

### 3. RESULTS AND DISCUSSION

This research shows that the naive bayes model can be quite effectively used for sentiment analysis on Twitter, especially in the context of the pre-employment card program.

#### 3.1 Comment Sentiment Result

- The analysis shows the results of positive, negative, and neutral sentiments towards the pre-employment card program with 836 tweets.
- The results of the analysis show that the majority of sentiment towards this pre-employment card program tends to be neutral, with the number of neutral tweets 800, positive tweets 17, and negative tweets 22 there are many people still confused about how to register for pre-employment cards.

#### 3.2 Classification of methods

- The performance of the Naive Bayes method is carried out to measure the performance of the model in classifying sentiment. The naive bayes method gets 95% comment prediction percentage accuracy. considered good enough for results in the google collab application.
- The graph shows the data of the research results made in can be seen in Figure 5.

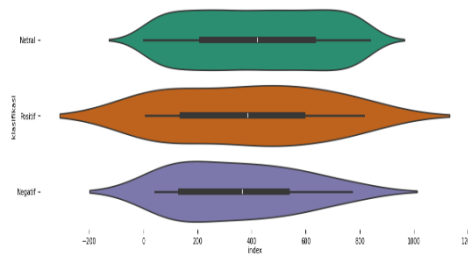


Figure 1. sentiment analysis chart

#### 3.3 Comparison with other methods

- SVM
- CNN
- RNN
- DATA MINING (CRISP-DM)

Comparison of accuracy in each method has different accuracy. the results can be seen in table

9.

Table 1. shows the accuracy comparison

NO	METHODS	ACCURACY
1.	Naive Bayes	95%
2.	SVM	98,34%.
3.	CNN	75,3%
4.	RNN	95,66%
5.	Data Mining	95.67%

The results of this research can serve as a basis for policy-making or program improvement based on the opinions and sentiments revealed through this analysis.

### Discussion

This research was conducted to analyze the sentiment of comments on Twitter about the Pre-Employment Card Program using the Naive Bayes method. The entire research process was implemented in Google Colab, which provides convenience and efficiency in data processing and model training. The following is a detailed discussion of the research results:

#### 1. Use of Google Colab in Research

Google Colab provides a platform that supports this research with sufficient computing resources, including GPUs and cloud-based storage. This allows researchers to access and process data at scale, and run machine learning algorithms without being limited by local capacity. In addition, collaboration features and integration with Google Drive make it easy to manage files, code and research results.

## 2. Effectiveness of Naive Bayes Method

The Naive Bayes method was chosen due to its simplicity and ability to handle text data with good performance. In this study, the Naive Bayes model was able to classify 800 out of 836 tweets correctly, resulting in an accuracy of about 95.69%. This accuracy rate shows that the model is quite reliable in identifying sentiment in the data used.

## 3. Optimal Data Pre-processing

One of the key success factors of this model is the comprehensive pre-processing of the data. The tweet text is cleaned from noise such as punctuation, URLs, and special characters. This step is followed by tokenization, removal of stop words, and stemming, which helps in reducing data complexity and improving model accuracy. The implementation of these steps in Google Colab also proved to be efficient, allowing the processing of thousands of tweets in a relatively short time.

## 4. Model Performance Evaluation

The evaluation of the model shows that Naive Bayes provides satisfactory results for sentiment classification with 95% accuracy. This result confirms that the model is able to capture the sentiment patterns present in the data well, although there are some tweets that are misclassified. These errors may be due to ambiguities in the text or context that are difficult to capture by the model.

## 5. Visualization and Interpretation of Results

Visualizations of the results, such as the confusion matrix and word cloud, are helpful in understanding how the model works and where it goes wrong. These visualizations show the distribution of prediction errors and provide insight into the dominant key words in each sentiment category. Using Google Colab, these visualizations can be generated quickly, providing immediate feedback on model performance.

## 6. Limitations and Development Potential

Although this research has achieved satisfactory results, there are some limitations:

- **Simple Model:** Naive Bayes is a simple model and may not be able to capture complex relationships in the data. The use of more sophisticated models, such as Random Forest or SVM, can improve accuracy.
- **Text Representation:** The use of CountVectorizer results in a simple text representation. deep learning based can be used to improve the performance of the model.
- **Data Size:** The number of tweets analyzed was limited to 836 tweets. Further research with a larger amount of data may provide more representative results.

## 7. Conclusion

This research shows that the Naive Bayes method implemented in Google Colab is effective for sentiment analysis on Twitter regarding the Pre-Employment Card Program. With high accuracy, this model can be a useful tool for measuring public perception of government policies. However, for application on a wider scale, it is recommended to explore more sophisticated methods and approaches to make the results more accurate and reliable.