

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Di era globalisasi dan kemajuan teknologi informasi, penyakit jantung dan diabetes melitus telah menjadi dua masalah kesehatan utama yang dihadapi masyarakat modern. Menurut World Health Organization (WHO), penyakit jantung adalah penyebab kematian terbesar di dunia. Penyakit jantung merupakan salah satu ancaman bagi umat manusia, selain itu penyakit jantung juga dapat terjadi karena bawaan dari lahir dengan jenis dan tingkat keparahan yang beragam (D.F.R. Makarim, 2022). Identifikasi dini dan akurat terhadap kedua penyakit ini sangat penting untuk mengurangi angka morbiditas dan mortalitas.

Menurut Dr. Radityo yang dilansir oleh portal *Kompas.com* penyakit jantung menduduki peringkat pertama berdasarkan data terbaru Organisasi Kesehatan Dunia (WHO) dengan jumlah kematian akan terus meningkat yakni 20,5 juta orang di tahun 2020 dan 24,2 juta orang akan meninggal karena penyakit ini di tahun 2030. Penyakit ini tidak hanya mempengaruhi negara maju tetapi juga negara berkembang, dengan peningkatan kasus yang signifikan di kawasan Asia dan Afrika. Selain itu sekjen PP Perki juga Dr. Oktavia Lilyasari ikut menambahkan bahwa 80% penyakit jantung ini dapat dicegah, dengan melakukan perubahan pola makan dan pola hidup (D.Pitara, 2022).

Diabetes telah menunjukkan tren peningkatan yang mengkhawatirkan. Jumlah penderita penyakit diabetes meningkat dari tahun ke tahun, baik dari jumlah kasus maupun prevalensi. Pada tahun 2019, jumlah penderita diabetes di

dunia sudah mencapai 463 juta orang dan diprediksi akan terus bertambah mencapai angka 700 juta orang pada tahun 2045. Penderita diabetes mayoritas tinggal di negara berpenghasilan rendah dan menengah dan 1,6 juta kematian secara langsung disebabkan oleh diabetes setiap tahunnya. Hal ini menjadikan diabetes sebagai salah satu dari sepuluh penyakit penyebab utama kematian di seluruh dunia (IDF, 2020). Pada tahun 2019, Indonesia menempati posisi ke-7 di dunia, setelah China, India, Amerika Serikat, Pakistan, Brazil, dan Meksiko, sebagai negara dengan jumlah penderita diabetes tertinggi, dengan jumlah penderita sebesar 10,7 juta orang (J. Elflein, 2019). Pada tahun 2020, angka ini meningkat menjadi 10,8 juta dengan angka prevalensi pasien pengidap diabetes mencapai 6,2% dan diperkirakan jumlah penderita diabetes di Indonesia meningkat menjadi 16,7 juta pada tahun 2045 (H. Nurhayati-Wolff, 2020).

Metode diagnostik konvensional untuk penyakit seperti diabetes dan jantung sering kali menemui keterbatasan dalam keakuratan dan efisiensi, menjadikannya proses yang memakan waktu dan biaya yang signifikan. Tingginya prevalensi dan kompleksitas penyakit ini menuntut solusi yang lebih inovatif dan efektif. Dalam rangka mencari pendekatan yang lebih canggih, teknologi informasi dan *machine learning* muncul sebagai kandidat yang menjanjikan untuk memperbaiki proses diagnostik.

Machine learning, dengan beragam modelnya seperti Support Vector Machine (SVM) dan Random Forest (RF), telah membuka peluang baru dalam analisis data kesehatan. Kedua algoritma ini, tanpa memerlukan penjelasan teknis mendalam tentang cara kerjanya, dikenal karena kemampuan mereka dalam

mengidentifikasi pola dan hubungan yang kompleks dalam kumpulan data besar, termasuk data kesehatan. Keunggulan ini memungkinkan deteksi dini dan lebih akurat dari penyakit kritis seperti diabetes dan jantung, yang dapat secara signifikan mengurangi waktu dan biaya yang terkait dengan metode diagnostik konvensional. Namun, penting untuk membandingkan kedua algoritma ini untuk menentukan mana yang lebih efektif dalam kondisi dan jenis data tertentu. Dengan melakukan perbandingan, kita dapat menemukan kelebihan dan kekurangan masing-masing metode, sehingga kita dapat memilih metode yang paling sesuai untuk aplikasi klinis tertentu.

Pentingnya meningkatkan kesadaran akan penyakit kronis ini tidak bisa dianggap remeh. Dengan prevalensi yang terus meningkat, kebutuhan untuk mendeteksi dan mengelola diabetes dan penyakit jantung menjadi lebih kritis. Pendekatan berbasis *machine learning*, khususnya melalui penggunaan algoritma seperti SVM dan RF, dapat memberikan solusi dalam menghadapi tantangan ini. Dengan memanfaatkan *machine learning* ini mendapatkan diagnostik yang lebih cepat, akurat, dan efisien, sehingga memberikan hasil yang lebih baik bagi pasien dan sistem kesehatan di Indonesia.

Penelitian ini bertujuan untuk mengisi celah tersebut dengan membandingkan efektivitas SVM dan *Random Forest* dalam identifikasi penyakit jantung dan diabetes menggunakan *dataset Kaggle*. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan bagi komunitas medis, terutama dalam pengembangan alat diagnostik yang lebih akurat dan efisien.

1.2 Rumusan Masalah

Rumusan masalah yang menjadi fokus dalam penelitian ini adalah:

- 1) Bagaimana analisis model menggunakan algoritma SVM mampu mendeteksi penyakit jantung dan diabetes?
- 2) Bagaimana analisis model menggunakan algoritma RF mampu mendeteksi penyakit jantung dan diabetes?
- 3) Bagaimana hasil perbandingan dari analisis model dengan menggunakan algoritma *machine learning* SVM dan RF dalam mendeteksi penyakit jantung dan diabetes?

1.3 Tujuan Penelitian

Adapun yang menjadi tujuan dari penelitian ini, yaitu:

- 1) Menganalisis dan memprediksi menggunakan algoritma Support Vector Machine (SVM) dan Random Forest (RF) klasifikasi penyakit diabetes dan jantung.
- 2) Membandingkan tingkat akurasi algoritma Support Vector Machine (SVM) dan Random Forest (RF) pada klasifikasi penyakit diabetes dan jantung.

1.4 Batasan Masalah

Batasan permasalahan dari penelitian ini adalah sebagai berikut:

- 1) Algoritma statistik yang digunakan adalah Support Vector Machine (SVM) dan Random Forest.
- 2) *Dataset* yang akan digunakan merupakan data gejala penyakit diabetes dan jantung yang diperoleh dari *website Kaggle.com*.

1.5 Manfaat Penelitian

Manfaat yang diinginkan dari penelitian ini adalah:

1) Bagi Peneliti:

Hasil penelitian dapat membantu para peneliti dalam melakukan pengembangan topik dari penelitian ini sebagai bahan referensi.

2) Bagi Akademis:

Membantu para akademis dalam menggunakan metode-metode klasifikasi statistika agar mengetahui metode klasifikasi *machine learning* mana yang memiliki nilai akurasi tinggi

3) Bagi Pelayan Medis:

Hasil penelitian dapat dijadikan sebagai saran agar dapat berinovasi lagi dan pelayanan dapat ditingkatkan. Penelitian ini juga dapat membantu meningkatkan kualitas pelayan medis dalam mendeteksi penyakit jantung dan diabetes

4) Bagi Masyarakat:

Hasil penelitian dapat membantu masyarakat untuk mengetahui gejala-gejala awal penyakit jantung dan diabetes, sehingga penderita dapat melakukan tindakan pencegahan.

1.6 Sistematika Penulisan

Sistematika penulisan berisikan susunan sebuah laporan yang dibuat agar mempermudah memahami permasalahan yang dibahas. Dengan begitu, berikut merupakan sistematika penyusunan studi ini:

BAB I PENDAHULUAN

Dalam bab ini berisikan latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini berisikan tentang penjelasan penyakit diabetes dan jantung, *machine learning*, *data mining*, metode klasifikasi, Support Vector Machine (SVM), Random Forest (RF) dan penelitian terdahulu sebagai acuan di dalam penelitian agar lebih terstruktur dan mudah dipahami.

BAB III METODOLOGI PENELITIAN

Pada bab ini berisikan mengenai kerangka pikir, metode pengumpulan data, hipotesis, tahapan penelitian, tahapan penyelesaian, prediksi algoritma Support Vector Machine (SVM), prediksi algoritma Random Forest (RF), dan *confusion matrix*.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menjelaskan mengenai hasil pengumpulan data, hasil penelitian, dan pembahasan sebuah klasifikasi dari prediksi penyakit diabetes dan jantung serta melakukan perbandingan tingkat akurasi dari algoritma Support Vector Machine (SVM) dan Random Forest (RF)

BAB V KESIMPULAN DAN SARAN

Dalam bab ini berisikan tentang hasil akhir yang didapatkan dari hasil penelitian yang telah dilakukan, serta bagian akhir juga diberikan saran untuk topik penelitian bagi studi selanjutnya.