

1. Pendahuluan

Latar Belakang

Diabetes Melitus (DM) didefinisikan sebagai suatu penyakit dengan *multi etiologic* atau gangguan metabolisme kronis ditandai dengan tingginya kadar gula darah disertai dengan gangguan metabolisme karbohidrat, lipid dan protein sebagai akibat dari defisiensi insulin [1]. Apabila terlalu banyak gula yang menetap pada aliran darah dalam waktu yang lama akan dapat mempengaruhi fungsi pembuluh darah, saraf, mata, ginjal serta sistem kardiovaskular [2].

Seringkali DM dianggap hanya diderita oleh orang dewasa saja. Pada kasus nya, DM tipe-1 banyak dialami oleh anak-anak, IDAI mencatat bahwa terjadi peningkatan yang cukup signifikan DM tipe-1 pada anak dan remaja dari 3,88 menjadi 28,19 per 100 juta penduduk pada tahun 2000 dan 2010. Data Ikatan Dokter Anak Indonesia (IDAI) menyatakan angka penderita DM pada anak usia 0-18 mengalami peningkatan 700% selama jangka 10 tahun tercatat sejak September 2009 hingga September 2008 terdapat 1213 kasus DM tipe-1 [3].

Untuk mendeteksi DM tipe-1 perlu dilakukan tes hemoglobin A_{1c} (HbA_{1c}) untuk mengukur kadar rata-rata gula darah pasien 2-3 bulan terakhir. DM tipe-1 dapat didiagnosa apabila kandungan keton terdeteksi didalam urine [4]. Namun, ada beberapa hal yang dapat mempengaruhi kadar HbA_{1c} yaitu HbA_{1c} dapat meningkat dari nilai sebenarnya apabila dipengaruhi oleh anemia, konsumsi alkohol berlebihan, penggunaan salisilat dosis tinggi dalam jangka panjang, dan lainnya. Dan akan menurunkan nilai kadar HbA_{1c} apabila setelah melakukan transfusi darah, kehilangan darah, penyakit ginjal, penyakit hati, infeksi HIV, dan lainnya [5].

Dikarenakan banyak hal yang mempengaruhi pendeteksian DM tipe-1 menggunakan tes hemoglobin A_{1c} (HbA_{1c}). Maka salah satu alternatif dilakukan yaitu deteksi diabetes dengan menggunakan data ekspresi gen. Data ekspresi gen adalah suatu rangkaian proses identifikasi banyak informasi gen dan mengetahui gambaran dari suatu gen dalam bentuk *microarray* (DNA dan RNA) [6]. Dengan menggunakan teknologi DNA microarray, yang memungkinkan untuk mengamati beribu-ribu ekspresi gen dalam waktu yang bersamaan [13]. *Machine learning* akan mengklasifikasi DM tipe-1 menggunakan data *microarray* ekspresi gen. Selain itu, teknik *machine learning* juga membantu menyelesaikan permasalahan-permasalahan pada data-data dengan dimensi tinggi seperti *microarray* [7].

Beberapa penelitian terkait implementasi *machine learning* pada data ekspresi telah dilakukan untuk mengidentifikasi diabetes. 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 baik dalam klasifikasi dengan nilai akurasi 91,30%. [7].

Kemudian, pada tahun 2020 S. Thenappan, dkk melakukan prediksi Diabetes Melitus (DM) menggunakan Modified Support Vektor Machine (MSVM) Cloud Security dengan menggunakan ukuran kinerja metode *accuracy*, *sensitivity*, *precision* dan *negative predictive value*. Hasil menunjukkan bahwa penklasifikasian MSVM lebih baik dengan akurasi tertinggi sebesar 97,13% [8]. Pada tahun 2018, Tao Fang, dkk melakukan penelitian tentang mendiagnosa diabetes dengan cepat menggunakan ATR-FTMIR dengan SVM dan beberapa metode. Dengan kesimpulan GA lebih baik dalam mengoptimalkan parameter SVM dan kombinasi ATR-FTMIR dengan SVM mampu mendiagnos diabetes tipe-2 secara cepat dan akurat tanpa reagen [9].

Kemudian dilakukan penelitian pada tahun 2018 oleh Quan Zou, dkk tentang memprediksi diabetes melitus menggunakan *decision tree*, *random forest*, dan *neural network*. Hasilnya, bahwa prediksi dengan *random forest* mencapai akurasi 80,84% [10]. Adapun pada tahun 2018, Arianna Dagliati, Phd dan rekan melakukan penelitian tentang prediksi komplikasi diabetes menggunakan RF untuk mengatasi missing data dan menerapkan strategi yang sesuai untuk menangani ketidakseimbangan kelas. Model akhir yang diperoleh sesuai dengan komplikasi diabetes menghasilkan akurasi hingga 83,8% [11].

Adapun penelitian terkait GSA pernah dilakukan pada tahun 2021 oleh Isman Kurniawan dan rekan, GSA digunakan untuk memprediksi aktivitas inhibitor protease HIV-1 dengan menggunakan *neural network* dan *hyperparameter tuning* untuk meningkatkan kinerja model. Hasil nya dari ke-empat model yang diprediksi, model 3 yang berisi 7 deskriptor memberikan hasil yang paling baik [12].

Penelitian ini bertujuan untuk mengidentifikasi diabetes pada anak berdasarkan ekspresi gen menggunakan pendekatan berbasis *microarray*, penelitian ini akan membangun model prediksi diabetes pada anak dengan menggunakan *Gravitational Search Algorithm-Ensemble*. GSA memiliki beberapa konsep dan operator unik yang tidak mirip dengan metaheuristik lainnya. [13]. *Ensemble* dipilih karena penggabungan beberapa single classifier untuk membentuk sebuah classifier baru sehingga didapatkan hasil yang lebih akurat [14].

Topik dan Batasannya

Topik dan Batasan yang dibahas pada tugas akhir ini adalah mengidentifikasi diabetes pada anak berdasarkan ekspresi gen berdasarkan dataset GSE9006. Adapun batasan masalah pada penelitian ini adalah *Gravitational Search Algorithm* sebagai seleksi fitur. Dan pengklasifikasian menggunakan *Ensemble Method* yakni *Random Forest*, *AdaBoost* dan *XGBoost*.

Rumusan Masalah

Rumusan masalah yang diangkat pada penelitian ini yaitu:

1. Apakah metode *Gravitational Search Algorithm* mampu melakukan seleksi dalam identifikasi *Diabetes Melitus* pada anak?
2. Bagaimana membangun model prediksi dengan menggunakan metode *Ensemble*?
3. Bagaimana performa metode *Gravitational Search Algorithm-Ensemble* dalam melakukan klasifikasi pada identifikasi Diabetes pada anak?

Tujuan

Tujuan yang ingin dicapai pada penelitian ini yaitu:

1. Menggunakan metode *Gravitational Search Algorithm* untuk melakukan seleksi fitur.
2. Membangun model prediksi dengan menggunakan metode *Ensemble*.
3. Mengetahui performa metode *Gravitational Search Algorithm-Ensemble* dalam melakukan klasifikasi pada identifikasi diabetes pada anak menggunakan pendekatan *microarray*.