UNIVERSITAS
GADJAH MADA
Weakly Supervised Emotion Classification for Indonesian Tweets
Pradipta Rakahartyanto, Drs. Edi Winarko, M.Sc.,Ph.D
Universitas Gadjah Mada, 2023 | Diunduh dari http://etd.repository.ugm.ac.id/

REFERENCES

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, *5*(2), 157–166. https://doi.org/10.1109/72.279181

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. https://doi.org/10.3115/v1/d14-1162

Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & Ré, C. (2019). Snorkel: Rapid Training Data Creation with weak supervision. *The VLDB Journal*, *29*(2-3), 709–730. https://doi.org/10.1007/s00778-019-00552-1

Saputri, M. S., Mahendra, R., & Adriani, M. (2018). Emotion Classification on Indonesian Twitter dataset. *2018 International Conference on Asian Language Processing (IALP)*. https://doi.org/10.1109/ialp.2018.8629262

Chatterjee, A., Gupta, U., Chinnakotla, M. K., Srikanth, R., Galley, M., & Agrawal, P. (2019). Understanding emotions in text using Deep Learning and Big Data. *Computers in Human Behavior*, *93*, 309–317. https://doi.org/10.1016/j.chb.2018.12.029

Bouchekif, A., Joshi, P., Bouchekif, L., & Afli, H. (2019). Epita-adapt at Semeval-2019 task 3: Detecting emotions in textual conversations using deep learning models combination. *Proceedings of the 13th International Workshop on Semantic Evaluation*. https://doi.org/10.18653/v1/s19-2035

Shah, F. M., Reyadh, A. S., Shaafi, A. I., Ahmed, S., & Sithil, F. T. (2019). Emotion detection from tweets using AIT-2018 dataset. *2019 5th International Conference on Advances in Electrical Engineering (ICAEE)*. https://doi.org/10.1109/icaee48663.2019.8975433

Ardiada, I. M., Sudarma, M., & Giriantari, D. (2019). Text mining pada sosial media untuk Mendeteksi Emosi Pengguna Menggunakan metode support vector machine Dan K-nearest neighbour. *Majalah Ilmiah Teknologi Elektro*, *18*(1), 55. https://doi.org/10.24843/mite.2019.v18i01.p08

UNIVERSITAS
GADJAH MADA
Weakly Supervised Emotion Classification for Indonesian Tweets
Pradipta Rakahartyanto, Drs. Edi Winarko, M.Sc.,Ph.D
Universitas Gadjah Mada, 2023 | Diunduh dari http://etd.repository.ugm.ac.id/

Fanesya, F. (2019). *Deteksi emosi Pada Twitter menggunakan metode naïve Bayes Dan Kombinasi ...* Deteksi Emosi Pada Twitter Menggunakan Metode Naïve Bayes Dan Kombinasi Fitur, from http://download.garuda.kemdikbud.go.id/article.php?article=1218473&val=10384&title=Deteksi%20Emosi%20Pada%20Twitter%20Menggunakan%20Metode%20Nave%20Bayes%20Dan%20Kombinasi%20Fitur

Junianto, E., & Rachman, R. (2020). Penerapan metode naïve Bayes classifier untuk Mendeteksi Emosi Pada KOMENTAR media sosial. *Jurnal Responsif : Riset Sains Dan Informatika*, *2*(1), 1–8. https://doi.org/10.51977/jti.v2i1.156

Karna, M., Juliet, D. S., & Joy, R. C. (2020). Deep Learning based text emotion recognition for chatbot applications. *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*. https://doi.org/10.1109/icoei48184.2020.9142879

Hilmiaji, N., Lhaksmana, K. M., & Purbolaksono, M. D. (2021). Identifying emotion on Indonesian tweets using convolutional neural networks. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(3), 584–593. https://doi.org/10.29207/resti.v5i3.3137

Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based Emotion Detection: A review of Bert-based approaches. *Artificial Intelligence Review*, *54*(8), 5789–5829. https://doi.org/10.1007/s10462-021-09958-2

Banerjee, S., Goswami, S., Das, A., Saha, N., Seth, S., Bhattacharya, S., Mandal, S., & Kolkata, U. E. M. (2021). A novel approach for emotion detection from text data using Natural Language Processing and machine learning. *American Journal of Electronics & Communication*, *2*(1), 16–24. https://doi.org/10.15864/ajec.2105

Kabir, M. Y., & Madria, S. (2021). EMOCOV: Machine learning for emotion detection, analysis and visualization using COVID-19 tweets. *Online Social Networks and Media*, *23*, 100135. https://doi.org/10.1016/j.osnem.2021.100135

Cortiz, D., Silva, J. O., Calegari, N., Freitas, A. L., Soares, A. A., Botelho, C., Rêgo, G. G., Sampaio, W., & Boggio, P. S. (2021). A weakly supervised dataset of fine-grained

UNIVERSITAS
GADJAH MADA
Weakly Supervised Emotion Classification for Indonesian Tweets
Pradipta Rakahartyanto, Drs. Edi Winarko, M.Sc.,Ph.D
Universitas Gadjah Mada, 2023 | Diunduh dari http://etd.repository.ugm.ac.id/

emotions in Portuguese. *Anais Do XIII Simpósio Brasileiro De Tecnologia Da Informação e Da Linguagem Humana (STIL 2021)*. https://doi.org/10.5753/stil.2021.17786

Goetz, S. J., Heaton, C., Imran, M., Pan, Y., Tian, Z., Schmidt, C., Qazi, U., Ofli, F., & Mitra, P. (2022). Food insufficiency and Twitter emotions during a pandemic. *Applied Economic Perspectives and Policy*. https://doi.org/10.1002/aepp.13258

Mahdikhani, M. (2022). Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of covid-19 pandemic. *International Journal of Information Management Data Insights*, *2*(1), 100053. https://doi.org/10.1016/j.jjimei.2021.100053

Stephan, A. and Roth, B. (2022) 'WeaNF":" weak supervision with normalizing flows', *Proceedings of the 7th Workshop on Representation Learning for NLP* [Preprint]. doi:10.18653/v1/2022.repl4nlp-1.27.

Saxena, S. (2023). *Learn about long short-term memory (LSTM) algorithms*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/

Vila-Suero, Daniel, and Francisco Aranda. 2023. "Argilla - Open-source framework for data-centric NLP." Version 1.2.0. Released on January 12, 2023. [Online]. Available: https://github.com/argilla-io/argilla

UNIVERSITAS
GADJAH MADA
Weakly Supervised Emotion Classification for Indonesian Tweets
Pradipta Rakahartyanto, Drs. Edi Winarko, M.Sc.,Ph.D
Universitas Gadjah Mada, 2023 | Diunduh dari http://etd.repository.ugm.ac.id/

APPENDICES

APPENDIX 1. Rulesets

List of Rulesets added in Rubrix

keyword	label	coverage	annotated_coverage	overlaps	conflicts	correct	incorrect	precision
sedih	{sadness}	0.002617	0.012500	0.000544	0.000480	9	2	0.818182
senang	{happy}	0.002554	0.000000	0.000177	0.000139	0	0	NaN
seneng	{happy}	0.001871	0.006818	0.000303	0.000253	2	4	0.333333
marah	{anger}	0.001669	0.012500	0.000303	0.000215	8	3	0.727273
takut	{fear}	0.005247	0.075000	0.000430	0.000367	55	11	0.833333
serem	{fear}	0.001264	0.022727	0.000126	0.000088	18	2	0.900000
sayang	{love}	0.003805	0.097727	0.000721	0.000442	55	31	0.639535
cinta	{love}	0.002857	0.071591	0.000493	0.000265	48	15	0.761905
astagfirullah	{anger}	0.000632	0.001136	0.000051	0.000025	1	0	1.000000
alhamdulillah	{happy}	0.004741	0.019318	0.000253	0.000139	14	3	0.823529
khawatir	{fear}	0.000683	0.004545	0.000051	0.000025	3	1	0.750000
kasihan	{sadness}	0.000582	0.003409	0.000013	0.000000	1	2	0.333333
kesel	{anger}	0.001429	0.004545	0.000164	0.000088	2	2	0.500000
anjing	(anger}	0.001669	0.005682	0.000228	0.000177	2	3	0.400000

UNIVERSITAS
GADJAH MADA
Weakly Supervised Emotion Classification for Indonesian Tweets
Pradipta Rakahartyanto, Drs. Edi Winarko, M.Sc.,Ph.D
Universitas Gadjah Mada, 2023 | Diunduh dari http://etd.repository.ugm.ac.id/

kontol	{anger}	0.000834	0.000000	0.000038	0.000000	0	0	NaN
love	{love}	0.003944	0.010227	0.000253	0.000139	9	0	1.000000
happy	{happy}	0.004640	0.013636	0.000341	0.000253	7	5	0.583333
sad	{sadness}	0.000594	0.000000	0.000025	0.000025	0	0	NaN
meninggal	{sadness}	0.000544	0.002273	0.000051	0.000051	1	1	0.500000
bucin	{love}	0.000367	0.001136	0.000013	0.000000	0	1	0.000000
bahagia	{happy}	0.001820	0.015909	0.000303	0.000215	6	8	0.428571
Innalillahi	{sadness}	0.000076	0.000000	0.000000	0.000000	0	0	NaN
astaga	{sadness}	0.000759	0.001136	0.000025	0.000025	0	1	0.000000
babi	{anger}	0.000961	0.001136	0.000051	0.000013	1	0	1.000000
ngentot	{anger}	0.000417	0.001136	0.000013	0.000000	1	0	1.000000
akit hat	{sadness}	0.000860	0.007955	0.000126	0.000076	4	3	0.571429
atah hat	{sadness}	0.000190	0.002273	0.000101	0.000076	2	0	1.000000
total	{love, anger, happy, fear, sadness}	0.044918	0.330682	0.002491	0.001707	249	98	0.717579