











TABLE VIII  
SIMILAR WORD SEARCH RESULTS

word	Similar word search results
<b>Disobedience</b>	World, disobedience, minority, disobedience, resistance, war, ownership, death penalty, argument, calling
<b>Philosopher shun</b>	Private Interest, Durure, Poverty, Marx, Orthopedic, Puda, Yoonsa, Death, Officials, Farmers
<b>democracy</b>	Democracy, Private, Common, Upper Class, Absolute, World, Justice, Hangout, Legal, USA
<b>Marx</b>	Spinoza, Jeong Yak-yong, Controversy, Nivoir, Jaspers, Sunja, Plato, Non-interference, Nationalism, Mencius
<b>Hobbs</b>	Reverse discrimination, SeonSeonSeol, Beopbo, First, Justice, Resistance, Mani, Substitution, SeongakSeol, Manjangil
<b>female</b>	Rescue, counterpart, incline, have, male, manual labor, servant, ethnicity, husband, market failure

Table IX is the result of the search quality of the question and answer search system implemented with the foundation of Doc2Vec. The result of Q&A Search evaluation with 30 pre-documented questions and answer queries, as shown in Table IX, presented that similar question and answer search average ranking 1st resulted in 67.9%, with a 60% probability of a very similar Q&A that ranked 1st place. The probability of a very similar question and answer searched within the top three of the search results was 70%.

TABLE IX  
QUALITY OF DOC2VEC-BASED Q&A SEARCH SYSTEM

Criteria	Average of Similarity of first results	Probability of very similar Q&A searched as first result	Probability of very similar Q&A searched within top three search results
<b>Score (Average)</b>	67.9%	60%	70%

#### IV. CONCLUSION

Q&A in e-learning is an important interactive tool that can motivate learners and enhance their learning effectiveness. However, in the current process where the educator identifies a learner's question and writes a response, limitations in immediate feedback do occur [16],[17]. This study aims to provide learners with immediate feedbacks on their queries based on word embedding algorithm technology: Doc2Vec, processing unsupervised machine learning, implementing questions and answer search system with database of previous queries.

A question and answer search based on the generated Doc2Vec model produced a 60% probability of providing a very similar answer with the first search result 70% for the top three search results. What was only available with expensive search engines, can now be simply applied with the results of this study in document similarity searches and if reflected in actual services, there is a 70% chance that students will be supplied with adequate answers to their queries.

If you want to provide real-time answers to the current question-and-answer system, you can use Doc2Vec algorithm to provide services with sufficiently good performance and quality. The use of big data and artificial intelligence has become more active in various fields. Natural language parsing is available in analyzing various unstructured data required by e-learning [18]-[20]. The study anticipates using

the results in analysing various unstructured data in e-learning field, which will further develop into an artificial intelligence question and answer service based on knowledge. Recently, the importance of non-face-to-face education is increasing worldwide due to covid-19. It is expected that this study will contribute to enhancing the learning satisfaction and enhancing the learning effect by resolving the curiosity of learners in real-time in increasing face-to-face education.

#### REFERENCES

- [1] National IT Industry Promotion Agency(NIPA), *Investigation on status of e-learning Industry*, Ministry of Trade and Industry(MTI), 2018.
- [2] T.H. Kang, *Case Studies on EBS CSAT*, Educational Broadcasting System, 2016.
- [3] LAFLEN, Angela, SMITH, and Michelle, "Responding to student writing online: Tracking student interactions with instructor feedback in a Learning Management System," *Assessing Writing*, vol 31, pp. 39-52, 2017.
- [4] W. Chang, Z. Xu, S. Zhou, Shenghan, and W. Cao, "Research on detection methods based on Doc2vec abnormal comments," *Future generations computer systems*, vol. 86, pp. 656-662, 2018
- [5] D.H. Kim, D.S. Seo, S. Y. Cho, and P. S. Kang, "Multi-co-training for document classification using various document representations: TF-IDF, LDA, and Doc2Vec," *Information sciences*, vol. 477 pp. 15-29, 2019.
- [6] B. Pan, C. C. Yu, O. C. Zhang, S. X. Xu, and S. Cao, "The Improved Model for word2vec Based on Part of Speech and Word Order," *ACTA ELECTRONICA SINICA* vol. 46, pp 1976-1982, 2018.
- [7] V. A.Nrusimha, W. Hannah, H. J. David, K. Matcheri, and T.J. Blake, "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape," *Canadian journal of psychiatry*, vol. 64, pp. 456-464, 2019.
- [8] M. Tomas, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv*, 2013.
- [9] M. Mudasar, J. Rafiyya, and S. Muzaffar, "Text document summarization using word embedding," *Expert systems with applications*, vol. 143, 2020.
- [10] (2017) Python Korean NLP website. [online]. Available: <https://replet.tistory.com/58>
- [11] (2015) KoNLPy website. [online]. Available: <https://konlpyko.readthedocs.io/ko/v0.4.3/morph/#comparison-between-pos-tagging-classes>
- [12] J. Yaser, A. A. Mahmoud, and B. Elhadj, "Advanced Arabic Natural Language Processing (ANLP) and its applications: Introduction to the special issue," *Information processing & management*, vol. 56, pp. 259-261, 2019.
- [13] C. Jingqiang, Z. Hai, "Extractive summarization of documents with images based on multi-modal RNN," *Future generations computer systems*, vol. 99, pp. 186-196, 2019.
- [14] Y. Cheng, Z. Ye, M. Wang, and Q. Zhang, "Document classification based on convolutional neural network and hierarchical attention network", *Neural Network World*, vol. 29, pp. 83-98, 2019.
- [15] F. Yang, F. Lidan, "Ontology semantic integration based on convolutional neural network," *Neural Computing And Applications*, vol. 31, pp. 8253-8266, 2019.
- [16] I. Alsmadi, H. G. Keng, "Term weighting scheme for short-text classification: Twitter corpuses," *Neural Computing And Applications*, vol. 31, pp. 3819-3831, 2019.
- [17] S. N. Bhushan, A. Danti, "Classification of text documents based on score level fusion approach," *Pattern recognition letters*, vol. 94, pp. 118-126, 2017.
- [18] M. Chowkwanyun, "Big Data, Large-Scale Text Analysis, and Public Health Research," *American journal of public health*, vol. 109, pp. S126-S127, 2019.
- [19] Y. Q. Song, U. Shyam, P. Haoruo, M. Stephen, and R. Dan. "Toward any-language zero-shot topic classification of textual documents" *Artificial intelligence*, vol. 274, pp. 133-150, 2019.
- [20] K. Hu, H. Wu, K. Qi, J. Yu, S. Yang, T. Yu, J. Zheng, and B. Liu, "A domain keyword analysis approach extending Term Frequency-Keyword Active Index with Google Word2Vec model," *Scientometrics*, vol. 114, pp. 1031-1068, 2018.