

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

*Stroke* adalah gangguan fungsi saraf yang disebabkan oleh gangguan aliran darah ke otak, yang dapat terjadi secara tiba-tiba dalam hitungan detik atau cepat dalam beberapa jam, hal ini disebabkan oleh terganggunya aliran pembuluh darah yang berada di otak [1]. Berdasarkan hal tersebut *stroke* dapat diklasifikasikan menjadi 2 jenis yaitu *stroke iskemik* dan *stroke hemoragik* yang keduanya dapat menghentikan suplai darah ke otak dan dapat menyebabkan munculnya gejala kematian jaringan otak [2].

Penyebab *stroke iskemik* diakibatkan adanya penyumbatan pada pembuluh darah otak, sering terjadi pada individu lanjut usia, tetapi tidak menutup kemungkinan terjadi pada individu usia di bawah 45 tahun, terdapat kurang dari 5% kasus *stroke* dialami oleh usia muda [3]. Sedangkan penyebab *stroke hemoragik* terjadi karena adanya pembuluh darah otak yang pecah, penderitanya lebih banyak terjadi pada seseorang yang berjenis kelamin laki-laki dibandingkan dengan perempuan. Hal tersebut karena adanya perbedaan faktor komponen darah dan hormonal antara laki-laki dan perempuan yang diduga berperan penting [4].

Faktor- faktor seperti usia, hipertensi, penyakit jantung, status perkawinan, jenis pekerjaan, jenis tempat tinggal, rata-rata kadar glukosa, BMI, status merokok, dan jenis kelamin dapat menyebabkan terjadinya *stroke*. Namun, hipertensi adalah salah satu faktor yang secara signifikan meningkatkan risiko *stroke* [5]. Hipertensi merupakan suatu keadaan dimana tekanan darah mengalami peningkatan, peningkatan hipertensi *systole* lebih dari 20 mmHg dan tekanan darah *diastole* lebih dari 10 mmHg mampu memicu terjadinya penyakit *stroke* [3]. Penderita *stroke* biasanya mengalami gejala seperti kelemahan anggota gerak sampai kelumpuhan, hilangnya sensasi di wajah, bibir tidak simetris, kesulitan berbicara (*afasia*) atau pelo, kesulitan menelan, penurunan

kesadaran, nyeri kepala, vertigo, mual muntah dan hilangnya penglihatan di satu sisi atau dapat terjadi kebutaan [2].

Organisasi Kesehatan Dunia (WHO) pada tahun 2019 mengungkapkan bahwa risiko seumur hidup terkena *stroke* telah meningkat sebesar 50% selama 17 tahun terakhir dan kini 1 dari 4 orang diperkirakan terkena *stroke* seumur hidupnya. Terjadi peningkatan kejadian *stroke* sebesar 70%, peningkatan kematian akibat *stroke* sebesar 43%, peningkatan prevalensi *stroke* sebesar 102%, dan peningkatan *Disability Adjusted Life Years* (DALY) sebesar 143%. Hal yang paling mencolok adalah sebagian besar beban *stroke* global (86% kematian akibat *stroke* dan 89% DALY) terjadi di negara-negara berpendapatan rendah dan menengah ke bawah [6]. Berdasarkan data dari *World Health Organization* (WHO) pada tahun 2020, terdapat sekitar 27.000 kasus *stroke* yang melibatkan sekitar 25.400 orang dari jumlah 100.000 penduduk, menderita *stroke* yang telah menurun lebih dari 40 % selama 15 tahun terakhir, dan sekitar 6.100 orang meninggal pada tahun 2020 dibandingkan tahun 2019 [7]. Kecacatan akibat *stroke* dapat berupa kecacatan jangka panjang, dengan lebih dari 89% pasien tidak dapat berfungsi secara mandiri dalam kehidupan sehari-hari dan 11% pasien *stroke* tidak dapat berjalan secara mandiri [8].

Di Indonesia kasus *stroke* menempati urutan kedua penyebab kematian, data yang dihimpun Riskesdas tahun 2019 menunjukkan angka kejadian *stroke* semakin bertambah dari 7% menjadi 10.9 per 1.000 penduduk Indonesia. Berdasarkan temuan Riskesdas, prevalensi *stroke* di Indonesia meningkat seiring bertambahnya usia. Sebagian besar kasus *stroke* yang didiagnosis oleh tenaga kesehatan berusia di atas 75 tahun (43,1%), dengan terendah pada kelompok usia 15-24 tahun sebesar 0,2%. Prevalensi *stroke* berdasarkan gender lebih tinggi pada pria (7,1%) dibandingkan wanita (6,8%). Prevalensi *stroke* lebih tinggi di perkotaan (8,2%) dibandingkan di pedesaan (5,7%). Masalah penyakit *stroke* di Indonesia memerlukan perhatian yang serius karena jumlah kasus yang terus meningkat dan mempunyai angka kematian yang tinggi [8].

*Stroke* membutuhkan perawatan segera, yang sangat dipengaruhi oleh deteksi dini yang tepat untuk mendapatkan penanganan medis yang optimal

adalah 3-6 jam setelah pasien *stroke* pertama kali ditemukan. Penanganan medis lebih dari 12 jam setelah *stroke* terjadi berisiko menyebabkan cacat permanen yang lebih besar [9]. Diagnosis yang cepat dan hemat biaya sangat penting dalam mengurangi komplikasi pasca *stroke* [10].

Minimnya pengetahuan masyarakat masih belum sepenuhnya memahami sifat dari kondisi *stroke* saat ini, dan banyak yang tidak menyadari bahwa tanda-tanda awal yang mungkin ada. Selain itu, kebanyakan masyarakat enggan pergi ke rumah sakit hanya untuk menanyakan gejala yang mereka alami. Sehingga menyebabkan peningkatan pesat dalam frekuensi *stroke* dan menghantui kehidupan Masyarakat [5].

Oleh karena itu pentingnya mengetahui gejala penyakit *stroke* sejak dini merupakan pencegahan awal terhadap karakteristik penyakit *stroke*. Maka dari itu, dilakukan penelitian untuk menganalisa data terkait dengan penyebab *stroke*. Deteksi dini dapat dilakukan melalui perancangan sebuah model pendekatan yang dapat digunakan untuk mengidentifikasi serta melakukan prediksi terhadap risiko *stroke*. Teknik *Machine Learning* merupakan model analisis data yang dapat digunakan untuk memprediksi penyakit *stroke* [7].

Memfaatkan perkembangan di bidang teknologi kecerdasan buatan (*AI*) salah satunya yaitu *Machine Learning* semakin disesuaikan, karena dengan pembelajaran mesin dapat mempermudah dalam prediksi dalam melakukan penanganan penyakit serta dapat mengurangi pembiayaan [11]. Sehingga dibuat pemodelan pembelajaran mesin yang mampu bekerja secara otomatis agar digunakan secara langsung dalam membantu pasien untuk menekan risiko penyakit *stroke*. Salah satu algoritma *Machine Learning* adalah *Random Forest* [12].

Algoritma *Random Forest* merupakan salah satu algoritma *Machine Learning* yang termasuk dalam kategori *Ensemble Learning*, dimana model terdiri dari sejumlah besar pohon keputusan (*Decision Trees*) yang bekerja secara bersama-sama. Keunggulan *Random Forest* antara lain kemampuannya dalam menangani data yang kompleks dan besar, menghindari *overfitting*, serta memberikan estimasi yang stabil terhadap variabel yang paling berpengaruh

dalam prediksi [10]. Dengan demikian peneliti berinisiatif menemukan solusi memprediksi seorang akan berisiko terkena penyakit *stroke* dengan menggunakan algoritma *Random Forest* secara lebih efisien dan lebih cepat .

## 1.2 Ruang Lingkup

Ruang lingkup penelitian ini dibatasi pada beberapa hal, antara lain sebagai berikut:

- a. Data yang digunakan merupakan data penyakit *stroke* dengan jumlah 4981 set data, data diambil dari *kaggle*.
- b. Penentuan prediksi faktor penyakit *stroke* menggunakan algoritma *random forest*, keluaran yang dihasilkan berupa model prediksi risiko penyakit *stroke* dalam lingkup simulasi *google Colaboratory*.
- c. Model yang dihasilkan hanya dapat memprediksi apakah seseorang terkena penyakit *stroke* atau tidak (*Binary Classification*).

## 1.3 Rumusan Masalah

Berdasarkan masalah dan alternatif solusi yang diusulkan, maka ditetapkan rumusan masalah sebagai berikut. “Apakah implementasi algoritma *random forest* dapat memprediksi seseorang yang akan terkena penyakit *stroke* secara lebih akurat dan cepat?”.

## 1.4 Tujuan Penelitian

Tujuan pada penelitian ini adalah sebagai berikut:

- a. Menentukan atribut yang sesuai dalam memprediksi risiko penyakit *stroke* agar dapat memberikan prediksi yang akurat dan dapat diandalkan terhadap kemungkinan seseorang mengalami *stroke*.
- b. Mendapatkan analisa terkait fitur-fitur yang paling berpengaruh.
- c. Menghasilkan model prediksi risiko penyakit *stroke* menggunakan algoritma *Random Forest*.

### 1.5 Manfaat Penelitian

Manfaat dari hasil penelitian ini adalah sebagai berikut:

- a. Berkontribusi terhadap ilmu pengetahuan tentang memprediksi risiko penyakit *stroke* menggunakan algoritma *random forest*.
- b. Tersedianya model prediksi yang membantu dan mempermudah dalam mengetahui prediksi risiko penyakit *stroke*.
- c. Meningkatnya efisiensi proses diagnosis penyakit *stroke* secara lebih cepat.

### 1.6 Sistematika Penulisan

Penulisan proposal ini dibagi menjadi 5 bab, dengan sistematika sebagai berikut:

#### a. Bab I Pendahuluan

Bab ini berisi antara lain latar belakang, ruang lingkup, rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

#### b. Bab II Tinjauan Pustaka

Pada bab ini menjelaskan tentang penelitian yang terkait dan teori pendukung yang digunakan penelitian ini

#### c. Bab III Metode Penelitian

Bab ini membahas tentang objek penelitian dan metode yang digunakan dalam menyelesaikan permasalahan.

#### d. Bab IV Hasil dan Pembahasan

Bab ini menjelaskan hasil penelitian dan pembahasan terkait hasil-hasil penelitian.

#### e. Bab V Kesimpulan dan Saran

Bab ini menyimpulkan hasil dari penelitian guna mendapatkan hasil analisis dari metode *random forest* yang dijadikan sebagai jawaban dari rumusan masalah yang telah disampaikan.