

Title: PAKE – PoS Tagger Augmented Keyword Extraction

Author: Md Saidur Rahman Kohinoor; Ayesha Loylus Miah; Shah Fayez Ali; Md Sabir Hossain; Md. Hasan Imam Bijoy; Shadman Sakib

Abstract: The state-of-the-art techniques for automatic keyword extraction majorly deal with the collection of long documents. However, for several reasons, these do not provide satisfactory results for shorter lengths of documents. Moreover, with the ever-increasing amounts of information available, a keyword extraction system that automatically deals with varying lengths of text can lessen the workload and make the entire process of manually assigning the keywords less time-consuming. For this, the widely used Natural Language Processing (NLP) techniques are examined in the context of extensive data. Therefore, this research introduces PAKE - PoStagger Augmented Keyword Extraction system as a practical amalgamation of statistical and textual-based features based on an unsupervised key phrase extracting algorithm to stand out as a suitable alternative to the existing solutions. The effectiveness is demonstrated by comparing it with six state-of-the-art unsupervised methods, and the results are illustrated using four datasets.

Keywords: NLP, Keyword Extraction, Unsupervised Method, Pos Tagger

DOI: <https://doi.org/10.1109/COMPAS60761.2024.10796100>