

Name: Jacob Ajak Makuach Abuol (2k20/SE/63)

Nabiyu Samuel (2k20/SE/63)

Sub: Empirical Software Engineering lab G3

Experiment No: 1

Collection of empirical studies.

- Churn prediction in telecommunication using logistic regression and logit boost (International Conference on Computational Intelligence and Data Science (ICCIDS 2019))
- Customer churn prediction system using machine learning approach. (Accepted: 12 January 2021. The Author(s), under exclusive licence to Springer-Verlag GmbH, AT part of Springer Nature 2021)
- Deep Learning as a Vector Embedding Model for Customer Churn (5th International Conference on Computer Science and Computational Intelligence 2020)
- Why you should stop predicting customer churn and start using uplift models (Data Analytics Laboratory, Faculty of Economic and Social Sciences and Solvay Business School, Vrije Universiteit Brussel, Pleinlaan 2, Brussels B-1050, Belgium, Available online 31 December 2019)
- Customer Churn Prediction in Influencer Commerce: An Application of Decision Trees (The 8th International Conference on Information Technology and Quantitative Management (ITQM 2020 & 2021))

Experiment No: 2

Identification of research gaps from empirical studies

Papers read	Research gaps
1. Churn prediction in telecommunication using logistic regression and logit boost	✓ This model used two standalone techniques that performed well but standalone techniques cannot have all

	<p>features that can raise accuracy. There is a need to use hybrid model that can be implemented to embed the characteristics of two or more techniques to enhance the model performance.</p> <ul style="list-style-type: none"> ✓ The experiment was carried on a small database. The experiment needs to be improved by targeting a real and big database and will work on the model efficiency on the big database.
2. Customer churn prediction system using machine learning approach.	<ul style="list-style-type: none"> ✓ This approach tried to solve the problem in a very convenient way. with the upcoming concepts and frameworks in the field of reinforcement learning and deep learning sector, it is crucial to use such frameworks and approaches for better accuracy and precision.
3. Deep Learning as a Vector Embedding Model for Customer Churn (5th International Conference on Computer Science and Computational Intelligence 2020)	<ul style="list-style-type: none"> ✓ Despite the satisfying result of this study, the vector embedding model can potentially be improved by applying more advanced techniques. One of the promising techniques to be employed in this scenario is to use cosine-distance-based loss functions, which have been proved for face recognition to improve the discrimination among embedding vectors in different classes.
4. Why you should stop predicting customer churn and start using uplift models	<ul style="list-style-type: none"> ✓ There is a need for more studies that will focus on generalizing the newly introduced MPU measure to facilitate its adoption in different fields, as there is a need for powerful, application-oriented evaluation measures for assessing the performance of uplift models. ✓ In fact, the paper does not provide conclusive empirical evidence that uplift modelling outperforms predictive modelling therefore, further empirical evidence—obtained by examining a broader variety of application types—of

	the added value of uplift versus predictive modelling is necessary to validate the findings of this paper.
5. Customer Churn Prediction in Influencer Commerce: An Application of Decision Trees (The 8th International Conference on Information Technology and Quantitative Management (ITQM 2020 & 2021))	✓ This study attempts to predict the customer churn in influencer commerce by using Decision Trees (DT) algorithms. The analysis of the result demonstrates the feasibility of the method used. The only problem is that this method used one algorithm. Therefore, there is a need a to incorporate some relevant algorithms for enhancing the applicability.

Research questions

1. Why, how to use hybrid model to predict customer churn?
2. Why, how to use cosine-distance-based loss functions to predict customer churn?
3. Why, how to use deep learning model to predict customer churn?
4. Are there enough conclusive empirical evidence that uplift modelling outperforms predictive modelling therefore, further empirical evidence—obtained by examining a broader variety of application types—of the added value of uplift versus predictive modelling
5. Try using many algorithms to perform predictive analysis of customer churn and compare the results versus one algorithm.