

## 1. Pendahuluan

### 1.1 Latar Belakang

Diabetes mellitus adalah penyakit kronis yang disebabkan oleh kurangnya produksi hormon insulin dalam tubuh dan ketidakmampuan untuk menggunakan insulin dengan baik.[1] Insulin adalah hormon yang diproduksi di pankreas yang memecah glukosa darah[2]. Diabetes adalah penyakit dimana kadar gula darah meningkat sedemikian rupa sehingga tubuh tidak dapat memproses kadar gula tersebut[3]. Penyakit ini biasanya ditandai dengan gula darah yang tinggi. Diabetes dibagi menjadi tiga jenis yaitu diabetes tipe 1, diabetes tipe 2 dan diabetes gestasional [1][2]. Diabetes tipe 1 masih merupakan penyakit metabolik tersering pada anak. Menurut perkiraan saat ini, 15.600 hingga 17.400 anak-anak dan remaja berusia 0-14 tahun hidup dengan diabetes tipe 1 di Jerman. Pada 1990-an, rata-rata tingkat penyakit baru tahunan (tingkat insiden) dilaporkan antara 12,9 (95% interval kepercayaan 12,4-13,4) dan 14,2 (interval kepercayaan 95% 12,9 – 15,5) per 100.000 anak berusia 0-14 tahun dan 17,0 (95% interval kepercayaan 15,2-18,8) per 100.000 anak usia 0-

19 tahun. Angka kejadian meningkat 3-4% per tahun. Peningkatan angka kejadian terutama mempengaruhi kelompok usia yang lebih muda [4]. Diabetes Tipe 1 hasil dari penghancuran autoimun sel-sel pankreas yang memproduksi insulin di pankreas [5]. Diabetes tipe 1 lebih sering terjadi pada anak – anak dan remaja. Diabetes tipe 1 dikategorikan oleh penghancuran sel autoimun, akibatnya, tubuh tidak menghasilkan insulin yang tidak mencukupi [6].

Sejauh ini, belum ada obat yang dapat menyembuhkan diabetes secara tuntas[1]. Salah satu cara yang mungkin untuk mengobati diabetes adalah dengan menghambat pertumbuhan enzim Diacylglycerol Acyltransferase-1 (DGAT-1). Dalam bidang kesehatan, diagnosis penyakit tergantung pada keputusan dokter untuk mengidentifikasinya sebagai penyebab yang paling mungkin pada gejala pasien. Penyakit ini dapat didiagnosis menggunakan tes fisik dan kimia[7]. Diagnosis diabetes tipe 1 didasarkan pada gejala klinis dan pemantauan glukosa darah. Komplikasi yang diakibatkan oleh diabetes adalah seperti penyakit jantung koroner, stroke, dan penyakit vaskular perifer, penyakit ginjal tahap akhir (ESRD), retinopati dan neuropati[8]. Diperkirakan 422 juta orang dewasa di seluruh dunia menderita diabetes, dimana 1,5 juta di antaranya meninggal pada tahun 2014 [9]. Diabetes merupakan penyebab kematian ketiga di Indonesia pada tahun 2016 saja, yaitu sebesar 6,7%, setelah stroke sebesar 21,1% dan penyakit jantung koroner sebesar 12,9%, dan terus meningkat setiap tahunnya [3].

Berdasarkan data tersebut deteksi dini diperlukan untuk mencegah kejadian tersebut. Dengan perkembangan teknologi saat ini, implementasi *machine learning* pada *microarray* banyak digunakan di segala bidang, termasuk bidang medis [3]. Machine Learning adalah pembelajaran berbasis komputer yang telah banyak digunakan untuk analisis data dalam kedokteran selama dekade terakhir[10][11]. Teknologi *microarray* diperkenalkan untuk mendefinisikan global pandangan fungsi sel dengan pendekatan gen. Selain itu, teknologi *microarray* digunakan untuk mengukur aktivitas ekspresi gen dari genom lengkap menjadi satu percobaan [12]. Teknologi ini juga telah berhasil diterapkan dalam banyak masalah dan telah mencapai hasil yang lebih unggul dibandingkan dengan teknik lain, khususnya di bidang medis. *Microarray* juga telah menunjukkan kemampuan untuk mendiagnosis pasien yang memiliki penyakit tertentu. Penggunaan teknik data mining dapat membantu mendiagnosis diabetes secara dini, yang dapat membantu mencegah penyakit dan berbagai komplikasinya, seperti penyakit kardiovaskular, gangguan penglihatan, dan penyakit nefogenik [13].

Identifikasi diabetes pada anak dan dewasa dengan *microarray* menggunakan metode machine learning beberapa kali sudah pernah dilakukan pada penelitian sebelumnya. Pada tahun 2020, syahmaa dan rekan melakukan penelitian untuk mengklasifikasi diabetes untuk type 1 dan 2 menggunakan metode Extended Binary Cuckoo Search dan K-Nearest Neighbor Classifier dengan akurasi 77,95% untuk type 1 dan 73,38% untuk type 2 [13]. Pada tahun 2019 Firda dan rekan melakukan penelitian untuk mengklasifikasi diabetes type 2 menggunakan SVM dengan hasil klasifikasi data uji adalah 100% dengan 1000 iterasi dan 100% AUC [14].

Pada tahun 2018, Haritha dan rekan melakukan penelitian untuk memprediksi diabetes type 1 dan type 2. Pada penelitian ini hasil terbaik dihasilkan oleh pengujian menggunakan cuckoo search yang dioptimalkan dengan fuzzy KNN dan nilai akurasi yang didapat mencapai 81 % [15].

Pada tahun 2017, Sonu dan rekan melakukan penelitian untuk mengklasifikasi diabetes menggunakan beberapa metode yaitu Naïve Bayes, Bayes Net, J48, SMO, dan Random Forest. Metode terbaik pada penelitian ini pengklasifikasian menggunakan metode SMO dengan hasil akurasi 77,34 dan memiliki rata – rata error terendah sebesar 22,65% [16]. Pada tahun 2019 Ramdaniah dan rekan melakukan penelitian untuk mengklasifikasi data *microarray* ekspresi gen ke dalam Diabetes Tipe 2 kelas mellitus dan kelas normal. Pada penelitian ini hasil terbaik dihasilkan oleh pengujian menggunakan metode SVM dengan kernel polinomial dengan akurasi yang didapat mencapai 100% [17].

Pada *microarray* terdapat tantangan yaitu seleksi fitur. Dataset *microarray* terdiri dari sejumlah kecil pengamatan, tetapi banyak gen [18]. Seleksi fitur dapat diatasi dengan metode *meta-heuristic* salah satunya adalah *Cuckoo Search* (CS). Berdasarkan literatur review di atas belum banyak implementasi metode *meta-heuristic* pada seleksi fitur *microarray* untuk prediksi diabetes pada anak. Salah satu metode *meta-heuristic* adalah *Cuckoo Search*.

Penelitian ini bertujuan untuk mengimplementasikan metode Cuckoo Search Algorithm – Support Vector Machine untuk prediksi diabetes pada anak berdasarkan data ekspresi gen. *Cuckoo search* (CS) dikenal karena efisiensinya sebagai algoritma berbasis swarm-intelligence dan bukan membangun algoritma metaheuristik baru, tetapi mengimprovisasi algoritma yang ada untuk memungkinkan efisiensi algoritma ditingkatkan [19]. Keuntungan dari algoritma CS adalah kesederhanaannya. CS memiliki lebih sedikit parameter yang perlu disetting sebelum memulai pencarian dibandingkan dengan teknik lain. Oleh karena itu, sangat mudah untuk diterapkan [20]. Untuk pengembangan model prediksinya menggunakan metode *Support Vector Machines* (SVM) karena SVM menawarkan hasil akurasi tinggi dan bekerja dengan baik dengan ruang dimensi tinggi. SVM pengklasifikasi pada dasarnya menggunakan subset dari train sehingga menggunakan memori yang sangat sedikit.

## 1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah:

1. Bagaimana efektifitas seleksi fitur dengan menggunakan metode *Cuckoo Search*?
2. Bagaimana pengaruh hyperparameter tuning diabetes pada anak menggunakan metode SVM?
3. Bagaimana hasil performa prediksi menggunakan metode *Cuckoo Search* dan *Support Vector Machines*?

## 1.3 Tujuan

Tujuan pada penelitian ini adalah:

1. Menerapkan metode *Cuckoo Search* (CS) untuk melakukan seleksi fitur.
2. Mengidentifikasi diabetes pada anak menggunakan metode *Support Vector Machines* (SVM).
3. Mengetahui hasil performa menggunakan metode *Cuckoo Search* (CS) dan *Support Vector Machines* (SVM).