BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

Analysis Of Public Sentiment Towards The Corruption Eradication Commission On Twitter

Siti Nurhaliza Sofyan, Sri Wahyuni

[1]Master of Information Technology, University Pembangunan Pancabudi, Medan, Indonesia
Email :[1] sitinurhalizas102@gmail.com, sriwahyuni@dosen.pancabudi.ac.id

Abstract - The Corruption Eradication Commission (KPK) is a state institution in Indonesia which was formed to eradicate corruption. The Corruption Eradication Committee (KPK) [1] has the main task of carrying out investigations, inquiries and prosecutions of criminal acts of corruption. This institution is independent and free from the influence of any power in carrying out its duties and authority [2]. This research explores the analysis of Indonesian people's sentiment towards the KPK in the current situation such as arrests for corruption and the policies and actions carried out by the KPK. Sentiment analysis used in the journal with data obtained from Twitter data and using Orange Data Mining, with multilingual sentiment analysis techniques to analyze Indonesian people's sentiment towards the KPK agency. The results of sentiment analysis are the emotion 'surprise' totals 148 or approximately 44%, the emotion 'joy' totals 96 or approximately 35%, and the emotion 'fear' totals 43 or approximately 16% visualized through box plots and scatter plots, which aim to classify Twitter users based on their emotional responses. The findings of this research provide valuable insight into the landscape of sentiment surrounding the Corruption Eradication Commission's as well as providing sustainable benefits and are expected to be used as material for evaluating the government's role [3]. The aim of this research is to find out how much public sentiment is towards the Corruption Eradication Commission on the Twitter application [4][5][6].

Keywords: Sentiment Analysis, Corruption Eradication Commission, Orange Data Mining, Multilingual Sentiment

1. INTRODUCTION

Nowadays, many people socialize not only when they meet but online is now much more widely used. There are many social media platforms, one of which is Twitter (X) where many people, especially young people and even the millennial generation, generation Z, use it because it is easy to express ideas and even comments on Twitter social media. The ease of social media now even means that people can easily get the latest information concisely and quickly so that one of them, the Corruption Eradication Commission (KPK), has not escaped the public's attention. The public really wants the KPK to be able to carry out its duties fairly, quickly and responsively, not just selectively in cases. The KPK or Corruption Eradication Commission is an independent institution in Indonesia that has a mandate to combat corruption. One of the main tasks of the KPK is to investigate corruption cases involving the public or public officials. Sentiment analysis is an automatic text processing or data extraction process to obtain sentiment information contained in a sentence. Sentiment analysis will classify the sentiment of textual documents applied to comments, certain products, or topics with positive or negative categories. Previous research on a related topic, namely analysis of Indonesian people's sentiment towards the transfer of the capital city of the archipelago using the Naive Bayes and K-Nearest Neighbor algorithms, presents the results of a Naive Bayes comparison of the performance of these methods that the Naïve Bayes method provides a sentiment analysis accuracy level of 82.27%, a precision value of 86.36% and the Recall value is 76.93%. The KNN method also presents analysis results with an accuracy level of 88.12%, precision of 93.%98 and recall value of 81.53%. Based on the results of this analysis, the analysis process using the KNN method outperformed the Naive Bayes method in this research [5]**.** In related research, namely Analysis of Public Opinion Sentiment towards Films on the Twitter Platform Using the Naive Bayes Algorithm, Based on the test results using the confusion matrix with orange tools, the average accuracy value was 0.65% and the precision value was 0.67%, and the recall was 0.65%, and the neutral percentage was 0.83% [7]. In another journal, namely Implementation of Data Mining with the Naïve Bayes Algorithm for Eligibility Classification of Basic Food Aid Recipients with 135 training data with 40 testing data and seven attributes resulting in 86% accuracy, 85% recall, and 88% precision [8]. In the form of negative sentiment related to the KPK Bill issue was 60.9% greater than positive sentiment of 39.1%. The SVM model classifies sentiment quite well because it has accuracy, sensitivity and specificity values of 81.32%, 71.47% and 87.64% respectively [1]. Related journals discussing sentiment towards electric vehicle batteries resulted in 37% responding positively, 42% responding neutrally, and 21% responding negatively [9].

Naive Bayes is a machine learning algorithm used to classify data based on probability. This algorithm is based on Bayes' Theorem, a simple probability theory. According to Olson Delen [7] Naive Bayes for each decision class, calculates the probability on the condition that the decision class is correct, given the object information vector. Sentiment analysis or known as opinion mining is an automatic process in understanding, extracting and processing textual data to obtain information [10]. Sentiment analysis is a field of science that analyzes[11] opinions, attitudes, evaluations, and assessments of an event, topic, organization, or individual. The classification method used in this study is the Naive Bayes Classifier. The Naive Bayes Classifier is combined with features to detect negation and weighting using term frequency and TF-IDF [12]. The classification model is used for testing data to carry out the classification process that produces

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

sentiment labels (positive/negative), this process is called the testing process [2], If these reviews are collected and then processed, the results can be used as one that has the best sentiment [7].

.2. RESEARCH METHODOLOGY

2.1 Research Stages

The research method used is the experimental method by observing the variables of the object being studied. The experimental method aims to test the effect of a variable on another variable or to test the causal relationship between variables. The steps in designing the research are as follows:

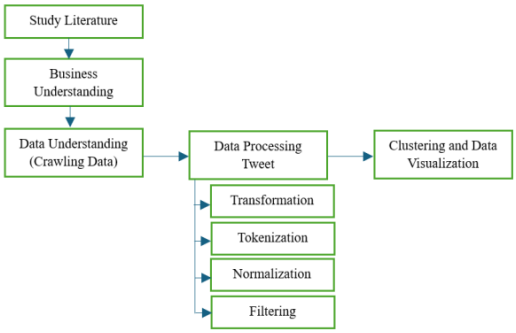


Figure 1. Research Stages Flow

1. Literature Study
   Previous research and literature study define and examine aspects related to text mining.
2. Business Understanding
   Analyze issues and facts surrounding the Corruption Eradication Commission published by the government that are currently circulating in the community.
3. Data Understanding (Crawling Data)
   Utilize comments given by the public on the Twitter social media platform. Opinion collection is carried out using the manual tweet collection method, namely entering keywords related to the Corruption Eradication Commission into the Twitter search column.
4. Tweet Data Processing
   Determining class attributes and loading dictionaries involves matching root words with sentiment word dictionaries to determine sentiment content (positive, neutral, negative). All tweet data is labeled according to class, with three classes to be used in this study: positive class, negative class, and neutral class.
5. Data Grouping and Visualization
   Grouping text mining data with Orange Data Mining involves Box Plot and Scatter Plot visualizations, which visualize text mining data processed with Twitter user emotions [3].

2.2 Data Mining

Data mining is a process of extracting[13] or mining large amounts of data and information, which are previously unknown, but can be understood and used from large databases and used to make very important business decisions [1], [14]. Data mining describes a collection of techniques with the aim of finding unknown patterns in the data that has been collected. In this study, Orange is used as a medium for data mining. Orange [2], [9], [15] or also known as Orange Data Mining is an open source software for carrying out data mining or data analytics processes through the concept of visual programming [16][8] [17] .

2.3 Classification

In machine learning, classification is the task of predicting the class label of a given sample based on a set of features or characteristics. It is the activity of converting tweet data into sentiment whether positive, negative or neutral using the Naive Bayes Classification method.

2.4 Algoritma Naive Bayes

The Naive Bayes method is a basic and well-known machine learning monitoring algorithm based on the Bayesian theorem and the assumption of autonomous characteristic conditions. This method postulates accessibility at the document stage with predefined views and labels. Naïve Bayes comes from two syllables, namely Naïve which comes from the assumption of the occurrence of another feature so that each feature contributes individually to classification without dependence on other features. Meanwhile, Bayes comes from the principle of Bayes' theorem. One of the advantages of using Naïve Bayes to classify is that classification uses a small amount of training data. In short, the naïve

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

Bayes classification algorithm is the prediction of all probabilities for classifying a set of data, an algorithm that is often used because it allows it to work more and is more complex than real situations.

Sample data from Twitter is as follows:

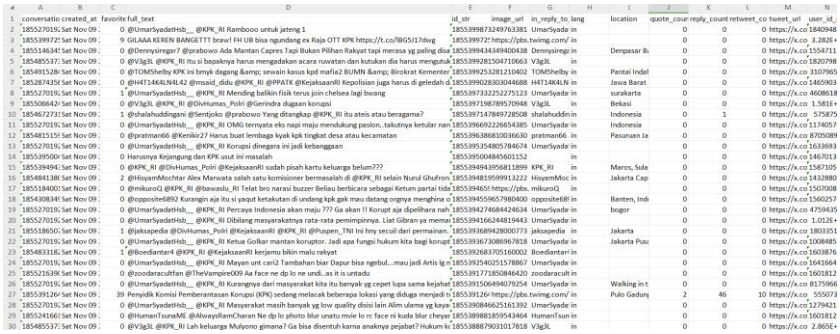


Figure 2. Dataset from Twitter

3. RESULT AND DISCUSSION

3.1 Research Analysis

The implementation of Orange Data Mining features a sentiment analysis widget interface design integrated into the workflow, as illustrated in Figure 3 below:

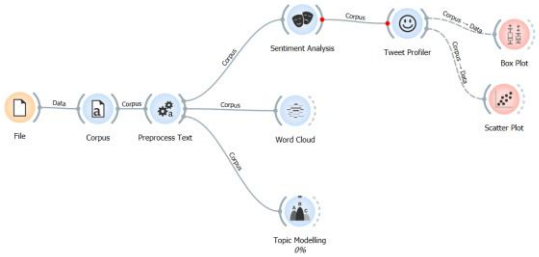


Figure 3. Sentiment Analysis Data Widget

The data crawled from the social media platform Twitter will be input and analyzed individually based on objects. Subsequently, it will be connected to the necessary widgets for research purposes, resulting in a widget design as shown in the figure above .

3.2 Crawling Data

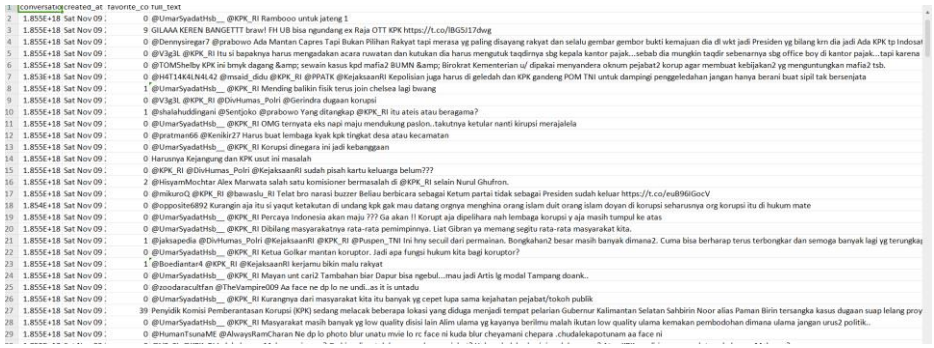
In this study, the research data consists of comments from Indonesian society on Twitter regarding electric motorcycles from Oktober 1, 2024, to November 1, 2024. The dataset for this research was obtained from a Python program written in Google Colab, as depicted in Figure 4 below.

```
[ ]  # Crawl Data

     filename = 'kpk.csv'
     search_keyword = 'kpk since:2024-10-01 until:2024-11-01 lang:id'
     limit = 300

     !npx -y tweet-harvest@2.6.1 -o "{filename}" -s "{search_keyword}" --tab "LATEST" -l {limit} --token {twitter_auth_token}
```

Figure 4. Python Code for Crawling Data

The code above describes that data crawling is performed using the keyword ‘kpk’ with comments written in Indonesian language from 2024-10-01 to 2024-11-01. A total of 300 tweets will be crawled, and the crawling results will be exported to a file named ‘kpk.csv’. From the data crawling results based on the above program, 300 comments were obtained. Subsequently, the file is imported into Orange Data Mining as shown in Figure 5 below.

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT



	conversation_created_at	favorite_co	full_text
2	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Rambooo untuk jateng 1
3	1.855E+18	Sat Nov 09 :	9 GILAAA KEREN BANGETTT brawl FH UB bisa ngundang ex Raja OTT KPK https://t.co/lBGSi17dwg
4	1.855E+18	Sat Nov 09 :	0 @Dennysiregar7 @prabowo Ada Mantan Capres Tapi Bukan Pilihan Rakyat tapi merasa yg paling disayang rakyat dan selalu gembar gembor bukti kemajuan dia di wkt jadi Presiden yg bilang krn dia jadi Ada KPK tp Indosat
5	1.855E+18	Sat Nov 09 :	0 @V3g3L @KPK_RI Itu si bapaknya harus mengadakan acara ruwatan dan kutukan dia harus mengutuk taqdirnya sbg kepala kantor pajak...sebab dia mungkin taqdir sebenarnya sbg office boy di kantor pajak...tapi karena
6	1.855E+18	Sat Nov 09 :	0 @TOMShelby KPK ini bnyk dagang & sewain kasus kpd mafia2 BUMN & Birokrat Kementerian u/ dipakai menyandera oknum pejabat2 korup agar membuat kebijakan2 yg menguntungkan mafia2 tsb.
7	1.853E+18	Sat Nov 09 :	0 @H4T14K4LN4L42 @msaid_didu @KPK_RI @PPATK @KejaksaanRI Kepolisian juga harus di geledah dan KPK gandeng POM TNI untuk dampingi penggeledahan jangan hanya berani buat sipil tak bersenjata
8	1.855E+18	Sat Nov 09 :	1 @UmarSyadatHsb__ @KPK_RI Mending balikin fisik terus join chelsea lagi bwang
9	1.855E+18	Sat Nov 09 :	0 @V3g3L @KPK_RI @DivHumas_Polri @Gerindra dugaan korupsi
10	1.855E+18	Sat Nov 09 :	1 @shalahuddingani @Sentjoko @prabowo Yang ditangkap @KPK_RI itu ateis atau beragama?
11	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI OMG ternyata eks napi maju mendukung paslon..takutnya ketular nanti kirupsi merajalela
12	1.855E+18	Sat Nov 09 :	0 @pratman66 @Kenikir27 Harus buat lembaga kyak kpk tingkat desa atau kecamatan
13	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Korupsi dinegara ini jadi kebanggaan
14	1.855E+18	Sat Nov 09 :	0 Harusnya Kejangung dan KPK usut ini masalah
15	1.855E+18	Sat Nov 09 :	0 @KPK_RI @DivHumas_Polri @KejaksaanRI sudah pisah kartu keluarga belum???
16	1.855E+18	Sat Nov 09 :	2 @HisyamMochtar Alex Marwata salah satu komisioner bermasalah di @KPK_RI selain Nurul Ghufron.
17	1.855E+18	Sat Nov 09 :	0 @mikuroQ @KPK_RI @bawaslu_RI Telat bro narasi buzzer Beliau berbicara sebagai Ketum partai tidak sebagai Presiden sudah keluar https://t.co/euB96iGocV
18	1.854E+18	Sat Nov 09 :	0 @opposite6892 Kurangin aja itu si yaqut ketakutan di undang kpk gak mau datang orgnya menghina orang islam duit orang islam doyan di korupsi seharusnya org korupsi itu di hukum mate
19	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Percaya Indonesia akan maju ??? Ga akan !! Korupt aja dipelihara nah lembaga korupsi y aja masih tumpul ke atas
20	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Dibilang masyarakatnya rata-rata pemimpinnya. Liat Gibran ya memang segitu rata-rata masyarakat kita.
21	1.855E+18	Sat Nov 09 :	1 @jaksapedia @DivHumas_Polri @KejaksaanRI @KPK_RI @Puspen_TNI Ini hny secuil dari permainan. Bongkahan2 besar masih banyak dimana2. Cuma bisa berharap terus terbongkar dan semoga banyak lagi yg terungkap
22	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Ketua Golkar mantan koruptor. Jadi apa fungsi hukum kita bagi koruptor?
23	1.855E+18	Sat Nov 09 :	1 @Boediantar4 @KPK_RI @KejaksaanRI kerjamu bikin malu rakyat
24	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Mayan unt cari2 Tambahan biar Dapur bisa ngebul...mau jadi Artis lg modal Tampang doank..
25	1.855E+18	Sat Nov 09 :	0 @zoodaracultfan @TheVampire009 Aa face ne dp lo ne undi..as it is untadu
26	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Kurangnya dari masyarakat kita itu banyak yg cepet lupa sama kejahatan pejabat/tokoh publik
27	1.855E+18	Sat Nov 09 :	39 Penyidik Komisi Pemberantasan Korupsi (KPK) sedang melacak beberapa lokasi yang diduga menjadi tempat pelarian Gubernur Kalimantan Selatan Sahbirin Noor alias Paman Birin tersangka kasus dugaan suap lelang proy
28	1.855E+18	Sat Nov 09 :	0 @UmarSyadatHsb__ @KPK_RI Masyarakat masih banyak yg low quality disisi lain Alim ulama yg kayanya berilmu malah ikutan low quality ulama kemakan pembodohan dimana ulama jangan urus2 politik..
29	1.855E+18	Sat Nov 09 :	0 @HumanTsunaME @AlwaysRamCharan Ne dp lo photo blur unatu mvie lo rc face ni kuda blur cheyamani cheparа .chudalekapotunam aa face ni

Figure 5. Dataset

3.3 Preprocess Text

Before conducting text analysis, the text will first undergo preprocessing. This involves segmenting the text into smaller units (tokens), followed by transformation, tokenization, normalization, and filtering. Sequential steps in the analysis can be enabled or disabled within the Preproces Text widget in Orange Data Mining. Figure 6 below shows the steps performed in the preproces text widget in the Orange Data Mining application.

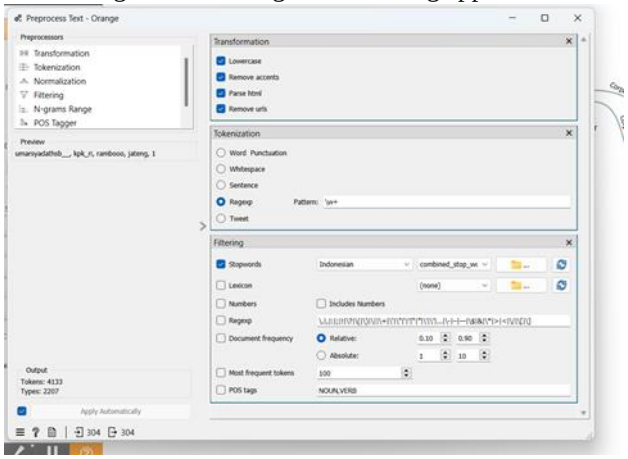


Figure 6. Preprocess Text

The steps carried out in the preprocess text in the Orange Data Mining application are as follows:

1. Transformation
   The first step is transformation, which involves transforming the entire text into lowercase, removing accents contained within the text, identifying HTML tags, parsing HTML tags, and removing URLs from the text.
2. Tokenization
   In this stage, sentences will be tokenized into words, preserving punctuation symbols.
3. Filtering
   In this stage, a process of removing or preserving selected words will be conducted. During this process, words that are not relevant to sentiment analysis will be removed. All words to be removed have been written into a file named 'combined_stop_words.txt' using the stopword widget. Additionally, the lexicon widget is utilized to extract tokens from the lexicon dictionary. The number widget is employed to remove meaningless numbers, while the regexp widget is used to eliminate tokens based on available regular expression patterns.

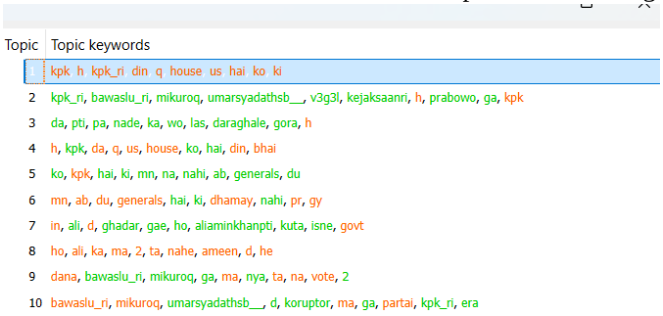
Once the preprocess text stage is completed, the text will be separated into individual words, and can observe the text distribution through a word cloud in Figure 7 below.

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

Figure 7. World Cloud

In the visualization above, a word cloud is displayed showing the frequency of the most commonly occurring words. The size of each word in the word cloud corresponds to its frequency of occurrence. The larger the word, the more frequent its occurrence. A word cloud serves as a visualization method to represent the results of text preprocessing. The visual settings and word variations enhance the attractiveness and comprehensibility of the visualization. The image depicts the preprocessing results of information previously comprised of comment lines on tweets about motorcycles from Oktober to November 2024.

In addition to the word cloud [18], after preprocessing the text, there is a widget in Orange Data Mining to view topic modeling. Topic modeling in Orange Data Mining is a text analysis process aimed at identifying and extracting topics or patterns that emerge within a collection of documents. For more details, please refer to Figure 8 below.



Topic	Topic keywords
	kpk, h, kpk_ri, din, q, house, us, hai, ko, ki
2	kpk_ri, bawaslu_ri, mikuroq, umarsyadathsb__, v3g3l, kejaksaanri, h, prabowo, ga, kpk
3	da, pti, pa, nade, ka, wo, las, daraghale, gora, h
4	h, kpk, da, q, us, house, ko, hai, din, bhai
5	ko, kpk, hai, ki, mn, na, nahi, ab, generals, du
6	mn, ab, du, generals, hai, ki, dhamay, nahi, pr, gy
7	in, ali, d, ghadar, gae, ho, aliaminkhanpti, kuta, isne, govt
8	ho, ali, ka, ma, 2, ta, nahe, ameen, d, he
9	dana, bawaslu_ri, mikuroq, ga, ma, nya, ta, na, vote, 2
10	bawaslu_ri, mikuroq, umarsyadathsb__, d, koruptor, ma, ga, partai, kpk_ri, era

Figure 8. Topic Modelling

3.4 Sentiment Analysis

The analysis process utilizes the MultiLingual Sentiment algorithm, which is capable of understanding opinions and viewpoints of users in various languages, in this case, using Bahasa Indonesia. The MultiLingual sentiment analyzer has been shown to provide a significant level of invariance compared to traditional sentiment analysis systems, such as the Vader variant, thus enhancing the accuracy and diversity of sentiment analysis. As depicted in Figure 9 below.

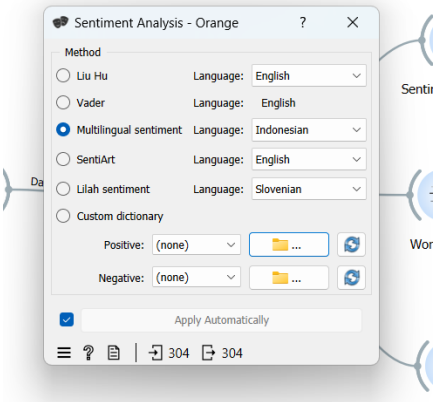


Figure 9. MultiLingual Sentiment Analysis

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

The MultiLingual sentiment approach enables a broader and more inclusive sentiment analysis, allowing text processing in various languages to understand the opinions, emotions, or sentiments contained within the text. Therefore, MultiLingual sentiment becomes crucial in the context of globalization and linguistic diversity in sentiment analysis and cross-cultural opinion understanding.

3.5 Tweet Profiler

Tweet Profiler is one of the features in the Orange Data Mining platform that enables this research to analyze sentiment from tweets or other text documents. By using Tweet Profiler, sentiment data can be retrieved from the available dataset through the server for each given tweet, and sentiment analysis can be conducted using various emotion classification methods provided, such as Ekman, Plutchik, and Profile of Mood States (POMS). Additionally, this feature allows for the utilization of specific attributes for analysis, such as content attributes, and performing emotion classification with multi-class options. Tweet Profiler is a valuable tool in text analysis and sentiment understanding in Orange Data Mining. In this study, a dataset of 300 tweets about motorcycles was utilized. The data was extracted using a widget from Orange Data Mining with Corpus and connected to Tweet Profiler using Ekman emotion. As depicted in Figure 10 below :

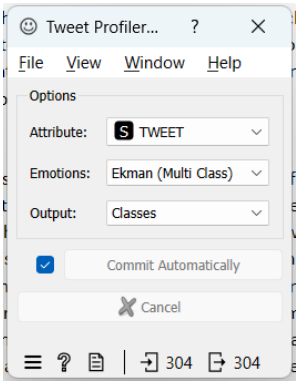


Figure 10. Tweet Profiler

3.6 Data Visualization

After performing tweet profiling in the Orange Data Mining widget, the next step is to connect the corpus to visualize the data and observe the results of sentiment analysis research using Box Plot and Scatter Plot. In the Box Plot data visualization, a diagram is displayed showing the results of 6 emotions: surprise,joy, fear, sadness, disgust and anger. From these 6 emotions, it can be observed that joy, surprise, and fear are the dominant emotional responses shown by Twitter users with the keyword ‘kpk’. For further details, please refer to Figure 11 below.

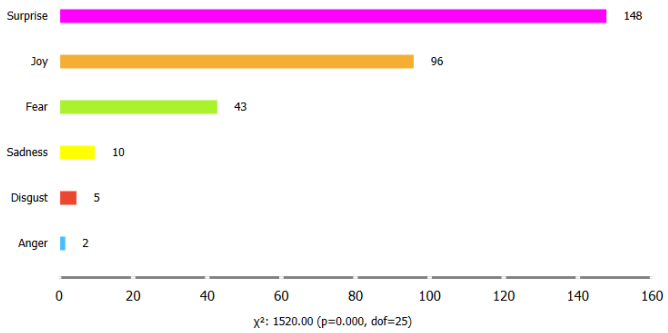


Figure 11. Box Plot Emotion

From the visualization above, it can be seen that the emotional responses exhibited by Twitter users with the keyword 'kpk' are as follows: the emotion 'surprise' totals 148 or approximately 44%, the emotion 'joy' totals 96 or approximately 35%, and the emotion 'fear' totals 43 or approximately 16%. Data visualization can also be observed through Scatter Plots to visualize patterns or relationships between two variables, such as positive correlation, negative correlation, or no correlation at all. In a scatter plot, each point represents one observation, where one axis indicates the value of one variable and the other axis indicates the value of another variable. In this study, the variables used are emotion and sentiment variables.

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

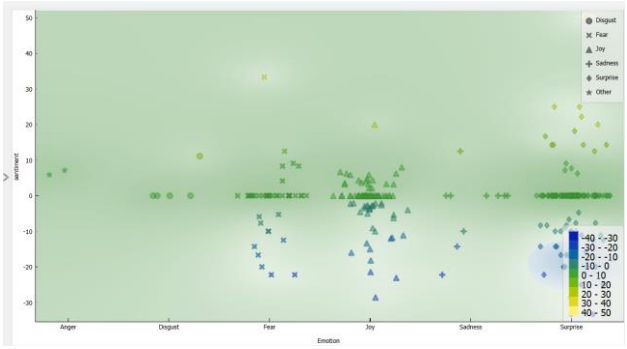


Figure 12. Scatter Plot

4. CONCLUSION

The conclusions obtained are Multilingual sentiment algorithm can be utilized for sentiment analysis on social media platforms such as Twitter, to comprehend users opinions and viewpoints across various languages. It has been proven to provide a significant level of invariance compared to traditional sentiment analysis systems, thus enhancing the accuracy and diversity of sentiment analysis. The emotion 'surprise' totals 148 or approximately 44%, the emotion 'joy' totals 96 or approximately 35%, and the emotion 'fear' totals 43 or approximately 16%. The analysis method utilizing tweet profiling enables the determination of the mood or emotions of Twitter users regarding trending topics in Indonesia, particularly concerning Komisi Pemberantasan Korupsi. By employing box plot and scatter plot visualizations, we can determine the classification of Twitter users with the visualization of emotions that have been input into each corpus within Orange Data Mining.

REFERENCES

[1] R. Nooraeni, A. Fikri Fadhilah I, H. Dwi, S. Fatimatul, S. Pertiwi, and Y. Ronaldias, "Analisis Sentimen Data Twitter Mengenai Isu RUU KPK Dengan Metode Support Vector Machine (SVM)," vol. 22, no. 1, 2020, doi: 10.31294/p.v21i2.

[2] A. Fathiarahma, A. Voutama, T. Ridwan, and N. Heryana, "Analisis Text Mining Klasifikasi Kegiatan Keluarga menggunakan Orange dengan Metode Naive Bayes," *J. Teknol. Terpadu*, vol. 9, no. 1, pp. 35–41, 2023, doi: 10.54914/jtt.v9i1.606.

[3] M. Saputra, S. Nurhaliza Sofyan, A. Aulia, A. Ernawati, A. Oftasari, and R. Farta wijaya, "Implementation of E-Commerce System as SME Development Strategy in the Digital Era," *Bull. Inf. Technol.*, vol. 5, no. 3, pp. 195–202, 2024, doi: 10.47065/bit.v5i2.1545.

[4] S. K. P. Barakbah, Ali Ridho, M. K. Karlita, Tita, S.Kom., and S. K. Ahsan, Ahmad Syauqi, *Logika dan Algoritma*, no. tahun 2013. Surabaya: PENS, 2012.

[5] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.

[6] P. Peralatan *et al.*, "J-SISKO TECH Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Penerapan Data Mining Untuk Menentukan Pola," ■, vol. 118, no. 1, pp. 118–136, 2020.

[7] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 4, pp. 411–423, 2022, doi: 10.55123/insologi.v1i4.770.

[8] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 219, 2021, doi: 10.30865/jurikom.v8i6.3655.

[9] T. Yuniarti, J. Astuti, F. Faujiyah, and M. Zaiyar, "Pendekatan Text Mining dalam Menilai Sentimen Publik pada Baterai Kendaraan Listrik," vol. IX, no. 4, pp. 10602–10612, 2024.

[10] F. Sulianta, "Buku Dasar Data Mining from A to Z," no. January, 2024.

[11] J. Han and M. Kamber, *Data Mining Concept and Technique*. San Francisco: Morgan Kauffman, 2006.

[12] F. SLN, "Buku Dasar Data Mining from A to Z," no. January, pp. 15–15, 2023.

[13] P. Bhargavi and S. Jyothi, "Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 9, no. 8, pp. 117–122, 2009.

[14] D. Sophia and L. Y. Banowosari, "Implementasi Metode Aturan Asosiasi Menggunakan Algoritma Apriori Pada Data Transaksi Penjualan Di Waroeng Spesial Sambal," *J. Inform. dan Komput.*, vol. 22, no. 1, pp. 44–56, 2017.

[15] R. A. raffaidy Wiguna and A. I. Rifai, "Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 5, No 4, Desember 2024, Hal. 287-294
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.1711
https://journal.fkpt.org/index.php/BIT

Menggunakan Orange Data Mining," *J. Inf. Syst. Informatics*, vol. 3, no. 1, pp. 1–12, 2021, doi: 10.33557/journalisi.v3i1.78.

[16] M. Muharrom, "Analisis Komparasi Algoritma Data Mining Naive Bayes, K-Nearest Neighbors dan Regresi Linier Dalam Prediksi Harga Emas," *Bull. Inf. Technol.*, vol. 4, no. 4, pp. 430–438, 2023, doi: 10.47065/bit.v4i4.986.

[17] E. Mardiani *et al.*, "Membandingkan Algoritma Data Mining Dengan Tools Orange untuk Social Economy," *Digit. Transform. Technol.*, vol. 3, no. 2, pp. 686–693, 2023, doi: 10.47709/digitech.v3i2.3256.

[18] Hafiz Aryan Siregar, Muhammad Zacky Raditya, Aditya Nugraha Yesa, and Inggih Permana, "Comparison of Classification Algorithm Performance for Diabetes Prediction Using Orange Data Mining," *Indones. J. Data Sci.*, vol. 4, no. 3, pp. 176–182, 2024, doi: 10.56705/ijodas.v4i3.103.