# Exploring Touch 'n Go eWallet user reviews using Stopwords and N-gram

Loh Wan Teng School of Computer Science Universiti Sains Malaysia (USM) Penang, Malaysia wantengloh@student.usm.my Lim Phei San School of Computer Science Universiti Sains Malaysia (USM) Penang, Malaysia pheisanlim@student.usm.my

Abstract—Touch 'n Go eWallet is a Malaysian digital wallet and online payment platform. The adoption of cashless transactions due to emergence of Covid-19 and variety of services provided by Touch 'n Go has make itself the most popular eWallet in Malaysia. This study performs sentiment analysis on Touch 'n Go eWallet user reviews using Natural Language Processing techniques: stop words and n-gram. A total of 35458 user reviews from 21st March 2018 to 25th November 2022 are scrapped from Google Play Store. The sentiment analysis can be served as a reference to improve the application. The results provide positive and negative user reviews of Touch 'n Go.

Keywords—sentiment analysis, Natural Language Processing, Stopwords, Concordance, N-gram, eWallet

# I. Overview

Covid-19 once plunged Malaysia into an economic recession. To facilitate customer spending while creating a secure, contactless payment experience, Malaysia's government has allotted a total of RM750 million to develop the e-wallet business in order to benefit 15 million Malaysians. E-wallets have gained popularity because of the government's action, especially Touch 'n Go eWallet, which is now widely used by Malaysians.

Touch 'n Go eWallet is a joint venture between Touch 'n Go and Ant Financial that was established in July 2017. It is the most popular eWallet in Malaysia with over 15 million registered users and over 689 thousand merchants cooperating on cashless payment via QR code. With its attractive features, including peer-to-peer money transfer, GO+, utility bill payment, car parking payment and mobile prepaid top-up, by January 2021, it had 4.6 and 4.7 user ratings on Google Play and App Store respectively.

The reasons Touch 'n Go eWallet is being widely used are Touch 'n Go eWallet allows users to transfer funds instantly to other people who has the Touch 'n Go eWallet account. Not only transfer to another Touch 'n Go eWallet account, but users are also able to transfer funds to bank account or other wallets directly through DuitNow and all the transaction has no minimum transfer amount and no transaction fee will be charged. [1]

To investigate the reasons for Touch 'n Go eWallet popularity, this study is carried out to identify users' satisfaction toward Touch 'n Go eWallet. The user reviews can be further used for strategy analysis for retaining customers and to improve shortcomings of the application.

The study uses 35458 Touch 'n Go user reviews (21st March 2018 to 25th November 2022) extracted from Google Play Store. Text preprocessing is performed to improve the dataset for sentiment analysis on users' experience with Touch 'n Go using concordance and n-gram.

#### II. LITERATURE REVIEW

G.S.N. Murthy and his team used Concordance to search for the Application Programming Interface (API) call sequence to detect malware behaviors. The authors created the dataset by selecting several samples from the files managed by several cybersecurity companies. In their study, Concordance is easier to use than n-gram because n-gram lists all possible outcomes based on the *n* value, which displayed a lot of outputs compared to concordance which only returned the results based on the queried keywords. With n-gram, the time taken to display output is longer than that of concordance which is crucial for malware detection. However, Concordance has its disadvantage when keyword is not known if the API call sequence for malicious behavior is not known. Comparing the performance of malware behavior detection methods, Concordance key word in context (KWIC) can identify more malicious behaviors than N-gram approach. [2]

The focus of another research described in paper [3] is on using n-gram representations of Web pages to automatically categorize them by genre. In this study, two sets of experiments were run to examine how Web sites were automatically categorized into genres. The results of this study show that smaller n-gram lengths outperform bigger n-gram lengths, with an optimal n-gram length of 2 for both the SVM and distance function approaches. An innovative, computationally effective, and unsupervised C-BOWinspired technique for training and inferring language embeddings has been proposed by Pagliardini, Gupta, and Jaggi. This method outperforms state-of-the-art unsupervised models on the majority of benchmark tests, highlighting the dependability of the created general-purpose sentence embeddings. Sent2Vec, a straightforward approach that enables sentence embedding composition using word vectors and n-grams has been suggested. [4] Research in the future might focus on enhancing the model's capacity to use data containing ordered sentences. The study suggests a strategy for identifying hot topics that assesses the topic's hotness using the ngram model. The importance of the keywords is used to gauge how hot a topic is. The results show that the proposed strategy yields promising results after being tested on a Chinese corpus. [5]

# III. METHODOLOGY

#### A. Data Collections

A total of 38458 Touch 'n Go user reviews dataset is collected from Google Play Store from 21<sup>st</sup> March 2018 to 25<sup>th</sup> November 2022. Reviews with 4- or 5- star ratings are considered positive while reviews with 1- or 2- star ratings are categorized as negative sentiment. Reviews with 3-star ratings are categorized as neural reviews and are not taken into consideration for the sentiment analysis in this study.

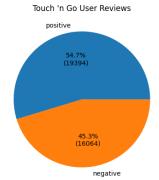
# B. Text Preprocessing

The following text preprocessing is done to eliminate noises in the dataset: removing emojis, removing punctuations and lowering word cases to eliminate discrimination between uppercase and lowercase letters.

### C. Categorize reviews

The reviews with 4- and 5- star ratings are categorized as positive sentiment while the reviews with 1- and 2- star ratings are categorized as negative ratings.

FIGURE I. Touch 'n Go User Reviews



54.7% of the user reviews are positive reviews and another 45.3% user reviews are negative.

#### IV. RESULTS AND DISCUSSION

# A. Concordance

The team had explored the positive and negative reviews and observed that the stop words (especially in negative reviews) occupied most of the ranks in both types of reviews as shown in **Figure II**.

FIGURE II. TOP 15 WORDS IN POSITIVE & NEGATIVE REVIEWS

	Type	Rank	Freq	Range		Type	Rank	Freq	Range
1	good	1	6224	1	1	i	1	10865	1
2	to	2	3344	1	2	to	2	10309	1
3	the	3	2203	1	3	the	3	9795	1
4	i	4	2138	1	4	my	4	7131	1
5	and	5	1946	1	5	and	5	5261	1
6	арр	6	1919	1	6	it	6	4899	1
7	easy	7	1647	1	7	арр	7	4767	1
8	very	8	1469	1	8	this	8	4206	1
9	use	9	1411	1	9	is	9	4055	1
10	for	10	1335	1	10	in	10	3732	1
11	it	11	1303	1	11	cant	11	3626	1
12	ok	12	1132	1	12	not	12	3600	1
13	this	13	1074	1	13	but	13	3483	1
14	my	14	1063	1	14	for	14	3263	1
15	is	15	1014	1	15	card	15	3022	1

(Positive Reviews)

(Negative Reviews)

Some interesting words such as "good", "app", "ok", and "card" aid in describing the user reviews in respective corpus. To find out more interesting words in the corpus, the "Keyword List" tool was used to provide the team to compare the current wordlist and the generic wordlist.

Note that the "Keyness (Likelihood)" implies that the target word occurs more frequently in the selected corpus compared to the general corpus.

# a) Positive Sentiment

**Figure III** shown below presents more useful words in expressing users' positive experience with Touch 'n Go, particularly the word "good" outperform the rest of the words by at least 5 times.

FIGURE III. TOP 15 POSTIVE REVIEWS WORDS IN KEYLIST

	Туре	Rank	Freq_Tar	Freq_Ref	Range_Tar	Range_Ref	Keyness (Likelihood)	Keyness (Effect)
1	good	1	6224	6910	1	1	6608.059	0.104
2	easy	2	1647	1815	1	1	1724.828	0.030
3	ok	3	1132	1461	1	1	1002.313	0.021
4	nice	4	903	984	1	1	952.789	0.017
5	very	5	1469	2428	1	1	946.713	0.027
6	best	6	633	703	1	1	654.930	0.012
7	great	7	578	686	1	1	559.090	0.011
8	convenient	8	532	661	1	1	489.969	0.010
9	terbaik	9	424	433	1	1	474.299	0.008
10	excellent	10	370	384	1	1	407.747	0.007
11	use	11	1411	3770	1	1	335.903	0.025
12	love	12	292	316	1	1	309.306	0.005
13	awesome	13	234	246	1	1	254.670	0.004
14	fast	14	280	356	1	1	251.113	0.005
15	useful	15	291	395	1	1	242.026	0.005

To figure out how "good" the users' experiences are with Touch 'n Go, concordance is used to analyze the word "good" as shown in **Figure IV**.

FIGURE IV. CONCORDANCE FOR "GOOD"

File	Left Context	Hit	Right Context
pos_corp	s led services with additional feature like go investment for future	good	good good mantap terbaik super fast transaction good
pos_corp	s b c great apps nice app love it very wel	good	good good excellant good best easy n convenient
pos_corp	s terbaik sangat memudahkan urusan kewangan kekalkan prestas	good	good good no comment good good just try
pos_corg	s i wont be hack may god bless you good oi	good	good good apps for saving good not bad
pos_corp	s good excellent app ok so okay so good senam guna	good	good good good fast and easy very good
pos_corp	s excellent app ok so okay so good senam guna good	good	good good fast and easy very good apps
pos_corg	s reinstall the app good ok good good buy stuff is	good	good good nice ill already download these apps
pos_corp	s good good app very goods nice good kedekut dpt ser	good	good good app gpod pls send me weed
pos_corp	s dalam go apabila baki ewallet tak mencukupi bila lalu to	good	good good good good very good ok lah
0 pos_corp	s go apabila baki ewallet tak mencukupi bila lalu tol good	good	good good good very good ok lah bih
1 pos_corp	s apabila baki ewallet tak mencukupi bila lalu tol good good	good	good good very good ok lah blh je
2 pos_corp	s payah bawah duit bayakitu good good good very nice and	good	good good luck the best app please put

From **Figure IV**, it is obvious that the word "good" is repeated in immediate succession (called Epizeuxis) which results in the abnormally high frequent of "good" word. Besides "good", one may observe words like "terbaik", "super fast", "excellent", "nice" and "best" are associated with sentences containing the "good" word indicating the positive experience users were having with Touch 'n Go.

FIGURE V. CONCORDANCE FOR "EASY"

	File	Left Context	Hit	Right Context
1	pos_corpus	quick and easy good saya sangat suka very convenient and	easy	to use currently when i put in an account
2	pos_corpus	everywhere touch n go ok easy trendy good easy user	easy	to use cashless more benefit ok k good supper
3	pos_corpus	in discount voucher not everyday faster services and transaction	easy	to use system saya suka super convenient to pay
	pos_corpus	good and convenient very good the best apps ever very	easy	to use this apps k laaa gud thorbaekkk easy
5	pos_corpus	app time save excellent ok good excellent and good service	easy	to use rasa lebih selamat untuk urusan online daripada
5	pos_corpus	ops good workinhg good easy and convenient cepat dan cergas	easy	to use fast and convenient option to choose the
,	pos_corpus	not scanned in the dark theme good service perfect good	easy	to use really recommended new coner nice easy supar
3	pos_corpus	best e wallet in play store good mantap cok best	easy	to use regarding the parking why dont you create
9	pos_corpus	tap and go simple ok good job nice very convenient	easy	to use many functions available and background images eas
10	pos_corpus	nt easy to use many functions available and background images	easy	to use good best ewallet good ezey to use
1	pos_corpus	to use good best ewallet good ezey to use if	easy	to use and make all the online works so
12	pos_corpus	anks cards salam good morning very good and convenient very	easy	to use and easy to use and easy to

The second most frequent word is "easy", from **Figure V**, the word "easy" is always followed by "to use" indicating the app is easy to use and provides user-friendly interface and services.

# b) Negative Sentiment

**Figure VI** shown below presents the top 15 users' negative experience words with Touch 'n Go. Compared to the top 15 positive reviews words in **Figure III**, the negative reviews did

not show a clear picture of how users reacted towards the app except the words "useless" and "stupid".

FIGURE VI. TOP 15 NEGATIVE REVIEWS WORDS IN KEYLIST

	Type	Rank	Freq_Tar	Freq_Ref	Range_Tar	Range_Ref	Keyness (Likelihood)	Keyness (Effect)
1	cant	1	3626	4322	1	1	76.508	0.020
2	my	2	7131	9171	1	1	59.313	0.038
3	even	3	1403	1603	1	1	42.602	0.008
4	useless	4	754	788	1	1	41.314	0.004
5	log	5	1319	1512	1	1	39.007	0.007
6	cannot	6	1731	2050	1	1	38.652	0.010
7	number	7	2273	2776	1	1	37.135	0.01
8	i	8	10865	14637	1	1	35.399	0.05
9	stupid	9	606	630	1	1	34.182	0.00
10	account	10	1999	2450	1	1	31.366	0.01
11	not	11	3600	4615	1	1	31.217	0.020
12	tak	12	1423	1699	1	1	29.372	0.008
13	call	13	591	644	1	1	25.144	0.003
14	register	14	848	975	1	1	24.451	0.00
15	pin	15	987	1156	1	1	24.336	0.00

Words such as "cant", "log", "number", "account", "call", "register" and "pin" might have different semantic meanings instead of showing strong negative sentiment. Hence, concordance is performed to further analyze these words.

FIGURE VII. CONCORDANCE FOR "CANT"

40	neg_corpus	kali pun masih tiada respon i have problem which is i	cant	even log into my acc the otp doesnt even got
41	neg_corpus	dont get it such a mistake for dwld this app i	cant	even log into my account i never ewallet before so
42	neg_corpus	ogin saying that mobile number does not exist correct pin but	cant	even log in verification part answered the question with right
43	neg_corpus	and to transfer to touch n go card download the app	cant	even log inwhile i can log in as normal on
44	neg_corpus	mu nuber already exist why can you take a look i	cant	even log to my account i try to forget my
45	neg_corpus	qr code for tng e wallet to receive payment pls help	cant	even register an account it always said something like auto
46	neg_corpus	i cant log in into my account please hire qa i	cant	even register an account it says account doesnt exist i
47	neg_corpus	company still in stone age why there is no transfer icon	cant	even register an account can you pls tell me why
48	neg_corpus	them please improve thank you seldom give star but this app	cant	even register and work properly i cant get my topup
49	neg_corpus	call sc xde dlm apps suck i dont know why i	cant	even register and login when i tried to register the
50	neg_corpus	so slow during registration process making it invalid everytime	cant	even register and use the apps unable to login eventhough
89	neg_corpus	whats the point of this app being available if we	cant	even use it properly tak boleh guna bila saya
90	neg_corpus	went through all the trouble to reload with rm but	cant	even use it sent emails about the different interface
91	neg_corpus	not valid for toll fares sometimes please improve your system	cant	even use it cuz they dont hv the sri
92	neg_corpus	i just installed it a few seconds earlier and i	cant	even use it because of this terrible keep on
93	neg_corpus	update it even there is no update to update i	cant	even use it cannot reload tng total waste of
94	neg_corpus	doesnt have my mobile carrier to top up my phone	cant	even use it for all Irt stations i wanted
95	neg_corpus	the good job well idk why but these days i	cant	even use the app everytime when i tried to
96	neg_corpus	asking to update but nothing to update ended up i	cant	even use the app and also it would be
97	neg_corpus	good disappointed with this appi cant even log in i	cant	even use the app it doesnt have my correct
98	neg_corpus	is weak even i was standing beside my router i	cant	even use the prepaid function slow pathetic apps with
99	neg_corpus	for nothing can even do the money cash out you	cant	even use the ewallet at a lot of tolls
100	neg_corpus	sebab saya belum tebus satu sen pun lagi e wallet	cant	even use the pay function even after multiple days

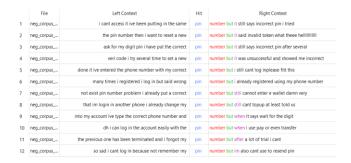
Apparently in **Figure VII**, the word "cant" with the highest frequency in **Figure VI** indicates the problem users were having with the app such as "cant even log in", "cant even register" and "cant even use".

FIGURE VIII. CONCORDANCE FOR "NUMBER"

	File	Left Context	Hit	Right Context
	neg_corpus	it even enter the app ssssfzssssszzbi i cant change my phone	number	i dont have my old phone number i cant log
	neg_corpus	ask you guys to solve i need to change my phone	number	but cannot because i log out and cannot log in
	neg_corpus	for hours few days later i try to change my phone	number	then i forgot the security question my account is locked
	neg_corpus	nerce get your priorities right cannot even change my phone	number	without sim i cant use it anymore very unstable and
	neg_corpus	using the old phone number then now ive change my phone	number	how to update thank you i cant understnd i have
	neg_corpus	duct rfid tu stop laa kekalkan tngsmartag i change my phone	number	so i cannot login how can i change to my
	neg_corpus	ike our information about our acc careful i change my phone	number	and i cant enter i try enter when jm underage
	neg_corpus	pdated arghhhh please help me how can i change my phone	number	i cannot log in into my account because the phone
	neg_corpus	masa nk log in i am trying to change my phone	number	i have tried to email but i have no response
0	neg_corpus	tak boleh log in eh good i have change my phone	number	and i cant to log in i hope theres a
1	neg_corpus	i cannot log in bad service i have change my phone	number	and an old wasnt available already my problem is the
2	neg_corpus	touched me and just gone i want to change my phone	number	to the new one but i cant my old number

Another interesting word is "number" where users were unable to log in the account after changing their phone number.

FIGURE VIII. CONCORDANCE FOR "PIN"



Users also suffer from not being able to use the application even with the correct pin number input.

#### B. N-Gram

The team has removed stop words using the Natural Language Toolkit (NLTK) with additional stop words taken from <a href="https://www.ranks.nl/stopwords">https://www.ranks.nl/stopwords</a> and Malay stop words before applying n-gram analysis on the dataset.

The top 10 most common words in positive and negative reviews are shown in Table  ${\bf I}$ .

TABLE I. TOP 10 Words Appeared In Unigram, Bigram and trigram

		Top 10 v	vords app	oeared in U	nigram	
#	Unigr	Unigr	Bigra	Bigram	Trigra	Trigram
	am	am	m	Negativ	m	Negative
	Positi	Negati	Positi	e	Positi	
	ve	ve	ve		ve	
1	good	app	good	tng card	good	reload
			good		good	tng card
					good	
2	app	card	good	phone	good	reload
			app	number	app	touch
					good	card
3	easy	tng	good	touch	good	top tng
			easy	card	good	card
					app	
4	apps	numbe	app	custome	good	top touch
		r	good	r service	good	card
					apps	
5	nice	ewalle	good	reload	good	call
		t	apps	tng	apps	customer
					good	service
6	ewalle	apps	good	card	good	tng card
	t		nice	balance	nice	balance
					good	
7	tng	reload	easy	create	app	topup tng
			good	account	good	card
					good	
8	best	accou	nice	qr code	good	tng card
		nt	good		good	app
					easy	
9	great	touch	apps	tng	easy	change
			good	ewallet	good	phone
					good	number
1	conve	topup	tng	number	good	phone
0	nient		card	exist	good	number
					nice	exist

From **Table I**, it is obvious that unigram can represent the users' positive sentiment, however, unigram does not provide much valuable information on users' negative experience with the app. Further investigation is being carried out in bigram. Based on the positive unigram, the word "easy" is used frequently by the users to

indicate that Touch 'n Go eWallet is easy to use. The ranks of the words "best" and "convenient" are number 8 and 10. This shows that most of the users consider Touch 'n Go eWallet is the best digital wallet compared to others due to its convenience.

Referring to the result of bigram, one may observe that Epizeuxis ("good good") and interchanging words ("good easy" and "easy good") are present in bigram positive sentiment. On the other hand, bigram negative shows more clue in users' negative experience with Touch 'n Go. The bigram negative words might be indicating the problems users were having with Touch 'n Go, specifically "tng card" which has the highest rank in the negative bigram. Some interesting words followed "tng card" are "phone number", "customer service", "card balance", "create account" and "qr code". The word "customer service" is ranked fourth due to the bad customer service of Touch 'n Go eWallet. One of the users, Shawal Abdul Rashid, complained that he was told by customer service that he had provided incorrect information 20 days after he contacted customer service to get his refund back. [6]

Apparently, users were frustrated over the absence of updated balance on "tng card" after topping up. Other issues users were facing were:

- 1. system showing "phone number exist" when users *intent to* change their forgotten pin number,
- 2. long waiting time for inefficient customer service
- 3. account number existed when creating account
- qr code payment crashing and long loading time to refresh qr code.

Epizeuxis and interchanging words also appeared in the trigram positive sentiment. On the other hand, one may observe that the words in trigram negative are showing almost the same context: "tng card", "customer service" and "phone number". These words can be obtained in Bigram negative in **Table I**. The top 4 words in Trigram negative are related to the problem of reloading the Touch 'n Go card. This may reflect the problem of being unable to reload the Touch 'n Go card through the Touch 'n Go eWallet. The Touch 'n Go eWallet card can only be reloaded though online banking or debit or credit card which is very inconvenient to the users.

Based on the result of **Table I**, unigram is sufficient to represent positive sentiment as bigram and trigram are emitting noises in showing Epizeuxis and interchanging words. On the other hand, bigram is sufficient to deliver the issues users were having with Touch 'n Go.

Referring to the positive unigrams, the words such as "good", "easy", "nice", "best", "great" and "convenient" are expressing the satisfaction of users towards Touch 'n Go while extra effort is needed to find out issues underlying the words shown in bigram negative. Due to this problem, the balance of Touch 'n Go card and Touch 'n Go eWallet are different. This problem may be the reason why the "tng card balance" ranks sixth. [7]

#### IMPROVEMENTS

Epizeuxis ("good good"), interchanging words ("good easy" and "easy good") and words with synonyms ("good", "nice", "best", and "terbaik") can be clustered and use one of the synonyms to represent them in each n-gram. While words such as "good app" and "good apps" can be improved by stemming the words.

Words such as "tng" and "ewallet" are highly used in this context due to the nature of the digital finance, which resulting in less valuable insights. To address the issue, concordance or other techniques are highly recommended to combine with n-gram for futher analysis.

#### V. CONCLUSION

The paper attempts to highlight the most positive and the most critical Touch 'n Go users reviews, but the negative user reviews were prone to issues users were having with the application. The results that we obtain by using concordance and n-gram methods, on our dataset which consists of 38458 users' reviews have fulfilled our objectives. The results could be used for the app improvement for Touch 'n Go eWallet. For example, based on the analysis results of negative reviews, the person in charge will know that the services related to tng card need to be improved to attract more customers.

The study has some limitations:

- Some reviews are a mix of multiple languages.
- Some reviews have spelling errors (e.g., cant instead of can't).
- Some reviews have an incorrect structure that makes it hard to understand the meanings.
- Some comments of the reviews do not match the rating (e.g., The review writes "excellent app" but the rating is only 1 star)

# VI. REFERENCES

- 1 "Touch 'n Go," [Online]. Available: https://www.touchngo.com.my/. [Accessed 18 November 2022].
- 2 P. v. Terlapu, "Exploring the API Calls for Malware Behavior Detection using Concordance and Document Frequency," August 2019. [Online]. Available: https://www.researchgate.net/publication/337414635\_Explor ing\_the\_API\_Calls\_for\_Malware\_Behavior\_Detection\_usin g\_Concordance\_and\_Document\_Frequency. [Accessed 22 November 2022].
- J. E. Mason, J. Duffy and M. Shepherd, "Classifying Web Pages by Genre: An n-Gram Approach," January 2009. [Online]. Available: https://www.researchgate.net/publication/221157666\_Classifying\_Web\_Pages\_by\_Genre\_An\_n-Gram\_Approach. [Accessed 20 November 2022].
- 4 M. Pagliardini, P. Gupta and M. Jaggi, "Unsupervised learning of sentence embeddings using compositional N-Gram features," 28 December 2018. [Online]. Available: https://arxiv.org/abs/1703.02507. [Accessed 10 December 2021]
- 5 X. Wang and J. Wang, "A method of hot topic detection in blogs using N-Gram model," January 2013. [Online]. Available: https://www.researchgate.net/publication/274662888\_A\_Method\_of\_Hot\_Topic\_Detection\_in\_Blogs\_Using\_N-gram\_Model. [Accessed 2022 November 20].
- 6 Actionline, "Not satisfied with Touch 'n Go mobile app," 4 November 2018. [Online]. Available: https://www.nst.com.my/actionline/2018/11/428179/not-satisfied-touch-n-go-mobile-app. [Accessed 1 December 2022].
- 7 I. Adreena, "Battle Of Malaysian E-Wallets: Who Offers The Best Service Among All?," 1 July 2019. [Online]. Available: https://gempak.com/rojakdaily/lifestyle/battle-malaysian-ewallets-who-offers-best-service-among-all-58949. [Accessed 2022 December 2].