JOURNAL OF WORLD SCIENCE
https://jws.rivierapublishing.id/index.php/jws
Volume 2 No. 6 June 2023
P-ISSN: 2828-8726
E-ISSN: 2828-9307

CLASSIFICATION OF TRAFFIC LOCK NESS IN INDONESIA USING THE NAIVE BAYES CLASSIFICATION METHOD

Abdul Robi Padri[1], Asro[2], Indra[3]
Universitas Raharja, Banten, Indonesia

abdulrobipadri06@gmail.com[1], asro@raharja.info[2], indra@budiluhur.ac.id[3]

ABSTRACT
The purpose of this research is to analyze the accuracy of congestion data using Google Colab in detecting congestion by the province in Indonesia the author tries to test strategies for dealing with congestion in the Indonesian region by utilizing the Naïve Bayes method. In this journal, apply with Google Collab . This research uses data that comes from crawling data on Twitter. Using the Naive Bayes method to find the shortest route is efficient and not congested. Implementation of online school transportation using the naive Bayes method in minimizing travel costs to pick up students can reduce traffic jams, reduce accidents, reduce student tardiness, and minimize travel costs. The Naive Bayes method can be used to identify relevant information about traffic jams in Indonesia through Twitter data with a good degree of accuracy. These results can assist decision-making and strategic planning in overcoming the problem of traffic congestion in Indonesia. Therefore, this research implies that it can help improve the accuracy of traffic congestion data in Indonesia. By using Google Colab, more advanced analysis methods and machine learning algorithms can be applied to process the existing traffic data. Additionally, utilizing Google Colab allows for fast and efficient data processing.

Keywords: congestion classification, naïve bayes, traffic.

Corresponding Author: Abdul Robi Padri
Email: abdulrobipadri06@gmail.com

INTRODUCTION

Traffic jams often occur in various and The increase in population growth in Indonesia every year has shown that Indonesia is experiencing a population surge; this is also the background of congestion (Janah et al., 2016). The ratio of the number of private vehicles and public transportation has also increased so that it can cause traffic jams which take many victims of accidents in Indonesia (Pratiwi, 2016). With the presence of Transportation Technology innovations that make it easier for people to travel to other places, such as going to work and school, shopping as well as traveling, around 2011 Indonesia, a trend emerged for online public transportation, such as online taxis and online motorcycle taxis (Gojek, Ubermotor, Blu-jek) which in fact has helped the problems of some people throughout Indonesia when traveling (Hariansyah, 2018). However, traffic jams in Indonesian territory have not been optimally resolved. In recent years, Deadlock has frequently used naive Bayes Google Collab Algorithm, then analysis. To support the research, the author made a survey of 70 respondents about traffic in various traffic in Indonesia. Starting from the aim of implementing online school transportation or vehicles that the local government has provided is to reduce traffic jams, reduce accidents, reduce tardiness for both students and students, and minimize costs. We hope that the results of implementing public school transportation can provide the best solutions and make useful contributions to all West Java people in avoiding traffic jams and being more efficient, safe, and comfortable.

This is especially true in big cities due to the increasing number of vehicles every year and the number of inadequate roads (Rasyid et al., 2020). Problems are also increasingly complex with the increasing number of accidents that occur (Desmira et al., 2015). Transportation is the process of distributing goods to several places-the high demand for transportation results in a lack of performance efficiency in serving the community or customers (Ardini & Lutfiyana, 2018). So we need a Transportation Method to complete and optimize Transportation performance to save costs and distances (Hermanto et al., 2017). Based on the discussion of these conditions, the authors try to offer a strategy for dealing with traffic jams in West Java with the concept of Business Intelligence. This article offers a model for implementing Business Intelligence-based online school transportation using the Naive Bayes algorithm method. The dataset is a recap of West Java road traffic jam data daily on Google Maps 2017, accessed online; this dataset is tested for accuracy (Darmawan & Makruf, 2023).

Based on the background above, this research aims to identify and analyze developing applications using the Naive Bayes classification method. Applications that will be developed in the future will classify tweet data that contains traffic information in Indonesia. After the data is classified, then data visualization is carried out (Dewa et al., 2021). The benefit of this research is that it can help develop applications that can accurately classify data. In the context of traffic congestion in Indonesia, this application can help users to obtain relevant information about real-time traffic conditions. This information can be used to make informed decisions and help reduce traffic congestion problems in Indonesia (Habiba et al., 2023). In addition, application development using the Naive Bayes method can be applied to various fields, such as email spam detection and text classification on social media.

METHODS

Research Stage

The stages of this research were carried out in 8 stages, each stage from obtaining data to testing the Naive Bayes method and the linear model and RandomForest. The steps taken**.**

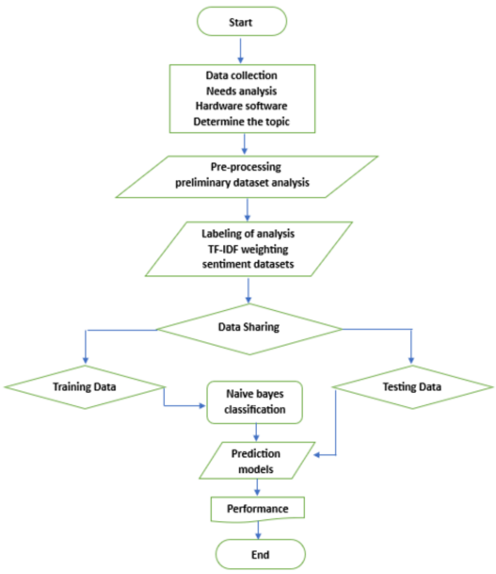


Figure 1. Trending Topic Analysis Flowchart

1. Data collection is a Twitter website, and by registering as a developer account, you will get an API where this API can be used for crawling Twitter data using orange software and by searching for certain topic keywords and producing the amount of data obtained is 1501 (Kasogi et al., 2020).
2. It is determining the Topics to be labeled in the analysis of the preliminary dataset, which is from the relevant and not relevant preliminary labels as well as the media or individuals that will be obtained from the manual labeling.
3. Preprocessing
   Preprocessing aims to change unstructured data into structured data because it will produce data that can be easily processed according to their needs (Harjanta, 2015). The first preprocessing process is deleting retweet data from tweets because retweets cannot explain the personality of the Twitter account user. They were then followed by deleting (URL, RT, HTTP, #, @, removing hashtags, removing character spaces from left and right text, and replacing newline into the space). Case folding is changing all the characters in the text to lowercase. Tokenization or string splitting, the text becomes a token list, then the tweet data is broken down into word units. Filtering is the selection of important words after the tokenizing process. Moreover, the final stage is steamer-reducing words and changing words affixed to basic words.
4. Polarity labeling, where the polarity score determine the polarity of Tweet Sentiment with the Indonesian Sentiment Lexicon Loading positive and negative and neutral data lexicons.
5. TF-IDF weighting
   In this weighting, each Twitter dataset that has been preprocessed will be given a weighting value using TF-IDF which will produce a value for each text dataset from the tweet (Wahyunita et al., 2020).
6. Naïve Bayes Algorithm Classification
   The data that has been preprocessed continues to enter the classification stage using the Naïve Bayes algorithm and RandomForest as a classifier.
7. Prediction Models
   At this stage, the data after all the preprocessing and processing with algotima using naïve Bayes RandomFores and using CountVectorizer only with trigrams or using TF-IDF vectors for text preprocessing. The prediction model will produce predictive results from the classification of the Algorithm and the data that has been tested, which has been output to an accurate accuracy value.
8. performance evaluation
   In this process, accuracy, precision, and Recall calculations will be carried out to measure the system that has been made.

Software and Hardware Requirements Analysis

At this stage, a software and hardware requirements analysis is needed to run a project researchers are developing using needed system design in this study (Hariansyah, 2018).

Table 1. Software Requirements

Software	Version
Operating system	Mac OS Monterey
Application 1	Orange3
Application 2	Google Colab

Table 2. Software Requirements

Software	Version
Processor	Intel(R) Core(TM) i5
Storage (RAM)	16GB DDR4
HDD/SSD	1TB

RESULTS AND DISCUSSION

The research results determine the fastest route, whether to continue or look for another way. Preliminary dataset analysis aims to obtain the results of the dataset with topics that have been used for research methods by conducting direct manual research of these topics and sample tables that have been adapted to the research topic dataset percentage of relevant tweets in the datasets.

Labeling Dataset

Based on the results of crawling data taken from Twitter, with a total of ***1501*** tweets on traffic jams, traffic jams, and traffic jams throughout Indonesia. Moreover, after preprocessing, the data will produce ***1182*** tweets. Moreover, where the data has resulted in polarity score labeling with the Indonesian lexicon dictionary, which gets an accuracy value by obtaining sentiment analysis labeling, Negative, Positive, and Neutral produce the analysis data automatically with the Indonesian dictionary. Moreover, do the preliminary plebeian manually from the relevant and the media or individuals (Febriyani & Februariyanti, 2023).

	Author	text_clean	text_preprocessed	Lis_Relevant	Lis_Media	polarity_score	polarity
0	@2o43i	aku soalnya jarak ke kampus jauh dan aku males...	['jarak', 'kampus', 'males', 'macetmacetan', '...	Relevant	Indivial	-6	negative
1	@rzimeldaa	legit setiap titik macet bikin mual pagi pagi ...	['legit', 'titik', 'macet', 'bikin', 'mual', '...	Relevant	Indivial	-9	negative
2	@girls_love_alot	asli nya macet ada kali deket situ banyak kang...	['asli', 'nya', 'macet', 'kali', 'deket', 'sit...	Relevant	Indivial	3	positive
3	@fuckcmpn	udah di jalan cuma macet	['udah', 'jalan', 'macet']	Relevant	Indivial	0	neutral
4	@devilscomerade	macet sattt	['macet', 'sattt']	Relevant	Media	0	neutral
...	...	...	...	...	...	...	...
1177	@mili_sekon	cara alternatif pembelah kemacetan ngekorin mo...	['alternatif', 'belah', 'macet', 'ngekorin', '...	Relevant	Indivial	-4	negative
1178	@tmusuh	udah mau ganti taun pikiran lu macet di hal ha...	['udah', 'ganti', 'taun', 'pikir', 'lu', 'mace...	Relevant	Indivial	-6	negative
1179	@cumantidursiang	udh macet aja	['udh', 'macet', 'aja']	Relevant	Indivial	1	positive
1180	@manusiamuscat	pagi pagi kabel depan puri beta sempet ngalin...	['pagi', 'pagi', 'kabel', 'puri', 'beta', 'sem...	Not Relevant	Indivial	-12	negative
1181	@bebyanisa27	lagi butuh uang tapi orderan lagi macet huhuhu	['butuh', 'uang', 'order', 'macet', 'huhuhu']	Not Relevant	Indivial	-7	negative

1182 rows × 7 columns

Figure 1. The dataset that has preprocessed the trending topic analysis flow

Table 3. Polarity Percentage of sentiment analysis Negative Neutral Positive

Type	Datasets	Negative	Neutral	Positive
Sentiment analysis	1182	683	182	317
Percentage Each		57.78%	15.40%	26.82%

Results The total number of sentiment analysis labels total **1182** /3, after which will result respectively (Negative 57.8% =683 Tweets, Neutral 26.82% = 317 Tweets, Positive 15.4% = 182 Tweets).

Visual Bar Chart and Donut Chart matplotib python

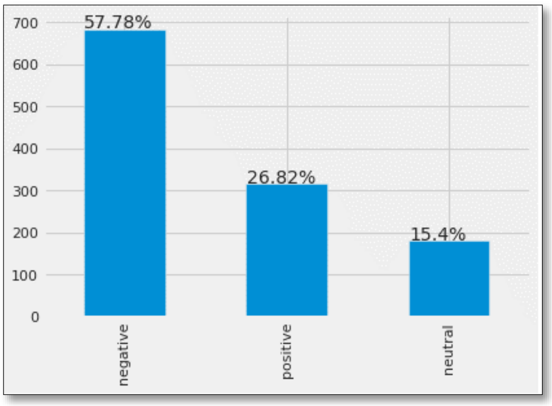


Figure 3. Generating Bar Visual Chart

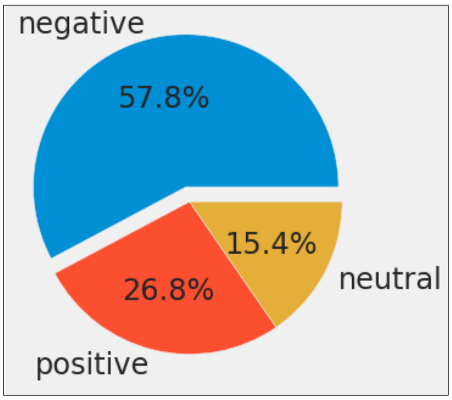


Figure 4. Visuals That Produce Pie and Donut Charts for Displaying Labeling Sentiment Analysis

Preliminary Datasets

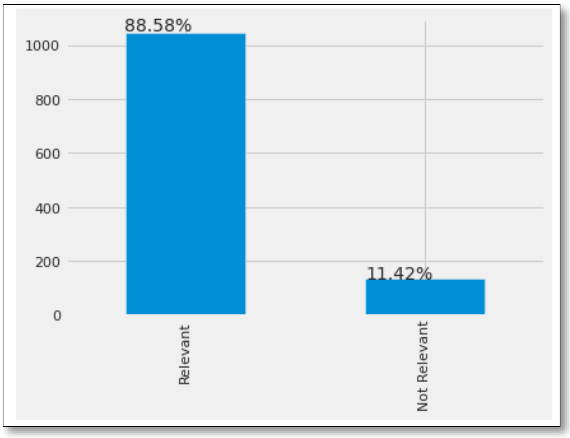
Table 4. Percentage of Relevant Tweets In Their Datasets

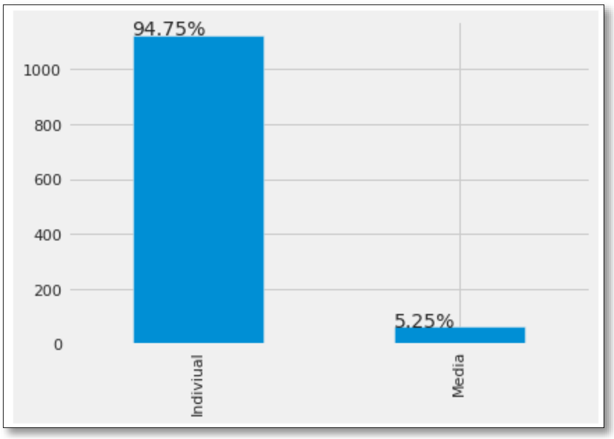
Type	Datasets	Relevant	Not Relevant	%Relevant
Relevant	1182	1047	135	88.58%

Table 5. The Proportion of Total Tweets from Media And Individual Accounts

Type	Datasets	Relevant	Not Relevant	%Relevant
Media	1182	62	1120	94.75%

Results The total amount of the preliminary dataset shown in Figure **5** is relevant data with total data (88.58% = **1047** tweets), while the Not relevant total data (11.42% = **135 Tweets)** and the results of the preliminary media dataset shown in Figure These **6** are the total data for Individuals (94.75% = 1120 Tweets), and the Media (5.25% = **62** Tweets).





Figures 5 and 6. Visuals That Produce A Bar Chart To Display The Preliminary Labeling Preliminary Analysis

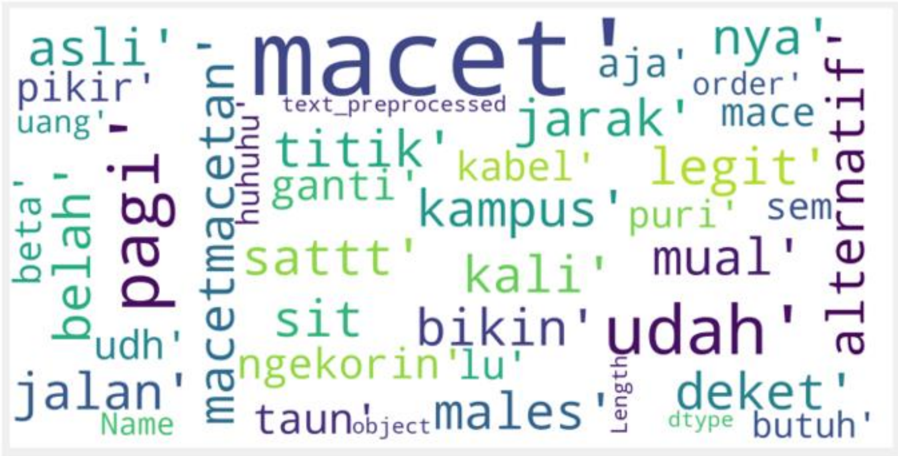


Figure 7. Visuals that generate WordCloud top topic stopwords Example

WordCloud with the topic 'stuck topic that has been successfully created! Moreover, the result is that the selected topic can be seen in the image with the largest font size (jammed), for example, if the topic I use is stuck or stuck. For example, in our case, it has a total data record equal to 1182 Tweets recorded with the mathematical calculation of the separation method (805 and 20%).

TF-DF bottling

Word weighting was done to assign accuracy values to certain data tweets using classic NLP classifying by topic text data, using gram bag techniques, n-grams, TF-IDF, etc., for text representation and applying different classification algorithms.

Calculation of training and testing data using TF-IDF Naive Bayes weighting

Every machine learning algorithm that is used requires a technique which is a division of some sort. In this process, the entire dataset will be divided [3] into two parts, namely Training and Testing. Each division will produce each percentage, which can be 80% or 20%. However, this method will certainly do the 80% and 20% division techniques that will be used to evaluate the performance of the machine learning algorithm that will be used. As said above, this processing method requires dividing a dataset into two subsets.

a. The training set is used to customize and train machine-learning models
b. A set of tests used to evaluate suitable machine learning models

This technique is an important step, and in most cases, in using the division algorithm, we always use 80% of the dataset given to the topic to learn from the training. The remaining 20% will be tested to see the accuracy of the classification, to check whether going well or not. If the suggested Machine Learning algorithm is not working properly, another classifier must be applied to the data to be tested; for example, the image below explains the procedure for splitting/splitting data in machine learning Tweets recorded with the mathematical calculation of the splitting method from (805 and 20%) for the three calcifiers, which are calculated as follows.

a. **Training sets:** 80% × 1182 = 80/100× 1182 = 0.8 × 1182 = 945 records
b. **Testing set:** 20% × 1182 = 20/100× 1182 = 0.2 × 1182 = 237 records

The procedure for dividing datasets in machine learning:

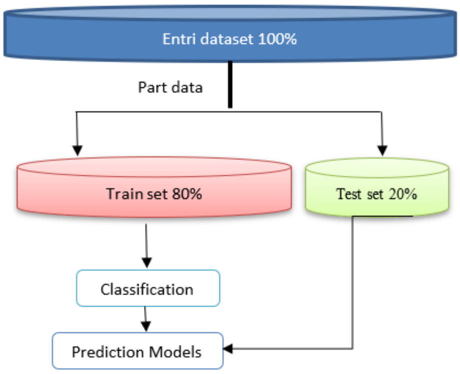
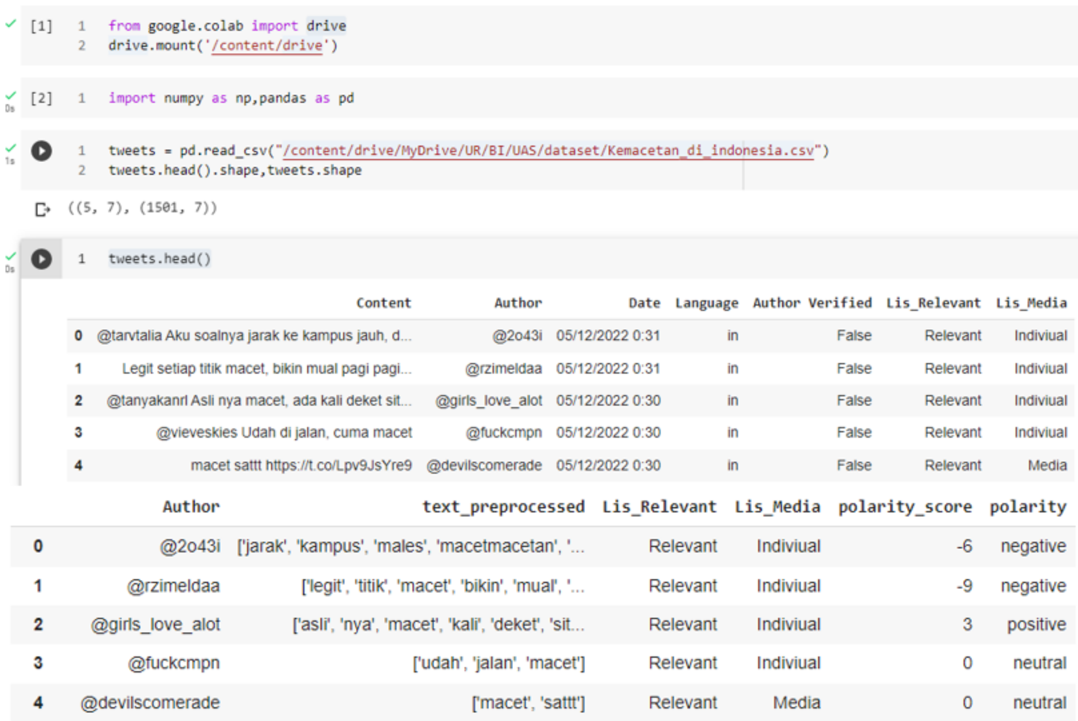


Figure 8. Entri Data Set

Classification Process 3.4

The classification process stage is sourced from Twitter data and crawled using orange3 software (Desmira et al., 2015). Testing this classification model is carried out on the developed system, and testing the accuracy of the classification model from the results of accurate data generated using the process of preprocessing and cleaning, which will produce data from 1501 will produce 1182 tweets.



```
[1]
from google.colab import drive
drive.mount('/content/drive')

[2]
import numpy as np,pandas as pd

tweets = pd.read_csv("/content/drive/MyDrive/UR/BI/UAS/dataset/Kemacetan_di_indonesia.csv")
tweets.head().shape,tweets.shape

((5, 7), (1501, 7))

tweets.head()
```

	Content	Author	Date	Language	Author Verified	Lis_Relevant	Lis_Media
0	@tarvtalia Aku soalnya jarak ke kampus jauh, d...	@2o43i	05/12/2022 0:31	in	False	Relevant	Indiviual
1	Legit setiap titik macet, bikin mual pagi pagi...	@rzimeldaa	05/12/2022 0:31	in	False	Relevant	Indiviual
2	@tanyakanrl Asli nya macet, ada kali deket sit...	@girls_love_alot	05/12/2022 0:30	in	False	Relevant	Indiviual
3	@vieveskies Udah di jalan, cuma macet	@fuckcmpn	05/12/2022 0:30	in	False	Relevant	Indiviual
4	macet sattt https://t.co/Lpv9JsYre9	@devilscomerade	05/12/2022 0:30	in	False	Relevant	Media

	Author	text_preprocessed	Lis_Relevant	Lis_Media	polarity_score	polarity
0	@2o43i	['jarak', 'kampus', 'males', 'macetmacetan', '...	Relevant	Indiviual	-6	negative
1	@rzimeldaa	['legit', 'titik', 'macet', 'bikin', 'mual', '...	Relevant	Indiviual	-9	negative
2	@girls_love_alot	['asli', 'nya', 'macet', 'kali', 'deket', 'sit...	Relevant	Indiviual	3	positive
3	@fuckcmpn	['udah', 'jalan', 'macet']	Relevant	Indiviual	0	neutral
4	@devilscomerade	['macet', 'sattt']	Relevant	Media	0	neutral

Figure 9. Image Starting Processing Data Cleaning and Testing

Classification is determining a record or Classification with naïve Bayes classifier google collab python, which will show the classification process with google collab and with the naïve Bayes method where experiments or tests using the sklearn pipeline module create a classification pipeline to generate data by one of them using Count Vectorizer with unigrams and bigrams and using Multinomial Naïve Bayes, classifier prints a classification report As for the random forest for testing the accuracy of the values that will be used to produce the classification of this research data so that accuracy is obtained in the rankings that produce data that has divided the Twitter data with (80% Trainset and 20% Test set) and Testing data is what will be tested for accuracy in this software (Indriyani, 2019).

Model prediction and evaluation of metrics 3.5

This metric measurement summed up the results as follows:

a. accuracy
b. Precision
c. Recall
d. F1 scores

Accuracy was calculated as the total number of correct predictions over the total number of data sets (i.e., all correct/all). Accuracy rule is

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

An example Of precision is,

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)}$$

An example For Recall is,

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)}$$

An example Of an F1 score is,

$$F1\ score = 2\ \times \frac{Precision\ \times Recall}{Precision + Recall}$$

The description of the understanding of the accuracy, precision, Recall, and f1 score model metrics is as follows; for example, a Comparison of a classification system that has been trained to classify will give results in predicting metrics that the total data results are 237 tweets that are to be tested for get Negative, Neutral, Positive results.

Table confusion matrix Manually Using Calculations with Microsoft Excel so that the calculations are more detailed and also make verification reference materials how accurate with Python and manual you can see the confusion matrix image below and the following example.

Total Test	237 Multinomial Naïve Bayes		
JENIS	**Negative**	**Neutral**	**Positive**
Negative	131	1	5
Neutral	33	1	2
Positive	44	2	18

Figure 10. Calculation Manual To Make Using Method Excels

JENIS	precission	recall	fi-score	support
Negative	63%	96%	76%	137
Neutral	25%	3%	5%	36
Positive	72%	28%	40%	64
Accuracy	63.2911392405063%			237
Macro avg	53%	42%	40%	237
weighted avg	60%	63.5%	56%	237

Figure 11. Results From the Naïve Bayes Test Accuracy

Examples of results from the confusion matrix include precision, sensitivity (Recall), and F1 scores (F1-score). We can use manual formulas to calculate these values using software such as Microsoft Excel. Here is an explanation.

$$Precission = \frac{TP}{TP + FP} = \frac{131}{131 + 77}$$

$$Recall = \frac{TP + FN}{TP} = \frac{131}{131 + 6}$$

$$F1\ score = 2 \times \frac{Precision\ \times Recall}{Precission + Recall} = 2 \times \frac{63\%\ \times 96\%}{63\% + 96\%}$$

Accuracy Determination Stage from Confusion Matrix and determine the results

Accuracy	63.2911392405063%

= (131+1+18)/ (237)

The average value of the results of the confusion matrix calculation that produces macro averages and weighted averages can be explained in the following example:

	precission	recall	fi-score
Macro avg	53%	42%	40%
weighted avg	60%	63.5%	56%

Figure 12. Total Result Of Confusion Matrix

$$Macro - Precision = \frac{Percision\ 1 + Percision\ 2}{2} = \frac{63\% + 25\% + 72\%}{3}$$

$$Macro - Recall = \frac{Recal\ 1 + Recall\ 2}{2} = \frac{96\% + 3\% + \ \ 28\%}{3}$$

$$Macro - F - Score = 2\ \frac{Macro - Precision. Macro - Recall}{Macro - Precision + Macro - Recall} = \frac{76\% + 5\% + 40\%}{3}$$

$$Weighted\ Precision = \sum_{i=1\ ton} wi\ \times Precision\ i = \frac{63\%\ \times 137 + 25\%\ \times 36 + 72\%\ \times 64}{237}$$

$$Weighted\ Recall = \sum_{i=1\ ton} wi\ \times Recall\ i = \frac{96\%\ \times 137 + 3\%\ \times 36 + 28\%\ \times 64}{237}$$

$$Weighted\ F1\ Score = \sum_{i=1\ ton} wi\ \times F1\ Score\ i = \frac{76\%\ \times 137 + 5{,}6\%\ \times 36 + 40\%\ \times 64}{237}$$

The explanation regarding the calculation formula presented is a formula that has been tested to assess its accuracy. This formula includes an example of calculating the results of each component, such as the total of negative, neutral, and positive values, which will result in accuracy by dividing the calculation using an Excel worksheet. In addition, this formula will also produce other values, such as macro values, averages, and weight values calculated from the total of the entire calculation table.

One example of a formula that produces an accuracy level using the Naive Bayes method is based on the results of testing 237 tweet data. The tweet data is divided into positive, neutral, and negative categories. For example, the total test is divided into 131, 1, 5, 33, 1, 2, 44, 2, and 18. From these numbers, the accuracy value can be calculated as follows: (131+1+18) / (237).

An example of a confusion matrix image that has been tested with Google Colab and the example is as follows:

```
[[131   1   5]
 [ 33   1   2]
 [ 44   2  18]]
              precision    recall  f1-score   support

    negative       0.63      0.96      0.76       137
     neutral       0.25      0.03      0.05        36
    positive       0.72      0.28      0.40        64

    accuracy                           0.63       237
   macro avg       0.53      0.42      0.40       237
weighted avg       0.60      0.63      0.56       237

' accuracy 0.6329113924050633 % '
```

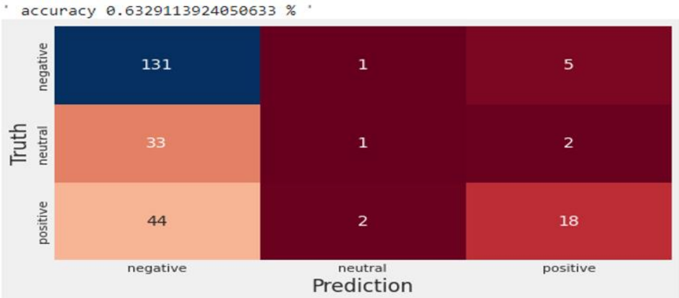


Figure 14. Confusion Matrix From Data Testing

This image shows all the stages tested on data using the manual method using an accurate Excel sheet formula. This includes preliminary, relevant, and irrelevant data and testing involving around 20% of the data to test accuracy through Naïve Bayes. This test uses software such as the Python notebook, Conda Python, or Orange, with particular emphasis on using Python Colab from A to Z.

From the initially mixed data, the cleaning text process in Python has removed irrelevant symbols to make the data more readable and relevant. After this process, it was decided to test the data by dividing it into 100%/20%, resulting in 237 data that would be used to process and classify accuracy.

Several different methods are applied to Python, covering all aspects of both manual and automated tests and the results of tests performed through Python.

CONCLUSION

This study shows that the Naive Bayes method can be used to analyze traffic jams in Indonesia by utilizing Twitter data. Through preprocessing, sentiment analysis, and classification, researchers can identify relevant information about traffic jams and classify them accurately. These results can be considered for decision-making and strategic planning in overcoming the problem of traffic.

REFERENCES

Ardini, A., & Lutfiyana, N. (2018). Metode Transportasi Untuk Mengoptimalkan Biaya Pengiriman Barang Pada PT Trimuda Nuansa Citra Jakarta. *Information System For Educators And Professionals: Journal Of Information System*, *3*(1), 55–66.

Darmawan, A., & Makruf, M. (2023). Deteksi Gaya Belajar Siswa SMA pada Virtual Based Learning Environment (VBLE) dengan Decision Tree C4. 5 dan Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, *3*(5), 532–544. https://doi.org/10.30865/klik.v3i5.760

Desmira, D., Kautsar, A., & Darmawan, A. W. (2015). Prototype Perancangan Informasi Kemacetan Jalan Tol Berbasis Mikrokontroler At89s52 Dengan Tampilan LCD. *PROSISKO: Jurnal Pengembangan Riset Dan Observasi Sistem Komputer*, *2*(2).

Dewa, W. A., Maknunah, J., & Putri, A. D. (2021). Penerapan Metode Naïve Bayes untuk Menentukan Pengajuan Polis Baru pada PT.“XYZ.” *Jurnal Ilmiah Komputasi*, *20*(1), 83–92. https://doi.org/10.32409/jikstik.20.1.2696

Febriyani, E., & Februariyanti, H. (2023). Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter. *Jurnal Tekno Kompak*, *17*(1), 25–38. https://doi.org/10.33365/jtk.v17i1.2061

Habiba, A., Isnanto, R. R., & Suseno, J. E. (2023). The Effect of Chi Square Feature Selection on the Naïve Bayes Algorithm on the Analysis of Indonesian Society’s Sentiment About Face-to-Face Learning During the Covid-19 Pandemic. *JST (Jurnal Sains Dan Teknologi)*, *12*(1).

Hariansyah, M. (2018). *Millenials “Bukan Generasi Micin.”* Guepedia.

Harjanta, A. T. J. (2015). Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining. *Jurnal Informatika Upgris*, *1*(1 Juni). https://doi.org/10.26877/jiu.v1i1%20Juni.804

Hermanto, N., Hermaliani, E. H., & Sutinah, E. (2017). Vogell’s Aproximation Method dalam Optimalisasi Biaya Transportasi Pengiriman Koran pada PT. Arah Medialog Pembangunan. *Jurnal Teknik Komputer AMIK BSI*, *3*(1), 30–36. https://doi.org/10.31294/jtk.v3i1.1340

Indriyani, L. (2019). Analisis penerapan Naïve Bayes untuk memprediksi resiko kredit anggota koperasi keluarga guru. *Jurnal Informatika*, *6*(2), 262–270.

https://doi.org/10.31294/ji.v6i2.5724

Janah, S. H., Nur, S., Emil Adly, S. T., & SH, L. (2016). Model Kebijakan Antisipatif Mengatasi Kemacetan Lalu Lintas Darat di Kota Batam. *Prosiding Seminar Nasional INDOCOMPAC*.

Kasogi, I., Setiawan, E., & Syauqy, D. (2020). Pengoptimalan Lampu Lalu Lintas menggunakan Metode Naïve Bayes Classifier. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *4*(6), 1725–1731.

Pratiwi, R. H. (2016). *Dampak Kemacetan Terhadap Kondisi Sosial Dan Ekonomi Pengguna Jalan Di Jakarta Utara (Studi Kasus: Pegawai Kantor Kecamatan Cilincing dan Pegawai Rumah Sakit Umum Kecamatan (RSUK) Cilincing Jakarta Utara)*. Fakultas Ekonomi dan Bisnis Unpas Bandung.

Rasyid, A. D. A., Aulia, R., & Fathurrachman, M. R. (2020). Penerapan Aplikasi Online pada Sistem Transportasi Umum Massal untuk Meningkatkan Minat Masyarakat dalam Upaya Mengurangi Kemacetan. *Sainteks*, *15*(2). DOI: 10.30595/sainteks.v15i2.6308

Wahyunita, S., Azhar, Y., & Hayatin, N. (2020). Analisa Sentimen Tweet Berbahasa Indonesia dengan Menggunakan Metode Pembobotan Hybrid TF-IDF pada Topik Transportasi Online. *Jurnal Repositor*, *2*(2), 185–192. https://doi.org/10.22219/repositor.v2i2.238