UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR PUSTAKA

Aharoni, R. & Goldberg, Y., 2020, Unsupervised Domain Clusters in Pretrained Language Models, Dalam, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 7747–7763., *https://www.aclweb.org/anthology/2020.acl-main.692*,.

Akmaliah, W., 2022, Reclaiming Moderate Islam in Nahdlatul Ulama: Challenging the Dominant Religious Authority in Digital Platform, *Journal of Indonesian Islam*, 16, 1, 223. *https://jiis.uinsa.ac.id/index.php/JIIs/article/view/2680*,.

Ali, H., Purwandi, L., Nugroho, H., Halim, T. & Firdaus, K., 2019, *Indonesia Moslem Report 2019: The Challenges of Indonesia Moderate Moslems*, Jakarta, Indonesia. *www.alvara-strategic.com*,.

Aljalbout, E., Golkov, V., Siddiqui, Y., Strobel, M. & Cremers, D., 2018, Clustering with Deep Learning: Taxonomy and New Methods, *arXiv preprint arXiv:11801.07648*, 1–12. *http://arxiv.org/abs/1801.07648*,.

Alturki, A., Bchir, O. & Ismail, M.M. Ben, 2021, Joint Deep Clustering: Classification and Review, *International Journal of Advanced Computer Science and Applications*, 12, 10, 2021. *https://thesai.org/Downloads/Volume12No10/Paper_96-Joint_Deep_Clustering_Classification_and_Review.pdf*,.

AlZuhair, M.S., Ben Ismail, M.M. & Bchir, O., 2023, Soft Semi-Supervised Deep Learning-Based Clustering, *Applied Sciences*, 13, 17, 9673. *https://www.mdpi.com/2076-3417/13/17/9673*,.

Amirizadeh, E. & Boostani, R., 2021, CDEC: a constrained deep embedded clustering, *International Journal of Intelligent Computing and Cybernetics*, 14, 4, 686–701. *http://www.emerald.com/ijicc/article/14/4/686-701/135405*,.

Babić, K., Martinčić-Ipšić, S. & Meštrović, A., 2020, Survey of neural text representation models, *Information (Switzerland)*, 11, 11, 1–32. *https://www.mdpi.com/2078-2489/11/11/511*.

Basu, S., Davidson, I. & Wagstaff, K.L., 2008, *Constrained Clustering*, S. Basu, I. Davidson, & K. Wagstaff, ed., Chapman and Hall/CRC. *https://www.taylorfrancis.com/books/9781584889977*,.

Baumann, P. & Hochbaum, D.S., 2025, An Algorithm for Clustering with Confidence-Based Must-Link and Cannot-Link Constraints, *INFORMS Journal on Computing*, 37, 4, 1044–1068. *https://pubsonline.informs.org/doi/10.1287/ijoc.2023.0419*,.

Bengio, Y., Vincent, R.D.P. & Jauvin, C., 2003, A Neural Probabilistic Language Model, *Journal of Machine Learning Research*, 3, 1137–1155. *https://www.tandfonline.com/doi/full/10.1080/1536383X.2018.1448388*,.

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Blei, D.M., Ng, A.Y. & Jordan, M.I., 2003, Latent dirichlet allocation, *Journal of machine Learning research*, 3, Jan, 993–1022. *https://dl.acm.org/doi/epdf/10.5555/944919.944937*,.

Bogdanov, S., Constantin, A., Bernard, T., Crabbé, B. & Bernard, E.P., 2024, NuNER: Entity Recognition Encoder Pre-training via LLM-Annotated Data, Dalam, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 11829–11841., *https://aclanthology.org/2024.emnlp-main.660*,.

Bojanowski, P., Grave, E., Joulin, A. & Mikolov, T., 2017a, Enriching Word Vectors with Subword Information, *Transactions of the Association for Computational Linguistics*, 5, 135–146. *https://aclanthology.org/Q17-1010/*,.

Bojanowski, P., Grave, E., Joulin, A. & Mikolov, T., 2017b, Enriching Word Vectors with Subword Information, *Transactions of the Association for Computational Linguistics*, 5, 135–146. *https://aclanthology.org/Q17-1010/*,.

Bsoul, Q., Salam, R.A., Atwan, J. & Jawarneh, M., 2021, Arabic Text Clustering Methods and Suggested Solutions for Theme-Based Quran Clustering: Analysis of Literature, *Journal of Information Science Theory and Practice*, 9, 4, 15–34. *https://koreascience.or.kr/article/JAKO202100551544940.page*,.

Chefrour, A. & Souici-Meslati, L., 2022, Unsupervised Deep Learning: Taxonomy and Algorithms, *Informatica (Slovenia)*, 46, 2, 151–168. *https://www.proquest.com/openview/dfe0a9f1319753eac6cbf6c9904090d6/*,.

Chelba, C., Norouzi, M. & Bengio, S., 2017, N-gram Language Modeling using Recurrent Neural Network Estimation, *arXiv preprint arXiv:1703.10724*, , 2, 1–10. *https://arxiv.org/abs/1703.10724.*

Chen, G., 2015, Deep Learning with Nonparametric Clustering, *arXiv ppreprint arXiv:1501.03084*, abs/1501.0, 1–14. *https://arxiv.org/abs/1501.03084.*

Chen, T. Xu, R., He, Y., dan Wang, X., 2015, *A Gloss Composition and Context Clustering Based Distributed Word Sense Representation Model*, Entropy, 17, pp. 6007–6024. Doi: 10.3390/e17096007.

Chen, Z. & Ren, J., 2019a, Short Text Embedding for Clustering Based on Word and Topic Semantic Information, Dalam, *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, IEEE, hlm. 61–70., *https://ieeexplore.ieee.org/document/8964215/*,.

Chen, Z. & Ren, J., 2019b, Short Text Embedding for Clustering Based on Word and Topic Semantic Information, Dalam, *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, IEEE, hlm. 61–70., *https://ieeexplore.ieee.org/document/8964215/*,.

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K., 2019a, BERT: Pre-training of deep bidirectional transformers for language understanding, Dalam, *Proceedings of the 2019 Conference of the North American Chapter of the*

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 4171–4186., *https://aclanthology.org/N19-1423.pdf.*

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K., 2019b, BERT: Pre-training of deep bidirectional transformers for language understanding, Dalam, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1*, Association for Computational Linguistics, Minneapolis, Minnesota, hlm. 4171–4186., *https://aclanthology.org/N19-1423.pdf.*

Diallo, B., Hu, J., Li, T., Khan, G.A., Liang, X. & Zhao, Y., 2021, Deep Embedding Clustering Based On Contractive Autoencoder, *Neurocomputing*, 433, 96–107. Doi: 10.1016/j.neucom.2020.12.094.

Dieng, A.B., Gao, J., Wang, C. & Paisley, J., 2017, TopicRNN: A recurrent neural network with long-range semantic dependency, *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 1–13. *https://openreview.net/pdf?id=rJbbOLcex*.

Dizaji, K.G., Herandi, A., Deng, C., Cai, W. & Huang, H., 2017, Deep Clustering via Joint Convolutional Autoencoder Embedding and Relative Entropy Minimization, Dalam, *Proceedings of the IEEE International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers Inc., hlm. 5747–5756., h*ttps://ieeexplore.ieee.org/document/8237874/.*

Dodda, R. & Alladi, S.B., 2024, Enhancing Document Clustering with Hybrid Recurrent Neural Networks and Autoencoders: A Robust Approach for Effective Semantic Organization of Large Textual Datasets, *EAI Endorsed Transactions on Intelligent Systems and Machine Learning Applications*, 1. *https://publications.eai.eu/index.php/ismla/article/view/4564/3019*

Fakhruroji, M., 2019a, Muslims Learning Islam on the Internet, *Handbook of Contemporary Islam and Muslim Lives*, 1–17. *https://link.springer.com/rwe/10.1007/978-3-319-73653-2_70-1*.

Fan, Y., Gongshen, L., Kui, M. & Zhaoying, S., 2018a, Neural Feedback Text Clustering With BiLSTM-CNN-Kmeans, *IEEE Access*, 6, 57460–57469. *https://ieeexplore.ieee.org/document/8478272/,*.

Fan, Y., Gongshen, L., Kui, M. & Zhaoying, S., 2018b, Neural Feedback Text Clustering with BiLSTM-CNN-Kmeans, *IEEE Access*, 6, 57460–57469. *https://ieeexplore.ieee.org/document/8478272/,*.

Fard, Maziar Moradi, Thonet, T. & Gaussier, E., 2020a, Deep k-Means: Jointly clustering with k-Means and learning representations, *Pattern Recognition Letters*, 138, 185–192. *http://arxiv.org/abs/1806.10069*,.

Fard, Mazar Moradi, Thonet, T. & Gaussier, E., 2020, Seed-Guided Deep Document Clustering, Dalam, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

UNIVERSITAS GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Bioinformatics), Springer International Publishing, hlm. 3–16., *https://link.springer.com/10.1007/978-3-030-45439-5_1*,.

Fard, Maziar Moradi, Thonet, T. & Gaussier, E., 2020b, *Seed-Guided Deep Document Clustering*, Springer International Publishing. *https://link.springer.com/10.1007/978-3-030-45439-5_1*,.

Gajendran, S., D, M. & Sugumaran, V, 2020, *Character level and word level embedding with bidirectional LSTM – Dynamic recurrent neural network for biomedical named entity recognition from literature*, Journal of Biomedical Informatics, 112. Doi: 10.1016/j.jbi.2020.103609.

Goldberg, Y., 2017, *Neural Network Methods in Natural Language Processing*, edisi ke 37, G. Hirst, ed., Morgan & Claypool Publisher, Toronto. *https://link.springer.com/book/10.1007/978-3-031-02165-7.*

González-Almagro, G., Peralta, D., De Poorter, E., Cano, J.R. & García, S., 2025, Semi-supervised constrained clustering: an in-depth overview, ranked taxonomy and future research directions, *Artificial Intelligence Review*, 58, 5. *https://link.springer.com/article/10.1007/s10462-024-11103-8.*

Goodfellow, I., Bengio, Y. & Courville, A., 2016a, *Deep Learning*, MIT Press, Cambridge, MA, USA. *https://www.deeplearningbook.org/.*

Goodfellow, I., Bengio, Y. & Courville, A., 2016b, *Deep Learning*, MIT Press, Cambridge, MA, USA. *https://www.deeplearningbook.org/.*

Graves, A., Mohamed, A. & Hinton, G., 2013, Speech recognition with deep recurrent neural networks, Dalam, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, hlm. 6645–6649., *http://ieeexplore.ieee.org/document/6638947/*,.

Graves, A. & Schmidhuber, J., 2005, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, Dalam, *Neural Networks*, hlm. 602–610., doi: 10.1016/j.neunet.2005.06.042.

Guan, R., Zhang, H., Liang, Y., Giunchiglia, F., Huang, L. & Feng, X., 2020a, Deep Feature-Based Text Clustering and Its Explanation, *IEEE Transactions on Knowledge and Data Engineering*, 14, 8, 1–1. *https://ieeexplore.ieee.org/document/9215004.*

Guan, R., Zhang, H., Liang, Y., Giunchiglia, F., Huang, L. & Feng, X., 2020b, Deep Feature-Based Text Clustering and Its Explanation, *IEEE Transactions on Knowledge and Data Engineering*, 14, 8, 1–1. *https://ieeexplore.ieee.org/document/9215004.*

Guo, X., Gao, L., Liu, X. & Yin, J., 2017, Improved deep embedded clustering with local structure preservation, *IJCAI International Joint Conference on Artificial Intelligence*, 1753–1759. *https://dl.acm.org/doi/10.5555/3172077.3172131*.

Gusmita, R.H., Firmansyah, A.F., Moussallem, D. & Ngonga Ngomo, A.-C., 2023, IndQNER: Named Entity Recognition Benchmark Dataset from the Indonesian Translation of the Quran, Dalam, *Natural Language*

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Processing and Information Systems, Springer, hlm. 170–185., *https://link.springer.com/10.1007/978-3-031-35320-8_12*,.

Hadi, M.N., 2022, The Narrative Of Protecting Polygamous Women In Indonesia's Digital World: Between Moderate and Conservative Muslims, *Al-Ahwal: Jurnal Hukum Keluarga Islam*, 15, 2, 161–186. *https://ejournal.uin-suka.ac.id/syariah/Ahwal/article/view/2700*,.

Hadifar, A., Sterckx, L., Demeester, T. & Develder, C., 2019, A Self-Training Approach for Short Text Clustering, Dalam, *Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019)*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 194–199., *https://www.aclweb.org/anthology/W19-4322*,.

Han, J., Kamber, M. & Pei, J., 2012, *Data mining concepts and techniques, third edition*, edisi ke 3rd, Morgan Kaufmann Publishers, Waltham, Mass. *https://www.sciencedirect.com/book/monograph/9780123814791/data-mining-concepts-and-techniques.*

Hasyim Asy'ari, A. & Hanafi, M., 2025, Pengembangan Model Klasterisasi Topik Hadis Bukhari Muslim Menggunakan BERT dengan Penambahan Fitur Semantik, *Journal Computer Science*, 4, 2. Doi: 10.31294/ijcs.v4i2.8931,.

Hazratgholizadeh, R., Balafar, M.A. & Derakhshi, M.R.F., 2022, Active constrained deep embedded clustering with dual source, *Applied Intelligence*, 53. *https://link.springer.com/10.1007/s10489-022-03752-5*,.

Hochreiter, S. & Schmidhuber, J., 1997, Long Short-Term Memory, *Neural Computation*, 9, 8, 1735–1780. *https://direct.mit.edu/neco/article/9/8/1735-1780/6109*,.

Honnibal, M., Montani, I., Van Landeghem, S. & Boyd, A., 2020, spaCy: Industrial-strength Natural Language Processing in Python, *https://spacy.io/*,.

Huang, P., Huang, Y., Wang, W. & Wang, L., 2014, Deep embedding network for clustering, *Proceedings - International Conference on Pattern Recognition*, 1532–1537. *https://ieeexplore.ieee.org/document/6976982.*

Huang, S., Kang, Z., Xu, Z. & Liu, Q., 2021, Robust deep k-means: An effective and simple method for data clustering, *Pattern Recognition*, 117, 107996. *https://doi.org/10.1016/j.patcog.2021.107996*,.

Huda, S., Maulana Mas'udi, M. & Muthohirin, N., 2022, The Rise of Muhammadiyah's Islamic Da'wah in the Contemporary Era: Transformation to Online Trend and Responses to Islamic Moderation, *Progresiva : Jurnal Pemikiran dan Pendidikan Islam*, 11, 01, 1–24. Doi: 10.22219/progresiva.v11i01.20889

Ienco, D. & Interdonato, R., 2023, Deep semi-supervised clustering for multi-variate time-series, *Neurocomputing*, 516, 36–47. *https://linkinghub.elsevier.com/retrieve/pii/S092523122201308X*,.

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Iswahyudi, I., Safala, U. & Agustina, D.A., 2021, Deradicalization Of Religious Understanding Through The NU Online and Suara Muhammadiyah Website in The Perspective of Critical Discourse Analysis, *Dialogia*, 19, 1, 1–31. *https://jurnal.iainponorogo.ac.id/index.php/dialogia/article/view/2782*,.

Jain, S., Jain, A.K. & Singh, S.P., 2021a, Building a Machine Learning Model for Unstructured Text Classification: Towards Hybrid Approach, Dalam, *Advances in Intelligent Systems and Computing*, hlm. 447–454., h*ttps://link.springer.com/chapter/10.1007/978-981-15-6014-9_51.*

Jain, S., Jain, A.K. & Singh, S.P., 2021b, Building a Machine Learning Model for Unstructured Text Classification: Towards Hybrid Approach, Dalam, *Advances in Intelligent Systems and Computing*, hlm. 447–454., *ttps://link.springer.com/chapter/10.1007/978-981-15-6014-9_51.*

Jiang, Z., Zheng, Y., Tan, H., Tang, B. & Zhou, H., 2017a, Variational deep embedding: An unsupervised generative approach to Clustering, Dalam, *IJCAI International Joint Conference on Artificial Intelligence*, hlm. 1965–1972., *https://dl.acm.org/doi/10.5555/3172077.3172161*.

Jiang, Z., Zheng, Y., Tan, H., Tang, B. & Zhou, H., 2017b, Variational deep embedding: An unsupervised generative approach to Clustering, Dalam, *IJCAI International Joint Conference on Artificial Intelligence*, hlm. 1965–1972., *https://dl.acm.org/doi/10.5555/3172077.3172161*.

Joshi, A. V, 2023, *Machine Learning and Artificial Intelligence 2nd Edition*, Springer. *https://link.springer.com/book/10.1007/978-3-031-12282-8*.

Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N. & Wu, Y., 2016, Exploring the Limits of Language Modeling, *arXiv preprint arXiv:1602.02410v2*. *https://arxiv.org/pdf/1602.02410*.

Jurafsky, D. & Martin, J.H., 2026, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, 3rd ed. draft, *https://web.stanford.edu/~jurafsky/slp3/*,.

Kim, J., Yoon, J., Park, E. & Choi, S., 2020a, Patent document clustering with deep embeddings, *Scientometrics*, 123, 2, 563–577. *https://doi.org/10.1007/s11192-020-03396-7*,.

Kim, J., Yoon, J., Park, E. & Choi, S., 2020b, Patent document clustering with deep embeddings, *Scientometrics*, 123, 2, 563–577. *https://doi.org/10.1007/s11192-020-03396-7*,.

Kim, Y., 2014, Convolutional Neural Networks for Sentence Classification, Dalam, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA, Qatar, hlm. 1746–1751., *http://aclweb.org/anthology/D14-1181*,.

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Kim, Y., Jernite, Y., Sontag, D. & Rush, A.M., 2016a, Character-Aware neural language models, *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, 2741–2749. *https://dl.acm.org/doi/10.5555/3016100.3016285*.

Koehn, P. and Knowles, R., 2017, *Six Challenges for Neural Machine Translation*, in Proceedings of the First Workshop on Neural Machine Translation. Association for Computational Linguistics, pp. 28–39. Doi: 10.18653/v1/W17-3204.

Kim, Y., Jernite, Y., Sontag, D. & Rush, A.M., 2016b, Character-Aware neural language models, *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, 2741–2749. https://dl.acm.org/doi/10.5555/3016100.3016285.

Koto, F., Rahimi, A., Lau, J.H. & Baldwin, T., 2020a, IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP, Dalam, *The 28th International Conference on Computational Linguistics*, International Committee on Computational Linguistics, Barcelona, Spain (Online), hlm. 757–770., *https://aclanthology.org/2020.coling-main.66/*.

Koto, F., Rahimi, A., Lau, J.H. & Baldwin, T., 2020b, IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP, Dalam, *The 28th International Conference on Computational Linguistics (COLING)*, International Committee on Computational Linguistics, Barcelona, Spain (Online), hlm. 757–770., *http://arxiv.org/abs/2011.00677*,.

Kushwaha, N. & Pant, M., 2020, Textual data dimensionality reduction - a deep learning approach, *Multimedia Tools and Applications*, 79, 15–16, 11039–11050. *https://link.springer.com/article/10.1007/s11042-018-6900-x*.

Lai, S., Xu, L., Liu, K. & Zhao, J., 2015a, Recurrent convolutional neural networks for text classification, *Proceedings of the National Conference on Artificial Intelligence*, 3, 2267–2273. *https://dl.acm.org/doi/10.5555/2886521.2886636*.

Lai, S., Xu, L., Liu, K. & Zhao, J., 2015b, Recurrent convolutional neural networks for text classification, *Proceedings of the National Conference on Artificial Intelligence*, 3, 2267–2273. *https://dl.acm.org/doi/10.5555/2886521.2886636*.

Landis, J.R. & Koch, G.G., 1977, The Measurement of Observer Agreement for Categorical Data, *Biometrics*, 33, 1, 159. *https://www.jstor.org/stable/2529310*.

Lanin, I., Hardiyanto, R. & Purnama, A., 2014, Kateglo API, *http://kateglo.lostfocus.org/api.php*,.

Le, Q. & Mikolov, T., 2014, Distributed Representations of Sentences and Documents, Dalam, E. P. Xing & T. Jebara, ed. *Proceedings of the 31st International Conference on Machine Learning*, PMLR, Bejing, China, hlm. 1188–1196., *https://proceedings.mlr.press/v32/le14.html.*

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Li, Y., Cai, J. & Wang, J., 2020a, A Text Document Clustering Method Based on Weighted BERT Model, Dalam, *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, IEEE, hlm. 1426–1430., *https://ieeexplore.ieee.org/document/9085059*.

Li, Y., Cai, J. & Wang, J., 2020b, A Text Document Clustering Method Based on Weighted BERT Model, Dalam, *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, IEEE, hlm. 1426–1430., *https://ieeexplore.ieee.org/document/9085059/*,.

Liu, X., Qin, Y., Xu, Q., Liu, Z., Guo, X. & Xu, W., 2025, Integrating Knowledge Graph Reasoning with Pretrained Language Models for Structured Anomaly Detection, Dalam, *Proceedings - 2025 International Conference on Artificial Intelligence, Human-Computer Interaction and Natural Language Processing, ICAHN 2025*, Institute of Electrical and Electronics Engineers Inc., hlm. 107–111., *https://ieeexplore.ieee.org/document/11216762*.

Lu, Y., Li, H., Li, Y., Lin, Y. & Peng, X., 2024, *A Survey On Deep Clustering: From The Prior Perspective*, Vicinagearth, 1, 1, no. 4. doi: 10.1007/s44336-024-00001-w.

Lu, Z. & Leen, T.K., 2007, Semi-supervised Clustering with Pairwise Constraints: A Discriminative Approach, Dalam, *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*, hlm. 299–306., *https://proceedings.mlr.press/v2/lu07a/lu07a.pdf*,.

Mayani, S. & Swarndeep, S., 2020, A Novel Approach of Text Document Clustering by using Clustering Techniques, *International Research Journal of Engineering and Technology*, , June, 4153–4158. *https://www.irjet.net/archives/V7/i6/IRJET-V7I6774.pdf.*

Mehta, V., Agarwal, M. & Kaliyar, R.K., 2024, A comprehensive and analytical review of text clustering techniques, *International Journal of Data Science and Analytics*, 18, 3, 239–258. *https://link.springer.com/article/10.1007/s41060-024-00540-x*.

Mehta, V., Bawa, S. & Singh, J., 2021, WEClustering: word embeddings based text clustering technique for large datasets, *Complex & Intelligent Systems*, 7, 3211–3224. *https://link.springer.com/article/10.1007/s40747-021-00512-9.*

Mikolov, T., 2012, *Statistical Language Models Based on Neural Networks*,. BRNO University of Technology, *https://www.fit.vut.cz/study/phd-thesis/283/.en*.

Mikolov, T., Chen, K., Corrado, G. & Dean, J., 2013, Efficient Estimation of Word Representations in Vector Space, *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, 1–12. *https://arxiv.org/pdf/1301.3781*,.

Mikolov, T., Karafiát, M., Burget, L., Černocký, J. & Khudanpur, S., 2010a, Recurrent neural network based language model, Dalam, *Interspeech 2010*,

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ISCA, ISCA, hlm. 1045–1048., *https://www.isca-archive.org/interspeech_2010/mikolov10_interspeech.html*.

Mikolov, T., Karafiát, M., Burget, L., Černocký, J. & Khudanpur, S., 2010b, Recurrent neural network based language model, Dalam, *Interspeech 2010*, ISCA, ISCA, hlm. 1045–1048., *https://www.isca-archive.org/interspeech_2010/mikolov10_interspeech.html*.

Min, E., Guo, X., Liu, Q., Zhang, G., Cui, J. & Long, J., 2018, A Survey of Clustering with Deep Learning: From the Perspective of Network Architecture, *IEEE Access*, 6, 39501–39514. *https://ieeexplore.ieee.org/document/8412085*.

Miyamoto, Y. and Cho, K., 2016, *Gated Word-Character Recurrent Language Model*, in Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, pp. 1992–1997, Doi: 10.18653/v1/D16-1209.

Mohammadi, M., Mundra, R., Socher, R., Wang, L. & Kamath, A., 2019, Language Models, RNN, GRU and LSTM, *Lecture Notes for CS224n: Natural Language Processing with Deep Learning in Standford University*. *http://web.stanford.edu/class/cs224n/*.

Mujibuddin, M. & Riza, M.F., 2022, Internet sebagai Ruang Baru Kontestasi Diskursus Agama dan Politik: Kajian atas Diskursus Islam, Demokrasi dan Pancasila pada Website Keislaman Populer di Indonesia, *JISPO Jurnal Ilmu Sosial dan Ilmu Politik*, 11, 2, 331–358. *https://journal.uinsgd.ac.id/index.php/jispo/article/view/14148*,.

Naf'an, M.Z., Sari, Y. & Suyanto, Y., 2021, Word Embeddings Evaluation on Indonesian Translation of Al-Quran and Hadiths, *IOP Conference Series: Materials Science and Engineering*, 1077. *https://iopscience.iop.org/article/10.1088/1757-899X/1077/1/012025*.

National Library of the Republic of Indonesia, 2017, List of Islamic Subject Headings and Islamic Classification (in Indonesian), *https://bintangpusnas.perpusnas.go.id/konten/BKA82YH/daftar-tajuk-subjek-islam-dan-klasifikasi-islam*,.

Niu, L., Dai, X., Zhang, J. & Chen, J., 2016, Topic2Vec: Learning distributed representations of topics, *Proceedings of 2015 International Conference on Asian Language Processing, IALP 2015*, 193–196. *https://ieeexplore.ieee.org/document/7451564/*.

Noaman, H.M., Sarhan, S.S. & Rashwan, M.A.A., 2018, Enhancing recurrent neural network-based language models by word tokenization, *Human-centric Computing and Information Sciences*, 8, 1, 1–13. *https://link.springer.com/article/10.1186/s13673-018-0133-x*.

Park, J., Park, C., Kim, J., Cho, M. & Park, S., 2019a, ADC: Advanced document clustering using contextualized representations, *Expert Systems with*

UNIVERSITAS GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Applications, 137, 157–166. *https://www.sciencedirect.com/science/article/pii/S0957417419304762*.

Park, J., Park, C., Kim, J., Cho, M. & Park, S., 2019b, ADC: Advanced document clustering using contextualized representations, *Expert Systems with Applications*, 137, 157–166. *https://www.sciencedirect.com/science/article/pii/S0957417419304762*.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J. & Chintala, S., 2019, PyTorch: An Imperative Style, High-Performance Deep Learning Library, *http://arxiv.org/abs/1912.01703*,.

Putra, D.I.A., Sayuti, S. & Rahim, R.A.A., 2024, Inconsistency of Salafi's Bid'ah Concept, *Al-Risalah: Forum Kajian Hukum dan Sosial Kemasyarakatan*, 24, 1, 116–134. *https://shariajournals-uinjambi.ac.id/index.php/al-risalah/article/view/1523*,.

RedaksiIB, 2020, 100 Situs Islam Indonesia, NU Online Peringkat Pertama, *https://ibtimes.id/100-situs-islam-indonesia-nu-online-peringkat-pertama/*, diakses 26 April 2022.

Řehůřek, R. & Sojka, P., 2010, Software Framework for Topic Modelling with Large Corpora, Dalam, *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, ELRA, Valletta, Malta, hlm. 45–50., *http://is.muni.cz/publication/884893/en*,.

Reinsel, D., Gantz, J. & Rydning, J., 2018, *The Digitization of the World From Edge to Core*, *https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf*,.

Ren, L., Qin, Y., Chen, Y., Bai, R., Xue, J. & Huang, R., 2023, Deep structural enhanced network for document clustering, *Applied Intelligence*, 53, 10, 12163–12178. *https://link.springer.com/10.1007/s10489-022-04112-z*,.

Ren, Y., Pu, J., Yang, Z., Xu, J., Li, G., Pu, X., Yu, P.S. & He, L., 2024, Deep Clustering: A Comprehensive Survey, *IEEE Transactions on Neural Networks and Learning Systems*, 1–21. *https://ieeexplore.ieee.org/document/10585323/*.

Ren, Y., Hu, K., Dai, X., Pan, L., Hoi, S.C.H. & Xu, Z., 2019, Semi-supervised deep embedded clustering, *Neurocomputing*, 325, 121–130. *https://www.sciencedirect.com/science/article/abs/pii/S0925231218312049.*

Robbani, H.A., 2016, Sastrawi, *https://pypi.org/project/Sastrawi*,.

Rousseeuw, P.J., 1987, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, 20, C, 53–65. *https://www.sciencedirect.com/science/article/pii/0377042787901257.*

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Schuster, M. & Paliwal, K.K., 1997, Bidirectional recurrent neural networks, *IEEE Transactions on Signal Processing*, 45, 11, 2673–2681. *http://ieeexplore.ieee.org/document/650093/*,.

Shah, S.A. & Koltun, V., 2018, Deep Continuous Clustering, *arXiv:1803.01449v1*. *https://arxiv.org/abs/1803.01449*.

Shi, H. & Wang, C., 2021, Self-supervised Document Clustering Based on BERT with Data Augment, *arXiv:2011.08523*. *http://arxiv.org/abs/2011.08523*,.

Slamet, C., Rahman, A., Ramdhani, M.A. & Dharmalaksana, W., 2016, Clustering the Verses of the Holy Qur'an using K-Means Algorithm, *Asian Journal of Information Technology*, 15, 24, 5159–5162. *https://digilib.uinsgd.ac.id/5117/*.

Sundermeyer, M., Ney, H. & Schluter, R., 2015, From feedforward to recurrent LSTM neural networks for language modeling, *IEEE Transactions on Audio, Speech and Language Processing*, 23, 3, 517–529. *https://ieeexplore.ieee.org/document/7050391*.

Sundermeyer, M., Schlüter, R. & Ney, H., 2012, LSTM neural networks for language modeling, Dalam, *Interspeech 2012*, ISCA, ISCA, hlm. 194–197., doi: 10.21437/Interspeech.2012-65.

Suryana, A., 2023, *Indonesia's moderate Muslim websites and their fight against online Islamic extremism*, ISEAS-Yusof Ishak Institute. *https://www.iseas.edu.sg/wp-content/uploads/2023/09/TRS15_23.pdf*,.

Tang, W., Wen, X. & Hu, Z., 2024, Named Entity Recognition for Crop Diseases and Pests Based on Gated Fusion Unit and Manhattan Attention, *Agriculture (Switzerland)*, 14, 9. *https://www.mdpi.com/2077-0472/14/9/1565*.

Uriawan, W., Priyajie, A.A., Gustian, A., Hidayat, F.N., Rafiudin, S.A. & Zaelani, M.F., 2025, Unsupervised Thematic Clustering Of hadith Texts Using The Apriori Algorithm, *arXiv:2512.16694*. *http://arxiv.org/abs/2512.16694*,.

Vahidnia, S., Abbasi, A. & Abbass, H.A., 2020a, Document Clustering And Labeling For Research Trend Extraction And Evolution Mapping, *CEUR Workshop Proceedings*, 2658, 54–62. *https://ceur-ws.org/Vol-2658/paper7.pdf*.

Vahidnia, S., Abbasi, A. & Abbass, H.A., 2020b, Document Clustering And Labeling For Research Trend Extraction And Evolution Mapping, *CEUR Workshop Proceedings*, 2658, 54–62. *https://ceur-ws.org/Vol-2658/paper7.pdf*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. & Polosukhin, I., 2017, Attention Is All You Need, Dalam, *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, hlm. 6000–6010., *https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf*.

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Wagstaff, K., Cardie, C., Rogers, S. & Schrödl, S., 2001, Constrained K-means Clustering with Background Knowledge, Dalam, : *Proceedings of the Eighteenth International Conference on Machine Learning*, hlm. 577–584., *https://dl.acm.org/doi/10.5555/645530.655669*.

Wang, B., Wang, A., Chen, F., Wang, Y. & Kuo, C.C.J., 2019, Evaluating word embedding models: Methods and experimental results, *APSIPA Transactions on Signal and Information Processing*, 8, doi:10.1017/ATSIP.2019.12.

Wang, Y., Zou, J., Wang, K., Liu, C. & Yuan, X., 2023, Semi-supervised deep embedded clustering with pairwise constraints and subset allocation, *Neural Networks*, 164, 310–322. Doi: 10.1016/j.neunet.2023.04.016.

Wang, Z., 2018, N-gram and LSTM based Language Models, Dalam, *ABCs 2018 - 1st ANU Bio-inspired Computing conference*, Canberra, Australia., *https://users.cecs.anu.edu.au/~Tom.Gedeon/conf/ABCs2018/paper/ABCs2018_paper_16.pdf*.

Wang, Z., Mi, H. & Ittycheriah, A., 2016, Semi-supervised clustering for short text via deep representation learning, Dalam, *CoNLL 2016 - 20th SIGNLL Conference on Computational Natural Language Learning, Proceedings*, Association for Computational Linguistics, Berlin, Germany, hlm. 31–39., *https://aclanthology.org/K16-1004.pdf*.

Xie, J., Girshick, R. & Farhadi, A., 2016a, Unsupervised Deep Embedding for Clustering Analysis, *33rd International Conference on Machine Learning, ICML 2016*, 1, 740–749. *https://dl.acm.org/doi/10.5555/3045390.3045442.*

Xie, J., Girshick, R. & Farhadi, A., 2016b, Unsupervised deep embedding for clustering analysis, *33rd International Conference on Machine Learning, ICML 2016*, 1, 740–749. *https://dl.acm.org/doi/10.5555/3045390.3045442.*

Xu, J., Wang, P., Tian, G., Xu, B., Zhao, J., Wang, F. & Hao, H., 2015a, Short Text Clustering via Convolutional Neural Networks, Dalam, *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 62–69., *http://aclweb.org/anthology/W15-1509,*.

Xu, J., Wang, P., Tian, G., Xu, B., Zhao, J., Wang, F. & Hao, H., 2015b, Short Text Clustering via Convolutional Neural Networks, Dalam, *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 62–69., *http://aclweb.org/anthology/W15-1509.*

Xu, Q., Gu, H. dan Ji, S., 2024, *Text clustering based on pre-trained models and autoencoders*, Frontiers in Computational Neuroscience, 17, *https://doi.org/10.3389/fncom.2023.1334436*.

Yang, B., Fu, X., Sidiropoulos, N.D. & Hong, M., 2017, Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering, Dalam, *International Conference on Machine Learning*, JMLR.org, Sydney, hlm. 3861–3870., *https://proceedings.mlr.press/v70/yang17b/yang17b.pdf.*

UNIVERSITAS
GADJAH MADA
Arsitektur Deep Constrained Clustering Berbasis Dual-Stream Rnn untuk Closed-Domain Teks Keislaman
Berbahasa Indonesia
Muhammad Zidny Naf'an, Prof. Drs. Edi Winarko, M.Sc., Ph.D.; Dr. Sigit Priyanta, S.Kom., M.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Yang, C., Cao, B. & Fan, J., 2024, HS-GC: Holistic Semantic Embedding and Global Contrast for Effective Text Clustering, Dalam, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, ELRA and ICCL, hlm. 8349–8359., *https://aclanthology.org/2024.lrec-main.732/*,.

Yang, Z., Hu, W., Tan, Y. & Luo, Z., 2023, Improved Training of Deep Text Clustering, Dalam, *Findings of the Association for Computational Linguistics: EMNLP 2023*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 2490–2499., *https://aclanthology.org/2023.findings-emnlp.163*,.

Zafarani-Moattar, E., Kangavari, M.R. & Rahmani, A.M., 2024, Neural Network Meaningful Learning Theory and its Application for Deep Text Clustering, *IEEE Access*, 12, 42411–42422, doi: 10.1109/ACCESS.2024.3375754.

Zaki, M.J. & Meira, Jr, W., 2020, *Data Mining and Machine Learning: Fundamental Concepts and Algorithms, 2nd Edition*, Cambridge University Press, Cambridge, England. *https://dataminingbook.info/book_html/*,.

Zaratiana, U., Tomeh, N., Holat, P. & Charnois, T., 2024, GLiNER: Generalist Model for Named Entity Recognition using Bidirectional Transformer, Dalam, *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, Association for Computational Linguistics, Stroudsburg, PA, USA, hlm. 5364–5376., *https://aclanthology.org/2024.naacl-long.300*,.

Zhang, H., Zhan, T., Basu, S. & Davidson, I., 2021, A framework for deep constrained clustering, *Data Mining and Knowledge Discovery*, 35, 2, 593–620. *http://link.springer.com/10.1007/s10618-020-00734-4*,.

Zhang, Y., Lu, J., Liu, F., Liu, Q., Porter, A., Chen, H. & Zhang, G., 2018, Does deep learning help topic extraction? A kernel k-means clustering method with word embedding, *Journal of Informetrics*, 12, 4, 1099–1117, doi: 10.1016/j.joi.2018.09.004.

Zhu, P., Lang, Q. & Liu, X., 2023, Word Embedding of Dimensionality Reduction for Document Clustering, Dalam, *Proceedings of the 35th Chinese Control and Decision Conference, CCDC 2023*, Institute of Electrical and Electronics Engineers Inc., hlm. 4371–4376., doi: 10.1109/CCDC58219.2023.10327354.