**Literature Review (Body and Format Plan) [Get reviewed]**

**Clinical Decision Support:**

Overview of what it is, its importance in healthcare, and its current use in medicine.

**Text Classification:**

Explanation of text classification methods and their relevance to medical prescriptions.

**Medical and Non-Medical Prescription Classification:**

* Review of research on the use of text classification in differentiating between medical and non-medical prescriptions.
* Discuss any key models or algorithms that have been used in this area.

**Datasets for Prescription Classification:**

Discussion of the different data sources used in this type of research. Specifically talk about the datasets (MTSamples and Kaggle).

Clinical Decision Support (CDS) systems assist health professionals by providing accessible health-related information, significantly through Natural Language Processing (NLP). Demner-Fushman et al., (2009) in their paper states that NLP enables the extraction of data from narrative text, such as radiology reports and discharge summaries, enhancing all three primary components of CDS systems: patient data, decision rules/knowledge base, and patient-specific assessments/recommendations. CDS systems have improved practitioner performance in about 60% of reviewed cases, with proactive and automatic provision of decision support, and offering recommendations rather than assessments at the point of decision-making being key factors for success. The literature suggests a future need for reliable and high-quality NLP performance, as well as modular, flexible, and rapid systems. However, further research is required to gauge the effectiveness and adaptability of NLP systems for CDS and to establish evaluation methods for their impact on healthcare. Future NLP development will be influenced by user readiness to adopt NLP.

In their seminal work, Al-Garadi et al. (2021) addressed the critical issue of prescription medication (PM) misuse and abuse, which has surged to the level of a national crisis in the United States. They recognized the potential of social media, particularly Twitter, as a resource for active monitoring of this issue. However, they acknowledged that automation of such monitoring posed significant challenges, necessitating the use of advanced Natural Language Processing (NLP) and machine learning techniques.

Al-Garadi and colleagues extensively experimented with state-of-the-art bi-directional transformer-based language models like BERT, RoBERTa, XLNet, AlBERT, and DistilBERT, leveraging their ability to utilize tweet-level representations for transfer learning. In contrast to traditional machine learning and deep learning approaches, the focus of their work was on developing more sophisticated models capable of detecting and classifying instances of PM misuse from Twitter data.

The team proposed and evaluated fusion-based models, which showed significant improvement over traditional models. They reported an impressive F1-score (0.67 [95% CI: 0.64–0.69]) for these models, significantly outperforming the F1-score of traditional models (0.45 [95% CI: 0.42–0.48]). Additionally, Al-Garadi et al. (2021) noted the stability of these transformer-based models, which required less annotated data compared to other models, contributing to their effectiveness and efficiency.

However, Al-Garadi et al. (2021) also pointed out the challenges associated with using BERT and BERT-like models. These primarily stemmed from the idiosyncratic nature of social media language, which often lacks context and is riddled with evolving slang and dialects. The unique ways in which information about nonmedical use is presented on social media also posed a challenge. The team identified these challenges as potential areas for future research, which could further enhance the accuracy and efficiency of these models.

In conclusion, Al-Garadi et al. (2021) made a significant contribution to the literature, demonstrating the potential of transformer-based models for detecting self-reports of PM misuse on Twitter, while also outlining future research directions to further enhance these models.

**References:**

* Al-Garadi, M.A. et al. (2021) Text classification models for the automatic detection of nonmedical prescription medication use from social media - BMC Medical Informatics and decision making, BioMed Central. Available at: <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01394-0#Sec3> (Accessed: 02 June 2023).
* Demner-Fushman, D., W. Chapman, W. and J. McDonald, C. (2009) What can natural language processing do for clinical decision support?, Journal of Biomedical Informatics. Available at: <https://doi.org/10.1016/j.jbi.2009.08.007> (Accessed: 28 May 2023).