with details of 25,281 data categorized as containing depression (Depression) and 25,242 data that does not contain depression (Non-Depression). In this study, FastText plays a role in feature expansion by categorizing similar words into three categories including Top 1, Top 5, and Top 10, based on corpus similarity totaling 151,117 data derived from three datasets, namely Tweet, IndoNews, and a combination of both. This research uses five evaluation scenarios for the hybrid deep learning model, namely Split Data, N-Gram, Max Features, Feature Expansion, and Optimization, where the best results from each scenario are used as the basis for the next scenario. Based on scenarios 1 to 4, each model achieved the best performance at 90:10 Split Data, the best N-Gram variation with Unigram + Bigram + Trigram, and the best number of max features, which is 10,000 for CNN and GRU models, and 5,000 for CNN + GRU and GRU + CNN hybrid deep learning models. Furthermore, the application of FastText Feature Expansion using the Top 1 category from the corpus similarity combination of the Tweet + IndoNews dataset resulted in optimal performance for the hybrid model. In the optimization scenario, Optimizer Nadam provides the best results for the CNN + GRU model, while Optimizer Adam provides optimal performance for the GRU + CNN model.

The results showed that the non-hybrid CNN and GRU models only experienced an accuracy improvement of 0.02% in the N-Gram scenario, with the best accuracy of 83.44% for CNN and 82.97% for GRU, respectively. This can be explained because the non-hybrid architecture has limitations in capturing complex patterns compared to the hybrid architecture. In contrast, in the FastText Feature Expansion scenario, the CNN + GRU and GRU + CNN hybrid models showed improved accuracy. The CNN + GRU model achieved the best accuracy of 83.19% with an increase of 1.36% from the baseline, while the GRU + CNN model achieved the best accuracy of 83.32% with an increase of 1.44%. The successful application of feature expansion using FastText and optimization on the hybrid model shows that this method is able to enrich the word information in the dataset and minimize the loss function resulting in better model performance. As a suggestion for future research, other feature expansion methods can be applied to increase the variety and complexity of data. In addition, the use of data from other social media can be considered to expand the scope of the research so as to allow for more optimal performance of the hybrid deep learning model.

REFERENCES

[1] S. Khan and S. Alqahtani, “Hybrid machine learning models to detect signs of depression,” *Multimed. Tools Appl.*, vol. 83, no. 13, pp. 38819–38837, 2024, doi: 10.1007/s11042-023-16221-z.

[2] Vandana, N. Marriwala, and D. Chaudhary, “A hybrid model for depression detection using deep learning,” *Meas. Sensors*, vol. 25, no. September 2022, p. 100587, 2023, doi: 10.1016/j.measen.2022.100587.

[3] J. Angskun, S. Tipprasert, and T. Angskun, “Big data analytics on social networks for real-time depression detection,” *J. Big Data*, vol. 9, no. 1, pp. 2–15, 2022, doi: 10.1186/s40537-022-00622-2.

[4] who.int, “Depressive disorder (depression),” *www.who.int*, 2023. https://www.who.int/news-room/fact-sheets/detail/depression (accessed Apr. 24, 2024).

[5] F. M. Shah *et al.*, “Early Depression Detection from Social Network Using Deep Learning Techniques,” *2020 IEEE Reg. 10 Symp. TENSYMP 2020*, vol. 1, no. June, pp. 823–826, 2020, doi: 10.1109/TENSYMP50017.2020.9231008.

[6] S. Ghosal and A. Jain, “Depression and Suicide Risk Detection on Social Media using fastText Embedding and XGBoost Classifier,” *Procedia Comput. Sci.*, vol. 218, pp. 1631–1639, 2023, doi: 10.1016/j.procs.2023.01.141.

[7] J. Cha, S. Kim, and E. Park, “A lexicon-based approach to examine depression detection in social media: the case of Twitter and university community,” *Humanit. Soc. Sci. Commun.*, vol. 9, no. 1, pp. 1–10, 2022, doi: 10.1057/s41599-022-01313-2.

[8] R. Safa, P. Bayat, and L. Moghtader, *Automatic detection of depression symptoms in twitter using multimodal analysis*, vol. 78, no. 4. Springer US, 2022.

[9] statista.com, “Leading countries based on number of X (formerly Twitter) users as of April 2024(in millions)[Graph],” *We Are Social, &DataReportal, &Meltwater*, 2024. https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ (accessed May 07, 2024).

[10] V. Tejaswini, K. S. Babu, and B. Sahoo, “Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep Learning Model,” *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 23, no. 1, pp. 1–20, 2024, doi: 10.1145/3569580.

[11] M. De Choudhury, S. Counts, and E. Horvitz, “Social media as a measurement tool of depression in populations,” *Proc. 5th Annu. ACM Web Sci. Conf. WebSci’13*, pp. 47–56, 2013, doi: 10.1145/2464464.2464480.

[12] Y. Cai, H. Wang, H. Ye, Y. Jin, and W. Gao, “Depression detection on online social network with multivariate time series feature of user depressive symptoms,” *Expert Syst. Appl.*, vol. 217, no. August 2022, p. 119538, 2023, doi: 10.1016/j.eswa.2023.119538.

[13] R. Skaik and D. Inkpen, “Using twitter social media for depression detection in the canadian population,” *ACM Int. Conf. Proceeding Ser.*, pp. 109–114, 2020, doi: 10.1145/3442536.3442553.

[14] H. Zogan, I. Razzak, S. Jameel, and G. Xu, “DepressionNet: Learning Multi-modalities with User Post Summarization for Depression Detection on Social Media,” *SIGIR 2021 - Proc. 44th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 133–142, 2021, doi: 10.1145/3404835.3462938.

[15] J. Hussain *et al.*, “Exploring the dominant features of social media for depression detection,” *J. Inf. Sci.*, vol. 46, no. 6, pp. 739–759, 2020, doi: 10.1177/0165551519860469.

[16] M. Hasan, E. Rundensteiner, and E. Agu, “Automatic emotion detection in text streams by analyzing Twitter data,” *Int. J. Data Sci. Anal.*, vol. 7, no. 1, pp. 35–51, 2019, doi: 10.1007/s41060-018-0096-z.

[17] M. Ahmad Wani, M. A. Elaffendi, K. A. Shakil, A. Shariq Imran, and A. A. Abd El-Latif, “Depression Screening in Humans With AI and Deep Learning Techniques,” *IEEE Trans. Comput. Soc. Syst.*, vol. 10, no. 4, pp. 2074–2089, 2023, doi: 10.1109/TCSS.2022.3200213.

[18] C. Lin *et al.*, “SenseMood: Depression detection on social media,” *ICMR 2020 - Proc. 2020 Int. Conf. Multimed. Retr.*, pp. 407–411, 2020, doi: 10.1145/3372278.3391932.

[19] Y. Zhang, Y. Wang, X. Wang, B. Zou, and H. Xie, “Text-based Decision Fusion Model for Detecting Depression,” *ACM Int. Conf.*

Proceeding Ser., pp. 101–106, 2020, doi: 10.1145/3421515.3421516.

[20] T. Gui *et al.*, “Cooperative multimodal approach to depression detection in twitter,” *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 110–117, 2019, doi: 10.1609/aaai.v33i01.3301110.

[21] M. Bañón *et al.*, “ParaCrawl: Web-scale acquisition of parallel corpora,” *Proc. Annu. Meet. Assoc. Comput. Linguist.*, pp. 4555–4567, 2020, doi: 10.18653/v1/2020.acl-main.417.

[22] K. U. Wijaya and E. B. Setiawan, “Hate Speech Detection Using Convolutional Neural Network and Gated Recurrent Unit with FastText Feature Expansion on Twitter,” *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 3, pp. 619–631, 2023, doi: 10.26555/jiteki.v9i3.26532.

[23] Crisanadenta Wintang Kencana, Erwin Budi Setiawan, and Isman Kurniawan, “Hoax Detection System on Twitter using Feed-Forward and Back-Propagation Neural Networks Classification Method,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 4, pp. 655–663, 2020, doi: 10.29207/resti.v4i4.2038.

[24] A. Hakim, A. Erwin, K. Eng, M. Galinium, and W. Muliady, “Automated document classification for news article in Bahasa Indonesia based on term frequency inverse document frequency (TF-IDF) approach,” *2014 6th Int. Conf. Inf. Technol. Electr. Eng.*, pp. 1–4, 2014, doi: 10.1109/ICITEED.2014.7007894.

[25] H. Liu, X. Chen, and X. Liu, “A Study of the Application of Weight Distributing Method Combining Sentiment Dictionary and TF-IDF for Text Sentiment Analysis,” *IEEE Access*, vol. 10, pp. 32280–32289, 2022, doi: 10.1109/ACCESS.2022.3160172.

[26] E. B. Setiawan, D. H. Widyantoro, and K. Surendro, “Feature expansion using word embedding for tweet topic classification,” *Proceeding 2016 10th Int. Conf. Telecommun. Syst. Serv. Appl. TSSA 2016 Spec. Issue Radar Technol.*, no. October, 2017, doi: 10.1109/TSSA.2016.7871085.

[27] R. A. Yahya and E. B. Setiawan, “Feature Expansion with FastText on Topic Classification Using the Gradient Boosted Decision Tree on Twitter,” *2022 10th Int. Conf. Inf. Commun. Technol. ICoICT 2022*, no. January, pp. 322–327, 2022, doi: 10.1109/ICoICT55009.2022.9914896.

[28] H. R. Alhakiem and E. B. Setiawan, “Aspect-Based Sentiment Analysis on Twitter Using Logistic Regression with FastText Feature Expansion,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 5, pp. 840–846, 2022, doi: 10.29207/resti.v6i5.4429.

[29] N. Nedjah, I. Santos, and L. de Macedo Mourelle, “Sentiment analysis using convolutional neural network via word embeddings,” *Evol. Intell.*, vol. 15, no. 4, pp. 2295–2319, 2022, doi: 10.1007/s12065-019-00227-4.

[30] B. Gupta, P. Prakasam, and T. Velmurugan, “Integrated BERT embeddings, BiLSTM-BiGRU and 1-D CNN model for binary sentiment classification analysis of movie reviews,” *Multimed. Tools Appl.*, vol. 81, no. 23, pp. 33067–33086, 2022, doi: 10.1007/s11042-022-13155-w.

[31] M. Umer *et al.*, “Impact of convolutional neural network and FastText embedding on text classification,” *Multimed. Tools Appl.*, vol. 82, no. 4, pp. 5569–5585, 2023, doi: 10.1007/s11042-022-13459-x.

[32] S. Soni, S. S. Chouhan, and S. S. Rathore, “TextConvoNet: a convolutional neural network based architecture for text classification,” *Appl. Intell.*, vol. 53, no. 11, pp. 14249–14268, 2023, doi: 10.1007/s10489-022-04221-9.

[33] M. S. Rani and S. Subramanian, “Attention Mechanism with Gated Recurrent Unit Using Convolutional Neural Network for Aspect Level Opinion Mining,” *Arab. J. Sci. Eng.*, vol. 45, no. 8, pp. 6157–6169, 2020, doi: 10.1007/s13369-020-04497-4.

[34] A. H. Uddin, D. Bapery, and A. S. Mohammad Arif, “Depression Analysis of Bangla Social Media Data using Gated Recurrent Neural Network,” *1st Int. Conf. Adv. Sci. Eng. Robot. Technol. 2019, ICASERT 2019*, vol. 2019, no. Icasert, pp. 1–6, 2019, doi: 10.1109/ICASERT.2019.8934455.

[35] S. Zhang, J. Luo, S. Wang, and F. Liu, “Oil price forecasting: A hybrid GRU neural network based on decomposition–reconstruction methods,” *Expert Syst. Appl.*, vol. 218, no. January, p. 119617, 2023, doi: 10.1016/j.eswa.2023.119617.

[36] M. A. K. Raiaan *et al.*, “A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks,” *Decis. Anal. J.*, vol. 11, no. April, p. 100470, 2024, doi: 10.1016/j.dajour.2024.100470.