

1. Pendahuluan

Latar Belakang

Penyakit ginjal kronis, yang sering disebut sebagai *chronic kidney disease* (CKD) merupakan isu masalah kesehatan dengan dampak yang cukup signifikan di dunia dan telah menjadi penyebab kematian cukup banyak di abad ke-21 [1]. Berbagai faktor berkontribusi pada peningkatan kejadian CKD, seperti usia, obesitas, dan diabetes. CKD dapat menyebabkan kesulitan dalam mengeluarkan cairan berlebih dari aliran darah secara efektif. Pada tahap lanjut CKD, komplikasi seperti hipertensi, anemia, dan kerusakan saraf dapat muncul akibat penumpukan cairan, elektrolit, dan limbah berbahaya dalam tubuh [2]. Sebuah studi Global Burden of Disease menyoroti CKD sebagai penyebab utama kematian global. Angka kematian akibat CKD telah meningkat sebesar 82,3% dalam dua dekade terakhir, menjadikannya peningkatan terbesar ketiga dalam tingkat kematian, setelah HIV dan diabetes [3]. Selain itu, Global Burden of Disease melaporkan peningkatan kematian sebesar 90% dari tahun 1999 hingga 2013, menjadikan CKD sebagai penyebab kematian ke-13 di dunia [4]. Sebuah studi tahun 2017 memperkirakan bahwa jumlah pasien CKD bisa mencapai 843,6 juta, dengan prevalensi global sebesar 11,1% (10,4% pada pria dan 11,8% pada wanita) [5]. Pada tahun 2040, CKD diproyeksikan menjadi penyebab kematian kelima di dunia [6]. Untuk mendiagnosis apakah seseorang terkena CKD, kunjungan ke rumah sakit untuk pemeriksaan diperlukan. Namun, proses ini memerlukan waktu yang signifikan dan biaya tambahan. Oleh karena itu, untuk menyederhanakan deteksi CKD dengan pengujian minimal dan biaya yang rendah, teknologi pembelajaran mesin (*machine learning*) digunakan dalam pengembangan model prediktor CKD yang efektif dan optimal.

Banyak penelitian telah dilakukan untuk meneliti prediksi CKD menggunakan pembelajaran mesin. Mengacu pada penelitian sebelumnya, seperti studi oleh Neves et al. pada tahun 2015 [7], yang membangun model menggunakan *artificial neural network* (ANN) untuk diagnosis penyakit ginjal kronis (CKD). Penelitian ini memperoleh nilai sensitivitas berkisar antara 93,1% hingga 94,9%. Penelitian yang dilakukan oleh Vasquez-Morales et al. [8] pada tahun 2019 menggunakan *neural network classifier* untuk melakukan klasifikasi CKD dengan hasil akurasi yaitu 95%. Pada tahun 2020, Bhaskar et al. [11] membandingkan kinerja metode pembelajaran mesin dalam memprediksi CKD. Kinerja terbaik dicapai oleh CorrNN-SVM dengan akurasi 98,67%. Pada tahun 2021, Krishnamurthy S. et al. [12] mengembangkan model prediksi CKD menggunakan teknik *convolutional neural network* dan *bi-directional long short-term memory* (BLSTM), mencapai akurasi masing-masing 95,7% dan 93%. Pada tahun 2021, Hamida et al. [13] bertujuan untuk membandingkan algoritma *random forest* dengan *decision tree* J48 untuk prediksi CKD, dengan akurasi sebesar 85,5%, lebih baik dibandingkan dengan *random forest*. Selain itu, Arroyo et al. [14] pada tahun 2022 melakukan optimasi ANN menggunakan algoritma genetika (GA) untuk prediksi kardiovaskular. Penelitian ini menghasilkan akurasi sebesar 73,43%, sedangkan penggunaan ANN saja tanpa optimasi memperoleh akurasi 68,35%.

Berdasarkan penelitian di atas [7], [8] dan [11] - [14], optimasi LSTM menggunakan GA dipilih untuk prediksi CKD karena optimasi menggunakan GA telah terbukti memberikan akurasi yang lebih baik dibandingkan tanpa optimasi. Selain itu, LSTM akan digunakan untuk prediksi CKD karena LSTM adalah versi optimal dari jaringan saraf berulang (RNN), yang dapat mengatasi masalah terkait *gradient disappearance and explosion* saat mempelajari data sekuensial yang besar [15]. LSTM juga tahan terhadap *noise* yang membuatnya berguna dalam data tabular yang terdapat *noise* atau fitur tidak relevan. Hal ini membuat LSTM menjadi metode yang kuat untuk menganalisis data dan mengungkap fitur penting untuk prediksi [16]. Penelitian ini bertujuan untuk meningkatkan akurasinya sebesar 1% hingga 5% dengan menerapkan optimasi GA dan diharapkan dapat meningkatkan kinerja model LSTM, yang mengarah pada prediksi CKD yang lebih andal. Oleh karena itu, penelitian ini akan membandingkan hasil akurasi LSTM yang dioptimalkan menggunakan GA dengan LSTM tanpa optimasi.

Topik dan Batasannya

Topik yang dibahas pada penelitian ini adalah melakukan deteksi dini penyakit ginjal kronis atau *chronic kidney disease* (CKD) pada seseorang dengan memanfaatkan metode pembelajaran mesin, yaitu LSTM. Tak hanya itu, untuk meningkatkan performa dari LSTM, optimasi menggunakan *genetic algorithm* (GA) diterapkan. Performa dari LSTM yang dioptimasi dengan GA dibandingkan dengan LSTM tanpa optimasi sehingga pengaruh penerapan optimasi GA diketahui.

Batasan yang ada pada Tugas Akhir ini adalah penggunaan dataset yang relatif kecil dikarenakan tidak menemukan dataset serupa yang dapat digunakan secara umum. Selain itu, platform yang digunakan untuk melatih dan mengembangkan model yaitu Google Colab Pro dengan batasan RAM 50gb. Penelitian ini juga memiliki batasan *environment* tensorflow yang hanya mendukung hingga versi 2.15 dikarenakan versi 2.16 ke atas memiliki perubahan *environment* yang cukup berdampak pada hasil penelitian.

Tujuan

Tujuan dari penelitian ini adalah untuk mengembangkan model yang dapat melakukan prediksi CKD pada seseorang menggunakan *long short-term memory* (LSTM) yang dioptimasi menggunakan *genetic algorithm* (GA). Selain itu, model tersebut juga akan dibandingkan performanya dengan LSTM tanpa optimasi GA.

Organisasi Tulisan

Bagian selanjutnya akan membahas terkait penelitian yang sudah ada (bab 2), metode dan rancangan sistem yang akan dibangun (bab 3), hasil dan evaluasi sistem yang telah dibangun (bab 4), dan terakhir kesimpulan dari penelitian yang telah dilakukan (bab 5).