**NLP FOR HYBRID TEXT CLASSIFICATION**

**Abstract:**

This systematic review examines methodologies for the classification of biomedical literature, a critical task addressing complex diagnostic and treatment-related questions that physicians frequently encounter. Our goal is to provide an analysis of recent studies on medical text classification, focusing on evaluation metrics, machine learning approaches, dataset selection, and challenges in achieving accurate representations of medical concepts such as medications, procedure codes, and patient data. We reviewed articles from several databases, including Web of Science, Scopus, MDPI, arXiv, IEEE, and ACM, covering research from January 1, 2016, to July 10, 2022. Out of 894 articles, we identified 33 that specifically address biomedical text categorization. Our analysis highlights two central issues in biomedical text classification: challenges related to dataset quality and methodological limitations. Using a Natural Language Processing (NLP) model, our study achieves an accuracy of 67%, underscoring ongoing challenges in this field and opportunities for advancement in both model design and data handling practices.

**Keywords:**biomedical literature, classification accuracy, dataset quality, diagnosis, evaluation metrics, machine learning, medical text classification, Natural Language Processing (NLP), systematic review, text categorization