

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kanker prostat adalah kanker kedua yang paling sering didiagnosis di antara pria di seluruh dunia, dengan perkiraan 1.414.000 kasus kanker baru dan 375.304 kematian pada tahun 2020 [1]. Kanker prostat adalah kanker yang paling sering didiagnosis di 112 negara, dan penyebab utama kematian akibat kanker di 48 negara. Selain itu ada juga faktor-faktor yang mempengaruhi angka kematian dari penyakit kanker prostat. Salah satu faktor kuncinya adalah perbedaan ras pada kanker prostat. Saat ini, teknik pembelajaran mesin telah digunakan dalam penelitian untuk mengembangkan lebih lanjut diagnosis kanker prostat [2], [3]. Teknik pembelajaran mesin dapat menemukan pola dari kumpulan data yang kompleks dan secara efektif memprediksi hasil dari kanker prostat [4].

Perbedaan ras memainkan peran penting dalam diagnosis kanker prostat, yang memengaruhi hasil deteksi di berbagai kelompok ras. Di Amerika Serikat, pria kulit hitam 1,76 kali lebih mungkin didiagnosis menderita kanker prostat dan memiliki tingkat kematian 2,14 kali lebih tinggi dibandingkan dengan pria kulit putih [5]. Selain itu, pria kulit hitam lebih mungkin didiagnosis pada stadium yang lebih lanjut dari penyakit ini [6]. Kesenjangan ini menyoroti perlunya analisis spesifik ras lebih lanjut untuk mengurangi risiko kesalahan diagnosis dan meningkatkan akurasi deteksi. Oleh karena itu, memahami perbedaan ras sangat penting untuk memastikan diagnosis kanker prostat yang lebih akurat.

Pada penelitian terdahulu, studi yang menggunakan data miRNA dengan algoritma XGBoost belum ada yang menggunakan. Studi yang dibahas oleh Kalaiyarasi M. et al. melakukan penelitian dengan menggunakan set data miRNA pada kanker prostat. Penelitian ini menggunakan seleksi fitur dari website untuk melihat miRNA yang paling signifikan yang mendapatkan 209 miRNA. Penelitian ini mencapai akurasi 92% dan AUC 95% dengan menggunakan algoritma SVM [7]. Dari penelitian sebelumnya, penelitian tersebut menggunakan website untuk

mencari miRNA yang signifikan dan menggunakan dataset yang berbeda. Penelitian ini mengambil data dari GDC TCGA XenaBrowser dan menggunakan algoritma XGBoost yang dimana belum digunakan untuk dalam penelitian deteksi stadium kanker prostat. Selain itu, penelitian ini menggunakan seleksi fitur yang berbeda dan untuk menemukan fitur miRNA signifikan seperti EdgeR, *Lasso* + RFE dan ROC serta melakukan validasi *AUC Score*. Penelitian menggunakan berbagai teknik penyeimbangan untuk mendapatkan data yang seimbang. Data yang digunakan berfokus pada pasien dengan ras kulit putih dan akan dilakukan tes pada ras kulit hitam untuk menganalisis pengaruh ras terhadap performa model. Dengan menggunakan semua algoritma dan seleksi fitur yang berbeda, diharapkan penelitian ini dapat membantu melakukan klasifikasi stadium kanker prostat yang lebih akurat sehingga pasien dapat mengetahui penyakit kanker prostat lebih awal.

1.2. Rumusan Masalah

1. Bagaimana cara membangun model deteksi dini kanker prostat menggunakan algoritma XGBoost?
2. Bagaimana performa data miRNA berdasarkan ras spesifik dalam mendeteksi kanker prostat secara akurat?

1.3. Tujuan Penelitian

1. Penelitian ini bertujuan untuk membangun model deteksi stadium kanker prostat untuk deteksi dini menggunakan algoritma Xgboost dan performa menggunakan metrik akurasi, presisi, *recall* dan *f1-score*.
2. Penelitian ini bertujuan untuk meneliti lebih dalam mengenai performa data miRNA pada ras spesifik dalam masalah deteksi kanker prostat dengan mengevaluasi performa model pada XGBoost

1.4. Urgensi Penelitian

Masih dibutuhkan pengembangan untuk deteksi kanker prostat khususnya *staging* menggunakan data miRNA pada ras spesifik. Penggunaan *staging* sekarang dilakukan tanpa AI menggunakan *TNM system*. Sistem tersebut digunakan oleh dokter untuk melakukan *staging* dengan mengandalkan tumor, *node* dan metastasis.

1.5. Luaran Penelitian

1. *Pipeline* pengembangan sistem AI dalam mendeteksi *staging* pada kanker prostat yang kedepannya jika diuji *valid* dan baik dapat dimanfaatkan oleh pihak rumah sakit
2. Menyusun artikel ilmiah dari hasil eksperimen dan evaluasi model AI untuk klasifikasi stadium kanker prostat menggunakan algoritma Xgboost dengan data nyata.

1.6. Manfaat Penelitian

1. Memberikan kontribusi ilmiah dalam pengembangan model deteksi dini kanker prostat dengan pendekatan machine learning, khususnya menggunakan algoritma XGBoost, serta memperkaya literatur terkait pemanfaatan data genomik dalam deteksi penyakit kanker.
2. Memberikan wawasan kepada praktisi di bidang kesehatan dan bioinformatika mengenai pentingnya mempertimbangkan faktor ras dalam analisis data genomik untuk meningkatkan akurasi deteksi kanker prostat.
3. Memberikan kontribusi terhadap pengembangan sistem deteksi dini kanker prostat yang lebih personal dengan mempertimbangkan faktor ras sebagai variable penting. Selain itu, pasien juga dapat diagnosis yang lebih cepat dan akurat sehingga meningkatkan peluang kesembuhan.

UNIVERSITAS
MULTIMEDIA
NUSANTARA