INTE 43216 – Research Project

Short Research Proposal

Research Topic: A study on developing a hybrid recommender system for Telecommunication Industry in Sri Lanka

Aligning Research Area(s) of the Department: Recommender systems, Data engineering, Machine learning/AI, Sentiment analysis, text classifications and text mining

Submitted by

Student Number: IM/2016/046

Name: Chiran Hewawitharana

A short research proposal submitted in partial fulfilment of the requirements for the degree of B.Sc. (Honours) in Management and Information Technology

Department of Industrial Management

Faculty of Science

University of Kelaniya

Sri Lanka

April 2021

A study on developing a hybrid recommender system for Telecommunication Industry in Sri Lanka

Background

The telecommunication industry has become the center for digital growth and will continue to play the role of a disruptor. Today, terms like mobile, cloud, analytics broadband etc. have become a common lingo. According to [Central Bank, 2020], 10.9 million (34 per 100 people) internet users and 30.41 million mobile connections (1499 per 1000 people), equivalent to 141.7% of the total population, have been reported in Sri Lanka in January 2021. For example [Sri Lanka Telecom PLC, 2020] discusses that, the Sri Lanka Telecom PLC, a key player in Sri Lankan telecommunication industry, was having 9 million subscribers with a growing 91,119 million revenue by 2020.

With the advancements of big data, artificial intelligence and machine learning, a number of recommender systems have been implemented for retail, entertainment, social and other domains. In general, a recommender system uses historical data of purchases of products by other individuals to determine which to recommend to a particular customer. As [Ivens et.al., 2018] described, the field of recommender systems has its origins in the mid-1990s with the introduction of the Tapestry system. Content-based and collaborative filtering methods are popular techniques adopted in recommender systems. According to [Ahmed et. al., 2019] and [Khadiev et. al., 2016], collaborative filtering works by collecting user ratings for items in a given domain and calculating similarities between users or items in order to provide relevant recommendations. Content-based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or feedback. Each approach has both advantages and limitations, leading to several unsolved problems such as limited information retrieval, “cold-start”, “sparsity”, scalability problems, and overspecialization as described in [Ivens et.al., 2018] and [Ahmed et. al., 2019].

Studies on recommendation systems have not much applied yet in the telecommunication domain. [Chen C., 2016] presents a recommendation algorithm in the mobile environment, but does not mention the corresponding architecture suitable for a mass-scale service provider. As [Soft et. al., 2017] and [Chen C., 2016] identified, telecommunication data are highly asynchronous and significantly different compared to retail and other industries.

Research Problem

According to [Soft et. al., 2017], the increasing flourish of available telecommunication services offer more choices to the end user, leading some services to pass unobserved even if useful. In [Chen C., 2016] argues that, due to this high availability of options, subscriber churn has significantly increased. Acquiring new customers has become multiple times more expensive than retaining a customer. Therefore, segmentation of subscribers on their behavioral patterns and recommending services and offers to these identified beneficial groups of subscribers, will increase the business value while reducing subscriber churn. Although traditional recommender system approaches, where only the user-service interactions are incorporated, will not be able to meet these challenges. Therefore considering subscribers' opinions and reviews of consumed services, would be helpful to provide more effective recommendations.

Proposed Solution

In [Chen C., 2016], discussed that every telecommunication service provider in the industry is seeking new opportunities to increase operational and marketing efficiency with available big data and machine learning techniques. Therefore, the proposed solution will compare the available recommender algorithms and select the most appropriate stack of models to develop a hybrid recommender system to recommend telecommunication services. Moreover, it will focus on segmenting and identifying beneficial subscriber groups to provide these personalised recommendations, to empower long-term customer relationships with reducing churn rates.

Data will be collected from a particular telecommunication operator, e.g. Sri Lanka Telecom PLC, including user demographics, service purchases and usage histories, interactions, CRM logs, and related network KPIs. An elementary survey will be conducted among a stratified sample selected from the mobile and broadband internet subscribers in Sri Lanka, to extract their opinions on offered services, and need of such recommendations. Data integration, cleansing and preprocessing tasks will be carried out on raw data retrieved, including extract, transform, load (ETL), and reconciliation processes as described in [Bursha et. al., 2019].

The uplift modeling technique will be used for subscriber segmentation considering their service usage patterns. As [Ahmed et. al., 2019] presented, it is a predictive modelling technique which calculates the probability of a potential subscriber response with positive impact, when a recommendation is provided, so that focus can be given to the subscribers with higher potential. Subscribers will be segmented as, “sure things” (who will purchase no matter what), “persuadables” (who will purchase only if a recommendation is given), “lost causes” (who will not purchase no matter what), and “sleepers” (who will not purchase even a recommendation is given).

According to [Sundermann et. al., 2019] considering rich information embedded in users’ reviews into the recommender systems can produce more precise recommendations. Extracting user opinions from reviews (opinion mining) can be achieved through applying Natural Language Processing(NLP) based text analysis techniques to available CRM logs, and/or social media to gauge the sentiment about specific characteristics of a service consumed, which can be considered for the proposed model, when recommending telcom services.

As [Ivens et.al., 2018] and [Yousef et. al., 2018] described, bayesian classifiers and decision trees are widely used in content-based approaches, while neighborhoods or the latent factor models, (e.g. matrix factorization, ALS) were used in collaborative filtering approaches. According to [Soft et. al., 2017] and [Afsar et. al., 2021] associative classification, Artificial Neural Networks(ANN), and Reinforcement Learning based agents have been proposed recently for recommender models. As [Ivens et.al., 2018] suggested, a confusion matrix, which illustrates the accuracy of the solution to a classification problem with the information about actual and predicted results, can be used to compare the performance of these algorithms. Then the most appropriate algorithms will be selected and improved. Therefore, an ensemble modeling approach is proposed to implement this hybrid recommender system. According to [Aggarwal CC, 2016], using ensemble modeling, the resulting outputs from the content-based, collaborative, and context-aware(opinion based) recommenders can be combined into a single and more robust output. Further, a web-based system with a REST API will be developed to demonstrate the predictions, insight visualizations and facilitate querying for recommendations for a specific user or a user group.

References

[Afsar et. al., 2021] Afsar, M. Mehdi, et al. Reinforcement Learning Based Recommender Systems: A Survey. *University of Calgary, Calgary, AB Canada*, deepai.org/publication/reinforcement-learning-based-recommender-systems-a-survey, 15 Jan. 2021.

[Ahmed et. al., 2019] Ahmed, A.A.Q., Maheswari, D. An enhanced ensemble classifier for telecom churn prediction using cost based uplift modelling. *Int. j. inf. tecnol. 11*, 381–391, 2019.

[Aggarwal CC, 2016] Aggarwal CC (2016) Ensemble-based and hybrid recommender systems. *In: Recommender systems, Springer International Publishing Switzerland, pp 199–224, 2016*

[Bursha et. al., 2019] Bushra Ramzan, Imran Sarwar Bajwa, Noreen Jamil, Riaz Ul Amin, Shabana Ramzan, Farhan Mirza, Nadeem Sarwar, "An Intelligent Data Analysis for Recommendation Systems Using Machine Learning", *Scientific Programming, vol. 2019*, Article ID 5941096, 20, 2019.

[Chen C., 2016] Chen, C. (2016). Use cases and challenges in telecom big data analytics. *APSIPA Transactions on Signal and Information Processing*, 5, E19, 2016.

[Central Bank, 2020] Central Bank of Sri Lanka, Sri Lanka Socio Economic Data 2020, *Statistics Department, Central Bank of Sri Lanka,* ISSN : 1391 – 3557, August 2020.

[Ivens et.al., 2018] Ivens Portugal, Paulo Alencar, and Donald Cowan. The use of machine learning algorithms in recommender systems: A systematic review. E*xpertSystems with Applications 97 (2018)*, 205–227, 2018.

[Khadiev et. al., 2016] Khadiev, K.; Zulfat, M.; Sidikov, M. Collaborative filtering approach in adaptive learning.Int. *J.Pharm. Technol*, 8, 15124–15132, 2016.

[Soft et. al., 2017] Soft, M., Zulu, D., & Mazhandu, R. Recommender System for Telecommunication Industries: A Case of Zambia Telecoms. *American Journal of Economics*, 7, 271-273, 2017.

[Sri Lanka Telecom PLC, 2020] Sri Lanka Telecom PLC. 2020 Annual Report Sri Lanka Telecom PLC. [online] Available at: <https://www.slt.lk/en/about-us/investors/financial-reports/annual> [Accessed 23 March 2021]

[Sundermann et. al., 2019] Sundermann, C.; Domingues, M.; Sinoara, R.; Marcacini, R.; Rezende, S. Using Opinion Mining in Context-Aware Recommender Systems: A Systematic Review.*Information 2019*, 10, 42, 2019.

[Yousef et. al., 2018] Yousef Kilani, Ahmed Fawzi Otoom, Ayoub Alsarhan, Manal Almaayah, A genetic algorithms-based hybrid recommender system of matrix factorization and neighborhood-based techniques, *Journal of Computational Science, Volume 28, Pages 78-93*, ISSN 1877-7503, 2018.

[Zui Zhang et. al., 2013] Zui Zhang, Hua Lin, Kun Liu, Dianshuang Wu, Guangquan Zhang, Jie Lu, A hybrid fuzzy-based personalized recommender system for telecom products/services, *Information Sciences, Volume 235, Pages 117-129*, ISSN 0020-0255,, 2013.