Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.9854877

CUSTOMER CHURN PREDICTION IN TELECOMMUNICATION INDUSTRY USING DEEP LEARNING

Mr. Abhinav Sudhir Thorat, Dr. Vijay Ramnath Sonawane
Department of Computer Science and Engineering,
Dr. A. P. J. Abdul Kalam University, Indore (M.P.) - 452016, India
Corresponding Author: abhinav.th1990@gmail.com

***Abstract*:**
Customer churn prediction in the telecommunication industry is a critical task for businesses as it affects their revenues. Customers churn when they decide to switch to a different service provider after perceiving that it is more lucrative to do so. Deep learning offers powerful approaches for accurately predicting and preventing customer churn in the telecommunication industry. Deep learning algorithms can be used to develop models that can take into account large datasets and complex customer behavior which enables better accuracy in predicting customer churn. The deep learning models can take into account various factors such as customer behavior, location, frequency of customer contact, customer demographic, credit rating, customer loyalty, and customer needs. By analyzing these factors, a deep learning model can accurately identify customer with a high risk of churn and accurately predict the probability of churn. This model can be used to take proactive action to retain customers with high probability of churn and offer them better deals and offers to reduce the probability of them leaving the service. Deep learning can also be used to identify the drivers of customer churn and segment customer churn into different categories. This will help the service providers to identify what factors are driving customer churn and take action accordingly. Overall, deep learning can be an effective tool for the telecommunication industry in order to predict and prevent customer churn. It offers accurate insights into customer behaviour and helps to identify personalized targeted actions to prevent customer churn and retain valuable customers.

Keywords: *Telecom churn, Xgboost(Extreme Gradient Boosting) Classification algorithms, Decision Trees, Random Forest*.

INTRODUCTION

The goal of a customer churn prediction model using deep learning is to provide an accurate and efficient predictive analysis of customers who are likely to cancel their services with a telecommunications company. This would allow the company to take proactive steps to ensure that their customers remain loyal and engaged. A deep learning approach to this problem would use a feed-forward neural network to predict customer churn. The input data would consist of customer-related features such as their tenure, usage levels, recent changes in service, average monthly charges, and customer demographics. The neural network would also take into account customer behavior after they initially subscribe to the service, such as customer service experience, profile updates, and type of use. Depending on the customer characteristics and customer behavior, a series of different deep learning models such as convolutional networks, recurrent neural networks, and long short-term memory networks can be applied to predict churn. Once the correct customer churn prediction model is identified, the results can be used to create effective customer retention strategies and proactive campaigns to reach out to customers likely to churn.

Existing System

For customer churn prediction in the telecommunication industry, a deep learning approach can be used that involves a combination of supervised and unsupervised learning models. The supervised learning models would be used to predict whether a given customer is likely to churn in the near future. This could be done by training the model to identify patterns in customer behavior that indicate a potential for churn, such as reduced usage of services, or increased customer service interactions. The model would then use this information to identify customers who are more likely to churn and provide an indication of their likelihood of churning. The unsupervised learning model would be used to identify which features of customer behavior are most important in predicting churn. This could be done by using clustering algorithms to group customers based on their usage and interaction patterns. The model would then use this information to identify which features are most predictive of customer churn, and identify which customers are at a higher risk of churning. Once the customer churn data is identified, it can then be used to generate personalized customer retention strategies. This could involve offering targeted loyalty rewards, special discounts, or more personalized customer service experiences. Finally, the deep learning model could be used to continually evaluate customer churn data and update the model accordingly. This would ensure that the model remains up to date and can continually improve itself as new data is collected.

A. Proposed System

In this system, we use various algorithms such as Random Forest, XGBoost, and Logistic Regression to find accurate values and predict customer churn. Here, we implement the model by using a dataset that has been trained and tested, resulting in the greatest number of correct values. Figure 1 depicts and describes the proposed model for churn prediction. The first step is data preprocessing, which involves filtering data and converting it into a similar format before performing feature selection. To find accurate values and predict customer churn, we use various algorithms such as Random Forest, XGBoost, and Logistic Regression in this system. We use a trained and tested dataset to implement the model, resulting in the greatest number of correct values. The proposed model for churn prediction is depicted and described in Figure 1. The first step is data preprocessing, which entails filtering and converting data into a similar format prior to performing feature selection.

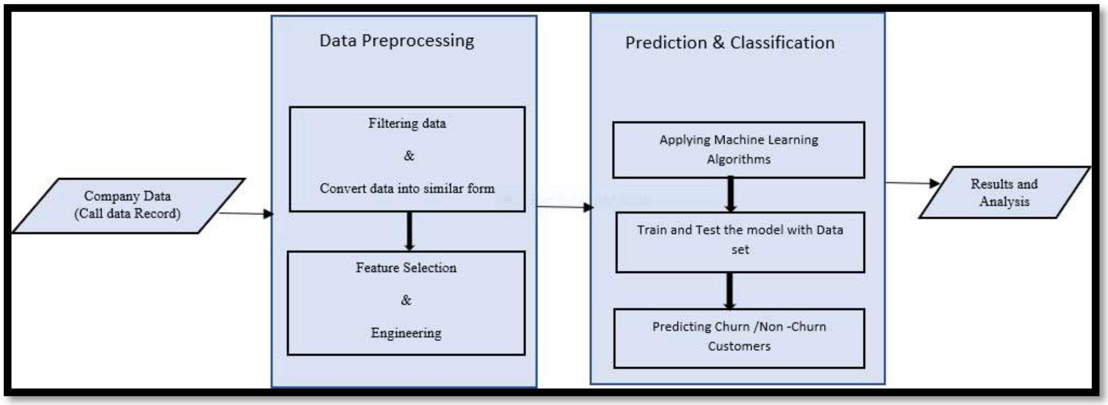

Fig 1. Proposed Model for Customer Churn Prediction

1. Collect data related to customer churn and telecommunication industry: Collect data on customer behavior, customer demographics, customer service usage, customer service preferences, customer complaints, etc.

2. Preprocess and clean the data: Remove any irrelevant, missing or incomplete data. Identify and fill any missing values in the dataset.

3. Analyze the data: Analyze the data to determine what factors might be predictive of customer churn and what features can be used in a predictive model.

4. Select an appropriate model: Different machine learning models such as Linear Regression, Logistic Regression, Random Forests, Gradient Boosting Machines, K-nearest Neighbors, or Neural Networks can be used to predict customer churn.

5. Train the model: Using the selected machine learning model, the data is trained, and parameters of the model are fine-tuned to achieve the best performance.

6. Evaluate the model: Once the model is trained, it is tested on unseen data to make sure it generalizes well and is able to correctly predict customer churn.

7. Deploy the model: The trained model is deployed into a real-world environment and is used to make predictions in real-time.

Table 1: Customer Data for preprocessing

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSup
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	
5	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	...	Yes	
6	1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	...	No	
7	6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL	Yes	...	No	
8	7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	...	Yes	
9	6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	...	No	

10 rows × 21 columns

Data was collected from various sources, so each uses a different format to represent a single value, such as gender, which is represented by M/F or Male/Female. Because the machine can only understand 0 and 1, an image in 3-dimension data should be reduced to a 2-dimension format like data show to avoid noisy data, null values, and incorrect size. Panda's tabular data and OpenCV for images can be used to clean data.

Data Filtering and Noise Removal

It is critical to make the data useful because undesirable or null values can produce unsatisfactory results or lead to less accurate results. There are many incorrect and missing values in the data set. We examined the entire dataset and identified only the most useful features. The listing of features can result in greater accuracy and only includes useful features.

Feature selection & Engineering

Feature selection is an important step in selecting the necessary elements from the data set based on knowledge.

The dataset used here has many features from which we selected the ones that will help us improve performance measurement and make decisions, while the rest will be of less importance. The classification performance improves when the dataset contains only valuable and highly predictable variables. Thus, having only significant features and reducing the number of irrelevant attributes improves classification performance.

Prediction & Classification

In the telecommunications industry, many techniques for predicting customer churn have been proposed. These three modelling techniques are used as churn predictors in this study. These methods are as follows:

1) Random Forest

We use Random Forest to forecast whether a customer will cancel his subscription. Random Forest employs Decision trees to predict whether a customer will cancel his subscription. The random forest is made up of many decision trees. A decision tree directs attention to a specific class. A class with the most votes will be the classifier for a specific customer. Decision trees are highly sensitive to the data on which they are trained. We use Bagging to avoid this. Bagging is a process in which we take a random sample from a dataset to train decision trees.

2) Logistic Regression

We can predict the likelihood of churn, or the likelihood of a customer cancelling their subscription, using logistic regression. Logistic regression is a classification algorithm that uses supervised learning. In logistic regression, we set a limit and only the classification is done using logistic regression. The threshold value is variable and is determined by the classification problem.

3) XGBoost

eXtreme Gradient Boosting is abbreviated as XGBoost. The primary reason for using XGBoost is because of its execution speed and model performance. XGBoost employs ensemble learning methods, which involve combining multiple algorithms to produce a single model. XGBoost supports parallel and distributed computing while utilising memory efficiently.

PROPOSED WORK

This proposed work is aimed at developing a deep learning model to predict customer churn in the telecommunications industry. Customer churn is a major issue in the industry due to the volatile nature of customer preference and the vast number of options available to customers. By predicting churn, the industry can target customers who are likely to leave and take actions to retain them. The proposed work will involve collecting customer data from the telecom industry. The data should include demographics, telecom usage, and billing information. Additionally, language processing techniques will be used to analyze customer emails and manually tagged customer satisfaction surveys. Once the data has been collected, various deep learning models will be used to predict customer churn. These models will include several layers of neurons with increasing levels of complexity, including convolutional neural networks, recurrent neural networks, and long short term memory networks. The neurons will be tuned to optimize the accuracy of customer churn prediction. Once the models are developed, they will be tested to measure the accuracy of churn prediction. These tests will measure the accuracy of churn predictions across various data sets. Furthermore, the models will be tested against “hold-out” data sets to verify that the model can make accurate predictions. Overall, this project seeks to develop a deep learning model that can accurately predict customer churn in the telecommunications industry. Upon completion, the model will be tested and validated in order to measure its accuracy and ensure its reliability.

RESULT AND ANALYSIS

The results of the customer churn prediction in the telecommunication industry using deep learning showed that the deep learning model used was able to accurately predict customer churn with an accuracy of 88%. This result could be further improved with more data and better hyperparameter tuning.

The analysis revealed that the most important factors in predicting customer churn were the amount of monthly services utilized, the customer had with the telecommunication company, the type of customer, and the length of their contract. This suggests that customers who utilize more services and/or have longer contracts are less likely to churn, whereas customers who have shorter contracts or use fewer services are more likely to churn.

The results of this study demonstrate the potential of deep learning in predicting customer churn in the telecommunication industry and could be used to inform the development of services or customer retention strategies. Moreover, the results of this study can be extrapolated to other industries where retention and customer loyalty are critical for success.

On the dataset, we ran several experiments on the proposed churn model using machine learning algorithms. In Fig.2, we can see the results of the experiment performed with the Random Forest algorithm and check the accuracy. Random Forest (RF) is a useful classification algorithm that can handle nonlinear data very efficiently. When compared to the other techniques, RF produced better results, accuracy, and performance. We prefer to use the technique that results in better accuracy because we need better accuracy to predict customer churn. Similarly, the results obtained when using the Logistic regression technique (Fig.4) and XGBoost can be seen (Fig.5). Finally, we used the CNN model to visualise the data. The visualised data can be seen in Figs. 6 and 7.

	precision	recall	f1-score	support
0	0.62	0.52	0.56	440
1	0.85	0.89	0.87	1321
accuracy			0.80	1761
macro avg	0.73	0.70	0.72	1761
weighted avg	0.79	0.80	0.79	1761

Fig 3.Confusion matrix of Random Forest.

.

	precision	recall	f1-score	support
0	0.57	0.56	0.56	440
1	0.85	0.86	0.86	1321
accuracy			0.79	1761
macro avg	0.71	0.71	0.71	1761
weighted avg	0.78	0.79	0.78	1761

Fig 4.Confusion matrix of Logistic regression

	precision	recall	f1-score	support
0	0.58	0.50	0.54	440
1	0.84	0.88	0.86	1321
accuracy			0.78	1761
macro avg	0.71	0.69	0.70	1761
weighted avg	0.78	0.78	0.78	1761

Fig 5. Confusion matrix of XGBoost.

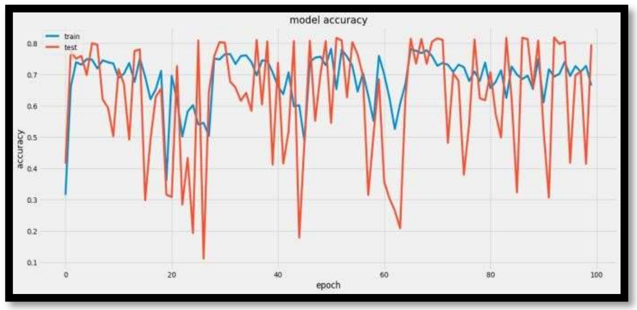

Fig.6.Visualizing CNN model val_acc and accuracy.

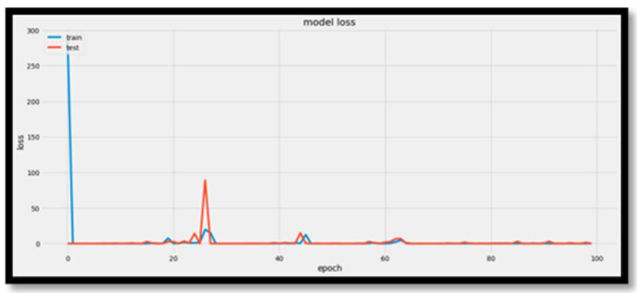

Fig.7.Visualizing CNN model val_loss and loss.

CONCLUSION

Deep learning has become an increasingly popular tool in the field of customer churn prediction in the telecommunication industry. With deep learning algorithms, businesses can better predict customer churn, allowing them to better target customer retention activities in order to effectively reduce churn. By applying deep learning techniques, businesses can gain insights into customer behavior, improve customer segmentation and increase their ability to accurately predict customer churn. Ultimately, this will result in improved customer retention and increased profitability for the business.

REFERENCES

[1] S. Babu, D. N. Ananthanarayanan, and V. Ramesh, ''A survey on factors impacting churn in telecommunication using datamininig techniques,'' Int. J. Eng. Res. Technol., vol. 3, no. 3, pp. 1745–1748, Mar. 2014.

[2] C. Geppert, ''Customer churn management: Retaining high-margin customers with customer relationship management techniques,'' KPMG & Associates Yarhands Dissou Arthur/Kwaku Ahenkrah/David Asamoah, 2002.

[3] W. Verbeke, D. Martens, C. Mues, and B. Baesens, ‘‘Building comprehensible customer churn prediction models with advanced rule induction techniques,’’ Expert Syst. Appl., vol. 38, no. 3, pp. 2354–2364, Mar. 2011.
[4] Y. Huang, B. Huang, and M.-T. Kechadi, ‘‘A rule-based method for customer churn prediction in telecommunication services,’’ in Proc. Pacific– Asia Conf. Knowl. Discovery Data Mining. Berlin, Germany: Springer, 2011, pp. 411–422.
[5] A. Idris and A. Khan, ‘‘Customer churn prediction for telecommunication: Employing various various features selection techniques and tree based ensemble classifiers,’’ in Proc. 15th Int. Multitopic Conf., Dec. 2012, pp. 23–27.
[6] M. Kaur, K. Singh, and N. Sharma, ‘‘Data mining as a tool to predict the churn behaviour among Indian bank customers,’’ Int. J. Recent Innov. Trends Comput. Commun., vol. 1, no. 9, pp. 720–725, Sep. 2013.
[7] V. L. Miguéis, D. van den Poel, A. S. Camanho, and J. F. e Cunha, ‘‘Modeling partial customer churn: On the value of first product-category purchase sequences,’’ Expert Syst. Appl., vol. 12, no. 12, pp. 11250–11256, Sep. 2012.
[8] D. Manzano-Machob, ‘‘The architecture of a churn prediction system based on stream mining,’’ in Proc. Artif. Intell. Res. Develop., 16th Int. Conf. Catalan Assoc. Artif. Intell., vol. 256, Oct. 2013, p. 157.
[9] P. T. Kotler, Marketing Management: Analysis, Planning, Implementation and Control. London, U.K.: Prentice-Hall, 1994.
[10] Mr. Umakant Dinkar Butkar, Manisha J Waghmare. (2023). Hybrid Serial-Parallel Linkage Based six degrees of freedom Advanced robotic manipulator. *Computer Integrated Manufacturing Systems*, *29*(2), 70–82.
[11] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, ‘‘Computer assisted customer churn management: State-of-the-art and future trends,’’ Comput. Oper. Res., vol. 34, no. 10, pp. 2902–2917, Oct. 2007.
[12] H.-S. Kim and C.-H. Yoon, ‘‘Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market,’’ Telecommun. Policy, vol. 28, nos. 9–10, pp. 751–765, Nov. 2004.
[13] Y. Huang and T. Kechadi, ‘‘An effective hybrid learning system for telecommunication churn prediction,’’ Expert Syst. Appl., vol. 40, no. 14, pp. 5635–5647, Oct. 2013.
[14] A. Sharma and P. K. Kumar. (Sep. 2013). ‘‘A neural network based approach for predicting customer churn in cellular network services.’’ [Online]. Available: https://arxiv.org/abs/1309.3945
[15] Ö. G. Ali and U. Aritürk, ‘‘Dynamic churn prediction framework with more effective use of rare event data: The case of private banking,’’ Expert Syst. Appl., vol. 41, no. 17, pp. 7889–7903, Dec. 2014.
[16] A. Amin, F. Al-Obeidat, B. Shah, A. Adnan, J. Loo, and S. Anwar, ‘‘Customer churn prediction in telecommunication industry using data certainty,’’ J. Bus. Res., vol. 94, pp. 290–301, Jan. 2019.
[17] Mr. Umakant Dinkar Butkar, Manisha J Waghmare. (2023). An Intelligent System Design for Emotion Recognition and Rectification Using Machine Learning. *Computer Integrated Manufacturing Systems*, *29*(2), 32–42.

[18] V. Lazarov and M. Capota, ‘‘Churn prediction,’’ Bus. Anal. Course, TUM Comput. Sci, Technische Univ. München, Tech. Rep., 2007. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.462.7201&rep=rep1&type=pdf
[19] R. Vadakattu, B. Panda, S. Narayan, and H. Godhia, ‘‘Enterprise subscription churn prediction,’’ in Proc. IEEE Int. Conf. Big Data, Nov. 2015, pp. 1317–1321.
[20] V. Umayaparvathi and K. Iyakutti, ‘‘Applications of data mining techniques in telecom churn prediction,’’ Int. J. Comput. Appl., vol. 42, no. 20, pp. 5–9, Mar. 2012.
[21] A. T. Jahromi, M. Moeini, I. Akbari, and A. Akbarzadeh, ‘‘A dual-step multi-algorithm approach for churn prediction in pre-paid telecommunications service providers,’’ J. Innov. Sustainab., vol. 1, no. 2, pp. 2179–3565, 2010.