

BAB I

PENDAHULUAN

1.1 Latar Belakang

Diabetes Retinopati merupakan penyakit pada mata atau yang biasa disebut Diabetes Mellitus (DM) pada umumnya. Penyakit diabetes adalah salah satu penyakit yang metabolik meningkat secara drastis didunia (Perveen et al., 2016). Seiring bertambahnya waktu dan semakin banyaknya orang yang terkena penyakit Diabetes Mellitus (DM), maka perlu adanya diagnosis lebih dini dan akurat sehingga dapat membantu meningkatkan penanganan dan perawatan medis (Science et al., 2018).

Penelitian sebelumnya, telah dilakukan classifier untuk deteksi penyakit Diabetes Mellitus (DM) menggunakan metode Support Vector Machine (SVM) dengan akurasi 76.30% dari 8 atribut 768 kasus (Sisodia & Sisodia, 2018). Ada juga yang menggunakan metode K-Nearest Neighbor (K-NN) untuk mendeteksi penyakit Diabetes Mellitus (DM) pada RSUD Puri Husada Tembilahan terdiri dari 6 variabel input dan target sebanyak 2 kelas dengan akurasi 96% (Yunita, 2016).

Kelebihan pada penelitian sebelumnya, yang menggunakan metode Support Vector Machine (SVM) memiliki kelebihan yaitu dapat memaksimalkan pemisah margin antara 2 kelas sehingga dapat memperoleh akurasi yang lebih tinggi. Sedangkan KNN dapat memecahkan masalah dari penelitian sebelumnya karena memiliki kelebihan yaitu ketangguhan terhadap training data yang menghasilkan data yang efektif jika digunakan pada data yang besar. Tetapi kelemahan yang terdapat pada metode SVM garis lurus titik suatu data tidak boleh saling berdekatan untuk tingkat akurasi yang lebih tinggi.

Sedangkan metode KNN juga memiliki kelemahan yaitu kurang optimal dalam menentukan nilai K dan perlu menggunakan semua atribut atau hanya 1 atribut yang pasti dengan kata lain perlu melakukan beberapa percobaan pada nilai k, serta metode knn ini tidak dapat menentukan nilai yang kosong atau missing pada suatu data yang besar.

Namun dalam penelitian ini metode yang digunakan yaitu Decision Tree Algoritma C4. 5. Decision tree adalah algoritma supervised machine learning yang digunakan untuk memecahkan suatu masalah klasifikasi dan mampu menghasilkan prediksi secara spesifik yang mudah diimplementasikan (Noviandi, 2018). Sedangkan metode mean merupakan salah satu metode imputasi yang mampu menyelesaikan missing value pada data dalam suatu variable dengan menggunakan nilai semua rata-rata yang telah diketahui pada suatu variabel (Pratama et al., 2017).

Pada penelitian ini metode mean akan diterapkan untuk mengatasi data missing value pada dataset yang terdapat pada algoritma C4.5.

1.2 Rumusan Masalah

1. Seberapa meningkat akurasi jika metode mean diterapkan pada data yang terdapat missing value?
2. Seberapa meningkat akurasi C4. 5 jika diterapkan metode mean?

1.3 Batasan Masalah

1. Dataset penyakit Diabetes Retinopati diambil dari UCI Machine Learning Repository Pima Indians Diabetes Data

2. Menghitung data yang terdapat missing value didalamnya untuk mengetahui tingkat akurasi yang dihasilkan

1.4 Tujuan Penelitian

1. Untuk mengetahui tingkat akurasi metode mean yang digunakan pada data yang terdapat missing value
2. Untuk mengetahui tingkat akurasi C4. 5 jika diterapkannya metode mean

1.5 Manfaat Penelitian

1. Dengan adanya penelitian ini diharapkan agar bisa menambah wawasan dan pembelajaran bagi peneliti
2. Dengan adanya penelitian ini diharapkan bisa bermanfaat bagi peneliti lainnya sebagai referensi, pembandingan, atau hanya sekedar menambah wawasan bagi peneliti lainnya
3. Dengan adanya tingkat akurasi yang lebih baik diharapkan bisa membantu dalam penanganan pasien diabetes retinopati

