

BAB 6 KESIMPULAN DAN SARAN

Bab ini menjelaskan tentang kesimpulan, kendala, dan saran dari eksperimen yang dilakukan untuk klasifikasi dokumen teks menggunakan metode Naïve Bayes dan ontologi.

6.1 Kesimpulan

Berdasarkan hasil eksperimen yang diperoleh, dapat ditarik beberapa kesimpulan, yaitu pertama, jumlah kategori yang ada mempengaruhi kinerja klasifikasi dokumen menggunakan metode Naïve Bayes dan ontologi. Secara umum, penambahan jumlah kategori dapat menurunkan tingkat akurasi klasifikasi dokumen. Selain itu, tingkat kemiripan diantara kategori juga dapat mempengaruhi tingkat akurasi klasifikasi dokumen. Jika tingkat kemiripan diantara dua kategori tinggi, maka akan sulit untuk membedakan kedua kategori tersebut sehingga tingkat akurasi klasifikasi dokumen akan turun. Kedua, penggunaan *stopwords* dan *stemming* dapat meningkatkan tingkat akurasi klasifikasi dokumen.

Klasifikasi dokumen dapat menggunakan ontologi dan memiliki nilai *f-measure* mencapai 94.02%. Meskipun nilai *f-measure* klasifikasi dokumen menggunakan ontologi tidak lebih tinggi daripada nilai akurasi klasifikasi dokumen menggunakan metode Naïve Bayes. Namun, metode klasifikasi dokumen dengan menggunakan ontologi memiliki kelebihan, yaitu tidak memerlukan proses pembelajaran atau data eksperimental sedangkan metode Naïve Bayes membutuhkan proses pembelajaran agar dapat mengklasifikasikan dokumen baru.

6.2 Kendala

Kendala yang dihadapi dalam melakukan eksperimen klasifikasi dokumen teks dalam penelitian ini antara lain: pertama, kurangnya dokumen (artikel berita berbahasa Indonesia) sehingga dalam penelitian ini tidak dapat menganalisis pengaruh jumlah

dokumen terhadap tingkat akurasi klasifikasi dokumen. Kedua, keterbatasan data yang digunakan untuk memodelkan ontologi olahraga.

6.3 Saran

Klasifikasi dokumen teks (metode Naïve Bayes dan ontologi) yang dilakukan pada penelitian ini masih memiliki kekurangan. Beberapa saran yang mungkin berguna untuk melakukan penelitian klasifikasi dokumen selanjutnya, antara lain:

1. Mencari lebih banyak dokumen, khususnya dokumen berbahasa Indonesia sehingga dapat menganalisis pengaruh jumlah dokumen terhadap tingkat akurasi klasifikasi dokumen.
2. Mengumpulkan lebih banyak data untuk memodelkan ontologi dan merancang metode klasifikasi yang lebih efisien dan akurat.
3. Pemodelan ontologi untuk sebuah konsep atau *class* dibuat seunik mungkin dari konsep atau *class* lain sehingga dapat meningkatkan nilai akurasi klasifikasi dokumen.