

BAB 2

LANDASAN TEORI

Bab ini menguraikan landasan teori yang menjadi dasar pemahaman terhadap permasalahan dan solusi yang diusulkan dalam penelitian. Pembahasan diawali dengan tinjauan medis mengenai kanker payudara, subtipe molekuler, dan pencitraan MRI. Selanjutnya, dipaparkan konsep-konsep komputasi yang digunakan, yaitu ekstraksi fitur, *Deep Learning* dengan arsitektur ResNet50, serta *Ensemble Learning*.

2.1 Kanker Payudara

Kanker payudara didefinisikan sebagai keganasan yang timbul akibat proliferasi selular yang tidak terkendali pada jaringan payudara, umumnya berasal dari epitel duktus (*ductal carcinoma*) atau lobulus (*lobular carcinoma*). Secara patofisiologis, kondisi ini ditandai dengan kemampuan sel neoplastik untuk melakukan invasi ke jaringan stroma sekitarnya dan bermetastasis ke organ distal melalui sistem limfatik atau hematogen [1].

2.2 Klasifikasi Subtipe Molekuler

Kanker payudara diklasifikasikan terhadap ekspresi tiga *biomarker* utama: *Estrogen Receptor* (ER), *Progesterone Receptor* (PR), dan *Human Epidermal Growth Factor Receptor 2* (HER2). Klasifikasi ini membagi kanker payudara menjadi empat subtipe intrinsik utama dengan implikasi terapeutik yang berbeda [6]:

1. *Luminal A* (ER+/PR+, HER2-,-): Subtipe ini ditandai dengan ekspresi reseptor hormon yang kuat dan indeks proliferasi selular yang rendah. Secara klinis, subtipe ini memiliki prognosis yang paling baik dan responsif terhadap terapi endokrin (hormonal), namun menunjukkan sensitivitas yang rendah terhadap kemoterapi sitotoksik.
2. *Luminal B* (ER+, HER2+/-): Memiliki fenotipe yang serupa dengan Luminal A namun dengan tingkat proliferasi sel yang lebih tinggi dan/atau ekspresi reseptor progesteron yang rendah. Subtipe ini memiliki prognosis yang

lebih buruk dibandingkan Luminal A dan sering kali memerlukan kombinasi kemoterapi.

3. *HER2-Enriched* (ER-/PR-, HER2+): Subtipe ini dicirikan oleh overekspresi onkogen HER2 yang memacu pertumbuhan tumor secara agresif. Meskipun memiliki tingkat keganasan tinggi, subtipe ini menunjukkan respons yang signifikan terhadap agen terapi target anti-HER2 (seperti Trastuzumab).
4. *Triple Negative Breast Cancer / TNBC* (ER-, PR-, HER2-): Merupakan subtipe yang tidak mengekspresikan reseptor hormon maupun HER2. TNBC memiliki sifat biologis yang sangat agresif, risiko kekambuhan dini yang tinggi, dan opsi terapi sistemik yang terbatas pada kemoterapi sitotoksik konvensional dikarenakan ketiadaan target molekuler spesifik [7].

2.3 Neoadjuvant Chemotherapy (NAC) dan pCR

Neoadjuvant Chemotherapy (NAC) merupakan prosedur kemoterapi yang diberikan sebelum dilakukan operasi pengangkatan tumor. Indikasi utama NAC meliputi upaya pengecilan massa tumor untuk memberikan opsi tindakan bedah konservasi payudara (*Breast-Conserving Surgery*) serta evaluasi sensitivitas tumor terhadap kemoterapi [2].

Keberhasilan NAC dievaluasi melalui pencapaian *Pathological Complete Response* (pCR). Secara definisi pCR adalah ketiadaan sel kanker invasif residu pada spesimen payudara dan kelenjar getah bening aksila setelah operasi dilakukan. Studi menunjukkan bahwa pencapaian pCR berkorelasi positif secara signifikan dengan peningkatan *Disease-Free Survival* (DFS) dan *Overall Survival* (OS), menjadikannya *endpoint* klinis yang valid [2].

2.4 Prinsip Citra DCE-MRI

Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) adalah modalitas pencitraan yang mengevaluasi karakteristik tumor. Teknik ini memanfaatkan citra *T1-weighted* sebelum dan sesudah injeksi kontras.

Penggunaan DCE-MRI dalam onkologi didasarkan pada fenomena angiogenesis tumor. Jaringan neoplastik ganas mensekresikan faktor angiogenik yang memicu pembentukan pembuluh darah baru dengan permeabilitas tinggi. Kondisi ini menyebabkan ekstravasasi agen kontras yang cepat ke dalam ruang

ekstraseluler tumor (*rapid wash-in*) dan eliminasi yang cepat (*rapid wash-out*). Kinetika penyangatan sinyal ini memberikan informasi kuantitatif mengenai morfologi dan fisiologi tumor yang tidak dapat dideteksi oleh pencitraan anatomi biasa [8].

2.5 Ekstraksi dan Rekayasa Fitur

Analisis citra medis berbasis komputasi memerlukan transformasi data visual mentah menjadi representasi numerik yang dapat diproses oleh algoritma pembelajaran mesin.

2.5.1 Ekstraksi Fitur Berbasis Deep Learning

Ekstraksi fitur dalam pengolahan citra medis bertujuan untuk mendapatkan representasi numerik yang merepresentasikan karakteristik intrinsik organ atau jaringan. Berbeda dengan metode konvensional yang mengandalkan fitur buatan tangan (*hand-crafted features*), pendekatan *Deep Learning* memungkinkan ekstraksi fitur dilakukan secara otomatis dan hierarkis.

Salah satu metode yang efektif dalam ekstraksi fitur pada *dataset* medis yang terbatas adalah *Transfer Learning*. Secara teoritis, *Transfer Learning* memanfaatkan bobot dari model yang telah dilatih pada *dataset* berskala besar (seperti ImageNet) untuk diterapkan pada domain baru [9]. Model arsitektur seperti ResNet50 [10] sering menjadi acuan dalam proses ini karena kemampuannya mempertahankan informasi fitur melalui *residual connections*, yang memungkinkan jaringan mempelajari fitur mendalam (*Deep Features*) tanpa mengalami degradasi gradien.

2.5.2 Rekayasa Fitur (*Feature Engineering*)

Untuk meningkatkan daya diskriminasi model pada *dataset* dengan dimensi tinggi, diterapkan teknik rekayasa fitur berupa *Interaction Features*. Metode ini mengonstruksi variabel baru melalui operasi matematis (perkalian polinomial) antara fitur radiomik (pencitraan) dan fitur klinis. Pendekatan ini bertujuan untuk memodelkan hubungan non-linier dan ketergantungan sinergis antara fenotipe radiologis dan profil biologis pasien [11].

Formulasi matematis untuk fitur interaksi ini ditunjukkan pada Persamaan 2.1. Jika X_{mri} adalah fitur visual dan X_{klinis} adalah variabel klinis, maka fitur interaksi baru ($X_{\text{interaksi}}$) didefinisikan sebagai:

$$X_{\text{interaksi}} = X_{\text{mri}} \cdot X_{\text{klinis}} \quad (2.1)$$

2.6 Ensemble Learning

Ensemble Learning adalah paradigma pembelajaran mesin yang menggabungkan prediksi dari beberapa model (*base learners*) untuk meningkatkan generalisasi dan mengurangi varians prediksi tunggal.

Metode *voting* yang digunakan pada penelitian ini adalah *Soft Voting*. Secara matematis, metode ini menghitung estimasi probabilitas rata-rata dari seluruh model dasar (N) untuk menentukan probabilitas prediksi akhir (\hat{y}), sebagaimana dinyatakan dalam Persamaan 2.2:

$$\hat{y} = \frac{1}{N} \sum_{j=1}^N P_j \quad (2.2)$$

Dimana P_j adalah probabilitas yang dihasilkan oleh model ke- j . Pendekatan ini memberikan keputusan yang lebih stabil dan konsisten dibandingkan metode *Hard Voting* dalam kasus klasifikasi dengan keputusan (*decision boundary*) yang rumit [12].

2.7 Metrik Evaluasi Kinerja

Evaluasi kinerja model prediksi pada *dataset* medis yang tidak seimbang (*imbalanced*) memerlukan penggunaan metrik komprehensif selain akurasi.

2.7.1 Sensitivitas (Recall)

Sensitivitas mengukur proporsi kasus positif (pCR) yang teridentifikasi secara benar. Metrik ini krusial untuk meminimalkan tingkat *False Negative*, dengan rumus perhitungan yang disajikan pada Persamaan 2.3:

$$\text{Sensitivitas} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.3)$$

Dimana TP adalah *True Positive* dan FN adalah *False Negative*.

2.7.2 Spesifisitas

Spesifisitas mengukur kemampuan model dalam mengeliminasi kasus negatif (Non-pCR) secara benar. Perhitungan spesifisitas dapat dilihat pada Persamaan 2.4:

$$\text{Spesifisitas} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2.4)$$

Dimana TN adalah *True Negative* dan FP adalah *False Positive*.

2.7.3 Area Under the ROC Curve (AUC)

Merepresentasikan probabilitas bahwa model akan memberikan peringkat lebih tinggi pada instans positif yang dipilih secara acak dibandingkan instans negatif. Nilai AUC memberikan ukuran agregat performa diskriminasi model yang invariant terhadap ambang batas klasifikasi (*threshold*).