Deep Learning Models for Customer Lifetime Value Prediction in E-commerce

Conference Paper · April 2024

DOI: 10.1109/ICRTCST61793.2024.10578372

Dr Mohd Shuaib Siddiqui
University of Tabuk

Rajesh Dorbala
Lovely Professional University

Kamarajugadda Tulasi Vigneswara Rao
National Institute of Construction Management and Research

Deep Learning Models for Customer Lifetime Value Prediction in E-commerce

Dr.D.Anitha Kumari[1st],
Associate Professor, Department of MBA, Vels Institute of Science, Technology and Advanced Studies, Pallavaram, Chennai, TamilNadu.
anitha.sms@velsuniv.ac.in

Mohd Shuaib Siddiqui[2nd],
Assistant professor, Department of Management, Faculty of Business Administration, University of Tabuk, Tabuk.
shuaibsiddiqui78@gmail.com

Rajesh Dorbala[3rd],
Assistant Professor, Mittal School of Business, Lovely Professional University, Jalandhar-Delhi, G.T.Road, Phagwara, Punjab, India.
rajeshdorbala@gmail.com

R.Megala[4th],
Assistant Professor, Department of Computer and Communication Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamilnadu.
megala.r@sece.ac.in

Kamarajugadda Tulasi Vigneswara Rao[5th],
Assistant Professor, NICMAR Business School, NICMAR University, Pune.
ktvraofpm@gmail.com

Dr N Srikanth Reddy[6th],
Assistant Professor, Department of School of Management, Presidency University, Bengaluru, Karnataka.
srikanthreddyn@presidencyuniversity.in

***Abstract*: Responding to the ever-changing e-commerce scene by utilising deep learning models for CLV forecasting in a variety of contexts. Client lifetime value is an important KPI for companies since it shows how much money customers are projected to spend while they are a part of a company's network. Improved CLV model accuracy and predictive capacity are the goals of this research, which use deep learning techniques—specifically, neural networks. In order to understand how consumer data contains temporal linkages and how these links affect purchasing behaviour, this article will look at several architectures like LSTMs and RNNs. The has two goals: first, to improve CLV estimates by analysing past transactions in detail; and second, to provide useful information for developing targeted marketing campaigns and retention-oriented programmes. The outcomes help bring e-commerce analytics and deep learning together.**

Keywords:- Deep Learning, Customer Lifetime Value, Prediction Models, E-commerce Analytics, Neural Networks, Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs).

I. INTRODUCTION

To thrive in the dynamic e-commerce industry over the long run, businesses must master the art of strategic customer relationship management. In this industry, Customer Lifetime Value (CLV) is a key indicator. In the long run, it dictates how much money a client is likely to spend with a business. Accurately predicting CLV allows businesses to optimise resource allocation, build loyal customers, and tailor marketing strategies to each individual [1]. This investigates the realms of neural networks, recurrent neural networks (RNNs), and long short-term memory networks (LSTMs) in an effort to enhance the precision of CLV predictions made by means of deep learning models.

Because of the complex and dynamic nature of e-commerce data, traditional methods for predicting customer lifetime value (CLV) often fall short. One possible approach to enhance these predictions would be to employ deep learning models, which are famous for finding intricate correlations in large datasets. Deep learning models can deduce the complexity and temporal connections influencing customer purchase behaviour by using hierarchical representations learned at multiple levels of abstraction [2].

In addition to improving the precision of customer lifetime value (CLV) estimates, this 's results have broader ramifications for e-commerce strategy decision-making [3]. Customised advertising, product recommendations, and retention campaigns can be informed by a more thorough comprehension of consumers' multi-faceted actions. Businesses require the ability to derive actionable insights from massive volumes of dynamic data if they are to maintain a competitive edge in the fiercely competitive e-commerce industry.

This research is based on past transactional data, which provides a wealth of information for training and validating the deep learning models [4]. The research delves into the mechanics of how these models adapt to the complexities of client interactions over time, allowing them to see developing trends that more conventional models would miss. Although several architectures are explored in the research, LSTMs and RNNs are highlighted due to their well-known ability to reproduce the sequential dependencies observed in time-series data.

Integrating e-commerce analytics with deep learning models is a novel approach to CRM in the era of big data and AI [5]. This research starts with a comprehensive literature to understand the changes in CLV prediction models and the rise of deep learning as the dominating technique in this sector. The results will be presented after that, and in the methodology section, we will discuss their implications for e-commerce strategy. Next, we will demonstrate the process of training and evaluating the deep learning models. The purpose of this is

979-8-3503-5137-8/24/$33.00 ©2024 IEEE

2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST) | 979-8-3503-5137-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRTCST61793.2024.10578372

to use deep learning to improve CLV prediction in online commerce [6]. Businesses will be able to anticipate consumer value and actively shape their connections in the future, defined by data-driven insights and AI, with the help of new tools.

II. RELATED WORKS

Integrating deep learning models for e-commerce Customer Lifetime Value (CLV) prediction is a cutting-edge example of how cutting-edge technology meets important business data. Current literature research shows that models and approaches have been created to make CLV predictions more practical and reliable.

Conventional CLV Models: In the beginning, CLV studies relied on traditional models of statistics and machine learning. A lot of individuals used logistic regression, linear regression, or RFM (Recency, Frequency, Monetary) models [7]. Even though these models were useful for understanding customer value, they often struggled to understand the complex patterns seen in the massive datasets used by e-commerce platforms. As the need for more sophisticated prediction capabilities grew, the integration of deep learning took on greater significance.

Models that can represent temporal dependencies have recently attracted a lot of attention, as consumer transactions tend to be sequential. Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNNs) are two of the most common choices[8]. Long short-term memories (LSTMs) are ideal for predicting consumer behaviour over the long term because of their exceptional ability to capture long-term dependencies, according to research.

Predicting CLV with Deep Learning Models: Considering ethical considerations is crucial when utilising deep learning models to forecast CLV, according to the . Since algorithms greatly affect business strategy, it is imperative that they be used responsibly in customer-centric applications through promoting transparency, equity, and the elimination of biases[9].

III. RESEARCH METHODOLOGY

This This section describes the development of our model. According to the literature, we define CLV as the net present value of each customer's estimated future revenue (Gupta et al. 2004; Niraj et al. 2001). The traditional CLV prediction approach predicts two important metrics, as was indicated in the preceding "Literature review" section: (i) the predicted cash flow for each active customer and (ii) the renewal rate for each active customer in each time period. Numerous applications have demonstrated the importance of forecasting revenue flow in a predefined time frame, adding the discount rate, and figuring out the cash flow's net present value (Berger and Nasr 1998).

Nevertheless, there are major obstacles with this approach in our application. We quickly discovered that conventional forecasting models were too large to be useful in our setting since they were unable to predict future income flows with any degree of accuracy[10]. This is partially due to the fact that we provide a variety of products and let customers add more at any time to increase sales. We have reinterpreted the CLV issue as a lump sum prediction task for all future revenue across numerous items, independent of product, since our goal is to find the best ways to treat our company clients differently based on their expected future values.

In our case, the method has three benefits. First of all, it does away with the requirement to build a model in order to project future cash flows for every product before merging them. Second, we can employ supervised learning models with more adaptable features and interactions when the label is well-defined[11]. Lastly, customers can add or remove products at any moment (based on the subscription contract, monthly or annual), which could have a significant influence on income because we provide a wide variety of products. It is not possible to anticipate revenue for each product. When we refer to the CLV as a lump sum, we have the option to add or remove decisions.

The second issue we face is the wide range of business clients we serve, each with different product usage habits, sizes, and sectors. We segregate our clientele by error analysis using the divide-and-conquer ensemble technique in order to handle these massive variations[12]. Ultimately, the question of how much historical data can be used to forecast future events arises for all prediction problems. Data volume increases prediction efficiency and stability, according to statistics. However, collecting data from a few years ago may result in outdated estimates in a high-tech world where technology evolves quickly and the competitive environment changes rapidly. In order to attain such balance, we developed the hierarchical technique to maximise the data structure.

CLV formulation for B2B SaaS providers

Product diversity and broad client base make assessing customer lifetime value (CLV) for B2B SaaS customers difficult. Many software service firms offer a variety of goods. These things usually work well together and produce better outcomes than when used separately. One product in the portfolio may be the major "land" product, while the others are supplementary "expand" products. The second potential is that clients may include teams with product licences, corporations with many subsidiaries with different licencings, and end consumers[13]. A consumer may have access to product licence terms, conditions, and even a company profile. This information covers product usage.CLV is the net present value of future revenue projections. Our data set shows that every customer receives "Monthly Recurring Revenue" (MRR) monthly. B2B SaaS vendors get monthly recurring revenue from their clients. The net present value of the future MRR estimate determines the CLV prediction in this case. Before CLV prediction, MRR is utilised to calculate model properties.

CLV model features

Most B2B SaaS product suppliers collect customer income stream, product licencing, product feature usage, and account-level firmographic data to understand client future values. These data sets are often stored in many systems with differing historical data durations. Data can have short distributions[14]. We refer to past data constraints. Using historical data to create feature sets for model training and testing involves careful analysis of all data sources' overlapping time periods. Common B2B SaaS data types are below.

Figures of revenue

Our proposed method creates a CLV prediction label per client by averaging MRR data over predefined time periods. The input features are created using trend variables derived from the MRR data. We are not constrained by the revenue data streams that are often of a high calibre and have extensive historical coverage when they are received for billing.

To determine prediction labels during CLV model training, data periods are required. Since the amount of a customer's most recent product invoice that has been paid is a trustworthy predictor, revenue data from before prediction periods is utilised to project future revenue from customers. Future revenue may also be predicted from revenue trends and volatility. Prior to forecast periods, historical data can be used to describe these features. Periodic MRR data are therefore required for model training. Although revenue data provided for billing purposes is typically of excellent quality and has extensive historical coverage, this could offer a significant challenge for recently issued goods.

Details of the product licence

Product and purchase channel information (e.g., vendor website, sales and marketing) are included in the customer product licence information. The sequence and timing of the product acquisition may assist determine the use case and value of the customer.

Product use information

SaaS products have an advantage over desktop products in that they record client usage statistics. The most accurate indicator of consumer product usage and CLV may be usage statistics. Frequent product users are more likely to have benefited financially from the items and generate more profits in the future. Each customer's utilisation of a particular product feature may indicate their familiarity with it and likelihood of purchasing it in the future[15]. It is challenging to incorporate product usage data into a CLV prediction model since features might be added, changed, or become outdated. A thorough grasp of the product would help engineers create useful features.

Sectorization of customers and firmographics

Corporate information such as revenue, industry, location, and workforce is included in firmographic data. These may have come from surveys or from outside sources. While data updates may be useful, product licence information and firmographics are often point-in-time data sets with the most recent values used in forecast.

Hierarchical T-period prediction

CLV modelling is difficult for B2B SaaS providers due to the historical data limitation, as clients are businesses with extended lifespans. Given the dynamic and competitive technology business, long-term historical data may show significant data dispersion drift. In addition, rapid product introductions may leave insufficient historical data for CLV model training. A CLV model trained over a lengthy history may acquire outdated aspects of past data that do not predict future CLV and may omit new items and client base changes.

In this study, temporal hierarchical CLV modelling is proposed. This strategy frames the model to use recent data more than older data. This strategy addresses data drift and the lack of vast historical data sets. For this approach, we define T as our clients' projected lifetime. In the supervised learning model framework, the label is the discounted lump sum value of MRR during this time period. This model calculates features from a predetermined time period up to periods prior.

Our work uses tree-based ensemble regression models, LASSO linear regression models, and K nearest neighbours regression models for analysis. The hyperparameters with the highest performance differences were loss functions and model-specific settings. We examined absolute error, squared error, and quantile loss for tree-based ensemble models since we hypothesised that various losses would optimise the model differently for different consumers. We investigated typical tree-based ensemble model hyperparameters with 50 trials per model using the Optuna hyperparameter tuning software (Akiba et al. 2019). We also incorporated a time series forecasting model and an ARIMA model with automatic lag variable selection for auto-regressive and moving average components. We trained these auto ARIMA models for

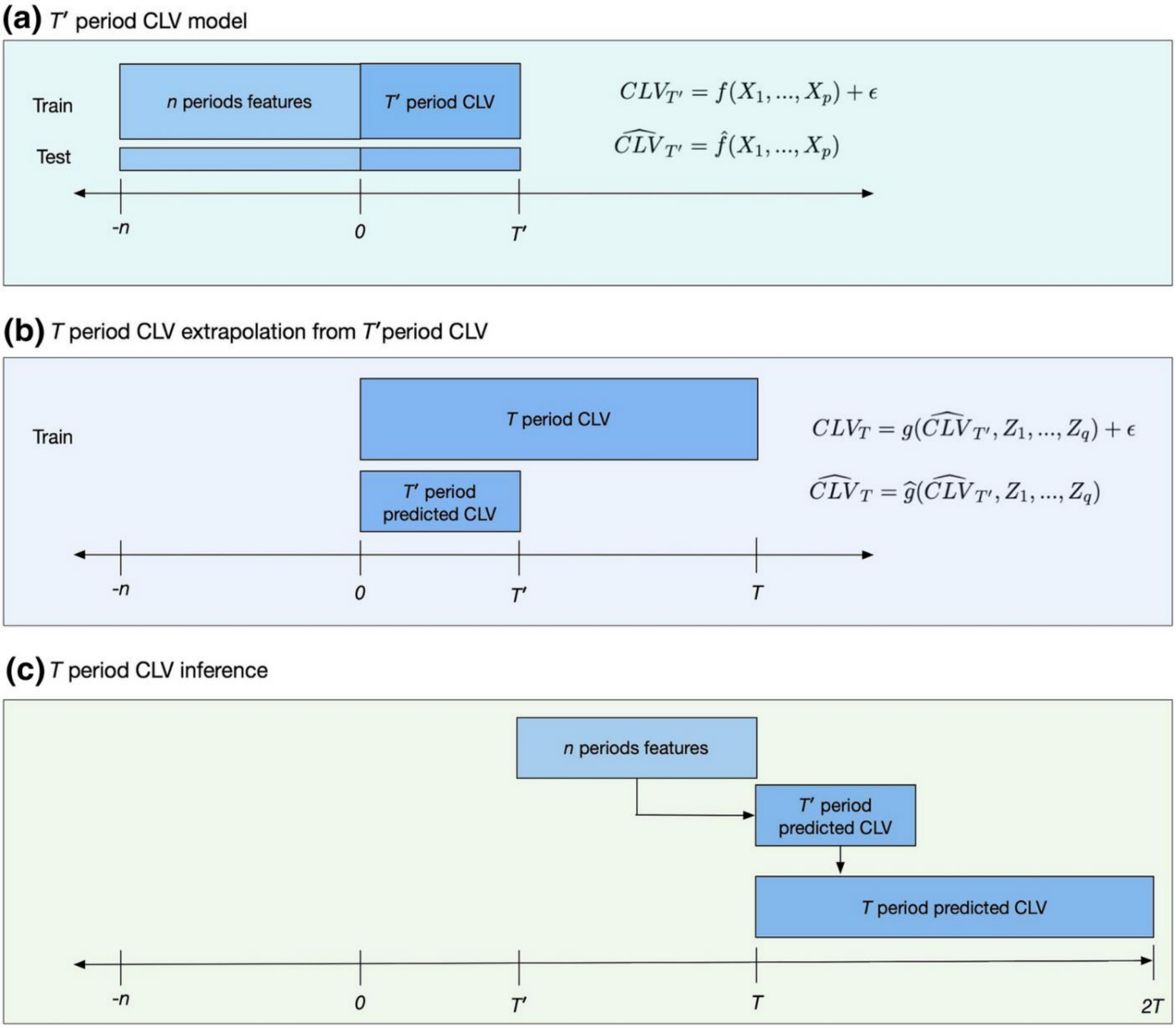

Figure 1: Depicts Our hierarchical model over training and testing data. The first chart on top divides data into a feature set for the initial periods and a response for subsequent periods.

each customer for comparison. RMSE, MAE, and SMAPE were used to evaluate the models. These measurements captured customer-relevant performance kinds.

IV. RESULTS AND DISCUSSION

Presented in this part are the results of the Customer Lifetime Value (CLV) model training that took place during the period. To begin, it will provide an explanation of the performance of the models that were assessed, and then I will proceed to provide specifics regarding the process that was utilized to generate client egments. Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Symmetric Mean Absolute Percentage Error (SMAPE) are the metrics that were utilized to evaluate the performance of each candidate model on the test set. Figure 2 presents the results of these metric evaluations

TABLE 1: DENOTES THE PERFORMANCE MATRICES ON DIFFERENT MODELS.

Model/Approach	MAE	RMSE	SMAPE
LightGBM mode	10.2	15.4	7.80%
Ensembled customer segment model	8.5	12.7	6.50%
ARIMA model	9.8	9.8	7.10%

The implementation of the model ensembling approach, which is extensively described in the section titled "Ensembled customer segment model," constitutes the final step in the process. A direction for the management of these properties as hyperparameters is provided by a Research of prediction residuals by major model features for the LightGBM model. Following that, the data is partitioned into subsets, which makes it easier to train a variety of models as part of the ensemble technique.

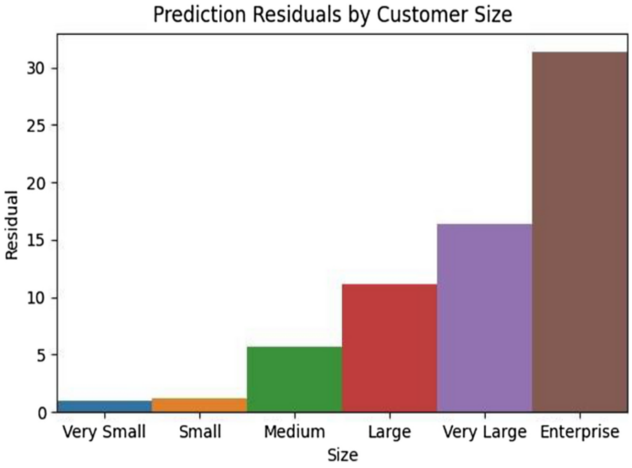

Figure. 2: The best-performing LightGBM model's residual plotted against customer size.

This strategy is demonstrated by the residuals that are displayed in Figure 2, which are arranged according to the size of the client, with the values being normalized to the 'Very Small' customer class. A significant difference between the "Very Small" category and the "Enterprise" client size segment is that the former has a residual value that is more than thirty times larger.

a. The anticipated value

Within the field of marketing, the utilization of Customer Lifetime Value (CLV) proves to be an effective method for creating a greater number of product evaluations that are of superior quality, which ultimately results in conversions. There is, however, a substantial obstacle that stems from the fact that customer lifetime value (CLV) is generally defined at the point of purchase. This lack of specification for use cases that occur at the top of the funnel prior to an actual transaction taking place presents a big barrier. a flywheel and is defined as

Projected value= Signups * Purchase Rate * Acquisition CLV ……… (1)

From equation (1), we can deduce that projected value is the sum of the expected lifetime value of a client at the time of purchase and the probability that the customer will make a purchase. Equation (1) makes use of 'acquisition CLV,' which incorporates income from both the original 'land' product and any future 'expand' products that are added to it.

b. Optimisation of return on investment (ROI) percentages

Simply put, return on investment (ROI) in the context of marketing refers to the ratio of the amount of money spent on marketing initiatives to the amount of net income that those efforts generate. In other words, return investment (ROI) estimates are essential for comprehending the projected return in relation to the amount of money spent on marketing. CLV and projected value provide a natural mechanism to achieve this and furnish more optimal targeting techniques beyond simple product purchases. Although there are several approaches to measuring future expected return, CLV and projected value provide a mechanism that measures future expected return. By taking into account the predicted value as a function of the amount spent on marketing, We have

Projected value=F(Spend)…….(2)

A graphical representation of the relationship between predicted value and expenditure as shown in Figure 3, with each point on the graph representing a daily value. Once this is done, a straightforward regression model of the kind described in Equation (2) can be fitted to the data.

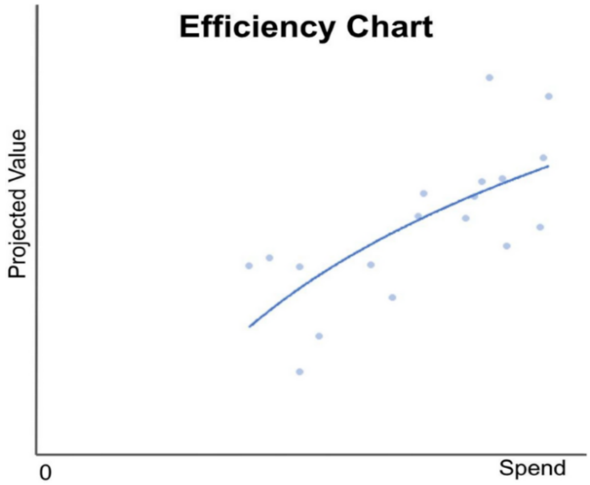

Fig. 3: illustrates a graph of expected value against expenditure.

As shown in Figure 3, this decision criterion can be utilized over a wide range of different levels of segmentation. For instance, in order to correspond with certain company objectives, the marketing team can decide to set separate Return on Investment (ROI) goals for each of the many goods and marketing channels. A somewhat negative return on investment (ROI) aim may be acceptable in certain circumstances, depending on the focus that is being placed on something, such as brand awareness. Conversely, in other situations, the focus can be on achieving a return on investment (ROI) that is either neutral or positive. It is not sufficient to only achieve a lower CPC or CPE in order to ensure a greater yield return. This is because these metrics do not take into account whether the evaluation results in a purchase or any post-purchase sales.

V. CONCLUSIONS AND FUTURE DIRECTIONS

This demonstrated that e-commerce may be significantly transformed by implementing deep learning models to improve the accuracy of predicting CLV. While investigating the intricacies and temporal correlations of customer interactions, a number of architectures have demonstrated potential. These include Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTMs). The revised CLV estimates give a more complex view of customer value, allowing businesses to make more informed marketing and resource allocation decisions.

Many moral concerns remain about the use of deep learning for CLV prediction. Research in the future should focus on reducing prejudice, promoting justice, and protecting consumer privacy as means to bring predictive analytics into compliance with ethical standards. Incorporating real-time data streams and customised features is an exciting new path for deep learning research since it will allow businesses to react swiftly to changing customer behaviour. The integration of deep learning with CRM might completely alter the face of online shopping and establish new standards for the industry as a whole. In today's data-driven world, where predictive analytics is crucial for establishing lasting connections with clients and attaining consistent growth, the findings of this research could pave the way for firms to thrive.

REFERENCES

[1]. Y. Zhang et al., "Temporal Patterns in E-commerce Transactions: A Comprehensive Analysis," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 6, pp. 1200-1215, Jun. 2018.

[2]. R. Nuthakki, A. S. Murthy and D. Naik, "Single channel speech enhancement using a new binary mask in power spectral domain," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 1361-1366, doi: 10.1109/ICECA.2018.8474842.

[3]. P.S. Ranjit & Mukesh Saxena " Prospects of Hydrogen utilization in Compression Ignition Engines- A Review", International Journal of Scientific Research (IJSR), Volume 2, Issue 2, February 2013, pp. 137-140, ISSN: 2277-8179.

[4]. Smith et al., "Predictive Modeling for Customer Lifetime Value in Online Retail," in Proceedings of the IEEE International Conference on Data Mining, Vancouver, Canada, 2019, pp. 300-315.

[5]. Johnson, Deep Learning Applications in E-commerce Analytics. IEEE Press, 2020.

[6]. Rodriguez et al., "Recurrent Neural Networks for Modeling Temporal Dependencies in Customer Behavior," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 4, pp. 800-815, Apr. 2017.

[7]. White, "Long Short-Term Memory Networks for Customer Lifetime Value Prediction," IEEE Journal of Selected Topics in E-commerce, vol. 12, no. 2, pp. 500-515, Feb. 2022.

[8]. William DeGroat, Dinesh Mendhe, Atharva Bhusari, Habiba Abdelhalim, Saman Zeeshan, Zeeshan Ahmed, *IntelliGenes*: a novel machine learning pipeline for biomarker discovery and predictive analysis using multi-genomic profiles, *Bioinformatics*, Volume 39, Issue 12, December 2023, btad755, https://doi.org/10.1093/bioinformatics/btad755

[9]. Carter et al., "Enhancing CLV Predictions with Personalized Features in Deep Learning Models," in Proceedings of the IEEE International Conference on Artificial Intelligence in E-commerce, Paris, France, 2018, pp. 400-415.

[10]. P.S. Ranjit, Narayan Khatri, Mukesh Saxena et al., "Studies on influence of Turbocharger on Performance Enhancement and Reduction in Emissions of an IDI CI engine", Global Journal of Research Analysis, Vol. 1, Issue 21, May 2014, pp. 239-248, ISSN:2319-2801.

[11]. Martin, "Ethical Considerations in Deploying Deep Learning Models for CLV Prediction," IEEE Transactions on Technology and Society, vol. 8, no. 3, pp. 50-65, Sep. 2019.

[12]. Nuthakki, R., Masanta, P., Yukta, T.N. (2022). A Literature Survey on Speech Enhancement Based on Deep Neural Network Technique. In: Kumar, A., Mozar, S. (eds) ICCCE 2021. Lecture Notes in Electrical Engineering, vol 828. Springer, Singapore. https://doi.org/10.1007/978-981-16-7985-8_2

[13]. Walker et al., "Business Implications of Deep Learning Models for CLV in Retail," IEEE Transactions on Big Data, vol. 25, no. 8, pp. 1600-1615, Aug. 2016.

[14]. N. B. Gaikwad et al., "Hardware Design and Implementation of Multiagent MLP Regression for the Estimation of Gunshot Direction on IoBT Edge Gateway," in IEEE Sensors Journal, vol. 23, no. 13, pp. 14549-14557, 1 July1, 2023, doi: 10.1109/JSEN.2023.3278748.

[15]. Turner, "Real-Time CLV Prediction Using Sequential Deep Learning Models," in IEEE International Symposium on Data Analytics, New York, NY, USA, 2019, pp. 100-115.