Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., … Zheng, X. (2015). *TensorFlow, Large-scale machine learning on heterogeneous systems*. https://doi.org/10.5281/zenodo.4724125

Adelaar, K. A. (1993). The internal classification of the malayic subgroup. *Bulletin of the School of Oriental and African Studies, University of London*, *56*, 566–581. http://www.jstor.org.ezproxy.ugm.ac.id/stable/620695

Aini, Q., Julianto, A. E., & Purbohadi, D. (2018). Development of a scoring application for indonesian language essay questions. *ACM International Conference Proceeding Series*, 6–10. https://doi.org/10.1145/3291078.3291099

Alkhatib, B., & Eddin, M. M. K. (2020). Voice identification using mfcc and vector quantization. *Baghdad Science Journal*, *17*. https://www.iasj.net/iasj/article/187369

Arslan, A., & Yildiz, O. (2018). Automated auscultative diagnosis system for evaluation of phonocardiogram signals associated with heart murmur diseases. *Gazi University Journal of Science*, *31*.

Blust, R. (2013). *The austronesian languages*. Asia-Pacific Linguistics, School of Culture, History; Language, College of Asia; the Pacific, The Australian National University.

Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., Roma, G., Salamon, J., Zapata, J. R., & Serra, X. (2013). Essentia: An audio analysis library for music information retrieval. *International Society for Music Information Retrieval Conference (ISMIR'13)*.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & Zhang, Q. (2018). *JAX: Composable transformations of Python+NumPy programs* (Version 0.2.5). http://github.com/google/jax

Bre, F., Gimenez, J., & Fachinotti, V. (2017). Prediction of wind pressure coefficients on building surfaces using artificial neural networks. *Energy and Buildings*, *158*. https://doi.org/10.1016/j.enbuild.2017.11.045

89

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., & Varoquaux, G. (2013). API design for machine learning software: Experiences from the scikit-learn project. *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 108–122.

Burrus, C. S., & Parks, T. (1985). Convolution algorithms. *Citeseer: New York, NY, USA*.

Butterworth, S., et al. (1930). On the theory of filter amplifiers. *Wireless Engineer*, *7*(6), 536–541.

Cai, Z., Lin, N., Ma, C., & Jiang, S. (2019). Indonesian automatic text summarization based on a new clustering method in sentence level. *ACM International Conference Proceeding Series*, 30–35. https://doi.org/10.1145/3341620.3341626

Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. *CoRR*, *abs/1405.3531*. http://arxiv.org/abs/1405.3531

Cole, P., Hara, Y., & Yap, N. T. (2008). Auxiliary fronting in peranakan javanese. *Journal of Linguistics*, *44*, 1–43. http://www.jstor.org/stable/40058027

Conners, T. J., & Vander Klok, J. (2016). On language documentation of colloquial javanese varieties. *Proceedings of 2016 Annual Conference of the Canadian Linguistics Association*, 1–12.

DeGraff, M. (2007). Kreyòl ayisyen, or haitian creole (creole french). *Comparative creole syntax: Parallel outlines of*, *18*, 101–126.

Draghici, A., Abeßer, J., & Lukashevich, H. (2020). A study on spoken language identification using deep neural networks. *Proceedings of the 15th International Conference on Audio Mostly*, 253–256.

El Naqa, I., & Murphy, M. J. (2015). What is machine learning? In *Machine learning in radiation oncology* (pp. 3–11). Springer.

Fausett, L. V. (2006). *Fundamentals of neural networks: Architectures, algorithms and applications*. Pearson Education India.

Garcia-Romero, D., & McCree, A. (2016). Stacked long-term tdnn for spoken language recognition. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *08-12-September-2016*, 3226–3230. https://doi.org/10.21437/Interspeech.2016-1334

Gonzalez-Dominguez, J., Lopez-Moreno, I., Sak, H., Gonzalez-Rodriguez, J., & Moreno, P. J. (2014). Automatic language identification using long short-term memory recurrent neural networks. *Fifteenth annual conference of the international speech communication association*.

Gris, L., & Junior, A. C. (2020). Automatic spoken language identification using convolutional neural networks. *Anais do XVII Congresso Latino-Americano de Software Livre e Tecnologias Abertas*, 16–20. https://doi.org/10.5753/latinoware.2020.18603

Hanani, A., & Naser, R. (2020). Spoken arabic dialect recognition using x-vectors. *Natural Language Engineering*, *26*, 691–700. https://doi.org/10.1017/S1351324920000091

Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S., Sengupta, S., Coates, A., et al. (2014). Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:1412.5567*.

Haryono, H. (2019). Identity politics and symbolic interactions between sundanese and javanese in indonesia. *JCIC : Jurnal CIC Lembaga Riset dan Konsultan Sosial*, *1*. https://doi.org/10.51486/jbo.v1i1.7

Hassine, M., Messaoud, H., & Boussaid, L. (2018). Tunisian dialect recognition based on hybrid techniques. *International Arab Journal of Information Technology*, 58–65. https://www.researchgate.net/publication/321973464_Tunisian_Dialect_Recognition_Based_on_Hybrid_Techniques

Haugen, E. (1966). Dialect, language, nation. *American Anthropologist*, *68*, 922–935. https://doi.org/10.1525/aa.1966.68.4.02a00040

Hidayatullah, A. F., Cahyaningtyas, S., & Pamungkas, R. D. (2020). Attention-based cnn-bilstm for dialect identification on javanese text. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 317–324.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Huzaifah, M. (2017). Comparison of time-frequency representations for environmental sound classification using convolutional neural networks. *CoRR*, *abs/1706.07156*. http://arxiv.org/abs/1706.07156

Ibrohim, M. O., Setiadi, M. A., & Budi, I. (2019). Identification of hate speech and abusive language on indonesian twitter using theword2vec, part of speech and emoji features. *ACM International Conference Proceeding Series*, 1–5. https://doi.org/10.1145/3373477.3373495

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A convolutional neural network for modelling sentences. *52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference*, *1*, 655–665. https://doi.org/10.3115/v1/p14-1062

Kehtarnavaz, N. (2008). Chapter 7 - frequency domain processing. In N. Kehtarnavaz (Ed.), *Digital signal processing system design (second edition)* (Second Edi-

tion, pp. 175–196). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-374490-6.00007-6

Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, *6*(1), 97–133. https://doi.org/https://doi.org/10.1075/ijcl.6.1.05kil

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger (Eds.), *Advances in neural information processing systems*. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Lever, J., Krzywinski, M., & Altman, N. (2016). Points of significance: Model selection and overfitting. *Nature methods*, *13*(9), 703–705.

Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*, *22*, 319–344. https://doi.org/10.1075/ijcl.22.3.02lov

Lozano-Diez, A., Zazo-Candil, R., Gonzalez-Dominguez, J., Toledano, D. T., & Gonzalez-Rodriguez, J. (2015). An end-to-end approach to language identification in short utterances using convolutional neural networks. *Sixteenth Annual Conference of the International Speech Communication Association*.

Masterson, K. (2010). *From grunting to gabbing: Why humans can talk : Npr*. https://www.npr.org/templates/story/story.php?storyId=129083762

McFee, B., Raffel, C., Liang, D., Ellis, D., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: Audio and music signal analysis in python. *Proceedings of the 14th Python in Science Conference*, 18–24. https://doi.org/10.25080/majora-7b98e3ed-003

Merriam-Webster. (n.d.-a). Accuracy. In *Merriam-webster.com dictionary*. Retrieved August 3, 2021, from https://www.merriam-webster.com/dictionary/accuracy.

Merriam-Webster. (n.d.-b). Model. In *Merriam-webster.com dictionary*. Retrieved August 3, 2021, from https://www.merriam-webster.com/dictionary/model.

mohd ali, Y., M P, P., Yaacob, S., Yusuf, R., & Abu bakar, S. (2013). Analysis of accent-sensitive words in multi-resolution mel-frequency cepstral coefficients for classification of accents in malaysian english. *International Journal of Automotive and Mechanical Engineering*, *7*, 1053–1073. https://doi.org/10.15282/ijame.7.2012.21.0086

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. MIT press.

Nagrani, A., Chung, J. S., & Zisserman, A. (2017). Voxceleb: A large-scale speaker identification dataset. *CoRR*, *abs/1706.08612*. http://arxiv.org/abs/1706.08612

Nothofer, B. (1986). The barrier island languages in the austronesian language family. *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics*, 87–109.

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *CoRR*, *abs/1511.08458*. http://arxiv.org/abs/1511.08458

Pal, S. K., & Mitra, S. (1992). Multilayer perceptron, fuzzy sets, classifiaction.

Park, D. S., Chan, W., Zhang, Y., Chiu, C.-C., Zoph, B., Cubuk, E. D., & Le, Q. V. (2019). SpecAugment: A simple data augmentation method for automatic speech recognition. *Interspeech 2019*. https://doi.org/10.21437/interspeech.2019-2680

Park, H., Lee, D., Lim, M., Kang, Y., Oh, J., & Kim, J. (2018). A fast-converged acoustic modeling for korean speech recognition: A preliminary study on time delay neural network. *CoRR*, *abs/1807.05855*. http://arxiv.org/abs/1807.05855

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox & R. Garnett (Eds.), *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Peddinti, V., Povey, D., & Khudanpur, S. (2015). A time delay neural network architecture for efficient modeling of long temporal contexts. *Sixteenth annual conference of the international speech communication association*.

Poedjosoedarmo, S. (1968). Javanese speech levels. *Indonesia*, (6), 54–81. http://www.jstor.org/stable/3350711

Prasad, N. V., & Umesh, S. (2013). Improved cepstral mean and variance normalization using bayesian framework. *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, 156–161. https://doi.org/10.1109/ASRU.2013.6707722

Putra, S. J., Khalil, I., Gunawan, M. N., Amin, R. I., & Sutabri, T. (2018). A hybrid model for social media sentiment analysis for indonesian text. *ACM International Conference Proceeding Series*, 297–301. https://doi.org/10.1145/3282373.3282850

Rahmah, L. S., & Kurniawan, E. (2020). Tense and aspect markers of cirebon javanese, 112–115. https://doi.org/10.2991/assehr.k.200406.022

Rahmawati, R., & Lestari, D. P. (2018). Java and sunda dialect recognition from indonesian speech using gmm and i-vector. *Proceeding of 2017 11th International Conference on Telecommunication Systems Services and Applications, TSSA 2017, 2018-January*, 1–5. https://doi.org/10.1109/TSSA.2017.8272892

Ren, Z., Yang, G., & Xu, S. (2019). Two-stage training for chinese dialect recognition. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2019-September*, 4050–4054. https://doi.org/10.21437/Interspeech.2019-1522

Reyes, F., & Fajardo, A. (2019). Building corpora of spoken filipino words using speech segmentation with automatic labeling. *Proceeding - 2019 International Conference on ICT for Smart Society: Innovation and Transformation Toward Smart Region, ICISS 2019*. https://doi.org/10.1109/ICISS48059.2019.8969813

Reynaldi, R. G. (2017). *Phonological differences between standard javanese and banyumas dialect of javanese* (1). http://ejournal2.undip.ac.id/index.php/compendium

Salam, K. M. A., Rahman, M., & Khan, M. M. S. (2019). Developing the bangladeshi national corpus- a balanced and representative bangla corpus. *2019 International Conference on Sustainable Technologies for Industry 4.0, STI 2019*. https://doi.org/10.1109/STI47673.2019.9068005

Salman, S., & Liu, X. (2019). Overfitting mechanism and avoidance in deep neural networks. *CoRR, abs/1901.06566*. http://arxiv.org/abs/1901.06566

Schelter, S., Biessmann, F., Januschowski, T., Salinas, D., Seufert, S., & Szarvas, G. (2018). On challenges in machine learning model management.

Shahab, A., & Lestari, D. (2017). An investigation of indonesian speaker identification for channel dependent modeling using i-vector. *2016 Conference of the Oriental Chapter of International Committee for Coordination and Standardization of Speech Databases and Assessment Techniques, O-COCOSDA 2016*, 151–155. https://doi.org/10.1109/ICSDA.2016.7919002

Shon, S., Ali, A., & Glass, J. (2018). Convolutional neural network and language embeddings for end-to-end dialect recognition. *Proc. Odyssey 2018 The Speaker and Language Recognition Workshop*, 98–104. https://doi.org/10.21437/Odyssey.2018-14

Singh, G., Sharma, S., Kumar, V., Kaur, M., Baz, M., & Masud, M. (2021). Spoken language identification using deep learning. *Computational Intelligence and Neuroscience, 2021*. https://doi.org/10.1155/2021/5123671

Sneddon, J. (2003). The indonesian language. *Its history and role in modern society, Sydney*.

Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., & Khudanpur, S. (2018). X-vectors: Robust dnn embeddings for speaker recognition. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. http://www.danielpovey.com/files/2018_icassp_xvectors.pdf

Sreeraj, V. V., & Rajan, R. (2018). Automatic dialect recognition using feature fusion. *Proceedings - International Conference on Trends in Electronics and Informatics, ICEI 2017*, *2018-January*, 435–439. https://doi.org/10.1109/ICOEI.2017.8300964

Sulaiman, A. (2018). Javanese language of banten dialects in tangerang regency. *Journal of English Language and Literature (JELL)*, *2*, 117–124. https://doi.org/10.37110/jell.v2i01.20

Tawaqal, B., & Suyanto, S. (2021). Recognizing five major dialects in indonesia based on mfcc and drnn. *Journal of Physics: Conference Series*, *1844*, 12003. https://doi.org/10.1088/1742-6596/1844/1/012003

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Von Hann, J. (1903). *Handbook of climatology*. MacMillan.

Wang, H., Ma, C., & Zhou, L. (2009). A brief review of machine learning and its application. *2009 International Conference on Information Engineering and Computer Science*, 1–4. https://doi.org/10.1109/ICIECS.2009.5362936

Wang, S.-C. (2003). Artificial neural network. In *Interdisciplinary computing in java programming* (pp. 81–100). Springer US. https://doi.org/10.1007/978-1-4615-0377-4_5

Warohma, A. M., Kurniasari, P., Dwijayanti, S., Irmawan & Suprapto, B. Y. (2018). Identification of regional dialects using mel frequency cepstral coefficients (mfccs) and neural network. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, iSemantic 2018*, 522–527. https://doi.org/10.1109/ISEMANTIC.2018.8549731

Yeh, S. L., Chao, G. Y., Su, B. H., Huang, Y. L., Lin, M. H., Tsai, Y. C., Tai, Y. W., Lu, Z. C., Chen, C. Y., Tai, T. M., Tseng, C. W., Lee, C. K., & Lee, C. C. (2019). Using attention networks and adversarial augmentation for styrian dialect continuous sleepiness and baby sound recognition. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2019-September*, 2398–2402. https://doi.org/10.21437/Interspeech.2019-2110

Zhang, Q., & Hansen, J. H. (2017). Dialect recognition based on unsupervised bottleneck features. *Interspeech*, 2576–2580.

Zhang, Q., & Hansen, J. H. (2018). Language/dialect recognition based on unsupervised deep learning. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *26*, 873–882. https://doi.org/10.1109/TASLP.2018.2797420

Zhou, W., Newsam, S., Li, C., & Shao, Z. (2017). Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval. *Remote Sensing*, *9*(5). https://doi.org/10.3390/rs9050489