

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Pada era perkembangan teknologi yang sekarang ini semakin pesat kemajuannya, peningkatan aliran informasi kian meningkat dalam jumlah besar setiap harinya, peningkatan aliran pada informasi ini dapat menyebabkan terjadinya penumpukan kumpulan-kumpulan data berupa teks, baik secara data *online* maupun data *offline*. Dokumen teks yang menumpuk setiap harinya menyebabkan terjadinya kesulitan mencari dokumen yang sesuai dengan kebutuhan. Ketersediaan pada data yang berlimpah, kemudian dihasilkan dari penggunaan teknologi informasi yang saat ini di hampir semua bidang kehidupan menimbulkan suatu kebutuhan untuk dapat menggunakan dan memanfaatkan informasi dan pengetahuan yang dimana terdapat banyak kandungan dalam keseluruhan bentuk keseluruhan pada data tersebut, yang kemudian melahirkan *Data Mining*. *Data Mining* merupakan sebuah konsep yang diperuntukkan untuk menemukan pengetahuan atau informasi berharga yang tersembunyi di dalam *database* besar atau data dalam jumlah yang sangat besar. *Data Mining* memiliki beberapa teknik dalam menghasilkan informasi penting dari tumpukan data dan teknik-teknik yang berbeda, seperti deskripsi, estimasi, prediksi, klasifikasi, pengklusteran dan asosiasi. Penerapan *Data mining* banyak sekali digunakan dalam berbagai jenis bidang, seperti bidang kesehatan, pendidikan, pemerintahan, perdagangan dan lain-lain.

Diabetes Melitus (DM) merupakan penyakit kronis progresif yang ditandai dengan ketidakmampuan tubuh untuk melakukan metabolisme karbohidrat, lemak dan protein, mengarah ke hiperglikemia (kadar glukosa darah tinggi) [1]. Diabetes Melitus adalah gangguan metabolisme karbohidrat ketika suplai insulin tidak ada, tidak cukup, atau tidak efektif karena adanya resistensi insulin [2].

Klasifikasi data mengenai penyakit Diabetes Melitus dalam dunia medis merupakan salah satu tugas penting dalam mengklasifikasikan jenis penyakit, dan bahkan dapat membantu pihak dokter dalam mengambil keputusan untuk diagnosa mengenai penyakit tersebut, dengan demikian maka sangat penting untuk melakukan diagnosis secara cepat agar dapat membantu mempermudah dalam pengobatan penyakit pada pasien diabetes melitus. Etiologi atau penyebab Diabetes Melitus (DM) adalah genetika atau keturunan pada penderita diabetes, dan berdasarkan etiologi adalah Diabetes tipe 1 yang disebabkan oleh kerusakan sel beta di pancreas. Diabetes tipe 2 disebabkan oleh resistensi insulin, dan Diabetes Gestasional adalah Diabetes Melitus yang muncul pada saat hamil. Dari beberapa penjelasan yang ada tentang penyakit Diabetes Melitus maka pengklasifikasian dengan metode *Naïve Bayes* adalah tugas penting untuk dilakukan untuk membantu memberikan informasi serta wawasan tentang penyakit Diabetes bukan penyakit yang menular melainkan cenderung oleh faktor keturunan. Klasifikasi penyakit juga merupakan sistem kategori tempat yang dimana jenis penyakit dikelompokkan sesuai dengan kriteria yang sudah ditentukan. Dan juga pengklasifikasian bertujuan memudahkan pencatatan data mortalitas dan morbiditas, serta analisis, interpretasi dan perbandingan sistematis data tersebut.

Pada penelitian ini penulis memilih algoritma *Naïve Bayes*. Algoritma *Naïve Bayes* merupakan sebuah metode klasifikasi dengan menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan asal Inggris yaitu Thomas Bayes. Algoritma *Naïve Bayes Classifier* dapat memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri-ciri utama dari *Naïve Bayes Classifier* ini adalah asumsi yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian.

UPTD Puskesmas Pakuan Baru yang beralamat di Jalan Sudirman No.75, Tambak Sari, Kec. Jambi Selatan Kota Jambi merupakan salah satu Puskesmas di Kota Jambi. Puskesmas Pakuan Baru memiliki banyak pasien, setiap bulannya pasien di Puskesmas Pakuan Baru selalu datang membuat data histori pasien menjadi menumpuk di dalam data rekam medis, dan tidak ada tindak lanjut manfaat dari data-data tersebut yang tersedia. Padahal data-data pada Puskesmas Pakuan Baru dapat di manfaatkan dan diolah kembali untuk menjadi sebuah pengetahuan dan informasi yang bisa bermanfaat sebagai bahan pertimbangan dalam sebuah keputusan menentukan tingkat resiko pada penyakit yang diderita oleh pasien. Selama ini pasien hanya mendapatkan diagnosa pada jenis penyakitnya. Dengan masalah yang ada, penulis ingin mengklasifikasikan data pasien menggunakan teknik *data mining* untuk menghasilkan sebuah informasi. Teknik pada *data mining* yang digunakan pada penelitian ini adalah Klasifikasi dengan menggunakan metode *Naïve Bayes*, karena klasifikasi dapat berguna sebagai pengelompokkan data berdasarkan *data training* dan *data testing* dan dapat membuat perbandingan antara data yang lama dengan data yang baru. Serta metode *Naïve Bayes* yang akan

menghasilkan perhitungan serta nilai akurasi dari setiap hasil yang telah di klasifikasikan. Pada penelitian ini data yang akan diteliti berasal dari Puskesmas Pakuan Baru dengan jumlah data sebanyak 220 data pasien penderita Diabetes Melitus dari bulan juli sampai september 2022. Dan pada data tersebut akan dilakukan *pre-processing* yang akan melalui tahap *cleaning* data dan perhitungan data *testing* pada data pasien, serta atribut-atribut yang akan digunakan pada data pasien.

Beberapa penelitian sebelumnya telah dilakukan dengan menggunakan salah satu metode *Data Mining* yaitu metode *K-Means Clustering*. Salah satunya pada penelitian Akmal Rustandi dengan judul “Penerapan Metode *K-Means Clustering* Untuk Menentukan Penderita Diabetes” pada penelitian yang serupa tentang penyakit Diabetes penulis mendapatkan data melalui situs dan *platform kaggle*, *Dataset* ini diberikan oleh suatu perusahaan atau para peneliti, dengan suatu deskripsi masalah tertentu. Dan pada penelitian ini penulis menggunakan *dataset* penderita penyakit Diabetes dengan jumlah data 101766 dan diambil 100 data sebagai sampel. Presentasi hasil akurasi klasifikasi *clustering* dengan menghasilkan 5 kelompok penderita penyakit Diabetes [3]. Dan juga penelitian yang dilakukan oleh Andi Maulida Argina dengan menggunakan metode *K-Nearest Neighbor* pada judul “Penerapan Metode Klasifikasi *K-Nearest Neighbor* Pada Dataset Penderita Penyakit Diabetes” dan Kesimpulan yang didapatkan dari penelitian tersebut adalah perhitungan dengan menggunakan algoritma *K-Nearest Neighbor* (KNN) mendapatkan hasil akurasi tertinggi yaitu 39% pada $K=3$, presisi tertinggi yaitu

65% pada K=3, dan K=5, *recall* tertinggi yaitu 36% pada K=3 dan *F-Measure* tertinggi yaitu 46% pada K=3 [4].

Berdasarkan permasalahan latar belakang diatas, maka penulis tertarik untuk melakukan pengklasifikasian penyakit Diabetes Melitus di Puskesmas Pakuan Baru. Yang Penulis tuangkan dalam judul “**Analisis dan Penerapan Algoritma *Naïve Bayes* Untuk Klasifikasi Penyakit Diabetes Melitus (Studi Kasus : UPTD Puskesmas Pakuan Baru)**”.

1.2 RUMUSAN MASALAH

Berdasarkan dari latar belakang yang telah dijelaskan diatas, maka rumusan masalah pada penelitian ini adalah “Bagaimana cara menghitung dan menerapkan algoritma *Naïve Bayes* dalam mengklasifikasikan penyakit diabetes melitus?”.

1.3 BATASAN MASALAH

Agar penelitian ini dapat lebih terarah dan tidak menyimpang dari permasalahan, maka penulis menetapkan batasan masalah sebagai berikut :

1. Jumlah data dari Puskesmas Pakuan Baru yang diteliti berjumlah sebanyak 220 data dari bulan juli-september tahun 2022.
2. Metode yang digunakan adalah Algoritma *Naïve Bayes*.
3. Alat bantu analisa menggunakan *Microsoft Excel*, *Rapid Miner* dan *Google Colaboratory*, *Anaconda Navigator*.

1.4 TUJUAN PENELITIAN

Berdasarkan permasalahan yang telah dijelaskan diatas, maka dapat disimpulkan bahwa tujuan dari penelitian ini adalah :

1. Menghasilkan data klasifikasi penyakit Diabetes Melitus menggunakan algoritma *Naïve Bayes*.
2. Menerapkan keakurasian klasifikasi menggunakan metode *Naïve Bayes*.
3. Untuk memberikan informasi mengenai seputar penyakit Diabetes Melitus dan tipe-tipe dari penyakit Diabetes Melitus.

1.5 MANFAAT PENELITIAN

Adapun manfaat penelitian yang dapat diperoleh dalam penelitian adalah sebagai berikut :

1. Dapat memberikan informasi bagi pembaca untuk mengetahui apa itu penyakit diabetes, gejala pada diabetes dan tipe pada penyakit diabetes.
2. Bisa dijadikan wawasan dan menambah pengetahuan peneliti tentang bagaimana mencari informasi penting yang tersembunyi dalam suatu data menggunakan teknik pada algoritma klasifikasian.
3. Dapat menggunakan hasil analisis untuk menentukan langkah dalam meningkatkan kualitas pelayanan pasien khususnya pada tahun-tahun berikutnya.
4. Dapat digunakan sebagai acuan untuk melakukan penelitian berikutnya.

1.6 SISTEMATIKA PENULISAN

Untuk membahas lebih lanjut dan terperinci dalam penelitian ini guna memberikan gambaran terhadap pembaca, maka dengan ini penulis membagi atas beberapa bab yang saling berhubungan satu sama lainnya dan sesuai dengan ruang lingkup judul. Sistematika penulisannya antara lain sebagai berikut :

BAB I : PENDAHULUAN

Pada bab ini menjelaskan secara umum mengenai latar belakang masalah, rumusan masalah dengan batasan-batasan masalah yang digunakan, tujuan dan manfaat penelitian, serta sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab ini penulis mengutip dan menuangkan ide serta pendapat para pakar yang berhubungan dengan permasalahan yang penulis angkat. Teori-teori yang dipakai adalah mengenai definisi analisis, definisi data mining, definisi algoritma *Naïve Bayes* dan definisi *rapid miner* serta tentang bahasa pemrograman *python*.

BAB III : METODOLOGI PENELITIAN

Pada bab ini membahas tentang tahapan-tahapan seperti kerangka kerja penelitian, metode pengumpulan data serta alat-alat dan bahan-bahan pendukung untuk melakukan penelitian ini.

BAB IV : ANALISIS DAN INTERPRETASI

Pada bab ini menjelaskan tentang gambaran umum objek penelitian, serta menganalisis dan seleksi data pasien dengan melakukan *preprocessing* data. Transformasi data tersebut kemudian diolah dengan teknik *data mining* untuk mengetahui keakurasian data menggunakan metode klasifikasi algoritma *Naïve Bayes*. Hasil dari analisis yang telah dilakukan diinterpretasikan menjadi sebuah informasi mengenai status tipe diabetes melitus pasien.

BAB V : PENUTUP

Bab ini merupakan penutup dari penelitian ilmiah ini yang berisikan kesimpulan dari pembahasan bab-bab sebelumnya dan saran-saran yang berguna bagi pihak-pihak yang bersangkutan dalam penelitian ilmiah ini.