



Sentiment Analysis of the Maps Application Uses the SVM Method and Predicts the Growth of Maps Application Users

Ilham Fariz Asya Mubarak ^{1(✉)}, Tukino ¹, Baenil Huda ¹, Aprilia Putri Nardilasari ¹, Risa Nur Islami ¹, and Muhammad Khaerudin ²

¹ Department of Information System, Faculty of Computer Science, Universitas Buana Perjuangan, Karawang, Indonesia

si19.ilhammubarak@mhs.ubpkarawang.ac.id

² Department of Informatics Engineering, Faculty of Engineering, Universitas Bhayangkara, Jakarta, Indonesia

Abstract. Nowadays, the directions application is often used by many people, this application is very helpful for its users in finding the best travel routes, by simply entering the address code, this application can provide routes that can be passed for various types of vehicles. In Indonesia itself there are several directions applications that are widely used with different positive and negative comments. There are several types of applications that will be used in this study, namely Google Maps, Maps.me, and Waze, the data collected is user comment data by means of web scraping. The purpose of this research is to find out the user reviews of each application used as research material. Data processing used the Support Vector Machine method, 750 comments were taken from each application, and the final result was Sygic Maps being the application with the highest score based on the results of 87.47% accuracy, 87.47% precision, and 99.85% on recall value. The Sygic Maps application has 31% positive comments, then Here We Go gets 29%, Maps.me gets 24%, then Waze has 14% positive comments, and google maps has 1% positive comments. With this, Sygic Maps became the application with the highest rating based on positive comments. In this study the Google Maps application became the application with the most users and comments with a total of 655,909, and the application with the least number of reviews was the Here We Go application with a total of 3,394 comments.

Keywords: Google Play, Sentiment Analysis, Google Play Support Vector Machine.

1 Introduction

In the current technological era, the application of directions is not a taboo for the community [1]. This application was created to make it easier for people to find an address or a location they want to visit, by simply entering the name of the address or location,

the application will provide the best route information to go through [2]. The application can be easily downloaded on the Google Play Store, the platform provides various kinds of applications, one of which is a directions application for Android users [3].

In Indonesia, there are several directions applications that are often used, such as Google Maps, Maps.me and Waze [4]. These three applications can be obtained by downloading through the Google Play Store, in this study these three applications are the objects to be studied. However, each application has its advantages and disadvantages, such as the placement of location points that are not appropriate, the route that is passed does not match the type of vehicle, and sometimes the navigation position delayed. These problems greatly affect user satisfaction.

In this study, data was required for reviewing application directions where the data was taken from the Google Play Store using a scraper [5]. These results are called sentiment analysis, the process is used to find user opinions in the form of topics or text [6][7]. The initial step in analyzing the data is to separate positive and negative user opinions from comments typed in the comments column [8], then compare the three applications to find out the level of each application based on the accuracy value [9].

Previous research was conducted by Fathurahman Bei and Sudin Saepudin, with the title Analysis of Sentiment for Online Ticket Applications in the Play Store Using the Support Vector Machine (SVM) Method. , then 75.03% owned by Traveloka, then 64.00% Mister Aladin, and 58.68% owned by the Tiket.com application. It can be concluded from the above results that the pegipegi platform has the highest accuracy value of 78.21% using the Support Vector Machine method [10].

2 Method

2.1 Data Recap

This stage is one of the initial processes for conducting sentiment analysis, the data used is taken from user comments [11]. This data collection process uses the help of a scraper by entering a link for each of these applications [12]. After carrying out data collection then giving sentiment labels to some of the comments, the labels are positive and negative to be used as training data and test data [13].

2.2 Pre-processing

Processing includes organizing and tidying up data, while several stages are tokenization, transform case, filter stopwords, and filter tokens [14].

2.3 Data classification

The following process uses the rapidminer application. comment data is used as an opinion that is negative and positive then data processing is carried out [15].

2.4 Application of the Algorithm

Support Vector Machine (SVM) is the method used in research. the results obtained through the Cross Validation stage are in the form of accuracy, a performance vector is needed to get accuracy, precision, and recall values [16].

2.5 Research Framework

The following are several processes in data collection using web scraping in each application (Figure 1).

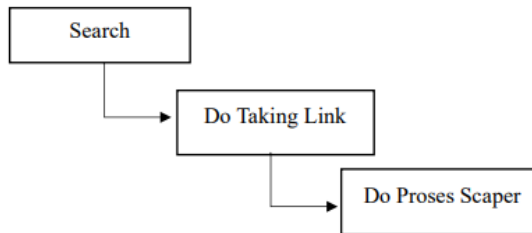


Fig. 1. Data collection flowchart

Data collection steps:

1. Search: researchers need to search for the type of application they want on google play store.
2. Links: before carrying out the scrap process it is necessary to take links from each application, this is done for the data collection process using a scraper/web scraping.
3. Scraping: the process of getting the desired comment data, in the scrap process you can set how much review data you want to take.

3 Results

To get commentary data, researchers need to take links from each of these applications from the Google Play Store. After getting the application link, the researcher needs the help of a scrapper to get a review of application comments in csv format so that they can be processed using the RapidMiner application. Comment data that was successfully collected was 750 from each application. The following is a Fig. 2 of the data collection process with the help of a scraper.

```
# Scrape All available reviews
#(DONT RUN THIS CELL!!! RUN ONLY IF YOU WANT TO SCRAPE ALL AVAILABLE REVIEWS)
#JANGAN DI RUN KODE INI, KECUALI ANDA INGIN SCRAPE SEMUA REVIEWS SEKALIGUS

from google_play_scraper import Sort, reviews_all

result = reviews_all(
    'com.bukalapak.android',
    sleep_milliseconds=0, # defaults to 0
    lang='id', # defaults to 'en'
    country='id', # defaults to 'us'
    sort=Sort.NEWEST, # defaults to Sort.MOST_RELEVANT , you can use Sort.NEWEST
)
```

Fig. 2. Picture 2. Data scraper process

In preparation for data processing, it is necessary to label each application user's comments into positive and negative. Then enter into the pre-processing process and use of the method (Fig. 3).

3.1 Pre-processing

The following are the stages of data processing:

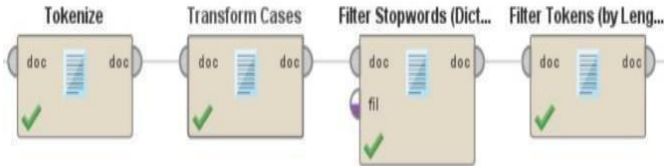


Fig. 3. Data Process

1. Tokenization, this stage is a process to distinguish between word separators and remove certain characters such as punctuation marks.
2. Transform case, which aims to automatically change text to lowercase.
3. Filter stopwords, functions in this analysis to remove text that is not used automatically without reducing the sentiment content of the text.
4. Filter Tokens, used to filter the length of letters, the word used is not more than one.

3.2 Prediction

Through the data collection stage per year, the number of users on the sygic maps, google maps, waze, maps.me, and here we go applications is obtained each year with the following numbers:

In Fig. 4, it can be seen that the most users were in 2022 with a total of 10.3316 and for users at least 80 users in 2011, with a total number of users as much as 655.909.

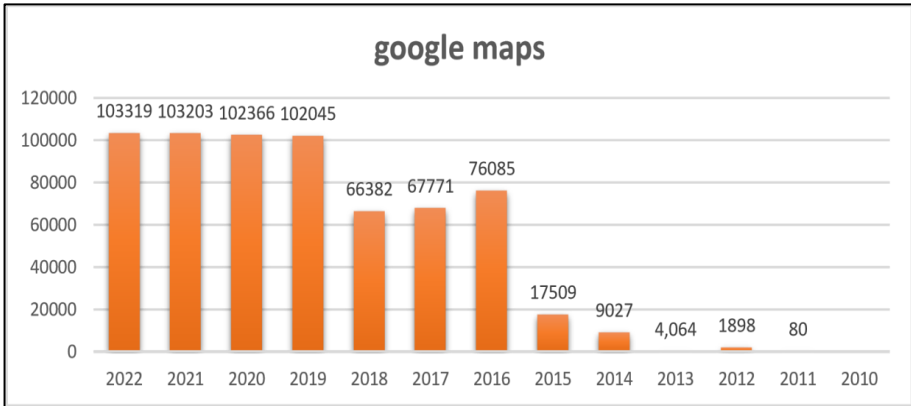


Fig. 4. The number of google maps users per year

On Fig. 5, it can be seen that the most users were in 2015 with a total of 2.780 and for users at least 71 users in 2014, with a total number of users as much as 6.409.

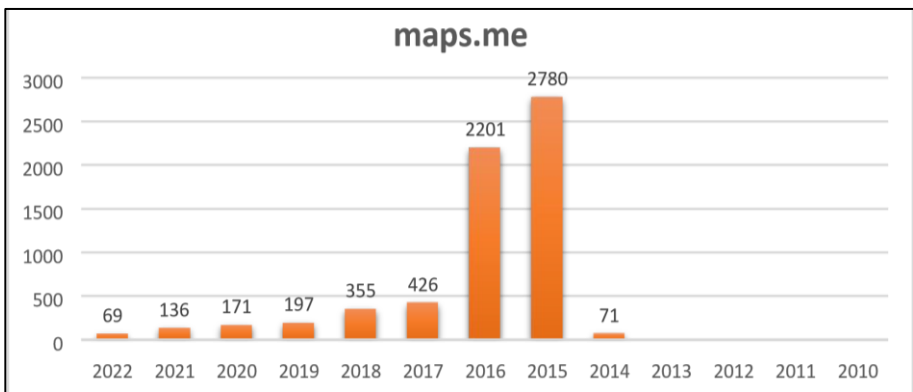


Fig. 5. The number of Maps.me users per year

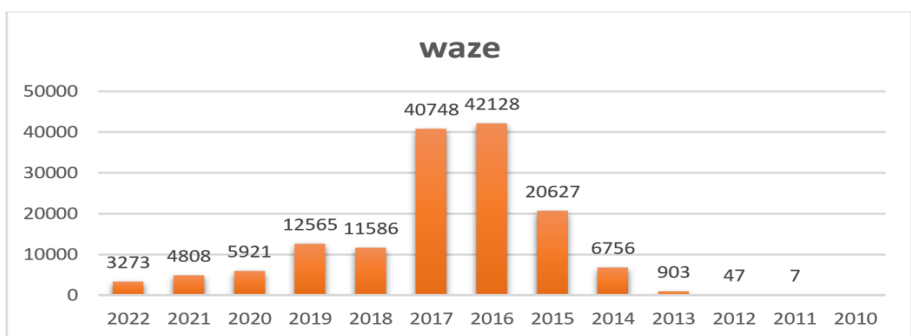


Fig. 6. The number of Waze users per year

On Fig. 6, it can be seen that the most users were in 2016 with a total of 42.128 and for users at least 7 users in 2011, with a total number of users as much as 149.453.

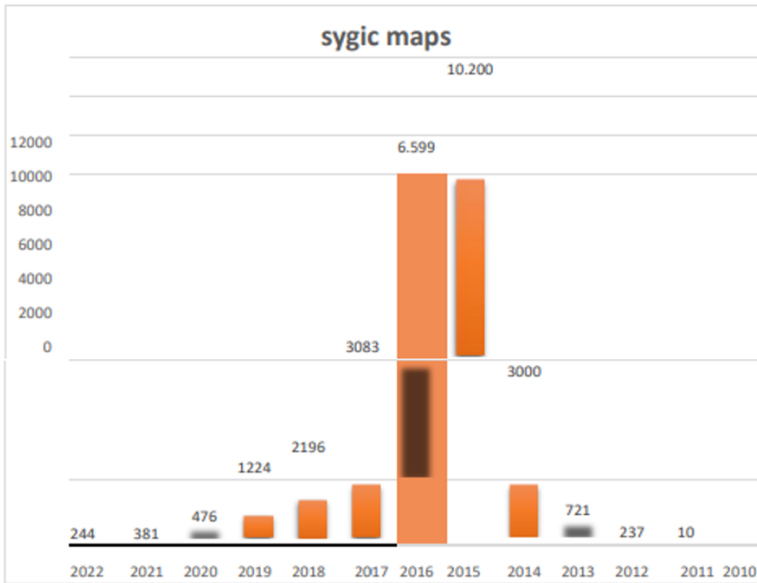


Fig. 7. The number of Sygic users per year

On Fig. 7, it can be seen that the most users were in 2015 with a total of 10.200 and for users at least 10 users in 2011, with a total number of users as much as 28.383.

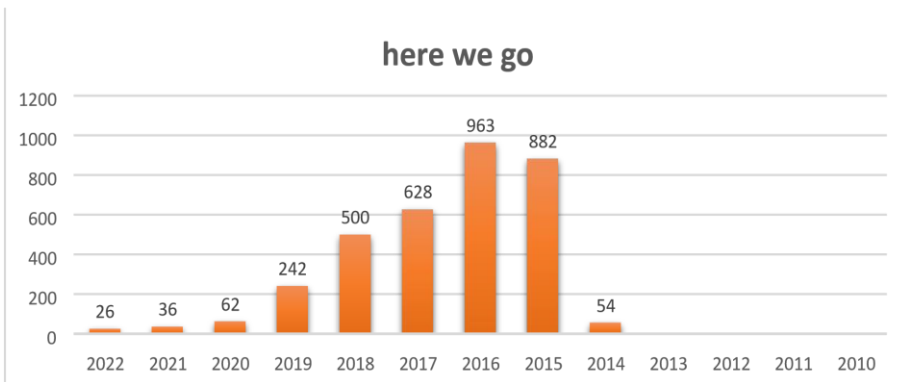


Fig. 8. The number of Here we go users per year

On Fig. 8, it can be seen that the most users were in 2016 with a total of 963 and for users at least 54 users in 2014, with a total number of users as much as 3.394

4 Conclusion

After analyzing and testing each application with the Support Vector Machine algorithm, it shows that the highest total score is found in the Here We Go application with results of 87.47% accuracy, 87.47% precision, and 99.85% on the recall value. The Sygic Maps application has 31% positive comments, then Here We Go gets 29%, Maps.me gets 24%, then Waze has 14% positive comments, and google maps has 1% positive comments. With this, Sygic Maps became the application with the highest rating based on positive comments. It's with that in mind that Sygic Maps is the app with the best value for now. This can be a recommendation for users who are looking for a directions application to travel to various places they want to go. In this study, Google Maps became the application with the highest number of users and reviews. From the data collected with the help of a scraper, researchers managed to obtain 655,909 user review data starting in 2011 to 2022. And the least number of users and comments are in the Here We Go application. with a total of 3,394, the data was successfully obtained from 2014 to 2022.

References

1. Giovani, A.P.; Ardiansyah, A.; Haryanti, T.; Kurniawati, L.; Gata, W. Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *J. Teknoinfo* 2020, *14*, 115, doi:10.33365/jti.v14i2.679.
2. Hananto, A.L.; Priyatna, B.; Rahman, A.Y. Penerapan Algoritma Dijkstra Pada Sistem Monitoring Petugas Lapangan Pemkab Bekasi Berbasis Android. *JOINTECS (Journal Inf. Technol. Comput. Sci.* 2019, *4*, 95, doi:10.31328/jointecs.v4i3.1078.
3. Herlinawati, N.; Yuliani, Y.; Faizah, S.; Gata, W.; Samudi, S. Analisis Sentimen Zoom Cloud Meetings Di Play Store Menggunakan Naïve Bayes Dan Support Vector Machine. *CESS (Journal Comput. Eng. Syst. Sci.* 2020, *5*, 293, doi:10.24114/cess.v5i2.18186.
4. Trapsilawati, F.; Wijayanto, T.; Jourdy, E.S. Human-Computer Trust in Navigation Systems: *Commun. Sci. Technol.* 2019, *4*, 38–43, doi:10.21924/cst.4.1.2019.112.
5. Yolanda, N.; Santi, I.H.; Fanny, D.; Permadi, H. Analisis Sentimen Popularitas Aplikasi Moodle Dan Edmodo Menggunakan Algoritma Support Vector Machine. 2022, *3*, 48–59.
6. Tineges, R.; Triayudi, A.; Sholihati, I.D. Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM). *J. Media Inform. Budidarma* 2020, *4*, 650, doi:10.30865/mib.v4i3.2181.
7. Kusnadi, R.; Yusuf, Y.; Andriantony, A.; Ardian Yaputra, R.; Caintan, M. Analisis Sentimen Terhadap Game Genshin Impact Menggunakan Bert. *Rabit J. Teknol. dan Sist. Inf. Univrab* 2021, *6*, 122–129, doi:10.36341/rabit.v6i2.1765.
8. Diki Hendriyanto, M.; Ridha, A.A.; Enri, U. Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Mola Application Reviews on Google Play Store Using Support Vector Machine Algorithm. *J. Inf. Technol. Comput. Sci.* 2022, *5*, 1–7.

12. Sari, B.W.; Haranto, F.F. Implementasi Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Terhadap Pelayanan Telkom Dan Biznet. *J. Pilar Nusa Mandiri* 2019, *15*, 171–176, doi:10.33480/pilar.v15i2.699.
13. Bei, F.; Sudin, S. Analisis Sentimen Aplikasi Tiket Online Di Play Store Menggunakan Metode Support Vector Machine (Svm). *Sismatik* 2021, *01*, 91–97.
14. Wahyudi, R.; Kusumawardana, G. Analisis Sentimen Pada Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine. *J. Inform.* 2021, *8*, 200–207, doi:10.31294/ji.v8i2.9681.
15. Nurhafida, S.I.; Sembiring, F.; Raya, J.; No, C. Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM). 2022, *6*, 317–327.
16. Siregar, A.M. Klasifikasi Untuk Prediksi Cuaca Menggunakan Esemble Learning. *Petir* 2020, *13*, 138–147, doi:10.33322/petir.v13i2.998.
17. Fatmawati, F.; Affandes, M. Klasifikasi Keluhan Menggunakan Metode Support Vector Machine (SVM) Pada Akun Facebook Group IRaise Helpdesk. *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.* 2018, *3*, 24, doi:10.24014/coreit.v3i1.3552.
18. Salim, A.; Gata, W.; Hilman Fakhri, M.; Sri Rhayu, C.; Budiarto, A.; Studi Magister Ilmu
19. Komputer STMIK Nusa Mandiri Jakarta, P. Analisis Sentiment Instagram Menggunakan Metode Support Vector Machine (SVM) Berbasis Grid Search Algorithm (GSA). 2022, 466–472.
20. Fikri, M.I.; Sabrila, T.S.; Azhar, Y. Perbandingan Metode Naïve Bayes Dan Support Vector Machine Pada Analisis Sentimen Twitter. *Smatika J.* 2020, *10*, 71–76, doi:10.32664/smatika.v10i02.455.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

