

1. Studi Terkait Latar Belakang

Diabetes adalah penyakit kronis yang terjadi ketika pankreas tidak memproduksi insulin dalam jumlah cukup atau ketika tubuh tidak dapat menggunakan insulin yang diproduksi secara efektif. Insulin adalah hormon yang mengatur glukosa darah. Hiperglikemia, yang juga disebut peningkatan glukosa darah atau gula darah, adalah efek umum dari diabetes yang tidak terkontrol dan seiring waktu menyebabkan kerusakan serius pada banyak sistem tubuh, terutama saraf dan pembuluh darah.[1]. Penyakit ini ditandai dengan hilangnya berat badan, poliuria, polidipsi, dan polipagia. Pada penyakit ini kekurangan insulin sehingga terjadi peningkatan kadar gula darah dan menimbulkan perubahan metabolisme protein dan lemak.[2]

DM tipe 1 merupakan salah satu penyakit kronis terbanyak yang terjadi pada anak-anak. Pengelolaan diabetes pada anak sangatlah kompleks sehingga membutuhkan dukungan dari orang tua untuk melakukannya[3]. Diabetes melitus tipe 1 (DMT1) 90% terjadi pada anak dan remaja. Insidennya bervariasi di seluruh dunia, tertinggi di Finlandia yaitu 40/100.000 populasi dan terendah di Cina 0,1/100.000 populasi.

Data Unit Kerja Koordinasi (UKK) Endokrinologi Anak Pengurus Pusat Ikatan Dokter Anak Indonesia (PPIDAI), jumlah pasien DM tipe 1 sebesar 731 pasien pada tahun 2012[4].

Gejala diabetes pada anak, khususnya diabetes tipe 1, dapat muncul dengan cepat dan sering kali lebih parah dibandingkan dengan diabetes tipe 2. Berikut adalah beberapa gejala yang umum ditemukan pada anak-anak dengan diabetes tipe 1: Sering Buang Air Kecil (Polyuria), Haus yang Berlebihan (Polidipsia), Nafsu Makan Meningkat (Polyphagia), Penurunan Berat Badan yang Tidak Dijelaskan, Kelelahan, Penglihatan Kabur, Luka yang Sulit Sembuh, Infeksi Berulang[5].

Salah satu cara untuk mengetahui DM tipe-1 perlu dilakukan tes hemoglobin A1c (HbA1c) untuk mengukur rata-rata gula dalam darah dengan melakukan pengecekan dalam 2-3 bulan terakhir[6]. Dengan banyaknya cara dalam pendeteksian diabetes, salah satu cara adalah dengan klasifikasi ekspresi gen dengan menggunakan machine learning. Ekspresi gen adalah sebuah rangkaian data yang digunakan untuk mengidentifikasi informasi gen dalam bentuk microarray (DNA dan RNA)[7]. Machine learning dapat mengklasifikasi data microarray pada ekspresi gen dengan kata lain Machine learning mempunyai kemampuannya dalam menangani data yang memiliki dimensi yang tinggi [8].

Dalam Upaya memudahkan klasifikasi dan mengatasi permasalahan pada Diabetes pada anak, penelitian ini menerapkan metode AdaBoost, dengan ditambahkan metode KNN, dan MLP sebagai pembanding, pada proses klasifikasinya dan dioptimalkan menggunakan Gravitational Search Algorithm. Beberapa penelitian terkait implementasi machine learning pada data ekspresi telah dilakukan untuk mengidentifikasi diabetes.

Pada tahun 2022, Putry, N. M, dkk, melakukan Penelitian ini melakukan perbandingan algoritma KNN dan Naïve Bayes dalam mengklasifikasikan diagnosis penyakit diabetes melitus. Selain itu, penelitian ini menggunakan lima pembagian data yang dilakukan, bahwa nilai akurasi dari Naïve Bayes lebih tinggi dibandingkan KNN. Dimana nilai akurasi paling tinggi yang didapatkan dari algoritma Naïve Bayes yaitu sebesar 80%. Sedangkan algoritma KNN nilai akurasi tertinggi yaitu sebesar 75%. Selain itu, diketahui bahwa nilai recall paling tinggi dihasilkan oleh algoritma KNN yaitu sebesar 0.92. Dan untuk nilai presisi lebih tinggi dihasilkan oleh algoritma Naïve Bayes yaitu 0.86[14].

Pada tahun 2020, Iqbal Fathur, dkk melakukan analisis klasifikasi pada data microarray hasil dari ekspresi gen pada sampel diabetes melitus tipe 2 dengan membandingkan tiga metode dalam mendeteksi DM tipe-2 dengan menggunakan SVM, Multilayer Perceptron (MLP) dan Xtreme Gradient Boosting (XBOOST) pada data ekspresi gen. Dari penelitian tersebut dapat disimpulkan bahwa SVM memperoleh hasil yang paling paling baik dalam klasifikasi dengan nilai akurasi 91,30%, Sedangkan metode klasifikasi Multilayer Perceptron mampu mengklasifikasikan data ekspresi gen dengan nilai akurasi sebesar 78,26% Metode klasifikasi terakhir Xtreme Gradient Boosting mampu mengklasifikasikan data ekspresi gen dengan nilai akurasi sebesar 73,91%. [8]

Pada tahun 2022, Madhubala dkk, melakukan penelitian untuk memprediksi diabetes. Penelitian ini menggunakan dataset PID untuk membangun pengklasifikasi untuk IBk, OneR, J48, Multilayer Perceptron, RandowForest, dan Naive Bayes. Model Nave Bayes menghasilkan kinerja yang baik dibandingkan dengan model lain, pada ANN yang dimodifikasi menggunakan multilayer perceptron, menghasilkan akurasi pelatihan 79% dan akurasi pengujian 77%[21].

Pada tahun 2021, Fallucchi dkk, melakukan pengujian untuk Prediksi Resiko Diabetes Menggunakan MLP, hasil yang didapatkan untuk model keseluruhan lebih besar dari 0,939, dengan hasil terbaik lebih dari 0,96 untuk 1500 elemen berlabel sebagai kumpulan data pelatihan[24].

Pada tahun 2020, Goudjerk dkk, melakukan Klasifikasi Otomatis Penderita Diabetes dengan MLP. Hasil yang didapat pada akurasi hingga 95% untuk Klasifikasi.[26].

Penelitian terkait GSA telah pernah dilakukan pada tahun 2021 oleh Priyadarshini, R. untuk memprediksi akurasi diabetes meitus. aklasifikasi untuk dataset pima dihitung menggunakan metode GSA-K-means. Hasil yang diperoleh kemudian dibandingkan dengan hasil algoritma pengelompokan k-means secara terpisah dan dijalankan selama 10 iterasi. Hasil yang ditunjukkan di atas adalah rata-rata dari semua hasil iterasi Hasil pengamatan menunjukkan bahwa teknik GSA-k-means yang diusulkan mampu mendeteksi diabetes melitus dengan tingkat akurasi yang tinggi sebesar 86% sedangkan untuk klasifikasi akurasi full train and full test pengujian ini mendapatkan hasil 92%[18].

Sehubungan dengan penelitian diatas, penelitian ini menggunakan dataset Diabetes pada anak berdasarkan ekspresi gen menggunakan AdaBoost dengan ditambahkan metode KNN, dan MLP sebagai pembandin dalam mengklasifikasi dan dioptimalkan dengan Gravitational Search Algorithm. Pengujian model dilakukan dengan menerapkan Confusion Matrix dalam menguji seberapa baik sebuah model dengan data yang memiliki. Tujuan dari penelitian ini adalah klasifikasi penyakit Diabetes Melitus pada anak.

Topik dan Batasannya

Topik dan Batasan yang dibahas pada tugas akhir ini adalah megidentifikasi diabetes pada anak berdasarkan ekspresi gen berdasarkan dataset GSE9006. yang berisi profil ekspresi gen dengan penelitian yang difokuskan untuk meneliti T1D yang terdiri dari 162 sampel. Adapun batasan masalah pada penelitian ini adalah Gravitational Search Algorithm sebagai seleksi fitur. Dan pengklasifikasian menggunakan metode metode utama yaitu Adaptive Boosting (AdaBoost), selanjutnya ditambahkan 2 metode ensemble sebagai pembanding yaitu KNeighbors (KNN), Multi- Layer Perceptron (MLP).

Rumusan Masalah

Rumusan masalah yang diangkat pada penelitian ini yaitu:

1. Bagaimana performa metode Gravitational Search Algorithm dalam melakukan klasifikasi pada identifikasi Diabetes pada anak?
2. Bagaimana pengaruh hyperparameter dalam melakukan klasifikasi pada identifikasi Diabetes pada anak?
3. Bagaimana performa AdaBoost dan 2 model pembanding yaitu KNN dan MLP dalam melakukan klasifikasi pada identifikasi Diabetes pada anak

Tujuan

Tujuan yang ingin dicapai pada penelitian ini yaitu:

1. Mengetahui performa metode Gravitational Search Algorithm dalam melakukan klasifikasi pada identifikasi diabetes pada anak.
2. Mengetahui pengaruh hyperparameter dalam melakukan identifikasi dan tuning Diabetes Melitus pada anak
3. Mengetahui performa metode AdaBoost dan 2 model pembanding yaitu KNN dan MLP dalam melakukan klasifikasi pada identifikasi diabetes pada anak.