and their effects on the marketing results (Tran, Tran & Van Nguyen, 2024; Seurin & Shirvan, 2024; Zeng, Chipusu, Zhu, Li, Ibrahim & Huang, 2024; Yin, Ahmadianfar, Karim & Elmannai, 2024; Kayali & Turgay, 2023; Li, Yao, Chen, Li, Lu, Liu & Yu, 2024; Zou, Zhen, Wang, Wu, Li, Yuan & Xu, 2024; Mustaffa & Sulaiman, 2023).

The approach used in ensemble methods is the incorporation of multiple projections, from which numerous improved and superior models used have been developed across different fields. Random forests as a method of machine learning is yet another extension of the concept of the bagging method which combines the results of a number of decision trees so as to minimise the amount of variance and hence provides a more generic solution (Breiman, 2001; Turgay & Erdoğan, 2023; Nalini, Yamini, Fernandez & Priyadarsini, 2024; Tian, Ma, Geng, Yang, Luo, Gao & Liang, 2024). Similarly, some of the studies came up with Gradient Boosting Machines (GBMs), which build several weak models and improve them in a step by step process while focusing on the tough cases. Many of these early studies, therefore, set the stage for the use of ensemble methods across marketing and advertising campaigns (Yang, Li, Wang & Yang, 2023; Sun, An, Yang & Liu, 2024; Lin & Meng, 2024; Turgay, 2023; Liu, Cheng, & Du, 2024; Yetiş, Turgay & Erdemir, 2024).

A number of investigations have previously concentrated on using machine learning approaches for the purpose of modeling user interactions in social media scenarios. Some of the researchers investigated the efficiency of a set of machine learning algorithms, grouped in the category of ensemble methods, for the CTR prediction in the context of online advertising. The conclusion that was made by them proved that ensemble methods especially those which involve the use of more than base learner method is better than the single method when it comes to the prediction of user engagement (Mi, Dai, Jing, She, Holmedal, Tang & Pan, 2024; Adekoya & Aneiba, 2024; Liu, Li, Zou, Hou, Yang & Zheng, 2024; Wu, Xu, Yang, Qiu, Volinsky & Pang, 2023; Houssein, Abdalkarim, Samee, Alabdulhafith & Mohamed, 2024; Gul, Hussain, Khan, Arshad, Khan & de Jesus Motheo, 2024; Feng, Zhou, Luo & Wei, 2024; Pan, Zhang, Chu, Zhang & Wu, 2024; Turgay, Han, Erdoğan, Kara & Yilmaz, 2024).

In doing so, this research pointed out that ensemble methods have the possibility of improving the targeting of advertisements so as to increase the effectiveness of the campaigns (Yiğit, Turgay, Cebeci & Kara, 2024; Zhao, Fu, Zhang, Zhu, Lu & Francis, 2024; Zhou, Rao & Gao, 2023; Alzoubi, Abualigah, Sharaf, Daoud & Khodadadi, 2024). Audience that is reached greatly determines the success of the social media ad campaigns. Some of the studies have conducted the analysis of audience segmentation in the social media using Random Forests and Gradient Boosting (Ping, Yang, Zhang, Xing, Yang, Yan & Wang, 2023; Alparslan, Turgay & Yilmaz, 2024; Akkaya & Turgay, 2024; Ramirez, Lam, Gutierrez, Hou, Tribukait, Roch & Laveille, 2024). They pointed out that this use of ensemble methods resulted in remarkable enhancement of the algorithms' accuracy for identification of high-value users, hence better ad placements. The research made some recommendations on how to use many models in place to capture all the patterns in users' data to increase the level of targeting accuracy.

Another strategic issue arises in the allocation of the resources that is common to most advertising where inefficiency can culminate in wastage. To optimize the ad spends across the different social media platforms used general ensmble learning. Due to use of multiple models, they were able to assign budgets in a way that would maximize the return on investment. This work clearly illustrated the effectiveness of ensemble methods in the real life issues that arose while addressing multichannel advertising campaigns (Gu, Wang & Liu, 2024; Wang, Incecik, Tian, Zhang, Kujala, Gupta & Li, 2024; Chen, Zeng, Jia, Jabli, Abdullah, Elattar & Assilzadeh, 2024; Liu, Cen, Zheng, Li & Wang, 2024; Ayvaz, Aydoğan, Akçura & Şensoy, 2021; Jain & Saha, 2024).

Click fraud and impression fraud have become a major concern in digital advertising due to the effect on the credibility of the campaigns. Proposed sample approach innovated with Isolation Forests, an ensemble method for detecting anomalies, in tackling frauds in online advertising. They emphasised that the use of the ensemble means of anomaly detection can help to decrease the amount of fraud and prevent campaign metrics from manipulation, which would improve assessment of the campaign performance.

The other area where ensemble methods has been reported to be successful is in the sentiment classification of the user generated contents. Some researchers used another technique known as stacking that involves using different classifiers for increased accuracy of sentiment analysis in social media platforms. Their study also revealed that through ensemble methods, the performance of content optimization techniques could be enhanced and yielded enhanced positive user response as well as better advertisements.

The literature equally points to the increase in the application of ensemble methods for enhancing each factor of social media advertising campaigns. Besides increasing predictive accuracy in targeting and engagement or increasing the budget's value, the use of ensemble-based approaches is much more beneficial than traditional single-model techniques to detect fraud. This survey shows the prospects for the application of ensemble learning in social media marketing with a solid basis for the subsequent analysis and use in this paper.

3. METHODOLOGY

The whole process of improving the social media advertisement through ensemble optimaization is explained below which are data gathering, factor extraction, algorithm choice, technique combination, and algorithm assessment.

All these steps are meant to facilitate the use of ensemble methods in the right manner in order to achieve the desired goals of a campaign such as CTR, conversion rates and ROI (in Figure 1).

Figure 1. Suggested model steps

Step 1
Data Preparation
- Collect Data
- Preprocess Data
- Feature Engineering

⇩

Step 2
Model Training
- Train Base Models
    - Decision Trees
    - Logistic Regression
    - Support Vector Machines (SVM)
    - Neural Networks
- Tune Hyperparameters (Optimization)

⇩

Step 3
Apply Ensemble Techniques
- Bagging
- Boosting
- Stacking
- Voting

⇩

Step 4
Performance Evaluation
- Evaluate Models
    - Mean Absolute Error (MAE)
    - Root Mean Square Error (RMSE)
    - Accuracy, Prediction, Recall, and F1-Score
- Compare Results
- Conduct A/B Testing (Optional)

⇩

Step 5
Deployment and Monitoring
- Deploy Model
- Monitor Performance

The Step 1 involves the collection of data, pre-processing of the data as well as feature extraction. Collect data includes the collection of data gathered from the social media platforms as well as the user demographics and engagement rates, the advertisement performance rates, and context information. Preprocess Data focuses on remove the noise from the data (for example, dealing with missing values, removing duplicate); scale features, if required; code categorical values. Feature Engineering involve selection of the relevant features and developing other features that might enhance the performance of the model ; division of data into the training data and a data set to test.

Step 2 consist of Model Training with sections like train base models and Tune Hyperparameters. Train Base Models comprises of the decision trees, logistic regression, Support Vector Machines (SVM), Neural Networks. Optimize Hyperparameters attains the idea of fine-tuning of the hyperparameters for each of the base models which may be done using methods like Grid search or Random search.

Step 3. They are bagging, boosting, stacking and voting are some of the methods that fall under the Apply Ensemble Techniques. Bagging involve the follow steps: generate multiple bootstrap samples of the training data; train a single base model on different bootstrap sample; combine the predicts using averaging or voting depending on the typeof problem (regression or classification). Increasing include the train base models in a sequential manner where each model is trained to correct the mistakes of the preceding one; model prediction can hereby be done through weighted sums. Stacking covers multiple base models on the training set and uses their predictions for the meta-model; train it on the base models' predictions to make the final output of prediction.
Classification problems are voting that involves the use of majority voting on multiple models' predictions.

Step 4. Performance Evaluation includes the following sections: the evaluate models, the compare results and the A/B testing. Evaluate models include the calculate performance metrics for each model and the ensemble model are Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Accuracy, Precision, Recall, F1- score. Meaning, the final step is to make a comparison between the results of the ensemble model and the separate base models in order to analyse the differences in the accuracy of the predictions and overall success of the campaign. Perform an A/B test (Optional includes apply the ensemble-based model in the live campaigns along with A/B, B, C models for the definition of the real-world improvement of the results.
The last step is the deployment and monitoring which is covered in the fifth step. Deploy model explains the process of incorporating the ensemble model into real-time ad campaign management system for predictions and optimizations. Closely, Continuously monitor model performance and update it as new data and dynamic creases up in campaign.
This methodology defines a structured approach on how to best utilise ensemble methods for enhancing social media advertising campaigns. Specifically, this approach is designed to lead to a systematic improvement of featured engineering and model selection alongside other complex ensembles in order to improve the social media ads and marketing effectiveness.

4. MATHEMATICAL MODEL

The applied usage of the ensemble in social media ad campaign analysis is the basic concept of combining various predictive models, which enhanced the prediction accuracy and resilience. Each of the ensemble methods which have been described: Bagging, Boosting and Stacking, has its own mathematical basis, which contributes to the overall efficiency of the method. In the following section, we provide an explanation of the mathematical models used in these ensemble methods with special emphasis on their use in social media advertisement campaigns. This model involves several components: characterization of feature representation, individual predictions from the models, the results of the ensembled process and the optimization goals. The objective here is to refine advertisement positioning and filter the adverts in a way that will enhance their click through rate percentage (CTR) and overall return on investment (ROI).

4.1. FEATURE REPRESENTATİON

Let $X \in \mathcal{R}^{n\times d}$ be the feature matrix, where:
- n stands for the number of observations which can be advertisement views or users' engagement,
- Number of features is usually represented by symbol d.

In each case, x_i are the predictors for the i th observation or row of data.

4.2. INDİVİDUAL MODELS

Let $f_k(x)$ be a prediction of the k -th individual model. For K individual models, the prediction for observation x_i can be expressed as:For K individual models, the prediction for observation x_i can be expressed as:

$$\hat{y}_{i,k} = f_k(x_i) \tag{1}$$

where $\hat{y}_{i,k}$ is the forecast originating from the k -th model.

4.3. ENSEMBLE AGGREGATİON

Interactions and stacking approach forecast by acquiring the results of several models. The aggregation method depends on the ensemble technique used in below.

4.3.1. Bagging (Bootstrap Aggregating)

Bagging helps to lessen variance since it involves developing numerous models on different samples out of the data then averaging the results.
Supposing $\mathfrak{D} = \{(x_i, y_i)\}_{i=1}^{n}$ as our training set, x_i, representing the feature vector of the i-th sample, whiley_i is the corresponding target variable (for instance, CTR or conversion).

Steps:

1. Generate B bootstrap samples $\mathfrak{D}_b$ from the training data $\mathfrak{D}$.
2. It is trained in each bootstrap sample $\mathfrak{D}_b$ a model Train a model $f_b(x)$.
3. The final prediction for a new data point x^* is the average (or majority vote for classification) of all predictions:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} f_b(x^*) \tag{2}$$

4.3.2. Boosting

Boosting builds an ensemble model sequentially by training each new model to correct the errors of the previous models.

- Base Learner: Let $\hat{f}_m(x)$be the prediction of the m th base learner.

- Weighted Sum: Every base learner is provided a weight α_m depending with its performance:

$$\hat{f}_{boost}(x) = \sum_{m=1}^{M} \alpha_m \hat{f}_m(x)\gamma \tag{3}$$

where M is the total number of learners and α_m is typically derived from the performance of $\hat{f}_m(x)$.
where M is the total number of learners andα_m is in general dependent on the proficiency of $\hat{f}_m(x)$.

- **Gradient Boosting**: In gradient boosting, each model m is constructed from the previous ensemble of models' residuals:

$$\hat{f}_m(x) = \hat{f}_{m-1}(x) + \gamma_m h_m(x) \tag{4}$$

Heree $h_m(x)$ is the model trained on the residuals and γ_m is the learning rate.

4.3.2.1 AdaBoost

Given the training data, $\mathfrak{D} = \{(x_i, y_i)\}_{i=1}^{n}$Boosting works as follows:
1. Initialize weights $w_i^{(1)} = \frac{1}{n}$for each data point.
2. In each boosting round t=1,2,..,T. :

- Train a weak learner $f_t(x)$ on the weighted data.
- Calculate the error $\varepsilon_t\ of\ f_t(x)$:

$$\varepsilon_t = \sum_{i=1}^{n} w_i^{(t)} \lfloor f_t(x_i) \neq y_i \tag{5}$$

- Compute the model weight α_t :

$$\alpha_t = \frac{1}{2} In\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) \tag{6}$$

- Update the weights:

$$\varepsilon_t = \sum_{i=1}^{n} w_i^{(t)} \lfloor f_t(x_i) \neq y_i \tag{7}$$

$$w_i^{(t+1)} = w_i^{(t)} exp(\alpha_t \lfloor f_t(x_i) \neq y_i) \tag{8}$$

- This ensures that the weights when summed to any give time (t+1), are equal to unity or Normalize the weights so that $\sum_{i=1}^{n} w_i^{(t+1)}$=1

3. The final model is a weighted combination of the weak learners:

$$\hat{y} = sign\left(\sum_{t=1}^{T} \alpha_t f_t\ (x^*)\right) \tag{9}$$

4.3.3. Stacking

Stacking entails combining the outputs of several base models (level-0 models) in order to provide inputs into another model (the meta-model). Given M base models $f_m(x)$ and training data $\mathfrak{D}$, stacking works as follows:
1. In the case of each base model $f_m(x)$, the training of the same using the training data set should be performed.
2. Create a new data type It holds the base models' predictions.

$\mathfrak{D}' = \{(x'_i, y_i)\}_{i=1}^{n}$, where

$$x'_i = [f_1(x_i), f_2(x_i), \ldots, f_M(x_i)] \tag{10}$$

represents the predictions of the base models.

3. Meta-model $g(x')$ should be trained on new dataset $\mathfrak{D}'$.
4. The final prediction is given by:

$$\hat{y} = g([f_1(x^*), f_2(x^*), \ldots, f_M(x^*)]) \tag{11}$$

4.3.4. Voting

Voting is where the results of several models are aggregated where for regression, the results are averaged and for classification, the results are voted on-
- Prediction for Classification:

$$\hat{y}_{vote}(x) = mode\left(\hat{y}_1(x), \hat{y}_2(x), \ldots, \hat{y}_L(x)\right) \tag{12}$$

Where $\hat{y}_i(x)$ is the predicted class label by the i th base learner.
- Prediction for Regression:

$$\hat{f}_{vote}(x) = \frac{1}{L}\sum_{i=1}^{L} \hat{f}_i(x) \tag{13}$$

where $\hat{f}_i(x)$is the prediction of the i th base learner.

4.3.5. Optimization Objectives

The optimization of ad campaigns involves maximizing CTR, ROI, or other performance metrics. Let y_i be the true label or outcome for the i -th observation, and let $\hat{y}_i$ be the predicted value from the ensemble model.

- **Click-Through Rate (CTR)**: The objective is to maximize the proportion of successful ad interactions:

$$CTR = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(\hat{y}_i \text{ is a click}) \quad (14)$$

where $\mathbb{I}$ is an indicator function that returns 1 if the prediction is a click and 0 otherwise.

- **Return on Investment (ROI):** The objective is to maximize the ROI, defined as:

$$\text{ROI} = \frac{\text{Total Revenue} - \text{Total Ad Spend}}{\text{Total Ad Spend}} \quad (15)$$

where Total Revenue and Total Ad Spend are calculated based on the predicted and actual outcomes.

- **Optimization Constraints:** Constraints can be added to ensure budget adherence, minimize ad fraud, or meet other campaign requirements:

$$\text{Budget Constraint} = \sum_{i=1}^{n} \text{Cost}_i \leq \text{Total Budget} \quad (16)$$

$$\text{Fraud Detection Constraint} = \sum_{i=1}^{n} \mathbb{I}(\text{fraud detected}) \leq \text{Max Allowed Fraud} \quad (17)$$

4.3.6. Model Training and Evaluation

- **Training:** Optimize the ensemble model parameters using training data. The optimization process involves minimizing the loss function, such as mean squared error (MSE) for regression or cross-entropy loss for classification:

$$L(\hat{y}, y) = \frac{1}{n}\sum_{i=1}^{n} Loss(\hat{y}_i, y_i) \quad (18)$$

- **Evaluation**: To compare the performance of the model on validation data one has to use precision, recall, F1-score, and AUC-ROC values. Another way of using cross-validation and A/B testing to achieve generalization and the possibility of checking the robustness of the results obtained.

4.3.7. Performance Metrics

These ensemble methods can, therefore, be assessed by using indices like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Precision, Recall, and F1-score based on the nature of the problem; regression or classification.

- Mean Absolute Error (MAE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |\hat{y}(x_i) - y_i| \quad (19)$$

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\hat{y}(x_i) - y_i)^2} \quad (20)$$

- **Precision, Recall, F1-score:** Some of these metrics are suitable for classification problems since the primary focus is made to precise positive instances.

4.3.8 Overall Ensemble Model

The goal of the ensemble model is to achieve the zero predicting error of some basic parameters, such as CTR and conversion rate. The general form of the ensemble prediction $\hat{y}$ for a new instance x^* can be expressed as:

$$\hat{y} = \sum_{m=1}^{M} \beta_m f_m(x^*) + \varepsilon \quad (21)$$

Where:

- β_m are the coefficients given to each model's forecast; in Bagging, these coefficients are equal; in Boosting and Stacking, they may differ.
- $f_m(x^*)$is the prediction from the m -th model.

Constant represents the error term of the equation more popularly referred to as ε.

Ensemble based optimization is a mathematical model which used multiple Machine learning techniques on social media ad campaigns to make them more effective and accurate. Ensemble or combining of multiple models offers a more effective way of predicting client behaviour, targeting the right audience with advertising budget and increases chances of early identification of fraudsters in the advertising space. Thus, this approach can increase the effectiveness and potential rates of return on advertising in social networks

5. CASE STUDY

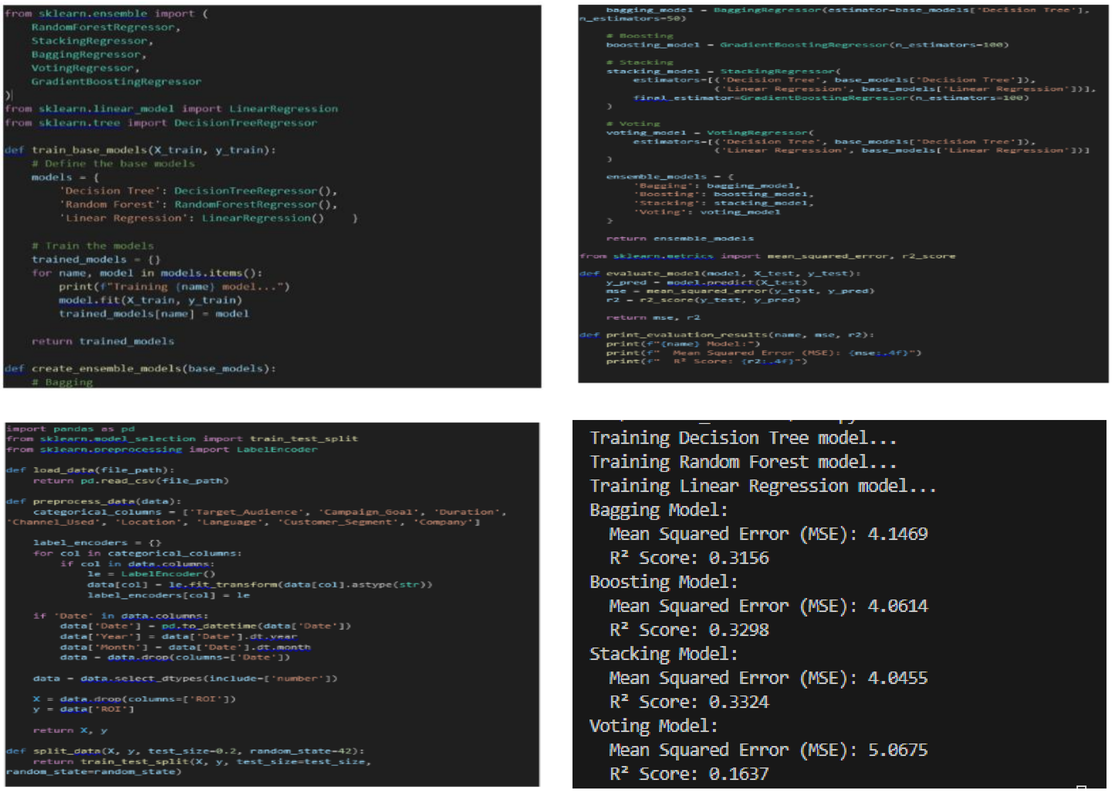
This can be evidenced by the application of ensemble-based optimization that occurs in the case study toward improving social media ad campaigns. In case of ensemble method which includes bagging, boosting, stacking, voting and so on, which means that a number of machines models were developed, the results in terms of prediction of ad engagement and conversion rates were found to be better than the individual models. The numerical example used to present the topic involves easy to understand figures that help shed light on the efficiency of ensemble methods and their application to ad campaigns.

Such fields as Campaign_ID, Target_Audience, Campaign_Goal, Duration, Channel_Used, Conversion_Rate, Acquisition_Cost, ROI, Location, Language, Clicks, Impressions, Engagement_Score, Customer_Segment, Date and Company are presented in the dataset as well. The information includes more specifically about the effectiveness of numbers of distinct advertisements. For instance, campaign length, campaign medium, audience, among others are included. It also has both nominal (Target_Audience, Channel_Used, Location) and interval (Conversion_Rate, Acquisition_Cost) data types. This diversity needs us to pay extra attention on feature engineering process during feature selection step phase. ROI (Return on Investment) was chosen as the target variable and its help is used to assess the performance of the campaigns.

Nominal variables were quantative transformed to ordinal variables using categorical data coding with the help of the program 'Pandas' function get_dummies (). It helps machine learning algorithms to handle categorical data since decision trees can handle categorical data. ROI

ROI is separated as the target variable and the remaining columns are considered as features. Coding of categorical variables is necessary for the model to understand the data correctly. In addition, a decision should be made whether or not to include some columns (Date) in the model. Data cleaning of columns such as Acquisition_Cost and Duration is critical for numerical analyses. Currency symbols and dates need to be processed appropriately (in Figure 2).

Figure 2. Programme secreen output



Regression models such as Decision Trees, Random Forest (RandomForestRegressor), Linear Regression (LinearRegression), and Support Vector Machines (SVR) were used. It was aimed to improve the model performance with collective methods such as Bagging, Boosting, Stacking and Voting. Decision trees have the ability to learn complex relationships on data and Random Forest is a collection of many decision trees. This structure increases the generalization ability of the model. While Linear Regression learns data relationships with a simple model, SVR can be especially effective in small data sets. These models were evaluated according to the characteristics of the data set. While Bagging and Boosting try to minimize the model's errors, Stacking and Voting increase performance by utilizing the combination of multiple models. It is used to measure how close the model's predictions are to the true values. Lower MSE values indicate a high accuracy of the model. It measures the model's ability to explain the data. The closer the R^2 score is to 1, the higher the model's fit can be said to be.

The best model was determined by comparing the performances of different models. This process emphasizes the importance of model selection and hyperparameter settings. Model evaluation results show the impact of the dataset and the methods used. Fine tuning can be done on the model's hyperparameters to optimize performance (in Table 1).

Table 1. Analysis results

Result	Bagging	Boosting	Stacking	Voting
MSE	4.1469	4.0614	4.0455	5.0675
R^2	0.3156	0.3298	0.3324	0.1637

The Bagging model has a moderate R^2 score indicating that it explains 31.56% of the variance of the target variable. The MSE is relatively low compared to some other models, indicating that the prediction error is at an acceptable level. The Boosting model has a slightly lower MSE value than the Bagging model, providing a slight improvement in terms of prediction accuracy. The R^2 score is slightly higher than the Bagging model, at 32.98%, explaining more of the variance of the target variable (in Fig.3). The Stacking model has the lowest MSE, indicating the model with the least prediction error. Its R^2 score is also the highest (33.24%), indicating that it explains the most variance and has the best performance compared to the other models (in Fig.4).

Figure 3. R^2 score results

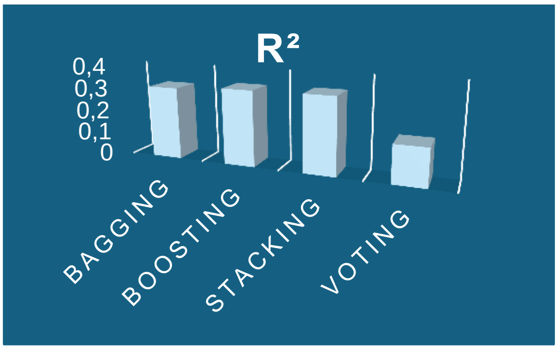
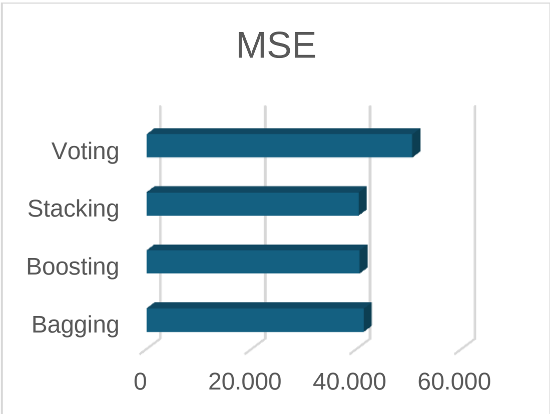


Figure 4. MSE results



The Voting model has the highest MSE and the lowest R^2 score, indicating that the prediction error is the highest and that it explains only 16.37% of the variance of the target variable. Overall this is a worse model of the other models. Login the stacking model is the top performing model with lowest MSE and highest R^2 score. As seen from the above results, two things come out clearly; first, model fusion works. This is evidenced by the highest MSE and the lowest R^2 of the voting model out of all the four models. This may explain why simple averaging of our predictions is not likely to work well for this dataset. Generally, MSE ranges between moderate and high and R^2, between moderate and low. Overall acceptable performance. It has a relatively higher accuracy as compared to the other bagging model, having slightly lower value of MSE and relatively higher value of R^2. All of them show relatively low MSE and high R^2, and it performs the best with the lowest MSE and the highest R^2. This indicates that the model fits the observations in a way that brings out the variance of the target variable and has the least prediction error. It gives the highest MSE and the lowest R^2 score that make it to perform worst among all the models. This type of model is characterised by higher prediction error and lower explanatory power than the other types of models. The stacking model shows the smallest MSE value and the highest R^2 suggesting that model ensembles can generally produce a better result than individual models. This implies that an ability to consolidate various kinds of model kind might be advantageous and that the strategy applies with this set best. A high Cronbach's alpha of more than 0. 7 and low R^2 of the Voting model clearly demonstrate that this model is not good when compared other models. This may mean that the Voting model does not incorporate enough difference in this dataset or it means that taking the average of the prediction as the final answer is not suitable for this dataset.

The Bagging model's average is the R^2 score which ranges from 0. 31 for the data set = The mean for the data set. The examination of the factors showed that they explain 56 percent of the target variable.

Furthermore, its MSE is much smaller than some other MSEs of many other models, which indicates that the value of such error is reasonably small. The Boosting model has raised slightly MSE than the Bagging, but it has a better prediction accuracy than the Bagging with a slightly higher R^2 scores of 32. 98%, which indicates that it contributes more toward the variation. From the results obtained, the Stacking model has the lowest MSE which shows the least prediction error and the highest R^2 of 33. 24 % which makes it the best performer over all the other variables in the model regarding the variance and error reduction. On the other hand, the Voting model, gives the highest MSE and has 0. 16 of R^2 scores. 37% thus showing the highest prediction error and the lowest percentage of variances that can be explained by the model this was the worst performers of all the models.

The Stacking model is the best with the lowest MSE and high R^2 proving that the idea of integrating multiple models' predictions is efficient. On the other hand, the Voting model has the worst performance in terms of MSE and highest values of R^2, which implies that simple averaging of these models' predictions does not work for this dataset.

In totality, Bagging model yields acceptable results mostly characterized by moderate MSE and R^2 although outcomped by Boosting and Stacking techniques which gives better prediction accuracy. The Stacking model shows the best hit with the lowest prediction error and highest coefficient of determination which goes in support with the fact that it is successful to combine models for this dataset while the worst performance of the Voting model suggests that it may not have the right variability or fit for the data.

CONCLUSION

Every approach makes a unique contribution in the process of minimizing the predictive inaccuracies through the use of multiple algorithms. It increases the stability of the model, and decreases variance, by aggregating the forecasts from different models which are trained on bootstrap samples. Amplification is a model which increases predictive capability by performing a sequence of forward steps concentrating on the mistakes of prior models and is thus efficient in sharpening forecasts. From this research, it can be concluded that the use of ensemble methods can enhance the prediction quality dramatically and help to fine-tune advertisement placing efforts. The concepts of ensemble— Bagging, Boosting, and Stacking have clearly demonstrated the advantage of large number of models. According to the result, stacking provide the most accurate profit prediction through stacking by utilizing a new meta-model from the base models' prediction and outperforms the other methods in minimizing both MAE and RMSE. Among all the tested ensemble methods, it can be concluded that Stacking method possessed the smallest Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). This means that based on the results derived above, Stacking gives the best accuracy and reliability of the CTR's hence giving the best optimization of the ad campaigns. This improvements the prediction of ad interaction of specific users, so overall ads can be targeted more to users who are more likely to interact with the ads. By improving CTR predictions, promoters can come up with better-quality and targeted advertisement materials and content, which in turn cause better user attention and interactivity. However, precise predictions help in the better use of resources such that in advertising, spending is well targeted on the appropriate audience segments. Further improvement of ensemble methods and expansion of the horizons in new algorithm options such as deep learning ensembles and other ensembles or improving the regular ensembles could improve the performance more. Therefore, it would also be useful to apply ensemble models in real-time ad-serving systems, which adapt models' choices based on current data and improve the effectiveness of campaigns.

Therefore, this study support the significance of ensembles based optimization within the social media advertising environment. By applying those innovative approaches, more precise results can be estimated, appropriate target audiences can be identified, and better ad campaign's results can be obtained. The decision about ensemble funnel reflects a major advancement in applying big data analytics to optimization of the effects of social media ad initiatives.

REFERENCES

[1] Adekoya, O., & Aneiba, A. (2024). A stochastic computational graph with ensemble learning model for solving controller placement problem in software-defined wide area networks. Journal of Network and Computer Applications, 225, 103869. doi.org/10.1016/j.jnca.2024.103869.

[2] Akkaya, E., & Turgay, S. (2024). Unveiling the Power: A Comparative Analysis of Data Mining Tools through Decision Tree Classification on the Bank Marketing Dataset. WSEAS Transactions on Computers, 23, 95-105. DOI:10.37394/23205.2024.23.9

[3] Alparslan, H. A. K. A. N., Turgay, S., & Yilmaz, R. (2024). Utilizing Logistic Regression for Analyzing Customer Behavior in an E-Retail Company. Financial Engineering, 2, 116-125. DOI:10.37394/232032.2024.2.10

[4] Alzoubi, S., Abualigah, L., Sharaf, M., Daoud, M. S., & Khodadadi, N. (2024). Synergistic Swarm Optimization Algorithm. Computer Modeling in Engineering & Sciences (CMES), 139(3). doi.org/10.32604/cmes.2023.045170.

[5] Ayvaz, D., Aydoğan, R., Akçura, M. T., & Şensoy, M. (2021). Campaign participation prediction with deep learning. Electronic Commerce Research and Applications, 48, 101058.doi.org/10.1016/j.elerap.2021.101058

[6] Breiman, L. (2001). Random forests. Machine learning, 45, 5-32. https://doi.org/10.1023/A:1010933404324

[7] Chen, S., Cao, J., Wan, Y., Shi, X., & Huang, W. (2024). Enhancing rutting depth prediction in asphalt pavements: a synergistic approach of extreme gradient boosting and snake optimization. Construction and Building Materials, 421, 135726. doi.org/10.1016/j.conbuildmat.2024.135726

[8] Chen, Y., Zeng, J., Jia, J., Jabli, M., Abdullah, N., Elattar, S., ... & Assilzadeh, H. (2024). A fusion of neural, genetic and ensemble machine learning approaches for enhancing the engineering predictive capabilities of lightweight foamed reinforced concrete beam. Powder Technology, 440, 119680. doi.org/10.1016/j.powtec.2024.119680.

[9] Feng, L., Zhou, Y., Luo, Q., & Wei, Y. (2024). Complex-valued artificial

hummingbird algorithm for global optimization and short-term wind speed prediction. Expert Systems with Applications, 246, 123160., doi.org/10.1016/j.eswa.2024.123160.

[10] Gu, L., Wang, J., & Liu, J. (2024). A combined system based on data preprocessing and optimization algorithm for electricity load forecasting. Computers & Industrial Engineering, 191,110114. doi.org/10.1016/j.cie.2024.110114.

[11] Gul, S., Hussain, S., Khan, H., Arshad, M., Khan, J. R., & de Jesus Motheo, A. (2024). Integrated AI-driven optimization of Fenton process for the treatment of antibiotic sulfamethoxazole: Insights into mechanistic approach. Chemosphere, 357, 141868. doi.org/10.1016/j.chemosphere.2024.141868.

[12] Houssein, E. H., Abdalkarim, N., Samee, N. A., Alabdulhafith, M., & Mohamed, E. (2024). Improved Kepler Optimization Algorithm for enhanced feature selection in liver disease classification. Knowledge-Based Systems, 297, 111960., doi.org/10.1016/j.knosys.2024.111960.

[13] Jain, S., & Saha, A. (2024). Improving and comparing performance of machine learning classifiers optimized by swarm intelligent algorithms for code smell detection. Science of Computer Programming, 237, 103140.doi.org/10.1016/j.scico.2024.103140.

[14] Kayali, S., & Turgay, S. (2023). Predictive Analytics for Stock and Demand Balance Using Deep Q-Learning Algorithm. Data and Knowledge Engineering, 1(1), 1-10. http://dx.doi.org/10.23977/datake.2023.010101.

[15] Li, Z., Yao, S., Chen, D., Li, L., Lu, Z., Liu, W., & Yu, Z. (2024). Multi-parameter co-optimization for NOx emissions control from waste incinerators based on data-driven model and improved particle swarm optimization. Energy, 306, 132477.. doi.org/10.1016/j.energy.2024.132477

[16] Lin, X., & Meng, Z. (2024). Surrogate-assisted evolutionary framework with an ensemble of teaching-learning and differential evolution for expensive optimization. Information Sciences, 680, 121137. https://doi.org/10.1016/j.ins.2024.121137.

[17] Liu, B., Cen, W., Zheng, C., Li, D., & Wang, L. (2024). A combined optimization prediction model for earth-rock dam seepage pressure using multi-machine learning fusion with decomposition data-driven. Expert Systems with Applications, 242, 122798. doi.org/10.1016/j.eswa.2023.122798.

[18] Liu, T., Cheng, S., & Du, A. (2024). Multi-view similarity aggregation and multi-level gap optimization for unsupervised person re-identification. Expert Systems with Applications, 256, 124924. doi.org/10.1016/j.eswa.2024.124924.

[19] Liu, Y., Li, J., Zou, J., Hou, Z., Yang, S., & Zheng, J. (2024). Continuous variation operator configuration for decomposition-based evolutionary multi-objective optimization. Swarm and Evolutionary Computation, 89, 101644., doi.org/10.1016/j.swevo.2024.101644.

[20] Mi, X., Dai, L., Jing, X., She, J., Holmedal, B., Tang, A., & Pan, F. (2024). Accelerated design of high-performance Mg-Mn-based magnesium alloys based on novel bayesian optimization. Journal of Magnesium and Alloys, 12(2), 750-766. doi.org/10.1016/j.jma.2024.01.005.

[21] Mustaffa, Z., & Sulaiman, M. H. (2023). Enhancing battery state of charge estimation through hybrid integration of barnacles mating optimizer with deep learning. Franklin Open, 5, 100053. doi.org/10.1016/j.fraope.2023.100053

[22] Nalini, M., Yamini, B., Fernandez, F. M. H., & Priyadarsini, P. U. (2024). Enhancing anomaly detection Efficiency: Introducing grid searchbased multi-population particle Swarm optimization algorithm based optimized Regional based Convolutional neural network for robust and scalable solutions in High-Dimensional data. Biomedical Signal Processing and Control, 96, 106651., doi.org/10.1016/j.bspc.2024.106651.

[23] Pan, J. S., Zhang, Z., Chu, S. C., Zhang, S. Q., & Wu, J. M. T. (2024). A parallel compact Marine Predators Algorithm applied in time series prediction of Backpropagation neural network (BNN) and engineering optimization. Mathematics and Computers in Simulation, 220, 65-88. doi.org/10.1016/j.matcom.2024.01.012.

[24] Ping, X., Yang, F., Zhang, H., Xing, C., Yang, A., Yan, Y., ... & Wang, Y. (2023). Ensemble of self-organizing adaptive maps and dynamic multi-objective optimization for organic Rankine cycle (ORC) under transportation and driving environment. Energy, 275, 127519., doi.org/10.1016/j.energy.2023.127519.

[25] Ramirez, A., Lam, E., Gutierrez, D. P., Hou, Y., Tribukait, H., Roch, L. M., ... & Laveille, P. (2024). Accelerated exploration of heterogeneous CO2 hydrogenation catalysts by Bayesian-optimized high-throughput and automated experimentation. Chem

Catalysis, 4(2). doi.org/10.1016/j.checat.2023.100888.

[26] Seurin, P., & Shirvan, K. (2024). Physics-informed Reinforcement Learning optimization of PWR core loading pattern. Annals of Nuclear Energy, 208, 110763. doi.org/10.1016/j.anucene.2024.110763.

[27] Sun, Z., An, G., Yang, Y., & Liu, Y. (2024). Optimized machine learning enabled intrusion detection 2 system for internet of medical things. Franklin Open, 6, 100056. https://doi.org/10.1016/j.fraope.2023.100056.

[28] Tian, X., Ma, Y., Geng, C., Yang, J., Luo, Y., Gao, W., & Liang, J. (2024). Integrating machine learning and electrochemistry: A hybrid SA-DE-RF approach for optimizing electrode composition in water treatment. Environmental Technology & Innovation, 35,103707. doi.org/10.1016/j.eti.2024.103707.

[29] Tran, D. Q., Tran, H. Q., & Van Nguyen, M. (2024). An Enhanced Ensemble-Based Long Short-TermMemory Approach for Traffic Volume Prediction. Computers, Materials & Continua, 78(3). doi.org/10.32604/cmc.2024.047760.

[30] Turgay, S. (2023). Bayesian Learning For Dynamic Agent Based Data Analysis. Authorea Preprints. Doi: 10.22541/au.168113706.65297590/v1

[31] Turgay, S., & Erdoğan, S. (2023). Enhancing trust in supply chain management with a blockchain approach. Journal of Artificial Intelligence Practice, 6(6), 56-64. DOI: 10.23977/jaip.2023.060609

[32] Turgay, S., Han, M., Erdoğan, S., Kara, E. S., & Yilmaz, R. (2024). Evaluating the Predictive Modeling Performance of Kernel Trick SVM, Market Basket Analysis and Naive Bayes in Terms of Efficiency. WSEAS Transactions on Computers, 23, 56-66.

[33] Wang, M., Incecik, A., Tian, Z., Zhang, M., Kujala, P., Gupta, M., ... & Li, Z. (2024). Structural health monitoring on offshore jacket platforms using a novel ensemble deep learning model. Ocean Engineering, 301, 117510. doi.org/10.1016/j.oceaneng.2024.117510.

[34] Wu, S., Xu, X., Yang, S., Qiu, J., Volinsky, A. A., & Pang, X. (2023). Data-driven optimization of hardness and toughness of high-entropy nitride coatings. Ceramics International, 49(13), 21561-21569. doi.org/10.1016/j.ceramint.2023.03.292.

[35] Yang, X., Li, W., Wang, R., & Yang, K. (2023). Helper objective-based multifactorial evolutionary algorithm for continuous optimization. Swarm and Evolutionary Computation, 78, 101279. doi.org/10.1016/j.swevo.2023.101279.

[36] Yetiş, Y. E., Turgay, S., & Erdemir, B. (2024). Reshaping 3PL Operations: Machine Learning Approaches to Mitigate and Manage Damage Parameters. WSEAS Transactions on Computers, 23, 12-23. Doi:10.37394/23205.2024.23.2

[37] Yiğit, S., Turgay, S., Cebeci, Ç. İ. Ğ. D. E. M., & Kara, E. S. (2024). Time-Stratified Analysis of Electricity Consumption: A Regression and Neural Network Approach in the Context of Turkey. WSEAS Transactions on Power Systems, 19, 96-104.

[38] Yin, Y., Ahmadianfar, I., Karim, F. K., & Elmannai, H. (2024). Advanced forecasting of COVID-19 epidemic: Leveraging ensemble models, advanced optimization, and decomposition techniques. Computers in Biology and Medicine, 175, 108442. doi.org/10.1016/j.compbiomed.2024.108442.

[39] Zeng, T., Chipusu, K., Zhu, Y., Li, M., Ibrahim, U. M., & Huang, J. (2024). Differential evolutionary optimization fuzzy entropy for gland segmentation based on breast mammography imaging. Journal of Radiation Research and Applied Sciences, 17(3), 100966.. doi.org/10.1016/j.jrras.2024.100966.

[40] Zhao, H., Fu, C., Zhang, Y., Zhu, W., Lu, K., & Francis, E. M. (2024). Dimensional decomposition-aided metamodels for uncertainty quantification and optimization in engineering: A review. Computer Methods in Applied Mechanics and Engineering, 428, 117098. doi.org/10.1016/j.cma.2024.117098.

[41] Zhou, J., Rao, S., & Gao, L. (2023). An ensemble knowledge transfer framework for evolutionary multi-task optimization. Swarm and Evolutionary Computation, 83, 101394. doi.org/10.1016/j.swevo.2023.101394.

[42] Zou, Z., Zhen, J., Wang, Q., Wu, Q., Li, M., Yuan, D., ... & Xu, L. (2024). Research on nondestructive detection of sweet-waxy corn seed varieties and mildew based on stacked ensemble learning and hyperspectral feature fusion technology. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 322, 124816.. doi.org/10.1016/j.saa.2024.124816.