A NOVEL FEATURE SELECTION STRATEGY BASED PREDICTIVE MODELING FOR CUSTOMER CHURN PREDICTION IN TELECOM INDUSTRY USING SWISH CNN

[1]BISWA RANJAN AGASTI, [2]Dr. SUSANTA KUMAR SATPATHY

[1]Ph.D Scholar, Department of CSE, VFSTR University, Guntur, AP, India.

[2]Associate Professor, Department of CSE, VFSTR University, Guntur, AP, India.

ABSTRACT:Customer Churn Prediction (CCP) is a crucial procedure for maintaining clients, according to a number of businesses. Customer retention is a significant difficulty in many business sectors. Churn analysis has been used for a long time to improve customer-company relationships and profitability. Since there are more providers of telecommunications services, CCP in the telecommunications industry has become an essential requirement. Large businesses face a significant problem with customer churn because it is much more important to retain existing customers than to acquire new ones.Businesses will be able to keep their current customers and acquire new customers by following a predictable customer model. The selection of customer attributes (feature selection) from the dataset for the model's construction is crucial to the model's efficiency. Effective CCP models have only recently started to be developed using Deep Learning (DL) and Machine Learning (ML) models.A novel feature selection strategy based predictive modeling for customer churn prediction in telecom industry using swish CNN (Convolutional Neural Network) is presented in this analysis.Finally, a modified version of the S-CNN (Swish Convolutional Neural Network) is used to distinguish between a CC (Churn Customer) and a normal customer. If there is a churn, the customer retention process looks at the network utilization history.This approach's performance is evaluated in terms of Precision, Accuracy, F1-score, and Sensitivity.

KEYWORDS: Customer Churn Prediction, Telecom Industry, Feature Selection, Convolutional Neural Network.

I. INTRODUCTION

Once clients using a company's services, that is known as customer churn. The banking industry, the insurance industry, the video gaming industry, and the telecommunications industry are all directly impacted by this issue. Retaining customers is an important factor in predicting churn, and telecom companiesare closely linked to financial institutions [4]. The telecommunications industries have gone through a lot of change in recent years, including new services, advancements in technology, and more competition. Telecom industry is continually developing and enhancing. Targeted advertising strategies are becoming more important for managing client churn as a result of the ever-increasing competition between companies. Today, customers expect the best services at reasonable prices. They quickly switch to another telecom network if they are dissatisfied.

In order to maintain user trustworthiness and attract new customers, telecommunications companies have utilized a wide range of strategies, including the creation of a collection of services. The provision of a number of ideas to meet user needs, and the provision of discounts. All of these strategies have allowed for gradual growth in profit. Ironically, these strategies make several businesses more competitive, and it's hard to predict user behavior.It would rather provide a modified service in order to retain customers, prevent customer churn, and maintain increased competition. The term "normal churn" describes to the loss of customers who frequently switch suppliers within a short time frame. Certainly, customer retention is done to reduce on customer churn,it is a major issue in the telecommunications industry [5].

The term "churn" refers to a customer who switches telephone service providers. A significant problem for the telecom industry is the amount of customers who will end their existing relationship with a company or network in the coming years. This issue may have an effect on revenues as well as the rapid growth of the business. As a result, the telecommunications industry's actors must be able to predict customer churn in order to protected their loyal clients [1]. To perform better in such a competitive market, businesses must develop innovative techniques to predict potential customer churn.

Hence, several Customer Churn Prediction (CCP) models have been utilized, but they don't produce the outcomes that are predicted, this is because there are still unknown variant variables that could influence Customer Churn (CC). Many factors can cause a customer to finished: The company doesn't discuss to its customers, they don't respond to their complaints quickly, negative comments about the provided services, software that doesn't address requirements, a competitor's new product of better quality and pricing, or a network that doesn't allow advanced network types (4G, 5G, 6G) are all common on social media, the press, etc.The majority of the time, the business is unaware of the causes of the customer's leaving. To prevent client churn in this situation, data analysis can help understand and predict churn [2].

Churn prediction techniques heavily based on classification algorithms developed by Artificial Intelligence (AI). Accurate churn prediction is impacted by dataset imbalance and high dimensionalityis presented by these classification techniques. The efficacy of certain data mining techniques is significantly influenced by the amount of data that can be analyzed and whether the dataset is balanced or unbalanced [3].

Using advanced machine learning methods, a number of experiments have been carried out to better evaluate churning affects the customer network. There are numerous approaches for predicting user churn. Many experiments have been carried out utilizing advanced machine learning approaches to better assess the impact of customer churn on consumer networks. It implementingmachine learning and data mining techniques like logical regression and decision tree, support vector machines, neural networks, and etc, to develop a prediction model of customer loss in telecommunications companies. To determine the factors that lead to the loss of customers and users. Although a number of algorithms have been developed, it is unclear which model is best suited for detecting churning customers.

Consequently, deep learning is currently the most used AI method for various uses. Also, currently it is the retail client churn model that is recommended., music streaming, mobile gaming, and telecommunications industries [7]. Convolutional neural networks, are specialized range of neural networks that are adaptable to a wide range of data types with varying dimensions. When compared to its predecessors, CNN has the primary advantage of automatically identifying the crucial features without the assistance of an expert. CNN becomes more attractive when a churn prediction model with numerous data features is used.

As a result, a novel Swish-Convolutional Neural Network-based feature selection strategy based predictive modeling for telecom industry customer churn prediction is presented in this analysis. The following is the order of the remaining analysis: Section II discusses the literature review. In this section III demonstrates the novel feature selection strategy based predictive modelling for customer churn prediction in telecom industry using S-CNN.Analyses of the results are evaluated in section IV. In this Section V provides the analysis of conclusion.

II. LITERATURE SURVEY

ShreyasRajesh Labhsetwaret. al. [6] describes the use of supervised learning in the telecom industry's predictive analysis of customer churn. This study attempts to define potential churning consumers based on their usage patterns, identify them using Machine Learning (ML) techniques, and then present the analysis conclusions. Low false negatives and average AUC scores of 0.843, 0.787, and 0.735, the results show that Extra Trees Classifier, XGBoosting Algorithm, and Support Vector Machine have the best churn modelling performance. According to the report, ML algorithms can help in the development of client retention programmes by properly forecasting possible customer churn.

SajjadShumaly, PedramNeysaryan, YanhuiGuoet. al. [9] Utilizing Sampling Methods, Bagging, and Boosting Trees to Handle Class Imbalance in Customer Churn Prediction in the Telecom Industry.Customers who want to stop using the company's services have been identified. The imbalanced data is one of the most significant issues in predicting customer churn, and various solutions have been tested. Multi-layer perceptron, decision tree, support vector machine, random forest, and gradient boosting are the machine learning methods used in this research. SMOTE (Synthetic Minority Oversampling Technique) techniques, random over-sampling, and random under-sampling were used to balance the data. It is clear from the information that the methods of under-sampling and over-sampling both demonstrated better specificity and sensitivity.

SanketAgrawal, AmitGaikwad, Aditya Das, SudhirDhage et.al. [12] discusses Deep Learning-Based Consumer Churn Prediction Modeling Based on Behavioural Patterns Analysis. Using data from a telecom company, a deep learning method is used in this study to predict churn.A non-linear classification model was created using a multilayered neural network. Using customer features, support features, usage features and contextual features, the churn prediction model operates. Both the possibility of churn and the deciding elements are predicted. The final weights are then applied to these features by the trained model, this additionally predicts the probability that a consumer will leave. The accuracy rate was 80.03%.

R. Prashant, K. Deepak, and Amit Kumar Meheret. al. [15] The telecom sector is defined as using high accuracy predictive modelling for customer churn prediction.Churn prediction was done with statistical and data mining methods in this study. For prediction, they utilize Random Forest, linear (logistic regression), and non-linear (deep neural network, deep belief network, and recurrent neural network) architectures of deep learning. This is the first time that a comparison of deep learning approaches and conventional machine learning methods has been done for churn estimation. Non-linear models were shown to be the most successful. The telecom sector could benefit from utilizing such predictive models for improved decision-making and customer management.

V. Umayaparvathi, K. Iyakuttiet. al. [16] provides an explanation for automated feature selection and churn predictions using deep learning models.The corresponding churn prediction model is utilising three deep neural network architectures and two telecom datasets. CrowdAnalytix and Cell2Cell, two real-world datasets, were used in the experiments. Without utilizing the hand-picked features, experimental results demonstrate that deep-learning based models are performing as well as traditional classification models.

Mohammad Ridwan Ismail, M Nordin A Rahman, Mohd Khalid Awang, and MokhairiMakhtaret. al., [20] A Multi-Layer Perceptron Method for Customer Churn Prediction discusses the method. In one of the top telecommunications companies in Malaysia, in order to predict customer churn, the research suggests employing a Multilayer perceptron (MLP) neural network method.The outcomes are examined with those acquired using the most popular churn prediction techniques, such as logistic regression analysis and multiple regression analysis.The outcomesshows that neural networks outperformed statistical models in prediction tasks (91.28% prediction accuracy). Overall, the results point to a neural network learning algorithm as a potential substitute for statistical predictive methods for predicting customer churn.

III. NOVEL FEATURE SELECTION STRATEGY BASED PREDICTIVE MODELLING

The process of predicting customer churn includes finding out whether or not a customer will change telecom networks. Customer churn happens when customers leave a telecom firm because they are dissatisfied with the support they are receiving.Clients begin to switch to other service providers as a result, which leads to service migration. The churned consumer must be identified the moment is practical and their expectations must be met in order to prevent. The Swish-Convolutional Neural Network is used in this part to propose an unique feature selection strategy-based predictive modelling for customer churn prediction in the telecom industry.The block diagram presented approach is shown in Fig. 1.

With the increase of telecom customer churns, this analysis makes use of the IBM (International Business Machines Corporation) dataset, which represents a fictional telecom firm that provided home phone and Internet services to 7043 customers in California. This dataset includes user information for 7043 clients, adds to a final classification of "Churned" or "Not Churned" for each customer.

This dataset gives basic demographic information about the customers as well as information on their subscription, both of which are useful for training the base model. The following key properties may be derived from this datasets:

i) Demographic Data: Customer ID, gender, and whether or not they have dependents and partners.

ii) The customer's subscriptions, it also provides phone service, internet security, online backup, multiple lines, device security, tech support, and on-demand TV and movie streaming.

iii) User billing profile: Method of payment, contract, paperless billing, monthly charges, and total charges

iv) The duration that the client was a customer of the business (Tenure) Each customer's information also includes a churn or non-churn classification, which, in this case, refers to employees who left the company within the past month.

Data pre-processing, which is a key stage in information discovery activities, is very important. It requires a number of stages, including data transformation and data reduction. The effectiveness and accuracy of learning algorithms will be harmed if raw information is transformed into poor-quality data. Consequently, the acquired data can be examined accurately by selecting the appropriate learning algorithms and performing the appropriate data preprocessing steps.

Telecom datasets include a number of problems that need to be solved, including non-numeric features, missing values, inconsistent feature scales, etc. Consequently, it is crucial to pre-process the data before applying a learning model. Processing of the acquired raw data is required before any algorithm can be applied. To do this, a number of different methods are used to clean the data and prepare it for the CNN. First, columns that are totally irrelevant to the model's training are eliminated. Columns like Client ID, which is particular to each data entry, were included. To eliminate any ambiguity or potential sources of error, incomplete entries are also eliminated from the dataset.

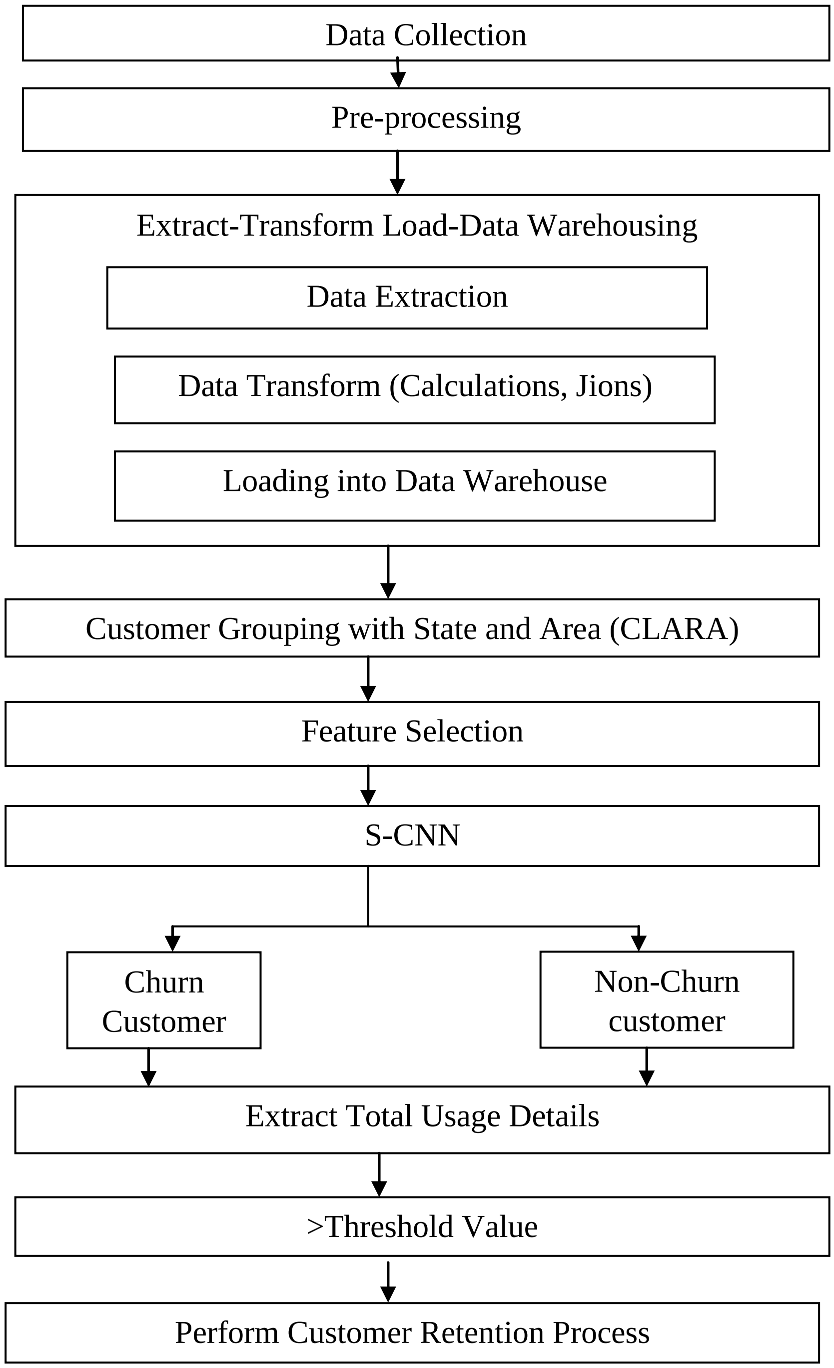

Fig. 1: Block Diagram of Presented novel feature selection strategy based predictive modelling

In this research, the data set (from IBM Telco) is pre-processed through various stages, including handling missing data and encoding categorical variables are using two techniques: "label encoding" and "one-hot encoding." The object labels are converted to numerical labels using label encoding, which are always between 0 and n classes-1. This is applied to several columns where the entries are classified as either yes or no responses. Categorical parameters are represented as binary vectors through one-hot encoding. As a result, the indexed columns are given a value of 1, while all other indices are set to 0. The data can be more expressive due to this technique. Since binary categorization provides better training over a wider numerical range, this is utilized more frequently in Deep Learning models. In addition, the normalization method is used to scale the high variance values, and any features that do not affect the dependent variable are removed.

Data marts are created by transferring data using the Extract Transform Load (ETL) process from one database to another. The act of extracting involves removing data files from databases and storing them in a big data platform. By specifying parameters through rules, lookup tables, or merging with other data, data can be transformed from its original state of unstructured raw data into structured data during the transformation process. Attempting to put data into the target database so that it may be accessed for additional analysis and the creation of prediction models is known as loading.

After the distinctive characteristic has been identified, the client records are grouped. The majority of the time, client records for telecom services contain information on customers from other states and countries. Analyzing worldwide consumer records is an extremely difficult task. Therefore, the client records from the respective states and regions are combined and placed into a cluster in order to reduce these responsibilities.

It is managed by the Clustering Large Applications (CLARA) technique. The CLARA is essentially a PAM (Partitioning AroundMedoids) clustering algorithm extension. For larger datasets, CLARA is a better technique because it can reduce computation time as well as the memory allocation problem.

The steps of the CLARA are explained as follows: Step 1:out of all the input data *D(i)*, the CLARA randomly selects a smaller portion (40 + 2k, where k represents the total number of clusters, or medoids), and then applies the PAM to that subset. Step 2: Using that subset, the PAM randomly selects k medoids and assigns them to the primary set of medoids M. Step 3: The nearest medoid is then connected with each information test I using the Euclidean distance for the entire dataset D(i), $\delta(i) \in D(i)$, i= 1, 2, 3, ..., n, where $\delta(i)$ is a non-selected object as of $D(i)(\delta(i) \neq M)$."

$$E_{dis} = \sqrt{(\delta(i) - M)^2 + (}\delta(i) - M)^2 \quad (1)$$

Step 4: The mean values of the dissimilarities (explicitly, the minimal Euclidean distance function as a gauge of dissimilarities) are calculated for each pair of the data sample and its equivalent medoid. A fundamental Cost Function (CF) is established from the obtained value. Step 5: The following iteration again randomly selects the newer subset of 40 + 2k data points. The PAM is used to obtain a more recent collection of medoids. The CF is determined for the entire dataset as well as the more recent group of medoids. The relationship, the following formula is used to determine the CF:

$$Cost\ (M, D(i) = \frac{\sum_{i=1}^{n} \widetilde{s_{dis}}\,(\delta(i), rep\ (\delta(i), M}{N} \quad (2)$$

whererep(δ(i),M) is δ(i) signifies the sample that corresponds to the cluster specified by M and N denotes the total amount of data.

Step 6:If this CF is lower than the existing CF, it is fixed as a newer variation on the present CF. The collection of medoids gets updated as a result. PAM then performs its subsequent run. Step 7: Then, the minimum CF and the equivalent medoids are achieved. Therefore, *M* set of medoids achieves the best clusterization. As a result, the generated cluster Ch^* is represented as,

$$ch^* = \{ch_1, ch_2, ch_3, \ldots . ch_n \ (3)$$

The feature selection phase effectively eliminates undesirable noisy features chooses the most important and helpful features for churn prediction. This can improve the accuracy, memory, and computation time of the model.Therefore, the Feature selection procedure considerably improves the customer churn prediction performance. Some of the most well-liked methods for selecting embedded features are called lasso regression.

When the data contains a number of inconsequential features, "Lasso Regression" is superior on the reducing variance since it has the capacity to completely remove the coefficient of an irrelevant variable from the data. "sum of the magnitude of coefficients absolute values " is the solution for the lasso in mathematical logic: = square residual sum+λ*

$$lasso{:} \sum_{i=1}^{n} (y_i - \sum_{j} x_{ij}\,\beta_j)^2 + \lambda \sum\nolimits_{j=1}^{p} |\beta_j\ | \ (4)$$

In that case, a shrinking amount λ. It λ = 0 denotes the equivalent of a linear regression, in which a prediction model is only constructed using the squared residual sum.λ = ∞ no attribute is considered, i.e., (λ as close to infinity). The variance increases with a decrease in λ, and the bias rises with an increase λ.

Subsequently, the Swish - Convolutional Neural Network algorithm receives the selected features as input.Swish is an activation function,

$$f(x) = x.sigmoid(\beta x)(5)$$

When the parameter is learnable.This activation function is x σ (x) (“Swish-1”) if the learnable parameter is not used in approximately all implementations. Swish activation is used in the neural networks, which are used extensively in sequence prediction and likelihood problems.If the parameters are set to the proper values, a convolutional neural network can carry out the relevant classification task with high accuracy. One significant advantage of CNN, in addition to the translation invariance property, it can benefit from shared weights and learn features in a hierarchical manner, reducing the number of required parameters without compromising performance.Once the ConvNet (Convolutional Network) has retrieved the feature, fully linked layers are utilized to categorise the input image to the target labels. The classified images refer to churner/non churner.

According to the S-CNN classifier, the "2" form of the last outcome is Customer Churn and Non-CC (Customer Churn). Someone prepared to transfer to another telecommunications network is referred to as a churn customer. Non-churn customers: Those who aremoving with the same telecommunications system.

In the case,once a CC result has been achieved, the network usage history of the particular client is examined. The corresponding threshold values are set at a fixed point in relation to their network consumption. The process of keeping a customer is carried out if their network consumption continues to be high, or if it exceeds the threshold value. The term "customer retention process" refers to the practice of keeping existing customers and existing users of a network on the same network by making them a few attractive offers and preventing them from transferring to any other telecommunication networks. On the other hand, if a client only uses a small part of the network, or if their network consumption is below a certain level, the customers are ignored.

IV. RESULT ANALYSIS

A novel feature selection strategy based predictive modelling for customer churn prediction in telecom industry using Swish- Convolutional Neural Network is implemented in this analysis.The accuracy, precision, sensitivity, and F1-score performance of the presented technique are evaluated using the confusion matrix parameters TP, TN, FP, and FN, which are defined as follows:

False Positive (FP) stands for incorrectly classified positive cases; False Negative (FN) stands for incorrectly classified negative cases; and finally, True Negative (TN) stands for incorrectly classified negative cases. The correct classification of positive cases is represented by the letters TP (true positive), FP (false positive), and FN (false negative). The following is a definition of the metrics that were derived from the confusion matrix:
Accuracy:It is defined as the proportion of correctly predicted outcomes. Overall number of predictions is represented as

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (6)$$

Sensitivity:It is the metric used to evaluate a model's capacity to predict the true positives for each class of data that is available. It is also known as True Positive Rate (TPR) or Recall.

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \quad (7)$$

Precision:Positive predictions' accuracy is measured by a metric called precision. It is determined by dividing the total number of correct predictions by the total number of false positive predictions and correct predictions.

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (8)$$

F1-score:Precision and Sensitivity are combined to get the F1 score, which is represented as

$$F1 - score = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity} \times 100 \quad (9)$$

The higher the precision and sensitivity, the higher the F1-score.The table 1 represents the performance metrics evaluation.

Table 1: Performance Metrics Evaluation

Performance Metrics	SVM	Presented Swish CNN
Precision (%)	79	92
Sensitivity (%)	85	94.21
F1-Score (%)	82	93.12
Accuracy	82	95.34

For the purpose of predicting customer churn, the performance of the presented approach is compared with that of the SVM algorithm in terms of Precision, Accuracy, F1-score andSensitivity. Swish CNN has higher precision, recall, f1-score, and sensitivity when compared to SVM algorithm. The Fig. 2 shows the comparative graph for precision.

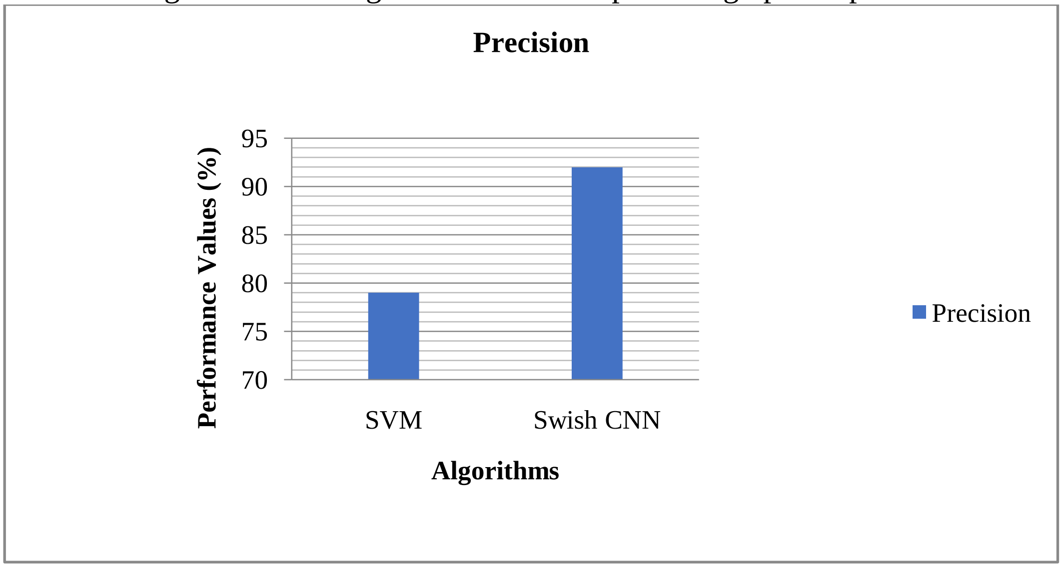

Fig. 2: Comparative Graph for Precision

In Fig. 2, the y-axis represents performance results expressed as percentages, and the x-axis shows various algorithms. Presented swish CNN has high precision than SVM algorithm. The Fig. 3 shows the sensitivity metrics comparison.

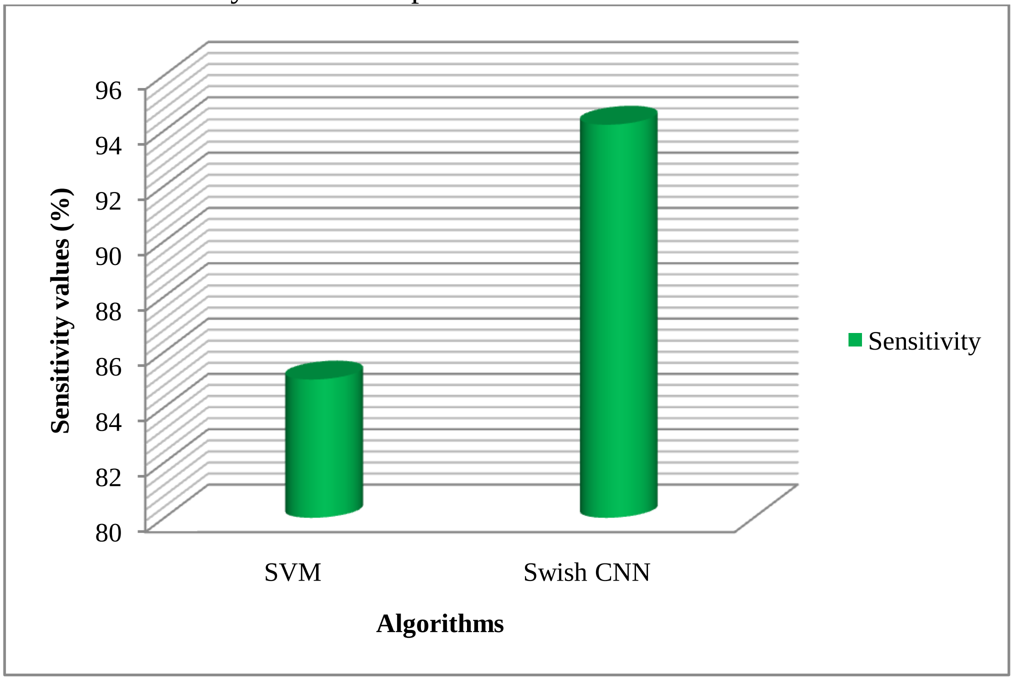

Fig. 3: Sensitivity Comparison

Compared to SVM algorithm, Presented swish CNN has high sensitivity. Also, the CLARA method utilised in this technique has its clustering time measured and compared with a well-known clustering algorithm, as shown in Fig. 4.

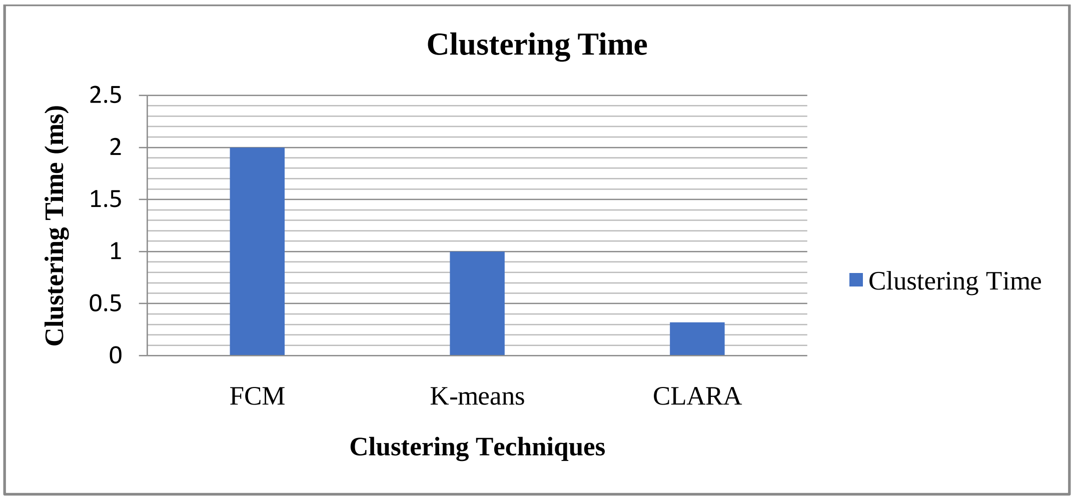

Fig. 4: Clustering Time Comparison

Compared to most popular FCM (Fuzzy c-means Clustering) and K-means clustering algorithms, CLARA algorithm has required less time for clustering. The Fig. 5 shows the performance comparison in terms of F1-score and Accuracy.

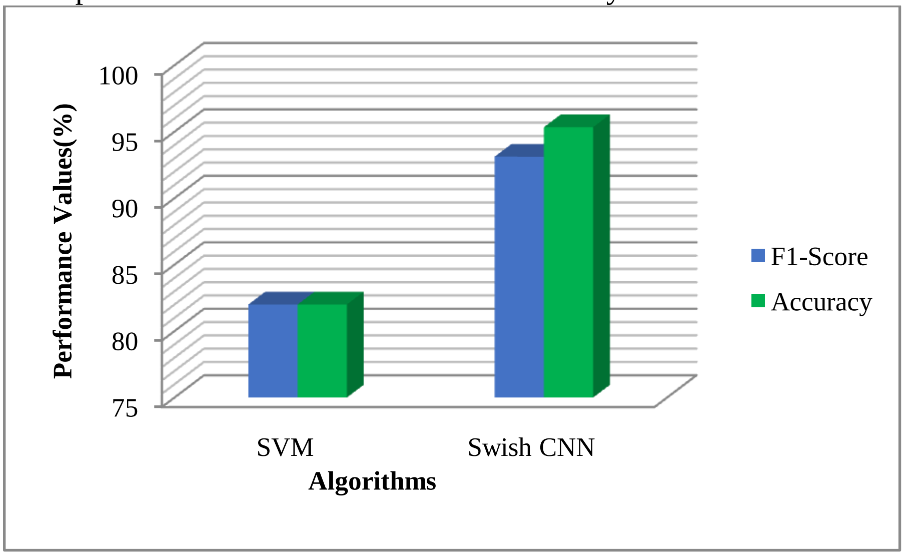

Fig. 5: Performance Comparison

Presented swish CNN has high accuracy and F1-score than SVM algorithm. Hence this approach has accurately predicted the customer as churn or non-churn. In addition customer retention is performed which plays a vital role in telecom industry.

V. CONCLUSION

A novel feature selection strategy based predictive modelling for customer churn prediction in telecom industry using Swish- Convolutional Neural Network is presented in this analysis. In this analysis, Telco customer churn IBM dataset is used which contains7043 customers and 33 attributes. Before the information involves a number of processes, such as dealing with missing data and encoding categorical features using two different strategies: label encoding and "one-hot encoding." Data marts can be developed using the Extract Transform Load (ETL) method for moving data between databases. CLARA algorithm is used to lessen the computation time. Lasso regression technique is used for feature selection. Swish-CNN is used to classify the customer as churn or non-churn. The performance of presented approach is measured in terms of Precision, Sensitivity, F1-score and Accuracy. In addition clustering time is also measured. Compared to popular clustering algorithms, CLARA requires less time for clustering. Compared to popular ML Algorithm, presented swish CNN has high accuracy, precision, sensitivity and F1-score.So, this approach will be a preferable alternative for customer retention and churn prediction in the telecom sector.

VI. REFERENCES

[1]AnouarDalli Impact of Hyper parameters on Deep Learning Model for Customer Churn Prediction in Telecommunication Sector",Hindawi Mathematical Problems in Engineering, Volume 2022, Article ID 4720539, 11 pages, doi:10.1155/2022/4720539

[2]SamahWaelFujo, Suresh Subramanianand Moaiad Ahmad Khder, "Customer Churn Prediction in Telecommunication Industry Using Deep Learning", Inf. Sci. Lett. 11, No. 1, 185 – 198 (2022), doi:10.18576/isl/110120

[3]Seyed Mohammad SinaMirabdolbaghi and BabakAmiri, "Model Optimization Analysis of Customer Churn Prediction Using Machine Learning Algorithms with Focus on Feature Reductions", Hindawi Discrete Dynamics in Nature and Society, Volume 2022, Article ID 5134356, 20 pages, doi:10.1155/2022/5134356

[4] Irina V. Pustokhina, Denis A. Pustokhin, PhongThanh Nguyen, Mohamed Elhoseny, K. Shankar, "Multi‑ objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector", Complex & Intelligent Systems, 2021, Springer, doi:10.1007/s40747-021-00353-6

[5] Jan Kunnen, MaximeDuchateau, Ziboud Van Veldhoven, Jan Vanthienen, "Benchmarking Stacking Against Other
Heterogeneous Ensembles in Telecom Churn Prediction", 2020 IEEE Symposium Series on Computational Intelligence (SSCI), December 1-4, 2020, Canberra, Australia, 978-1-7281-2547-3/20.

[6] Shreyas Rajesh Labhsetwar, "Predictive Analysis of Customer Churn in Telecom Industry using Supervised Learning", Predictive Analysis of Customer Churn In Telecom Industry Using Supervised Learning Doi: 10.21917/Ijsc.2020.0291

[7] TjengWawanCenggoro, RadityaAyuWirastaria, EdyRudianto, MochamadIlhamMohadi, DinneRatj, Bens Pardamean, "Deep Learning as a Vector Embedding Model for Customer Churn",
5th International Conference on Computer Science and Computational Intelligence 2020, Elsevier.

[8] AwodeleOludele, Adeniyi Ben, Ogbonna A.C., Kuyoro S.O., EbiesuwaSeun, "Enhanced Churn Prediction in the Telecommunication Industry",International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume-8, Issue-2, March 2020, doi:10.21276/ijircst.2020.8.2.1

[9] SajjadShumaly, PedramNeysaryan, YanhuiGuo, "Handling Class Imbalance in Customer Churn Prediction in Telecom Sector Using Sampling Techniques, Bagging and Boosting Trees", 10th International Conference on Computer and Knowledge Engineering (ICCKE2020)October 29-30, 2020 – Ferdowsi University of Mashhad – Iran.

[10]BuddhikaLivera, D D M Ranasinghe, "Naive Bayes Based Approach on Telecommunication Churn Prediction Using Financial Data",SLAAI - International Conference on Artificial Intelligence, 2020

[11] Ahmad Hammoudeh, MalakFraihat, Mahmoud Almomani, "Selective Ensemble Model for Telecom Churn Prediction", 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 978-1-5386-7942-5/19.

[12] SanketAgrawal, Aditya Das, AmitGaikwad, SudhirDhage, "Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using Deep Learning",International Conference on Smart Computing and Electronic Enterprise. (ICSCEE2018), 2018 IEEE,

[13] Muhammad Ali, Aziz Ur Rehman, ShamazHafeez, “Prediction of Churning Behavior of Customers in Telecom Sector Using Supervised Learning Techniques”, 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE),978-1-5386-4123-1/18

[14]Abinash Mishra; U. Srinivasulu Reddy, “A Novel Approach for Churn Prediction Using Deep Learning”, 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC),doi:10.1109/ICCIC.2017.8524551

[15] R. Prashant, K. Deepak, and Amit Kumar Meher, “High Accuracy Predictive Modelling for Customer Churn Prediction
in Telecom Industry”, MLDM 2017, LNAI 10358, pp. 391–402, 2017, Springer International Publishing, DOI: 10.1007/978-3-319-62416-7 28

[16]V. Umayaparvathi, K. Iyakutti, “Automated Feature Selection and Churn Prediction using Deep Learning Models”, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056, Volume: 04 Issue: 03, Mar -2017

[17]Abinash Mishra, U. Srinivasulu Reddy, “A Comparative Study of Customer Churn Prediction in Telecom Industry Using Ensemble Based Classifiers”,Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017), 978-1-5386-4031-9,

[18] Tan Yi Fei, Lam HaiShuan, Lai Jie Yan GuoXiaoning, SooWooi King, “Prediction on Customer Churn in the Telecommunications Sector Using Discretization and Naïve Bayes Classifier”, Int. J. Advance Soft Compu. Appl, Vol.9, No. 3, Nov 2017, ISSN 2074-8523

[19] Preeti K. Dalvi, Siddhi K. Khandge, AshishDeomore, AdityaBankar, Prof. V. A. Kanade, “Analysis of Customer Churn Prediction in Telecom Industry using Decision Trees and Logistic
Regression”, 2016 Symposium on Colossal Data Analysis and Networking (CDAN), 978-1-5090-0669-4

[20] Mohammad Ridwan Ismail, Mohd Khalid Awang, M Nordin A Rahman and MokhairiMakhtar, “A Multi-Layer Perceptron Approach for Customer Churn Prediction”,International Journal of Multimedia and Ubiquitous Engineering Vol.10, No.7 (2015), pp.213-222, doi:10.14257/ijmue.2015.10.7.22.