

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pembelajaran mesin (*machine learning*) dianggap sebagai cabang kecerdasan buatan, yang dianggap sebagai metode analisis data yang mengotomatiskan proses pembuatan model. Identifikasi pola yang mudah di antara data dan kemampuan untuk meningkatkan pembelajaran mesin dari waktu ke waktu adalah dua keunggulan utama pembelajaran mesin dibandingkan teknik klasifikasi data tradisional (Lemons, 2023). Membuat prediksi yang akurat mengenai tingkat keparahan hepatitis dan harapan hidup pasien sangat penting untuk pengelolaan dan pengobatan penyakit yang efektif. Hepatitis sulit didiagnosis, dan dokter perlu mengandalkan pengetahuan dan pengalaman mereka untuk mengambil keputusan yang tepat. Namun, penggunaan teknologi canggih, seperti pembelajaran mesin (*machine learning*) dan kecerdasan buatan (*artificial intelligence*), dapat membantu meningkatkan keakuratan diagnosis dan prognosis penyakit. (Yulhendri, 2023)

Data *Mining* merupakan proses kegiatan mengumpulkan sekumpulan data yang memiliki kapasitas yang besar untuk diolah sehingga menjadi informasi yang dapat digunakan. Data *Mining* juga merupakan proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang digunakan untuk keperluan kerja, prediksi dan meningkatkan keuntungan (Zidane, *et al.*, 2022). Kerangka kerja yang disajikan dibagi menjadi empat fase. Untuk tahap pengumpulan data penelitian, *database* Pima Indian Diabetes diubah untuk membuat prediksi diabetes. Visualisasi data digunakan pada tahap kedua untuk menyoroti proporsi orang yang memiliki masalah diabetes dan membuat data lebih mudah diinterpretasikan dengan menampilkannya dalam bentuk diagram batang. Tahap ketiga, pra-pemrosesan, terdiri dari menghilangkan *outlier* dan standarisasi

data. Fase klasifikasi, yang keempat, adalah diagnosis diabetes menggunakan berbagai teknik pembelajaran mesin. Penulis mengadopsi algoritma pembelajaran mesin karena kesederhanaan dan popularitasnya (Okikiola, *et al.*, 2023).

Algoritma Naïve Bayes adalah merupakan metode klasifikasi yang membagi suatu permasalahan ke dalam sebuah kelas label, menggunakan metode probabilitas dan statistik dengan persamaan dan perbedaannya sehingga hasilnya dapat di jelaskan (Nurjanah, *et al.*, 2023). Metode pengklasifikasi dianggap sebagai metode yang berpotensi lebih baik untuk mengklasifikasikan data dibandingkan metode klasifikasi lainnya dalam hal akurasi. Metode klasifikasi juga dipilih sebagai pengklasifikasi karena mudah dan cepat dalam memprediksi data. Oleh karena itu, berdasarkan latar belakang yang telah diuraikan, penulis berencana menggunakan metode klasifikasi dengan beberapa kelas, dengan mempertimbangkan akurasi berdasarkan penelitian sebelumnya, mempunyai potensi yang baik, mudah dan cepat dalam penerapannya, serta banyak digunakan untuk klasifikasi (Jefriyanto, *et al.*, 2023).

Menggunakan metode algoritma KNN dan Naïve Bayes dapat diterapkan pada berbagai objek dataset lainnya seperti menggunakan analisis sentimen film untuk mengkategorikan data sebagai negatif atau positif serta menentukan mana yang paling akurat. Hasil pengujian komparatif kedua algoritma menunjukkan bahwa algoritma Nave Bayes memiliki akurasi rata-rata yang lebih tinggi yaitu 99,63% dengan AUC sekitar 1.000, sedangkan algoritma KNN memiliki akurasi rata-rata yaitu 99,25% dengan AUC sebesar 1.000 (Muharrom, 2023). Kekurangan ini dapat dijembatani dengan pendekatan modern berbasis data *mining* melalui pembelajaran mesin seperti Naïve Bayes dan Tree. Pembelajaran mesin adalah metode analisis yang menggunakan informasi masa lalu, biasanya dalam bentuk data elektronik yang tersedia, untuk membuat prediksi yang akurat. Pada prinsipnya pendekatan ini dilakukan dengan memanen *big data*. Kemudian, *machine learning* akan menghasilkan perkiraan apakah desa wisata tersebut berhasil dikembangkan atau tidak. Selain itu, algoritma pembelajaran mesin seperti pohon keputusan dapat digunakan untuk menentukan faktor apa (faktor anteseden) yang menghasilkan pola keberhasilan pengembangan desa wisata (Ariyani, *et al.*, 2023).

Tugas pokok dari algoritma Naïve Bayes adalah untuk memprediksi kelas atau label dari suatu data yang diberikan, berdasarkan peluang dari fitur-fitur yang dimilikinya. Algoritma Naive Bayes dapat diartikan sebagai sebuah metode yang

tidak memiliki aturan, Naive Bayes menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi dengan cara melihat frekuensi tiap klasifikasi pada data training (Halila Nasution & Humaira, 2024).

Sebagai seorang praktisi Teknologi Informasi yang juga aktif di dunia Makeup Artis, penulis berinisiatif mempelajari dan menerapkan Algoritma Naive Bayes untuk membantu para pengguna aktif *skincare* dan kosmetik di Inti Makeup Store Padang dalam memilih dan menentukan produk yang cocok dan sesuai jenis kulit wajahnya. Berdasarkan latar belakang yang telah penulis jelaskan, penulis melakukan penelitian yang berjudul “Penerapan Metode Naive Bayes Untuk Memilih Produk *Skincare* Dan Kosmetik Yang Sesuai Dengan Jenis Kulit Wajah (Studi Kasus di Inti Makeup Store)”.

## 1.2 Perumusan Masalah

Berdasarkan uraian latar belakang masalah yang dijelaskan di atas, maka dapat dirumuskan pokok permasalahan dalam penelitian ini sebagai berikut :

1. Bagaimana cara mengelompokkan *skincare* dan kosmetik yang cepat dan tepat dengan menggunakan metode Naive Bayes?
2. Bagaimana cara meenerapkan metode Naive Bayes untuk menentukan pemilihan produk bagi pengguna *skincare* dan kosmetik dan pelanggan Inti Makeup Store?
3. Bagaimana penggunaan aplikasi *RapidMiner* pada metode Naive Bayes dapat menentukan pemilihan produk *skincare* dan kosmetik?

## 1.3 Batasan Masalah

Agar permasalahan dalam penelitian ini lebih terarah dan jelas, perlu adanya batasan masalah demi tercapainya tujuan yang diinginkan, maka peneliti mengambil batasan masalah dalam penelitian ini adalah :

1. Data sampel yang akan diolah data *skincare* dan kosmetik pada toko Inti Makeup Store.
2. Data yang diproses dalam penelitian ini menggunakan metode *Naive Bayes* berdasarkan kecocokan *skincare* dan kosmetik dengan jenis kulit wajah.
3. Pengujian data dalam penelitian ini menggunakan aplikasi *RapidMiner*.

#### 1.4 Tujuan Penelitian

Berdasarkan perumusan masalah dan batasan masalah di atas maka tujuan penelitian yang ingin dicapai dari penelitian ini adalah sebagai berikut :

1. Menerapkan metode *Naive Bayes* dapat menentukan pemilihan produk bagi pengguna *skincare* dan kosmetik dan pelanggan Inti Makeup Store.
2. Membantu pihak toko untuk menentukan stok produk *skincare* dan kosmetik yang paling cocok untuk jenis kulit wajah pelanggan Inti Makeup Store.
3. Membantu mengelompokkan *skincare* dan kosmetik sesuai dengan jenis kulit wajah yang paling diminati dan meminimalisir penumpukan barang.

#### 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Mampu memberikan saran pengambilan keputusan yang tepat untuk pengguna *skincare* dan kosmetik dalam pemilihan produk sesuai jenis kulit wajahnya.
2. Memberikan gambaran produk *skincare* dan kosmetik yang paling cocok dan diminati oleh pelanggan Inti Makeup Store.
3. Memberikan masukan penelitian lain tentang Data Mining dengan penerapan metode *Naive Bayes*.

## **1.6 Sistematika Penulisan**

Agar penelitian ini terarah dan sistematis, penulis membuat sistematika penulisan sebagai berikut :

### **BAB 1 PENDAHULUAN**

Bab ini berisikan pendahuluan dan latar belakang masalah dalam melakukan penelitian, gambaran umum tentang permasalahan yang akan dibahas pada bab-bab berikutnya, yaitu perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian yang diakhiri dengan sistematika penulisan.

### **BAB II LANDASAN TEORI**

Berisi pembahasan tentang konsep-konsep dan penerapan metode dari literatur jurnal, artikel, makalah, dan lain-lain yang berkaitan dengan penelitian agar dapat hasil yang didapatkan sesuai dengan.

### **BAB III METODOLOGI PENELITIAN**

Pada bab ini berisi tentang penjelasan kerangka kerja (*framework*) yang dipakai dalam penelitian ini sebagai berikut :

1. Pembelajaran Literatur Metode ini dilaksanakan dengan melakukan studi ke perpustakaan dengan membaca buku-buku maupun artikel-artikel yang dapat mendukung penulisan tesis ini.
2. Perancangan pada tahap ini dilakukan dengan mengambil sampel data *skincare* dan kosmetik untuk mengelompokkannya berdasarkan jenis kulit wajah.
3. Penyusunan laporan dan kesimpulan akhir membuat laporan hasil analisa perancangan kedalam format penulisan tesis dengan disertai kesimpulan akhir.

### **BAB IV ANALISA DAN PERANCANGAN**

Pada Bab ini diperlihatkan tahapan analisa dan pengolahan data sehingga didapat metode simulasi prediksi yang cepat, tepat, akurat dan mudah dalam penggunaannya.

**BAB V IMPLEMENTASI DAN HASIL**

Berisi tentang menguraikan cara implementasi dan hasil pengujian terhadap sistem yang dibangun.

**BAB VI KESIMPULAN DAN SARAN**

Berisi kesimpulan dan saran dari seluruh tahapan penelitian yang dilakukan.