

# Customer Stress Prediction in Telecom Industries using Machine Learning

Dr. Pathan Hussain Basha<sup>1</sup>, V.Sriharsha<sup>2</sup>, and S. Giri Babu<sup>3</sup>

<sup>1</sup>Professor, Department of Computer Science & Engineering, PACE Institute of Technology & Sciences, Ongole, Andhra Pradesh, India

<sup>2</sup>Assistant.Professor, Department of Computer Science & Engineering, PACE Institute of Technology & Sciences, Ongole, Andhra Pradesh, India

<sup>2</sup>Associate.Professor, Department of Computer Science & Engineering, PACE Institute of Technology & Sciences, Ongole, Andhra Pradesh, India

Copyright © 2022 Made Dr.Pathan Hussain Basha et al. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**ABSTRACT-** In the competitive world especially in enterprises market maintaining valuable customers is becoming a difficult task. In one situation losing a customer is like decrease in profits for telecom industry growth, in another situation the cost of acquiring new customers is much higher than the cost of retaining the existing customers, for this critical situation the telecom industries should focus on retaining existing customers. This project will analyze the customer data which was collected as open dataset and predict the customer stress by applying supervised machine learning algorithms mainly using Linear Discriminant Analysis, Support Vector Machine, K Nearest Neighbor and Random Forest.

**KEYWORDS-** Customer Stress, Support Vector Machine, K-Nearest Neighbor, Linear Discriminant Analysis, One Hot Encoding.

## I. INTRODUCTION

Running a Telecom industry is not an easy matter. The telecom industry is mainly based on the information and communication services, which was provided in different ways like Phone services, Internet services, Television, etc. By providing all services to the thousands of customers will may cause stress among the services. The stress may cause by Industry or some other changes in climate so on. The stress will may lose customer by moving to another industry. The customer stress will leads fall down of the market and causes great loss, for this cause the telecom industry must and should retain the existing customers, because the cost of acquiring new customer is greater than retaining existing customers. The machine learning algorithms will helps to predict the customer stress, accurate percentage value. At first we must collaborate with customers and collect data into an open dataset and start the process.

## II. LITERATURE SURVEY

In these days many corporations collaborating with customers and collecting the information where the customer was stressed and launching the services with an updated version to the last version. Which was provided by old Telecom industry and attracting customers by their advertisement which leads to major loss to other telecom industries. By observing this problem Reich held and Sasser in 1990 was started Customer Relationship Management which can impact stress rates overall telecom industry. Several studies combine several different types of predictors to develop a customer stress model. This model can take demographic characteristics, environmental changes and other factors into account.

## III. RELATED WORK

At beginning, customer stress prediction process started by Abbas Keramatiin 2014 featuring Demographic, Usage pattern, Customer service using Binomial Logistic and regression models. Ismail Mohammad in 2016 predicted customer stress by featuring Demographic billing data, usage pattern and customer relationship and using Neural network [1-5] and regression models. Chih Fong Tsai and Yu Hsin Lu in 2019 predicted customer stress by using hybrid neural networks. In this project we mainly discuss about:

- Linear Discriminant Analysis (LDA): In this algorithm mainly discuss about dimensionality reduction and used for supervised classification problem. It will convert from Higher into Lower dimension space.
- Support Vector Machine (SVM): In this algorithm it will construct hyper plane to classify the data and regression analysis and leads to optimal hyper plane in nonlinear manner.
- 3.K-Nearest Neighbor (KNN): In this algorithm it will captures the similar data points by calculating distance between them.

#### IV. EXISTING SYSTEM

Earlier in customer stress prediction, the authors used hybrid neural networks which leads to back propagation, the accuracy value of the models was not much accuracy, the confusion matrix will leads to error I and error II and the clustering process was the main problem to missing of data, so the accuracy was less.

#### V. PROPOSED SYSTEM

In this project the customer stress was predicted by supervised machine learning algorithms mainly using Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and the accuracy value will take by comparing best of each accuracy values [6-16]. The figure 1 shows the system architecture.

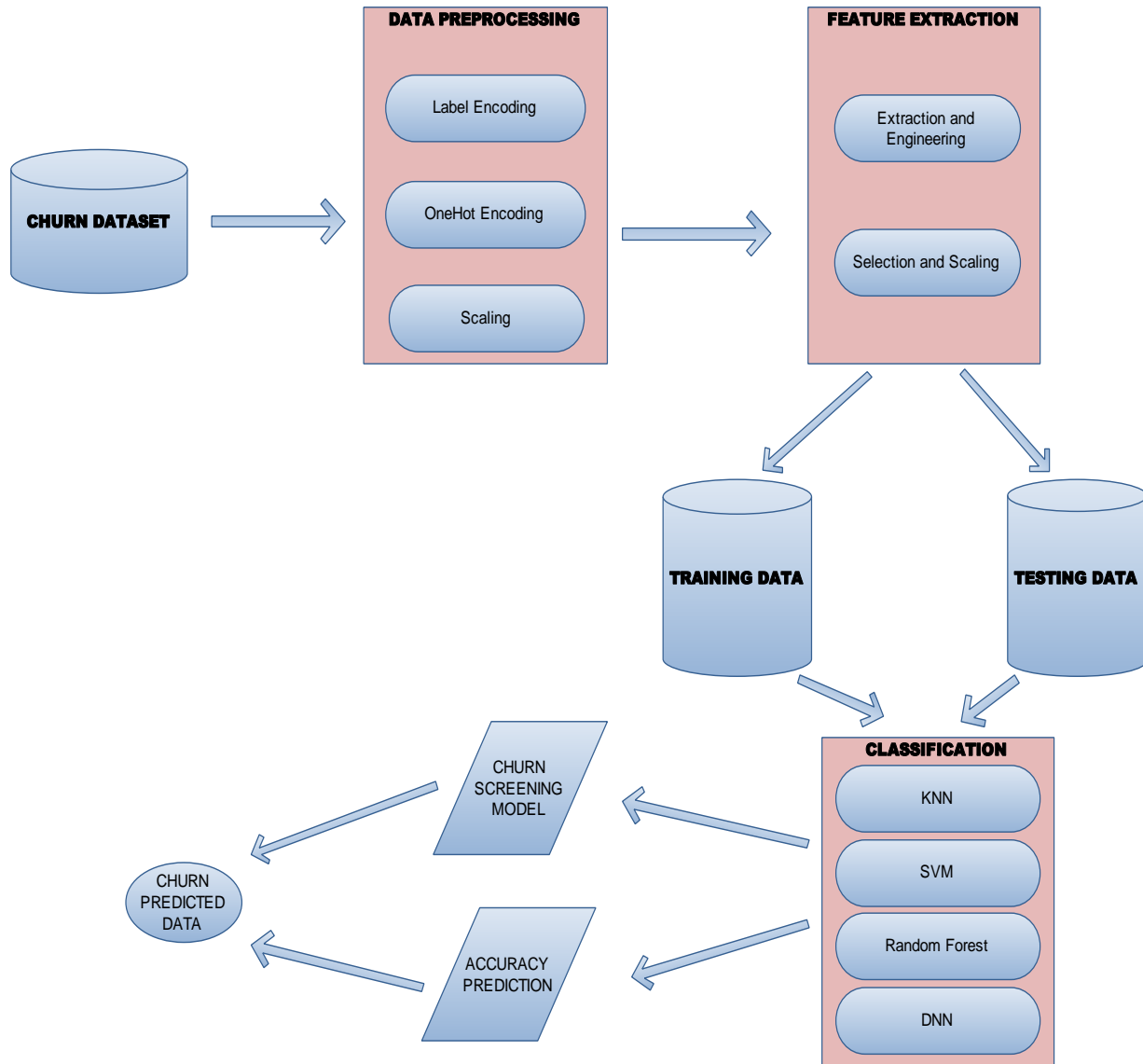


Figure.1. System architecture

##### A. Algorithm

The datasets are collected from customers by using different attributes along customer stress. The datasets will splits into test and train data. The 80% of datasets will splitsto train and 20% of datasets will splits to test. After the training and testing, the data was preprocessed. The data should be classified based on the conditions. After the classification the data is transferring to three different algorithms, each algorithm will check different accuracy values up to

hundred times. By comparing last accuracy value of the each algorithm, the best accuracy value to predict customer stress was taken.

#### VI. RESULTS AND DISCUSSION

The data collection and split the data as shown in figure 1 and figure 2.

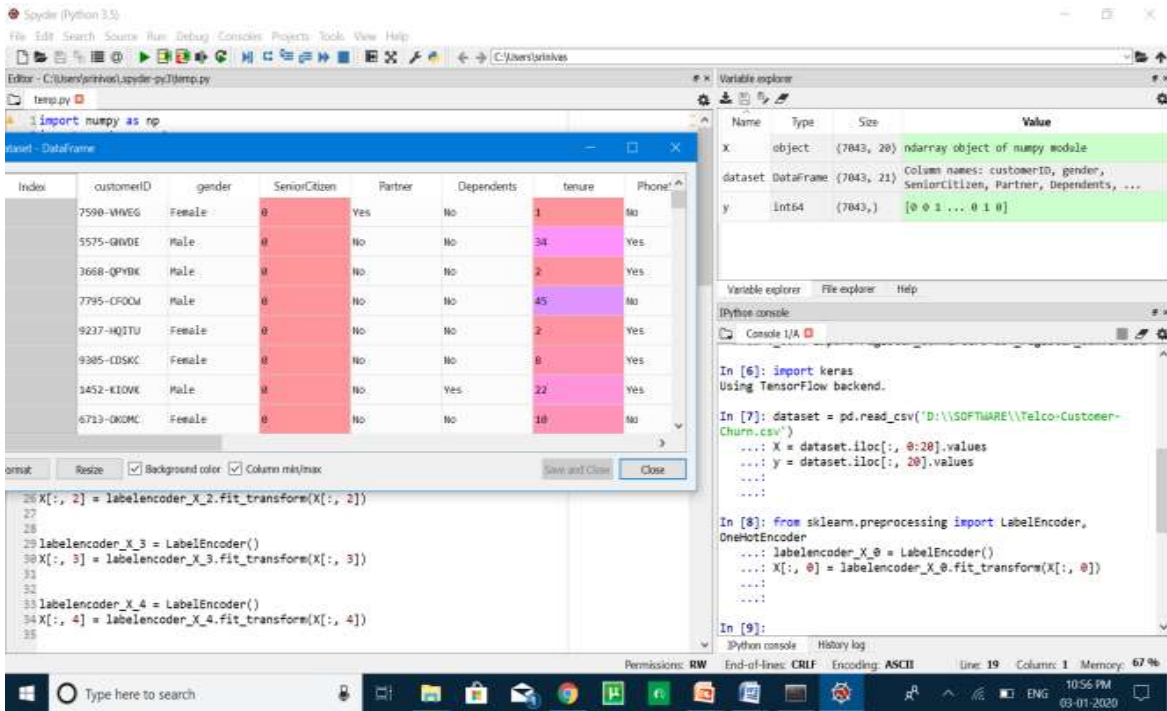


Figure 2: The data was collected from the customers

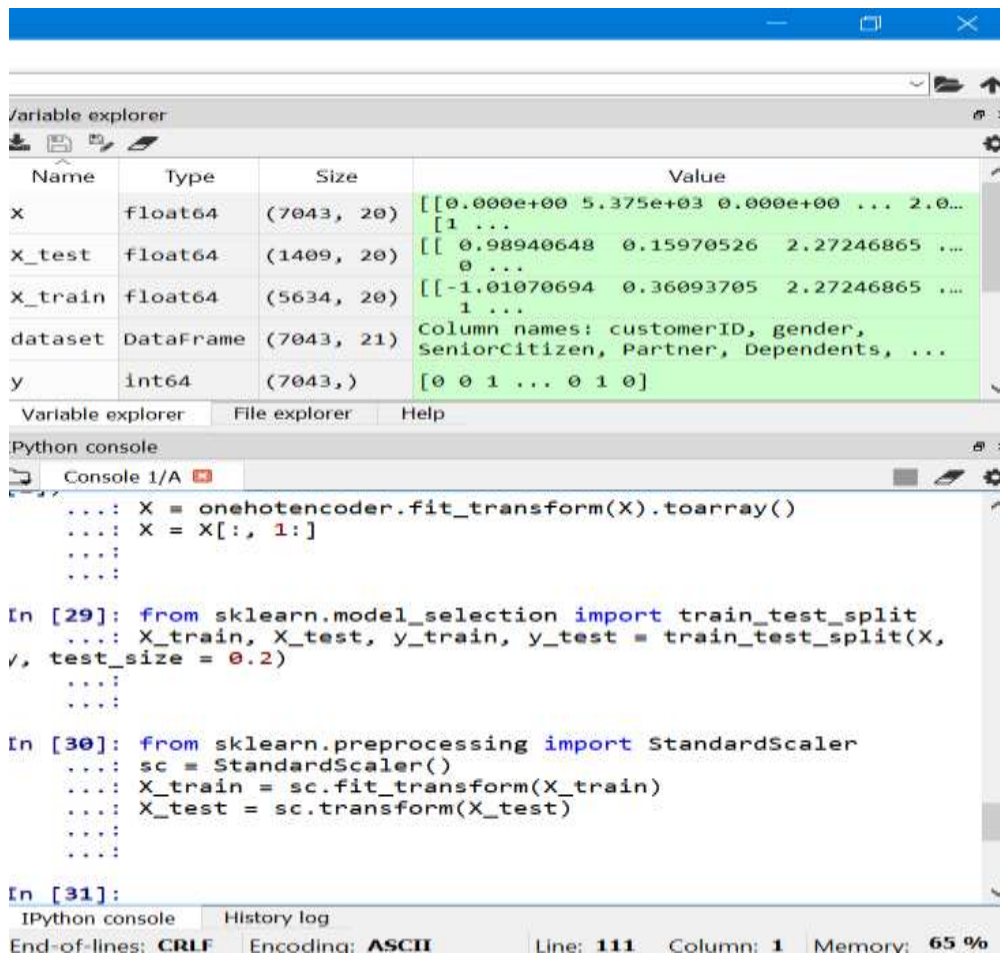


Figure 3: The data was split into training and testing

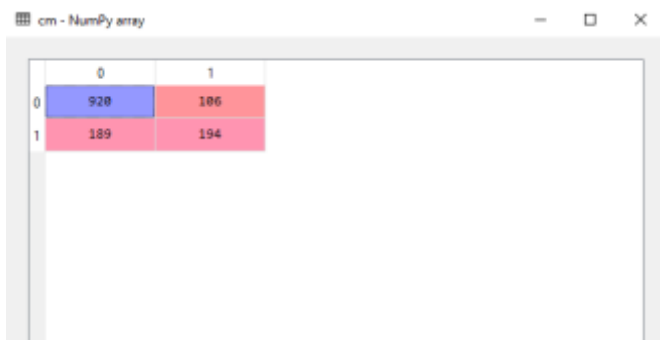


Figure 4: Confusion Matrix of Stress Prediction

The above figure 4 shows that the confusion matrix which is used to calculate the accuracy value of stress prediction.

## VII. CONCLUSION

In the project we considered the best of the three algorithms namely Support Vector Machine, K-Nearest Neighbor and Linear Discriminant Analysis according to the accuracy values given by these algorithms.

The highest percentile among three algorithms shows the maximum customers affecting on their respective services which they were in the long lasting relationship to the industry. Finally the telecom industry must satisfy old customers without interrupting the services using the stress prediction.

## REFERENCES

[1] Tou, J. T., & Gonzalez, R. C. (1974). Pattern recognition principles. Addison-Wesley.

[2] Van den Poel, D., & Larivie' re, B. (2004). Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research*, 157, 196–217.

[3] Wei, C. P., & Chiu, I. T. (2002). Turning telecommunications call details to churn prediction: A data mining approach. *Expert Systems with Applications*, 23, 103–112.

[4] Wong, B. K., Bodnovich, T. A., & Selvi, Y. (1997). Neural network applications in business: a review and analysis of the literature (1988–1995). *Decision Support Systems*, 19, 301–320.

[5] Zhan, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural network: The state of the art. *International Journal of Forecasting*, 14, 35–62.

[6] Zhang, X., Edwards, J., & Harding, J. (2007). Personalised online sales using web usage data mining. *Computers in Industry*, 58, 772–782.

[7] Berry, M. J. A., & Linoff, G. S. (2003). Data mining techniques: For marketing, sales, and customer support. John Wiley & Sons.

[8] Buckinx, W., & Van den Poel, D. (2005). Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research*, 164(1), 252–268.

[9] Burez, J., & Van den Poel, D. (2007). Crm at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. *Expert Systems with Applications*, 32, 277–288.

[10] Chiang, D. A., Wang, Y. F., Lee, S. L., & Lin, C. J. (2003). Goal-oriented sequential pattern for network banking churn analysis. *Expert Systems with Applications*, 25, 293–302.

[11] Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Systems with Applications*, 34, 313–327.

[12] Cybenko, H. (1998). Approximation by super positions of sigmoidal.

[12] S. Patibandla, M. Archana, and R. C. Tanguturi, "Object Tracking using Multi Adaptive Feature Extraction Technique," *International Journal of Engineering Trends and Technology*, vol. 70, no. 6, pp. 279–286, Jun. 2022, doi: 10.14445/22315381/ijett-v70i6p229.

[13] G. Sadineni, A. M, and R. C. Tanguturi, "Optimized Detector Generation Procedure for Wireless Sensor Networks based Intrusion Detection System," *International Journal of Engineering Trends and Technology*, vol. 70, no. 6, pp. 63–72, Jun. 2022, doi: 10.14445/22315381/ijett-v70i6p208.

[14] S. Patibandla, Dr. M. Archana, and Dr. R. C. Tanguturi, "DATA AGGREGATION BASED HYBRID DEEP LEARNING TECHNIQUE FOR IDENTIFYING THE UNCERTAINTIES AND ACCURATE OBJECT DETECTION," *Indian Journal of Computer Science and Engineering*, vol. 13, no. 3, pp. 697–708, Jun. 2022, doi: 10.21817/indjcs/2022/v13i3/221303049.

[15] Dr. S. R. Anand, Dr. R. C. Tanguturi, and S. D S, "Blockchain Based Packet Delivery Mechanism for WSN," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 2, pp. 1112–1117, Jul. 2019, doi: 10.35940/ijrte.b1627.078219.

[16] M. V. Bharathi, R. C. Tanguturi, C. Jayakumar, and K. Selvamani, "Node capture attack in Wireless Sensor Network: A survey," 2012 IEEE International Conference on Computational Intelligence and Computing Research, Dec. 2012, doi: 10.1109/iccic.2012.6510237.