

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Kanker kandung kemih (*bladder cancer*) merupakan salah satu penyakit kanker yang paling sering ditemukan di dunia, termasuk di Indonesia. Berdasarkan data statistik kesehatan global terbaru dari GLOBOCAN, jumlah kasus penyakit ini terus meningkat seiring dengan bertambahnya usia penduduk dan paparan faktor risiko lingkungan, seperti kebiasaan merokok dan paparan zat kimia berbahaya di tempat kerja [1]. Penyakit ini memberikan beban yang besar baik dari sisi kesehatan maupun biaya pengobatan karena karakternya yang berkelanjutan dan memerlukan pengawasan jangka panjang yang ketat [2].

Dalam penanganan medis, pengelompokan pasien yang tepat pada saat diagnosis awal adalah faktor penentu yang paling penting untuk mengetahui peluang kesembuhan dan menentukan langkah operasi. Secara medis, kanker kandung kemih dibagi menjadi dua kelompok besar yang sangat berbeda penanganannya: Kanker Kandung Kemih Non-Otot Invasif (*NMIBC*) dan Kanker Kandung Kemih Otot Invasif (*MIBC*) [3], [4].

Kanker kandung kemih non-invasif otot (*NMIBC*) merupakan tipe kanker awal yang belum menembus lapisan otot detrusor. *Transurethral resection of bladder tumour (TURBT)* berfungsi sebagai metode diagnostik serta berpotensi menyembuhkan *NMIBC*, namun terdapat beberapa kendala, seperti kesulitan dalam memastikan seluruh tumor terangkat selama reseksi bertahap dan kemungkinan terjadinya reimplantasi tumor setelah prosedur. Penanganan *NMIBC* masih jauh dari ideal, di mana tingkat kekambuhan satu tahun berkisar antara 15 hingga 61 persen, dan tingkat kekambuhan lima tahun antara 31 hingga 78 persen. Berbagai faktor penyebab kekambuhan pada *NMIBC* telah diidentifikasi, termasuk tumor yang terlewat saat *cystoscopy*, reseksi yang tidak tuntas selama *TURBT*, risiko reimplantasi tumor sesudah *TURBT*, serta metastasis yang berasal dari *karsinoma urotelial* di bagian atas saluran kemih, dan perubahan lapangan yang berkontribusi pada *karsinogenesis*. [5]. Sebaliknya, *MIBC* adalah jenis yang jauh lebih agresif di mana akar kanker telah menembus lapisan otot. Diagnosis *MIBC* sering kali mengharuskan operasi besar berupa pengangkatan seluruh kandung kemih (*kistektomi radikal*) untuk menyelamatkan nyawa pasien [6].

Masalah utama yang dihadapi saat ini adalah keterbatasan alat diagnosis standar dalam membedakan secara tepat antara kanker stadium awal yang berisiko tinggi dan kanker yang sudah mulai masuk ke otot. Pemeriksaan standar menggunakan *biopsi* sering kali tidak memberikan hasil yang akurat karena sampel otot yang diambil kurang memadai, menyebabkan dokter mengira kanker tersebut masih ringan padahal sudah berbahaya (*understaging*) [7]. Selain itu, interpretasi citra *CT Scan* maupun MRI secara visual oleh radiolog masih bergantung pada pengalaman dan keahlian klinis, sehingga dapat menimbulkan variasi antar-pengamat (*inter-observer variability*) yang berpotensi memengaruhi konsistensi diagnosis [8].

Untuk mengatasi masalah ini, teknologi *Radiomics* muncul sebagai solusi inovatif. *Radiomics* bekerja dengan mengubah gambar medis digital menjadi ribuan data angka (*quantitative features*) yang dapat diolah oleh komputer [9]. Data ini mampu mendeteksi pola tekstur mikroskopis pada tumor yang tidak terlihat oleh mata manusia. Namun, *Radiomics* menghasilkan data yang sangat banyak dan kompleks, yang dapat membingungkan sistem komputer jika tidak disaring dengan benar. Oleh karena itu, diperlukan metode seleksi fitur seperti algoritma *Boruta*. *Boruta* dikenal sebagai metode *all-relevant* yang menyaring fitur tidak relevan dan hanya mempertahankan fitur yang benar-benar penting untuk diagnosis, yang terbukti efektif dalam studi data medis terkini [10]. Selain itu, dataset MRI *multi-center* sering menunjukkan heterogenitas spasial akibat perbedaan resolusi *voxel*, yang dapat berdampak negatif terhadap kestabilan fitur dan kinerja model [11]. Maka dari itu pada penelitian ini dilakukan *resampling voxel* 1x1x1 mm untuk setiap *center*.

Di sisi lain, keberhasilan klasifikasi berbasis *radiomics* tidak hanya dipengaruhi oleh kualitas fitur yang diekstraksi, tetapi juga oleh algoritma pembelajaran mesin yang digunakan. Berbagai model klasifikasi tunggal seperti *Extreme Gradient Boosting* (XGBoost), *Support Vector Machine* (SVM), dan *Multilayer Perceptron* (MLP) telah banyak diterapkan pada klasifikasi citra medis dan menunjukkan karakteristik serta keunggulan yang berbeda-beda. Untuk meningkatkan performa prediksi, pendekatan *ensemble* seperti *stacking* dikembangkan dengan mengombinasikan prediksi dari beberapa model dasar sehingga diharapkan mampu memanfaatkan keunggulan masing-masing model [12]. Meskipun demikian, peningkatan kompleksitas melalui *stacking* tidak selalu menghasilkan kinerja yang lebih baik dibandingkan model tunggal, terutama pada dataset *radiomics* yang memiliki karakteristik berdimensi tinggi dan jumlah sampel

yang relatif terbatas. Oleh karena itu, diperlukan perbandingan antara model tunggal dan model *stacking* untuk mengevaluasi apakah peningkatan kompleksitas model benar-benar memberikan manfaat yang signifikan terhadap klasifikasi kanker kandung kemih berbasis *radiomics*.

Sebagai respons terhadap berbagai tantangan tersebut, penelitian ini bertujuan untuk menginvestigasi pengaruh kombinasi seleksi fitur *Boruta* dan *isotropic voxel resampling* terhadap klasifikasi berbasis *radiomics* dalam membedakan *Non-Muscle Invasive Bladder Cancer* (NMIBC) dan *Muscle Invasive Bladder Cancer* (MIBC) menggunakan citra MRI T2-weighted multi-center. Tiga algoritma pembelajaran mesin, yaitu *Extreme Gradient Boosting* (XGBoost), *Support Vector Machine* (SVM), dan *Multilayer Perceptron* (MLP), dievaluasi melalui empat skenario eksperimen untuk menganalisis secara sistematis pengaruh seleksi fitur dan prapemrosesan citra terhadap kinerja klasifikasi.

Selain itu, penelitian ini juga mengeksplorasi apakah peningkatan kompleksitas model melalui pendekatan *stacking-based ensemble* mampu memberikan peningkatan kinerja yang signifikan dibandingkan model klasifikasi tunggal pada konfigurasi prapemrosesan terbaik. Pengaruh jumlah dimensi fitur turut dianalisis untuk mengevaluasi kestabilan klasifikasi, dengan tujuan mengidentifikasi jumlah fitur hasil seleksi *Boruta* yang optimal dalam mencapai keseimbangan antara performa model dan robustitas prediksi.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana performa algoritma *XGBoost*, *SVM*, dan *MLP* (*Multilayer Perceptron*)/*Neural Network* dalam membedakan jenis kanker kandung kemih (NMIBC dan MIBC).
2. Apakah pendekatan *stacking-based ensemble* mampu meningkatkan performa klasifikasi dibandingkan model tunggal pada tugas klasifikasi kanker kandung kemih?
3. Bagaimana pengaruh variasi jumlah fitur hasil seleksi *Boruta* terhadap performa dan kestabilan klasifikasi, serta bagaimana tingkat sensitivitas model terhadap perubahan dimensi fitur yang digunakan?

1.3 Batasan Permasalahan

Penelitian ini dibatasi pada ruang lingkup sebagai berikut:

1. Data gambar yang digunakan adalah citra *MRI (T2-weighted imaging)* dari dataset FedBCa.
2. Klasifikasi penyakit hanya dibagi menjadi dua kategori: *Non-Muscle Invasive (NMIBC)* dan *Muscle Invasive (MIBC)* bukan dalam bentuk *staging*.
3. Klasifikasi model yang digunakan pada penelitian ini hanya *XGBoost*, *SVM*, dan *MLP (Multilayer Perceptron)/Neural Network*
4. Seleksi fitur (pemilahan data penting) hanya menggunakan metode *Boruta*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Membandingkan dan menganalisis performa algoritma *XGBoost*, *SVM*, dan *MLP (Multilayer Perceptron)/Neural Network* dalam membedakan jenis kanker kandung kemih (*NMIBC* dan *MIBC*).
2. Membandingkan performa klasifikasi antara metode *stacking ensemble* dan model *machine learning* tunggal untuk mengevaluasi sejauh mana pendekatan *stacking* memberikan peningkatan kinerja.
3. Menganalisis pengaruh variasi jumlah fitur hasil seleksi *Boruta* terhadap performa dan kestabilan klasifikasi untuk mengetahui tingkat sensitivitas model terhadap perubahan dimensi fitur.

1.5 Manfaat Penelitian

Manfaat yang akan diperoleh dari penelitian ini adalah:

1. **Manfaat bagi bidang medis** adalah memberikan pemahaman yang lebih komprehensif mengenai pengaruh transformasi citra, seleksi fitur, serta pendekatan *ensemble learning* terhadap kinerja algoritma klasifikasi seperti *XGBoost*, *SVM*, *MLP*, dan metode *stacking* dalam membedakan kanker kandung kemih tipe *NMIBC* dan *MIBC* berbasis *radiomics MRI*. Hasil

penelitian ini diharapkan dapat menjadi dasar ilmiah dalam pengembangan sistem pendukung keputusan klinis yang lebih robust dan stabil.

2. **Manfaat bagi bidang akademis** adalah memperkaya literatur ilmiah terkait integrasi teknik transformasi citra (seperti *Wavelet*, *Laplacian of Gaussian*, dan transformasi intensitas non-linear) dengan metode seleksi fitur Boruta serta pendekatan *ensemble stacking* pada data *radiomics* multidimensional. Penelitian ini juga memberikan kontribusi metodologis dalam memahami bagaimana kombinasi prapemrosesan, seleksi fitur, dan strategi klasifikasi memengaruhi stabilitas serta performa model pada data MRI.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN

Bab ini membahas latar belakang penelitian mengenai klasifikasi kanker kandung kemih berbasis radiomik dari citra MRI, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan skripsi.

- Bab 2 LANDASAN TEORI

Bab ini membahas teori-teori yang mendukung penelitian, meliputi konsep kanker kandung kemih, radiomik, ekstraksi fitur, seleksi fitur menggunakan Boruta, isotropic voxel resampling pada citra MRI, algoritma machine learning seperti Support Vector Machine (SVM), XGBoost, dan Multilayer Perceptron (MLP), metode stacking, serta metrik evaluasi yang digunakan dalam penelitian.

- Bab 3 METODOLOGI PENELITIAN

Bab ini menjelaskan metodologi penelitian yang digunakan, meliputi sumber data MRI, tahapan preprocessing, ekstraksi fitur radiomik, seleksi fitur menggunakan Boruta, pembangunan model machine learning tunggal dan stacking, serta metode evaluasi performa model.

- Bab 4 HASIL DAN DISKUSI

Bab ini menyajikan hasil eksperimen dan analisis dari model yang dibangun.

Selain itu, dibahas perbandingan performa model machine learning tunggal dan stacking, pengaruh isotropic resampling, seleksi fitur, serta analisis terhadap hasil klasifikasi kanker kandung kemih berbasis radiomik dari MRI.

- Bab 5 KESIMPULAN DAN SARAN

Bab ini berisi simpulan dari hasil penelitian yang telah dilakukan berdasarkan tujuan penelitian, serta saran untuk penelitian selanjutnya yang dapat digunakan sebagai pengembangan lebih lanjut.

