

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Penyakit kardiovaskular adalah penyebab utama kematian secara global, diambil dari data WHO diperkirakan 17,9 juta manusia meninggal setiap tahunnya [1]. Menurut WHO lebih dari empat dari lima kematian kardiovaskular disebabkan oleh serangan jantung dan stroke, dan sepertiga dari kematian berusia dibawah 70 tahun [1]. Penyakit kardiovaskular adalah sekelompok penyakit jantung dan pembuluh darah yang dimana penyakit kardiovaskular di golongkan sebagai penyakit sistem sirkulasi darah dan termasuk penyakit tidak menular. Penyakit kardiovaskular meliputi : penyakit jantung koroner (*coronary heart disease*), penyakit serebrovaskular (*cerebro-vascular disease*), penyakit arteri perifer (*peripheral arterial disease*), penyakit jantung rematik (*rheumatic heart disease*), penyakit jantung bawaan (*congenital heart disease*), trombosis vena dalam (*deep vein thrombosis*) dan emboli pulmonal (*pulmonary embolism*)[2].

Menurut data dari kemenkes kematian di Indonesia yang disebabkan oleh penyakit kardiovaskular mencapai 651.481 penduduk per tahunnya [3]. Meliputi penyakit jantung koroner mencapai 243.343 kematian setiap tahunnya [3], dan biaya pelayanan kesehatan yang dikeluarkan akibat dari penyakit kardiovaskular pun tergolong cukup besar. Sebesar Rp. 8.2 Triliun untuk biaya pelayanan kesehatan dari penyakit kardiovaskular yang diambil dari data BPJS [3].

ECG (Electrocardiogram) merupakan alat yang penting untuk memantau kesehatan jantung. ECG sendiri merupakan sinyal yang digunakan untuk memantau akan kesehatan jantung namun penganalisis sinyal ECG secara manual sering kali memakan waktu dan rawan kesalahan. Hal ini menjadi tantangan besar, dikarenakan penyakit kardiovaskular merupakan penyebab sepertiga kematian di seluruh dunia. Banyak orang mengalami gangguan irama jantung (aritmia) yang bisa berakibat fatal. Oleh karena itu deteksi artimia secara akurat dan otomatis sangat dibutuhkan agar diagnosis bisa dilakukan dengan cepat dan efisien [4].

Identifikasi penyakit pada gangguan irama jantung (aritmia) saat ini masih dilakukan dengan metode tradisional, yaitu melalui pengamatan langsung terhadap signal ECG. Metode ini menyebabkan identifikasi penyakit menjadi tidak efisien, dan rentan terhadap kesalahan diagnosis. Di era yang semakin maju perkembangan akan teknologi, terutama di bidang *artificial intellegent*, terdapat peluang untuk meningkatkan efisiensi dalam identifikasi penyakit jantung melalui data yang berbentuk sinyal ECG (*Electrocardiogram*). Penggunaan teknologi *artificial intelegent* seperti *deep learning*, menjadi relevan untuk mengatasi permasalahan ini. Dengan penerapan teknologi *deep learning*, proses identifikasi penyakit dapat dilakukan lebih cepat dan akurat.

Deep Learning merupakan bagian dari *machine learning* yang algoritmanya sering digunakan dikarenakan struktur dan cara kerja nya seperti struktur otak manusia atau lebih di kenal sebagai jaringan syaraf tiruan ANN (*Aritficial Neural Network*). Akibat dari hal itu algoritma ini sangat cocok digunakan untuk pemrosesan video, gambar, maupun teks [5]. RNN (*Recurrent Neural Network*) adalah jaringan saraf buatan yang sangat cocok dengan data berurutan. RNN memiliki koneksi berulang yang memungkinkannya mempertahankan status internal, yang secara efektif memungkinkan jaringan untuk menyimpan memori masukkan sebelumnya [6]. CNN (*Convolutional Neural Networks*), unggul dalam tugas terkait gambar maupun data berbentuk signal dikarenakan CNN menangkap *spatial hierarchy* melalui *convolution layer*, *Activation ReLU*, *pooling layer*, *loss function*, dan *fully connected layer* [7]. GRU (*Gated Recurrent Units*) adalah varian LSTM yang disederhanakan yang secara efisien menangkap ketergantungan diberbagai skala waktu [8]. LSTM (*Long Short Term Memory*) adalah jenis RNN yang dirancang menangkap ketergantungan jangka panjang dalam data berurutan. Setiap Sel LSTM mempertahankan status memori yang diatur oleh tiga *gate* [8]. Penelitian yang dilakukan oleh P.E Martin dan lain lain, yang berjudul “*Three-Stream 3D/1D CNN for Fine-Grained Action Classification and Segmentation in Table Tennis*” didapatkan hasil pada model CNN dengan menggunakan *three-stream 3D/1D*, dengan *accuracy training* sebesar 95.8%, *accuracy validation* sebesar 86.5% lalu *testing* sebesar 87.3% [9]. Pada penelitian yang dilakukan oleh Brett Naul dan lain lain, yang berjudul “*A recurrent neural network for*

classification of unevenly sampled variable stars” menggunakan model arsitektur dari model RNN dan GRU Didapatkan hasil *accuracy* rata-rata sebesar 98.6% untuk dataset dari *ASAS* dan juga 96.7% untuk dataset dari *Linear* [10]. Pada penelitian yang dilakukan oleh Alex Graves yang berjudul “*Generating Sequences With Recurrent Neural Networks*” menggunakan model LSTM dalam melakukan *Text Prediction* didapatkan hasil rata – rata *classification error rate* sebesar 26.9 % untuk *character-level*, sedangkan untuk *word-level* sebesar 76.2% [11].

Dataset yang digunakan dalam penelitian ini diambil dari Kaggle dengan judul “*ECG Heartbeat Categorization Dataset*” yang juga di tulis di dalam penelitian yang berjudul “*ECG Heartbeat Classification: A Deep Transferable Representation*” yang ditulis oleh M. Kachuee dan lain lain. Pada penelitian ini peneliti melakukan *mapping* tiap kelas dengan kategori *Normal, Atrial Premature, Premature Ventricular Contraction, Fusion of Ventricular and Normal, Fusion of paced and normal*. Memiliki frekuensi sebesar 125 Hz untuk input , dan memiliki panjang sampel signal sebesar 1800 *ms* atau 1,8 detik [4].

Adanya penyakit pada manusia seperti penyakit jantung, penggunaan teknologi seperti *deep learning* menjadi relevan dalam melakukan klasifikasi akan penyakit jantung. Penelitian ini bertujuan untuk menerapkan 4 model *deep learning* dalam melakukan klasifikasi akan penyakit jantung. Menggunakan dataset dari *MIT-BIH Arrhythmia* yang diambil dari *kaggle* dengan judul “*ECG Heartbeat Categorization Dataset*” dan ditulis di dalam penelitian yang berjudul “*ECG Heartbeat Classification: A Deep Transferable Representation*” yang ditulis oleh M. Kachuee dan lain lain [4]. Oleh karena itu, penulis menerapkan metode *deep learning* dalam penelitian yang berjudul “Klasifikasi Penyakit Jantung Berdasarkan Nilai Amplitudo ECG Signal Menggunakan Metode Deep Learning”. Diharapkan penelitian ini memberikan kontribusi atau manfaat signifikan untuk perkembangan teknologi, sebagai referensi untuk analisis performa model *deep learning*.

1.2 Rumusan Masalah

Berdasarkan penjelasan pada latar belakang sebelumnya rumusan masalah yang didapatkan pada penelitian ini. Model *deep learning* seperti CNN, RNN, LSTM, dan GRU dari keempat model itu bagaimana performa model dalam

melakukan klasifikasi gangguan irama jantung berdasarkan dataset signal ECG dari *MIT-BIH Arrhythmia* yang diambil dari *kaggle* dengan judul “*ECG Heartbeat Categorization Dataset*” dan ditulis di dalam penelitian yang berjudul “*ECG Heartbeat Classification: A Deep Transferable Representation*” yang ditulis oleh M. Kachuee dan lain lain [4]?

1.3 Tujuan dan Manfaat

Berdasarkan penjelasan pada latar belakang sebelumnya, tujuan dari penelitian ini yaitu menganalisis dan mengevaluasi performa model *Deep Learning* yaitu CNN, GRU, LSTM, dan RNN dalam melakukan klasifikasi sinyal ECG berdasarkan *dataset MIT-BIH Arrhythmia*.

Adapun manfaat yang diharapkan dari penelitian ini adalah penelitian ini diharapkan memberikan kontribusi atau manfaat signifikan untuk perkembangan teknologi, sebagai acuan perbandingan model *deep learning*.

1.4 Batasan Masalah

Penelitian ini memiliki beberapa batasan untuk memperjelas ruang lingkup permasalahan yang dibahas. Pertama, data yang digunakan adalah sinyal *electrocardiogram* (ECG) dari dataset “*ECG Heartbeat Categorization Dataset*” yang tersedia di *platform Kaggle* dan ditulis di dalam penelitian yang berjudul “*ECG Heartbeat Classification: A Deep Transferable Representation*” yang ditulis oleh M. Kachuee dan lain lain [4]. Kedua, sinyal yang digunakan memiliki panjang 1800 *ms* dengan frekuensi *sampling* sebesar 125 *Hz*. Ketiga, klasifikasi irama jantung yang dilakukan hanya terbatas pada lima kelas, yaitu *Normal*, *Atrial Premature*, *Premature Ventricular Contraction*, *Fusion of Ventricular and Normal*, dan *Fusion of Paced and Normal*.

1.5 Metode Penelitian

Metode penelitian dalam studi ini dirancang berdasarkan rumusan masalah yang telah ditentukan sebelumnya, dengan mempertimbangkan batasan sebagai berikut :

1. Perumusan Masalah

Perumusan masalah adalah langkah awal dalam penelitian yang berujuan untuk mengidentifikasi fokus penelitian merumuskan latar belakang, Menyusun rumusan masalah, menetapkan tujuan penelitian, menjelaskan manfaat penelitian, dan juga menentukan batasan penelitian.

2. Studi Literatur

Studi literatur bertujuan menganalisis penelitian terdahulu terkait metode-metode *deep learning* dalam melakukan klasifikasi, membandingkan teknik yang ada, dan mengevaluasi kelebihan serta kelemahan metode tersebut untuk menemukan hasil analisis baru yang dapat memberikan kontribusi baru dalam penelitian ini.

3. Pengumpulan Dataset

Tahap pengumpulan data bertujuan untuk memperoleh dataset signal ECG akan penyakit jantung yang diperlukan untuk penelitian ini yang didapatkan dari *kaggle*.

4. EDA (Exploratory Data Analysis)

Pada tahap EDA (*Exploratory Data Analysis*) bertujuan untuk memahami struktur data dari dataset, dan melihat distribusi kelas pada dataset dari *MIT-BIH Arrythmia*. Serta melihat panjang tiap sampel signal, dan isi dari dataset itu sendiri.

5. Preprocessing Dataset

Tahap *Preprocessing dataset* bertujuan untuk dilakukan nya *label encoding*. *label encoding* mengubah label kelas yang berbentuk teks atau kategori menjadi representasi angka (numerik), lalu dilakukannya *splitting dataset*. Untuk *training*, *validation*, dan *testing*. Serta melihat distribusi kelas setelah dilakukannya *splitting dataset*.

6. Pembuatan Model

Pada tahap training model juga dilakukan pembuatan model yang bertujuan untuk membuat model *deep learning*, seperti CNN, RNN, LSTM, dan GRU, dimana mengatur dari segi dimesi *convolutional*, *kernel size*, *stride* dan juga *padding*, untuk model CNN. Sedangkan untuk model RNN, LSTM, dan GRU diatur dari segi

input_size, *hidden_size*, *layer_size*, dan *output_size*, untuk mencapai hasil yang optimal.

Tahap *training* model bertujuan untuk melatih 4 model dengan *dataset training* dan *validation* yang telah dilakukan pada tahap *preprocessing dataset*, dan mengatur *parameter* seperti *epoch*, *batch size*, dan *learning rate* untuk mencapai hasil yang optimal.

7. *Testing Model*

Tahap *testing* model dilakukan untuk mengevaluasi akurasi dan performa model menggunakan data *testing*, setelah dilakukan nya pada tahap training model. Hasil *Testing* model digunakan untuk menilai kemampuan model dalam melakukan klasifikasi akan signal ECG akan irpenyakit jantung.

8. Analisis Hasil Model

Tahap Analisis hasil modle dilakukan untuk mengevaluai kinerja tiap model dalam melakukan klasifikasi signal ECG akan penyakit jantung. Pada tahap ini melibatkan perhitungan metrik seperti *precison*, *recall*, *F1-Score* serta *Accuracy* untuk masing-masing model. Hasil ini digunakan untuk mengindentifikasi kinerja model dan juga membandingkan dengan 4 model lainnya yang telah dilakukan pada tahap *testing model*.