RESEARCH ARTICLE | MAY 09 2023

PSO optimization for analysis of online marketplace products on the SVM method

Sucitra Sahara 🖾; Annida Purnamawati; Sulaeman Hadi Sukmana; ... et. al

Check for updates

AIP Conference Proceedings 2714, 030010 (2023) https://doi.org/10.1063/5.0129404



Articles You May Be Interested In

Individual-based simulation for diffusion of online marketplace among trade SMEs

AIP Conference Proceedings (June 2019)

Modeling the milling tool wear by using an evolutionary SVM-based model from milling runs experimental data

AIP Conference Proceedings (December 2015)

Mining Twitter data on Covid-19 for sentiment analysis using SVM algorithm

AIP Conference Proceedings (May 2023)





PSO Optimization For Analysis of Online Marketplace Products on the SVM Method

Sucitra Sahara,^{1, a)} Annida Purnamawati,^{2, b)} Sulaeman Hadi Sukmana,^{1, c)} Mely Mailasari,^{2, d)} Erma Delima Sikumbang,^{3, e)} and Nurlaela Eva Puji Lestari^{4, f)}

¹⁾Tecnology Informatics, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450, Indonesia.

²⁾System Informatics, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450, Indonesia

³⁾Accounting Information System, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450, Indonesia

⁴⁾Management, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450, Indonesia

a)Corresponding author: sucitra.scr@bsi.ac.id
b)Electronic mail: annida.npr@bsi.ac.id
c)Electronic mail: sulaeman.sdu@bsi.ac.id
d)Electronic mail: mely.myl@bsi.ac.id@bsi.ac.id
e)Electronic mail: erma@bsi.ac.id
f)Electronic mail: nurlaela.nep@bsi.ac.id

Abstract. The presence of online marketplace nowadays makes people shop via cellphones or laptops effortlessly. Customers are able to make transaction process easier. However, many internet users are still hesitant about choosing a marketplace. Some sellers still carry out a sale system which leads to customers' dissatisfaction. Therefore, researchers will make a selection of the product marketplace based on reviews and comments. Researchers carry out the stage of selecting marketplace products based on an opinion or public opinion in the comments column on the online marketplace being used. From the number of comments, samples were taken which were processed and grouped into datasets. The researcher classified the data in the form of text by applying the Support Vector Machine (SVM) algorithm which would be compared using Particle Swarm Optimization (PSO). After a applied to SVM method, it produces a pretty good value, however this value may still increase as compared with PSO optimization. It is concluded that by applying the SVM method with PSO optimization the accuracy value is more maximal in processing the dataset in the form of text classification in marketplace reviews. The results of testing these data; sellers are able to find out which words are related to the sentiments that often appear and have the highest weight.

INTRODUCTION

Currently, many online retailers use third parties (online marketplaces) as an alternative to selling other than their website[1]. Online trading and online marketplaces are places to make transactions quickly[2]. During Covid-19 pandemic has had an impact in various private sectors, especially in the economic sector. However, in the field of electronic commerce (e-commerce), it has incredibly escalated. The revenue of several marketplaces has shown a significant increase during the pandemic. It happened because the convenience given to people in shopping. There is no need to spend time and energy by visiting the shopping stores. Online shopping demands customers to carefully have a look at multiple reviews of a certain product to see the reputation of the stores [3]. Nowadays, product reviews are very important on e-commerce sites which tend to be a resource in evaluating customer behavior and needs[4]. The application of technology to the marketplace has attracted the attention of researchers to analyze sentiment on reviews from customers. Customers influence changes to new actions and information that must be carried out by the marketing department[5]. One branch of research in Text Mining is sentiment analysis or called opinion mining. The data source used is the difference in sentiment analysis, the level that is often used is sentiment analysis at the document level and sentence level[6]. According to Nugroho, Crisnanto and Wahana (2015), sentiment analysis is one of the computational sciences of various opinions, sentiments and emotions that are expressed in the text[7]. A research conducted by M. Iqbal Ahmadi (2020) classify positive and negative user reviews on online shopping applications using the Support Vector Machine (SVM) method. The result showed that Tokopedia obtained the highest accuracy value of 90.67%, JD.ID 75.33%, Blibli 74%, Shopee 70% and Lazada 69% [8]. A research by Sharazita Dyah Anggita and Ikmah (2020), having conducted a sentiment analysis based on user opinions about ship and delivery services via Twitter. The research was conducted by applying the Naive Baiyes classifier algorithm

> 2nd International Conference on Advanced Information Scientific Development (ICAISD) 2021 AIP Conf. Proc. 2714, 030010-1–030010-6; https://doi.org/10.1063/5.0129404 Published by AIP Publishing. 978-0-7354-4520-8/\$30.00

and Support Vector Machine (SVM) optimized with Particle Swarm Optimization (PSO). The results obtained were an increase in accuracy of 15.11% in the application of PSO-based Naive Bayes while the SVM PSO increased the highest accuracy by 1.74% [9]. A research conducted by Wajiyono et al (2019), research was conducted by analyzing the sentiments of Fintech users, especially OVO application. The data were taken from reviews of OVO at Google Playstore. This research uses the Support Vector Machine and PSO methods which are processed using RapidMiner. The results of testing the OVO application, it acquires the best accuracy of 82.33%. Therefore, it can be concluded that OVO review can convince users of the positive comments obtained in the application[10]. According to studies above, researchers are interested in finding out how to conduct sentiment analysis on reviews in marketplace to determine whether existing reviews are negative, positive or neutral using the Support Vector Machine (SVM) method compared to Particle Swarm Optimization (PSO). In this study, SVM is able to process datasets which have been collected and grouped into positive text and negative text, especially in selecting a text. By using the SVM method compared to PSO, it is expected to be able to produce high accuracy values in order that the processing of the dataset becomes stronger and more valid.

RESEARCH METHODS

The method used is a Support Vector Machine which is compared with Particle Swarm Optimization (PSO) to produce a high degree of accuracy in analyzing sentiment labels in this study. The main principle of using SVM is to find the best hyperplane that functions as a separator of two classes in the input space. SVM is a machine learning tool that conducts training using training datasets and generalizes and makes predictions from new data [11]. The Support Vector Machine is an improved learning classification algorithm for the solution of regression, developed by Vladimir Vapnik and his group[12]. SVM can generalize relatively invisible cases with the aim of minimizing errors so that it can predict more accurately in multi-class cases[13]. SVM advantages are high dimensional input space and vector document space[14]. Meanwhile, Particle Swarm Optimization (PSO) technique is used in the selection of optimal parameters of the SVM classification used to improve prediction performance[15]. The following is a picture of the method proposed by researchers designed in the Online Marketplace product review:



FIGURE 1. Proposed Method

The followings are the process of the stages of research carried out as follows: a). Tokenizazion is a researcher using Tokenize to separate words or letters from punctuation marks and symbols. b). Stopwards Removal is a word that is deemed unnecessary in processing review sentiment data, for example if, the, of, or, etc. c). Steamming is the process of changing the form of words into basic words. This method of changing the form of words into basic words adjusts the structure of the language used in the steamming process.

RESULT AND DISCUSSIONS

The dataset or training data used in text classification is in the form of marketplace product reviews consisting of 100 positive reviews and 100 negative reviews taken from online sample reviews of several marketplaces. Processed review data is in the form of a collection of text or sentences separated in the form of text documents. Before being classified, the data must go through a step-by-step process in order to make the data can be processed properly and the results can reach maximum accuracy. The followings are the stages of processing the dataset:

a. Preprocessing Tokenization Stage, Stopwords Removal, Stemming data before the data testing process.

doc	Vector Creation	doc doc

FIGURE 2. Data Preprocessing Stage

Some of the processes carried out in Figure 2 are Tokenizazion. Researchers use Tokenize to separate letters from punctuation and symbols. Stopwards Removal will process conjunctions or connectors that are deemed unnecessary in processing review sentiment data in the form of comments, for example is, a, the, if, of, or, etc. And Steamming is the process of changing the form of a previous word into a root word. This method of changing the form of a word into a root word adjusts the structure of the language used in the stemming process.

b. The results of the experiment testing the dataset using the SVM method. In the first stage the researchers experimented by testing the data on the Rapidminer application using the Support Vector Machine (SVM) algorithm. After preprocessing the new data, the method will be tested. In the validation process, it is divided into two frames, namely Training and Testing. In the training section, researchers use the Support Vector Machine (SVM) algorithm which is then linked to the Apply Model and Performance in the Testing section. The following is an image of testing the SVM method using Rapid Miner:



FIGURE 3. Testing Support Vector Machine (SVM)

The followings are the results obtained when testing data using the SVM method, an accuracy value of 75.92%, here are the results in the figure:

accuracy: 75.92% +/- 7.83% (mikro: 75.90%)				
	true Review_Negative	true Review_Positif	class precision	
pred. Review_Negative	61	10	85.92%	
pred. Review_Positif	37	87	70.16%	
class recall	62.24%	89.69%		

FIGURE 4. Results of the Accuracy Value

The following is the results of the ROC curve (AUC), the value that is equal to 0.880%, following the results of the AUC on Rapid Miner:



FIGURE 5. ROC (AUC) curve on SVM

c. After the researcher tested the dataset on the SVM method, it turned out that the accuracy value could still be improved by using the Particle Swarm Optimization feature (PSO) the following is Main Process of testing the dataset on the comparison of the SVM method with PSO optimization:



FIGURE 6. Testing Support Vector Machine (SVM) with PSO Optimization

The followings are the results obtained when testing the data by using the SVM method which has been compared with PSO optimization, the accuracy value is 81.05%, here are the results in the figure:

accuracy: 81.05% +/- 7.18% (mikro: 81.03%)				
	true Review_Negative	true Review_Positif	class precision	
pred. Review_Negative	88	27	76.52%	
pred. Review_Positif	10	70	87.50%	
class recall	89.80%	72.16%		

FIGURE 7. Results of the Accuracy Vector Machine (SVM) with PSO Optimization

And following the results of the ROC curve (AUC), the value is equal to 0.881%, the results of the AUC on Rapid-Miner is as follows:



FIGURE 8. ROC (AUC) curve on SVM with PSO

Classification or grouping of data by applying the Particle Swarm Optimization (PSO) method resulted in a value of 81.05% and an AUC value = 0.881.

TABLE 1. Kano Calculations 5 v M whill 1 50

- SVM SVM-PSO			
5 (11) 5 (11) 50	SVM-PSO	SVM	-
Accuracy 75.92% 81.05%	81.05%	75.92%	Accuracy
AUC Value 0.880 0.881	0.881	0.880	AUC Value

The researchers provide an application to test the model with a different dataset from commented reviews on one of the online marketplaces that have not been classified. Then, it will be selected on words related to sentiments that often appear and have the highest weight. Researchers apply it to the Web-based PHP programming language. It can be seen in the figure below:

The application above will determine the classification based on the sample comments inputted into positive, negative and neutral groups.

CONCLUSIONS

Classification or grouping of data by applying the Support Vector Machine (SVM) method is able to produce a fairly good accuracy value and the value of AUC quite sufficient, while the classification or grouping of data by applying the Particle Swarm Optimization (PSO) method produces a better accuracy values and has increased. It was concluded that by applying the SVM method and given PSO optimization for maximum accuracy in processing the dataset in the form of text classification in the marketplace review. From the results of testing these data, it will be able to help customers in determining a good and trusted marketplace by knowing which words are related to the sentiments that often appear and have the highest weight.



FIGURE 9. Marketplace Product Assessment Results

REFERENCES

- 1. E. Maier and J. Wieringa, "Acquiring customers through online marketplaces? The effect of marketplacesales on sales in a retailer's own channels," Int. J. Res. Mark, (2020).
- 2. Y. Y. Chang, S. C. Lin, D. C. Yen, and J. W. Hung, "The trust model of enterprise purchasing for B2B e-marketplaces," Comput. Stand. Interfaces, vol. 70, no. February, p. 103422, (2020).
- 3. S. Ayu and A. Lahmi, "Peran E-commerce Terhadap Perekonomian Indonesia Selama Pandemi Covid-19," J. Kaji.Manaj. Bisnis, vol. 9, no. 2, pp. 114–123, (2020).
- S. Smetanin and M. Komarov, "Sentiment Analysis of Product Reviews in Russian Using Convolutional NeuralNetworks," in Proceedings - 21st IEEE Conference on Business Informatics, CBI 2019, 2019, vol. 1, pp. 482–486, (2019)
- 5. E. Kipnis et al., "Institutionalizing Diversity-and-Inclusion-Engaged Marketing for Multicultural MarketplaceWell-Being," J. Public Policy Mark., vol. 40, no. 2, pp. 143–164, (2021).
- R. Meifitrah, I. Darmawan, and O. NurulPratiwi, "Sentiment analysis of tokopedia application review to service product recommender system using neuralcollaborative filtering for marketplace in Indonesia," IOP Conf. Ser. Mater. Sci. Eng., vol. 909, no. 1, (2020).
- M. Azhar, N. Hafidz, B. Rudianto, and W. Gata, "Marketplace Sentiment Analysis Using Naive Bayes And SupportVector Machine," PIKSEL Penelit. Ilmu Komput. Sist. Embed. Log., vol. 8, no. 2, pp. 91–100, (2020).
- M. I. Ahmadi, F. Apriani, M. Kurniasari, S. Handayani, and Gustian D, "Sentiment Analysis Online Shop on thePlay Store Using Method Support Vector Machine (Svm)," in Seminar Nasional Informatika 2020, X,vol. 2020,no. Semnasif, pp. 196–203.(2020).
- S. D. Anggita and Ikmah, "Komparasi Algoritma Klasifikasi Berbasis Particle Swarm Optimization Pada AnalisisSentimen Ekspedisi Barang," J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 4, no. 2, pp. 362– 369, (2020).
- A. Z. Rasyida, I. D. Wijaya, and Y. Yunhasnawa, "Analisis Sentimen Kualitas Layanan Online Marketplace DiIndonesia Menggunakan Metode Support Vector Machine," Semin. Inform. Apl. Polinema 2020, vol. 5, no. 1, pp.70–75, (2020).
- 11. S. Yadav, A. Ekbal, and S. Saha, "Information theoretic-PSO-based feature selection: an application in biomedicalentity extraction," Knowl. Inf. Syst., vol. 60, no. 3, pp. 1453–1478, (2018).
- 12. S. Ray, M. M. Rahman, M. Haque, M. W. Hasan, and M. M. Alam, "Performance evaluation of SVM and GBMin predicting compressive and splitting tensile strength of concrete prepared with ceramic waste and nylon fiber," J.King Saud Univ. Eng. Sci., no. xxxx, (2021).
- N. A. Utami, W. Maharani, and I. Atastina, "Personality Classification of Facebook Users According to Big FivePersonality Using SVM (Support Vector Machine) Method," Procedia Comput. Sci., vol. 179, no. 2020, pp. 177–184, (2021)
- 14. A. P. Giovani, Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," J. TEKNOINFO, vol. 14, no. 2, pp. 116–124, 2020.
- 15. S. Raj and K. C. Ray, "ECG Signal Analysis Using DCT-Based DOST and PSO Optimized SVM," IEEE Trans.Instrum. Meas., vol. 66, no. 3, pp. 470–478, (2017).