|  |  |  |  |
| --- | --- | --- | --- |
| Title | Study of Keyword Extraction Techniques for Electric Double-Layer Capacitor Domain Using Text Similarity Indexes: An Experimental Analysis. | | |
| Author(s) Name | Miah, M. S. U.; Sulaiman, J.; Sarwar, T. B.; Zamli, K. Z.; and Jose, R | | |
| Contact Email(s) | saef@aiub.edu | | |
| Published Journal Name | Complexity | | |
| Type of Publication | journal | | |
| Volume |  | Issue |  |
| Publisher | Hindawai | | |
| Publication Date | Dec 2021 | | |
| ISSN |  | | |
| DOI | http://doi.org/10.1155/2021/8192320 | | |
| URL |  | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| Keywords perform a significant role in selecting various topic-related documents quite easily. Topics or keywords assigned by humans or experts provide accurate information. However, this practice is quite expensive in terms of resources and time management. Hence, it is more satisfying to utilize automated keyword extraction techniques. Nevertheless, before beginning the automated process, it is necessary to check and confirm how similar expert-provided and algorithm-generated keywords are. This paper presents an experimental analysis of similarity scores of keywords generated by different supervised and unsupervised automated keyword extraction algorithms with expert-provided keywords from the electric double layer capacitor (EDLC) domain. The paper also analyses which texts provide better keywords such as positive sentences or all sentences of the document. From the unsupervised algorithms, YAKE, TopicRank, MultipartiteRank, and KPMiner are employed for keyword extraction. From the supervised algorithms, KEA and WINGNUS are employed for keyword extraction. To assess the similarity of the extracted keywords with expert-provided keywords, Jaccard, Cosine, and Cosine with word vector similarity indexes are employed in this study. The experiment shows that the MultipartiteRank keyword extraction technique measured with cosine with word vector similarity index produces the best result with 92% similarity with expert-provided keywords. This study can help the NLP researchers working with the EDLC domain or recommender systems to select more suitable keyword extraction and similarity index calculation techniques. | |