

INTISARI

REKOMENDASI ISTILAH-ISTILAH *MEDICAL SUBJECT HEADINGS* (MeSH) BERBAHASA INDONESIA MENGGUNAKAN BIOBERT

Oleh

Joko Susilo

20/466410/PPA/05983

Sistem rekomendasi untuk istilah-istilah MeSH saat ini sudah banyak dilakukan penelitian baik yang diambil dari abstrak maupun dari keseluruhan artikel. Penelitian-penelitian tersebut keseluruhannya adalah untuk mencari atau merekomendasikan MeSH dari abstraksi atau artikel biomedis berbahasa Inggris sehingga istilah-istilah MeSH yang terkandung di dalamnya juga berbahasa Inggris. Dari penelitian-penelitian tersebut terdapat kelemahan yaitu belum dilakukan penelitian terkait dengan pencarian istilah-istilah MeSH yang terkandung dalam artikel-artikel biomedis berbahasa Indonesia. Beberapa metode pendekatan dilakukan untuk mengatasi permasalahan tersebut.

Pada penelitian sebelumnya dilakukan beberapa metode dan pendekatan beberapa diantaranya menggunakan model RNN dan LSTM kemudian penelitian terakhir menggunakan transformer model dan *multi-knn*. Selanjutnya penelitian ini memiliki tujuan untuk dapat menjawab kelemahan dari penelitian sebelumnya yaitu untuk membangun model *pre-trained* BioBERT dan LSTM untuk menghasilkan rekomendasi istilah-istilah MeSH yang telah diberikan label pada artikel Biomedis berbahasa Indonesia.

Pada penelitian ini digunakan model *Long Short Term Memory* (LSTM) dan model *pre-trained* BioBERT. Penelitian ini juga melakukan pengujian hasil kinerja menggunakan *Confussion Matrix* dengan memperhatikan dari sisi *accuracy*, *precision*, *recall*, dan *f-measure*. Hasil kinerja pengujian model LSTM adalah *accuracy* 0.6622, *precision* 0.7261, *recall* 0.6456 dan *f-measure* 0.6835. Sedangkan model BioBERT adalah *accuracy* 0.8238, *precision* 0.84, *recall* 0.81 dan *f-measure* 0.82

Kata Kunci: *Long Short Term Memory*, *MeSH Recommendation*, *Biomedical*.

ABSTRACT

RECOMMENDATIONS FOR MEDICAL SUBJECT HEADING (MeSH) TERMS IN INDONESIAN LANGUAGE USING BIOBERT

By

Joko Susilo

20/466410/PPA/05983

The recommendation system for MeSH terms has now been extensively researched both from abstracts and from entire articles. These studies are entirely aimed at finding or recommending MeSH from biomedical abstracts or articles in English so that the MeSH terms contained therein are also in English. From these studies there is a weakness, namely that research has not been carried out related to the search for MeSH terms contained in Indonesian language biomedical articles. Several approaches have been made to overcome these problems.

In the previous research, several methods and approaches were carried out, some of which used the RNN and LSTM models, then the last research used the transformer model and multi-knn. Furthermore, this study aims to be able to answer the weaknesses of previous research, namely to build pre-trained BioBERT and LSTM model to produce a recommendation system for MeSH terms that have been labeled in Indonesian Biomedical articles.

In this research study, the Long Short Term Memory (LSTM) model and the BioBERT pre-trained model were used. This study also tested the performance results using the Confussion Matrix will give result to accuracy, precision, recall, and f-measure. The performance results of testing the LSTM model have accuracy 0.6622, precision 0.7261, recall 0.6456 and f-measure 0.6835. While the BioBERT model is accuracy 0.8238, precision 0.84, recall 0.81 and f-measure 0.82

Key Words : *Long Short Term Memory, MeSH Recommendation, Biomedical.*