2023 2nd International Conference for Innovation in Technology (INOCON)
Bangalore, India. Mar 3-5, 2023

Customer Behavior Prediction using Deep Learning Techniques for Online Purchasing

Nisha
Associate Professor, Department of CSE(Data Science), Noida Institute of Engineering and Technology,
Greater Noida, Uttar Pradesh, India
Email id- nishaverma@niet.co.in

Ajay Shanker Singh
Professor, Department of Computer Science and Engineering, Galgotias University,
Greater Noida, Uttar Pradesh, India,
Email id- ajay.shankersingh@galgotiasuniversity.edu.in

2023 2nd International Conference for Innovation in Technology (INOCON) | 979-8-3503-2092-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/INOCON57975.2023.10101102

***Abstract*— E-commerce success may be improved with a deeper knowledge of how customers make purchases online. The primary emphasis of prior research has been on consumers' intent to buy, with sales rank serving as a proxy. Furthermore, there is no guarantee that a shopper's retail buy intent would translate into action. Our goal is to get a deeper comprehension of how online shoppers use an e-commerce platform by anticipating their actions based on a huge multidimensional data sample consisting of more than 50,000 individual web sessions. In this investigation, we identified two groups of variables—platform engagement and customer characteristics—as significant predictors of retail consumers' propensity to shop online. We also compared our deep learning approach's predictive efficacy to that of some of the most popular supervised learning methods for forecasting, such as Decision Trees, Support Vector Machines,Random Forests, and ANNs. When compared to other machine learning methods, the deep learning approach performed better on the same dataset. These studies will aid in the planning stages of platform creation, and they will also contribute to the growth of the academic literature on the topic of purchase forecasting for online e-commerce platforms.**

Keywords — Customer Behavior Prediction, Deep Learning Techniques, and Online Purchasing

I. INTRODUCTION

In order to better serve their clientele, most businesses nowadays use consumer behaviour mining frameworks. Various sales, forecasting, and marketing trends may be anticipated with the use of these models. Machine learning and data mining techniques have been used to make predictions about consumer habits. Traditional methods, on the other hand, take longer to implement and can't accurately forecast users' actions [1]. Insurance firms may chart their success by tracking the amount of new policies they sell each year. The customer defection prediction model is vital to the survival of the business if it is to continue expanding and attracting new clients. Customer behaviour analysis and churn prediction are essential even if a business provides excellent service. Our primary contribution is a probabilistic classifier that can anticipate the client who will quit the insurance provider, allowing the business to take preventative measures to retain them[2]. Businesses today are making great strides in providing individualised services, smart care, and a stress-free buying experience for their clientele. The success of a business' interactions with its consumers hinges on its capacity to observe, analyse, and anticipate their activities. An organization's success and expansion may be greatly aided by its ability to accurately foresee its customers' wants and needs. Data about customers has been gathered and used with a variety of machine learning methods to forecast their likely actions. When conditions are optimal, traditional approaches fail to uncover underlying patterns and must be enhanced to provide more correct estimates [3]. Automated human age estimate from face images is a challenging but important task with many potential legitimate uses, particularly with the rise of social media and online communities. When it comes to combating negative behaviour and mental abuse, hearty faced affirmations structure are in very high demand. Various programmes collaborate to confirm a customer's identity before granting them access to restricted areas (both online and off). However, it has become a very challenging problem to match a man's face from a data stand-in photo with his face photos from a database [4]. Since customers increasingly rely on internet evaluations to help them make decisions, dishonest companies have a strong incentive to fabricate positive feedback about their products and services. False reviews, often known as "deceptive reviews," are posted online with the intent to mislead customers. Common techniques for identifying fake reviews rely on linguistic and psychological characteristics of the reviews themselves [5]. Online retail purchasing has become more popular since it saves time and eliminates the need to wait in long checkout lines. What if shoppers could pick up their essentials and get out without wasting time in long lines at the registers? As a result, this may encourage more people to shop at retail establishments, which might boost earnings. "i-Shopping," a hybrid of deep learning and analytical processing, plans to do away with cash registers at retail establishments. Using RFID tags and sensor fusion, this is achievable. It is possible to use deep learning algorithms in retail settings to track a customer's movements around a store and identify the products they have selected or abandoned. Customers get a digital invoice via the application and the purchases are charged to their account instantly [6]. All major e-commerce sites now provide some kind of recommendation algorithm. Even while standard recommender systems benefit from knowing previous interactions with a consumer, this tactic is impractical for "cold start" situations. Suggestions for fresh or unregistered users, as well as the introduction of brand-new products, fall within this category [7]. A valuable asset for marketing automation used by client relationship managers is the ability to accurately anticipate the Client Lifetime Value (CLV) of a customer. Advances in machine learning (ML) approaches have been applied to this problem in recent years due to the low predictive capacity of traditional methods based on recency, frequency, and historical data of profits and revenues. In spite of this, current methods typically fall short when it comes to simulating specific temporal patterns that are prevalent in reality, such as consumers' cyclical purchase behaviour [8]. Conversion rates from promotional purchases rely heavily on the specifics of customers' desire to buy at the promotion's micro level. Customers may easily access and compare pricing and services across channels to help in their purchasing choice in the context of combined promotions on several online channels. One further way to foretell consumers' purchasing decisions in response to sales events is to examine their interactions with various marketing channels [9]. Customers' deciding factors are heavily influenced by online reviews. Scammers and spammers may now influence consumer behaviour by spreading false information in the form of reviews, either by pushing nonexistent goods or by smearing competitors' offerings. As a result, it is more crucial than ever to tell bogus from genuine evaluations. When it comes to text classification, the standard technique relies on a bag-of-words model to depict text, which leads to sparsity and phrase representations learned from neural networks that are incapable of handling unfamiliar terms [10].

II. Literature Review

In this research, we use smart methods, such fuzzy grouping and machine learning approaches, to probe client portfolios in search of shopper preferences. Achieving this goal required using hierarchical fuzzy clustering to the problem of computing the connection between items and purchase criteria. The maximum error categorization issue is alleviated since data that is comparable are put together, as determined by the analysis. The prediction accuracy is then enhanced by adding a deep neural network that has been trained with the best possible parameters. The effectiveness of the mentioned system is measured across many datasets with corresponding performance indicators. Comparisons with existing single prototype and hybrid model-based techniques revealed that the suggested methodology achieved the highest accuracy and the lowest error rate [11]. This research makes use of machine learning methods to analyse lion insurance records. The study's second major addition is its use of an unsupervised approach to classify data on 12007 rows with 9 attributes, from which the K-means++ algorithm then formed 2 clusters. The training dataset was subjected to synthetic minority oversampling because of the unbalanced cluster findings. With an impressive 98.81% accuracy rate, the Deep Neural Networks method emerges as a leading model for churn prediction. Lion Insurance's two years' worth of client data was utilised to train, test, and assess the model. Each method was optimised with the help of a random selection process. However, a deep convolutional neural network with the structure of produced the greatest results (9-55-55-55-55-55-1). In this research aiming to forecast customer attrition, this classification technique was used [12]. In this paper, we present a unique hybrid model that combines a novel grouping module based on an optimised fuzzy deep network with a customer behaviour prediction module based on a deep recurrent network. We used learning parameters to examine customers' past buying habits and portfolio information. The butterfly optimization strategy, which seeks to reduce the maximum error in classification, was used to enhance the deep learning strategies used in this work. In order to gauge the system's efficacy, we conducted an experimental investigation. The suggested method outperformed both standalone and combined models [13]. This is due to the fact that people's faces change significantly under varying operating settings. Examples include learning, rebellion, outward appearance, camera shops, photo labs, film development, cosmetics, and eyewear. In this study, we show that by using convolutional neural networks (CNNs) for feature extraction, we can greatly improve the precision of gender prediction. The state-of-the-art result was accomplished by using a Convolutional Neural Network (CNN) that was fully integrated with deep learning techniques. Extensive testing on the biggest publicly accessible datasets of face photos with Gender labels (IMDB-WIKI dataset [14]) is used to assess the accuracy of image-based Gender determination. When compared to more conventional techniques, the suggested deep learning approaches perform much better because of their superior versatility in automatically extracting the required information. We have developed and evaluated many Deep Neural Network (DNN) based systems for spotting fake reviews, and we encourage you to do the same. We also found that a widespread issue is dealing with the widely varying durations of these assessments. To deal with the wide range of review texts, we have presented two approaches: multi-instance learning and a hierarchical structure. Several benchmark datasets of fake reviews have shown promising experimental outcomes, outperforming the current state-of-the-art. We also tested the suggested approach on a secondary job involving reviews, sentiment detection, and found that it reached state-of-the-art accuracy on two benchmark datasets [15]. A content-based algorithm for predicting purchase intent for cold-start session-based situations is presented here: the Purchase Intent Session-bAsed (PISA) algorithm. Both the content modelling and the prediction of the user's desire to buy are deep-learned processes in our system. Based on our research, PISA performs better than a commonly used deep learning base when novel things are presented. Not only that, but our method does well in severely unbalanced datasets, where content-based methods often fall short. Finally, our tests demonstrate that integrating PISA with baseline further increases performance in non-cold start conditions [16]. This study provides a deep learning-based approach to forecasting customers' purchasing behaviour depending on mouse click event. In order to better understand the online shopping habits of consumers, businesses often keep detailed logs of every click a user makes on a website. These data are increasing in volume and variety on the internet and need an advanced machine learning technique for analysis. .. For predictive modelling, we've settled on a Wide and Deep Learning Networks that's been tweaked to take use of the greatest features of both types of networks in a combined learning neural network. With its capacity to learn behaviour from both high- and low-order interactions of features, as well as to take use of the memorising skill of linear models and the generalisation capability of deep learning networks, the Deep and Wide neural network beats previous models of its kind. Predictions made using the Wide and Deep model are more accurate (up to 78.26 percent) than those made using other models (such as the deep neural network, the product-based neural network, and the factorization-machine support neural network) according to experiments conducted on a real-world dataset of 33 million clicks [17]. To overcome these drawbacks, we provide an alternative strategy for CLV prediction that utilises a hybrid of several ML methods. Our solution is grounded in a custom deep learning methodology using sequential recurrent networks with enhanced temporal convolutions, implemented as an encoder-decoder pair. Then, in a hybrid setup, we add gradient boosting machines (GBMs) and several other new characteristics to this model. Real-world data from a big e-commerce firm and a publicly available dataset from the field of online retail support empirical assessments showing that the sequence-based model already yields competitive performance outcomes. The GBM model stacks well with it and increases accuracy further, suggesting that the two concepts capture distinct data patterns. [18] To predict the interactions between consumers and promotions channel, as well as nonlinear sequence correlations and cumulative impacts between customers' browsing behaviour, we offer a feature-combined deep learning framework that employs a full-connected long short-term networks (FC-LSTM). The framework integrates other aspects of the client profile, such as past purchases and demographics, to improve the effectiveness of the forecast. In a real-world prediction challenge, we use our technique to promote concert tickets across many internet channels. Extensive studies indicate that the suggested strategy outperforms state-of-the-art approaches on common measures including accuracy, recalls, f-measure, area under the curve (AUC), and lift [19]. In this research, we suggest a method that uses an ensemble of models built via the use of an aggregation methodology to provide final predictions; the method is based on the use of three separate models, each of which was trained using the notion of a multi-view learning method. In order to achieve this, we combine bag-of-n-grams and parallel convolution neural networks, both two techniques that have been shown to be effective in extracting information from unstructured text (CNNs). Using a modest kernel size in an n-gram embedding layer allows us to train local CNNs with the same amount of computational resources as would be needed to train a deep and sophisticated CNN. For better feature extraction from text, we employ an architecture based on convolutional neural networks (CNNs) that takes n-gram embeddings as input. We adopt a hybrid technique, combining linguistic aspects of the review text with elements of the reviewer's behaviour outside of the text, to identify fraudulent reviews. Our method is tested on the publicly accessible Yelp Filtered Dataset, where we get F1 scores as high as 92% for identifying fraudulent reviews [20].

III. Proposed Model

The purpose of this research was to investigate the efficacy of a DL strategy in finding important determinants of online purchases linked to platform engagement and customers. The DL method is applicable to this investigation because of the way its dense network model allows it to identify subtle patterns in datasets. In order to answer the research issue, this work combines a variety of analytic approaches, such as both classic ML methods and the cutting-edge DL method.

A. *Problem Statement*

- Predict a customer's happiness (positive or negative) with a purchase made on the Brazilian e-commerce site Olist based on their past behaviour when placing orders and reading product reviews.

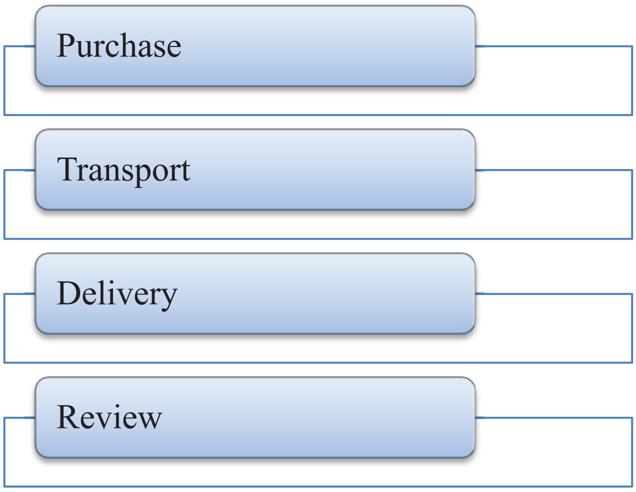

Fig. 1. Customer Behavior

Lots of information is needed for the process of predictive behaviour modelling. In-depth information on the consumers more specifically. The data of both the individual client and thousands of many other customers is necessary for a business to develop the computational intelligence model that can determine whether a certain consumer will perform a given action in the future.

The DNN is a refined DL method that originated from the ANN. A DNN differs from a conventional ANN in that its network topology includes both a fully connected layer and many sparsely connected intermediary layers. When compared to ANNs, its more sophisticated structure enables it to learn representations at numerous levels. The improved prediction performance of the DNN is a direct result of this improved representation learning. Image categorization, emotion recognition, and sales prediction are just some of the many data kinds and areas of study that have benefited from DL's versatility. As a consequence, several DL architectures tailored to certain tasks have emerged. A convolutional neural network (CNN) is a kind of neural network that has found use in image processing and is particularly well-suited to dealing with multi-dimensional data. For optimal performance, feed sequential data with temporal dynamics into a recurrent neural network (RNN). This research made use of feed-forward DNN technology. Due to the one-dimensional nature of the inputs in the dataset utilised for this investigation, we choose to use this DL method. DNN's data-driven self-adaptive approach, in contrast to the classic approaches' insistence on assumptions about the data's functional shape, further solidified our decision to deploy it here. The DNN is able to capture nonlinear connections in the dataset because its network topology is constructed in such a manner that the beginning layers learn that simple data characteristics while the hidden layers handled the more complicated features [47]. Additionally, the DNN has been demonstrated to be less susceptible to the dimensionality curse than ML regression-based models. Since the dataset includes both multi-class category and numerical variables, a DNN is an appropriate choice for this investigation.

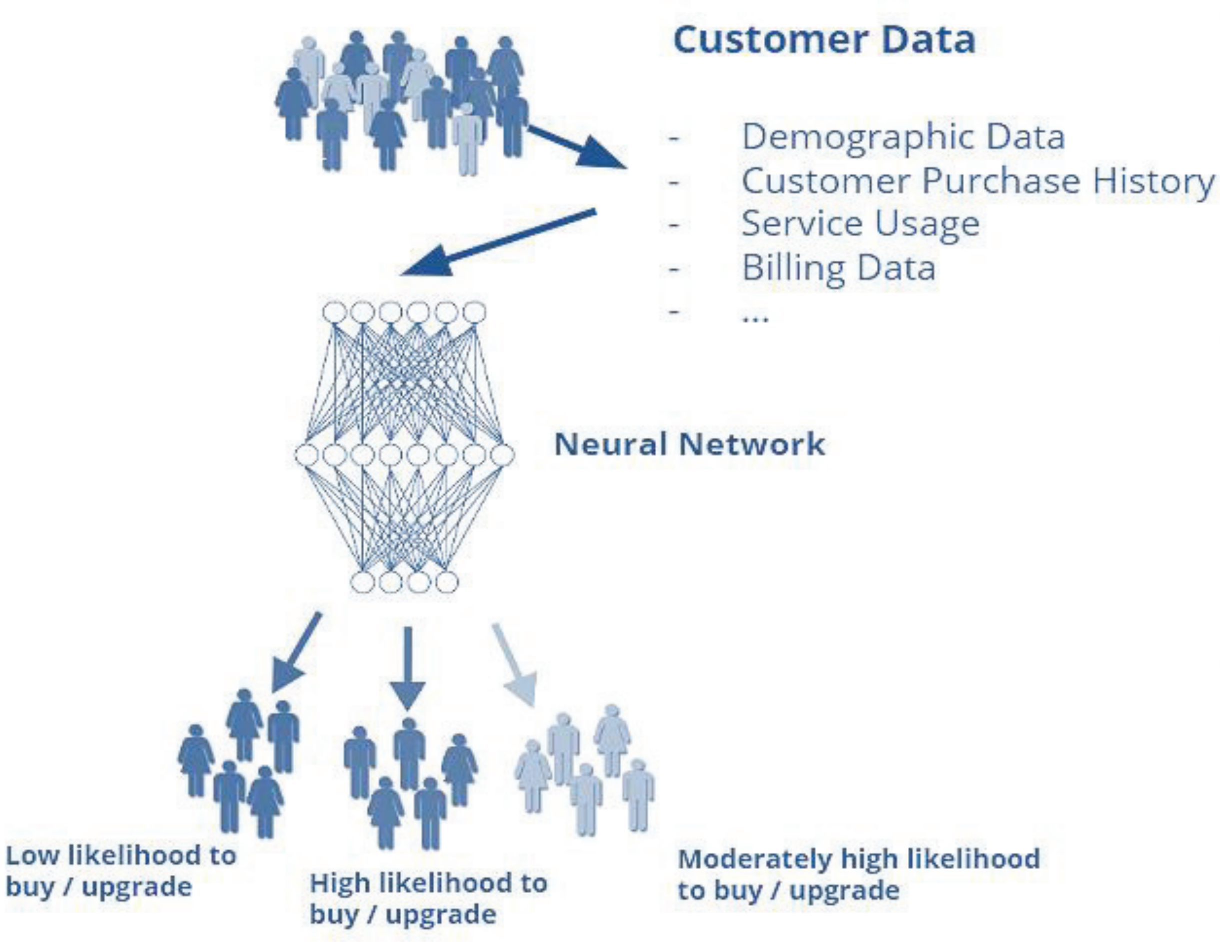

Fig. 2. Architecture

The behavior attribute features are follows:

- review id ID of the report submitted for the specified order.
- order id
- It's a special number associated with each customer's order.
- review score
- customer evaluation on a scale from 1 to 5 stars for each purchase.
- review comment title
- The Name of the Critical Analysis
- review comment message
- Customer feedback is displayed with each purchase.
- review creation date
- Whenever this review was made in time.
- review answer timestamp
- Date/time stamp of review response.

- If y hat is indeed a binary label, then it may be possible to get away with only a single neuron for output. In this case, A might be represented by the number 0, and B by 1, or vice versa..
- On the other hand, it is customary and reasonable to employ as many activation functions as there are actions.
- Two output neurons would be required if there were two potential courses of action. For this scenario, the network would spit out a two-element vector as its final result. When we talk about the label vector, we're talking about a binary vector where each element is either 1 or 0.
- • Probabilities between 0 and 1 would populate the elements of the output vector y. The Softmax function is suggested for usage as the functional in the last layer of a classification network because of the existence of mutually exclusive classes throughout the classification process.
- The probabilities in the output sequence y would fall between 0 and 1, as indicated by their numerical values. Since there are no overlapping classes in classification, the Softmax function is a good candidate for the output of the final hidden layer.

IV. EXPERIMENTAL RESULTS

Based on our comparisons, we conclude that DL methods significantly outperform traditional ML methods.

A crucial part of every machine learning application is the process of selecting features to use. The predicted accuracy of ML and DL systems may be enhanced by removing redundant variables, which also has the added benefits of speeding up the training procedure and decreasing the total cost of computation. Filter-based and wrapper-based approaches are two of the most common ways to pick features. As a part of the condition characterized phase, filter-based approaches (such as the correlation coefficient) are often used. Without relying on more fundamental prediction models, these approaches are prone to volatility. It has been discovered that wrapper-based approaches outperform filter-based methods by making use of prior information about the underlying learning methods. To develop an ideal subset of characteristics which can most effectively predict purchases, we employed a wrapper-based sensitive analysis-based feature selection strategy used in prior research. To do this, we redeployed the DL and ML methods using the characteristics presented in decreasing order. Therefore, the first input parameter was the most crucial, and so on. Variations in ML and Dl model average accuracy across 5 runs (corresponding to the five-fold cross-validation process)are shown in Fig. 2

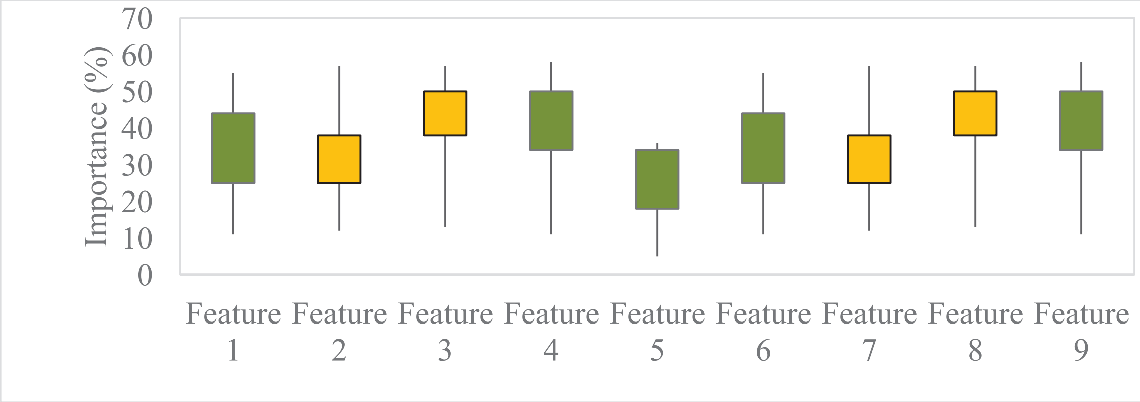

Fig. 3. Feature Selection for Amazon using proposed model

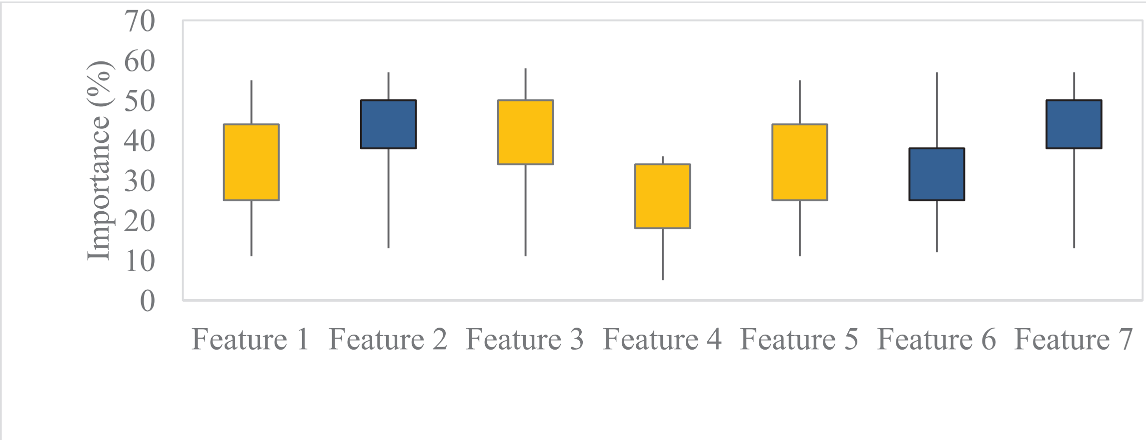

Fig. 4. Feature Selection for Flipkart using proposed model

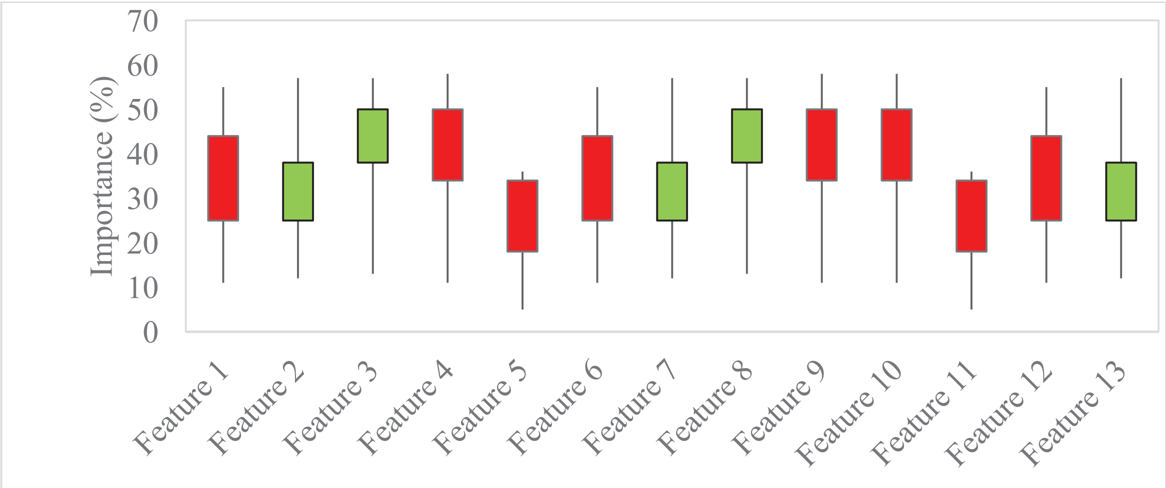

Fig. 5. Feature Selection for Snapdeal using proposed model

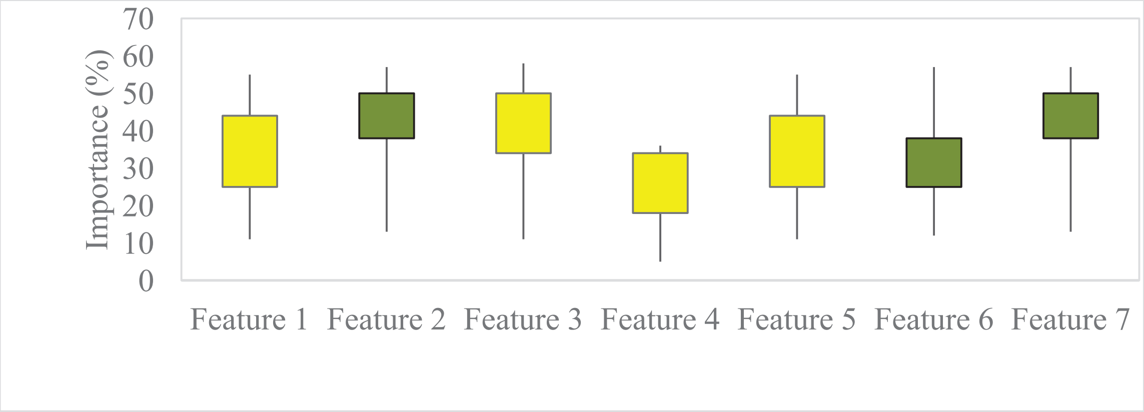

Fig. 6. Feature Selection for Meesho using proposed model

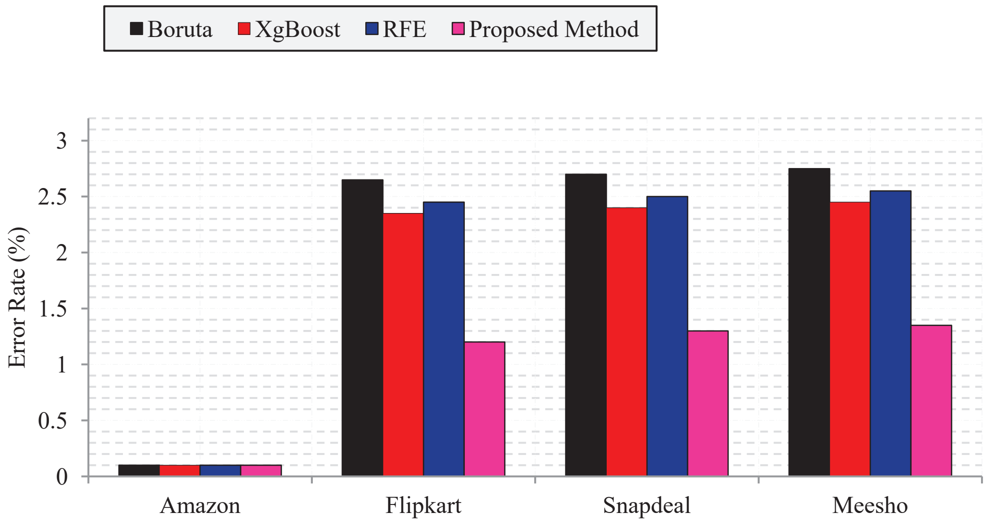

Fig. 7. Error Rate Analysis

V. CONCLUSION

The purpose of this research was to determine how well predictions of consumers' real buy behaviour on an online platform tracked with their actual purchase behaviour. To do so, it used a one-of-a-kind anonymised web surfing dataset including both purchase history and information about customers' interactions with the e-commerce website. Researchers found that two groups of factors—platform engagement and consumer characteristics—were the strongest predictors of purchases. We found that the time of day a session started, how long a customer has been using the platform, how long it has been since the user made a purchase, and the client's previous transaction all play a role in accurately predicting whether or not a customer would make a purchase. Additionally, the DNN fared better than other ML methods on the same datasets. These results will help retailers and developers of e-commerce platforms enhance the reliability of their recommendation systems. By giving these factors a higher priority in the current recommendation engines, these results may be used in retail IT systems. Predictions should be more accurate when these sometimes overlooked elements are taken into account, such as platform involvement and client characteristics. The results of this research will be used to enhance DNN-based purchase prediction on other platforms like it. However, there are still caveats to this research that prevent it from being completely conclusive. To begin, the focus of this research is on a specific e-commerce platform and its impact on retail sales. The dataset is big enough to use ML and DL methods for forecasting consumer behaviour, but a larger dataset may need adjusting the model to get optimal prediction results. Second, the information is representative of shopper actions in a certain product category while online shopping in Europe. There is a possibility that the results cannot be extrapolated to other populations or used with other kinds of products, even if they are accurate representations of the situation. To ensure that these results are applicable to a wide range of purchasing situations, further research is required. Statistically analysing the efficiency of different prediction models is one way we may enhance the results of this and future research to help us draw better conclusions. Third, the predicted precision of the model was not compared in real-time while the data was being created. This would have significantly greater practical ramifications, since it would have enabled us to provide advice to boost retail sales at precisely the right moment. It is our sincere desire that future research make use of such integrated methods in order to provide timely advice with significant practical ramifications for businesses. Finally, in the future, researchers might investigate the possibility of creating a deep learning-based frequent pattern mining approach that can be used in similar scenarios and then evaluating its efficacy in relation to the current standards.

REFERENCES

[1] Altameem, A.A., & Hafez, A.M. (2022). Behavior Prediction Scheme Using Hierarchical Clustering and Deep Neural Networks. Journal of Nanoelectronics and Optoelectronics.

[2] Kingawa, E.D., & Hailu, T.T. (2022). Customer Churn Prediction Using Machine Learning Techniques: the case of Lion Insurance. Asian Journal of Basic Science & Research.

[3] Altameem, A.A., & Hafez, A.M. (2022). Behavior Analysis Using Enhanced Fuzzy Clustering and Deep Learning. Electronics.

[4] Bhat, S.F., Lone, A.W., & Dar, T.A. (2019). Gender Prediction from Images Using Deep Learning Techniques. 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), 1-6.

[5] Jain, N., Kumar, A., Singh, S., Singh, C., & Tripathi, S. (2019). Deceptive Reviews Detection Using Deep Learning Techniques. International Conference on Applications of Natural Language to Data Bases.

[6] Venkatraman, S., R, G.N., & Machado, H.K. (2018). i-Shopping with Sensor Fusion for finding Customer Behavior using Deep Learning Algorithm. 2018 4th International Conference for Convergence in Technology (I2CT), 1-6.

[7] Shekasta, M., Katz, G., Greenstein-Messica, A., Rokach, L., & Shapira, B. (2019). New Item Consumption Prediction Using Deep Learning. ArXiv, abs/1905.01686.

[8] Nguyen, K., Nguyen, A., Vu, L., Mai, N., & Nguyen, B.P. (2018). AN EFFICIENT DEEP LEARNING METHOD FOR CUSTOMER BEHAVIOUR PREDICTION USING MOUSE CLICK EVENTS. KỶ YẾU HỘI NGHỊ KHOA HỌC CÔNG NGHỆ QUỐC GIA LẦN THỨ XI NGHIÊN CỨU CƠ BẢN VÀ ỨNG DỤNG CÔNG NGHỆ THÔNG TIN.

[9] Bauer, J., & Jannach, D. (2021). Improved Customer Lifetime Value Prediction With Sequence-To-Sequence Learning and Feature-Based Models. ACM Transactions on Knowledge Discovery from Data (TKDD), 15, 1 - 37.

[10] Ling, C., Zhang, T., & Chen, Y. (2019). Customer Purchase Intent Prediction Under Online Multi-Channel Promotion: A Feature-Combined Deep Learning Framework. IEEE Access, 7, 112963-112976.

[11] Javed, M., Majeed, H., Mujtaba, H., & Beg, M.O. (2021). Fake reviews classification using deep learning ensemble of shallow convolutions. Journal of Computational Social Science, 4, 883-902.

[12] Kao, L., Chiu, C., Lin, Y., & Weng, H.K. (2022). Inter-Purchase Time Prediction Based on Deep Learning. Comput. Syst. Sci. Eng., 42, 493-508.

[13] Baderiya, S.H., & Chawan, P. (2018). Customer buying Prediction Using Machine-Learning Techniques: A Survey.

[14] SadighZadeh, S., & Kaedi, M. (2022). Modeling user preferences in online stores based on user mouse behavior on page elements. J. Syst. Inf. Technol., 24, 112-130.

[15] Shah, S.M., Usman, S.M., Khalid, S., Rehman, I.U., Anwar, A., Hussain, S., Ullah, S.S., Elmannai, H., Algarni, A.D., & Manzoor, W. (2022). An Ensemble Model for Consumer Emotion Prediction Using EEG Signals for Neuromarketing Applications. Sensors (Basel, Switzerland), 22.

[16] Damian, A.I., Piciu, L., Turlea, S., & Tapus, N. (2019). Advanced Customer Activity Prediction Based on Deep Hierarchic Encoder-Decoders. 2019 22nd International Conference on Control Systems and Computer Science (CSCS), 403-409.

[17] Akbarabadi, M., & Hosseini, M. (2018). Predicting the helpfulness of online customer reviews: The role of title features. International Journal of Market Research, 62, 272 - 287.

[18] Sun, J., Tian, Z., Fu, Y., Geng, J., & Liu, C. (2020). Digital twins in human understanding: a deep learning-based method to recognize personality traits. International Journal of Computer Integrated Manufacturing, 34, 860 - 873.

[19] Yalan, Y., & Wei, T. (2021). Deep Logistic Learning Framework for E-Commerce and Supply Chain Management Platform. Arabian Journal for Science and Engineering.

[20] Jacob, M.S., Rajendran, S., Michael Mario, V., Sai, K.T., & Logesh, D. (2019). Fake Product Review Detection and Removal Using Opinion Mining Through Machine Learning. International Conference Artificial Intelligence, Smart Grid and Smart City Applications.