

BAB V

PENUTUP

5.1 KESIMPULAN

Berikut kesimpulan dari penelitian yang dilakukan terkait prediksi penyakit gagal jantung menggunakan metode K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM), serta visualisasi data yang digunakan dalam proses prediksi:

1. Metode Penelitian dan Pengolahan Data: Penelitian ini menggunakan dataset *Heart Failure Prediction* dari Kaggle, yang berisi atribut seperti usia, jenis kelamin, tekanan darah, dan lain-lain yang relevan dengan penyakit gagal jantung. Data ini telah diproses dan dibersihkan untuk menghilangkan noise, dengan total 918 data yang digunakan untuk pelatihan model.
2. Evaluasi dan Performa Model:
 - KNN: Model KNN menghasilkan akurasi yang baik dengan nilai precision sebesar 0.87 untuk klasifikasi yang tidak memiliki penyakit gagal jantung dan 0.91 untuk klasifikasi yang memiliki penyakit gagal jantung. F1-score yang dihasilkan sebesar 0.89 menunjukkan performa yang solid dalam memprediksi.
 - SVM: Model SVM juga memberikan hasil yang baik, meskipun sedikit lebih rendah dibanding KNN, dengan precision 0.80 dan 0.89 untuk klasifikasi yang tidak memiliki dan memiliki penyakit, serta F1-score 0.83.

3. Perbandingan Hasil:

- Melalui perbandingan menggunakan confusion matrix, dapat dilihat bahwa KNN cenderung memberikan hasil yang lebih baik dalam mengenali kasus positif (yang memiliki penyakit gagal jantung) dibandingkan SVM. Namun, keduanya memiliki kelebihan dan kelemahan masing-masing tergantung pada pola data yang ada.

4. Visualisasi Data:

- Data divisualisasikan menggunakan histogram, box plot, dan scatter plot untuk memahami distribusi dan hubungan antar variabel numerik. Hal ini membantu dalam mengeksplorasi data sebelum melangkah ke tahap pemodelan dan mengevaluasi hasil prediksi yang merupakan tujuan utama pada penelitian ini.

Berdasarkan hasil evaluasi, meskipun SVM memberikan hasil yang baik, namun KNN menunjukkan performa yang lebih superior dalam memprediksi penyakit gagal jantung pada dataset yang digunakan. Visualisasi data juga memberikan wawasan tambahan dalam pemahaman karakteristik dataset sebelum diterapkan pada model machine learning.