

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Obesitas adalah suatu penyebab dari penyakit yang tidak menular yang dapat meningkatkan mortalitas dan Indonesia dengan persebaran terbanyak kesepuluh di dunia yang berisiko terkena penyakit kanker yang disebabkan oleh obesitas (Dwiky Riantoro & Ari Kristina, 2017). Faktor penyebab obesitas sangat banyak, salah satunya adalah masuknya asupan energi yang melebihi pengeluaran energi (Putri & Isti, 2015). Tingkat obesitas yang terjadi di Indonesia meningkat pesat pada masyarakat kalangan atas ataupun bawah sekalipun, dikarenakan mereka merubah pola makan tradisional ke produk hasil olahan yang memiliki lemak dan gula yang tinggi (Karana, 2021). Menurut data dari (Risikesdas, 2018) menyatakan bahwa pada tahun 2007 dibandingkan dengan 2018 memiliki angka yang cukup memprihatinkan, dimana obesitas pada orang dewasa pada tahun 2007 sebesar 10.5% terjadi peningkatan menjadi 21.8%, berat badan lebih pada tahun 2007 sebesar 8.6% terjadi peningkatan menjadi 13.6% dan obesitas sentral (lingkar perut) pada tahun 2007 sebesar 18.8% terjadi peningkatan menjadi 31%.

Dari hasil data diatas dapat dibuktikan bahwa penyakit obesitas berbahaya, maka penelitian ini dilakukan menggunakan data pasien penyakit obesitas yang diperoleh dari *online website* yaitu [www.kaggle.com](http://www.kaggle.com). Data dari pasien obesitas yang cukup banyak dalam melakukan pengolahan, akan dibantu lewat pemrosesan *data mining* menggunakan *machine learning* (Sulastri & Gufroni, 2017). Teknik tersebut

memiliki kaitan dengan ilmu pengetahuan teknologi seperti *machine learning*, *data statistic*, dan *artificial intelligence* (Nasir, 2021). Pada penelitian ini, dilakukan klasifikasi dalam guna memprediksi pasien obesitas sudah masuk ke tingkat yang berbahaya atau normal agar dapat dilakukan penanganan lebih awal.

Berdasarkan penelitian yang dilakukan oleh (Sibi & Widiarti, 2022) dalam melakukan klasifikasi terhadap penyakit obesitas dengan algoritma *K-Nearest Neighbor* (KNN) dengan nilai k yang bermacam-macam memberikan hasil akurasi k-1 sebesar 79.96%, k-3 sebesar 78.96%, dan k-5 sebesar 78.63%. Penelitian lainnya yang dilakukan oleh (Sitanggang & Sherly, 2022) dalam melakukan prediksi terhadap penyakit obesitas dengan algoritma *Support Vector Machine* (SVM) yang memberikan nilai akurasi sebesar 71.80%.

*Support Vector Machine* (SVM) merupakan teknik dalam menentukan *hyperplane* terbaik yang dapat memisahkan dua set data ke dua kelas dimensi yang berbeda. Kelebihan dari algoritma *Support Vector Machine* (SVM) yaitu menentukan suatu jarak akan lebih cepat dalam proses komputasinya, serta kekurangan dari algoritma tersebut yaitu cukup sulit digunakan dalam masalah dengan skala yang besar pada sampel yang diolah (Desiani et al., 2022). Algoritma *K-Nearest Neighbor* (KNN) merupakan suatu metode klasifikasi pada *data mining*, yang dimana algoritma tersebut mengklasifikasikan suatu *dataset* berdasarkan data pelatihan yang diberikan label, metode tersebut sangat umum digunakan dalam klasifikasi pada penelitian lainnya dan mudah digunakan dalam tahap pemrosesannya (Rifaldi et al., 2022).

Oleh sebab itu, penelitian ini dilakukan dengan tahap awal yaitu mengumpulkan data dari pasien obesitas yang diperoleh dari *online website Kaggle*, dan akan dilakukan pemrosesan klasifikasi algoritma *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM), dengan bantuan aplikasi *Orange Data Mining*. Hasil dari tahap tersebut akan dilakukan perbandingan tingkat akurasi pada kedua algoritma agar dapat mengetahui algoritma mana yang baik dalam mengolah data pasien obesitas.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, dapat ditarik rumusan dari permasalahan yang harus di selesaikan pada penelitian ini yaitu :

- 1) Bagaimana cara dalam menganalisis *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) pada klasifikasi penyakit obesitas?
- 2) Bagaimana tahap implementasi *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) pada klasifikasi penyakit obesitas?
- 3) Bagaimana hasil tingkat akurasi yang diberikan oleh *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) pada klasifikasi penyakit obesitas?

## 1.3 Tujuan Penelitian

Adapun berbagai tujuan dari dilakukannya penelitian ini adalah sebagai berikut :

- 1) Melakukan analisis serta prediksi menggunakan *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) pada klasifikasi penyakit obesitas.
- 2) Melakukan perbandingan hasil tingkat akurasi *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) pada klasifikasi penyakit obesitas.

#### **1.4 Batasan Masalah**

Adapun permasalahan yang akan dibatasi guna menghindari permasalahan yang akan meluas, batasan masalah dalam penelitian tersebut sebagai berikut :

- 1) *Dataset* yang digunakan merupakan data gejala penyakit obesitas yang didapatkan dari *Kaggle.com*.
- 2) *Format file* yang digunakan yaitu *Comma-Separated Values* (CSV).
- 3) Penelitian tersebut hanya menggunakan metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM).

#### **1.5 Manfaat Penelitian**

Manfaat yang diberikan pada penelitian ini yaitu sebagai berikut :

- 1) Memberikan sebuah pengetahuan cara klasifikasi penyakit obesitas dengan metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM).
- 2) Mengetahui hasil dari tingkat akurasi yang lebih unggul dalam metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) pada penyakit obesitas.

## 1.6 Sistematika Penulisan

Adapun sistematika pada penulisan yang terdiri dari 5 (lima) bab yang akan dijelaskan yaitu sebagai berikut :

BAB I : PENDAHULUAN, bab ini berisikan tentang penjelasan dari latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORI, bab ini berisikan tentang penjelasan penyakit obesitas, *dataset*, *machine learning*, teknik klasifikasi, *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM), dan penelitian terdahulu sebagai sebuah acuan pada penelitian agar tampak lebih terstruktur dan mudah dipahami.

BAB III : METODOLOGI PENELITIAN, bab ini berisikan mengenai kerangka pikir, metode pengumpulan data, tahapan penelitian, tahapan penyelesaian, prediksi metode *K-Nearest Neighbor* (KNN), prediksi metode *Support Vector Machine* (SVM), dan juga *confusion matrix*.

BAB IV : HASIL PEMBAHASAN, bab ini akan menjelaskan mengenai hasil dari keseluruhan cara kerja diatas seperti pengumpulan data, hasil penelitian, dan pembahasan klasifikasi penyakit obesitas dan akan dilakukan perbandingan hasil tingkat akurasi pada metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM).

BAB V : PENUTUP, bab ini akan menjelaskan kesimpulan dari hasil penelitian yang dilakukan diatas, saran dan keterbatasan dari penulis.