Evaluating the Effectiveness of Machine Learning Techniques for Customer Churn Prediction in the Telecom Industry.

A Focus on the British Market.

Research Title.

An Investigation into the Effectiveness of Machine Learning Techniques for Customer Churn Prediction in the Telecom Industry:
A British Context.

Abstract.

This study intends to identify customers most likely to cancel their subscriptions by analysing and forecasting customer turnover. To achieve this, the project will explore machine learning algorithms such as **Decision Tree**, **Random Forest**, **Gradient Boosted Machine Tree**, **Logistic Regression**, and **Extreme Gradient Boosting** to develop a predictive churn model to forecast customer churn in the telecom industry.

Research Problem.

Recent advancements in machine learning (ML) offer promising avenues for enhancing churn prediction models by leveraging large datasets and identifying intricate patterns that elude traditional statistical methods. However, applying ML techniques in the British telecom sector's churn prediction still needs to be explored, with existing studies providing limited insights into their effectiveness, adaptability, and practical implementation challenges. This research aims to **bridge this gap by systematically investigating the application and performance of various ML techniques** in predicting customer churn within this context with additional **Social Network Analysis parameters**.

Research Question.

How effective are machine learning techniques in predicting customer churn in the telecom industry?

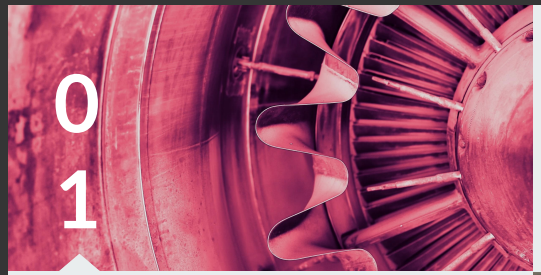
Research Aim.

This research aims to investigate the use of machine learning techniques in developing predictive models that can effectively determine customer churn.

Research Objectives.

1. To analyse the factors and reasons contributing to customer churn in the telecom industry.
2. **To build a churn prediction model based on customer service usage history.**
3. To develop a new way of feature engineering and selection to enhance the performance of the churn prediction model (Ahmad et al., 2019).
4. **To evaluate the performance of different machine learning algorithms (such as Decision Tree, Random Forest, Gradient Boosted Machine Tree, and Extreme Gradient Boosting) in predicting customer churn in the telecom industry.**
5. Incorporate social network analysis features into the churn prediction model to enhance its performance further (Ahmad et al., 2019).

Key Literatures.



Krishnaveni et al., 2022

Telecom Churn Prediction using Machine Learning

Ahmad et al. (2019)

Customer Churn Prediction in Telecom Using Machine Learning in Big Data Platform

Marin, M. & Goran, K. 2022

Churn Prediction Model Improvement Using Automated Machine Learning with Social Network Parameters

Methodology & Design

The proposed methodology for this research involves utilising machine learning techniques for churn analysis in the telecommunications industry. Specifically, the following steps will be taken:

01

Data Collection

Gather a dataset from telecommunications companies that includes relevant variables such as customer demographics, usage patterns, service complaints, and churn status.

02

Data Preprocessing

Clean and preprocess the collected data by handling missing values, encoding categorical variables, and scaling numerical variables.

03

Feature Selection

Use methods such as correlation analysis, chi-square tests, and recursive feature elimination to identify the most relevant features for churn prediction.

04

Model Development

Implement and evaluate multiple machine learning algorithms, including Logistic Regression, Random Forest, Decision Tree, and Gradient Boosting, to develop predictive models for customer churn.

05

Model Evaluation & Cross-validation

Assess the performance of the developed models using evaluation metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve.

06

Results Analysis

Analyse the results of the developed models and compare their performance to determine the most accurate algorithm for customer churn prediction in the telecommunications industry.

Challenges and Limitations.

Predicting customer churn in the telecommunications industry is a complex task. Various challenges and limitations come with it. Some challenges include dealing with imbalanced datasets, where the number of churned customers is much smaller than that of non-churned customers

Datasets

Imbalanced, noisy & incomplete

Market

external factors/conditions

Customer

Customer behaviour

Overcoming Limitation.

Investigating the use of advanced machine learning techniques, such as deep learning or ensemble methods and Incorporating additional external data sources, such as customer social media activity, could offer further insights and enhance the predictive power of the models.

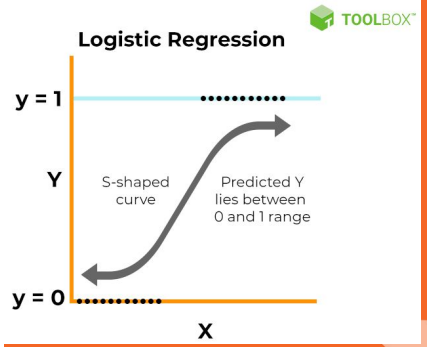
Ethical Consideration and Risk Assessment.

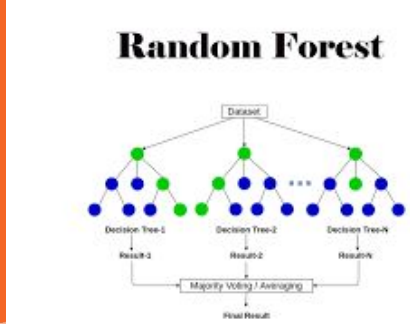
Customer privacy and data protection regulations

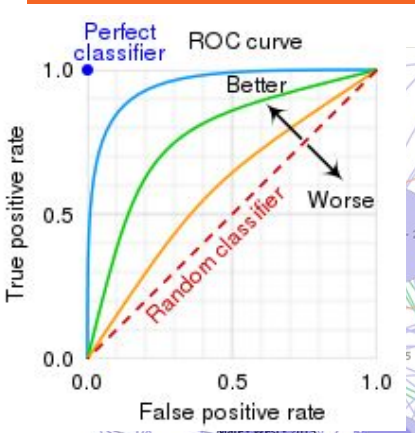
Misclassifying customers as churners or overlooking those likely to churn

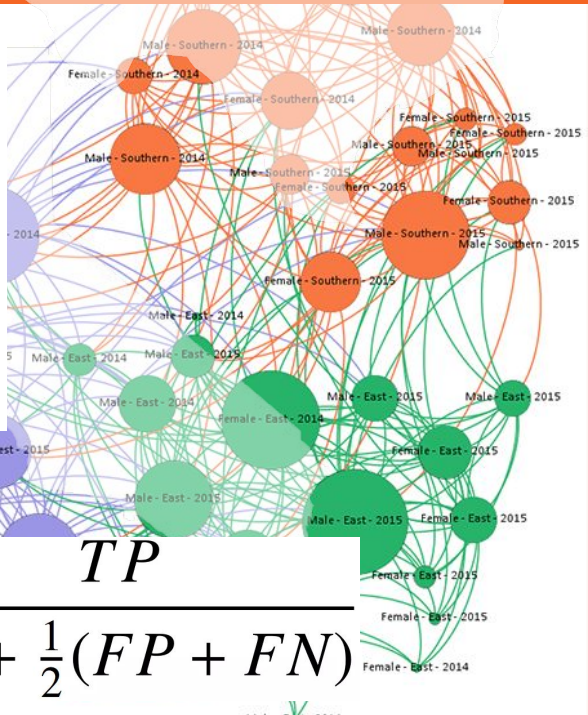
Combine human expertise with algorithmic predictions .

Description of Artefacts.









$$\text{F1 Score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

Timelines

Month 1.
Background Research, Literature Review & Data Collection

Month 2.
Data Preprocessing & Feature Selection

Month 3.
Model Development, Model Evaluation & Cross-validation.

Month 4.
Results Analysis & Conclusion

Month 5.
Writing & Revision

References.

Marin, M. & Goran, K. (2022). Telecom Churn Prediction using Machine Learning. World Journal of Advanced Engineering Technology and Sciences Vol. 07, pp. 87-96, https://doi.org/10.30574/wjaets.2022.7.2.0130

Krishnaveni, CV., et al. (2022). Churn Prediction Model Improvement Using Automated Machine Learning with Social Network Parameters. Revue d'Intelligence Artificielle Vol. 36, No. 3, pp. 373-379, https://doi.org/10.18280/ria.360304

Vafeiadis, T., et al. (2015). A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, Vol. 55, pp. 1-9, https://doi.org/10.1016/j.simpat.2015.03.003.

Nagaraju, J. et al. (2023). Predicting Customer Churn in Insurance Industry Using Big Data and Machine Learning. First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI), Tiruchengode, India, 2023, pp. 1-8, doi: 10.1109/ICAEECI58247.2023.10370876.

References.

Ahmed, M., et al. (2017). A Survey of Evolution in Predictive Models and Impacting Factors in Customer Churn. Advances in Data Science and Adaptive Analysis, Vol. 09, https://doi.org/10.1142/S2424922X17500073

Dahiya, K. & Bhatia, S. (2015). Customer churn analysis in telecom industry. 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Noida, India, 2015, pp. 1-6, doi: 10.1109/ICRITO.2015.7359318.

Yeshwanth, V. et al. (2011). Evolutionary Churn Prediction in Mobile Networks Using Hybrid Learning. 24th International Florida Artificial Intelligence Research Society Conference, pp. 471-476 https://cdn.aaai.org/ocs/2538/2538-11222-1-PB.pdf

Höppner, S., et al. (2020). Profit driven decision trees for churn prediction. European Journal of Operational Research. Vol. 284, Issue 3, pp. 471-476, https://doi.org/10.1016/j.ejor.2018.11.072.

References.

Idris, A., et al. (2012). Genetic Programming and Adaboosting based churn prediction for Telecom, IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, Korea (South), 2012, pp. 1328-1332, doi: 10.1109/ICSMC.2012.6377917.

Amin, A., et. al. (2015). Churn Prediction in Telecommunication Industry Using Rough Set Approach. In: Camacho, D., Kim, SW., Trawiński, B. (eds) New Trends in Computational Collective Intelligence. Studies in Computational Intelligence, vol 572. Springer, Cham. https://doi.org/10.1007/978-3-319-10774-5_8

Gaur, A. & Dubey, R. (2018). Predicting Customer Churn Prediction In Telecom Sector Using Various Machine Learning Techniques, International Conference on Advanced Computation and Telecommunication (ICACAT), Bhopal, India, 2018, pp. 1-5, doi: 10.1109/ICACAT.2018.8933783.

References.

Yiqing, H. et. al, (2015). Telco Churn Prediction with Big Data. In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15). Association for Computing Machinery, New York, NY, USA, 607–618. https://doi.org/10.1145/2723372.2742794

Mishra, K. & Rani, R. (2017). Churn prediction in telecommunication using machine learning. International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 2017, pp. 2252-2257, doi: 10.1109/ICECDS.2017.8389853.

Prashanth, R., et. al. (2017). High Accuracy Predictive Modelling for Customer Churn Prediction in Telecom Industry. In: Perner, P. (eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2017. Lecture Notes in Computer Science(), vol 10358. Springer, Cham. https://doi.org/10.1007/978-3-319-62416-7_28

References.

Srinivasan, R. et al. (2023). Customer Churn Prediction Using Machine Learning Approaches. International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-6, doi: 10.1109/ICECONF57129.2023.10083813.

Andrews, R. et al. (2019). CHURN PREDICTION IN TELECOM SECTOR USING MACHINE LEARNING. International Journal of Information Systems and Computer Sciences, Vol. 8. Issue 2. pp. 132-134, doi: https://doi.org/10.30534/ijiscs/2019/31822019

Ahmad, A.K., et al. (2019). Customer churn prediction in telecom using machine learning in big data platform. J Big Data 6, 28 https://doi.org/10.1186/s40537-019-0191-6

Umayaparvathi, V., & Iyakutti, K. (2012). Applications of Data Mining Techniques in Telecom Churn Prediction. International Journal of Computer Applications, 42, 5-9.

References.

Ewieda, M., Shaaban, E.M., & Roushdy, M.I. (2021). Review of Data Mining Techniques for Detecting Churners in the Telecommunication Industry. Future Computing and Informatics Journal.

Bugajev, A., Kriauziene, R., Vasilecas, O., & Chadysas, V. (2022). The Impact of Churn Labelling Rules on Churn Prediction in Telecommunications. Informatica, 33, 247-277.

Kumar, M.N., & Naik, M.C. (2017). Comparative Analysis of Machine Learning Algorithms for their Effectiveness in Churn Prediction in the Telecom Industry.

Aldalan, A. & Almaleh. A. (2023). Customer Churn Prediction Using Four Machine Learning Algorithms Integrating Feature Selection and Normalization in the Telecom Sector. World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering. Vol. 17. Issue 3. pp. 76-83.

References.

Paulose. J. et al.(2021). Application of Machine Learning in Customer Churn Prediction," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia., pp. 1-7, doi: 10.1109/i-PACT52855.2021.9696771.

Kaur, I. & Kaur, J. (2020). Customer Churn Analysis and Prediction in Banking Industry using Machine Learning. Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, India, 2020, pp. 434-437, doi: 10.1109/PDGC50313.2020.9315761.