

BAB I

PENDAHULUAN

1.1 Latar Belakang

Migrain adalah salah satu gangguan *neurologis* paling umum yang ditemukan di seluruh dunia. Kondisi ini biasanya ditandai dengan sakit kepala berulang yang berdenyut, seringkali menyerang satu sisi kepala, dan disertai gejala seperti mual, sensitivitas terhadap cahaya, dan gangguan penglihatan. Menurut laporan *Global Burden of Migrain*, (2025) penyakit ini menempati peringkat kedua sebagai penyebab utama disabilitas akibat gangguan neurologis secara global, dengan lebih dari 1 miliar orang di seluruh dunia yang terkena dampaknya. Berdasarkan laporan *World Health Organization*. (2022) migrain menjadi penyebab kedua terbesar disabilitas neurologis di dunia setelah stroke. Di Indonesia, berdasarkan data dari Kementerian Kesehatan RI. (2023), sekitar 10-12% populasi dewasa mengalami migrain, dengan tingkat prevalensi yang cenderung lebih tinggi pada wanita dibandingkan pria. Hal ini menunjukkan bahwa migrain tidak hanya merupakan masalah kesehatan, tetapi juga memiliki dampak sosial dan ekonomi, karena dapat mengurangi produktivitas kerja dan menurunkan kualitas hidup bagi mereka yang mengalaminya.

Pada praktik klinis, diagnosis migrain masih sangat bergantung pada wawancara subjektif dengan pasien dan pengamatan gejala oleh tenaga kesehatan. Pendekatan tradisional ini seringkali berisiko menyebabkan diagnosis yang salah, karena gejala migrain dapat mirip dengan gejala jenis sakit kepala lainnya, seperti sakit kepala tegang atau sakit kepala *cluster*. Masalah ini semakin parah di fasilitas kesehatan yang tidak memiliki ahli saraf. Oleh karena itu, diperlukan metode berbasis teknologi yang dapat menganalisis data pasien secara lebih objektif dan akurat. Salah satu alternatif adalah *machine learning*, teknik kecerdasan buatan yang dapat mengidentifikasi pola data dan menghasilkan prediksi otomatis berdasarkan informasi gejala pasien Khan dkk. (2024)

Beberapa penelitian terkini telah mengeksplorasi penggunaan *machine learning (ML)* untuk klasifikasi migrain. “*Klasifikasi Penyakit Migrain Menggunakan Metode Support Vector Machine*”.

Model SVM pada penelitian ini memiliki kinerja dengan akurasi 75,00% serta juga pada penelitian sebelumnya “*Klasifikasi Penyakit Migrain dengan Metode Naïve Bayes pada Dataset Kaggle*” dengan Model *Naïve Bayes* menghasilkan akurasi sebesar 88,51% pada tahap pelatihan dengan 280 data *training*, dan 89,02% pada tahap pengujian dengan 120 data *testing*. Meskipun hasil-hasil ini menunjukkan prospek yang menjanjikan, terdapat beberapa kelemahan dalam studi-studi sebelumnya. “*Decision Tree and Random Forest Classification Algorithms for Mangrove Forest Mapping in Sembilang National Park, Indonesia*” Untuk gambar Landsat-7 ETM+, akurasi klasifikasi keseluruhan algoritma RF (99,12%) lebih tinggi dibandingkan algoritma DT (92,82%), tetapi untuk gambar Landsat-8 OLI, akurasi algoritma DT (98,34%) sedikit lebih tinggi daripada algoritma RF (97,79%). Secara keseluruhan, algoritma RF cenderung unggul dibandingkan algoritma DT karena menggunakan parameter yang memberikan akurasi produsen yang lebih tinggi dalam pemetaan hutan *mangrove*. Sebagian besar studi masih mengandalkan jumlah data yang kecil, belum menerapkan metode penyeimbangan data untuk mengatasi ketidakseimbangan kelas, Kekurangan-kekurangan ini telah menciptakan kebutuhan akan penelitian baru yang dapat memberikan analisis perbandingan dan hasil yang lebih representatif Petrušić dkk. (2025)

Penelitian ini bertujuan untuk mengklasifikasikan gangguan migrain menggunakan algoritma *Random Forest* guna memperoleh model dengan kinerja optimal dalam mendukung proses diagnosis migrain secara otomatis berdasarkan gejala pasien. Pendekatan ini diharapkan dapat menentukan metode terbaik yang memberikan kinerja optimal dalam proses diagnosis migrain secara otomatis, berdasarkan informasi gejala pasien.

Penerapan algoritma klasifikasi seperti *Random Forest*, yang dipadukan dengan teknik (*Synthetic Minority Oversampling Technique*) SMOTE, mampu mengatasi ketidakseimbangan data pada kasus diagnosis penyakit seperti migrain. Metode ini memungkinkan model untuk mengenali karakteristik pasien dengan gejala minoritas secara lebih akurat, sehingga mendukung pengembangan sistem pendukung keputusan (*Decision Support System*) di bidang neurologi (Kumar & Jain, 2023).

Pada konteks klasifikasi penyakit migrain, data yang digunakan umumnya bersifat tidak seimbang (*imbalanced data*), sehingga berpotensi menurunkan performa model, khususnya dalam mengenali kelas minoritas. Oleh karena itu, penelitian ini menerapkan *Synthetic Minority Oversampling Technique* (SMOTE) untuk menyeimbangkan distribusi kelas sebelum proses klasifikasi dilakukan. Penerapan SMOTE diharapkan mampu meningkatkan kemampuan model dalam mendeteksi kasus migrain secara lebih akurat dan adil.

Selain itu, penelitian ini secara khusus mengevaluasi kombinasi *hyperparameter max_features* dan *min_samples_leaf* pada algoritma *Random Forest*. *Hyperparameter max_features* berfungsi untuk mengatur jumlah atribut yang dipertimbangkan pada setiap pemisahan *node*, sedangkan *min_samples_leaf* mengontrol jumlah minimum data pada daun pohon keputusan agar model tidak terlalu kompleks. Evaluasi dilakukan untuk menentukan kombinasi nilai *hyperparameter* yang menghasilkan performa terbaik berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score* setelah penerapan SMOTE. Dengan pendekatan ini, diharapkan dapat diperoleh model *Random Forest* yang optimal dan andal dalam melakukan klasifikasi penyakit migrain berdasarkan gejala pasien.

Berdasarkan perspektif teoretis, penelitian ini diharapkan dapat memperkaya literatur mengenai penerapan *machine learning* di bidang neurologi, khususnya dalam penanganan kasus migrain. Sementara itu, dari perspektif praktis, hasil penelitian ini diharapkan dapat menjadi landasan untuk pengembangan sistem pendukung keputusan klinis yang dapat membantu tenaga medis dalam melakukan diagnosis awal migrain secara lebih cepat dan akurat. Oleh karena itu, penelitian ini tidak hanya berkontribusi pada kemajuan ilmu pengetahuan, tetapi juga berpotensi meningkatkan kualitas layanan kesehatan berbasis data di Indonesia.

1.2 Rumusan Penelitian

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah:

1. Berapa akurasi, presisi, *recall*, dan *F1-score* dari algoritma *Random Forest* dalam melakukan klasifikasi penyakit migrain setelah penerapan metode SMOTE?

2. Kombinasi *hyperparameter* (*max_features*, *min_samples_leaf*, *max_depth* dan *n_estimator*) mana yang menghasilkan model terbaik (berdasarkan nilai akurasi, presisi, *recall*, dan *F1-score*) dalam klasifikasi penyakit migrain setelah penerapan SMOTE (*Synthetic Minority Oversampling Technique*)?
3. Apakah *deployment* sistem klasifikasi penyakit migrain berbasis web menggunakan algoritma *Random Forest* dan metode SMOTE dapat membantu proses prediksi jenis migrain berdasarkan gejala pasien secara objektif dan akurat?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas tujuan penelitian ini adalah untuk:

1. Untuk mengetahui nilai akurasi, presisi, *recall*, dan *F1-score* dari algoritma *Random Forest* dalam melakukan klasifikasi penyakit migrain, khususnya setelah dilakukan penyeimbangan data menggunakan metode SMOTE (*Synthetic Minority Oversampling Technique*).
2. Untuk mengetahui kombinasi optimal *hyperparameter* *max_features*, *min_samples_leaf*, *max_depth* dan *n_estimator* yang memberikan performa terbaik dalam klasifikasi penyakit migrain setelah penerapan SMOTE, serta mengidentifikasi pengaruh kombinasi tersebut terhadap hasil evaluasi model.
3. Untuk mengetahui penerapan *deployment* sistem klasifikasi penyakit migrain berbasis web menggunakan algoritma *Random Forest* dan metode SMOTE sehingga dapat membantu proses prediksi jenis migrain berdasarkan gejala pasien secara objektif, dan akurat.

1.4 Manfaat Penelitian

Penelitian ini memiliki manfaat untuk:

- a. Penelitian ini berkontribusi pada kemajuan ilmu pengetahuan di bidang *machine learning*, khususnya dalam penggunaan algoritma *Random Forest* dan teknik SMOTE untuk mengatasi ketidakseimbangan data medis dalam klasifikasi migrain. Temuan penelitian ini menambah literatur tentang pemodelan gejala migrain yang kompleks menjadi data yang lebih terorganisir untuk diagnosis bantu komputer, dan dapat digunakan sebagai

acuan dalam pengembangan model prediksi medis lainnya yang memanfaatkan data gejala atau sinyal fisiologis.

- b. Dari sudut pandang praktis, penelitian ini bermanfaat bagi tenaga medis, peneliti, dan lembaga kesehatan dalam mengenali gejala utama migrain, mendukung pengembangan sistem pendukung keputusan untuk deteksi dini migrain, serta memberikan panduan dalam penerapan metode *Random Forest* dan SMOTE pada kondisi penyakit dengan distribusi kelas yang tidak seimbang. Penelitian ini diharapkan dapat mempercepat proses diagnosis, meningkatkan kualitas hidup pasien migrain, dan membuka peluang kolaborasi antara bidang teknologi informasi dan kesehatan dalam inovasi layanan medis.

1.5 Batasan Penelitian

Penelitian ini memiliki batasan dalam hal:

1. Penelitian ini hanya berfokus pada klasifikasi penyakit migrain, tidak mencakup jenis gangguan neurologis lain seperti *tension-type headache* atau *cluster headache*.
2. *Dataset* yang digunakan bersifat sekunder (Data yang digunakan dalam penelitian ini merupakan *dataset* penyakit migrain yang diperoleh dari *platform* terbuka *Kaggle*) dengan jumlah data 400, tanpa melibatkan data rekam medis langsung dari fasilitas kesehatan. Waktu Pengumpulan pada *platform kaggle* dilakukan pada bulan Juni 2020.
3. Evaluasi performa model dibatasi pada penggunaan nilai akurasi, presisi, *recall*, dan *F1-score*.
4. Pengujian *hyperparameter* difokuskan pada empat parameter utama, yaitu *hyperparameter max_features*, *min_samples_leaf*, *max_depth* dan *n_estimator* sedangkan parameter lainnya dipertahankan pada nilai default yang digunakan oleh pustaka *scikit-learn*.
5. Algoritma yang digunakan terbatas pada *Random Forest*, dengan metode penyeimbangan data SMOTE.