

BAB I

PENDAHULUAN

1.1 Latar Belakang

Algoritma *Naïve Bayes* adalah teknik klasifikasi berdasarkan pada teorema *Bayes* dan digunakan untuk memprediksi kemungkinan bahwa suatu kelas akan memiliki anggota. Memiliki keyakinan yang kuat bahwa setiap situasi atau keadaan independen adalah ciri utama *Naïve Bayes*. (Halimah dkk., 2020). Akan tetapi terkadang *overfitting* juga sering terjadi pada model *Naïve Bayes* yang terlalu kompleks, dalam konteks ini *Particle Swarm Optimization (PSO)* dapat membantu mengurangi risiko *overfitting* pada model *Naïve Bayes* dengan mengoptimalkan parameter dan memilih fitur yang relevan (Putra, 2022). *Particle Swarm Optimization (PSO)* akan melakukan pembobotan pada atribut-atribut yang relevan pada data dengan melakukan seleksi atribut pada pembobotan data yang paling berpengaruh terhadap hasil klasifikasi sehingga meningkatkan akurasi model *Naïve Bayes* (Rakhmawati, 2023). Beberapa penelitian menunjukkan dengan mengintegrasikan *Particle Swarm Optimization (PSO)* hasil klasifikasi model *Naïve Bayes* meningkat secara signifikan (Geetha dkk, 2020). *Particle Swarm Optimization (PSO)* juga relatif cepat dan mudah diimplementasikan, yang memungkinkan untuk pengolahan data dalam jumlah besar dengan efisiensi yang baik (D. Pajri dkk, 2020). *Naïve Bayes* yang dioptimasi dengan *Particle Swarm Optimization (PSO)* juga menghasilkan hasil yang konsisten dan dapat digunakan dalam berbagai studi kasus dan menunjukkan keefektifan yang cukup baik (T. Hidayat dkk, 2020).

Dalam konteks ini, penggunaan algoritma pembelajaran mesin untuk klasifikasi faktor risiko stroke menjadi sangat penting. Salah satu algoritma yang sering digunakan adalah *Naïve Bayes*, yang berdasarkan pada Teorema Bayes dengan asumsi independensi antar fitur. Meskipun asumsi independensi ini sering tidak sepenuhnya akurat dalam aplikasi dunia nyata, *Naïve Bayes* telah terbukti efektif dalam berbagai tugas klasifikasi, termasuk dalam bidang kesehatan. Namun, akurasi dari algoritma ini masih dapat ditingkatkan untuk memberikan

hasil yang lebih baik dalam identifikasi faktor risiko stroke. Meskipun *stroke* sudah dikenal dalam dunia medis, masyarakat pada umumnya masih kurang mengenal atau memahami. Hal ini sering mengakibatkan penundaan dalam mencari bantuan medis yang tepat, sehingga angka kejadian *stroke* semakin meningkat. Banyak faktor yang menyebabkan peningkatan jumlah kasus *stroke*, termasuk gaya hidup masyarakat saat ini yang cenderung tidak terkontrol. Faktor-faktor tersebut meliputi gaya hidup tidak sehat, kerja berlebihan, stres, kurang berolahraga, merokok, penggunaan obat yang tidak sesuai, dan faktor lainnya. (Nugroho dkk, 2023).

Salah satu metode yang dapat digunakan untuk mengoptimalkan performa Naive Bayes adalah *Particle Swarm Optimization (PSO)*. PSO merupakan teknik optimasi yang terinspirasi dari perilaku sosial hewan, seperti burung yang terbang dalam kawanan. PSO bekerja dengan mengoptimalkan masalah melalui iterasi sejumlah partikel yang bergerak dalam ruang pencarian solusi. Dengan mengimplementasikan PSO, parameter-parameter pada *Naive Bayes* dapat dioptimalkan untuk meningkatkan akurasi klasifikasi. Dalam upaya membantu masyarakat mengidentifikasi *stroke* lebih dini, diperlukan pengembangan suatu model sistem yang mampu mengklasifikasikan potensi *stroke* dan memberikan informasi mengenai risiko *stroke*. Model sistem berbasis algoritma *Naive Bayes* dapat menjadi solusi yang efektif dalam konteks ini. *Data mining* menjadi terobosan yang membantu masyarakat sebagai penyelesaian permasalahan yang sedang dihadapi. *Naive Bayes* dapat diartikan algoritma yang berdasarkan pada Teorema *Bayes* dengan asumsi bahwa fitur-fitur yang digunakan bersifat independen satu sama lain. Walaupun asumsi ini sering kali tidak sepenuhnya tepat dalam banyak aplikasi di dunia nyata, *Naive Bayes* terbukti efektif dalam berbagai tugas klasifikasi, termasuk di bidang kesehatan. Dalam konteks prediksi *stroke*, *Naive Bayes* dapat digunakan untuk memodelkan hubungan antara berbagai faktor risiko dan gejala dengan kemungkinan terjadinya *stroke*. *Naive Bayes* adalah salah satu algoritma dalam jenis Jaringan Bayesian yang paling sering digunakan untuk mengklasifikasikan dataset. Metode *Naive Bayes* memiliki tingkat akurasi yang sangat tinggi, mencapai 99,51%. *Naive Bayes*

membuat asumsi bahwa kemunculan atribut kelas yang ada tidak terkait dengan keberadaan atribut lain yang membuat kelas lebih efisien (Geetha dkk, 2020). Algoritma *Naive Bayes* dengan *Particle Swarm optimization* memiliki keakuratan yang tinggi ini terbukti dalam penelitian klasifikasi dataset dermatologi. Klasifikasi dataset tersebut berdasarkan dari beberapa penyakit dermatologi yang berbeda dengan nilai akurasi yang sangat bagus. Oleh karena itu Metode *Naive Bayes* dengan *Particle Swarm optimization* cocok untuk klasifikasi data (Nuranisah & Yusman, 2023).

Metode *Naive Bayes* dipadukan dengan algoritma *Particle Swarm Optimization* (PSO) dengan tujuan memperoleh akurasi terbaik. Metode optimasi stokastik berbasis populasi diusulkan pada tahun 1995 oleh Kennedy dan Eberhart. Keuntungan dari PSO adalah mudah diterapkan dan ada beberapa parameter untuk menyesuaikan-nya (Putra, 2022). PSO dikembangkan bergantung pada metafora komunikasi dan interaksi sosial dari pergerakan kawanan burung atau ikan. Algoritma PSO adalah algoritma optimasi yang bertujuan untuk menemukan solusi terdekat dengan minimum global atau maksimum global (T. Hidayat dkk, 2020). PSO banyak menyelesaikan masalah seleksi fitur dan optimasi (D. Pajri dkk, 2020).

Penyakit *stroke* merupakan salah satu penyakit yang paling mematikan dan menjadi momok bagi kesehatan masyarakat seluruh dunia. *Stroke* terjadi karena kerusakan pada otak yang disebabkan oleh gangguan mengalirnya darah ke otak, yang bisa disebabkan oleh penyumbatan atau *stroke* iskemik dan pecahnya pembuluh darah atau *stroke* hemoragik (NIH, 2023). *Stroke* menjadi penyakit dengan penyebab kematian nomor satu di Indonesia dengan angka kematian akibat *stroke* di Indonesia pada tahun 2023 mencapai 131,8 kasus kematian per 100 ribu penduduk (Kautsar, 2023). Gejala *stroke* yang muncul bisa sangat bervariasi, mulai dari kelemahan tiba-tiba di satu sisi tubuh, kesulitan berbicara, gangguan penglihatan, hingga penurunan kesadaran. *Stroke* merupakan penyakit dengan tingkat kematian yang paling tertinggi di seluruh dunia. *Stroke* dapat menyerang siapa saja tidak peduli usia ataupun jenis kelamin, dan sering terjadi secara tiba-tiba dengan angka kematian yang cukup tinggi (Byna dkk, 2020) . Penyakit yang menyerang bagian otak secara mendadak disebut *stroke*. Penyakit ini

dapat mengganggu fungsi otak dan mengganggu aliran darah, yang dapat menyebabkan kematian dalam waktu singkat. *Stroke* umumnya disebabkan oleh dua hal yaitu gumpalan darah pada pembuluh darah di otak dan pecahnya pembuluh darah pada otak. Penyempitan maupun pecahnya pembuluh darah ini dapat disebabkan oleh berbagai faktor, seperti tekanan darah yang tinggi, penggunaan obat untuk pengencer darah, aneurisma otak, dan cedera pada otak. (Pittara, 2022).

Penelitian ini bertujuan untuk mengembangkan metode optimasi *Naive Bayes* menggunakan *Particle Swarm Optimization* dalam klasifikasi faktor risiko stroke otak. Dengan menggabungkan keunggulan kedua metode ini, diharapkan akurasi prediksi faktor risiko stroke dapat ditingkatkan, sehingga memungkinkan deteksi dini dan intervensi yang lebih tepat. Optimalisasi ini tidak hanya akan memberikan kontribusi signifikan dalam bidang kesehatan, tetapi juga menawarkan pendekatan baru dalam pengembangan sistem prediksi medis yang lebih cerdas dan efektif. Berdasarkan rujukan tersebut, maka penelitian ini perlu dilakukan untuk mengetahui kemungkinan terjadinya penyakit stroke, dengan menerapkan *Naive Bayes* yang dioptimasi dengan *Particle Swarm Optimization* untuk mengklasifikasikan penyakit *stroke* berdasarkan dataset faktor risiko *stroke*.

1.2 Rumusan Masalah

Rumusan masalah berikut dapat dibuat berdasarkan latar belakang di atas :

“Bagaimana mengembangkan model sistem berbasis *Naive Bayes* dan membuat sistem berbasis *Naive Bayes* yang dioptimasi dengan *Particle Swarm Optimization* (PSO) dan didukung oleh teknik *data mining* pada dataset faktor risiko *stroke*?”

1.3 Tujuan Penelitian

Tujuan dari penelitian ini :

1. Metode *Naive Bayes* yang dioptimasi menggunakan *Particle Swarm Optimization* (PSO) bertujuan untuk menganalisis klasifikasi dan prediksi berdasarkan data faktor risiko *stroke* otak.

2. Untuk menganalisis deskriptif terhadap dataset dan mengidentifikasi pola dan hubungan antara berbagai faktor risiko *stroke*.
3. Mengembangkan dan membuat model prediksi menggunakan algoritma *Naive Bayes* yang dioptimasi menggunakan *Particle Swarm Optimization (PSO)*.
4. Melakukan evaluasi kinerja model.

Dengan mengembangkan model sistem berbasis *Naive Bayes* yang dioptimasi dengan *Particle Swarm Optimization (PSO)* dan didukung oleh teknik *data mining*, penelitian ini bertujuan untuk menerapkan algoritma *Particle Swarm Optimization (PSO)* yang bertujuan untuk memilih atribut *input* pada metode *Naive Bayes*. Implementasi sistem ini diharapkan dapat memberikan informasi tambahan terkait algoritma *Naive Bayes* dan *Particle Swarm Optimization (PSO)*.

1.4 Manfaat Penelitian

Dengan mempertimbangkan tujuan di atas, keuntungan yang diharapkan dari penelitian ini adalah sebagai berikut :

1.4.1 Manfaat Teoritis

Penelitian terhadap *Naive Bayes* dan *Particle Swarm Optimization* ini diharapkan menambah pengetahuan dan wawasan mengenai pengembangan model sistem berbasis *Naive Bayes* yang dioptimasi dengan *Particle Swarm Optimization (PSO)* dan didukung oleh teknik *data mining* dengan dataset *stroke* otak, serta juga diharapkan menjadi bahan evaluasi pada pengembangan ilmu pengetahuan.

1.4.2 Manfaat Praktis

1. Bagi Penulis

Diharapkan bahwa penelitian ini akan menjadi alat yang bermanfaat untuk menerapkan pengetahuan penulis tentang pengembangan model sistem berbasis *Naive Bayes* yang dioptimasi dengan *Particle Swarm*

Optimization (PSO) dan didukung oleh teknik *data mining* dengan dataset *stroke* otak.

2. Bagi Peneliti Selanjutnya

Diharapkan penelitian ini akan membantu proses pembuatan model sistem berbasis Naive Bayes yang dioptimasi dengan *Particle Swarm Optimization* (PSO).

1.5 Batasan Masalah

Terdapat beberapa batasan yang terkait dengan masalah penelitian ini adalah sebagai berikut:

1. Penelitian akan menghitung tingkat akurasi model sistem berbasis Naive Bayes yang dioptimasi dengan *Particle Swarm Optimization* (PSO) dengan dataset faktor risiko *stroke* otak.
2. Data pada penelitian ini menggunakan adalah dataset faktor risiko *stroke* yang bersumber dari *Kaggle*.
3. Penelitian ini akan membuat model sistem berbasis Naive Bayes yang dioptimasi dengan *Particle Swarm Optimization* (PSO) dan didukung oleh teknik *data mining* untuk mengklasifikasikan *stroke* otak.