

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, bahwa sistem ekstraksi informasi menggunakan *Hidden Markov Model (HMM)* telah berhasil diimplementasikan dengan perolehan akurasi token-kelas sebesar 72.70%. Perolehan akurasi token-kelas disebabkan oleh penggunaan fitur pembobotan yang tidak signifikan, sehingga algoritma *HMM* tidak dapat melakukan klasifikasi secara maksimal.

5.2 Saran

Pada penelitian ini masih memiliki beberapa kekurangan, sehingga nilai akurasi pada konsep token-kelas tidak dapat diperoleh dengan maksimal. Dengan demikian, beberapa saran akan dipaparkan untuk pengembangan lebih lanjut mengenai sistem ekstraksi informasi menggunakan *machine learning*, diantaranya:

1. Menggunakan data masukan dengan format *.docx* atau *.html* agar dapat menggunakan beberapa fitur tambahan seperti mendeteksi *bold*, *italic*, *underline* dan beberapa *font style* lainnya.
2. Diperlukan suatu fitur tambahan yang signifikan untuk pembobotan kelas Fakultas, Program Studi dan Universitas. Hal tersebut dilakukan agar ketika proses *testing* dilakukan, algoritma *HMM* dapat menentukan kelas yang tepat untuk fakultas, Program Studi dan Universitas.
3. Perlu dibuktikan algoritma *HMM* dengan perbandingan parameter dalam melakukan ekstraksi informasi pada dokumen karya tulis ilmiah.
4. Dibutuhkan pengecekan dan pengkoreksian kesalahan ejaan atau typo pada hasil konversi *file PDF*.