

ABSTRAK

Penyakit yang semakin beragam dan kompleks, memerlukan cara pendekatan diagnosis yang lebih adaptif serta cerdas dari sebelumnya. Perkembangan teknologi informasi dalam bidang kesehatan membuka peluang dalam membantu proses diagnosis penyakit. Kasus ini bertujuan membuat sistem prediksi penyakit berdasarkan gejala atau keluhan dari pasien, menggunakan algoritma *Random Walk* yang menggunakan struktur *Knowledge Graph*. *Knowledge Graph* digunakan untuk merepresentasikan hubungan antara gejala atau keluhan dengan penyakit secara lebih terstruktur, sehingga memudahkan sistem untuk melakukan identifikasi keterkaitan diantara entitas. *Dataset* didapat dari *Kaggle* dengan nama "*DiseaseAndSymtoms.csv*" yang berisi 4921 baris dan 18 kolom, yang berisi berbagai penyakit beserta gejala atau keluhannya. Sistem ini meliputi *preprocessing data*, pembangunan struktur *graph* berbasis *Knowledge Graph*, prediksi menggunakan algoritma *Random Walk*, kemudian perhitungan evaluasi dan pengembangan sistem antarmuka berbasis *GUI*. Hasil pengujian dalam 3 percobaan menunjukkan bahwa *Random Walk* dengan 1000 langkah adalah model yang paling optimal dengan penyakit dengan kemungkinan paling tinggi dapat dikenali dengan baik, walaupun evaluasi klasifikasi masih perlu ditingkatkan, tetapi evaluasi berbasis peringkat memiliki hasil yang cukup baik yaitu *Mean Reciprocal Rank (MRR)* dan *Mean Average Precision (MAP)* sebesar 58.10% & dan *HitRate@3* sebesar 71.74%. Ini menunjukkan potensi model dalam mendukung diagnosis penyakit berdasarkan keluhan.

Kata Kunci : prediksi penyakit, gejala pasien, *Random Walk*, *Knowledge Graph*, evaluasi peringkat.

ABSTRACT

Diseases are increasingly diverse and complex, requiring a more adaptive and intelligent diagnostic approach than before. The development of information technology in the health sector opens up opportunities to assist the disease diagnosis process. This study aims to create a disease prediction system based on symptoms or complaints from patients, using the Random Walk algorithm that uses the Knowledge Graph structure. Knowledge Graph is used to represent the relationship between symptoms or complaints with diseases in a more structured manner, making it easier for the system to identify the relationship between entities. The dataset was obtained from Kaggle with the name "DiseaseAndSymtoms.csv" which contains 4921 rows and 18 columns, which contain various diseases and their symptoms or complaints. This system includes data preprocessing, building a Knowledge Graph-based graph structure, prediction using the Random Walk algorithm, then calculating the evaluation and developing a GUI-based interface system. The test results show that in 3 experiments, Random Walk with 1000 steps is the most optimal model with the disease with the highest possibility of being recognized well, although the classification evaluation still needs to be improved, but the ranking-based evaluation has quite good results, namely Mean Reciprocal Rank (MRR) and Mean Average Precision (MAP) of 58.10% & and Hit Rate@3 of 71.74%. This shows the potential use of this model in supporting disease diagnosis based on symptoms or complaints.

Keywords : disease prediction, patient symptoms, Random Walk, Knowledge Graph, ranking evaluation.