International Applied Science
UMJember Proceeding Series (2022) Vol. 1, No. 2 : 67-73

ICRD 2022

Post Pandemic Public Acceptance Map Toward Ijen Geopark Using Aspect Based Sentiment Analysis Technique

Bagus Setya Rintyarna[1]

[1]Universitas Muhammadiyah Jember, Indonesia

DOI: https://doi.org/10.32528/ias.v1i2.163
*Correspondence: Bagus Setya Rintyarna
Email: bagus.setya@unmuhjember.ac.id

Published: Juni, 2022

Abstract: Tourism is one of the sectors that has been seriously hit by the impact of the Covid-19 pandemic indicated by the sharp decrease in the number of both domestic and foreign tourists visit. The decrease has also been experienced by the popular tourist destination located in the border of Bondowoso and Banyuwangi Regency namely Ijen Geopark. Since the tourism sector has shown a significant role as a driver of economic activity, the aim of this work is to build public acceptance map toward Ijen Geopark as the government reduce the limitation of the prohibition policy to handle the spread of Covid-19. Sentiment analysis is adopted to build the public acceptance map by using TripAdvisor dataset. To gain a fine grained, results we extent the technique using our proposed aspect-based sentiment analysis making use cosine similarity algorithm and several keywords. Using several keywords and cosine similarity algorithm, we succesfully extract aspect of the topic. Employing Naïve Bayes, we then labeled the sentiment orientation of the associated aspect extracted from sentence. During 2022, we highlight the fact that most tourist positively accept Ijen Geopark. Some parts that need to be evaluated by the decision maker and related party is Sulfur Mining and the cost since both aspect received more negative comment compared to the positive ones.

Keywords: Ijen Geopark; Public Acceptance; Sentiment Analysis

INTRODUCTION

Since World Health Organization has anounced Coronavirus disease (Covid-19) as a pandemic in early 2020, countries worlwide including Indonesia has faced major consequences [1]. The consequences include political, economic, social and many other public and private sectors [2]. Tourism is one of the sectors that has been seriously hit by the impact of the pandemic. Following the massive hit of Covid-19 especially in 2020, the central government of Indonesia ruled to limit movement of people troughout the country to face Covid-19 spread. Central Bureau of Statistics (BPS) has reported the sharp decrease in the number of foreign tourists visit during 2020 and 2021. Throughout 2020, the number of foreign tourists who visit Indonesia was only about 4 million people. Compared to the number of tourists visit in 2019, the number was only 25%. The condition in 2021 is even worst. Indicated by the sharp decrease in the number of foreign tourists visit, the number achieved only 1,1 million people.

The sharp decrease of the number of tourists visit during the spread of Covid-19 has also been experienced by a wellknown tourism object in Bondowoso namely Ijen Geopark or Ijen Volcano Complex that will be the object of this study [3]. The geopark consists of a cluster of composites volcanoes situated in the border of Bondowoso and Banyuwangi Regency of Jawa Timur Province. There are three most popular tourist destination located in the geopark i.e: acidic crater lake, traditional sulfur mining and the popular blue fire. The tourism sector has shown a significant role, especially as a driver of economic activity through its contribution to employment, increased investment and business opportunities including Ijen Geopark to both

Bondowoso and Banyuwangi Regency. Tourism is very influential on the development of economic life in the region. The decrease of the tourism sector is therefore will significantly impact the economy of both regencies.

Accordingly, this work attempt to build a public acceptance map toward Ijen Geopark as a popular tourist destination during post pandemic i.e.: 2022 as the government reduce the prohibition policy [4] to face Covid-19 infection spread. By mapping public acceptance, this work aims at foreseeing how the tourism sector especially Ijen Geopark will recover in the near future. The adopted technique is called sentiment analysis. Sentiment analysis is a text mining based technique to automatically extracts people's opinion, sentiment, perspective or emotion expressed in an online platform [5]. The popularity of sentiment analysis has made it as a major technique adopted to reveal trends in the community [6].

There are two common techniques to apply sentiment analysis i.e.: machine learning based sentiment analysis and lexicon-based sentiment analysis [7]. Machine learning based sentiment analysis is lied on a machine learning algorithm involving training and testing step. Meanwhile, lexicon-based sentiment analysis is lied on a sentiment lexicon, a digital library comprises collection of words along with their associated sentiment polarity, that is, positive, negative or neutral [8]. In this work, we adopt a machine learning-based sentiment analysis using naïve bayes algorithm [9]. To get a fine- grained result of the sentiment analysis, we extent to extract features of the dataset using our proposed aspect-based sentiment analysis technique. Aspect-based sentiment analysis aims at revealing the aspects that received positive, neutral, or negative ratings from the users. In applying sentiment analysis, we employ TripAdvisor dataset grabbed during 2022 representing the time after the government reduce the prohibition policy.

METHOD

All the proposed framework in extracting public acceptance map using TripAdvisor dataset can be shown in Fig. 1. In the first step, we grabbed dataset from TripAdvisor cloud. After collecting the datasets, the common text pre-processing technique is then applied involving case folding, tokenizing, normalization, removal of stop words, and also stemming [10]. With case folding, the letter is changed into lowercase [11]. Meanwhile tokenizing deals with splitting sentence into terms/words. Stop word removal is needed to filter the word from meaningless term using stop word algorithms.

Table 1. Example of the dataset

Raw data	Split to Sentence
Menakjubkan sekali. Kawah aktif yang sangat indah pemandangannya. Kami berangkat tengah malam berburu blue fire, walau belum beruntung.	menakjubkan sekali
	kawah aktif yang sangat indah pemandangannya
	kami berangkat tengah malam berburu blue fire, walau belum beruntung
Perjalanan melelahkan terobati dengan keindahan kawah ijen. Hamparan seperti danau tenang. Luar biasa landscape yang terbentang dengan tebing batu, sangat bagus untuk sesi fotografi. Pemandangan menakjubkan.	perjalanan melelahkan terobati dengan keindahan kawah ijen
	hamparan seperti danau tenang
	luar biasa landscape yang terbentang dengan tebing batu, sangat bagus untuk sesi fotografi.
	pemandangan menakjubkan.

In the next step, we then split sentence from the document assuming that every sentence contains an aspect to be commented by the user. Example of the aspect and its associated opinion word is presented in

Fig. 2. The sentence is then grouped to build dataset. Example of the dataset is presented in Table 1. A sentence is now assigned as a document to classify. To extract aspect, we use keywords as presented in Table 2.

Table 2. Keywords and their associated aspects

Aspects	Keywords
food	warung, restoran, makanan, kuliner, jajangan
hotel	penginapan, gues house, losmen, hotel
transportation	pesawat, angkutan, tiket, kereta api, bis, mobil, jalur, pendakian, jalan lingkar
ijen crater	kawah, kawah ijen, debu, pemandangan, landscape
blue fire	blue, fire, blue fire, api, biru, api biru
sulfur mining	penambang, belerang, sulfur, masker, sarung tangan
people	warga, lokal, keramahan, agensi, pemandu wisata, pelayanan
Cost	biaya, harga, charge

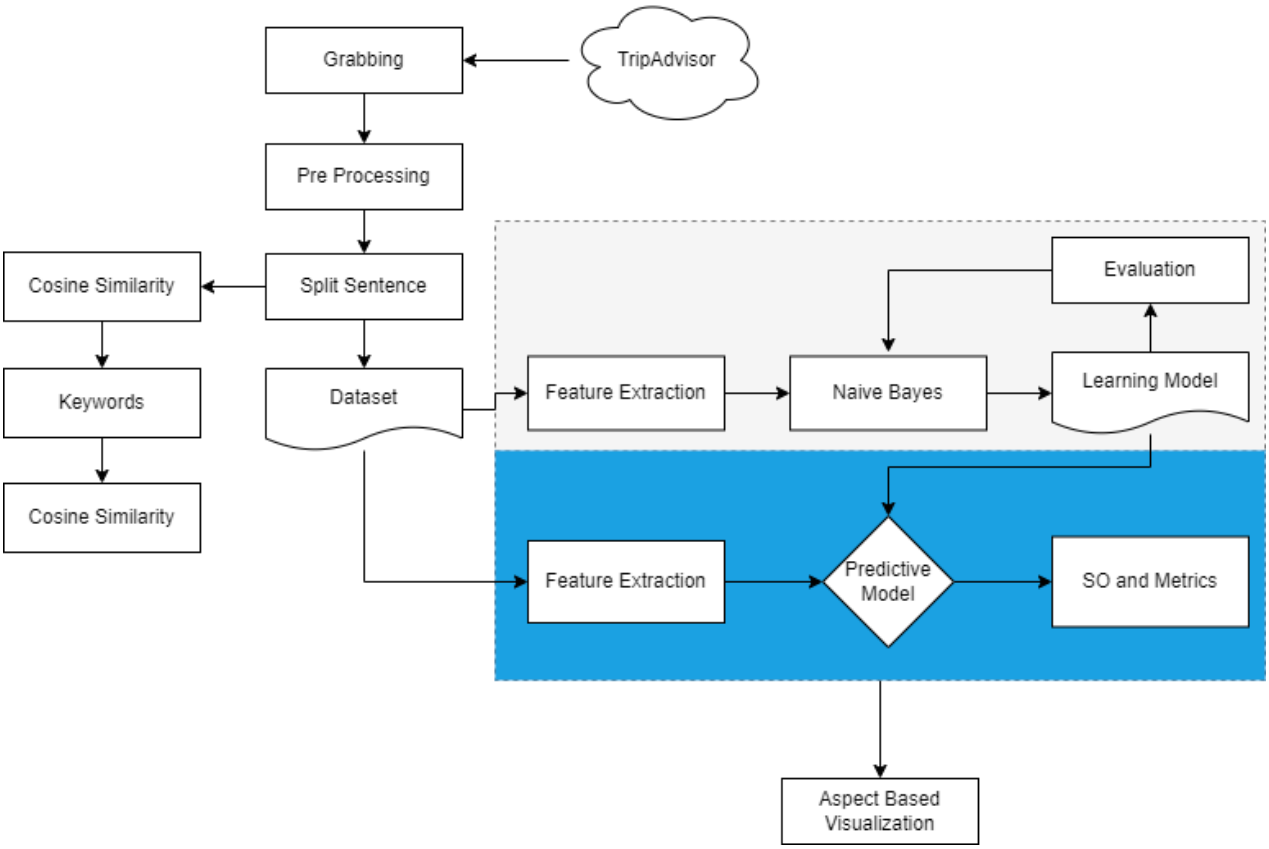


Fig 1. Framework of the proposed method

To extract aspect, the similarity between keywords (Table 1) and terms in the sentence is then calculated using cosine similarity algorithms as presented in formula (1). The extracted aspect is indicated by the highest value of similarity between keywords and term [7]. Aspect is then assigned to the associated sentence.

$$\cos(\phi) = \frac{A \,.\, B}{||A||||B||} = \frac{\sum_{i=1}^{n} A_i.B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (1)$$

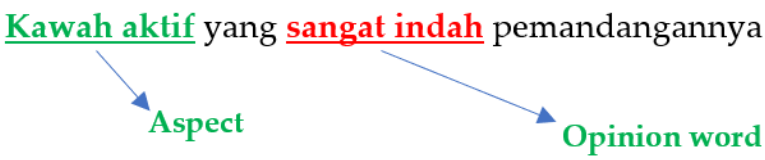


Fig. 2 Example of the aspect and its associated opinion word.

Feature extracted for classificatin is bag of words (BOW) as a common feature for text classification in the field of Natural Language Processing (NLP). Document, that in this work is sentence is represented by multiset of its words/terms omitting grammar, context and even word order. The extracted feature of BOW is used to build learning model for classification using Naïve Bayes algorithm.

In assigning sentiment orientation of a sentence that has been assigned by the aspect, we make use Naïve Bayes algorithm. To be applied to solve a classification task, Naïve Bayes Classifier makes use a simple probabilistic machine learning model based on a Bayes Theorem that assumes independence among predictors. It is simply conditional probability model, probabilistics classifier that is able to build fast machine learning model. Given $x = (x_1, \dots, x_n)$ is a vector representing instance to be categorized with n is some features of independent variables, Naïve Bayes assign probabilities p for each k possible results C_k as represented in (2) and (3). The main reason of using Naïve Bayes in this work is that it is simple and fast to predict sentiment of the datasets.

$$p(C_k\,|x_1, \dots, x_n) = p(x_1\,|C_k).p(x_2|C_k) \dots . p(x_n|C_k).p(C_k) \quad (2)$$

$$p(C_k\,|x) = \frac{p(C_k)p(x\,|C_k)}{p(x)} \quad (3)$$

The sentiment orientation assigned by Naïve Bayes to the sentence is then established as the sentiment of the associated aspect of the sentence. To validate the results, we also calculate precision, recall, f measure, and accuracy in the sentence basis.

RESULTS AND DISCUSSION

In the first step, the data is succesfully grabbed from TripAdvisor cloud as many as 669 review documents. Based on the metadata, the review document can be categorized as presented in Fig. 3. Before splitting into sentence, we also apply document classification using Naïve Bayes, namely document level sentiment analysis. The result is presented in Fig. 4 revealing the fact that during 2022, the visit is still dominated by domestic tourist. The foregn visit is not yet recovered. This information would be very helpful for both the government as the policy maker and the related party to encourage foreign visit to Ijen.

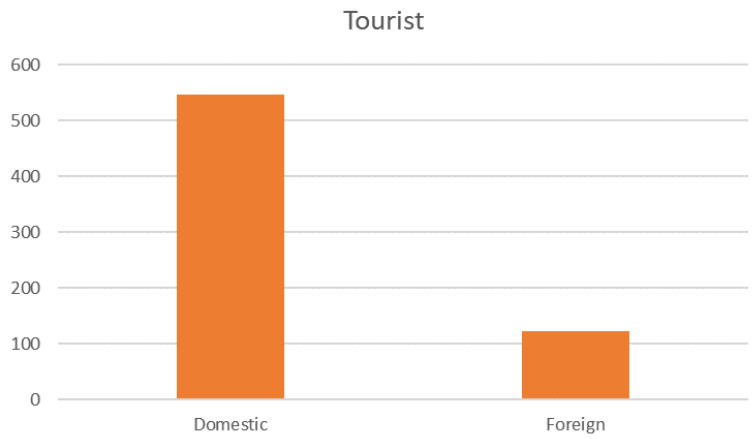


Fig. 3 Domestic and Foreign tourist

Fig. 4. Document level sentiment analysis

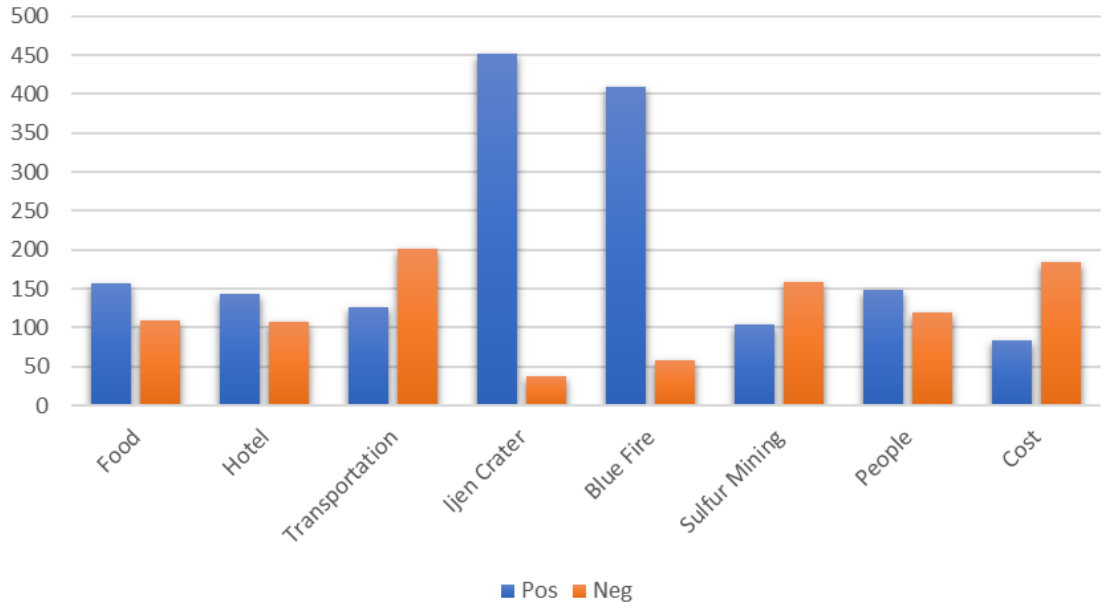
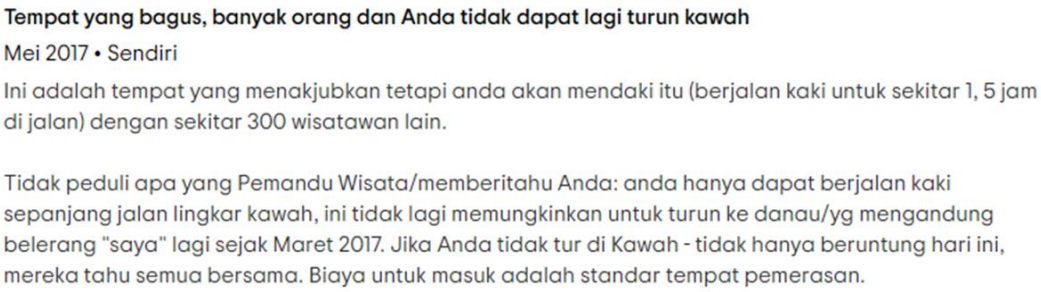


Fig 5. The result of aspect based sentiment analysis

Bar chart presented in Fig 4 highlights the fact that positive comment is still dominated the overall comment by 78% indicated that tourist both domestic and foregn accept positively Ijen Geopark during 2022. The fine-grained result of aspect-based sentiment analysis also confirms more likely the same result as portrayed in Fig 5. Fig 5 reveals that the aspect that mostly accepted by the tourist is Ijen Crater and Blue Fire. Yet, negative comment is mostly given to two aspects namely Sulfur mining and the cost. Here in Fig 6, we present the example of negative comment to the associated aspects.


Tempat yang bagus, banyak orang dan Anda tidak dapat lagi turun kawah

Mei 2017 • Sendiri

Ini adalah tempat yang menakjubkan tetapi anda akan mendaki itu (berjalan kaki untuk sekitar 1, 5 jam di jalan) dengan sekitar 300 wisatawan lain.

Tidak peduli apa yang Pemandu Wisata/memberitahu Anda: anda hanya dapat berjalan kaki sepanjang jalan lingkar kawah, ini tidak lagi memungkinkan untuk turun ke danau/yg mengandung belerang "saya" lagi sejak Maret 2017. Jika Anda tidak tur di Kawah - tidak hanya beruntung hari ini, mereka tahu semua bersama. Biaya untuk masuk adalah standar tempat pemerasan.

Fig. 6. Example of Negative comments

CONCLUSION

The aim of this work is to build public acceptance map toward Ijen Geopark as the government reduce the limitation of the prohibition policy to handle the spread of Covid-19, especially during 2022. Using a simple cosine similarity measure, this work is succesfully extract aspect from review sentence using several pre-defined associated keywords. During 2022, we highlight the fact that most tourist positively accept Ijen Geopark. Some parts that need to be evaluated by the decision maker and related party is Sulfur Mining and the cost since both aspects received more negative comment compared to the positive ones.

REFERENCES

[1] X. Wang, L. Wang, X. Zhang, and F. Fan, "The spatiotemporal evolution of COVID-19 in China and its impact on urban economic resilience," *China Econ. Rev.*, vol. 74, no. June 2020, p. 101806, 2022, doi: 10.1016/j.chieco.2022.101806.

[2] B. S. Rintyarna, "Mapping acceptance of Indonesian organic food consumption under Covid-19 pandemic using Sentiment Analysis of Twitter dataset," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 5, pp. 1009–1019, 2021.

[3] I. K. P. Widhiantari, I. N. G. Ustriyana, and I. M. Sudarma, "Dampak Pengembangan Objek Wisata Kawah Ijen terhadap Peningkatan Kesejahteraan Petani," *Ojs.Unud.Ac.Id*, vol. 10, no. 2, pp. 503–511, 2021, [Online]. Available: https://ojs.unud.ac.id/index.php/nandur/article/download/64811/40636.

[4] A. Ristyawati, "Efektifitas Kebijakan Pembatasan Sosial Berskala Besar Dalam Masa Pandemi Corona Virus 2019 oleh Pemerintah Sesuai Amanat UUD NRI Tahun 1945," *Adm. Law Gov. J.*, vol. 3, no. 2, pp. 240–249, 2020, doi: 10.14710/alj.v3i2.240-249.

[5] B. S. Rintyarna, H. Kuswanto, R. Sarno, and E. K. Rachmaningsih, "Modelling Service Quality of Internet Service Providers during COVID-19 : The Customer Perspective Based on Twitter Dataset," pp. 1–12, 2022.

[6] F. H. Rachman, Imamah, and B. S. Rintyarna, "Sentiment Analysis of Madura Tourism in New Normal Era using Text Blob and KNN with Hyperparameter Tuning," *2021 Int. Semin. Mach. Learn. Optim. Data Sci. ISMODE 2021*, pp. 23–27, 2022, doi: 10.1109/ISMODE53584.2022.9742894.

[7] B. S. Rintyarna, R. Sarno, and C. Fatichah, "Evaluating the performance of sentence level features and domain sensitive features of product reviews on supervised sentiment analysis tasks," *J. Big Data*, vol. 6, no. 1, 2020, doi: 10.1186/s40537-019-0246-8.

[8] B. S. Rintyarna, R. Sarno, and C. Fatichah, "Enhancing the performance of sentiment analysis task on product reviews by handling both local and global context," *Int. J. Inf. Decis. Sci.*, vol. 11, no. xxxx, 2018.

[9] B. S. Rintyarna, "Sentiment Analysis pada Data Twitter dengan Pendekatan Naïve Bayes Multinomial," *J. Sist. Teknol. Inf. Indones.*, pp. 1–6, 2017.

[10] B. S. Rintyarna, R. Sarno, and A. L. Yuananda, "Automatic ranking system of university based on technology readiness level using LDA-Adaboost.MH," *2018 Int. Conf. Inf. Commun. Technol. ICOIACT 2018*, vol. 2018-Janua, pp. 495–499, 2018, doi: 10.1109/ICOIACT.2018.8350706.

[11] A. A. Suryani, I. Arieshanti, B. W. Yohanes, M. Subair, S. D. Budiwati, and B. S. Rintyarna, "Enriching English Into Sundanese and Javanese Translation List Using Pivot Language," in *2016 International Conference on Information, Communication Technology and System (IC*, 2016, pp. 167–171.