

## BAB I PENDAHULUAN

### 1.1 Pendahuluan

Diabetes adalah penyakit kronis yang menimbulkan gangguan metabolisme dan dikenali melalui kadar gula darah yang tinggi. Ini menjadi penyebab utama masalah jantung, kebutaan dan kegagalan ginjal. (Kementerian Kesehatan RI., 2020). Adapun teknik diagnosis diabetes yang sering dilakukan adalah dengan pengukuran tingkat glukosa dalam darah. Apabila konsentrasi glukosa dalam darah melampaui batas normal maka orang tersebut akan dikategorikan sebagai penderita diabetes. Namun pada penyakit diabetes terdapat fase asimtomatik yang cukup lama yaitu kondisi penyakit sudah terdeteksi tetapi belum menunjukkan gejala klinis pada pasien (Karyadiputra & Setiawan, 2022). Antisipasi secara dini merupakan hal yang tepat untuk dapat mencegah penyakit diabetes. Dengan memanfaatkan teknologi informasi dan komunikasi maka saat ini proses diagnosis secara dini penyakit diabetes dapat dilakukan dengan cara mengolah *dataset* rekam medis orang yang memiliki penyakit diabetes dengan parameter gejala umum dan tidak umum yang ditemukan pada fase awal sampai fase diagnosa dan hasil dari pengelolaan data tersebut bisa digunakan guna membuat suatu sistem klasifikasi penyakit diabetes (Anwar et al., 2022).

Alternatif yang diambil dalam penelitian ini adalah dengan menggunakan data diabetes untuk mengekstrak informasi atau pengetahuan baru darinya, yang selanjutnya dapat digunakan untuk membangun model klasifikasi pasien baru untuk mengetahui apakah pasien menderita diabetes. Proses mengekstraksi informasi baru dari sekumpulan data disebut data mining (Written et al., 2011).

Terdapat berbagai macam algoritma yang dapat digunakan dalam proses klasifikasi ini, seperti *Decision Trees*, *Naïve Byes*, *K-Nearest Neighbor*, dan lain-lain. Dalam penelitian ini, algoritma Naïve Bayes digunakan untuk mengklasifikasikan diabetes. Algoritma Naïve Bayes merupakan metode klasifikasi yang menggunakan probabilitas dan statistik sederhana berdasarkan teorema Bayes dengan asumsi

independensi yang kuat (Prasetyo, 2014). *Naïve Bayes* merupakan algoritma klasifikasi sederhana dimana setiap atribut bersifat independent dan dapat mempengaruhi keputusan akhir (Testiana, 2018). Algoritma *Naïve Bayes* dapat digunakan untuk memprediksi probabilitas milik kelas. Keunggulan dari algoritma *Naïve Bayes* yaitu algoritma yang sederhana dan mempunyai nilai performa yang baik (Siddik et al., 2020).

Untuk mendapatkan hasil terbaik pada penelitian ini dilakukan beberapa referensi terhadap beberapa penelitian terkait klasifikasi dengan menggunakan algoritma *Naïve Bayes*. Seperti pada penelitian (Doni et al., 2021) yang berjudul “Penerapan Data Mining untuk klasifikasi penyakit *Hepatocellular Carcinoma* Menggunakan Algoritma *Naïve Bayes*“. Penelitian tersebut bertujuan untuk mengklasifikasikan tingkat kemungkinan hidup pasien yang telah di diagnosis menderita penyakit *Hepatocellular Carcinoma* dengan menggunakan penerapan metode data mining serta melakukan pengukuran terhadap performa algoritma *Naïve Bayes* dengan mengacu kepada *Confusion Matrix* dan kurva ROC dan didapatkan hasil akurasi sebesar 70,30%, Presisi sebesar 73,53% dan *Recall* sebesar 77,32%. Penelitian lainnya dilakukan oleh (Riany & Testiana, 2023) yang berjudul “Penerapan Data Mining untuk Klasifikasi Penyakit Stroke Menggunakan Algoritma *Naïve Bayes*“. Pada penelitian ini bertujuan untuk mengklasifikasi data terkait penyakit stroke menggunakan algoritma *Naïve Bayes* untuk mengetahui apakah pasien tersebut menderita penyakit stroke atau tidak dan didapatkan hasil akurasi sebesar 92,48%.

Berdasarkan uraian konteks masalah diatas, penelitian ini bertujuan untuk mengklasifikasikan pasien berdasarkan menderita diabetes atau tidak menggunakan algoritma *Naïve Bayes* dengan menghitung probabilitas setiap atribut. Tujuan selanjutnya adalah untuk mengetahui akurasi, presisi, dan recall dari algoritma *Naïve Bayes*. Berdasarkan permasalahan yang telah dijelaskan diatas maka peneliti akan melakukan penelitian dengan judul “Klasifikasi Diabetes Menggunakan Algoritma *Naïve Bayes*”.

## 1.2 Perumusan Masalah

Adapun perumusan masalah dari penelitian ini adalah berapa tingkat akurasi Algoritma *Naïve Bayes* pada klasifikasi penyakit diabetes?

## 1.3 Tujuan

Tujuan dari penelitian ini adalah untuk menerapkan metode Algoritma *Naïve Bayes* dalam pengklasifikasian penyakit diabetes.

## 1.4 Manfaat

Penelitian ini mempunyai beberapa manfaat yaitu berikut :

1. Memperoleh wawasan dan pemahaman yang lebih mengenai optimasi kinerja algoritma *Naïve Bayes*.
2. Mahasiswa akan memahami konsep dan cara kerja dari algoritma *Naïve Bayes* dalam melakukan klasifikasi terhadap data
3. Memberikan kontribusi terhadap pengembangan pengetahuan mengenai suatu optimasi yang dapat meningkatkan tingkat keakuratan suatu pemodelan algoritma.
4. Dapat digunakan sebagai landasan atau bahan untuk melakukan penelitian lanjutan guna mencapai hasil yang lebih baik dalam hal pengetahuan.

## 1.5 Batasan Masalah

Adapun beberapa cakupan penelitian ini adalah sebagai berikut :

1. Penelitian memfokuskan pada mengimplementasikan model algoritma *Naïve Bayes* dalam mengklasifikasi data.
2. Data yang digunakan yaitu *dataset* penyakit diabetes yang berasal dari *Kaggle* (<https://www.Kaggle.com/code/prasadshingare/diabetes-hypertension-and-stroke-prediction/data>)