obtained using the DT, LR, and LR-based models, SVM- and RF-based models have the highest or in the worst case, are similar. Based on the KNN, in the case of neutral text, the SVM-based model's precision score is lower than the LR-based model's precision score, which is the only exception. The same holds true for RF-based models' precision scores and recall. In most cases, these models had higher precision and recall values than the DT-based, LR-based, and KNN-based models. There are, however, some exceptions to this rule. Precision values for positive text are lower with the RF-based model compared to DT and KNN models, respectively. For negative texts, the recall score obtained using the DT-based, LR-based, and KNN-based models is lower than that obtained from the RF-based model. For both neutral and negative texts, the RF-based model's recall scores were lower than the corresponding recall scores from the DT-based model.

TABLE V
PRECISION, RECALL, F1-SCORE OF SVM-BASED MODELS

	Precision	Recall	f1-score
Negative	0.87	0.87	0.87
Neutral	0.81	0.72	0.77
Positive	0.93	0.95	0.94

TABLE VI
PRECISION, RECALL, F1-SCORE OF RF-BASED MODELS

	Precision	Recall	f1-score
Negative	0.90	0.73	0.80
Neutral	1.00	0.48	0.64
Positive	0.82	0.99	0.90

TABLE VII
PRECISION, RECALL, F1-SCORE OF DT-BASED MODELS

	Precision	Recall	f1-score
Negative	0.81	0.79	0.80
Neutral	0.77	0.73	0.75
Positive	0.91	0.93	0.92

TABLE VIII
PRECISION, RECALL, F1-SCORE OF LR-BASED MODELS

	Precision	Recall	f1-score
Negative	0.84	0.79	0.82
Neutral	0.87	0.38	0.53
Positive	0.84	0.97	0.90

TABLE IV
PRECISION, RECALL, F1-SCORE OF KNN-BASED MODELS

	Precision	Recall	f1-score
Negative	0.64	0.83	0.75
Neutral	0.65	0.65	0.65
Positive	0.91	0.83	0.87

When compared to the scores obtained using the DT-based, LR-based, and KNN-based models, the SVM- and RF-based models typically have the highest or, in the worst case, similar results. In the case of neutral text, the SVM-based model's precision score is lower than the LR-based model's precision score, which is the only exception. The same holds true for RF-based models' precision scores and recall. In most cases, these models had higher precision and recall values than the DT-based, LR-based, and KNN-based models. There are, however, some exceptions to this rule. Precision values for positive text are lower with the RF-based model compared to DT and KNN models, respectively. For negative texts, the recall score obtained using the DT-based, LR-based, and KNN-based models is lower than that obtained from the RF-based model. For both neutral and negative texts, the RF-based model's recall scores were lower than the corresponding recall scores from the DT-based model.

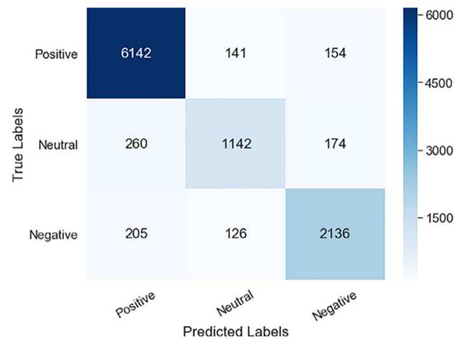

Fig. 4 Confusion matrix from SVM.

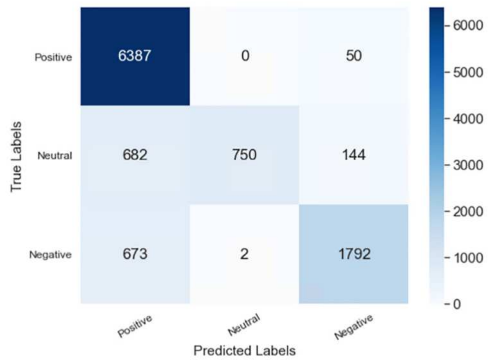

Fig. 5 Confusion matrix from RF.

The hyper-parameters were set by cross-validation with the SVM method, which took about 2 hours and 36 minutes. The RF-based model required 2 hours 46 minutes of training time. To train these models, the training time ranged from 1 to 9 seconds for each of the DT, KNN, and LR models, respectively.

Figure 4 shows the SVM-based model's confusion matrix results. Classifier correctly predicted 6142 out of 6437 positive tweets, but incorrectly predicted neutral and negative tweets in the same number of cases. The classifier correctly predicted 1142 of the 1576 neutral tweets, while incorrectly classifying 260 and 174 as positive and negative, respectively. The classifier correctly predicted 2136 of the 2467 tweets as negative, while 205 and 126 of the classifier incorrectly classified as positive and neutral.

Similar results are shown in Fig. 5, which shows the confusion matrix for the RF-based model. The classifier correctly predicted 6387 of the 6437 positive tweets, while only 50 of the negative tweets were incorrectly classified. The classifier correctly predicted 750 of the 1576 neutral tweets while incorrectly classifying 682 and 144 as positive and negative, respectively. Negative tweets accounted for 1792 of the 2467 tweets classified by the classifier as such, while positive tweets accounted for 673 and neutral tweets accounted for 2.

Each positive and negative tweet yielded 13,000 labeled samples. After 100-fold cross-validation, the best KNN-based model yields a f1 score of approximately 0.73. As part of our research, we used KNN to separate the tweets into three categories: negative, positive, and neutral. Our KNN-based model's f1 scores were 0.75, 0.65, and 0.87 for negative, neutral, and positive tweets, respectively, which are marginally better than Rumelli et al. [9]'s average. Using SVM and RBF to build our model further enhances classification performance. To represent these results, we calculated f1 scores of 0.87 for negative tweets, a neutral 0.77, and a positive 0.94 for our SVM-based model. In addition, our RF-based model's f1 scores were 0.80, 0.64, and 0.90 for negative and neutral tweets.

IV. Conclusion

Numerous educational institutions use sentiment analysis to process large amounts of data more efficiently and cost-effectively. An educational institution's ability to quickly gauge the general opinion of their students' wants and needs is made possible by using sentiment analysis. People's arguments, social media chatter, and more can be sorted out automatically so that one can make faster and more accurate decisions. Data and competitive analysis are powered by sentiment analysis. To discover new knowledge types or anticipate future trends, sentiment analysis can be extremely useful. There are numerous benefits to employing a sentiment analysis tool, including a reduction in both time and money. A supervised machine learning method was used in this study to develop a model for predicting the sentiment expressed in the text on social media. A number of different classifiers have been used to sort tweets into three categories: positively received, negatively received, and neutral. The f1 values of our KNN-based model were measured at 75%, 65%, and 87% for negative, neutral, and positive tweets, respectively, which are slightly more accurate than previous studies with the same method and purpose. People's tweets on Twitter data are used to predict the sentiment of their tweets using sentiment analysis. Prediction models based on sentiment analysis SVM and RF were found to outperform other sentiment analysis models built based on KNN. SVM and RF can be used to classify students' needs and opinions into positive, negative, and neutral groups within an acceptable error rate.

As a follow-up project, we intend to divide our training dataset into a k-fold number to enhance the sentiment analysis model's efficiency. It is possible to improve the accuracy of the training data by incorporating more words with polarity labels into it. There will be a search for new words to be added to the training data, and data from all Indonesian abbreviated words will be used to improve accuracy in future research.

References

[1] M. Ghiassi and S. Lee, "A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach," Expert Syst. Appl., vol. 106, pp. 197–216, 2018, doi: https://doi.org/10.1016/j.eswa.2018.04.006.

[2] Aryum Bibi et al., "A Novel Unsupervised Ensemble Framework using Concept-based Linguistic Methods and Machine Learning for Twitter Sentiment Analysis," Pattern Recognit. Lett., 2022, doi: https://doi.org/10.1016/j.patrec.2022.04.004.

[3] S. Yousefinaghani, R. Dara, S. Mubareka, A. Papadopoulos, and S. Sharif, "An analysis of COVID-19 vaccine sentiments and opinions on Twitter," Int. J. Infect. Dis., vol. 108, pp. 256–262, 2021, doi: https://doi.org/10.1016/j.ijid.2021.05.059.

[4] H. Vanam and J. Retna Raj R, "Analysis of twitter data through big data based sentiment analysis approaches," Mater. Today Proc., 2021, doi: https://doi.org/10.1016/j.matpr.2020.11.486.

[5] R. Riyanto, "Application of the Vector Machine Support Method in Twitter Social Media Sentiment Analysis Regarding the Covid-19 Vaccine Issue in Indonesia," J. Appl. Data Sci., vol. 2, no. 3, pp. 102–108, 2021, doi: 10.47738/jads.v2i3.40.

[6] Q. Aini, "Classification of Tweets Causing Deadlocks in Jakarta Streets with the Help of Algorithm C4.5," J. Appl. Data Sci., vol. 2, no. 4, pp. 143–156, 2021, doi: 10.47738/jads.v2i4.43.

[7] T. Hariguna and V. Rachmawati, "Community Opinion Sentiment Analysis on Social Media Using Naive Bayes Algorithm Methods," IJIIS Int. J. Informatics Inf. Syst., vol. 2, no. 1, pp. 33–38, 2019.

[8] A. Hananto, "COVID-19 Vaccination: A Retrospective Observation and Sentiment Analysis of the Twitter Social Media Platform in Indonesia," IJIIS Int. J. Informatics Inf. Syst., vol. 5, no. 1, pp. 56–69, 2022, doi: 10.47738/ijiis.v5i1.126.

[9] M. Rumelli, D. Akkuş, Ö. Kart, and Z. Isik, "Sentiment Analysis in Turkish Text with Machine Learning Algorithms," in 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), 2019, pp. 1–5, doi: 10.1109/ASYU48272.2019.8946436.

[10] P. Sharma and A. K. Sharma, "Experimental investigation of automated systems for twitter sentiment analysis to predict the public emotions using machine learning algorithms," Mater. Today Proc., 2020, doi: https://doi.org/10.1016/j.matpr.2020.09.351.

[11] J. R. Saura, D. Ribeiro-Soriano, and P. Zegarra Saldaña, "Exploring the challenges of remote work on Twitter users' sentiments: From digital technology development to a post-pandemic era," J. Bus. Res., vol. 142, pp. 242–254, 2022, doi: https://doi.org/10.1016/j.jbusres.2021.12.052.

[12] S. H. Biradar, J. V Gorabal, and G. Gupta, "Machine learning tool for exploring sentiment analysis on twitter data," Mater. Today Proc., vol. 56, pp. 1927–1934, 2022, doi: https://doi.org/10.1016/j.matpr.2021.11.199.

[13] T. Astuti and I. Pratika, "Product Review Sentiment Analysis by Artificial Neural Network Algorithm," IJIIS Int. J. Informatics Inf. Syst., vol. 2, no. 2, pp. 61–66, 2019, doi: 10.47738/ijiis.v2i2.15.

[14] S. Liu and J. Liu, "Public attitudes toward COVID-19 vaccines on English-language Twitter: A sentiment analysis," Vaccine, vol. 39, no. 39, pp. 5499–5505, 2021, doi: https://doi.org/10.1016/j.vaccine.2021.08.058.

[15] R. Endsuy, "Sentiment Analysis between VADER and EDA for the US Presidential Election 2020 on Twitter Datasets," J. Appl. Data Sci., vol. 2, no. 1, pp. 8–18, 2021, doi: 10.47738/jads.v2i1.17.

[16] T. Hariguna, W. M. Baihaqi, and A. Nurwanti, "Sentiment Analysis of Product Reviews as A Customer Recommendation Using the Naive Bayes Classifier Algorithm," IJIIS Int. J. Informatics Inf. Syst., vol. 2, no. 2, pp. 48–55, 2019, doi: 10.47738/ijiis.v2i2.13.

[17] A. Perti, M. C. Trivedi, and A. Sinha, "Development of intelligent model for twitter sentiment analysis," Mater. Today Proc., vol. 33, pp. 4515–4519, 2020, doi: https://doi.org/10.1016/j.matpr.2020.08.004.

[18] T. H. Jaya Hidayat, Y. Ruldeviyani, A. R. Aditama, G. R. Madya, A. W. Nugraha, and M. W. Adisaputra, "Sentiment analysis of twitter data related to Rinca Island development using Doc2Vec and SVM and logistic regression as classifier," Procedia Comput. Sci., vol. 197, pp. 660–667, 2022, doi: https://doi.org/10.1016/j.procs.2021.12.187.

[19] T. Hariguna, H. T. Sukmana, and J. Il Kim, "Survey Opinion using Sentiment Analysis," J. Appl. Data Sci., vol. 1, no. 1, pp. 35–40, 2020.

[20] Z. Bokaee Nezhad and M. A. Deihimi, "Twitter sentiment analysis from Iran about COVID 19 vaccine," Diabetes Metab. Syndr. Clin. Res. Rev., vol. 16, no. 1, p. 102367, 2022, doi: https://doi.org/10.1016/j.dsx.2021.102367.

[21] D. Sunitha, R. K. Patra, N. V Babu, A. Suresh, and S. C. Gupta, "Twitter Sentiment Analysis Using Ensemble based Deep Learning Model towards COVID-19 in India and European Countries," Pattern Recognit. Lett., 2022, doi: https://doi.org/10.1016/j.patrec.2022.04.027.