

ChatGPT vs state-of-the-art models: a benchmarking study in keyphrase generation task

A.J. López López; J. Portela González; R. Martínez Cruz

Abstract-

Transformer-based language models, including ChatGPT, have demonstrated exceptional performance in various natural language generation tasks. However, there has been limited research evaluating ChatGPT's keyphrase generation ability, which involves identifying informative phrases that accurately reflect a document's content. This study seeks to address this gap by comparing ChatGPT's keyphrase generation performance with state-of-the-art models, while also testing its potential as a solution for two significant challenges in the field: domain adaptation and keyphrase generation from long documents. We conducted experiments on eight publicly available datasets spanning scientific, news, and biomedical domains, analyzing performance across both short and long documents. Our results show that ChatGPT outperforms current state-of-the-art models in all tested datasets and environments, generating high-quality keyphrases that adapt well to diverse domains and document lengths.

Index Terms- ChatGPT · Text generation · Keyphrase generation · Natural language processing · Deep learning · Domain adaptation · Long documents · Large language models

Due to copyright restriction we cannot distribute this content on the web. However, clicking on the next link, authors will be able to distribute to you the full version of the paper:

[Request full paper to the authors](#)

If your institution has an electronic subscription to Applied Intelligence, you can download the paper from the journal website:

[Access to the Journal website](#)

Citation:

López López, A.J.; Martínez-Cruz, R.; Portela, J. "ChatGPT vs state-of-the-art models: a benchmarking study in keyphrase generation task", *Applied Intelligence*, vol.55, no.1, pp.50-1-50-25, January, 2025.