BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis pemodelan klasifikasi keterbacaan abstrak jurnal Informatika, diperoleh tiga kesimpulan utama sebagai berikut :

1. Pelabelan tingkat keterbacaan abstrak ilmiah berhasil dilakukan secara objektif menggunakan algoritma *K-Means Clustering* berdasarkan pembobotan fitur TF-IDF. Proses klasterisasi menghasilkan tiga kategori keterbacaan, yaitu Mudah, Sedang, dan Sulit, yang selanjutnya digunakan sebagai label pada proses klasifikasi. Evaluasi menggunakan *Silhouette Score* memperoleh nilai 0,47, yang menunjukkan bahwa hasil pengelompokan memiliki struktur *Cluster* yang cukup baik untuk digunakan pada proses pelabelan otomatis.

2. Model klasifikasi *Multinomial Naive Bayes* berhasil dibangun menggunakan fitur TF-IDF dengan menerapkan *Synthetic Minority Over-sampling Technique* (SMOTE) pada data latih untuk mengatasi ketidakseimbangan jumlah data antar kelas. Penerapan SMOTE menghasilkan distribusi data latih yang seimbang, yaitu 1.688 dokumen pada setiap kategori dengan total 5.064 data latih, sehingga model memperoleh kesempatan yang sama dalam mempelajari karakteristik setiap tingkat keterbacaan.

3. Pengujian terhadap model klasifikasi menunjukkan bahwa *Multinomial Naive Bayes* memperoleh akurasi sebesar 94,83% disertai nilai *precision*, *recall*, dan *F1-score* yang tinggi pada seluruh kategori keterbacaan. Validasi menggunakan *Automated Readability Index* (ARI) dengan batas kuantil sebesar 22,48 untuk kategori Mudah dan 23,97 untuk kategori Sulit menghasilkan tingkat kesesuaian sebesar 63,83%. Hasil tersebut menunjukkan bahwa sebagian besar prediksi model selaras dengan kategori keterbacaan berdasarkan ARI, meskipun masih terdapat perbedaan karena kedua pendekatan menggunakan dasar penilaian yang berbeda.

5.2 Saran

Berdasarkan batasan dan hasil evaluasi sistem yang telah diuraikan, disarankan beberapa poin pengembangan sebagai berikut :

1. Mengingat metode TF-IDF yang digunakan dalam penelitian ini terbatas pada statistik frekuensi kata, peneliti selanjutnya dapat mengeksplorasi penggunaan arsitektur *Deep Learning* seperti *Transformer* misalnya IndoBERT. Penggunaan model ini diharapkan mampu menangkap konteks semantik dan hubungan antar kata yang lebih kompleks pada abstrak jurnal Informatika, sehingga klasifikasi dapat menjadi lebih akurat.

2. Penelitian ini baru diuji menggunakan korpus abstrak Informatika. Untuk mengukur kemampuan generalisasi model, disarankan melakukan pengujian serupa pada dataset dari disiplin ilmu lain guna mengamati adaptasi algoritma terhadap variasi istilah teknis yang berbeda.

3. Pengembangan ke arah pemodelan *hybrid* dapat menjadi peluang riset berikutnya. Peneliti dapat menggabungkan fitur linguistik permukaan, seperti rasio kompleksitas kalimat, dengan matriks TF-IDF sebagai atribut masukan gabungan untuk melatih model klasifikasi yang lebih komprehensif.

4. Secara praktis, model hasil penelitian ini dapat diimplementasikan menjadi aplikasi berbasis *web*. Adanya fitur pengecekan tingkat keterbacaan secara *real-time* akan sangat membantu para akademisi melakukan revisi mandiri terhadap abstrak mereka sebelum proses pengiriman naskah.