ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

Rita Susanti, Alvito Aryo Pangestu, Haydar Arsy Firdaus and M. Fariz Fadillah Mardianto, S.Si., M.Si.

RITA SUSANTI

Universitas Airlangga

4 - 5 AUGUST 2021

CONTENTS

Introduction

One of the goals in the SDGs, which is to ensure a healthy life and promote the welfare for all people of all ages, has become difficult to maintain since the emergence of Covid-19 in Indonesia.

Introduction

NEWS • NATIONAL

Decision to punish anti-vaxxers under regional administrations: Task force

Health workers inject a civil servant with a COVID-19 vaccine during a vaccine drive simulation at the Udayana Military Command in Denpasar, Bali, on Dec. 10. (Antara/Nyoman Hendra Wibowo)

THE CONVERSATION

Disiplin ilmiah, gaya jurnalistik

COVID-19 Bisnis + Ekonomi Kesehatan Kota Pendidikan + Anak Muda Politik + Masyarakat Sains + Teknologi Budaya Lingkungan

Indonesia's decision to prioritise COVID-19 vaccination to citizens aged 18-59 years old questionable

Februari 2, 2021 7.43am WIB

Health workers are preparing COVID-19 vaccine Sinovac during first stage vaccination in Health Center, South Tagerang City, Indonesia, Januari 15, 2021. More than 8.000 health workers there are vaccinated. ANTARA FOTO/Fauzen/foc

NEWS • NATIONAL

Indonesia allows emergency use of Sinovac vaccine

Workers unload containers carrying the Sinovac vaccine for COVID-19 under National Police security at Soekarno-Hatta International Airport in Tangerang, Banten, on Dec. 31, 2020. As many as 1.8 million doses of Chinese firm Sinovac Biotech's COVID-19 vaccine, CoronaVac, arrived in Indonesia as the government eyes to roll out vaccination for medical workers in January. (Antara/Muhammad Iqbal)

Thus, the Indonesian government has issued a policy regarding the procurement of vaccines and the implementation of vaccinations through Presidential Regulation Number 99 of 2020

Introduction

NEWS · NATIONAL

Decision to punish anti-vaxxers under regional administrations: Task force
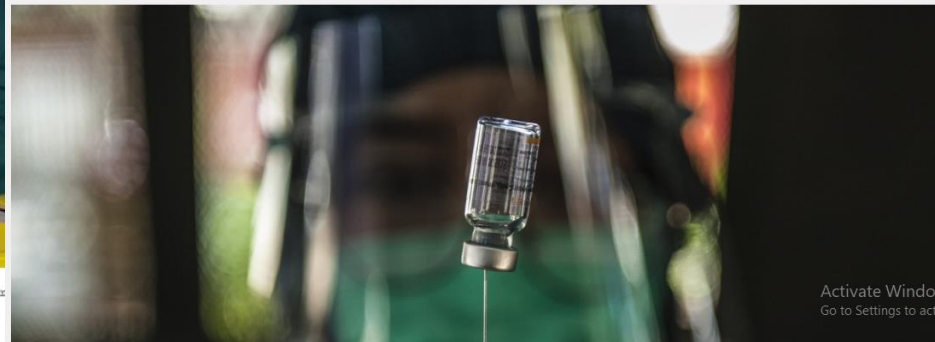
Health workers inject a civil servant with a COVID-19 vaccine during a vaccine drive simulation at the Udayana Military Command in Denpasar, Bali, on Dec. 10. (Antara/Nyoman Hendra Wibowo)

HEALTH

Indonesia's anti-vaxxers: Between religion and lack of information

ADI RENALDI
THE JAKARTA POST

PREMIUM Jakarta / Thu, February 4, 2021 / 05:43 pm

Activate Windows
Go to Settings to activa

There are different perceptions in the community regarding the Covid-19 vaccination policy which is using provocative narratives and hashtags on social media.

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

5

Introduction

The Naïve Bayes Classifier approach is used to classify public sentiment regarding the Covid-19 vaccination policy in Indonesia.

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

6

Introduction

Wongkar and Angdresey (2019) has classified the public sentiment on Indonesia's 2019 presidential candidates. This study compared the Naïve Bayes Classifier, SVM and KNN which concluded that Naïve Bayes Classifier has the highest accuracy value of 75.58%.

Introduction

The novelty of this research is use text mining with the Naïve Bayes Classifier approach to determine public sentiment towards the Covid-19 vaccination policy.
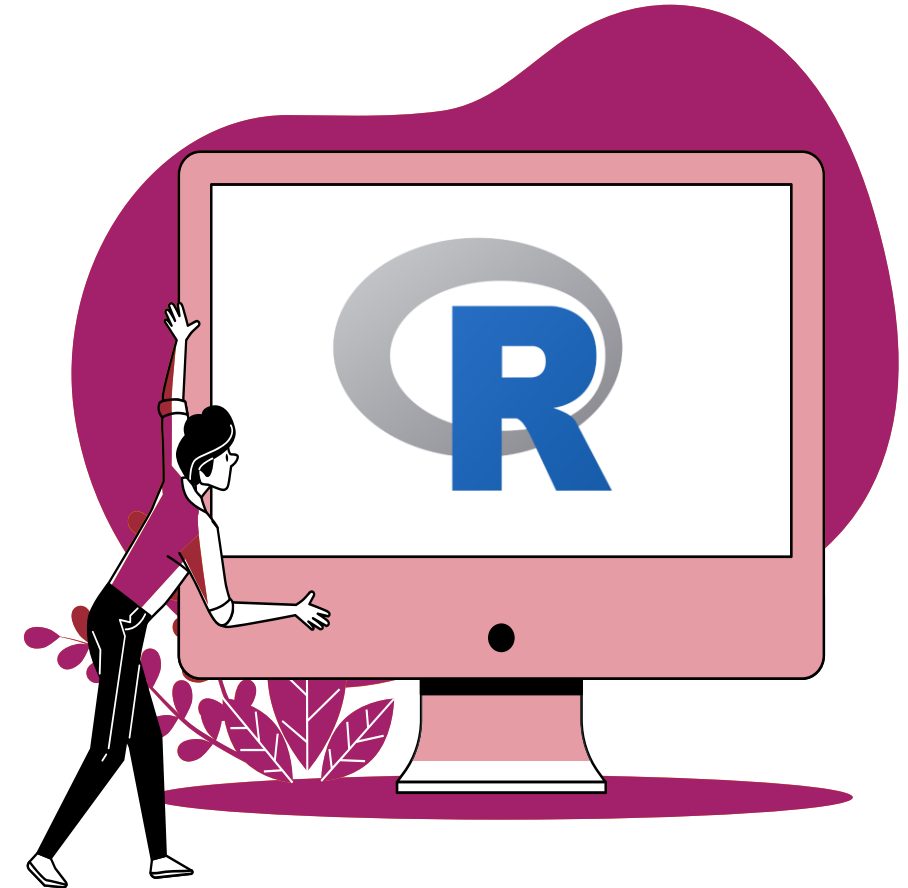
The results of this study are expected to be used as a reference for the government in convincing the public regarding the Covid-19 vaccination process towards a healthy Indonesia.

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

8

Method: Data

Data Type	Sentiment		Total
	Positive	Negative	
Training	198	202	400
Testing	52	48	100
Total	250	250	500

The research data is in the form of public tweets regarding the Covid-19 vaccination policy by the government taken from Twitter from August 4, 2020 - February 2, 2021, totaling 500 tweets.

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

Method: Analysis Procedure

01 Describe an overview of public sentiment

02 Preprocess the data

03 Create a bar chart to show frequently occurring words

04 Classify public comments based on sentiment categories using the Naïve Bayes Classifier

05 Calculate the sensitivity, specificity, Apparent Error Rate (APER), accuracy and Press' Q values

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

Results: Sentiment Classification

There are 1540 words produced after the preprocessing stage. All words in the training data will be used to create a Naïve Bayes Classifier model.

Training Data Classification

The sensitivity of 87.23% indicates that the model is very good for predicting positive sentiments, while the specificity of 100% indicates that the model is very good for predicting negative sentiments.

Prediction	Actual		Total
	Positive	Negative	
Positive	198	29	227
Negative	0	173	173
Total	198	202	400

Results: Sentiment Classification

Training Data Classification

Then, the accuracy percentage of 92.75% indicates that the probability value generated by the Naïve Bayes Classifier method on training data is very good for classifying the public sentiment.

The Press' Q value obtained is 292.41 which is greater than $\chi^2_{0.05(1)} = 3.841$. Thus, it can be concluded that the results of the classification on the training data are stable or statistically consistent.

Prediction	Actual		Total
	Positive	Negative	
Positive	198	29	227
Negative	0	173	173
Total	198	202	400

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

12

Results: Sentiment Classification

Testing Data Classification

The sensitivity of 70.59% indicates that the model is good enough to predict positive sentiment, while the specificity of 87.5% indicates that the model is very good for predicting negative sentiment.

Then, the accuracy percentage of 76% indicates that the probability value generated by the Naïve Bayes Classifier method on testing data is good enough to be used for classifying the public sentiment.

Prediction	Actual		Total
	Positive	Negative	
Positive	48	20	68
Negative	4	28	42
Total	52	48	100

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

Results: Sentiment Classification

Testing Data Classification

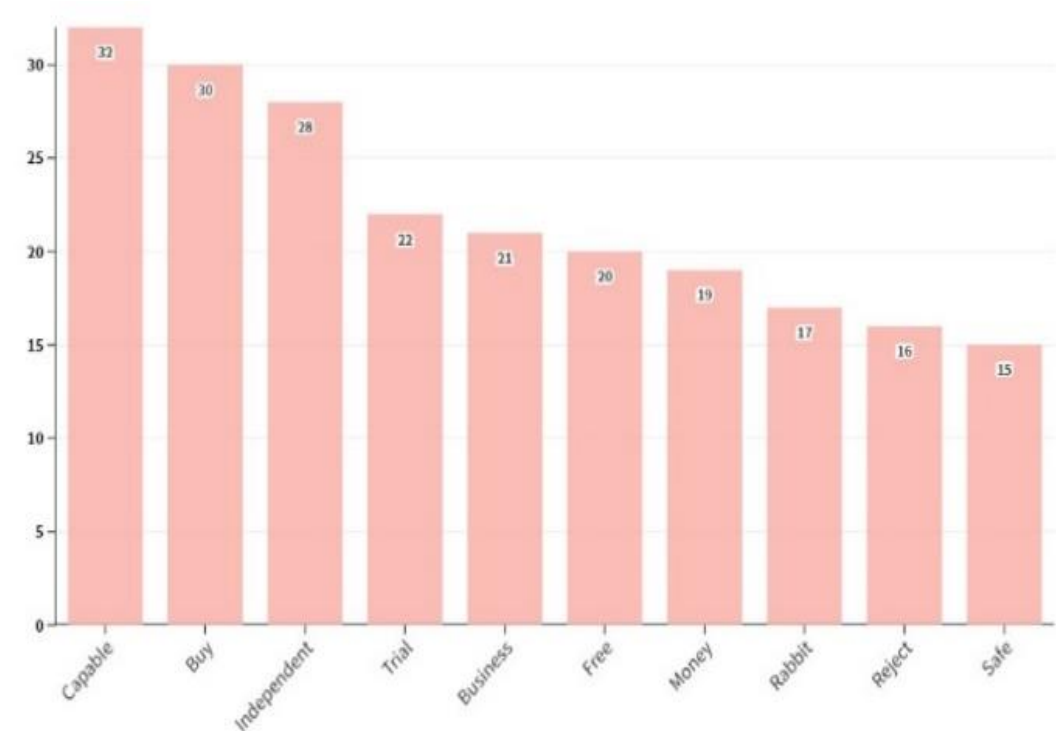The Press' Q value obtained is 27.04 which is greater than $\chi^2_{0.05(1)} = 3.841$. Thus, it can be concluded that the results of the classification on the testing data are stable or statistically consistent.

Prediction	Actual		Total
	Positive	**Negative**	
Positive	48	20	68
Negative	4	28	42
Total	52	48	100

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

14

Results: Sentiment Classification

Based on this figure, the word with the highest frequency of negative sentiment is the word "capable". Some people still doubt the ability of the Sinovac vaccine in dealing with Covid-19 because the efficacy of this vaccine in Indonesia is only 65.3% (Syakriah, 2021a).

Word that appear on negative sentiment

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

Results: Sentiment Classification

In addition, the words "buy" and "independent" rank second and third in negative sentiment. This was triggered by the government's plan to allow independent vaccination (Handayani and Kurniawan, 2021). Apart from the independent vaccination policy, the government is also trying to work with the private sector to cut costs (Syakriah, 2021b).

Word that appear on negative sentiment
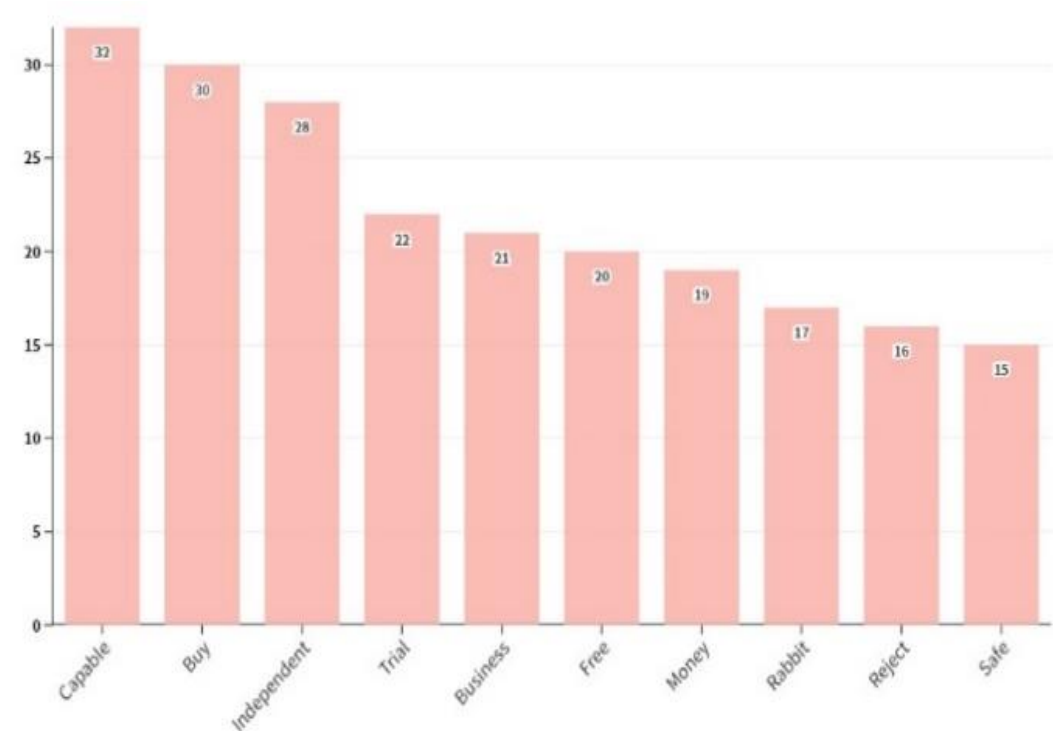
Results: Sentiment Classification
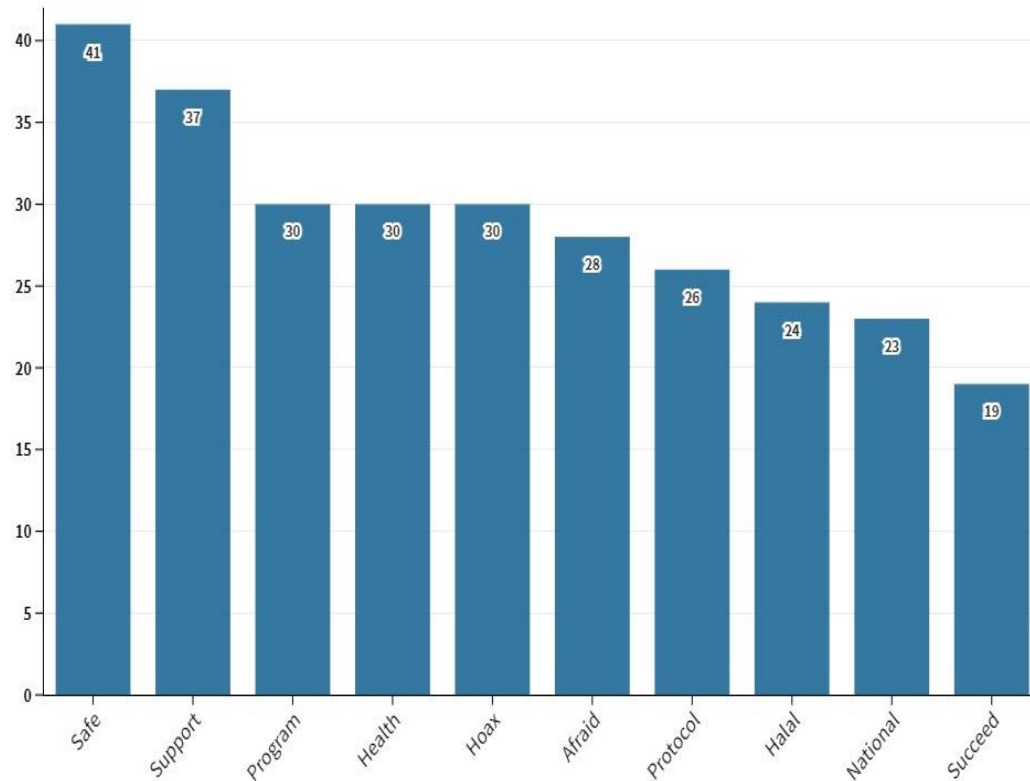
However, the word "safe" is the word that has the highest frequency of positive sentiment. This means that some people still believe that the Sinovac vaccine is safe to use.

Word that appear on positive sentiment

Results: Recommendation

The government and society should be able to be open to each other in listening to and giving criticism regarding the Covid-19 vaccination policy to accelerate the end of the Covid-19 pandemic towards a healthy Indonesia which is in line with the SDGs achievement targets in the health sector

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

18

Conclusion

The results of the classification of training and testing data were stable or statistically consistent

The model is good enough to be used for the classification of positive sentiment for training data and negative sentiment for testing data.

Overall, the Naïve Bayes Classifier method is a method that is quite appropriate to use to classify public sentiment regarding the Covid-19 vaccination policy because the accuracy value in training and testing data is quite high.

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

References

Bach, M. P., Krstic, Z., Seljan, S. and Turulja, L. (2019). Text Mining for Big Data Analysis in Financial Sector: A Literature Review. *Sustainability, 11*(1277):1 – 27.

Djaafara, B., Andiwijaya, F. R., Verisqa, F., Fadilah, I., Saraswati, K. and Maulida, R. (2021). Indonesia's Decision to Prioritise Covid-19 Vaccination to Citizens Aged 18 – 59 Years Old Questionable. Retrieved from https://theconversation.com/indonesias-decision-to-prioritise-covid-19-vaccination-to-citizens-aged-18-59-years-old-questionable-153883.

Feng, X., Li, S., Yuan, C., Zeng, P. and Sun, Y. (2018). Prediction of Slope Stability Using Naïve Bayes Classifier. *KSCE Journal of Civil Engineering, 22*(3): 941 – 950.

Gomes, S. R., Saroar, S. G., Telot, M. M. A., Khan, B. N., Chakrabarty, A. and Mostakim, M. M. (2017). A Comparative Approach to Email Classification Using Naïve Bayes Classifier and Hidden Markov Model. *Proceeding of the 2017 4th International Conference on Advance in Electrical Engineering* (ICAEE), 482 – 487.

Handayani, M. and Kurniawan, D. (2021, January 25). Vaccination Remains Free, Companies Must Buy and Are Not Allowed to Cut Salaries. Retrieved from https://voi.id/en/news/28736/coordinating-minister-airlangga-independent-vaccination-remains-free-companies-must-buy-and-are-not-allowed-to-cut-salaries.

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

20

References

Hasyim, M., Rahayu, D. S., Muliawati, N. E., Hayuhantika, D., Puspasari, R., Anggreini, D., Hastari, R. C., Hartanto, S. and Utomo, F. H. (2018). Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) to Analyse the Lecturer Research Performance in Private University. *Journal of Physics: Conf. Series, 1114*(1): 012117.

Navas, C. M., Bermejo, J. V. D., McLean, A. K., Jurado, J. M. L., Torres, A. B. B. R. D. and Gonzales, F. J. V. (2021). Discriminant Canonical Analysis of the Contribution of Spanish and Arabian Purebred Horses to the Genetic Diversity and Population Structure of Hispano-Arabian Horses. *Animals, 11*(269): 1 – 27.

Ranjan, N. M. and Prasad, R. S. (2017). Automatic Text Classification Using BPLion-neural Network and Semantic Word Processing. *The Imaging Science Journal,* 66(2): 69 – 83.

Sun, P., Lu, X., Xu, C., Sun, W. and Pan, B. (2020). Understanding of Covid-19 Based on Current Evidence. *Journal of Medical Virology, 92*(6):548 – 551.

Syakriah, A. (2021a, January 11). Indonesia Allows Emergency Use of Sinovac Vaccine. Retrieved from https://www.thejakartapost.com/news/2021/01/11/indonesia-allows-emergency-use-of-sinovac-vaccine.html.

Syakriah, A. (2021b, January 9). Indonesia on Track for Jan. 13 Vaccination Drive. Retrieved from https://www.thejakartapost.com/news/2021/01/09/indonesia-on-track-for-jan-13-vaccination-drive.html.

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

21

References

The Jakarta Post. (2020, December 25). Decision to Punish Anti-Vaxxers Under Regional Administrations: Task Force. Retrieved from https://www.thejakartapost.com/news/2020/12/25/decision-to-punish-anti-vaxxers-under-regional-administrations-task-force.html.

Wongkar, M. and Angdresey, A. (2019). Sentiment Analysis Using Naïve Bayes Algorithm of The Data Crawler: Twitter. *International Conference on Informatics and Computing* (ICIC), 1 – 5. IEEE. doi: 10.1109/ICIC47613.2019.8985884.

Zulfikar, W. B., Irfan, M., Alam, C. N. and Indra, M. (2017). The Comparation of Text Mining with Naïve Bayes Classifier, Nearest Neighbor, and Decision Tree to Detect Indonesian Swear Words on Twitter. *International Conference on Cyber and IT Service Management* (CITSM), 1 – 5. IEEE.

iCMS2021: 080-063

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

22

THANK YOU

ANALYSIS OF PUBLIC SENTIMENT ON COVID-19 VACCINATION POLICY BASED ON TEXT MINING WITH THE NAÏVE BAYES CLASSIFIER APPROACH

23