E-COMMERCE MARKET SPEND OPTIMIZATION

V.Senthilkumar[1] S.S.Dhanya[2] S.Harishini[3] P.P.Kanissha[4]M.Kashika[5]S.Keerthika[6]
[1,2,3,4,5,6] Kumaraguru College of Technology,
CoimbatoreSenthilkumar.v.cse@kct.ac.in

Abstract—Nowadays, Advertising online helps to drive more traffic, and their chance of showing interest in the brand's product will be increased. But brands or businesses are unaware of where and how much to spend on online advertising to promote their product. Hence they need a predictive optimizing model to identify, which all platforms are suitable for their product to be advertised and the budget to be allocated for posting ads on different social media platforms based on the customer's views and interest

Machines must be taught to understand input to do the necessary analysis. Here we have done optimization of budget allocation for marketing. The analysis that we have done is particularly for the well-known social media platform which is Facebook and the reason why we have chosen that is that Facebook currently has 2.895 billion monthly active users (MAUs). Furthermore, the number of Facebook's daily active users (DAUs) currently stands at 1.908 billion people. The optimization is done on the values given by the user which includes age, gender, interest, impression, clicks, the amount spent on ads, total conversion, and approved conversion rate. The strategy used is E-commerce-commerce optimization which is a holistic way to improve a website it is done by E-Commerce certified experts and allows visitors to easily convert into customers.

Keywords— Market spend optimization, Machine Learning, K-nearest neighbors(KNN), Support vector Machine(SVM).

I. INTRODUCTION

E-commerce platforms such as Amazon, and eBay run hundreds of millions of auctions to sell advertising opportunities. Online advertising plays a crucial role in connecting advertisers and audiences and generates tremendous value for E-commerce platforms. There are different types of online advertising which include performance-based advertising, and branding-guaranteed advertising. We focus on performance-based advertising in the paper. In this marketplace, advertisers can specify the maximum daily amount they are willing to pay and get the audience through various pages in the E-commerce platform. The objectives of performance-based advertisers are usually to spend out the budget to maximize the performance goals (e.g., clicks, conversions as many as possible). Meanwhile, the ad-serving system is optimizing revenue on behalf of the platform.

One of the central issues for the ad-serving system of the E-commerce platform is matching ads to requests with these objectives above, which can be formulated as a constrained optimization problem. There are many challenges to achieving all the objectives simultaneously in a complex competitive environment. Each individual campaign has its own budget and performance goal, and there are hundreds of thousands of campaigns that compete with each other to acquire inventory in the marketplace. These varieties make optimization extremely difficult. In this paper, we present our work on optimal delivery in an E-commerce platform, which can be formulated as a constrained optimization problem that maximizes specified goals and is subject to budget constraints.

II. LITERATURE REVIEW

Nie Chen(2022) author of Research on E-Commerce Database Marketing Based on Machine Learning Algorithm discussed how Effectively e-commerce uses the information to maintain and enhance the emotional bond between stores and consumers while avoiding open confrontation with competitors, and exploring more refined database marketing applications strategies of e-commerce. The proposed solution is based on Customer type Classification Algorithms - Logistic Regression, Random Forest, support vector machine, and GBDT. This shows that the classification Algorithm GBDT (gradient lifting decision tree) has proved to be better than other algorithms. Its merits are If the number of features is too much, it is easy to overfit and train the model for too long. Demerits are the selection of characteristic variables is difficult, if the number of feature variables is too small, the classification characteristics are not obvious, and the classification effect will be reduced. Needs future analysis of the in-depth understanding of e-commerce database marketing status and improve feature construction from multiple perspectives based on e-commerce practical experience.

Jin Chen, Ju Xu, Gangwei Jiang, Tiezheng Ge, Zhiqiang Zhang, Defu Lian, and Kai Zheng proposed a problem statement for merchants and e-commerce companies, the increase in click-through rate (CTR) can be considered as an indicator of an increase in revenue. Therefore, much attention has been paid to creative design for improving the visual experience of advertisements. Proposed solution The proposed solution framework consists of the following phases: CTR Estimation using an efficient search strategy, Efficient exploration via Thompson Sampling. Output is An automated creative optimization framework to model complex interaction between creative elements and to strike a balance between exploration and exploitation. Merits are Both offline and online experiments, showing that our proposed approach performs better than the competing baselines and verifying the effectiveness of the
proposed algorithms. Automated creative optimization of advertisements help to promote any company or a product in effective manner based on the previous sales or preference data is future scope of this proposed model.

Hui Yang, Zhuohang Zheng , and hu Sun, in E-Commerce Marketing optimization of Agricultural Products Based on Deep learning and data mining proposed about Usage of e-commerce effectively in agricultural product marketing which reduces the difficulties encountered by the customers, solution was the Creation of sales prediction model –super crown model (SICM) of deep learning, and application of data mining technology in the field of eCommerce to promote the transformation of marketing optimization. Output derived was The convenience of e-commerce of agricultural products can stimulate customers to consume and further increase the sales of rural online stores. The Case analysis, and data analysis puts forward an optimization scheme of e-commerce agricultural products based on deep learning and data mining. Merits are The SICM model is more flexible and the data used to support the findings of the study upon the author's request. Demerits involve the Classification of data is difficult for the big data. Its future development is in Online selling and purchasing offering innumerable benefits to both sellers and buyers, and these advantages are also the reasons for the rising scope of eCommerce. Address notable challenges and inefficiencies in the agriculture supply chain by streamlining farmers' access to the customer and creating new links between steps in the value chain.

Chao Wei, Weiru Zhang Shengjie Sun Fei Li ,Xiaonan Meng in Optimal Delivery with Budget Constraint in E-Commerce Advertising. Focused on performance-based advertising by specifying the maximum daily amount they are willing to pay and get the audience through various pages in the E-commerce platform. Proposed a solution using linear programming, simulation system, and auction. Linear programming to optimize the overall revenue of the platform and improve different performance goals can be implemented in both display and search advertising. Simulation system to evaluate the results of different allocation plans by replaying auctions from online requests. Conjunction with pricing and ranking schemes it can effectively improve both revenue and various performance goals. Focus on incorporating real-time data into the problem setup to improve accuracy.
The offline simulation system needs to be improved.

Lee, J.; Jung, O.; Lee, Y.; Kim, O.; Park in A Comparison and Interpretation of Machine Learning Algorithm for the Prediction of Online Purchase Conversion discussed about Using the suitable machine learning algorithm to predict the consumer preference towards a particular product and the methodology to retarget them for future advertisement. Proposed the solution with XGB model for predicting online consumers purchase conversion. Applying XGB Explainer from the perspective of individual consumers, to analyse the prediction results of individual consumers by decomposing the contribution of each predictor. The XGB model is the most suitable machine learning model for predicting online consumers purchase conversion
It is necessary to analyse the effects of machine learning models through field experiments in future studies. In addition, by utilizing causality machine learning, new implications can be provided for machine learning research in the marketing field.

III. PROBLEM DEFINITION

Generating accurate and reliable sales forecasts is crucial in the E-commerce business nowadays. An E-commerce platform contains a large number of related products, in which sales demand patterns in different advertising platforms can be correlated. Optimization of Budget Allocation for marketing especially, advertisements based on CPC(cost per click), YTD(year to date)etc. , that is indeed the previous dataset. An accurate forecast can lead to huge savings and cost reduction thereby increasing business profit.

IV. PROPOSED SYSTEM

Our fundamental objective is to optimize the budget for marketing a specific product in marketing platforms. Initially, the dataset is collected and the necessary data pre-processing steps are done to make the dataset suitable. Once the data is processed, Data analysis using Exploratory data analysis techniques is done for better understanding and knowledge of the dataset characteristics. Then dataset is trained using various machine learning algorithms like Gaussian Naive Bayes, Linear Regression, KNN, SVM, etc. In the end, it is noted that, the accuracy rate of SVM is higher and chosen as the optimized model as it gives a 97%accuracy level.

Data deployment is done using Gradio GUI – a customizable GUI interface for deploying python machine learning codes.

Finally, a UI to display an optimized budget for a specified product on a social media platform(Facebook) is displayed using the gradio tool

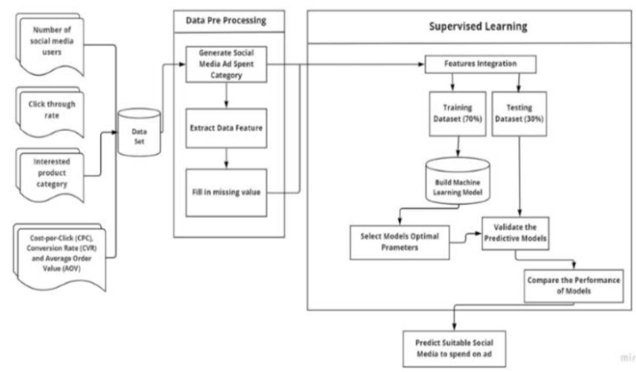

Fig 1: Market spend predictor architecture

V. IMPLEMENTATION OF THE SYSTEM

Collection of dataset, data importation:
exploratory , descriptive, statistical data analysis, building machine learning models:

Gaussian naive Bayes:

Naive Bayes is a basic but effective probabilistic classification model in machine learning that draws influence from Bayes Theorem.

Bayes theorem is a formula that offers a conditional probability of an event A taking happening given another event B has previously happened. Its mathematical formula is as follows: p(A/B)

Gaussian Naive Bayes

```
In [ ]: x = kag[['interest', 'Impressions']]
        x

In [ ]: y = kag[['Total_Conversion']]
        y

In [ ]: from sklearn import metrics
        from sklearn.metrics import accuracy_score
        from sklearn.model_selection import train_test_split

In [ ]: from sklearn.naive_bayes import GaussianNB
        GNB_classifier = GaussianNB()

In [ ]: GNB_classifier.fit(x, y)
        predicted = GNB_classifier.predict(x)

In [ ]: disp = metrics.plot_confusion_matrix(GNB_classifier, x, y)
        disp.figure_.suptitle("Confusion Matrix")
        print("\nAccuracy of the Algorithm: ", GNB_classifier.score(x, y))
```

Fig 2: Gaussian naïve bayes code

Linear regression:

Linear regression algorithm shows a linear relationship between a
dependent (y) and one or more independent (y) variables, hence called as
linear regression. Since linear regression shows the linear
relationship, which means it finds how the value of the dependent
the variable is changing according to the value of the independent variable

Linear Regression

```
In [ ]: from sklearn.linear_model import LinearRegression

In [ ]: linear = LinearRegression()

In [ ]: linear.fit(x,y)

In [ ]: linear.intercept_

In [ ]: linear.coef_

In [ ]: linear.predict([[45, 13]])

In [ ]: linear.score(x, y)
```

Fig 3: Linear regression code

KNN:

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on the Supervised Learning technique-KNN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.
K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for Classification problems.

KNN

```
In [ ]: from sklearn.neighbors import KNeighborsClassifier
        KNN_classifier = KNeighborsClassifier(n_neighbors=5, metric='euclidean')
        # KNN_classifier.fit(x_train, y_train)
        # predicted = KNN_classifier.predict(x_test)

        # disp = metrics.plot_confusion_matrix(KNN_classifier, x_test, y_test)
        # disp.figure_.suptitle("Confusion Matrix")
        # print("\nAccuracy of the Algorithm: ", KNN_classifier.score(x_test, y_test))

In [ ]: KNN_classifier.fit(x, y)
        predicted = KNN_classifier.predict(x)

In [ ]: print("\nAccuracy of the Algorithm: ", KNN_classifier.score(x, y))
```

Fig 4: KNN code

SVM:

Support Vector Machine(SVM) is a supervised machine learning algorithm used for both classification and regression. SVM works by mapping data to a high-dimensional feature space so that data points can be categorized, even when the data are not otherwise linearly separable.

SVM

```
In [ ]: from sklearn import svm

        svm_classifier = svm.SVC(gamma=0.001)
        svm_classifier.fit(x, y)

In [ ]: predicted = svm_classifier.predict(x)


        disp = metrics.plot_confusion_matrix(svm_classifier, x, y)
        disp.figure_.suptitle("Confusion Matrix")
        print("\nAccuracy of the Algorithm: ", svm_classifier.score(x, y))
```

Fig 5: SVM code

VI. IMPLEMENTATION SCREENSHOTS

Gaussian Naïve bayes:

ACCURACY OF GAUSSIAN NAIVE BAYES: 61.94%

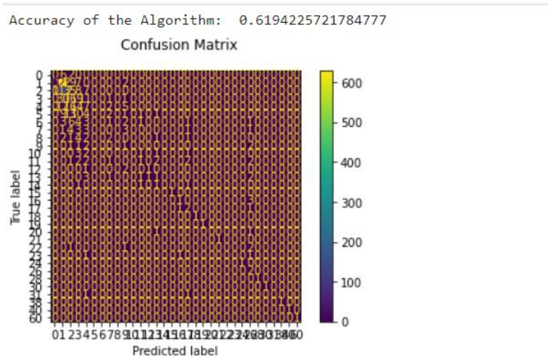

Fig 6: GNB confusion matrix

Support Vector Machine

ACCURACY OF SVM: 97.112%

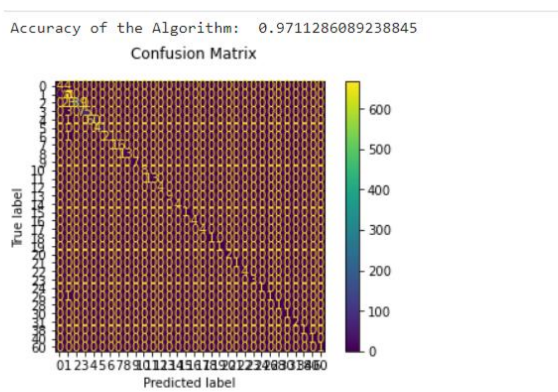

Fig 7: SVM confusion matrix

Gradio is a GUI library that allows is used to create customizable GUI components for your Machine Learning model. Thus the model deployment of the SVM algorithm is executed in Gradio open-source framework that is integrated with TensorFlow or Pytorch models.

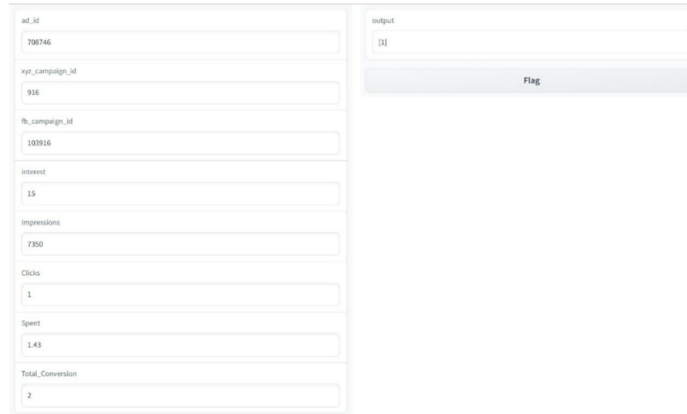

Fig 8: Model output to show Approved_conversion

VII. CONCLUSION

In this project, we have done the marketing optimization concept in the social media platform. The optimization is done on the values given by the user which includes age, gender, interest, impression, clicks, the amount spent on ads, total conversion and approved conversion rate, etc. Thus we have presented our work on optimal delivery in an E-commerce platform, which can be formulated as a constrained optimization problem that maximizes specified goals and is subject to budget constraints. We have explored the gradio gui tool for the output display.

REFERENCES

[1] Chen, N., 2022. Research on E-Commerce Database Marketing Based on Machine Learning Algorithm. Computational Intelligence and Neuroscience, 2022

[2] Yang, H., Zheng, Z. and Sun, C., 2022. E-Commerce Marketing Optimization of Agricultural Products Based on Deep Learning and Data Mining. Computational Intelligence and Neuroscience, 2022.

[3] Wei, C., Zhang, W., Sun, S., Li, F., Meng, X., Hu, Y., Lee, K.C. and Wang, H., 2019, October. Optimal Delivery with Budget Constraint in E-Commerce Advertising. In 2nd Workshop on Online Recommender Systems and User Modeling (pp. 46-58). PMLR.

[4] Lee, J., Jung, O., Lee, Y., Kim, O. and Park, C., 2021. A Comparison and Interpretation of Machine Learning Algorithm for the Prediction of Online Purchase Conversion. Journal of Theoretical and Applied Electronic Commerce Research, 16(5), pp.1472-1491.

[5] Y. Qian, “Exploration of machine algorithms based on deep learning model and feature extraction,” Mathematical Biosciences and Engineering, vol. 18, no. 6, pp. 7602–7618, 202.

[6] H. Lee, S.-H. Chung, and E.-J. Choi, “A case study on machine learning applications and performance improvement in learning algorithm,” Journal of Digital Convergence, vol. 14, no. 2, pp. 245–258, 2016.

[7] O. F. Althuwaynee, A. Aydda, I.-T. Hwang et al., “Uncertainty reduction of unlabeled features in landslide inventory using machine learning t-SNE clustering and data mining apriori association rule algorithms,” Applied Sciences, vol. 11, no. 2, p. 556, 2021.

[8] P. Jisu, Hyuk, Park, and Jong, “Enhanced machine learning algorithms: deep learning, reinforcement learning, and q-learning,” Journal of Information Processing Systems, vol. 16, no. 5, pp. 1001–1007, 2020.

[9] L. Zhou, W. Wang, J. Xu, T. Liu, and J. Gu, “Perceived information transparency in B2C e-commerce: an empirical investigation,” Information & Management, vol. 55, no. 7, pp. 912–927, 2018.

[10] G. Giantamidis, S. Tripakis, and S. Basagiannis, “Learning Moore machines from input-output traces,” International Journal on Software Tools for Technology Transfer, vol. 23, no. 1, pp. 1–29, 2019.

[11] J. Anson, M. Boffa, and M. Helble, “Consumer arbitrage in cross-border e-commerce,” Review of International Economics, vol. 27, no. 4, pp. 1234–1251, 2019.

[12] G. Walsh and D. Brylla, “Do product returns hurt relational outcomes? some evidence from online retailing,” Electronic Markets, vol. 27, no. 4, pp. 329–339, 2017.