

BAB I

PENDAHULUAN

1.1. Latar Belakang

Diabetes melitus (DM) merupakan suatu kelompok penyakit metabolik dengan karakteristik hiperglikemia yang terjadi karena kelainan sekresi insulin, kerja insulin, atau kedua-duanya. Hiperglikemia kronik pada diabetes berhubungan dengan kerusakan jangka panjang, disfungsi atau kegagalan beberapa organ tubuh, terutama mata, ginjal, saraf, jantung, dan pembuluh darah (Sudoyo, 2009). Diabetes mellitus merupakan suatu penyakit yang ditandai oleh kadar glukosa darah melebihi normal dan gangguan metabolisme karbohidrat, lemak dan protein yang disebabkan kekurangan hormon insulin secara relatif maupun absolut. Bila hal ini dibiarkan tak terkendali dapat terjadi komplikasi metabolik akut maupun komplikasi vaskuler jangka panjang, baik mikroangiopati maupun makroangiopati (Darmono dalam Hasdianah, 2007 : 1).

Klasifikasi diabetes melitus mengalami perkembangan dan perubahan dari waktu ke waktu. Dahulu diabetes diklasifikasikan berdasarkan waktu munculnya (time of onset). Diabetes yang muncul sejak masa kanak-kanak disebut “*juvenile diabetes*”, sedangkan yang baru muncul setelah seseorang berumur di atas 45 tahun disebut “*adult diabetes*”. Namun klasifikasi ini sudah tidak layak dipertahankan lagi, sebab banyak sekali kasus-kasus diabetes yang muncul pada usia 20-39 tahun, yang menimbulkan kebingungan untuk mengklasifikasikannya. Beberapa penelitian mengenai prediksi atau analisis metode terhadap penderita penyakit diabetes dilakukan dengan tujuan mencari metode terbaik dan paling cocok. *Data mining* pun menjadi salah satu didalamnya.

Pada tahun 1968, ADA (*American Diabetes Association*) mengajukan rekomendasi mengenai standarisasi uji toleransi glukosa dan mengajukan istilah-istilah *prediabetes*, *suspected diabetes*, *chemical* atau *latent diabetes* dan *overt diabetes* untuk pengklasifikasiannya. *British*

Diabetes Association (BDA) mengajukan istilah yang berbeda, yaitu *potential diabetes*, *latent diabetes*, *asymptomatic* atau *subclinical diabetes*, dan *clinical diabetes*.

Naïve Bayes merupakan suatu metode klasifikasi yang menggunakan perhitungan probabilitas. Penentuan kelas dari suatu data pada dataset dilakukan dengan cara membandingkan nilai probabilitas suatu sampel berada di kelas yang satu dengan nilai probabilitas suatu sampel berada di kelas yang lain. Metode klasifikasi *Naïve Bayes* adalah metode pembelajaran Bayesian yang ditemukan sangat berguna dalam berbagai aplikasi. *Naïve Bayes* merupakan salah satu metode *supervised document classification*. Metode ini dikenal memiliki tingkat akurasi yang tinggi dengan perhitungan sederhana.

Penelitian sebelumnya terhadap problem yang sama yaitu diabetes menyebutkan, dalam penelitian I Putu Dody Lesmana (2012) Pengembangan Decision Tree J48 Untuk Diagnosis Penyakit Diabetes yaitu pada data pasien diabetes dengan jumlah 625, efektivitas metode di uji dengan 10-fold cross validation, dimana dari hasil pengukuran matrik confusion didapatkan akurasi sebesar 74,72%. Sedangkan Yavis Kurniawan (2015) dalam penelitiannya yaitu Deteksi Dan Klasifikasi Tingkat Keparahan Retinopati Diabetes Dengan Menggunakan Metode Klasifikasi K – Nearest Neighbor menyimpulkan Sistem ini optimal pada level 5 curvelets, orientasi sudut pada 32, nilai $k = 5$, dan cosine sebagai metode penghitung jarak pada K – Nearest Neighbor. Dari 40 data yang diujikan, didapat 26 gambar yang diklasifikasi dengan tepat, sehingga tingkat akurasi sistem menjadi 65% serta sistem membutuhkan waktu komputasi sebesar 3.54 detik untuk mengklasifikasikan satu data uji. Dari latar belakang yang telah diuraikan, maka penulis mengajukan penelitian dengan judul “Klasifikasi Penyakit Diabetes Pada Wanita Menggunakan Metode *Naive bayes*” guna menemukan metode terbaik dari perbandingan dengan penelitian terdahulu.

1.2. Rumusan Masalah

Bagaimana mengukur tingkat akurasi dan presisi terhadap klasifikasi penderita Diabetes pada wanita menggunakan metode *Naive bayes*?

1.3. Tujuan

Mengetahui hasil klasifikasi metode *naive bayes* terhadap penyakit diabetes pada wanita serta tingkat akurasi dan presisi pada klasifikasi penyakit diabetes terhadap wanita menggunakan metode *naive bayes*.

1.4. Manfaat

Pengembangan penelitian menggunakan *data mining* dalam mengklasifikasi pasien penderita diabetes.

1.5. Batasan Masalah

1. *Output* yang dihasilkan ada dua yaitu pasien menderita diabetes dan pasien tidak menderita diabetes.
2. Dataset yang digunakan dalam penelitian ini adalah berasal dari laman <https://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes> dengan atribut yaitu; *Glucose, Blood Pressure, Skin Thickness, Body Mass Indeks, Diabetes Pedigree Function*.
3. Dataset yang diteliti hanya terhadap wanita berkeluarga atau sudah menikah.
4. Usia pada wanita minimal 21 tahun.
5. Jumlah data yang digunakan 730 *record*. Data latih 584 *record* dan data uji 146 *record*.
6. Validasi data uji dan data latih menggunakan *K-Fold Cross Validation* dengan 5 skenario percobaan.