

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan proses implementasi metode Natural Language Processing (NLP) dengan algoritma Support Vector Machine (SVM) untuk klasifikasi genre novel bahasa Indonesia.

1. Hasil klasifikasi menunjukkan bahwa model SVM dengan kernel linear dapat mengklasifikasikan data novel yang belum memiliki label romantis atau non romantis dari data genre dengan tingkat akurasi yang sangat baik yaitu 92%.
2. Tingkat akurasi yang tinggi ini membuat kesembilan data novel memiliki label romantis sebanyak enam dan label non romantis sebanyak tiga berdasarkan klasifikasi dengan data genre romantis. Model ini menunjukkan bahwa penggunaan *tokenizing*, *case folding*, *stemming*, *stop word*, *handling number* dan TF-IDF sebagai fitur dapat meningkatkan kinerja model dalam klasifikasi teks. Hal ini menunjukkan bahwa teknik yang digunakan efektif dalam mengidentifikasi dan mengklasifikasikan teks berdasarkan genre yang telah ditentukan. Proses ini merupakan bagian dari teknik data mining, khususnya dalam aspek klasifikasi teks, yang membantu mengungkap informasi tersembunyi dari data teks yang tidak terstruktur.

5.2 Keterbatasan

Berdasarkan penelitian yang telah dilakukan memiliki keterbatasan yang harus dipahami dalam proses penelitian ini. Berikut batasan-batasan yang perlu dipahami :

1. Dataset yang digunakan terbatas dikarenakan membutuhkan spesifikasi device yang tinggi, sehingga dataset dari novel hanya menggunakan sebanyak sembilan novel saja dari yang seharusnya 92 berdasarkan teori

- Slovin. Dan dataset genre yang digunakan hanya dari genre romantis saja, sehingga klasifikasi memiliki keterbatasan variasi dalam menentukan genre.
2. Proses klasifikasi yang dilakukan hanya untuk menentukan data novel memiliki peluang positif atau negatif untuk membuat novel tersebut bergenre romantis.
 3. Proses yang dilakukan juga hanya sampai pemodelan saja belum sampai diaplikasikan kedalam software ataupun hardware.

5.3 Saran

Apabila ada peneliti yang ingin menjadikan penelitian ini sebagai referensi atau pun melanjutkan penelitian ini untuk kepentingan akademik maupun individu berikut saran yang dapat dipertimbangkan :

1. Hal yang paling penting dan sangat utama untuk melakukan penelitian sejenis pre-processing data, memerlukan device dengan spesifikasi yang tinggi, apabila menggunakan dataset dalam jumlah besar.
2. Menggunakan dataset yang bersifat open-source dan aksesnya mudah, dikarenakan proses pengambilan data dari wattpad bisa dibilang tergolong sulit.
3. Implementasi teknik data mining yang lebih lanjut, seperti clustering atau association rule mining, dapat membantu dalam eksplorasi lebih dalam terhadap pola-pola tersembunyi dalam teks.