

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Kanker ginjal merupakan salah satu jenis kanker dengan tingkat mortalitas yang tinggi dan hingga saat ini masih sulit dideteksi secara dini melalui metode klinis konvensional. Meskipun termasuk dalam sepuluh besar penyebab kematian akibat kanker, penelitian terkait kanker ginjal masih relatif terbatas dibandingkan dengan jenis kanker lainnya. Kondisi ini menyebabkan kanker ginjal kerap terabaikan dalam komunitas medis, sehingga menghambat pengembangan metode diagnosis dan terapi yang lebih efektif [1]. Selama beberapa dekade, pilihan pengobatan bagi pasien kanker ginjal umumnya terbatas pada tindakan pembedahan, dan pada banyak kasus, harapan hidup pasien pascadiagnosis masih tergolong rendah [2]. Oleh karena itu, deteksi dini dan diagnosis yang akurat menjadi faktor krusial dalam meningkatkan peluang kelangsungan hidup pasien kanker ginjal [3].

Seiring dengan keterbatasan metode klinis tradisional, penelitian berbasis biomarker molekuler mulai berkembang dan memberikan kontribusi penting dalam memahami mekanisme awal terjadinya kanker ginjal. Studi biomarker memungkinkan diagnosis yang lebih presisi serta membantu mengurangi ketidakpastian dalam penentuan karakteristik penyakit [4]. Namun demikian, meskipun pemahaman genomik dan kemajuan teknologi bioinformatika terus mengalami perkembangan yang signifikan, masih terdapat banyak aspek yang belum sepenuhnya dipahami, khususnya terkait heterogenitas stadium dan subtipo kanker ginjal [5]. Oleh sebab itu, diperlukan pengembangan pendekatan baru yang mampu mendeteksi kanker ginjal secara lebih dini serta mengklasifikasikan stadium dan subtipo penyakit secara akurat. Pendekatan tersebut diharapkan dapat membantu tenaga medis dalam menentukan strategi terapi yang tepat bagi setiap pasien, meminimalkan efek samping pengobatan, serta meningkatkan tingkat kelangsungan hidup pasien [3].

Dalam beberapa tahun terakhir, berbagai upaya telah dilakukan untuk membedakan subtipo kanker ginjal, salah satunya melalui analisis informasi genetik pasien. Pendekatan berbasis ekspresi microRNA (miRNA) telah menarik perhatian banyak peneliti, terutama sejak kemajuan teknologi ekstraksi dan analisis miRNA

dalam dua dekade terakhir [6]. miRNA diketahui berperan penting dalam regulasi ekspresi gen, dan ketidakseimbangan ekspresi miRNA (*up-regulation and down-regulation*) telah terbukti berkaitan erat dengan perkembangan berbagai jenis kanker, termasuk kanker ginjal. Salah satu sumber utama data genomik kanker yang banyak dimanfaatkan adalah *The Cancer Genome Atlas* (TCGA), hasil kolaborasi antara *National Cancer Institute* (NCI) dan *National Human Genome Research Institute* (NHGRI). TCGA menyediakan data genomik multidimensional yang komprehensif untuk 33 jenis kanker, termasuk kanker ginjal [7].

Penelitian ini berfokus pada deteksi dan klasifikasi subtipe kanker ginjal dengan memanfaatkan variasi ekspresi miRNA sebagai dasar pembeda antar subtipe. Berbagai penelitian sebelumnya menunjukkan adanya hubungan yang kuat antara variasi regulasi miRNA dan karakteristik biologis subtipe kanker [8]. Oleh karena itu, variasi ekspresi miRNA yang disebabkan oleh perbedaan pola *up-regulation* dan *down-regulation* antar subtipe kanker ginjal dimanfaatkan sebagai fitur utama dalam proses klasifikasi. Melalui pendekatan yang diusulkan, penelitian ini berhasil melakukan seleksi fitur berbobot tinggi sehingga diperoleh 35 miRNA yang paling diskriminatif dari total 1881 miRNA yang teridentifikasi dalam dataset TCGA untuk kanker ginjal. Selanjutnya, melalui proses pelatihan model yang mendalam (*deep training*), sistem yang dibangun mampu membedakan tiga subtipe kanker ginjal yang terdapat dalam proyek TCGA.

Secara lebih spesifik, penelitian ini menggunakan LASSO sebagai metode seleksi fitur untuk mengidentifikasi miRNA yang paling berkontribusi dalam membedakan subtipe kanker ginjal [9]. miRNA terpilih kemudian digunakan sebagai masukan ke dalam model klasifikasi berbasis *machine learning*. Metode klasifikasi yang diterapkan adalah *Long Short-Term Memory* (LSTM), salah satu arsitektur *Recurrent Neural Network* yang mampu menangani data sekuensial maupun non-sekuensial secara efektif [10, 11]. Evaluasi kinerja model dilakukan menggunakan *Data Analysis Protocol* dan *Matthews Correlation Coefficient* (MCC) sebagai metrik penilaian [12]. Meskipun hasil yang diperoleh menunjukkan kinerja yang menjanjikan, efektivitas klinis dari subset miRNA terpilih dalam mendiagnosis subtipe kanker ginjal secara spesifik masih memerlukan validasi lebih lanjut melalui studi klinis.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana performa pendekatan *machine learning* multikelas dalam mengklasifikasikan subtipe kanker ginjal menggunakan data ekspresi miRNA dari TCGA?
- b. Bagaimana proses seleksi fitur menggunakan LASSO dalam mengidentifikasi miRNA yang paling diskriminatif untuk klasifikasi subtipe kanker ginjal?
- c. Sejauh mana model *Long Short-Term Memory* (LSTM) mampu mengklasifikasikan subtipe kanker ginjal berdasarkan subset miRNA terpilih?
- d. Bagaimana tingkat akurasi dan keandalan model klasifikasi yang diusulkan berdasarkan metrik evaluasi *Matthews Correlation Coefficient* (MCC)?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

- a. Mengembangkan model *machine learning* multikelas untuk klasifikasi subtipe kanker ginjal menggunakan data ekspresi miRNA dari TCGA.
- b. Mengimplementasikan metode LASSO untuk melakukan seleksi fitur miRNA yang paling relevan dan diskriminatif.
- c. Menerapkan algoritma *Long Short-Term Memory* (LSTM) sebagai model klasifikasi subtipen kanker ginjal berbasis data miRNA.
- d. Mengevaluasi kinerja model klasifikasi yang diusulkan menggunakan metrik akurasi dan *Matthews Correlation Coefficient* (MCC).
- e. Menganalisis kemampuan subset miRNA terpilih dalam merepresentasikan karakteristik subtipen kanker ginjal.

1.4 Urgensi Penelitian

Penelitian ini memiliki urgensi yang tinggi karena mengintegrasikan analisis data molekuler dengan pendekatan kecerdasan buatan untuk mendukung deteksi dan klasifikasi subtipe kanker ginjal secara lebih akurat. Pemanfaatan data ekspresi miRNA memungkinkan analisis yang lebih spesifik terhadap mekanisme biologis kanker pada tingkat molekuler, yang tidak dapat sepenuhnya dicapai melalui metode diagnostik konvensional.

Penerapan pendekatan *machine learning* multikelas memungkinkan identifikasi pola kompleks dan nonlinier dalam data genomik berukuran besar, sehingga meningkatkan ketepatan dalam membedakan subtipe kanker ginjal. Kemampuan ini menjadi sangat penting mengingat heterogenitas biologis kanker ginjal yang sering kali menyulitkan proses diagnosis dan penentuan terapi yang tepat.

Penelitian ini berpotensi mendukung pengembangan sistem diagnosis berbasis *precision medicine*, di mana keputusan klinis dapat disesuaikan dengan karakteristik molekuler masing-masing pasien. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam upaya peningkatan akurasi diagnosis, optimalisasi strategi pengobatan, serta peningkatan tingkat kelangsungan hidup pasien kanker ginjal.

1.5 Luaran Penelitian

Luaran yang diharapkan dari penelitian ini meliputi:

1. Artikel ilmiah yang dipublikasikan pada jurnal atau prosiding konferensi.
2. Model klasifikasi dan *pipeline* analisis data miRNA yang dapat digunakan sebagai referensi penelitian lanjutan.

1.6 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

- a. Memberikan kontribusi ilmiah dalam pengembangan pendekatan *machine learning* berbasis miRNA untuk klasifikasi multikelas subtipe kanker ginjal.

- b. Menambah referensi akademik terkait pemanfaatan metode LASSO dan *Long Short-Term Memory* (LSTM) dalam analisis data genomik, khususnya pada kasus kanker ginjal.
- c. Memberikan alternatif pendekatan komputasional yang dapat membantu tenaga medis dalam mendukung proses diagnosis subtipe kanker ginjal secara lebih akurat.
- d. Menjadi dasar pengembangan sistem pendukung keputusan (*decision support system*) berbasis data genomik untuk kanker ginjal.
- e. Menyediakan kerangka kerja (*pipeline*) analisis data miRNA mulai dari seleksi fitur hingga klasifikasi multikelas menggunakan *machine learning*.
- f. Menjadi acuan bagi penelitian selanjutnya dalam pengembangan model klasifikasi kanker berbasis data ekspresi gen atau miRNA.

