

BAB 1

PENDAHULUAN

1.1 LATAR BELAKANG

Jantung ialah salah satu organ dalam tubuh yang memiliki kewajiban untuk memompa dan menyalurkan darah yang mengangkut oksigen atau O₂ dan karbon dioksida atau CO₂ ke seluruh organ yang berada di dalam tubuh. Penyakit jantung biasanya disebut juga dengan penyakit kardiovaskular. Penyakit jantung yang tidak segera ditindak lanjuti oleh dokter akan menyebabkan terjadinya komplikasi yang mampu menghilangkan nyawa seseorang. Ketika jantung tidak mendapatkan suplai darah yang cukup untuk memompa darah, maka kinerja jantung akan menurun sehingga mengakibatkan terjadinya gagal jantung yang bisa terjadi secara tiba-tiba maupun bertahap [1].

Gagal jantung dapat diartikan juga sebagai kelainan struktur atau fungsi jantung yang mengakibatkan kegagalan jantung untuk menyebarkan oksigen ke seluruh tubuh. Gagal jantung adalah masalah kesehatan yang progresif dengan angka mortalitas dan morbiditas yang tinggi di negara maju maupun negara berkembang termasuk Indonesia. Di Indonesia, usia pasien gagal jantung relatif lebih muda dibandingkan dengan Eropa dan Amerika. *World Health Organization* (WHO) menyatakan bahwa meningkatnya angka gagal jantung di dunia disebabkan oleh meningkatnya jumlah manusia yang merokok, obesitas, diabetes, dan bertambahnya umur manusia [2].

Berdasarkan analisis dokter, prevalensi penyakit gagal jantung di Indonesia pada tahun 2013 sebesar 0,13% atau diprediksi sekitar 229.696 orang. Berdasarkan analisis dokter, perkiraan angka pasien penderita penyakit gagal jantung terbanyak berada di Provinsi Jawa Timur sebanyak 54.826 orang (0,19%), sedangkan Provinsi Maluku Utara mempunyai angka pasien penderita paling sedikit, yaitu sebanyak 144 orang (0,02%) [3].

Lalu berdasarkan [4] penyakit gagal jantung pada sistem informasi PTM berbasis *website* mencapai 4.161 kasus, dimana angka terbesar diperoleh

kelompok perempuan mencapai 2.247 kasus. Sedangkan menurut kelompok umur, analisis gagal jantung terbesar pada kelompok lanjut usia (umur ≥ 60 tahun) sebesar 1.880 kasus. Jumlah kasus dengan analisis gagal jantung yang dirawat inap di Rumah Sakit di Indonesia lebih banyak pada kelompok laki-laki yaitu sebesar 25.508 kasus daripada kelompok perempuan yaitu sebesar 24.507 kasus. Menurut umur, kasus gagal jantung terbanyak diperoleh oleh kelompok usia 45-64 tahun sebesar 24.283 kasus. Sedangkan jumlah kasus meninggal akibat gagal jantung sebanyak 4.996 orang. Jumlah kasus gagal jantung yang dirawat inap di Rumah Sakit di Indonesia berdasarkan provinsi, terbanyak diperoleh oleh Provinsi Jawa Tengah sebanyak 8.658 kasus [4].

Oleh karena itu dibutuhkan upaya terbaru untuk mendeteksi awal pasien gagal jantung sehingga dapat mengurangi angka kematian dan memperpanjang hidup pasien. Dengan bantuan teknologi terkini, data riwayat pasien yang mengalami gagal jantung bisa diolah sedemikian rupa hingga membentuk pola korelasi antara data terbaru dengan data riwayat pasien yang mengalami gagal jantung. Data-data tersebut bisa digunakan untuk membuat suatu program yang mampu mengetahui kelangsungan hidup pasien kedepannya. Penelitian ini sudah pernah dilakukan oleh peneliti-peneliti sebelumnya dengan menggunakan algoritma klasifikasi yang berbeda-beda diantaranya *Random Forest*, *Naive Bayes*, *Support Vector Machine* dan *K-Nearest Neighbor* dengan tingkat keakurasian tertinggi diraih oleh *Random Forest* sebesar 0.7947. Dalam penelitian ini menggunakan SMOTE+ENN untuk menyeimbangkan data serta menggunakan 4 fitur diantaranya *ejection_fraction*, *serum_creatinine*, *age* dan *smoking* [5]. Dalam penelitian lain pernah menggunakan algoritma klasifikasi yang berbeda-beda pula diantaranya *Random Forest*, *Decision Tree*, *K-Nearest Neighbor*, *Support Vector Machine*, *Artificial Neural Network*, dan *Naive Bayes* dengan tingkat keakurasian tertinggi diraih oleh *Random Forest* sebesar 94.31%. Dalam penelitian ini menggunakan semua fitur tanpa adanya *Feature Importances* untuk mengetahui fitur mana sajakah yang paling berpengaruh untuk mengklasifikasikan pasien gagal jantung [6].

Berdasarkan pembahasan sebelumnya, maka untuk tugas akhir ini akan dibuat program klasifikasi pasien gagal jantung dengan menggunakan algoritma

klasifikasi *Random Forest*. Dalam penelitian ini peneliti menggunakan *Feature Importances* untuk mengetahui fitur mana sajakah yang sangat mempengaruhi pengklasifikasian pasien gagal jantung, *feature scaling* untuk menormalisasi data yang menyimpang jauh, dan data akan *di-sampling* dengan menggunakan algoritma *Adaptive Synthetic* (ADASYN) untuk menyeimbangkan kelas minoritas dengan kelas mayoritas. Data-data tersebut juga nantinya akan di uji dengan menggunakan *K-Fold Cross Validation* dengan $k=10$ dan *Confusion Matrix*.

1.2 RUMUSAN MASALAH

Rumusan masalah dari penelitian ini adalah:

1. Bagaimana hasil *Feature Importances*, *feature Scaling*, dan *Adaptive Synthetic* (ADASYN) dalam *preprocessing* data?
2. Bagaimana hasil akurasi, *precision*, *recall*, dan *f1-score* dalam pengklasifikasian pasien gagal jantung dengan menggunakan algoritma *Random Forest*?

1.3 BATASAN MASALAH

Batasan masalah dari penelitian ini adalah:

- 1) Program klasifikasi pasien gagal jantung hanya mengambil dataset berupa *file csv* dari situs [Heart failure prediction / Kaggle](#) [7].
- 2) Menggunakan *Random Forest* sebagai pemodelan.
- 3) Data dipisah menjadi 10 *folds* dengan menggunakan *K-Fold Cross Validation* dan dari 10 *folds* tersebut setiap *folds* akan dibagi menjadi dua yaitu data *training* dan data *testing*.
- 4) Menggunakan algoritma *Adaptive Synthetic* (ADASYN) untuk menyeimbangkan kelas minoritas dan mayoritas.
- 5) Menggunakan *Feature Importances* untuk mengetahui fitur mana sajakah yang paling berpengaruh dalam pengklasifikasian pasien gagal jantung.
- 6) Menggunakan *Min Max Normalization* untuk menormalisasi data yang menyimpang.

- 7) Parameter yang digunakan hanya nilai akurasi, *precision*, *recall*, dan *f1-score*.
- 8) Bahasa pemrograman yang digunakan adalah *Python*.
- 9) *Tools* yang digunakan hanya *google colab*.

1.4 TUJUAN

Tujuan dari penelitian ini adalah:

1. Mengetahui hasil *Feature Importances*, *feature Scaling*, dan Adaptive Syntethic (ADASYN) dalam *preprocessing* data.
2. Mengetahui nilai akurasi, *precision*, *recall*, dan *f1-score* dalam pengklasifikasian pasien gagal jantung dengan menggunakan algoritma *Random Forest*.

1.5 MANFAAT

Penelitian ini diharapkan dapat membantu masyarakat untuk dapat mengembangkan gaya hidup sehat dan mempermudah tenaga kesehatan dalam mendeteksi awal pasien yang mengalami gagal jantung sehingga tenaga kesehatan dapat melakukan upaya awal untuk memperpanjang hidup pasien dan mengurangi angka kematian akibat gagal jantung dengan menggunakan program klasifikasi pasien gagal jantung. Selain itu supaya memberikan gambaran klasifikasi yang dilakukan oleh pemodelan *Random Forest* dan hasilnya berdasarkan kondisi pasien. Diharapkan pula dapat menjadi referensi peneliti lainnya yang ingin mengulas topik terkait.

1.6 SISTEMATIKA PENULISAN

Sistematika penulisan penelitian ini dibagi menjadi 5 bagian:

1. BAB 1: PENDAHULUAN

Bagian pendahuluan berisi mengenai latar belakang, rumusan masalah yang diangkat, batasan masalah, manfaat, tujuan penelitian, dan sistematika penulisan.

2. BAB 2: DASAR TEORI

Bagian dasar teori membahas mengenai pengertian *Artificial Intelligence*, *Machine Learning*, *Random Forest*, *Min Max Normalization*, *Adaptive Synthetic*, dll.

3. BAB 3: METODE PENELITIAN

Bagian metode penelitian membahas mengenai sumber data, *tools* yang akan digunakan, alur penelitian, alur perancangan dan pengujian program.

4. BAB 4: HASIL DAN PEMBAHASAN

Bagian hasil dan pembahasan membahas mengenai parameter penelitian dan analisis hasil penelitian yang terdiri dari 2 sub bab yaitu analisis *preprocessing* data dan analisis hasil klasifikasi.

5. BAB 5: PENUTUP

Bagian penutup membahas mengenai kesimpulan yang didapatkan dari penelitian yang dilakukan dan saran untuk penelitian.