

BAB V

PENUTUP

5.1. Kesimpulan

Pada akhirnya, hasil pengklasifikasian jurnal tugas akhir berdasarkan judul maupun abstrak sangat berbeda jauh dari segi keakurasiannya. Dengan menggunakan *ten fold validation*, rata-rata akurasi pengklasifikasian jurnal lebih baik daripada berdasarkan abstrak. Angka akurasi **86,59%** dihasilkan dari pengklasifikasian menggunakan *ten fold validation* untuk judul jurnal. Sedangkan untuk abstrak, nilai rata-rata akurasi mencapai angka **49,77%**.

Berikut adalah hasil perbandingan pengklasifikasian data pengujian (belum terkategori) jurnal tugas akhir berdasarkan judul dan abstrak jurnal tugas akhir:

Uji coba ke-	Akurasi Pengklasifikasian berdasarkan Judul (dalam %)	Akurasi Pengklasifikasian berdasarkan abstrak (dalam %)
1.	93,18%	86,36%
2.	84,21%	73,45%
3.	79,73%	61,71%
4.	70,42%	56,91%
5.	61,50%	36,25%

Tabel 5.1. Perbandingan Tingkat Akurasi untuk Pengklasifikasi Jurnal Tugas Akhir Berdasarkan Judul dan Abstrak dengan berbagai Proporsi antara Data Pelatihan dan Data Pengujian (merujuk pada tabel 4.5 dan 4.6)

Hasil pengklasifikasian jurnal tugas akhir berdasarkan judul ataupun abstraksi sangat mengejutkan karena algoritma yang digunakan dalam pengklasifikasian jurnal tugas akhir ini yaitu *Naïve Bayes Classifier* (NBC) cocok dalam mengklasifikasikan dokumen teks. Biasanya, NBC ini diterapkan pada data-data numerik, tetapi NBC juga bisa diterapkan pada data-data nominal (text) seperti pada contoh kasus pengklasifikasian jurnal tugas akhir ini.

Penggunaan banyaknya kuantitas dari data pelatihan dalam proses pengklasifikasian, membawa pengaruh yang sangat nyata terhadap akurasi yang dihasilkan melalui SISKLATA. Pengurangan kuantitas data pelatihan akan berdampak pada akurasi pengklasifikasian yang dihasilkan baik berdasarkan judul maupun abstraknya. Jadi, semakin sedikit (kuantitas) data pelatihan yang digunakan untuk proses pengklasifikasian, semakin rendah pula hasil pengklasifikasian yang dihasilkan baik berdasarkan judul maupun berdasarkan abstrak dari suatu jurnal.

Berdasarkan tabel 5.1 diketahui bahwa pengklasifikasian lebih efektif bila dilakukan berdasarkan judul jurnal. Hal ini disebabkan karena di dalam judul, kata-kata atau frase yang merangkai suatu judul terdiri dari beberapa kata yang konsisten. Mengingat bahwa judul jurnal adalah representasi dari keseluruhan isi dari sebuah jurnal. Sedangkan dalam abstrak, kata-kata atau frase yang merangkai suatu abstrak terdiri dari beberapa kata yang tidak konsisten, contohnya kata-kata seperti “*citra*”, “*website*” terdapat dalam kategori “*basis data*”. Sedangkan kata-kata yang relevan dalam kategori “*basis data*” adalah seperti “*database*”, “*mining*”. Satu hal yang

mempengaruhi pengklasifikasian berdasarkan abstrak adalah kemunculan kata-kata yang tidak konsisten itu lebih dari satu kali. Hal ini juga yang mempengaruhi perhitungan NBC.

5.2. Saran

Penerapan sistem pengklasifikasian jurnal tugas akhir (SISKLATA) dalam sebuah universitas atau perguruan tinggi hendaknya disesuaikan dengan kapasitas IT yang ada pada universitas atau perguruan tinggi tersebut. Sebaiknya SISKLATA ini berhubungan dengan sistem yang lain, seperti sistem *digital library* (perpustakaan) yang menyimpan berbagai jenis jurnal tugas akhir khususnya jurnal yang berhubungan dengan informatika atau ilmu komputer. Adapun metode yang telah penulis paparkan dalam skripsi ini merupakan salah satu cara yang dapat digunakan dalam memecahkan permasalahan pengklasifikasian jurnal tugas akhir. Tidak menutup kemungkinan bahwa adanya metode lain yang lebih baik atau bahkan perbaikan terhadap metode yang sudah ada. Semoga apa yang sudah penulis paparkan dalam skripsi ini dapat menginspirasi pihak lainnya dalam meneliti permasalahan di sekitar, khususnya lebih jauh lagi tentang pengklasifikasian jurnal tugas akhir.