

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Model *machine learning* dapat diterapkan untuk klasifikasi subtipe kanker payudara berbasis data RNA dengan menggunakan pendekatan *One-vs-Rest* (OvR). Berdasarkan pemodelan yang telah dilakukan, metode *OnevsRest* terbukti efektif dalam mengklasifikasikan subtipe kanker payudara berbasis data RNA, karena memungkinkan pembentukan model klasifikasi secara terpisah untuk setiap subtipe. Dengan pendekatan ini, analisis pengaruh masing-masing fitur (gen) pada setiap subtipe menjadi lebih terfokus dan mudah diinterpretasikan. Hal ini sangat penting dalam konteks biologis, karena setiap subtipe kanker payudara memiliki pola ekspresi yang khas, sehingga pemodelan terpisah dapat membantu mengidentifikasi gen-gen kunci yang relevan untuk masing-masing subtipe.
2. Dari hasil perbandingan performa model, algoritma *XGBoost* menunjukkan performa yang paling unggul dengan nilai akurasi sebesar 93%, *precision* 94%, *recall* 90% dan *F1-Score* 91%. Nilai-nilai tersebut menunjukkan bahwa *XGBoost* memiliki kemampuan dalam menghasilkan prediksi yang tidak hanya akurat, tetapi juga seimbang antara sensitivitas dan presisinya. Dengan demikian, dapat disimpulkan bahwa *XGBoost* merupakan model paling optimal yang dapat digunakan dalam klasifikasi subtipe kanker payudara berbasis data RNA.
3. Berdasarkan hasil identifikasi dan analisis gen melalui pendekatan *feature importance* dan *differential expression analysis*, diperoleh sepuluh gen kunci untuk masing-masing subtipe kanker payudara yang dianggap memiliki pengaruh signifikan, baik terhadap proses klasifikasi

maupun mekanisme biologis spesifik yang membedakan masing-masing subtipe. Berikut adalah daftar 10 gen kunci dari masing-masing subtipe kanker payudara:

- Kanker Payudara Subtipe Basal : FOXA1, MLPH, ABLIM3, FOXC1, FAM69C, BCAS1, CYP4F8, C16orf57, XBP1, TSPAN2.
- Kanker Payudara Subtipe HER2 : C6orf155, MPHOSPH6(MPP6), ZP2, TMEM45B, ERBB2, GPBP1, CEP55, C17orf37(MIEN1), UFC1, TPBG.
- Kanker Payudara Subtipe Luminal A : CEP55, TPX2, TTK, CENPA, AURKA, BCL2, FZD10, HELLS, BTG2, CENPN.
- Kanker Payudara Subtipe Luminal B : TRIM29, ESPL1, TIMELESS, KRT5, EXO1, MIA, PDE9A, SFRP1, UBE2T, SGOL2.

5.2 Saran

Berikut merupakan beberapa rekomendasi yang dapat dijadikan bahan evaluasi dan pengembangan untuk penelitian-penelitian berikutnya:

1. Sebagian besar gen kunci yang berhasil teridentifikasi dalam penelitian ini menunjukkan kesesuaian dengan karakteristik masing-masing subtipe kanker payudara. Namun, terdapat beberapa temuan yang menunjukkan perbedaan hasil dengan penelitian sebelumnya, seperti gen FOXC1 dan MIA yang pada penelitian terdahulu sering ditemukan mengalami peningkatan ekspresi (over ekspresi) pada berbagai subtipe justru menunjukkan penurunan ekspresi dalam penelitian ini. Selain itu, beberapa gen kunci lainnya juga belum diketahui terlibat dalam jalur biologis tertentu dan belum banyak dibahas dalam literatur ilmiah terkait kanker payudara. Temuan-temuan ini menunjukkan bahwa diperlukan analisis lanjutan, terutama melalui uji eksperimental di laboratorium, untuk memvalidasi peran gen-gen tersebut. Hal ini penting dilakukan, mengingat karakteristik kanker payudara sangat bervariasi antar pasien

dan bersifat individual (*case by case*), sehingga tidak terdapat satu pola ekspresi genetik yang sepenuhnya sama antar kasus.

2. Menggunakan dataset dengan jumlah sampel pasien yang lebih besar dan seimbang antar subtype, guna meningkatkan keakuratan model serta memastikan hasil klasifikasi lebih representatif dan dapat digeneralisasi pada populasi yang lebih luas.
3. Melakukan optimasi parameter pada algoritma *machine learning* yang digunakan menggunakan metode seperti *Grid Search* dan *Random Search* untuk memperoleh nilai parameter terbaik, sehingga model dapat mencapai performa klasifikasi yang optimal.
4. Melakukan seleksi fitur (gen) dengan memfokuskan pada proses biologis tertentu, sehingga jumlah gen yang digunakan dalam analisis lebih relevan, terarah, dan hasilnya lebih mudah diinterpretasikan.