
| RESEARCH ARTICLE

An AI-Driven Classification System for Early Detection of Customer Churn in E-Commerce Platforms

Dhuli Shyam¹, Prabu Manoharan², Muzaffer Hussain Syed³, Uday Kumar Ragireddy⁴, Prasanth Varma Addepalli⁵ and Sridhar Reddy Bandaru⁶

¹Business Application, IT, University or Client: Nagase Holdings America Corp, Manager, Application & Software Development, NYC, NY

²Information Technology, University or Client: Bourns Inc, HRIS Manager, California, USA

³Director of IT Projects & Programs, Powersys Inc

⁴Sr Technical Program Manager, Vdrive IT Solutions, Inc, Richardson, Texas

⁵Lead Data Architect/ Engineer, Federal Motor Carrier Safety Administration, Atlanta, Georgia

⁶University or Client: Discover Financial Services, Application Architect for AI/ ML Platforms

Corresponding author: Dhuli Shyam, **Email:** dhulishyam@gmail.com

| ABSTRACT

Customer churn is a key issue for e-commerce organizations since it lowers revenue and hampers long-term growth. E-commerce companies may reduce expensive customer attrition and implement proactive retention measures by predicting client churn. This study proposes an AI-driven classification system for early customer churn detection using the List online retail dataset. The proposed CNN has a higher ability in capturing complex and nonlinear patterns of customer behavior as opposed to the conventional machine learning methods. The results of the experiment display excellent performance, with 99.62% accuracy, 99.32% precision, 98.66% recall, and 98.78% F1-score and with the AUC of 0.9893. The findings demonstrate that the model is robust, has generalization and small misclassification with respect to the baseline models like Random Forest, Logistic Regression, AdaBoost and ANN-MLP. In addition to predictive strength, the paper highlights an importance of DL in extracting latent features representations and provides actionable information on customer retention strategies. The results determine CNNs as a strong and justified solution in detecting churn that provides academic value and practical applicability to the e-commerce industry.

| KEYWORDS

Customer Churn Prediction, Business Analytics, E-Commerce, Machine Learning, List Online Dataset, CNN Model

| ARTICLE INFORMATION

ACCEPTED: 01 November 2025

PUBLISHED: 30 November 2025

DOI: 10.32996/jcsts.2025.7.11.42

I. INTRODUCTION

In the age of Big Data, organizations can access large amounts of customer-generated data in form of transactional, browsing, and feedback data as well as online interactions [1]. The need to derive valuable insights on this data has become critical in making sound business decisions, especially in customer-focused businesses such as e-commerce [2]. A key example of big data analytics use is the creation of classification mechanisms that may recognize the pattern of behaviour of its customers so that early churn of customers can be detected [3]. Customer churn, which is characterized as the propensity of customers to stop interacting with a service or platform, is a serious threat to business development. In e-commerce sites, customer retention is less expensive than acquisition due to the high competition and low switching costs [4][5]. This has led to early identification of customer churn

being a strategic priority with the aim of enabling businesses to minimize customer loss, enhance customer satisfaction and maximize revenue [6].

To address this challenge, the use of ML and Algorithms has gained momentum. These intelligent systems are capable of analyzing complex customer data, identifying hidden patterns, and making accurate predictions [7][8]. ML techniques have demonstrated remarkable success in churn prediction tasks [9]. The List online retail dataset is used in this study to present an AI-driven classification approach for detecting customer churn in e-commerce platforms. The system combines both conventional ML and DL methods for data preparation, feature selection, and model training. By applying techniques such as SMOTE for handling class imbalance and Neighbourhood Components Analysis (NCA) for Features election, the proposed framework improves prediction accuracy and interpretability. Ultimately, this research aims to support e-commerce businesses in deploying effective retention strategies through intelligent churn prediction.

A. *Motivation and Contribution of study*

This research is motivated by the ever-increasing requirement to utilise e-commerce platforms to sustain clients in a continually competitive online industry. As customer acquisition costs are increasing and consumer loyalty is becoming difficult to sustain, identifying potential churners at an early stage has become a key to maintaining business growth. The traditional rule-based systems are usually not effective in modelling the dynamic and complicated behaviour of the customers. Thus, the objective of this research is to harness the strength of AI, specifically, ML and DL algorithms, to create an intelligent and automated classification framework that can predict churn with high precision. The study's overarching goal is to provide online retailers with a proactive tool for client retention; this will aid in the happiness of online shoppers and the sustainability of their business by addressing issues such as data imbalance, feature selection, and model optimisation. Some of contributions of paper are as:

- Cleaned List Online dataset, engineered datetime features, encoded categorical variables, and standardized numeric
- Applied NCA to reduce dimensionality and retain 12 key features, while using SMOTE to address class imbalance among churned and non-churned customers.
- Convolutional, pooling, and fully connected layers were included in the proposed customised CNN, and its performance was confirmed using acc, prec, rec, F1score, and ROCanalysis.
- Identified critical features influencing churn, enhancing transparency and supporting actionable business decisions

B. *Justification and Novelty*

The CNN model proposed has shown to be much more advanced in customer churn prediction and is superior according to acc, prec, rec and F1score than those models that have been used traditionally in predicting churn to the customer. The strength of the architecture is demonstrated by its good generalization capabilities, capability to stay out of overfitting, and its consistency in loss values. The high AUC also justifies why it has been effective in classifying churned and non-churned customers with low misclassification. These findings represent the fact that deep learning, especially CNNs, has greater feature extraction and prediction capabilities over the classical models, thus represent a confirmed and meaningful decision when it comes to churn detection in large-scale online e-commerce databases such as List Online.

C. *Structure of Paper*

This is how the study is organised The previous research on churn detection is reviewed in Section II, the proposed method and models are explained in Section III, the results are shown in Section IV, and the future scope is concluded in Section V.

II. LITERATURE REVIEW

This section presents research on Customer Churn Detection in E-Commerce Platforms utilizes diverse ML techniques; the overview of these investigations is presented in Table I.

Ahmad, Jafar and Aljoumaa (2019) created a churn prediction algorithm to help telecom companies identify consumers prone to leaving. Using ML methods in a big data setting, the model created for this research presents a fresh perspective on feature engineering and selection. An AUC value of 93.3% was achieved as a standard metric for evaluating the model's performance. The prediction model's integration of consumer social networks via the extraction of SNA features is another significant development. In comparison with the AUC criteria, the model was able to achieve a higher percentage of 84 to 93.3% with the inclusion of SNA [10].

Ebrah and Elnasir (2019) three ML methods, NB, SVM, and DT, were employed to forecast churn employing two benchmark datasets. There are 71,047 observations and 57 attributes in the cell2cell dataset, compared to 7033 observations and 21 attributes in the IBM Watson. The models were evaluated using the AUC, which was 0.82, 0.87, 0.77 for the IBM dataset and 0.98, 0.99, 0.98 for the cell2cell dataset. Additionally, compared to earlier research utilising the same datasets, the recommended models were more accurate [11].

Chauhan et al. (2019) suggests a Customer Aware Recommender System (CUARS) that can provide push alerts in an online store. The created system uses an analytical methodology and is socially conscious. CUARS generates effective push notifications by utilising collaborative filtering techniques and a variety of components such as predictive analytics, Smart Scheduling, Feature Extraction, and ranking factorisation recommenders. CUARS is an extremely effective means of maintaining customer groups and minimising churn. Results from a comparison with a standard recommendation system that makes use of a predetermined timetable and event-based notification system show that CUARS improves CTR by about 30% [12].

Rahaman (2018) using a number of advanced ML techniques, including clustering, classification, and ensemble approaches, various groups of clients will be separated based on their behaviour, preferences, and purchasing patterns. The benchmark various algorithms for their performance, assess their scalability, and provide actionable insights for businesses on better ways to target their customers, optimize resource allocation, and improve customer retention. These results also point to the importance of refining segmentation techniques in order to drive up customer engagement and profitability in the competitive environment of ecommerce [13].

Subramanya and Somani (2017) use feature selection techniques and an expanded feature set to show how classifier efficiency has improved. The contributions made by this study may be generally divided into three categories. First, we go over how numerous implicit consumer contact indicators, such as sales and clickstream logs produced as a outcomes of customer interaction, may be extracted using well-known data mining techniques in the Hadoop architecture. Second, via Feature Engineering approaches we prove that some of the new features we offer have a demonstrable influence on customer churn. Finally, we illustrate how Regularised LR, SVM and GB Random Forests are the highest performing models for forecasting customer churn proven by rigorous cross-validation procedures [14].

Even though current research indicates good performance in machine learning, big-data processing, and feature engineering to predict churn, there are still significant gaps. All the models are restricted to certain datasets and are not very generalized with little application of deep learning or real-time prediction. Combination of multi-source customer data remains unhealthy and the segmentation techniques are not often tested on their real retention influence. Other works focus more on the engagement measures than the reduction of the churn. Generally speaking, there is a need for more scalable, real-time, and data-rich churn prediction frameworks that can operate well in different consumer contexts.

TABLE I. COMPARATIVE ANALYSIS OF MACHINE LEARNING TECHNIQUES FOR CUSTOMER CHURN DETECTION IN E-COMMERCE PLATFORMS

Author	Methodology	Data	Key Findings	Limitation	Future Work
Ahmad, Jafar & Aljoumaa (2019)	ML (Decision Tree, RF, GBM, XGBoost); Big Data (Spark); Feature engineering; Social Network Analysis (SNA)	SyriaTel telecom customer data (9 months); Large transformed big data	XGBoost gave 93.3% AUC; SNA improved performance	Domain-specific SNA features may not generalize; No deep learning comparison	Test on other markets; add deep learning
Ebrah & Elnasir (2019)	Naive Bayes, SVM, Decision Trees	IBM (7033), Cell2Cell (71k)	AUC up to 0.99; outperformed prior studies	Focused only on traditional ML (no ensemble or DL);	Use advanced models, more features
Chauhan et al., (2019)	(CUARS); Collaborative filtering; Predictive analytics; Smart scheduling; Ranked factorization	E-commerce push-notification and user behaviour logs	30% CTR improvement; better personalization	Focused only on CTR improvement—not actual churn prediction; Limited evaluation metrics	Add churn prediction; deep recommenders
Rahaman (2018)	Clustering, Classification, Ensemble methods; Multi-source data integration (transaction + browsing + demographics)	E-commerce customer behavior and transactional datasets	Improved segmentation quality; Provided actionable insights for customer retention; Demonstrated importance of multi-source integration	Lack of quantitative performance results; Generalized discussion without algorithm-level comparisons	Add metrics (AUC, F1-score); Implement deep learning-based segmentation; Analyze temporal purchase patterns

Subramanya & Somani (2017)	Models: Regularized LR, SVM, Gradient Boosted Random Forests	Logs from Sales & Clickstream data; Big data environment	Improved accuracy; key churn features identified	Dataset limited; Feature selection not automated	Apply advanced feature engineering; Use gradient boosting and DL; Deploy real-time churn pipeline
----------------------------	--	--	--	--	---

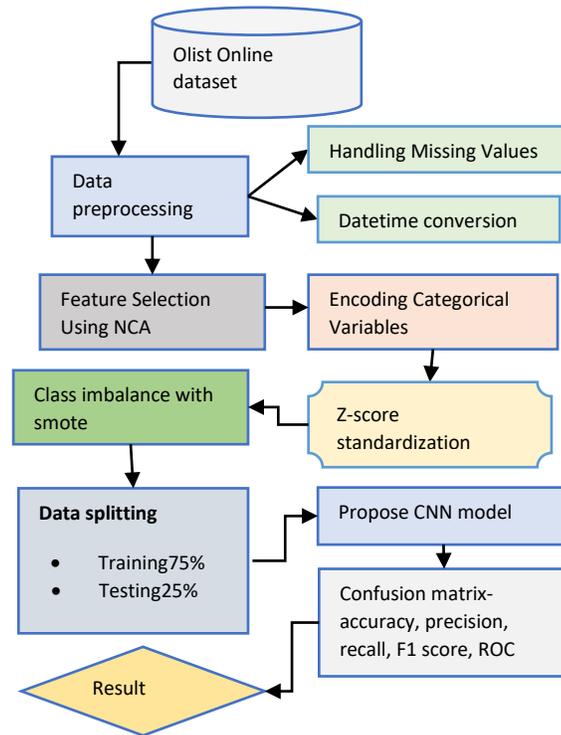


Fig. 1. Flowchart of the Customer Churn in E-Commerce Platforms

III. METHODOLOGY

The approach offers a foundation for developing an Olist e-commerce data-based customer churn prediction system. The data was highly processed to assurance data quality and model preparation, such as the elimination of the missing values, the conversion of the temporal features into the datetimes, feature subset selection under supervision, and the encoding of the categorical variables. Numerical variables have been scaled and the issue of class imbalance has been handled with the help of SMOTE to balance the churned and non-churned customers. The data was divided into training, validation and testing set (75: 25 ratio) and a CNN framework was suggested to forecast churn. The acc, prec, rec, F1score, and ROC analysis were utilized to determine the model's performance to ensure it was balanced and reliable. Figure 1 show the implementation steps.

The following sections provide each step description that also shows in methodology and proposed flowchart:

A. Data Collection

The Olist Online dataset, sourced from a Brazilian e-commerce platform, includes 111,468 records and 25 features covering transactions from 2016 to 2018. It contains order details, payments, reviews, and customer behaviour data, making it suitable for churn prediction tasks.

B. Data Preprocessing

The model experiment uses a wide range of feature selection approaches and pre-processing methods to aggressively pre-process the data. The many steps that make up the pre-processing pipeline are summarised in the following list.

- **Handling Missing Values:** To ensure data integrity and maintain dataset quality, rows with missing values were removed. This approach avoids the risk of introducing biases or errors that can arise from imputing or ignoring missing data during model training.

- **Datetime conversion:** Time-based insights were obtained by extracting granular variables like day, month, hour, minute, and second from the List Online dataset and converting pertinent date columns to datetime field types.

C. Feature Selection Using NCA

Neighbourhood Components Analysis (NCA), a supervised feature selection technique, was employed to reduce dimensionality. It projected the OriginalFeatureSpace into a new lower-dimensional space that better separates churned and non-churned classes. As a result, the Olist dataset was reduced to 12 most informative features as shows in Figure 2, improving model performance and interpretability.

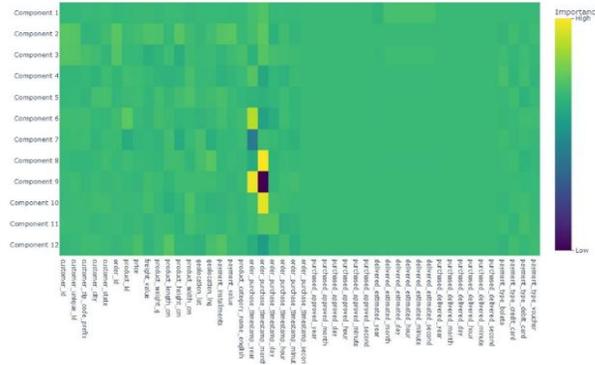


Fig. 2. Feature Importance Score

The heatmap of feature importance produced by a dimensionality reduction or feature interpretation technique is shown in Figure 2 using 12 components. All the components are latent variables based on the data set and the colours represent the relative value of the features, yellow implies great importance and dark purple means low relevance. The components are placed in the vertical axis, and the original names of the features in the horizontal axis. The heatmap displays that there is a limited number of features around the center that are much more important in several components. This suggests that the characteristics contribute most to the model's predictive power and are crucial in influencing the churn categorisation outcomes. It is possible to identify and emphasize such elements to improve model performance and interpretability.

D. Encoding Categorical Variables

Two encoding strategies were used to transform categorical features to a numerical format, which is conveniently used to input the model. High-cardinality categorical variables were frequency encoded to preserve their distributional value. With the rest of the categorical features, one-hot encoding was adopted to convert them to binary indicator variables without creating multicollinearity. A churned customer was any customer who did not make any purchase during a period of 90 days. This rule formed two groups 90,416 churned customers and 21,052 non-churned.

E. Standardization of Z-score

To ensure consistency across features and prevent dominance of any single feature, the dataset underwent Z-score standardization. Each feature was rescaled by subtracting its mean and dividing by its Standard Deviation. This process standardized all numerical attributes, improving the model's training stability and convergence. The Equation (1), for Z-score standardization.

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

- Z = Standardized Value
- X = Original Value
- μ= mean of the feature
- σ = Standard Deviation of the feature

This transformation ensures that the mean of the data is zero and the Standard Deviation is one.

F. Class Imbalance with Smote

The dataset exhibited a significant imbalance between churned and non-churned classes. To remedy this, SMOTE was used. SMOTE used interpolation to create synthetic instances for the Minority Class (non-churned). Figure 3 present the before and after applying smote technique.

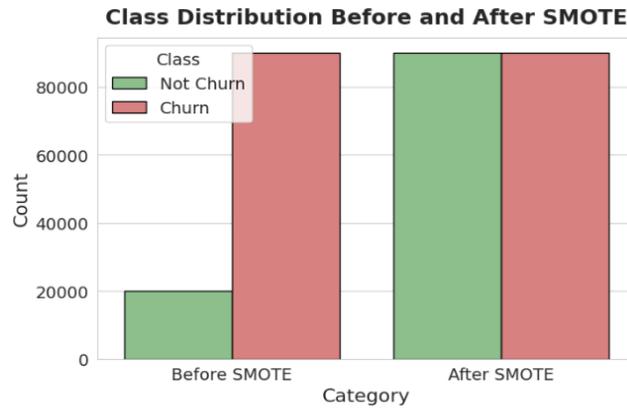


Fig. 3. Count Plot for Class Distribution

Figure 3 depicts the dataset's class distribution before and after implementing the SMOTE approach. First, the dataset is extremely one-sided and the class of Churn (red) is much overrepresented compared to the Not Churn (green) one, which may cause the biased model training. When SMOTE is applied, the dataset is balanced, each of the classes has an equal number of instances (approximately 100,000). This process of balancing assists in enhancing the performance of classification models because it learn equally in both classes, which helps them to reduce bias and detect the instances of the minority classes better during the prediction.

G. Data Splitting

A training set and a testing set were created from the Olist Online dataset using a 75:25 split. After standardisation, the training set was further separated into a basic training set and a validation set. This allows for successful model training, hyperparameter adjustment, and impartial assessment.

H. Propose Convolutional Neural Networks (CNN)

CNNs are a type of DL model created specifically for processing and analysing grid-like input, such as picture data. Their capacity to automatically uncover hierarchical features from raw pixel data has revolutionised picture categorisation and a range of other computer vision applications [15][16]. A CNN has 3 layers: a Convolutional Layer, a Pooling Layer, and a fully connected layer. Fig 2 displays all layers combined.

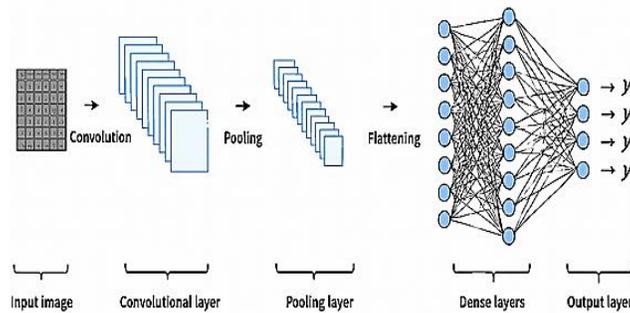


Fig. 4. Architecture of Propose CNN Model

Figure 4 shows a CNN architecture where the input image passes through convolution and pooling layers, is flattened, and then processed by dense layers to produce the final output prediction. The following CNN layers are:

- **Convolutional layer:** An activation map is generated by scanning the images few pixels at a time utilising a filter.
- **Pooling layer:** decreases the quantity of data generated by the convolutional layer and stores it more effectively.
- **Fully Connected Layer:** Fully linked input layer The output of the preceding layers is "flattened" and transformed into a singular vector; this is then utilized as input for the subsequent phase. The first fully connected layer adds weights to the feature analysis inputs in order to predict the appropriate label. Fully linked output layer - returns the probability for each label at the end.

1. Performance Matrix

The classification performance of our model was evaluated employing F1score, recall, validation acc, and prec, which are suitable measures for assessing models on a balanced dataset.

Accuracy: An assessment is made of the model's forecast accuracy rate. To get there, we use Equation (2) to divide the sum of all forecasts by the sum of all correct predictions.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{2}$$

Precision: Indicates the ratio of the number of positive predictions (TP and FP combined) to the number of accurately categorised positive samples, as determined by Equation (3):

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

Recall: It is the proportion of true positive forecasts to all actual positives, including both false negatives (FN) and true positives (TP). Equation (4) yields the recall formula.

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

F1_Score: integrates recall and precision to produce a fair metric that takes both into consideration. in Equation (5):

$$F1 - score = 2 * \frac{(precision*recall)}{(precision+recall)} \tag{5}$$

ROC: The ROC curve compares the TPR and FPR to assess the model's performance at different thresholds.

IV. RESULTS AND DISCUSSION

In this experimental setting, we utilised a robust computer environment to perform our research efficiently. The host machine ran Windows 10, with 8 GB of RAM and an Intel Core i7 CPU running at 2.50 GHz. This high-performance hardware arrangement offered the required processing capacity for studies. The Table II shows the performance indicators of CNN model in predicting customer churn. This model has an impressive accuracy of 99.62 meaning that the model is very successful in classifying churn and non-churn cases. The model supports accurate detection of real cases of churn with a high precision of 99.32 and a low false positive. The high recall of 98.66% demonstrates the high capability to capture almost all true churns whereas the F1score of 98.78% proves a positive balanced performance according to prec and rec. All these metrics prove that CNN model is very effective at identifying churn patterns in the e-commerce sphere.

TABLE II. PROPOSE MODEL PERFORMANCE ON OLIST ONLINE DATASET

Measures	CNN
Accuracy	99.62
Precision	99.32
Recall	98.66
F1-score	98.78

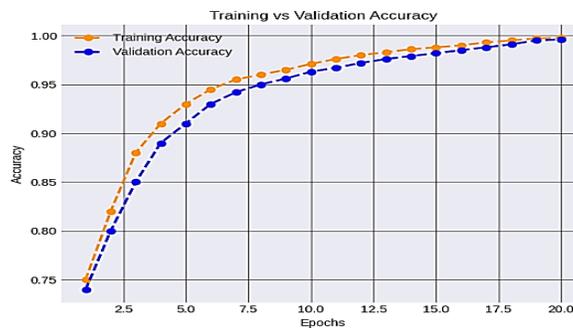


Fig. 5. Plot Accuracy curve of CNN Model

Figure 5 displays the accuracy curve of the CNN model after 20 training and validation epochs. The Blueline represents validation accuracy while the Orange Line represents training accuracy. At the beginning, the two curves grow at a fast rate, meaning that the model acquires knowledge fast. At approximately the 10th epoch, the two accuracies are greater than 95 and they have perfect learning performance. The model's ability to generalise and its resistance to overfitting are both shown by the

fact that the accuracy gap among training and validation is consistently small. The CNN model's efficacy and robustness are shown by the training and validation accuracy, which approaches 99 by the final epoch.



Fig. 6. Plot Loss curve of CNN Model

The CNN model's training and loss curves across 20 epochs are shown in Figure 6. Training loss is denoted by the red curve, whereas validation loss is denoted by the green curve. The declines in both curves are quite regular and steep, particularly during the early periods which mean successful learning and decrease of prediction error. The lack of overfitting and the model's ability to generalise to unknown data are both shown by the close proximity of the two curves. The loss values of the CNN model are also highly stable, and by the conclusion of training, the values are zero, which also supports high performance.

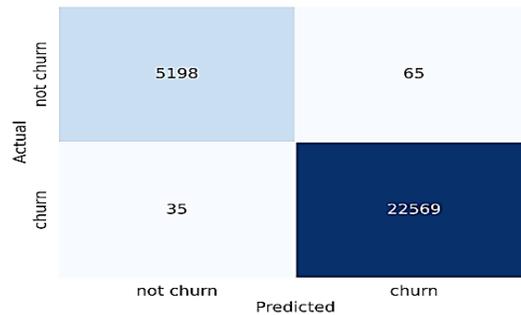


Fig. 7. Confusion matrix of CNN model

Figure 7 displays the confusion table of the CNNmodel. The model had a high predictive power such that it predicted 5,198 not churn cases and 22,569 churn cases correctly. It only falsely classified 65 cases that were not churn as churn (false positives) or 35 cases that were churn as not churn (false negatives). The very low misclassification and large values of TP and TN indicate that the model is sensitive and specific. On the whole, CNN model has proven to be very accurate and reliable in making discrimination between churn and non-churn customers.

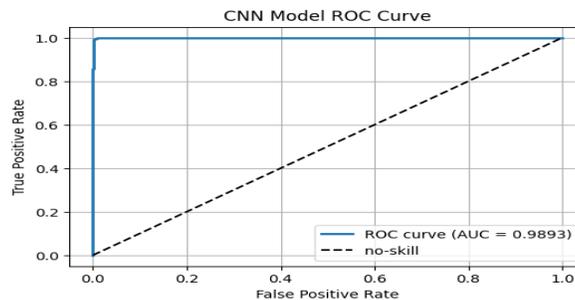


Fig. 8. AUC-ROC curve of CNN Model

The CNN classifier used to predict churn provides the AUC-ROC curve shown in Figure 8. The curve takes the TPR versus the FPR at different threshold levels. The blue line depicts the performance of the model and the no-skill classifier is just depicted by the diagonal dotted line. The CNN model has AUC 0.9893, which is almost equal to 1.0 showing excellent classification performance. As the large AUC demonstrates, the model does an excellent job of discriminating among the churn and non-churn classes with the highest possible accuracy and the fewest possible false positives, demonstrating its excellent generalisability.

TABLE III. COMPARISON BETWEEN BASE AND PROPOSE MODEL PERFORMANCE FOR CUSTOMER CHURN DETECTION

Matrix	CNN	Ada[17]	RF[18]	LR[19]	ANN-MLP[20]
Accuracy	99.62	-	65.26	95.15	68.08
Precision	99.32	83.9	46.42	-	79.31
Recall	98.66	83.5	74.07	-	74.19
F1-score	98.78	82.2	-	-	76.67

As the comparison Table III shows clearly, the proposed CNN model greatly surpasses all of the baseline models in all the key metrics of evaluation. CNN can provide a much more reliable and consistent churn prediction performance with an accuracy of 99.62, and always high values of prec, rec, and F1score. By contrast, the standard ML models like the RF (65.26% accuracy), or the Logistic Regression (95.15% accuracy) demonstrate poorer or inconsistent results, and ANN-MLP and AdaBoost present moderate results but still do not demonstrate the strength of CNN. All in all, the findings indicate that deep learning, in particular, CNN architecture, has a better feature extraction and prediction ability to detect churn than classical methods.

The suggested CNN model has explicit benefits and innovation of customer churn prediction. In contrast to the classical methods of machine learning, CNNs are superior at automatically learning more complex nonlinear features representations, which have better accuracy and generalization. Neighbourhood Components Analysis (NCA) combined with SMOTE balancing guarantees dimensionality reduction as well as fairness in classification, and thus, improves interpretability and reliability. The innovation is to run a deep learning architecture, which is typically applied in image recognition, on structured e-commerce data and it is claimed to have state of the art performance with little misclassification. This pre-processing, feature optimization, and hierarchical learning feature of CNN, creates a powerful and innovative framework of churn detection on large datasets.

V. CONCLUSION AND FUTURE SCOPE

"Customer churn," the process by which a company's clients cease doing business with it, is a major issue for online marketplaces. Businesses may identify their most at-risk clients using churn prediction and then utilise that information to develop retention tactics that will increase revenue and minimise churn.

The current research shows that CNNs can provide the state-of-the-art performance in customer churn prediction, making it more accurate, precise, and recalling, and with a higher F1-score than traditional machine learning models. with 99.62% accuracy, and 0.9893 AUC, which are significantly better than baseline models like Random Forest (65.26% accuracy), Logistic Regression (95.15% accuracy), AdaBoost (precision 83.9) These values support the higher capability of the CNN to record complicated patterns of customer behaviour and provide highly trustworthy churn prediction. Regardless of good performance, the model is over reliant on one dataset, which restricts the level of generalizability, and CNNs require large amounts of computational power and have low interpretability. Future research aims to enhance the robustness and utility of the model in e-commerce settings via multi-dataset validation, hybrid deep learning models, explainable AI integration, and real-time and economical churn prediction.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1]. A. Townsend and M. Professor, "Customer Churn: A Study of Factors Affecting Customer Churn using Machine Learning," *Creat. Compon.*, 2019.
- [2]. A. Dingli, V. Marmara, and N. S. Fournier, "Comparison of deep learning algorithms to predict customer churn within a local retail industry," *Int. J. Mach. Learn. Comput.*, 2017, doi: 10.18178/ijmlc.2017.7.5.634.
- [3]. G. Gupta and H. Aggarwal, "Improving Customer Relationship Management Using Data Mining," *Int. J. Mach. Learn. Comput.*, vol. 2, no. 6, 2012, doi: 10.7763/ijmlc.2012.v2.256.
- [4]. A. Dursun and M. Caber, "Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis," *Tour. Manag. Perspect.*, vol. 18, pp. 153–160, Apr. 2016, doi: 10.1016/j.tmp.2016.03.001.
- [5]. R. A. Daoud, A. Amine, B. Belaid, and R. L. R. Lbibb, "Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 4, p. 2367, Aug. 2018, doi: 10.11591/ijece.v8i4.pp2367-2383.
- [6]. L. Yang, "Predictions model of customer churn in E-commerce based on online sequential optimization extreme learning machine," *Nanjing Li Gong Daxue Xuebao/Journal Nanjing Univ. Sci. Technol.*, 2019, doi: 10.14177/j.cnki.32-1397n.2019.43.01.015.

- [7]. V. M. L. G. Nerella, "Automated cross-platform database migration and high availability implementation," *Turkish J. Comput. Math. Educ.*, vol. 9, no. 2, pp. 823–835, 2018.
- [8]. J. Cao, X. Yu, and Z. Zhang, "Integrating OWA and data mining for analyzing customers churn in E-commerce," *J. Syst. Sci. Complex.*, vol. 28, no. 2, pp. 381–392, Apr. 2015, doi: 10.1007/s11424-015-3268-0.
- [9]. K. G. M. Karvana, S. Yazid, A. Syalim, and P. Mursanto, "Customer Churn Analysis and Prediction Using Data Mining Models in Banking Industry," in *2019 International Workshop on Big Data and Information Security, IWBIS 2019*, 2019. doi: 10.1109/IWBIS.2019.8935884.
- [10]. A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0191-6.
- [11]. K. Ebrah and S. Elnasir, "Churn Prediction Using Machine Learning and Recommendations Plans for Telecoms," *J. Comput. Commun.*, vol. 07, no. 11, pp. 33–53, 2019, doi: 10.4236/jcc.2019.711003.
- [12]. G. Chauhan, D. V. Mishra, M. Farida Begam, and Akhilaa, "Customer-Aware Recommender System for Push Notifications in an e-commerce Environment," in *2019 Global Conference for Advancement in Technology, GCAT 2019*, 2019. doi: 10.1109/GCAT47503.2019.8978330.
- [13]. S. U. Rahaman, "Optimizing Predictive Models for Customer Segmentation in E-commerce: A Data Science Approach," *Int. J. Innov. Res. Creat. Technol.*, vol. 4, no. 4, pp. 1–11, 2018.
- [14]. K. B. Subramanya and A. K. Somani, "Enhanced feature mining and classifier models to predict customer churn for an e-retailer," in *Big Data Analytics: Tools and Technology for Effective Planning*, 2017. doi: 10.1201/b21822.
- [15]. S. Achouche and U. Yalamanchi, "Method, apparatus, and computer-readable medium for performing a data exchange on a data exchange framework," 2019
- [16]. M. Saghir, Z. Bibi, S. Bashir, and F. H. Khan, "Churn Prediction using Neural Network based Individual and Ensemble Models," in *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, IEEE, Jan. 2019, pp. 634–639. doi: 10.1109/IBCAST.2019.8667113.
- [17]. I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," *IEEE Access*, vol. 7, pp. 60134–60149, 2019, doi: 10.1109/ACCESS.2019.2914999.
- [18]. A. P. Patil, M. P. Deepshika, S. Mittal, S. Shetty, S. S. Hiremath, and Y. E. Patil, "Customer churn prediction for retail business," *2017 Int. Conf. Energy, Commun. Data Anal. Soft Comput. ICECDS 2017*, pp. 845–851, 2018, doi: 10.1109/ICECDS.2017.8389557.
- [19]. P. P. Shetty, V. C. M, V. D. Vadone, S. Sarode, and P. K. D, "Customers churn prediction with rfm model and building a recommendation system using semi-supervised learning in retail sector," *Int. J. Recent Technol. Eng.*, vol. 8, no. 1, 2019.
- [20]. S. Khodabandehlou and M. Zivari Rahman, "Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior," *J. Syst. Inf. Technol.*, 2017, doi: 10.1108/JSIT-10-2016-0061.
- [21]. Gupta, A. K., Polu, A. R., Narra, B., Buddula, D. V. K. R., Patchipulusu, H. H. S., & Vattikonda, N. (2024). Leveraging deep learning models for intrusion detection systems for secure networks. *Journal of Computer Science and Technology Studies*, 6(2), 199-208.
- [22]. Narra, B., Buddula, D. V. K. R., Patchipulusu, H., Vattikonda, N., Gupta, A., & Polu, A. R. (2024). The integration of artificial intelligence in software development: Trends, tools, and future prospects. Available at SSRN 5596472.
- [23]. Achuthananda, R. P., Bhumeka, N., Dheeraj Varun Kumar, R. B., Hari Hara, S. P., & Navya, V. (2024). Evaluating machine learning approaches for personalized movie recommendations: A comprehensive analysis. *J Contemp Edu Theo Artific Intel: JCETAI-115*.
- [24]. Waditwar, P. (2024) The Intersection of Strategic Sourcing and Artificial Intelligence: A Paradigm Shift for Modern Organizations. *Open Journal of Business and Management*, 12, 4073-4085. doi: 10.4236/ojbm.2024.126204.
- [25]. Bitkuri, V., Kendyala, R., Kurma, J., Mamidala, J. V., Attipalli, A., & Enokkaren, S. J. (2024). A Survey on Blockchain-Enabled ERP Systems for Secure Supply Chain Processes and Cloud Integration. *International Journal of Technology, Management and Humanities*, 10(04), 126-135.
- [26]. Mamidala, J. V., Bitkuri, V., Attipalli, A., Kendyala, R., Kurma, J., & Enokkaren, S. J. (2024). Machine Learning Approaches to Salary Prediction in Human Resource Payroll Systems. *Journal of Computer Science and Technology Studies*, 6(5), 341-349.
- [27]. Waditwar, P. (2024) AI for Bathsheba Syndrome: Ethical Implications and Preventative Strategies. *Open Journal of Leadership*, 13, 321-341. doi: 10.4236/ojl.2024.133020
- [28]. Attipalli, A., Kendyala, R., Kurma, J., Mamidala, J. V., Bitkuri, V., & Enokkaren, S. J. (2024). Privacy Preservation in the Cloud: A Comprehensive Review of Encryption and Anonymization Methods. *International Journal of Multidisciplinary on Science and Management IJMSM*, 1(1).
- [29]. Tamilmani, V., Maniar, V., Singh, A. A., Kothamaram, R. R., Rajendran, D., & Namburi, V. D. (2024). A Review of Cyber Threat Detection in Software-Defined and Virtualized Networking Infrastructures. *International Journal of Technology, Management and Humanities*, 10(04), 136-146.

- [30]. Singh, A. A. S., Kothamaram, R. R., Rajendran, D., Deepak, V., Namburi, V. T., & Maniar, V. (2024). A Review on Model-Driven Development with a Focus on Microsoft PowerApps. *International Journal of Humanities, Science Innovations and Management Studies*, 1(1), 43-56.
- [31]. Padur, S. K. R. (2024). AI-augmented platform engineering: Redefining developer experience through autonomous, self-optimizing enterprise systems. *International Journal of Science, Engineering and Technology*.
- [32]. Gangineni, V. N., Tyagadurgam, M. S. V., Pabbineedi, S., Penmetsa, M., Bhumireddy, J. R., & Chalasani, R. (2024). AI-Powered Cybersecurity Risk Scoring for Financial Institutions Using Machine Learning Techniques (Approved by ICITET 2024). *Journal of Artificial Intelligence & Cloud Computing*.
- [33]. S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," 2024 9th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2024, pp. 1149-1154, doi: 10.1109/ICCES63552.2024.10859381.
- [34]. S. R. Sagili and T. B. Kinsman, "Drive Dash: Vehicle Crash Insights Reporting System," 2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA), Pune, India, 2024, pp. 1-6, doi: 10.1109/ICISAA62385.2024.10828724.
- [35]. Padur, S. K. R. (2024). Securing Oracle Integration Cloud ERP ecosystems, zero trust architecture, data governance, and compliance automation. *International Journal of Science, Engineering and Technology*, 12(4), 10-5281.
- [36]. S. R. Sagili, S. Chidambaranathan, N. Nallametti, H. M. Bodele, L. Raja and P. G. Gayathri, "NeuroPCA: Enhancing Alzheimer's disorder Disease Detection through Optimized Feature Reduction and Machine Learning," 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 2024, pp. 1-9, doi: 10.1109/ICEEICT61591.2024.10718628.
- [37]. S. R. Sagili, V. K. B. Puli, P. Sundaramoorthy, M. R and K. N V, "Advancing Cervical Cancer Identification using Generative-based Adversarial Networks: An Integrative Learning Methodology," 2025 6th International Conference for Emerging Technology (INCET), BELGAUM, India, 2025, pp. 1-5, doi: 10.1109/INCET64471.2025.11140170.
- [38]. Routhu, K. K. (2024). Beyond Automation: AI-Powered Employee Engagement Journeys in Oracle HCM Cloud. *KOS Journal of AIML, Data Science, and Robotics*, 1(1), 1-6.
- [39]. Routhu, K. K. (2024). The future of HCM: Evaluating Oracle's and SAP's AI-powered solutions for workforce strategy. *Journal of Artificial Intelligence, Machine Learning & Data Science*, 2(2), 2942-2947.
- [40]. Sannapureddy, R., Nadella, V. M., & Nelavelli, S. (2024). Edge-Cloud Continuums for Latency-Sensitive Tasks. *International Journal of AI, BigData, Computational and Management Studies*, 5(4), 189-201.
- [41]. Arigela, A. K., Brahmareddy, A., Sreenivas, T. S., Selvan, M. P., Venu, N., & Lal, D. K. (2024, December). Optimizing Energy Efficiency and Latency in IoT Devices Through AI-Based Adaptive Protocols in Fog-Edge Computing Environments. In *Congress on Smart Computing Technologies* (pp. 595-607). Singapore: Springer Nature Singapore.
- [42]. Nadella, V. M. (2024). AI-Native 6G Network Management. *American International Journal of Computer Science and Technology*, 6(1), 23-37