Analysing Digital Banking Reviews Using Text Mining

Li Chen Cheng,
Department of Information and Finance
Management
National Taipei University of
Technology
Taipei, Taiwan
Lijen.cheng@gmail.com

Legaspi Rhea Sharmayne
Department of Information and Finance
Management
National Taipei University of
Technology
Taipei, Taiwan
Lijen.cheng@gmail.com

Abstract—**Digital banks are new entrants in the banking industry in the Philippines as they only started late 2018. Since then, a handful of players have and are still emerging. With more and more people becoming technologically savvy, it is very critical for financial institutions to develop a digital banking application that will stand out from the competition. This paper aims to use text mining methods to analyse digital banking application reviews. This study will perform topic modelling using LDA to explore customer concerns and will mine association rules between the digital banking features with the review score. The results will reveal which areas the digital banking application can further optimize for customer satisfaction and retention.**

Keywords—Online Reviews, Text Mining, LDA, Association Analysis

I. INTRODUCTION

With the advent of the internet and social media, information has become more widespread and easier to access. In the case of product reviews, a number of platforms have been made available to consumers to provide them with the needed information before they decide to avail of a company's service or product. (Kim and Chun, 2019). For mobile applications, users usually give the application a rating and provide feedback on their experience while using the application after downloading. (Al-Hawari et al, 2020). The feedback provided by the users usually reflect their satisfaction on the application. Satisfaction has a direct impact on the user's loyalty (Wang et al, 2018).

This has paved the way for companies to innovate marketing strategies to keep up with a more competitive environment. Now, more than ever, it is critical for companies to understand their market. Consumer reviews is one of the avenues that the company can get to know their customers however due to the unstructured nature and the huge number of user reviews, it may be difficult for companies to keep track of all of them and some may not even provide them with meaningful insights. (Al-Hawari et al, 2020). Furthermore, Martens and Johann (2017) has previously stated in their research that the emotional sentiment of the user is only weakly correlated to the app rating, which proves that the app rating is not an accurate representation of the customers' sentiment.

The purpose of this paper is to obtain valuable knowledge that was initially unrevealed in online reviews.

First, this research will apply LDA to pre-processed text reviews to obtain the relevant topics and supporting related terms. Second, keywords are extracted from the text based on their count frequency and labelled into specific digital banking features. The reviews are then divided into two clusters based on the current review rating ("positive" and "negative") and association rule mining is applied to determine the relationship between the above-mentioned features and the given rating.

This paper aims to answer these two research questions:

1)   What are users most concerned about?

2)   Which features are viewed positively and which are viewed negatively?

Analysing which features in digital banking applications are interested in has significant managerial implications in the digital finance industry as developers will be able to identify which features need to be prioritized and improved upon. By understanding the views and opinions of the users, they will be able to focus their efforts and better target their strategies to meet customers' expectations.

II. LITERATURE REVIEW

A. Latent Dirichlet Allocation (LDA)

To delve deeper into consumer insights, topic modelling is usually applied to understand the factors that drive customer satisfaction. (Sutherland and Kiatkawsin, 2020). LDA is a popular topic modelling technique under natural language processing. In LDA, each text corpus will be structured in a number of topics, with each topic containing similar terms that are correlated with a certain likelihood. (Guzman and Maalej, 2014)

The topics extracted by the algorithm help to illustrate which of the topics are relevant to the customers. The customers usually only write about their most unforgettable experiences with the product therefore, if the subject is not relevant, they also will not share any opinion about it. Moreover, the algorithm will only extract the words that are frequently mentioned. (Kiatkawsin et al., 2020).

Previous studies have used LDA in a number of industries. Among which are the (a) tourism industry: Song et al. (2020) aimed to understand the perceptions and experiences of citizens of an urban park in New York City while Yan et al. (2020) used topic modelling to identify the issues with regards to post-disaster tourism in Lombok and Bali; (b) health and safety industry: Bahng and Lee (2020) examined patients' concerns regarding hearing loss and Min et al. (2020) analysed the issues on occupational accidents; (c) sharing economy industry: Kiatkawsin et al. (2020) provided new findings on Airbnb guests' experiences while Sutherland and Kiatkawsin (2020) studied the relevant topics that drive customer satisfaction in the accommodation sector.

B. Association Analysis

Association Analysis was first introduced by Agrawal et al. (1993) to find the transaction patterns in data. Association rules are extracted from the frequent patterns in a given itemset. It aims to explore the correlations of certain

.

transactions while complying with minimum support and maximum confidence constraints (Wu and Lei, 2019).

III. METHODOLOGY
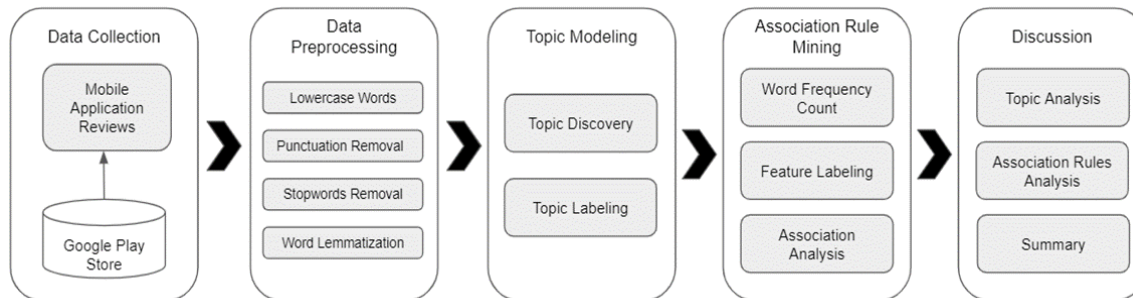
A. Data Collection

For this study, a digital banking application from the Philippines named CIMB Bank PH was chosen as the dataset. CIMB Bank PH. The application prides itself on being the country's "first all-digital and mobile -first bank."

Reviews from the digital bank application were collected from the Google Play Store Website and mined through the Google Play Scraper package in Python. The scraper provides an easier way to scrape data from the Google Play Store by providing APIs without external dependencies. [1] After scraping the data, variables such as: (1) Review ID, (2) Username, (3) User Image, (4) Content, (5) Score, (6) Thumbs Up Count, (7) Review Created Version, (8) At, (9) Reply Content, (10) Replied At, (11) Sort Order, (12) App ID were obtained as shown in Fig 2.

All results were exported and saved in a CSV file. A summary of the datasets is shown in Fig. 3.

Application	Number of Reviews	Review Date Period
CIMB PH	10,583	Dec 2018 - Sept 2020*

Fig. 3 *Data Description (The review period starts from the application's first launch in the Playstore)*

B. Data Cleaning and Data Pre-processing through Text Mining

To further analyse the data, only the following variables will be utilized: (1) Content, and (2) Score. All other variables will not be considered in this study.

To pre-process the data, text mining techniques were applied with the use of Python. The Natural Language Toolkit (NLTK) Platform was utilized. First, all review content was transformed into lowercase characters to avoid redundant words. Next, all punctuation marks were removed from the text as these were unnecessary information. Furthermore, stop words were also removed from the text using the default NLTK package' English stopwords corpus coupled with a predefined list of stopwords manually identified. Lastly, word lemmatization was applied using the Textblob package. Lemmatization transforms the words in the sentence into root words to make it easier to analyse.

C. Topic Discovery via Latent Dirichlet Allocation (LDA)

To extract the topics behind the reviews, Topic Modelling was adopted with the use of the LDA model tool from the Gensim package. To find the optimal number of topics, a CV Coherence analysis was conducted.

■ Topic Modelling and Analysis

From the results, all observed words were evaluated and interpreted based on the terms contained in each group and were given a topic label.

■ Word Frequency Count / Keywords Identification

Next, the top 1000 frequently occurring words were extracted from the reviews and were ranked and labelled according to its closely resembled banking feature.

■ Feature Labelling

All keywords were matched with the text reviews. These observed words were compared with the original review content and only those reviews containing the above-mentioned keywords were retained and labelled with the corresponding feature. After manual inspection, only 9,670 reviews remained out of 10,583.

■ Association Analysis

To acquire the association rules of each dataset, the features identified in each review sentence were then tokenized and separated into each data column.

The transformed review content was also labelled based on the score. Reviews with a 4-5 rating were deemed as "positive", while reviews with a 1-3 rating were deemed as "negative".

$$\text{If review rating is} >= 4, L = \text{positive}$$
$$\text{If review rating} <= 3, L = \text{negative}$$

Where L is the group label. After labelling all reviews, the results are as follows:

Positive	Negative
5,953	3,717

In association analysis, we usually refer to the support, confidence and lift metrics to quantify the strength of association between two items. Support is calculated as follows:

$$Support\ of\ XY = \frac{Transactions\ with\ both\ X\ and\ Y}{Total\ Number\ of\ Transactions} \quad (1)$$

In equation (1), support represents how often both X and Y occurs in the dataset. Confidence in equation (2), on the other hand, is a measurement of the likelihood that the consequent (Y) will occur given the antecedent (X).

$$Confidence\ of\ XY = \frac{Transactions\ with\ both\ X\ and\ Y}{Transactions\ with\ X}$$
$$(2)$$

In equation (3), lift refers to how probable the consequent (Y) will appear knowing that the antecedent (X) occurs divided by the possibility of Y without knowing that X is present. It can also be summarized as the confidence of XY divided by the support of Y.

$$Lift\ of\ XY = \frac{Transactions\ with\ both\ X\ and\ Y / Transactions\ with\ X}{Fraction\ of\ Transactions\ with\ Y}$$
$$(3)$$

Wu and Lei (2019) considers the lift to be a better measure of the model as this metric assumes that the occurrence of an itemset does not depend on the existence of the antecedent. For this study, we will utilize all measures, but will only mine association rules with lift greater than 1.

Association analysis was performed using the Efficient-Apriori Package, an Apriori Algorithm implementation in Python [3].

$$a + b = \gamma \quad (1)$$

[7].

IV. RESULTS

The chosen model was 7 topics with the highest coherence score of 0.4308.

Topic No.	Keywords	Topic Label
1	*loan, application, time, approval, pending*.....	Loan Experience
2	*user-friendly, convenient, fast, amazing, friendly*.....	General Application Usage Experience
3	*customer, service, poor, phone, email*.....	Customer Service Experience
4	*account, password, username, login, error*.....	Account Login Experience
5	*verification, process, selfie, smooth, difficult*.....	Verification Experience
6	*deposit, withdraw, transfer, credit, problem*.....	Experience with Banking Transactions (Deposit, Transfers, Withdrawal)
7	*interest, rate, better, high, happy*.....	Interest Rate Feedback

Topic 1 shows the words "loan", "application", "approval", "pending", "waiting". From these words, we can assume that the customers' main concern for the loan application is the processing time.

Topic 2 contains "user-friendly", "convenient", "fast", "amazing", "reliable". These words represent how the customers generally view the mobile application. The application could be designed in a very friendly manner that makes it easy for the customers to use.

Topic 3 with words such as "customer", "service", "poor", "waste", "phone", "email" reveals that the usual contact methods of the users are through phone and email. It also reveals that the application's customer service is currently unsatisfactory.

Topic 4 represents the account log in experience of the user with words such as "account", "password", "number", "username", "error". These are usually problems that arise from external factors such as if the user has a stable internet connection or possibly from the user forgetting the password or login details. It does not directly represent the application functions.

Topic 5 contains "verification", "process", "selfie", "smooth", "difficult". The verification process of the application resulted in mixed reactions and would need to be analysed further.

Topic 6 with words like "account", "money", "deposit", "withdraw", "credit", "problem", "responsive" represents the banking transaction features of the application. Same with

Topic 5, it also resulted in mixed reactions and will need to be examined further.

Topic 7 contains "good", "great", "interest", "rate", "high" which discloses that in general, users are very satisfied with the interest rate offered by the digital bank..

A. Association Analysis Results

To further explore which features are usually perceived as positive or negative, association rule mining was conducted to examine the association between the features with the review score..

LHS	RHS	Support	Confidence	Lift
Loan, Application	Negative	0.023	0.906	2.356
General, Application	Negative	0.082	0.875	2.277
Deposit, Application	Negative	0.022	0.872	2.268
General, Loan	Negative	0.024	0.857	2.229
Interest Rate, Withdrawal	Positive	0.005	0.875	1.421
Deposit, Interest Rate	Positive	0.010	0.844	1.371
Interest Rate, General	Positive	0.027	0.835	1.356
Transfer, General	Negative	0.021	0.488	1.270
Payment, Withdrawal	Positive	0.011	0.711	1.156
Interest Rate, Loan	Positive	0.007	0.707	1.148
Application, Interest Rate	Positive	0.007	0.703	1.142
Deposit, General	Negative	0.041	0.423	1.102
Withdrawal, General	Positive	0.025	0.671	1.091
Deposit, Withdrawal	Positive	0.012	0.625	1.015

The Apriori Algorithm was utilized and all rules with a lift greater than 1 were mined. From the results, we can observe that most features including "application", "loan", "transfer", "deposit", and "payment" are associated with a negative rating with a few exceptions, while only "interest rate" and "withdrawal" can be generally attributed to a positive rating.

V. CONCLUSION

Based on the LDA and association rule results, there seems to be great interest in the initial adoption of digital banking especially due to the fact that it provides higher interest rates compared to traditional banks. The fact that the application provides users with a more convenient way to access their money anywhere also seems to be a commendatory point. However, the registration/verification process and the current bank offerings may need to be further improved especially in relation to the deposit, transfer, and payment features. As digital banks also do not have any physical branch, providing better customer service is also a significant issue to focus on to troubleshoot customer problems.

For future research, other methods can be adopted to analyse and examine the relationship between text reviews and the review score. As this study only used one bank, competitor banks can also be studied to gain different perspectives.

REFERENCES

[1] (September). Available: https://github.com/JoMingyu/google-play-scraper

[2] (November). Available: https://github.com/tommyod/Efficient-Apriori

[3] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in Proc. 20th int. conf. very large data bases, VLDB, 1994, vol. 1215, pp. 487-499.

[4] A. Al-Hawari, H. Najadat, and R. Shatnawi, "Classification of application reviews into software maintenance tasks using data mining techniques," (in English), Software Quality Journal, Article; Early Access p. 37.

[5] J. Bahng and C. H. Lee, "Topic Modeling for Analyzing Patients' Perceptions and Concerns of Hearing Loss on Social Q&A Sites: Incorporating Patients' Perspective," (in English), International Journal of Environmental Research and Public Health, Article vol. 17, no. 17, p. 14, Sep 2020, Art. no. 6209.

[6] A. Garg. (2018, November 30). Complete guide to Association Rules Available: https://towardsdatascience.com/association-rules-2-aa9a77241654

[7] J. Guerreiro and P. Rita, "How to predict explicit recommendations in online reviews using text mining and sentiment analysis," (in English), Journal of Hospitality and Tourism Management, Article vol. 43, pp. 269-272, Jun 2020.

[8] E. Guzman and W. Maalej, "How do users like this feature? a fine grained sentiment analysis of app reviews," in 2014 IEEE 22nd international requirements engineering conference (RE), 2014, pp. 153-162: IEEE.

[9]   K. Kiatkawsin, I. Sutherland, and J. Y. Kim, "A Comparative Automated Text Analysis of Airbnb Reviews in Hong Kong and Singapore Using Latent Dirichlet Allocation," (in English), Sustainability, Article vol. 12, no. 16, p. 17, Aug 2020, Art. no. 6673.

[10]   E. G. Kim and S. H. Chun, "Analyzing Online Car Reviews Using Text Mining," (in English), Sustainability, Review vol. 11, no. 6, p. 22, Mar 2019, Art. no. 1611.

[11]   S. W. Lee, G. B. Jiang, H. Y. Kong, and C. Liu, "A difference of multimedia consumer's rating and review through sentiment analysis," (in English), Multimedia Tools and Applications, Review; Early Access p. 18, 2020.

[12]   D. Martens and T. Johann, "On the emotion of users in app reviews," in 2017 IEEE/ACM 2nd International Workshop on Emotion Awareness in Software Engineering (SEmotion), 2017, pp. 8-14: IEEE.

[13]   K. B. Min, S. H. Song, and J. Y. Min, "Topic Modeling of Social Networking Service Data on Occupational Accidents in Korea: Latent Dirichlet Allocation Analysis," (in English), Journal of Medical Internet Research, Article vol. 22, no. 8, p. 12, Aug 2020, Art. no. e19222.

[14]   Y. Song, J. Fernandez, and T. Wang, "Understanding Perceived Site Qualities and Experiences of Urban Public Spaces: A Case Study of Social Media Reviews in Bryant Park, New York City," (in English), Sustainability, Review vol. 12, no. 19, p. 15, Oct 2020, Art. no. 8036.

[15]   I. Sutherland and K. Kiatkawsin, "Determinants of Guest Experience in Airbnb: A Topic Modeling Approach Using LDA," (in English), Sustainability, Article vol. 12, no. 8, p. 16, Apr 2020, Art. no. 3402.

[16]   D. D. Tao, P. K. Yang, and H. Feng, "Utilization of text mining as a big data analysis tool for food science and nutrition," (in English), Comprehensive Reviews in Food Science and Food Safety, Review vol. 19, no. 2, pp. 875-894, Mar 2020.

[17]   Y. R. Wang, X. Lu, and Y. J. Tan, "Impact of product attributes on customer satisfaction: An analysis of online reviews for washing machines," (in English), Electronic Commerce Research and Applications, Review vol. 29, pp. 1-11, May-Jun 2018.

[18]   S. A. H. Wu and X. X. N. Lei, "The Analysis of the Influencing Factors on the Problems of Bike-Sharing System in China," (in English), Ieee Access, Article vol. 7, pp. 104000-104010, 2019.

[19]   Y. W. Yan, J. F. Chen, and Z. Y. Wang, "Mining public sentiments and perspectives from geotagged social media data for appraising the post-earthquake recovery of tourism destinations," (in English), Applied Geography, Article vol. 123, p. 13, Oct 2020, Art. no. 102306.