

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan dengan membandingkan dua klasifikasi data mining yaitu algoritma *Naive bayes* dan algoritma *Support Vector Machine*, menghasilkan Nilai Confusion Matriks yang dihasilkan dengan menggunakan algoritma *Naive Bayes* sebesar 91,34% nilai *accuracy*, 91,39% nilai *precision*, dan *recall* sebesar 91,39%. Sedangkan untuk algoritma *Support Vector Machine* menghasilkan nilai 85,04% *accuracy*, 85,29% *precision*, dan 84,94% untuk *recall*. Dari hasil penelitian ini dapat disimpulkan bahwa algoritma *Naive Bayes* lebih baik dari algoritma *Support Vector Machine*.

Algoritma *Naive Bayes* lebih baik dari Algoritma *Support Vector Machine* karena:

1. Kesederhanaan Model dan Kecepatan:

Naive Bayes bekerja dengan prinsip probabilitas sederhana, yang membuatnya cepat untuk pelatihan dan inferensi, terutama pada dataset kecil atau menengah. SVM, sebaliknya, membutuhkan lebih banyak waktu untuk melatih model karena memaksimalkan margin antara kelas.

2. Efektivitas pada Data Berdimensi Rendah:

NB sangat baik pada data dengan sedikit fitur. Ini karena NB tidak membutuhkan eksplorasi hubungan kompleks antar fitur, sedangkan SVM membutuhkan tuning parameter yang lebih cermat untuk menangani dimensi yang rendah.

3. Ketangguhan Terhadap Dataset Kecil:

NB tetap memberikan performa baik bahkan pada dataset kecil karena sifat probabilitiknya. SVM memerlukan lebih banyak data untuk mengoptimalkan hyperplane yang memisahkan kelas.

4. Efisiensi Dalam Kasus Multiklasifikasi:

Dalam kasus klasifikasi multi-label, NB lebih efisien karena perhitungan independensi antar fitur mempermudah proses dibandingkan SVM yang lebih cocok untuk binary classification .

5.2 Saran

1. Lebih memperbanyak data film, tidak hanya film Indonesia saja data film dari luar negeri juga bisa dimasukan.
2. Melakukan perbandingan algoritma lain, seperti algoritma *Decession Tree*, *Random Fores* dan yang lainnya.

