

Prediksi Diagnosis Hepatitis B menggunakan Gated Graph Neural Network

Fadhil Wisnu Ramadhan, Kemas Rahmat Saleh Wiharja, S.T., M.Eng., Ph.D.

Fakultas Informatika, Universitas Telkom, Bandung
fadhliwisnur@students.telkomuniversity.ac.id, bagindokemas@telkomuniversity.ac.id

Abstrak

Hepatitis merupakan infeksi virus pada hati dan dapat menyebabkan komplikasi terhadap penyakit lain yang dialami oleh pasien. Diagnosis dini dan penanganan yang tepat sangat penting untuk mencegah progresi penyakit dan komplikasi lebih lanjut. Diperlukan sebuah sistem prediksi diagnosis hepatitis yang akurat untuk menangani dan mengatasi kemungkinan terjangkitnya seseorang akan hepatitis. Penelitian ini melakukan prediksi model *Gated Graph Neural Network* terhadap pilihan data Hepatitis UCI *Machine Learning Repository*. Pada penelitian ini dilakukan pemodelan dan penelitian model dengan dua model *graph neural network* lainnya dan menghasilkan evaluasi yang baik pada prediksi klasifikasi *node* Hepatitis, dengan menggunakan *Gated Graph Neural Network* model menunjukkan nilai yang superior terhadap 2 metode lain yaitu GAT dan GCN. Dimana GGNN mendapatkan nilai *Accuracy*, *Precision*, dan *Recall* diatas 90%.

Kata Kunci: Hepatitis, Gated Graph Neural Network, Prediksi.

Abstract

Hepatitis is a viral infection of the liver that can lead to complications with other diseases experienced by the patient. Early diagnosis and proper treatment are essential to prevent disease progression and further complications. An accurate hepatitis diagnosis prediction system is necessary to address and manage the risk of someone contracting hepatitis. This research predicts hepatitis diagnosis using a *Gated Graph Neural Network* model on the Hepatitis UCI *Machine Learning Repository* dataset. The study models and compares this approach with two other *graph neural network* models, demonstrating strong evaluation results in predicting hepatitis *node* classification. The *Gated Graph Neural Network* model shows superior performance compared to the other two methods, GAT and GCN, achieving accuracy, precision, and recall scores above 90%.

Keywords: Hepatitis, Gated Graph Neural Network, Prediction.

1. Pendahuluan

Latar Belakang

Hepatitis adalah kondisi inflamasi pada hati yang disebabkan oleh berbagai virus infeksi, yang dapat menyebabkan berbagai masalah kesehatan, dengan beberapa kasus yang berpotensi fatal. Terdapat lima jenis utama virus hepatitis, yang dikenal sebagai tipe A, B, C, D, dan E. Meski semuanya menyebabkan penyakit hati, mereka memiliki perbedaan penting dalam hal cara penularan, tingkat keparahan, distribusi geografis, dan metode pencegahan. Secara khusus, tipe B dan C dapat menyebabkan penyakit kronis yang mempengaruhi ratusan juta orang dan menjadi penyebab utama sirosis hati, kanker hati, serta kematian akibat hepatitis virus. Diperkirakan sekitar 354 juta orang di seluruh dunia hidup dengan hepatitis B atau C, dan bagi kebanyakan orang, pengujian dan pengobatan masih belum dapat dijangkau. [1]. Prediksi yang akurat tentang perkembangan dan hasil infeksi Hepatitis menjadi hal yang penting untuk pengobatan dan penanganan penyakit. Terutama dengan kemajuan teknologi dalam bidang *Artificial Intelligence* dan *Machine Learning*, sebuah sistem dimana diagnosis yang lebih cepat dan akurat dapat sangat membantu penanganan pasien yang terjangkit Hepatitis.

Sudah ada beberapa penelitian sebelumnya yang menggunakan *Machine Learning* dalam melakukan prediksi virus hepatitis, dengan menggunakan klasifikasi pembelajaran mesin yang diantaranya adalah menggunakan *decision tree*, *logistic regression*, *super vector machine*, *random forest*, *Xgboost*, dan juga *AdaBoost*. Obaido et al. [5] melakukan penelitian tersebut menggunakan model-model yang dipilih dikarenakan oleh popularitas dan penggunaan reguler di beberapa kasus deteksi penyakit. Popularitas tersebut didapatkan karena model tersebut dapat melakukan pembelajaran model terhadap dataset dengan dimensi yang banyak dan *imbalance*, mengetahui paper ini menggunakan dataset UCI sebagai data untuk modelnya. Penggunaan metode tersebut ini menunjukkan bahwa model tersebut dapat melakukan performa prediksi yang baik dan efektif terhadap data klinis yang tersedia.

Ibrahim et al.[14] melakukan prediksi model terhadap infeksi hepatitis b dengan menggunakan *artificial*

neural network, dan *genetic algorithm*. Dimana data yang digunakan pada penelitian ini diambil dari 900 data medis pada . Beberapa fitur yang digunakan dalam penelitian ini dilakukan dengan mengambil gejala gejala yang dialami oleh pasien. Dari hasil penelitian tersebut dihasilkan bahwa *accuracy* yang didapat pada model menunjukkan akurasi sebesar 66,30%, spesifisitas 66,33%, dan sensitivitas 77,53%. Prediksi tersebut dianggap memiliki ukuran kinerja yang dapat diterima, yang diharapkan dapat membantu para ahli medis merespons lebih cepat di masa depan dan mengurangi wabah yang mungkin terjadi di seluruh dunia.

Dengan berkembangnya model *neural network* dalam melakukan prediksi model terhadap suatu penyakit, representasi data dalam bentuk graf mulai mendapatkan perhatian yang cukup baik beberapa tahun ini[15], data graf dapat membantu merepresentasikan data dalam bentuk yang lebih kompleks dengan dimensi yang tinggi, dimana relasi antar node dan edge dari sebuah data dapat di tentukan sebagaimana kuat relasinya terhadap node yang lain.

Dengan itu penelitian ini mengajukan sebuah model untuk memprediksi infeksi Hepatitis pada sebuah pasien dengan menggunakan Gated Graph Neural Network (GGNN). Gated graph neural networks (GGNN) adalah jenis GNN tertentu yang telah menunjukkan hasil yang menjanjikan dalam memodelkan grafik yang kompleks dan memprediksi hasil dalam berbagai bidang. GGNN menggunakan mekanisme gerbang dimana gate dalam GRU berfungsi seperti sebuah filter yang memungkinkan model untuk mempelajari representasi yang lebih kaya dan dinamis dengan menyesuaikan kontribusi dari informasi yang datang pada setiap iterasi. Dengan demikian, GGNN mampu mempertahankan konteks informasi lama sekaligus menyesuaikan informasi baru yang datang, memberikan model kemampuan untuk memproses informasi pada jaringan graf secara lebih efektif dan akurat.

Dalam konteks prediksi infeksi Hepatitis, GGNN dapat digunakan untuk memodelkan interaksi antara pasien dengan hasil tes laboratorium maupun demografisnya. Dengan adanya keterkaitan antara hasil tes diagnosis laboratorium seorang, GGNN yang berbentuk graf dapat melakukan prediksi dengan *node* sebagai pasien dan data pasien sebagai fitur *node*.

Perumusan Masalah

Berdasarkan latar belakang yang telah diurai, maka dapat dirumuskan bahwa permasalahan yang dicakup adalah:

- Bagaimana implementasi model *Gated Graph Neural Network* dalam melakukan prediksi diagnosis hepatitis?
- Bagaimana analisis performa akurasi klasifikasi yang dihasilkan oleh model *Gated Graph Neural Network* terhadap dataset Hepatitis UCI?

Batasan Masalah

Berikut adalah ruang lingkup penelitian tugas akhir ini:

- Dataset yang digunakan memiliki beberapa nilai yang hilang.
- Data yang digunakan adalah data survey Hepatitis UCI *Machine Learning Repository*, dengan hanya terdiri dari 19 fitur.

Tujuan

Tujuan dari Tugas Akhir ini adalah untuk melakukan implementasi model prediksi diagnosis pasien yang terjangkau Hepatitis dengan menggunakan Gated Graph Neural Network, dengan hasil prediksi yang akurat sebagai acuan utama dalam perancangan model.

2. Studi Terkait

Setelah dilakukan telaah lebih lanjut terhadap beberapa penelitian, ada beberapa penelitian yang memiliki keterkaitan terhadap penelitian yang akan dilakukan.

Sun et al. [8] mengusulkan model GNN dalam prediksi penyakit, yang menggunakan rekam medis elektronik(EMR). Model ini belajar membuat node *embeddings* yang sangat representatif untuk pasien, penyakit, dan gejala dari grafik konsep medis serta grafik rekam medis pasien, yang masing-masing dibangun dari basis pengetahuan medis dan EMR. Dengan menggabungkan informasi dari node tetangga yang terhubung langsung,