

BAB 2

TINJAUAN PUSTAKA

2.1 Kanker Ginjal

Kanker ginjal atau *Renal Cell Carcinoma* (RCC) adalah neoplasma ganas yang berasal dari sel epitel tubulus ginjal, khususnya tubulus proksimal. *Renal Cell Carcinoma* (RCC) menyumbang sekitar 85–90% dari seluruh kasus kanker ginjal dan merupakan salah satu jenis kanker urologi paling umum pada orang dewasa [13].

Di Amerika Serikat, *Renal Cell Carcinoma* (RCC) menyumbang sekitar 5% dari seluruh kasus kanker baru pada pria dan 3% pada wanita. Diperkirakan pada tahun 2025 akan terdapat 52.410 kasus baru pada pria dan 28.570 pada wanita, menunjukkan relevansi klinis dan epidemiologis penyakit ini hingga saat ini [13]. Faktor risiko utama meliputi merokok, obesitas, hipertensi, paparan bahan kimia tertentu (misalnya trikloroetilen), serta predisposisi genetik seperti *Von Hippel–Lindau Syndrome* (VHL).

Secara histopatologis, *Renal Cell Carcinoma* (RCC) dibagi menjadi beberapa sub tipe utama:

1. *Clear Cell Renal Cell Carcinoma* (ccRCC) merupakan bentuk paling umum, mencakup sekitar 70–80% dari seluruh RCC.
2. *Papillary Renal Cell Carcinoma* (pRCC) mencakup 10–15% dari kasus RCC.
3. *Chromophobe Renal Cell Carcinoma* (chRCC) mencakup sekitar 5% dari kasus RCC.

Perbedaan sub tipe ini penting karena memengaruhi prognosis, gambaran histopatologis, serta respon terhadap terapi [14].

Secara umum, tahapan kanker ginjal (berdasarkan sistem TNM/nomor) diklasifikasikan sebagai berikut:

1. Stadium I: Kanker berukuran kecil, yaitu 7 cm atau lebih kecil, dan sepenuhnya terbatas di dalam ginjal. Tumor belum menyebar ke kelenjar getah bening atau organ lain (T1, N0, M0) [15][16].

2. Stadium II: Kanker berukuran lebih besar dari 7 cm, namun masih sepenuhnya berada di dalam ginjal dan belum menyebar (T2, N0, M0) [15][16].
3. Stadium III: Kanker berukuran berapa pun, tetapi telah tumbuh dan menyebar ke pembuluh darah utama ginjal (seperti vena ginjal atau vena kava) dan/atau ke setidaknya satu kelenjar getah bening di dekat ginjal namun belum menyebar jauh (T3 atau N1, M0) [15][16].
4. Stadium IV: Kanker dianggap paling lanjut, yaitu ketika telah tumbuh keluar dari lapisan tebal di sekitar ginjal (Gerota's fascia), menyerang kelenjar adrenal, atau menyebar ke bagian tubuh jauh seperti paru-paru atau tulang (T4 atau M1) [15][16].

2.2 microRNA (miRNA)

miRNA merupakan RNA kecil non-koding dengan panjang sekitar 19–25 nukleotida, sehingga tergolong relatif pendek. Meskipun demikian, miRNA memiliki peranan biologis yang sangat penting pada tingkat molekuler dan berfungsi dalam mengatur ekspresi gen secara pascatranskripsi pada tumbuhan, hewan, dan beberapa jenis virus. Dalam genom manusia, terdapat ribuan miRNA yang telah terklasifikasi, dan miRNA bertanggung jawab dalam menargetkan sekitar 60% gen pengkode protein [17, 18].

miRNA berperan penting dalam berbagai proses biologis dasar di dalam tubuh, seperti proliferasi sel, pengendalian siklus sel, apoptosis, diferensiasi, migrasi, dan metabolisme. Meskipun masih terdapat beberapa aspek dalam proses biogenesis miRNA yang belum sepenuhnya dipahami, mekanisme utamanya telah berhasil dikarakterisasi. Ketidakteraturan regulasi miRNA (*miRNA dysregulation*) dapat menyebabkan berbagai jenis kanker. Data ekspresi mikroarray yang dikumpulkan dari beragam jenis kanker telah menunjukkan bahwa ekspresi miRNA yang abnormal merupakan salah satu faktor awal yang memengaruhi perkembangan kanker [19, 20]. Studi terkini pada strain tikus juga menunjukkan bahwa miRNA tunggal maupun kluster miRNA bertanggung jawab terhadap berbagai jenis penyakit. Selain itu, telah dilaporkan bahwa tingkat miRNA matang yang rendah berkaitan dengan penyakit kanker [21], dan miRNA terbukti menjadi salah satu biomarker paling menjanjikan dalam memberikan informasi terkait diferensiasi subtype kanker dan prognosis penyakit [22]. Untuk pembahasan lebih lanjut

mengenai hubungan antara miRNA dan kanker, dapat merujuk pada [19].

Para peneliti terus berupaya mengidentifikasi faktor bio-miRNA yang signifikan serta menentukan miRNA dengan daya diskriminatif tertinggi yang dapat digunakan sebagai biomarker untuk diagnosis yang presisi, khususnya dalam identifikasi sub tipe dan stadium kanker ginjal. Berikut ini dirangkum beberapa penelitian terdahulu yang membahas hubungan antara miRNA dengan identifikasi sub tipe dan stadium kanker ginjal. *White et al.* [23] melakukan analisis kombinatorial terhadap miRNA yang sebelumnya dilaporkan mengalami disregulasi dan berhasil mengidentifikasi 62 miRNA dari total 133 miRNA yang telah dilaporkan oleh peneliti lain.

Pada penelitian *Juan et al.* [24] ditemukan 35 miRNA yang mampu membedakan antara *clear cell Renal Cell Carcinoma* (ccRCC) dan jaringan ginjal normal dari pasien yang sama. Dari 35 miRNA tersebut, sembilan miRNA mengalami *up-regulation* dan 26 miRNA mengalami *down-regulation*, dengan miRNA-106b dipilih sebagai titik referensi atau kontrol endogen. *Samaan et al.* [25] melaporkan bahwa tingkat ekspresi miR-210 yang tinggi ditemukan pada *clear cell Renal Cell Carcinoma* (ccRCC), dengan ekspresi yang lebih tinggi pada tumor metastatik. Tingginya ekspresi miR-210 dianggap sebagai biomarker independen yang menunjukkan prognosis buruk pada ccRCC. Selain itu, ekspresi miR-210 lebih tinggi pada sub tipe *clear cell* dan *papillary* dibandingkan dengan *chromophobe renal cell carcinoma* dan *oncocytoma*.

Dapat disimpulkan bahwa sebagian besar penelitian sebelumnya berfokus pada kelompok miRNA yang terbatas atau hanya pada sub tipe kanker ginjal tertentu. Berbeda dengan pendekatan tersebut, penelitian ini mempertimbangkan seluruh 1881 miRNA yang tersedia, dengan mengecualikan nilai kosong sehingga diperoleh sub set akhir sebanyak 1627 miRNA. Selanjutnya, sebanyak 35 miRNA terpilih ditentukan menggunakan metode seleksi fitur yang akan dibahas pada Bagian 3.1. Selain itu, penelitian ini mempertimbangkan TIGA sub tipe kanker ginjal sebagaimana diklasifikasikan dalam proyek kanker ginjal *The Cancer Genome Atlas* (TCGA).

2.3 Data Ekspresi miRNA dari TCGA

The Cancer Genome Atlas (TCGA) merupakan proyek kolaboratif berskala besar yang menyediakan data genomik, transkriptomik, ekspresi miRNA, serta informasi klinis pasien kanker. Dataset yang disediakan oleh *The Cancer Genome*

Atlas (TCGA) telah banyak dimanfaatkan dalam penelitian bioinformatika dan biologi komputasi karena kualitas data yang tinggi, standarisasi proses akuisisi data, serta cakupan sampel yang luas dan beragam. Keunggulan tersebut menjadikan TCGA sebagai salah satu sumber data utama dalam pengembangan dan evaluasi model komputasional untuk analisis kanker.

Dalam penelitian ini digunakan dataset *The Cancer Genome Atlas* (TCGA), khususnya data ekspresi miRNA yang telah diproses dan dipadankan dengan label kelas sub tipe kanker ginjal. Data ekspresi miRNA ini digunakan sebagai dasar dalam pengembangan model klasifikasi berbasis *machine learning* untuk membedakan sub tipe kanker ginjal secara multikelas.

2.3.1 Karakteristik Data Ekspresi miRNA

Data ekspresi miRNA pada *The Cancer Genome Atlas* (TCGA) umumnya disajikan dalam bentuk *quantitative read counts* yang merepresentasikan tingkat ekspresi masing-masing miRNA pada setiap sampel pasien. Setiap sampel memiliki ribuan fitur miRNA, di mana nilai ekspresi tersebut mencerminkan aktivitas regulasi gen pada tingkat molekuler. Data ini bersifat numerik, berdimensi tinggi, serta memiliki distribusi yang kompleks dan tidak linier.

Selain itu, data *The Cancer Genome Atlas* (TCGA) bersifat heterogen karena berasal dari berbagai latar belakang klinis pasien, termasuk perbedaan usia, stadium kanker, dan kondisi biologis lainnya. Heterogenitas ini menjadikan data ekspresi miRNA sangat kaya akan informasi, namun sekaligus menuntut metode analisis yang mampu menangkap pola kompleks antar fitur secara efektif.

2.3.2 Keunggulan Dataset TCGA

Dataset *The Cancer Genome Atlas* (TCGA) memiliki beberapa keunggulan utama yang menjadikannya relevan untuk penelitian klasifikasi kanker berbasis machine learning, antara lain:

1. Kualitas dan Standarisasi Data Data yang tersedia dalam *The Cancer Genome Atlas* (TCGA) dikumpulkan dan diproses menggunakan protokol yang terstandarisasi, sehingga menjamin konsistensi, akurasi, dan reliabilitas data antar sampel.
2. Cakupan Data yang Komprehensif *The Cancer Genome Atlas* (TCGA) menyediakan data *multi-omics* yang mencakup ekspresi gen, ekspresi

microRNA (miRNA), mutasi genom, variasi jumlah salinan gen, serta informasi klinis pasien, sehingga memungkinkan analisis kanker secara menyeluruh dari berbagai aspek molekuler.

3. Ketersediaan Label Klinis Setiap sampel dalam *The Cancer Genome Atlas* (TCGA) dilengkapi dengan anotasi klinis dan label subtype kanker yang tervalidasi, sehingga mendukung penerapan metode klasifikasi terawasi (*supervised learning*) dalam penelitian berbasis pembelajaran mesin.
4. Akses Terbuka untuk Penelitian *The Cancer Genome Atlas* (TCGA) dapat diakses secara bebas oleh peneliti, sehingga mendukung transparansi, replikasi, validasi hasil, serta pengembangan penelitian lanjutan di bidang bioinformatika dan onkologi.

Keunggulan-keunggulan tersebut menjadikan *The Cancer Genome Atlas* (TCGA) sebagai sumber data yang ideal untuk mengembangkan model klasifikasi subtype kanker ginjal berbasis ekspresi miRNA.

2.3.3 Tantangan Data Berdimensi Tinggi

Meskipun memiliki keunggulan yang signifikan, data ekspresi miRNA dari *The Cancer Genome Atlas* (TCGA) juga menghadirkan sejumlah tantangan, khususnya terkait dengan dimensi data yang sangat tinggi. Jumlah fitur miRNA yang mencapai ribuan sering kali jauh lebih besar dibandingkan jumlah sampel, sehingga berpotensi menyebabkan permasalahan *curse of dimensionality* dan *overfitting* pada model *machine learning*.

Selain itu, tidak semua miRNA memiliki kontribusi yang signifikan dalam membedakan subtype kanker ginjal. Keberadaan fitur yang tidak relevan atau redundan dapat menurunkan performa model serta meningkatkan kompleksitas komputasi. Oleh karena itu, diperlukan metode seleksi fitur yang efektif untuk mengidentifikasi miRNA yang paling diskriminatif dan relevan.

Dalam konteks penelitian ini, tantangan data berdimensi tinggi tersebut menjadi dasar penting dalam penerapan metode seleksi fitur berbasis *Least Absolute Shrinkage and Selection Operator* (LASSO), yang mampu mengurangi jumlah fitur sekaligus mempertahankan informasi penting untuk proses klasifikasi subtype kanker ginjal.

2.4 Seleksi Fitur pada Data Genomik

Seleksi fitur merupakan salah satu tahapan penting dalam pemodelan *machine learning*, khususnya pada data berdimensi tinggi seperti data ekspresi miRNA. Tujuan utama seleksi fitur adalah untuk mengidentifikasi fitur-fitur yang paling relevan terhadap permasalahan klasifikasi, mengurangi kompleksitas model, meningkatkan efisiensi komputasi, serta memperbaiki kemampuan generalisasi model terhadap data baru.

Pada data bioinformatika, setiap sampel umumnya merepresentasikan satu pasien, sedangkan jumlah fitur yang digunakan dapat mencapai ribuan miRNA. Kondisi ini menyebabkan ketidakseimbangan antara jumlah fitur dan jumlah sampel, yang dikenal sebagai permasalahan *curse of dimensionality*. Semakin tinggi dimensi data, semakin besar ruang pencarian yang harus dipelajari oleh model, sehingga dapat menurunkan performa klasifikasi dan meningkatkan kebutuhan komputasi secara signifikan.

Selain itu, data berdimensi tinggi juga meningkatkan risiko terjadinya *overfitting*, yaitu kondisi di mana model terlalu menyesuaikan diri dengan data pelatihan sehingga kehilangan kemampuan untuk melakukan prediksi yang akurat pada data yang belum pernah dilihat sebelumnya. *Overfitting* sering terjadi ketika model mempelajari pola yang bersifat kebetulan atau *noise* akibat banyaknya fitur yang tidak relevan atau redundan. Oleh karena itu, seleksi fitur menjadi langkah krusial untuk mengurangi jumlah fitur yang tidak informatif sekaligus mempertahankan fitur yang memiliki kontribusi signifikan terhadap proses klasifikasi.

Dalam konteks analisis data genomik, seleksi fitur tidak hanya bertujuan untuk meningkatkan kinerja model, tetapi juga untuk meningkatkan interpretabilitas hasil. Dengan mengurangi jumlah fitur, peneliti dapat lebih mudah mengidentifikasi miRNA yang berpotensi berperan sebagai biomarker penting dalam membedakan subtype kanker ginjal. Proses ini juga membantu menghubungkan hasil komputasional dengan pemahaman biologis yang lebih mendalam.

Berbagai pendekatan telah dikembangkan untuk melakukan seleksi fitur, salah satunya adalah metode berbasis regularisasi. Metode ini mengintegrasikan proses seleksi fitur secara langsung ke dalam proses pelatihan model dengan memberikan penalti terhadap kompleksitas model. Pendekatan ini sangat sesuai untuk data genomik yang memiliki dimensi tinggi, karena mampu menekan koefisien fitur yang kurang relevan hingga bernilai nol, sehingga hanya fitur-fitur

yang paling informatif yang dipertahankan. Oleh karena itu, metode seleksi fitur berbasis regularisasi, seperti *Least Absolute Shrinkage and Selection Operator* (LASSO), banyak digunakan dalam penelitian bioinformatika untuk mengatasi permasalahan *curse of dimensionality* dan *overfitting* secara simultan.

2.4.1 *Least Absolute Shrinkage and Selection Operator* (LASSO)

Least Absolute Shrinkage and Selection Operator (LASSO) merupakan metode regularisasi berbasis regresi yang mengintegrasikan seleksi fitur dengan proses pembelajaran model melalui penalti L1. Penalti ini mendorong koefisien fitur yang kurang relevan menjadi nol, sehingga hanya fitur dengan kontribusi signifikan yang dipertahankan. Dalam konteks data genomik seperti ekspresi miRNA yang memiliki dimensi tinggi, LASSO sangat efektif dalam mengurangi kompleksitas model, mencegah *overfitting*, serta meningkatkan interpretabilitas hasil. Oleh karena itu, LASSO banyak digunakan dalam studi bioinformatika sebagai metode seleksi fitur untuk identifikasi biomarker kanker.

2.5 Pembelajaran Mesin (*Machine Learning*)

Perkembangan *machine learning* telah memberikan kontribusi signifikan dalam bidang biomedis, khususnya dalam analisis data genomik untuk klasifikasi penyakit kanker. Kemampuan *machine learning* dalam mempelajari pola kompleks dari data berdimensi tinggi menjadikannya sangat sesuai untuk menangani data ekspresi gen dan miRNA yang memiliki karakteristik nonlinier serta heterogen. Dalam konteks kanker ginjal, pendekatan *machine learning* memungkinkan pemanfaatan informasi molekuler untuk membedakan subtype kanker secara lebih akurat dibandingkan metode statistik konvensional.

2.5.1 Klasifikasi Multikelas

Klasifikasi multikelas merupakan permasalahan klasifikasi di mana setiap sampel dapat diklasifikasikan ke dalam lebih dari dua kelas. Pada kasus kanker ginjal, permasalahan ini muncul karena adanya beberapa subtype kanker yang memiliki karakteristik biologis yang berbeda, seperti *clear cell renal cell carcinoma*, *papillary renal cell carcinoma*, *chromophobe renal cell carcinoma*, serta subtype lainnya. Oleh karena itu, pendekatan klasifikasi multikelas menjadi lebih relevan dibandingkan klasifikasi biner.

Dalam klasifikasi multikelas berbasis data genomik, tantangan utama terletak pada tingginya dimensi data dan kemiripan pola ekspresi antar subtype tertentu. Model klasifikasi harus mampu membedakan pola ekspresi miRNA yang halus dan kompleks agar dapat menghasilkan prediksi yang akurat. Berbagai algoritma *machine learning* telah digunakan untuk permasalahan ini, mulai dari metode konvensional seperti *Support Vector Machine* hingga pendekatan *deep learning* yang mampu memodelkan hubungan nonlinier secara lebih mendalam.

2.5.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) merupakan salah satu arsitektur *deep learning* yang termasuk dalam keluarga *Recurrent Neural Network* (RNN). *Long Short-Term Memory* (LSTM) dirancang untuk mengatasi permasalahan vanishing gradient yang sering terjadi pada *Recurrent Neural Network* (RNN) konvensional, sehingga mampu mempelajari ketergantungan jangka panjang dalam data. Meskipun awalnya dikembangkan untuk data sekuensial, *Long Short-Term Memory* (LSTM) juga telah banyak diterapkan pada data non-sekuensial berdimensi tinggi, termasuk data ekspresi gen dan miRNA.

Arsitektur *Long Short-Term Memory* (LSTM) terdiri dari beberapa komponen utama, yaitu *input gate*, *forget gate*, dan *output gate*, yang bekerja bersama untuk mengatur aliran informasi dalam jaringan. Mekanisme ini memungkinkan LSTM untuk mempertahankan informasi penting dan mengabaikan informasi yang kurang relevan selama proses pembelajaran. Kemampuan tersebut menjadikan *Long Short-Term Memory* (LSTM) efektif dalam menangkap pola kompleks pada data genomik yang memiliki hubungan tidak linier antar fitur.

Dalam penelitian klasifikasi kanker, *Long Short-Term Memory* (LSTM) telah menunjukkan kinerja yang menjanjikan dalam membedakan subtype kanker berdasarkan data molekuler. Dengan memanfaatkan fitur miRNA terpilih hasil seleksi fitur, *Long Short-Term Memory* (LSTM) dapat mempelajari representasi yang lebih diskriminatif untuk setiap subtype kanker ginjal. Oleh karena itu, *Long Short-Term Memory* (LSTM) dipilih sebagai model klasifikasi dalam penelitian ini karena kemampuannya dalam menangani kompleksitas data, meningkatkan akurasi klasifikasi multikelas, serta mendukung pengembangan pendekatan diagnosis berbasis *precision medicine*.

2.5.3 Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) merupakan jenis jaringan saraf tiruan dasar yang banyak digunakan untuk tugas klasifikasi dan regresi [26]. MLP belajar dengan menyesuaikan bobot yang dikenakan pada setiap neuron untuk meminimalkan perbedaan antara hasil prediksi dengan nilai aktual. Proses pembelajaran ini umumnya menggunakan teknik *backpropagation*. Algoritma *backpropagation* menghitung turunan fungsi error terhadap setiap bobot dan memperbarui bobot tersebut menggunakan persamaan berikut:

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \eta \frac{\partial E}{\partial \omega_{ij}} \quad (2.1)$$

di mana η merupakan *learning rate* dan E adalah fungsi error.

Multi-Layer Perceptron (MLP) terdiri dari beberapa lapisan neuron, yaitu *input layer*, *hidden layer*, dan *output layer*. Lapisan masukan (*input layer*) memiliki jumlah neuron yang bergantung pada jumlah fitur data yang digunakan. Lapisan ini berfungsi untuk menerima dan meneruskan data mentah ke lapisan selanjutnya [27].

Lapisan tersembunyi (*hidden layer*) berperan dalam proses pembelajaran dan komputasi untuk mempelajari pola-pola kompleks dalam data menggunakan fungsi aktivasi (*activation function*). Output neuron pada *hidden layer* pertama dihitung menggunakan persamaan berikut:

$$h_{ij} = f \left(\sum_{k=1}^{n_0} \omega_{kj} x_k \right) \quad (2.2)$$

di mana x_k merupakan nilai fitur masukan, ω_{kj} adalah bobot dari neuron input ke- k ke neuron tersembunyi ke- j , dan n_0 adalah jumlah neuron input.

Output neuron pada *hidden layer* selanjutnya dihitung menggunakan persamaan:

$$h_j^{(i)} = f \left(\sum_{k=1}^{n_{i-1}} \omega_{kj}^{(i)} h_k^{(i-1)} \right) \quad (2.3)$$

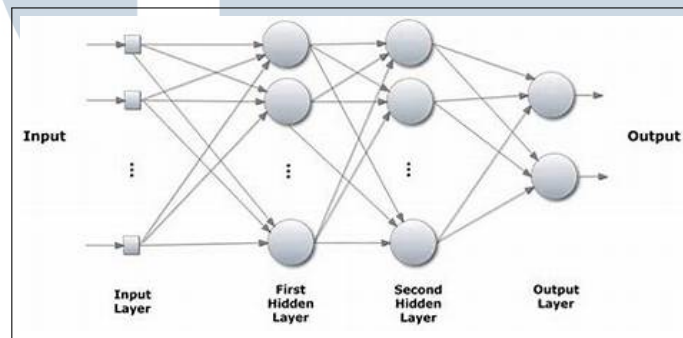
di mana f merupakan fungsi aktivasi, $\omega_{kj}^{(i)}$ adalah bobot dari neuron ke- k pada lapisan sebelumnya ke neuron ke- j pada lapisan ke- i , dan $h_k^{(i-1)}$ adalah keluaran neuron ke- k pada lapisan sebelumnya.

Hasil akhir pada lapisan keluaran (*output layer*) dihitung dengan persamaan berikut:

$$y_l = f \left(\sum_{k=1}^{n_N} \omega_{kl}^{(N+1)} h_k^N \right) \quad (2.4)$$

di mana h_k^N adalah output dari *hidden layer* terakhir dan $\omega_{kl}^{(N+1)}$ merupakan bobot dari neuron ke- k pada *hidden layer* terakhir ke neuron keluaran ke- l .

Lapisan keluaran bertugas untuk menghasilkan prediksi akhir dari *Multi-Layer Perceptron* (MLP). Jumlah neuron pada lapisan ini bergantung pada jenis permasalahan yang diselesaikan. Untuk regresi dan klasifikasi biner digunakan satu neuron keluaran, sedangkan untuk klasifikasi multikelas digunakan satu neuron untuk setiap kelas. Seluruh lapisan *Multi-Layer Perceptron* (MLP) bekerja secara iteratif untuk mempelajari pola kompleks dalam data dan menghasilkan prediksi kelas yang akurat.



Gambar 2.1. Kerangka *Multi-Layer Perceptron* (MLP)

2.6 Evaluasi Kinerja Model Klasifikasi

Evaluasi kinerja model klasifikasi merupakan tahapan penting untuk menilai sejauh mana model pembelajaran mesin atau *deep learning* mampu membedakan kelas-kelas yang terdapat pada data secara akurat dan konsisten. Evaluasi ini dilakukan menggunakan sejumlah metrik kuantitatif dan visualisasi pendukung yang memberikan gambaran menyeluruh mengenai keunggulan dan keterbatasan model.

Penggunaan beberapa metrik evaluasi menjadi penting terutama pada kasus klasifikasi multikelas dengan distribusi data yang tidak seimbang, seperti pada klasifikasi subtype kanker ginjal berbasis ekspresi miRNA. Dalam kondisi tersebut, metrik tunggal seperti *accuracy* belum tentu mencerminkan performa model secara komprehensif. Oleh karena itu, penelitian ini menggunakan beberapa metrik evaluasi utama, yaitu *accuracy*, *confusion matrix*, dan *Matthews Correlation*

Coefficient (MCC).

2.6.1 Accuracy

Accuracy merupakan metrik evaluasi yang mengukur proporsi prediksi yang dilakukan secara benar oleh model terhadap keseluruhan jumlah data yang dievaluasi. Metrik ini memberikan gambaran umum mengenai tingkat keberhasilan model dalam mengklasifikasikan data ke dalam kelas yang sesuai.

Secara matematis, *accuracy* dirumuskan sebagai berikut:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (2.5)$$

dengan *TP* (*true positive*) menyatakan jumlah data yang diprediksi benar sebagai kelas positif, *TN* (*true negative*) menyatakan jumlah data yang diprediksi benar sebagai kelas negatif, sedangkan *FP* (*false positive*) dan *FN* (*false negative*) masing-masing menyatakan kesalahan prediksi model.

Meskipun *accuracy* sering digunakan karena kemudahannya dalam interpretasi, metrik ini memiliki keterbatasan pada dataset dengan distribusi kelas yang tidak seimbang. Nilai *accuracy* yang tinggi dapat diperoleh meskipun model gagal mengklasifikasikan kelas minoritas dengan baik. Oleh karena itu, *accuracy* perlu dikombinasikan dengan metrik evaluasi lain untuk memperoleh penilaian performa model yang lebih representatif.

2.6.2 Confusion Matrix

Confusion matrix merupakan alat evaluasi yang menyajikan hasil prediksi model dalam bentuk matriks yang membandingkan antara label aktual dan label hasil prediksi. Setiap baris pada *confusion matrix* merepresentasikan kelas aktual, sedangkan setiap kolom merepresentasikan kelas prediksi.

Melalui *confusion matrix*, dapat diketahui jumlah prediksi yang benar maupun kesalahan klasifikasi pada masing-masing kelas. Informasi ini memungkinkan analisis yang lebih mendalam terhadap pola kesalahan model, seperti kelas-kelas yang sering tertukar atau sulit dibedakan.

Pada permasalahan klasifikasi multikelas subtype kanker ginjal, *confusion matrix* berperan penting dalam mengidentifikasi performa model pada setiap subtype, yaitu *Kidney Chromophobe (KICH)*, *Kidney Renal Clear Cell Carcinoma*

(KIRC), dan *Kidney Renal Papillary Cell Carcinoma* (KIRP), sehingga evaluasi tidak hanya bergantung pada nilai agregat seperti accuracy.

2.6.3 Matthews Correlation Coefficient (MCC)

Matthews Correlation Coefficient (MCC) merupakan metrik evaluasi yang mengukur kualitas klasifikasi dengan mempertimbangkan seluruh elemen pada *confusion matrix*, yaitu TP , TN , FP , dan FN . *Matthews Correlation Coefficient* (MCC) menghasilkan nilai dalam rentang -1 hingga 1 , di mana nilai 1 menunjukkan klasifikasi sempurna, nilai 0 menunjukkan performa setara dengan tebakan acak, dan nilai -1 menunjukkan kesalahan klasifikasi total.

Secara matematis, *Matthews Correlation Coefficient* (MCC) untuk klasifikasi biner dirumuskan sebagai berikut:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (2.6)$$

Pada kasus klasifikasi multikelas, *Matthews Correlation Coefficient* (MCC) dihitung menggunakan generalisasi dari *confusion matrix* sehingga tetap mampu memberikan evaluasi performa yang stabil meskipun terdapat ketidakseimbangan jumlah data antar kelas. Oleh karena itu, *Matthews Correlation Coefficient* (MCC) dianggap sebagai metrik yang lebih robust dan representatif dibandingkan *accuracy* dalam mengevaluasi kinerja model klasifikasi multikelas berbasis data genomik.

