

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini berhasil mengintegrasikan metode seleksi fitur berbasis *Fuzzy Mutual Information* (FMI) dengan tiga model klasifikasi, yaitu *Support Vector Machine* (SVM), *Random Forest* (RF), dan *K-Nearest Neighbor* (KNN), untuk mengklasifikasikan stadium kanker lambung berdasarkan ekspresi miRNA. Analisis performa model menunjukkan bahwa *Random Forest* memiliki sensitivitas dan F1-score tertinggi, menandakan kemampuan yang baik dalam mendeteksi stadium lanjut, sedangkan SVM memiliki kemampuan diskriminatif terbaik berdasarkan nilai AUC, dan KNN menunjukkan spesifisitas tertinggi untuk stadium awal. Hasil ini menegaskan adanya *trade-off* klasik antara sensitivitas dan spesifisitas yang perlu diperhatikan dalam konteks klinis.

Evaluasi stabilitas fitur menunjukkan bahwa beberapa miRNA terpilih secara konsisten muncul di seluruh fold dan lintas model. Analisis korelasi menegaskan bahwa sebagian besar miRNA memberikan informasi komplementer, dengan hanya sedikit pasangan yang berkorelasi tinggi, sehingga risiko *multicollinearity* rendah. Relevansi biologis dari miRNA terpilih didukung oleh literatur terkait peran mereka dalam tumorigenesis, progresi, dan outcome klinis pada kanker lambung, memperkuat nilai biomarker yang diidentifikasi. Secara keseluruhan, pipeline yang dikembangkan terbukti stabil, *robust*, dan mampu menyeleksi biomarker miRNA yang relevan, sekaligus memberikan performa klasifikasi yang dapat diandalkan dalam membedakan stadium awal dan lanjut kanker lambung.

5.2 Saran

Hasil penelitian menunjukkan bahwa integrasi metode seleksi fitur berbasis *Fuzzy Mutual Information* dengan model SVM, RF, dan KNN mampu menghasilkan performa klasifikasi yang stabil serta menyeleksi biomarker miRNA yang relevan untuk membedakan stadium kanker lambung. Meskipun demikian, terdapat beberapa keterbatasan yang dapat dioptimalkan untuk meningkatkan kualitas prediksi, validitas, dan generalisasi model. Keterbatasan ini mencakup jumlah sampel yang relatif terbatas, kemungkinan adanya fitur yang kurang informatif atau redundan, serta potensi perbaikan pada strategi pemilihan model dan metode klasifikasi. Oleh karena itu, beberapa saran pengembangan penelitian disusun untuk memperluas cakupan, meningkatkan akurasi, dan memperkuat relevansi klinis dari *pipeline* yang diusulkan:

1. Memperluas dataset dengan jumlah sampel yang lebih besar dan anotasi klinis yang lebih lengkap. Hal ini dapat meningkatkan kemampuan generalisasi model dan memperkuat validitas hasil klasifikasi pada populasi pasien yang lebih luas.
2. Mengadopsi metode seleksi fitur yang lebih canggih atau *hybrid*, misalnya kombinasi *filter* dan *wrapper*, atau integrasi dengan metode *embedded* pada model *ensemble*, untuk meningkatkan stabilitas dan kemampuan diskriminatif fitur.

3. Mengeksplorasi model klasifikasi tambahan atau teknik *ensemble learning* yang lebih kompleks, seperti XGBoost atau LightGBM, untuk meningkatkan performa klasifikasi terutama pada dataset tidak seimbang.
4. Melakukan validasi eksternal menggunakan dataset independen dari sumber lain atau data *multi-omics*, untuk menilai robustitas *pipeline* dan potensi translasi ke aplikasi klinis.

Implementasi saran tersebut diharapkan dapat memperkuat kualitas penelitian, meningkatkan akurasi prediksi, dan mendukung identifikasi biomarker miRNA yang lebih tepat dan relevan secara klinis untuk kanker lambung.

