

A Comprehensive Survey on Customer Churn Prediction Model for Telecommunication Sector

Dr. Animesh kumar sharma

S V College of Agricultural Engineering & Technology and Research Station- Jora,
Chhattisgarh

ABSTRACT

In current telecommunication business, it is vital to differentiate users aren't reluctant for moving towards a participant. Churn arises mostly because of customer dissatisfaction. Recognizing customer dissatisfaction needs many variables. Generally, user doesn't churn because of individual dissatisfaction condition. Examining this information will existing a clear understanding of the customer's present condition. Thus, customer churn prediction has to turn into a major problem in telecommunication business. In this competitive business a consistent user because of increasing amount of telecom suppliers. But because of this sparsity, hugeness, and imbalanced nature of the information, the churn predictive in telecom are often difficult to process. This paper has reviewed several CCP techniques that exist in the literature.

Keywords: Churn prediction, machine learning, Feature selection, telecommunication.

I. INTRODUCTION

An increasing amount of telecom suppliers have become a massive increase in competition and customer churn. Presently organization has emphasis on decreasing the churn by concentrating on customers autonomously. Churn [1] is determined by the tendency of a customer for ceasing business transactions with an organization. An essential need is the recognition of customers who have higher likelihoods of leaving this. The capability of an organization to interfere in a timely could efficiently decrease the churn. Churn arises mostly because of customer dissatisfaction. Recognizing customer dissatisfaction needs many variables. Generally, user doesn't churn because of individual dissatisfaction condition. These generally occur in numerous dissatisfaction scenarios earlier a customer fully stops to perform transactions with an organization. Various features are related to the customer and their manner of processes with an organization are noted by the organization. This denotes the customer's behavioral information. Examining this information will existing a clear understanding of the customer's present condition. Henceforth this is utilized as base data for predicting churn. The main problem that occurs from this mode of process is that the information in consideration is often large. The hugeness is featured in the behavior nature of the data, representing every product line is handled by the organization. Additionally, it is because of an essential structure illustration of the information, each instance is certain to have every feature equivalent to a general customer in the organization. This will result in data sparseness, as customers would be related to some features, and not every feature has pertained to the organization. The hugeness of sparsity and data plays as the main

problems in the procedure of churn prediction. Huge companies will communicate with their customer to give various facilities to them. Customer service is a major differentiator for companies. The capability of predicting whether the customer would leave to interfere in a timely could be substantial for preempting challenges and offering higher level of customer facility. The issue turns highly difficult as customer behavioral data is consecutive and varied. Churn is an inevitable procedure in some industries. But, by this complexity, it is probable for identifying the reasons for churn by numerous methods.

Customer churn prediction has to turn into a major problem in telecommunication business. In this competitive business a consistent user because of increasing amount of telecom suppliers. But because of this sparsity, hugeness, and imbalanced nature of the information, the churn predictive in telecom are an often difficult process. This paper has reviewed several CCP techniques that exist in the literature.

II. EXISTING WORKS ON CUSTOMER CHURN PREDICTION

[2] stimulates the examination of the probability churn by analyzing customer behaviors. The SVM, RF KNN, and DT classifications are utilized in this research. Similarly, few FS approaches have been made to discover related features and to authenticate system efficiency. The research was accompanied by the churn modeling dataset from Kaggle. [3] develops a churn predictive module that supports telecom operators for predicting customers that are expected subject to churn. The module established in this study utilizes ML

method on big data architecture and creates a novel method of FS and implementation.

[4] predict customer churn earlier thus appropriate customer retention stages could be occupied using investigative data analyses and create personalized proposals for the target. For predicting churn, this operation contains relative analyses of 4 modules, such as XGBoost, RF, LR, and SVM, on 3 various fields such as IT, banking, and telecom. [5] proposed a customer churn predictive module using unstructured data that is the spoken content in phone transmission. They collect larger scaling call centre dataset with 2 million calls from many half a million customers and accompanied widespread research. The outcomes display that this module could predict accurate customer churn threats and create useful perceptions by the interpretable ML method with personalities and customer sectors.

[6] determined the module of churn prediction. The sequence in which individual clients leave single company and enter into other is named churn. This study will examine to recognize the customer that can churn, utilize ML method for predicting and assist to denote huge datasets in graph formation. In [7], various modules of ML methods like RF, KKNN, DT, LR, and so on are employed in the bank dataset for predicting the likelihood of customer operating to churn. In [8], churn analyses are executed on real-time data from a Software as a Service (SaaS) company trading an innovative cloud based business phone system like Aircall. This application has certain feature that the presented dataset collects customer information on every month and contains an imbalanced distribution of the target: a huge number of customers don't churn. Thus, some approaches are attempted to reduce the effect of the imbalance when it is nearest to the real-time and temporal architecture. This technique includes over and under sampling (SMOTE and Tomek's link) and time sequence cross validation. Later the RF and LR techniques are utilized with an objective to explain and predict churns.

[9] improves a churn predictive module that supports telecom operators for predicting customers are most expected subject to churn. The module is improved in this study utilizes ML method on big data architecture and creates a novel manner of FS and implementing. The aim of this study is to give a prediction architecture of customer churn by 6 phases for precise predictions and prevent customer churn in the business field [10]. The 6 phases are given by initially, set of customer behavior and creation of the data's; next, the creation of acquired parameters and influential parameter selection utilized a technique of discriminative analyses; subsequently, selecting of trained and tested information and review their quantity; then, the improvement of predictive module utilize boosting, simple and bagging forms

of supervised ML method; after, the equivalence of churn predictive module is depending upon various forms of ML techniques and chosen parameters and lastly they provide suitable approach that is depending upon the presented module.

[11] concentrates on several ML methods to predict customer churn by building the classification modules like Gradient boosted tree, SVM, and LR, and relates the efficiency of this module. [12] proposed relative research on the most common ML technique that is employed to the complex issue of the customer churn predictions in the telecommunication field. In initial stage of this research, every module has been employed and calculated by the cross validation on a common, public domain dataset. In next stage, the efficiency enhancement is provided by boosting.

III. CONCLUSION

Customer churn prediction has to turn into a major problem in telecommunication business. In this competitive business a consistent user because of increasing amount of telecom suppliers. But because of this sparsity, hugeness, and imbalanced nature of the information, the churn predictive in telecom are an often difficult process. This paper has reviewed several CCP techniques that exist in the literature.

REFERENCES

- [1] Pustokhina, I.V., Pustokhin, D.A., Nguyen, P.T., Elhoseny, M., and Shankar, K., 2021. Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector. *Complex & Intelligent Systems*, pp.1-13.
- [2] Rahman, M. and Kumar, V., 2020, November. Machine Learning Based Customer Churn Prediction In Banking. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1196-1201). IEEE.
- [3] Ahmad, A.K., Jafar, A. and Aljoumaa, K., 2019. Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, 6(1), pp.1-24.
- [4] Jain, H., Yadav, G. and Manoov, R., 2021. Churn Prediction and Retention in Banking, Telecom and IT Sectors Using Machine Learning Techniques. In *Advances in Machine Learning and Computational Intelligence* (pp. 137-156). Springer, Singapore.
- [5] Vo, N.N., Liu, S., Li, X. and Xu, G., 2021. Leveraging unstructured call log data for customer churn prediction. *Knowledge-Based Systems*, 212, p.106586.

- [6] Mukhopadhyay, D., Malusare, A., Nandanwar, A. and Sakshi, S., 2021. An Approach to Mitigate the Risk of Customer Churn Using Machine Learning Algorithms. In *Machine Learning for Predictive Analysis* (pp. 133-142). Springer, Singapore.
- [7] Kaur, I. and Kaur, J., 2020, November. Customer Churn Analysis and Prediction in Banking Industry using Machine Learning. In *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 434-437). IEEE.
- [8] Sergue, M., 2020. Customer Churn Analysis and Prediction using Machine Learning for a B2B SaaS company.
- [9] Ahmad, A.K., Jafar, A. and Aljoumaa, K., 2019. Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, 6(1), pp.1-24.
- [10] Almana, A.M., Aksoy, M.S. and Alzahrani, R., 2014. A survey on data mining techniques in customer churn analysis for telecom industry. *International Journal of Engineering Research and Applications*, 4(5), pp.165-171.
- [11] S. Namasudra, S. Dhamodharavadhani, and R. Rathipriya, "Nonlinear neural network based forecasting model for predicting COVID-19 cases", *Neural Processing Letters*, 2021. DOI: 10.1007/s11063-021-10495-w
- [12] Elhoseny, M., & Shankar, K. (2019). Reliable data transmission model for mobile ad hoc network using signcryption technique. *IEEE Transactions on Reliability*, 69(3), 1077-1086.
- [13] Mukherjee, R., Kundu, A., Mukherjee, I., Gupta, D., Tiwari, P., Khanna, A., & Shorfuzzaman, M. (2021). IoT-cloud based healthcare model for COVID-19 detection: an enhanced k-Nearest Neighbour classifier based approach. *Computing*, 1-21.
- [14] Li, L., Sun, L., Xue, Y., Li, S., Huang, X., & Mansour, R. F. (2021). Fuzzy Multilevel Image Thresholding Based on Improved Coyote Optimization Algorithm. *IEEE Access*, 9, 33595-33607.
- [15] Alzubi, O. A. A deep learning-based frechet and dirichlet model for intrusion detection in IWSN. *Journal of Intelligent & Fuzzy Systems*, (Preprint), 1-11.
- [16] Kavitha, M., & Palani, S. (2014). Blood vessel, optical disk and damage area-based features for diabetic detection from retinal images. *Arabian Journal for Science and Engineering*, 39(10), 7059-7071.
- [17] Sangeetha J., Jayasankar T. (2019) Emotion Speech Recognition Based on Adaptive Fractional Deep Belief Network and Reinforcement Learning. In: Mallick P., Balas V., Bhoi A., Zobaa A. (eds) *Cognitive Informatics and Soft Computing. Advances in Intelligent Systems and Computing*, vol 768. Springer, Singapore. https://doi.org/10.1007/978-981-13-0617-4_16
- [18] Gaur, A. and Dubey, R., 2018, December. Predicting Customer Churn Prediction In Telecom Sector Using Various Machine Learning Techniques. In *2018 International Conference on Advanced Computation and Telecommunication (ICACAT)* (pp. 1-5). IEEE.
- [19] Vafeiadis, T., Diamantaras, K.I., Sarigiannidis, G. and Chatzisavvas, K.C., 2015. A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*, 55, pp.1-9.
- [20] Alzubi, O. A. (2015, September). Performance evaluation of AG block turbo codes over fading channels using BPSK. In *Proceedings of the The International Conference on Engineering & MIS 2015* (pp. 1-6).
- [21] Kavitha, M., & Palani, S. (2014). Hierarchical classifier for soft and hard exudates detection of retinal fundus images. *Journal of Intelligent & Fuzzy Systems*, 27(5), 2511-2528.
- [22] Anuradha, M., Jayasankar, T., Prakash, N. B., Sikkandar, M. Y., Hemalakshmi, G. R., Bharatiraja, C., & Britto, A. S. F. (2021). IoT enabled cancer prediction system to enhance the authentication and security using cloud computing. *Microprocessors and Microsystems*, 80, 103301.
- [23] S. Kumari, R. J. Yadav, S. Namasudra, and C. H. Hsu, "Intelligent deception techniques against adversarial attack on industrial system", *International Journal of Intelligent Systems*, vol. 36, no. 5, pp. 2412-2437, 2021. DOI: 10.1002/int.22384
- [24] Uthayakumar, J., Elhoseny, M., & Shankar, K. (2020). Highly reliable and low-complexity image compression scheme using neighborhood correlation sequence algorithm in WSN. *IEEE Transactions on Reliability*, 69(4), 1398-1423.
- [25] Chavhan, S., Gupta, D., Nagaraju, C., Rammohan, A., Khanna, A., & Rodrigues, J. J. (2021). An Efficient Context-Aware Vehicle Incidents Route Service Management for Intelligent Transport System. *IEEE Systems Journal*.
- [26] Mansour, R. F., El Amraoui, A., Nouaouri, I., Díaz, V. G., Gupta, D., & Kumar, S. (2021). Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model

- for Smart Healthcare Systems. IEEE Access, 9, 45137-45146.
- [27] Alzubi, O. A. (2016). An empirical study of irregular ag block turbo codes over fading channels. arXiv preprint arXiv:1604.00564.
- [28] Kavitha, M., & Palani, D. S. (2012). A New Fast Curvelet Transform with Morphological Operations based method for Extraction of Retinal blood vessels using Graphical User Interfacel. International Journal of Scientific & Engineering Research, 3(6).
- [29] Ramesh, S., Yaashuwanth, C., Prathibanandhi, K., Basha, A. R., & Jayasankar, T. (2021). An optimized deep neural network based DoS attack detection in wireless video sensor network. Journal of Ambient Intelligence and Humanized Computing, 1-14.
- [30] P. Pavithran, S. Mathew, S. Namasudra and P. Lorenz, "A novel cryptosystem based on DNA cryptography and randomly generated Mealy machine", Computers & Security, vol. 104, 2021. DOI: <https://doi.org/10.1016/j.cose.2020.102160>
- [31] Le, DN., Parvathy, V.S., Gupta, D. et al. IoT enabled depthwise separable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification. Int. J. Mach. Learn. & Cyber. (2021). <https://doi.org/10.1007/s13042-020-01248-7>
- [32] Sekaran, R., Goddumarri, S. N., Kallam, S., Ramachandran, M., Patan, R., & Gupta, D. (2021). 5G Integrated Spectrum Selection and Spectrum Access using AI-based Framework for IoT based Sensor Networks. Computer Networks, 186, 107649.
- [33] Zhang, Y. H., Li, Z., Zeng, T., Chen, L., Li, H., Gamarra, M., ... & Cai, Y. D. (2021). Investigating gene methylation signatures for fetal intolerance prediction. Plos one, 16(4), e0250032.
- [34] Alzubi, J. A. (2021). Blockchain-based Lamport Merkle Digital Signature: Authentication tool in IoT healthcare. Computer Communications, 170, 200-208.
- [35] Kavitha, M., & Palani, S. (2012). Retinal blood vessel segmentation algorithm for diabetic retinopathy and abnormality classification by supervised machine learning. Int. J. Neural Netw. Appl, 5(1), 47-53.
- [36] Jayanthi, J., Jayasankar, T., Krishnaraj, N., Prakash, N. B., Sagai Francis Britto, A., & Vinoth Kumar, K. (2021). An Intelligent Particle Swarm Optimization with Convolutional Neural Network for Diabetic Retinopathy Classification Model. Journal of Medical Imaging and Health Informatics, 11(3), 803-809.
- [37] S. Kumari and S. Namasudra, "System reliability evaluation using budget constrained real d-MC search", Computer Communications, vol. 171, 2021. DOI: <https://doi.org/10.1016/j.comcom.2021.02.004>
- [38] Shankar, K., Perumal, E. A novel hand-crafted with deep learning features based fusion model for COVID-19 diagnosis and classification using chest X-ray images. Complex Intell. Syst. (2020). <https://doi.org/10.1007/s40747-020-00216-6>
- [39] Abukharis, S., Alzubi, J. A., Alzubi, O. A., Alamri, S., & O'Farrell, T. (2016). Packet error rate performance of IEEE802. 11g under bluetooth interface. arXiv preprint arXiv:1602.05556.
- [40] Kavitha, M., & Palani, S. (2015). Hierarchical Classifier For Microaneurysm Detection. International Journal of Applied Engineering Research, 10(1), 1449-1458.
- [41] Anuradha, M., Ganesan, V., Oliver, S., Jayasankar, T., & Gopi, R. (2020). Hybrid firefly with differential evolution algorithm for multi agent system using clustering based personalization. Journal of Ambient Intelligence and Humanized Computing, 1-10.
- [42] Hnatiuc, M., Geman, O., Avram, A. G., Gupta, D., & Shankar, K. (2021). Human Signature Identification Using IoT Technology and Gait Recognition. Electronics, 10(7), 852.
- [43] Miled, A. B., Dhaouadi, R., & Mansour, R. F. (2020). Knowledge Deduction and Reuse Application to the Products' Design Process. International Journal of Software Engineering and Knowledge Engineering, 30(02), 217-237.
- [44] S. Namasudra, R. Chakraborty, A. Majumder and N. R. Moparthy, "Securing multimedia by using DNA based encryption in the cloud computing environment", ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 16, no. 3s, 2020. DOI: <https://doi.org/10.1145/3392665>
- [45] Shankar, K., & Elhoseny, M. (2019). Trust Based Cluster Head Election of Secure Message Transmission in MANET Using Multi Secure Protocol with TDES. J. UCS, 25(10), 1221-1239.
- [46] Chen, T. M., Blasco, J., Alzubi, J. A., & Alzubi O. A. (2014). Intrusion detection. IET, 1(1), 1-9.
- [47] Kavitha, M., Syedakbar, S., Meenal, T., Kumar, R. S., & Stonier, A. A. (2021, February). Enhanced Algorithm for Bio

- Metric Based Secret Data Hiding. In IOP Conference Series: Materials Science and Engineering (Vol. 1055, No. 1, p. 012126). IOP Publishing.
- [48] Punarselvam, E., Sikkandar, M. Y., Bakouri, M., Prakash, N. B., Jayasankar, T., & Sudhakar, S. (2020). Different loading condition and angle measurement of human lumbar spine MRI image using ANSYS. *Journal of Ambient Intelligence and Humanized Computing*, 1-14.
- [49] S. Namasudra, "An improved attribute-based encryption technique towards the data security in cloud computing", *Concurrency and Computation: Practice and Exercise*, vol. 31, no. 3, 2019. DOI: 10.1002/cpe.4364
- [50] Shankar, K., Lakshmanprabu, S. K., Khanna, A., Tanwar, S., Rodrigues, J. J., & Roy, N. R. (2019). Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. *Computers & Electrical Engineering*, 77, 230-243.
- [51] Alrabea, A., Alzubi, O. A., & Alzubi, J. A. (2019). A task-based model for minimizing energy consumption in WSNs. *Energy Systems*, 1-18.
- [52] KAVITHA, M., GANESH, R., & RAJKUMAR, A. FACILITIES NAVIGATION ANDPATIENT MONITORING SYSTEM USING IBEACON TECHNOLOGY.
- [53] Kumar, K. V., Jayasankar, T., Eswaramoorthy, V., & Nivedhitha, V. (2020). SDARP: Security based Data Aware Routing Protocol for ad hoc sensor networks. *International Journal of Intelligent Networks*, 1, 36-42.
- [54] S. Namasudra and P. Roy, "Time saving protocol for data accessing in cloud computing", *IET Communications*, vol. 11, no. 10, pp. 1558-1565, 2017.
- [55] Shankar, K., Lakshmanprabu, S. K., Gupta, D., Khanna, A., & de Albuquerque, V. H. C. (2020). Adaptive optimal multi key based encryption for digital image security. *Concurrency and Computation: Practice and Experience*, 32(4), e5122.
- [56] Mansour, R. F. (2017). Evolutionary computing enriched ridge regression model for craniofacial reconstruction. *Multimedia Tools and Applications*, 1-18.
- [57] Pustokhina, I. V., Pustokhin, D. A., Kumar Pareek, P., Gupta, D., Khanna, A., & Shankar, K. (2021). Energy-efficient cluster-based unmanned aerial vehicle networks with deep learning-based scene classification model. *International Journal of Communication Systems*, e4786.
- [58] Sholiyi, A., Alzubi, J. A., Alzubi, O. A., Almomani, O., & O'Farrell, T. (2016). Near capacity irregular turbo code. arXiv preprint arXiv:1604.01358.
- [59] Muthumayil, K., Buvana, M., & Jayasankar, T. (2021). Energy Utilization using Artificial Bee Colony Algorithm for Network Life Time Enhancement of Homogeneous WSNs. *International Journal of Modern Agriculture*, 10(2), 1649-1656.
- [60] S. Namasudra, R. Chakraborty, S. Kadry, G. Manogaran and B. S. Rawal, "FAST: Fast accessing scheme for data transmission in cloud computing", *Peer-to-Peer Networking and Applications*, 2020. DOI: 10.1007/s12083-020-00959-6
- [61] Mansour, R. F. (2015). Using adaptive mutation to accelerate the convergence of immune algorithms for prediction of 3D molecular structure. *International Journal of Computers and Applications*, 37(3-4), 127-133.
- [62] Shankar, K., Elhoseny, M., Kumar, R. S., Lakshmanprabu, S. K., & Yuan, X. (2020). Secret image sharing scheme with encrypted shadow images using optimal homomorphic encryption technique. *Journal of Ambient Intelligence and Humanized Computing*, 11(5), 1821-1833.
- [63] Mansour, R. F., & Aljehane, N. O. (2021). An optimal segmentation with deep learning based inception network model for intracranial hemorrhage diagnosis. *Neural Computing and Applications*, 1-13.
- [64] Alzubi, J. A. (2020). Bipolar fully recurrent deep structured neural learning based attack detection for securing industrial sensor networks. *Transactions on Emerging Telecommunications Technologies*, e4069.
- [65] Kavitha, M., Lavanya, G., & Janani, J. (2018). Enhanced SVM classifier for breast cancer diagnosis. *International Journal of Engineering Technologies and Management Research*, 5(3), 67-74.
- [66] Parvathy, P., Subramaniam, K., Venkatesan, G. P., Karthikaikumar, P., Varghese, J., & Jayasankar, T. (2020). Development of hand gesture recognition system using machine learning. *Journal of Ambient Intelligence and Humanized Computing*, 1-8.
- [67] S. Namasudra, "Data access control in the cloud computing environment for bioinformatics", *International Journal of Applied Research in Bioinformatics (IJARB)*, vol. 11, no. 1, pp. 40-50, 2021. DOI: 10.4018/IJARB.2021010105
- [68] K. Shankar, Y. Zhang, Y. Liu, L. Wu and C. Chen, "Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification," in *IEEE Access*, vol.

- 8, pp. 118164-118173, 2020, doi: 10.1109/ACCESS.2020.3005152.
- [69] Mansour, R. F., & Abdelrahim, E. M. (2019). An evolutionary computing enriched RS attack resilient medical image steganography model for telemedicine applications. *Multidimensional Systems and Signal Processing*, 30(2), 791-814.
- [70] Alzubi, O. A., Alzubi, J. A., Dorgham, O., & Alsayyed, M. (2020). Cryptosystem design based on Hermitian curves for IoT security. *The Journal of Supercomputing*, 76(11), 8566-8589.
- [71] Kavitha, M., & Palani, S. (2020). A comprehensive analysis for retinal image classification methods using real-time database. *International Journal of Business Information Systems*, 34(2), 229-252.
- [72] Nair, L. R., Subramaniam, K., PrasannaVenkatesan, G. K. D., Baskar, P. S., & Jayasankar, T. (2020). Essentiality for bridging the gap between low and semantic level features in image retrieval systems: an overview. *Journal of Ambient Intelligence and Humanized Computing*, 1-13.
- [73] S. Namasudra, "Fast and secure data accessing by using DNA computing for the cloud environment", *IEEE Transactions on Services Computing*, 2020. DOI: 10.1109/TSC.2020.3046471
- [74] Shankar, K., Sait, A. R. W., Gupta, D., Lakshmanprabu, S. K., Khanna, A., & Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recognition Letters*, 133, 210-216.