Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 2, April 2023
DOI : https://doi.org/10.33395/sinkron.v8i2.11994
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Implementation Opinion Mining For Extraction Of Opinion Learning In University

Mariana Purba[1)], Yadi[2)*]
[1]Informatika, Universitas Sjakhyakirti, Palembang, Indonesia
[2]Teknik Informatika, Institut Teknologi Pagar Alam, Pagar Alam, Indonesia
[1)]riagalihprasojo@gmail.com,[2)]Yadimkom@gmail.com

Submitted : Dec 13, 2022 | **Accepted** : Dec 16, 2022 | **Published** : Apr 1, 2023

Abstract: Opinion mining is a field of Natural Language Processing (NLP) that is used to carry out the process of extracting and processing textual data which functions to obtain information through sentiment analysis from a document in the form of text, among others, to detect attitudes towards objects or people. Sub-processes in opinion mining can use documents of subjectivity, opinion orientation, and detection targets to find out the data used as sentiment analysis, sentiment analysis aims to assess emotions, attitudes, opinions, and evaluations conveyed by a speaker or writer towards a product or towards a public figure. In this study, an opinion mining system was developed to analyze learning in college. The methodology used is quantitative descriptive, while the processing of sentiment analysis data uses Azure machine learning. Sentiment analysis results are very good at assessing opinions or opinions and emotions, and attitudes conveyed by someone. The learning process is the main element that must run well so that competency and achievement in learning can be maximally conveyed to students. Documents that identified opinions were then classified into negative, neutral, and positive opinions based on the results. In general, it can be concluded that the value obtained by sentiment analysis using Azure Machine Learning tools is quite good, judging from the results of a positive class of 0.79 and a neutral class of 0.53. The use of cleaning and labeling techniques and other classifications is still very possible to use. To get a better accuracy value.

Keywords: Data Mining, Extraction, Machine Learning, Opinion, Sentiment Analysis.

INTRODUCTION

Opinion mining is a field of natural language processing (NLP) such as the process of understanding, extracting and processing textual data automatically to obtain sentiment information contained in an opinion sentence (Santoso & Nugroho, 2019). The learning process is carried out to provide knowledge and improve abilities in both soft skills and hard skills. Learning competencies are achievements that must be possessed by all students, therefore the learning process must be able to run well (Singgalen, 2021). Evaluation of the learning process is carried out to improve the quality and measurement of the implementation of the learning program by using opinions from students on the ongoing learning process through sentiment analysis. Sentiment analysis is carried out to see the opinion or tendency of an opinion towards a problem or object by a person, by labeling opinions such as positive, neutral and negative (Budi, H, R, R, & D. T, 2019).

The use of sentiment analysis serves to identify the tendency of opinion on an object (Alifia, 2020). The influence and great benefit on the sentiment analysis process have been widely used to make a decision, this happens because sentiment analysis as a combination of data used comes from text mining (A. Nisa, 2019). Sentiment Analysis is one of the branches of text mining research, that

*name of corresponding author

deals with a broader field (Alita & Isnain, 2020). Text Mining is a stage of the process of analysis of data in the form of text where the data source is obtained from a document such as word data sentences, The concept of text mining is usually used in the classification of textual documents (Latifah, 2018).

One of the methods of text mining that can be used to solve opinion mining problems using Azure Machine Learning tools, which functions very well as a method of classifying text to be used as labeling on documents by looking at scores on opinions that have been processed (Dharmapatni, 2020)

This research is aimed at developing an opinion mining system to process opinion data on the learning process at a university (Laurensz & Sediyono, 2021)(Rusydiana & Marlina, 2020). The opinion processing process can use several sub-process documents, namely, sub-process document subjectivity, opinion orientation and target detection. The document subjectivity sub-process is aimed at recognizing the subjectivity of a text document (which text documents include opinions and do not include opinions) (D. Alita, 2020). The opinion orientation sub-process is used to determine the orientation of a sentence of opinion, whether it belongs to a positive or negative orientation (Mahendrajaya & Buntoro, 2019).

The data will be analyzed using rules to determine which documents include opinions and to determine the object that is the target of the opinion. In the target detection sub-process used Azure Machine Learning (Fikri et al., 2020).

LITERATURE REVIEW

Sentiment Analysis

Sentiment analysis is an ongoing field of text-based research. Sentiment analysis or opinion mining is the study of how to solve the problem of opinion society (Darwis et al., 2021) (Kurniawan & Susanto, 2019), the attitudes and emotions of such an entity can represent the individual. Sentiment analysis was done to see the opinion or propensity of the opinion towards a problem or object by a person, whether prone to negative or positive opinion (Alifia, 2020).

Preprocessing

Preprocessing is a process to prepare raw data before other processes are carried out. In general, data preprocessing is done by eliminating inappropriate data or transforming data into easier forms processed by the system (Nasution & Hayaty, 2019)(Fauziyyah, 2020).

Text mining

Text mining is the process by which data sources are usually obtained from documents to search words that can represent the contents of the document so that analysis can be carried out (Yulita, 2021). The purpose of text mining exists to get useful information from a set of documents, data used in text mining is a set of text that has an unstructured format or minimal semi-structure. The specific task of text mining is the grouping of text (Lengkong et al., 2021).

METHOD

The research approach uses quantitative descriptive with the help of Azure Machine Learning as a tool to perform sentiment analysis of the extraction of opinion mining of the learning process (Rachman & Pramana, 2020). The stages in the study can be seen in Figure 1.

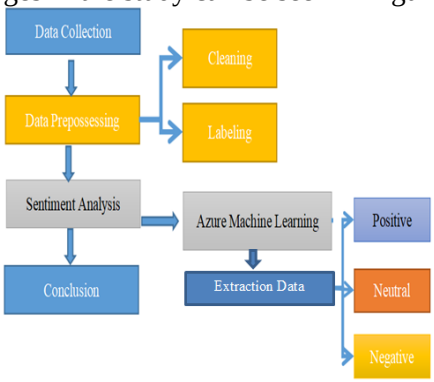


Fig 1. Stages of Research

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 2, April 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i2.11994 p-ISSN : 2541-044X

Data Collection

The data used in this study is learning process evaluation data taken using the Google form instrument from students who are respondents. This study only used data in Indonesian.

Data Prepossessing

The prepossessing stage is carried out after the data collection stage through the instrument deployment process. Prepossessing data includes cleaning and labeling. Cleaning to eliminate identity data, study program, semester, and gender. Labeling to provide a post label for positive value sentiment and a neg label for negative value, as well as a net for neutral sentiment.

Sentiment Analysis

Sentiment analysis is performed based on the stages according to Figure 1, which is processed using Azure Machine Learning tools with some sentiment analysis results (Fairuz & Ramadhani, 2021)(Giovani et al., 2020). The working principle of Azure Machine Learning that will be used in this study is to automatically classify based on text data. F-measure is principally used in retrieval systems to measure document search classification as well as classification query performance (Tuhuteru, 2020). The F-measure is focused only on the calculation of value. However, as it developed, the F-measure placed more emphasis on precision and recall performance The F-measure is expressed in the equation (Samsir et al., 2021):

$$f - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (14)$$

Precision is expressed in the equation:

$$Precision = \frac{TruePositive}{Falsenegative + Truepositive} \quad (14)$$

Recall is expressed in the equation:

$$recall = \frac{TruePositive}{Falsenegative + Truepositive} \quad (14)$$

RESULT

Sentiment analysis is excellent in assessing opinions or opinions as well as emotions, and attitudes conveyed by a person. The learning process is the main element that must run very well so that competencies and achievements in learning can be fully conveyed to students. The learning evaluation process needs to be carried out using various data sources related to the learning process, approach strategies in collecting evaluation data can use student opinion mining data. The opinion mining extraction process to find out positive, neutral, and negative responses goes through several stages including:

Data Exposure

Opinion data were taken using instruments spread with pedagogy indicators. Data was obtained involving 39 Respondents. The data is raw data including opinions on learning readiness, learning motivation, and the relevance of competencies in learning. Figure 2 is the data obtained at the data collection stage:

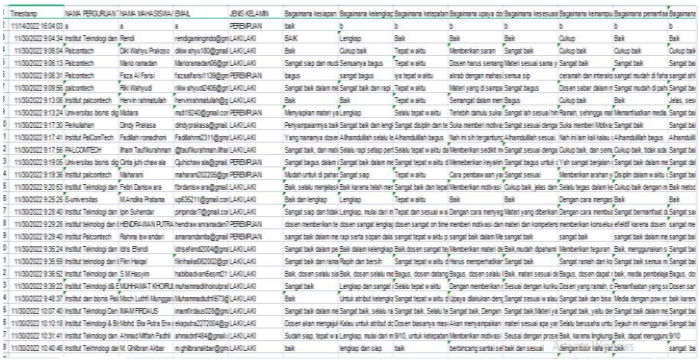


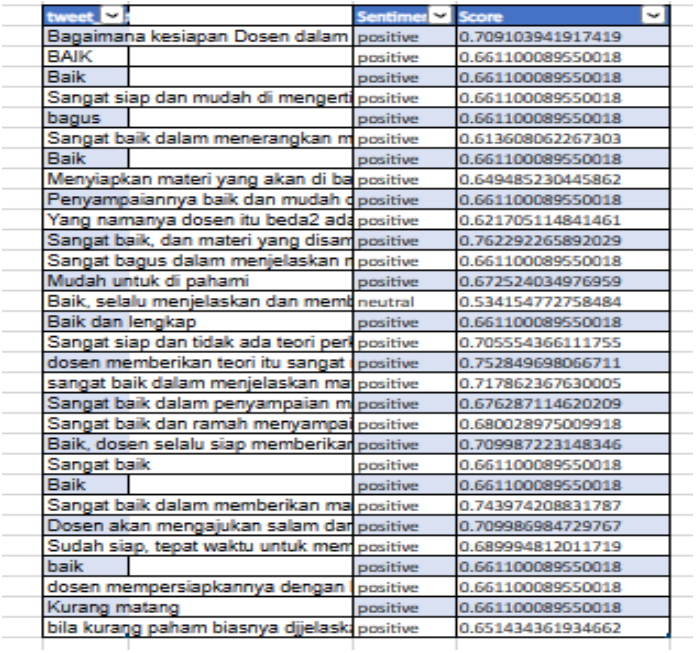
Fig 2. Data Opinion Mining

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 2, April 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i2.11994 p-ISSN : 2541-044X

Data Prepossessing

Prepossessing data to data includes cleaning and labeling. The cleaning process by removing opinions on learning readiness, learning motivation, and the relevance of competencies in learning leaves ready-made data. Among these data, there are 29 opinions data with positive labels and 10 opinions with neutral labels. Figure 3 is the result of prepossessing data processed using Azure Machine Learning:



tweet	Sentimer	Score
Bagaimana kesiapan Dosen dalam	positive	0.709103941917419
BAIK	positive	0.661100089550018
Baik	positive	0.661100089550018
Sangat siap dan mudah di mengerti	positive	0.661100089550018
bagus	positive	0.661100089550018
Sangat baik dalam menerangkan m	positive	0.613608062267303
Baik	positive	0.661100089550018
Menyiapkan materi yang akan di ba	positive	0.649485230445862
Penyampaiannya baik dan mudah d	positive	0.661100089550018
Yang namanya dosen itu beda2 ada	positive	0.621705114841461
Sangat baik, dan materi yang disam	positive	0.762292265892029
Sangat bagus dalam menjelaskan n	positive	0.661100089550018
Mudah untuk di pahami	positive	0.672524034976959
Baik, selalu menjelaskan dan memb	neutral	0.534154772758484
Baik dan lengkap	positive	0.661100089550018
Sangat siap dan tidak ada teori peri	positive	0.705554366111755
dosen memberikan teori itu sangat	positive	0.752849698066711
sangat baik dalam menjelaskan ma	positive	0.717862367630005
Sangat baik dalam penyampaian m	positive	0.676287114620209
Sangat baik dan ramah menyampai	positive	0.680028975009918
Baik, dosen selalu siap memberikan	positive	0.709987223148346
Sangat baik	positive	0.661100089550018
Baik	positive	0.661100089550018
Sangat baik dalam memberikan ma	positive	0.743974208831787
Dosen akan mengajukan salam dan	positive	0.709986984729767
Sudah siap, tepat waktu untuk mem	positive	0.689994812011719
baik	positive	0.661100089550018
dosen mempersiapkannya dengan	positive	0.661100089550018
Kurang matang	positive	0.661100089550018
bila kurang paham biasnya dijelask	positive	0.651434361934662

Fig 3. Data Prepossessing

Sentiment Analysis

At the sentiment analysis stage, using Azure Machine Learning with documents from respondents, the results of the sentiment analysis process with positive, neutral, and negative labels are shown in Table 1.

Table 1. Document Data

Document	Label
29	Positive
10	Neutral
0	Negative

Based on the document data used as many as 39 data sources from respondents, the analysis was carried out to see positive, neutral, and negative labels on opinion mining. The test results are shown in Table 2.

Table 2. Test Result

Test	Positive Class	Neutral Class
Precision	0.794555	0.534154
Recall	0.794555	0.534154

Tabel 3. Result F-measure

Test	Positive Class	Neutral Class
F-measure	0.794555	0.534154

A high precision value indicates a small False Positive value (predictably positive, actually negative). A high recall value indicates a small false negative value (predictably negative, actually positive). A good F-measure value indicates that the sentiment classification model has good precision and recall.

*name of corresponding author

DISCUSSIONS

In Sentiment analysis is very well used to process data on opinions in the form of text, by using sentiment analysis we can find out the response of an object through text by using Azure Machine Learning The processing of each text data is getting better, this can be seen from the labeling results that are automatically carried out in software with several labeling criteria such as positive responses, neutral and negative. Opinion data extraction is highly appropriate using azure machine learning so that stakeholders can easily make a decision.

CONCLUSION

Based on the results of the discussion, it can be concluded that sentiment analysis using Azure machine learning can be applied to predict the direction of a person's sentiment toward learning, whether positive, neutral or negative sentiment. A high precision value indicates a small False Positive value (predictably positive, actually negative). In general, it can be concluded that the value obtained by sentiment analysis using Azure Machine Learning tools is quite good to be seen in opinion positive class results of 0.79 and opinion neutral classes of 0.53 on the learning process that has been carried out by lecturers at university in preparation for lecture. The use of cleaning and labeling techniques and other classifications is still very possible to use. To obtain better accuracy values, subsequent research can use more datasets.

ACKNOWLEDGMENT

We as researchers would like to thank our fellow authors and colleagues who have helped in completing this article and we hope that this article can provide benefits to all of us and colleagues, and can be applied and implemented in every institution that wants to use it as a reference. We hope that the publisher can accept and publish it and thank you.

REFERENCES

Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, *8*(2), 50–58. https://doi.org/https://doi.org/10.30864/eksplora.v9i1.254

Alifia, R. P. (2020). Sentiment Analysis for Customer Review. Journal of Information System Engineering and Business Intelligenc , 1-8.

A. Nisa, E. D. (2019). Analisis Sentimen Menggunakan Naive Bayes Classifier dengan Chi-Square Feature Selection Terhadap Penyedia Layanan Telekomunikasi. eProceeding Eng., vol. 6, no. 2, (pp. 8650–8659).

Budi, H, Y., R, F., R, J., & D. T, R. (2019). Facebook Analysis of Community Sentiment on 2019 Indonesian. The Fifth Information Systems International Conference, (pp. 715-722).

Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, *15*(1), 131–145. https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/view/744

Dharmapatni, P. &. (2020). Penerapan Algoritma Support Vector Machine Dalam Sentimen Analisis Terkait Kenaikan Tarif BPJS Kesehatan. Jurnal BITe. Vol 2 No 2. , 1-8.

D. Alita, Y. F. (2020). Implementasi Algoritma Multiclass SVM pada Opini Publik Berbahasa Indonesia di Twitter. J. Tekno Kompak, vol. 14, no. 2 , 86–91.

Fairuz, A. L., & Ramadhani, R. D. (2021). Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Jurnal DINDA (Indonesian Journal of Data Science, IOT, Machine Learning and Artificial Intelligence)*, *1*(1), 42–51. https://doi.org/DOI: https://doi.org/10.20895/dinda.v1i1.180

Fauziyyah, A. K. (2020). Analisis sentimen pandemi Covid19 pada streaming Twitter dengan text mining Python. *Jurnal Ilmiah SINUS*, *18*(2), 31–42. https://p3m.sinus.ac.id/jurnal/index.php/e-jurnal_SINUS/article/view/491

Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *Smatika Jurnal*, *10*(2), 71–76. https://doi.org/https://doi.org/10.32664/smatika.v10i02.455

*name of corresponding author

Giovani, A. P., Ardiansyah, A., & Haryanti, T. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*, *14*(2), 116–124. https://doi.org/DOI: https://doi.org/10.33365/jti.v14i2.679

Kurniawan, I., & Susanto, A. (2019). Implementasi Metode K-Means dan Naive Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019. *Jurnal Eksplora Informatika*, *9*(1), 1–10. https://doi.org/https://doi.org/10.30864/eksplora.v9i1.237

Laurensz, B., & Sediyono, E. (2021). Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19. *Jurnal Nasional Teknik Elektro Dan …*, *10*(2), 118–123. https://doi.org/https://doi.org/10.22146/jnteti.v10i2.1421

Latifah, E. F. (2018). Perbandingan Kinerja Machine Learning Berbasis Algoritma Support Vector Machine dan Naive Bayes (Studi Kasus: Data Tanggapan Mengenai Traveloka Melalui Media Sosial Twitter) .

Lengkong, N. C., Safitri, O., Machsus, S., Putra, Y. R., & ... (2021). Analisis Sentimen Penerapan Psbb Di Dki Jakarta Dan Dampaknya Terhadap Pergerakan Ihsg. *Jurnal Teknokrat*, *15*(1), 20–25. https://doi.org/DOI: https://doi.org/10.33365/jti.v15i1.866

Mahendrajaya, R., & Buntoro, G. A. (2019). Analisis Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based Dan Support Vector Machine. *Jounal Komputek*, *3*(2), 52–63. https://doi.org/DOI: 10.24269/jkt.v3i2.270

Nasution, M. R. A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *Jurnal Informatika*, *6*(2), 226–235. https://pdfs.semanticscholar.org/f0d1/5870ff2ab5995bf3040dd7c0939e2897a94c.pdf

Rachman, F. F., & Pramana, S. (2020). Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. *Indonesian of Health Information Management Journal*, *8*(2), 100–109. https://doi.org/DOI: https://doi.org/10.47007/inohim.v8i2.223

Rusydiana, A. S., & Marlina, L. (2020). Analisis sentimen terkait sertifikasi halal. *JEBA (Journal of Economics and …*, *5*(1), 69–85. https://system4.yarsi.ac.id/index.php/jeba/article/view/1405

Samsir, S., Ambiyar, A., Verawardina, U., & ... (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes. *Jurnal Media Informatika Budidarma*, *5*(1), 157–163. https://doi.org/DOI: http://dx.doi.org/10.30865/mib.v5i1.2580

Santoso, E. B., & Nugroho, A. (2019). Analisis sentimen calon presiden indonesia 2019 berdasarkan komentar publik di facebook. *Jurnal Eksplora Informatika*, *9*(1), 60–69. https://doi.org/https://doi.org/10.30864/eksplora.v9i1.254

Singgalen, Y. A. (2021). Pemilihan metode dan algoritma dalam analisis sentimen di media sosial: sistematic literature review. *Journal of Information Systems and Informatics*, *3*(2), 278–302. https://doi.org/DOI: 10.33557/journalisi.v3i2.125

Tuhuteru, H. (2020). Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine. *Journal Information System Development (ISD)*, *5*(2), 7–13. https://ejournal.medan.uph.edu/index.php/isd/article/view/381

Yulita, W. (2021). Analisis sentimen terhadap opini masyarakat tentang vaksin covid-19 menggunakan algoritma naïve bayes classifier. *Jurnal Data Mining Dan Sistem Informasi*, *2*(2), 35–45. https://ejurnal.teknokrat.ac.id/index.php/JDMSI/article/view/1344

*name of corresponding author