

BAB 6 PENGUJIAN

Bab ini menjelaskan tentang pengujian yang dilakukan terhadap sistem yang telah dibuat. Jenis pengujian yang akan dilakukan pada sistem ini yaitu pengujian akurasi sistem.

6.1 Pengujian Akurasi Sistem

Pengujian akurasi sistem dilakukan pada sistem ini untuk mengetahui tingkat keberhasilan sistem dalam melakukan klasifikasi kepribadian untuk mengetahui karakter calon karyawan perusahaan menggunakan Twitter. Pengujian ini dilakukan pada 160 data dengan pembagian yaitu 50% untuk data latih 80 data dan 50% untuk data uji atau sekitar 80 data (Ariadi & Fithriasari, 2015). Pengujian ini akan dijelaskan pada skenario pengujian dan analisis hasil pengujian.

6.1.1 Skenario Pengujian Akurasi Sistem

Pada pengujian akurasi sistem pada sistem klasifikasi kepribadian dengan menggunakan data Twitter dan algoritme KNN ini akan dilakukan perhitungan jumlah data yang mendapatkan hasil klasifikasi yang benar atau sama dengan kelas kepribadian yang sebenarnya dimiliki data uji tersebut dan perhitungan jumlah data yang mendapatkan hasil klasifikasi yang salah atau tidak sama dengan kelas kepribadian yang sebenarnya dimiliki data uji tersebut serta perhitungan jumlah seluruh data uji. Nilai k yang menjadi parameter dari algoritme KNN diisi dengan 3 karena angka tersebut merupakan nilai k yang paling optimal (Nurjanah, 2017). Kemudian dilakukan perbandingan antara jumlah data yang mendapatkan hasil klasifikasi yang benar dan jumlah seluruh data uji kemudian dikali 100% hingga didapatkan nilai akurasi sistem dalam bentuk persentase. Nilai akurasi sistem ini menunjukkan tingkat keakuratan sistem dalam melakukan klasifikasi kepribadian untuk menentukan karakter calon karyawan untuk calon perusahaan. Hasil pengujian akurasi sistem dapat dilihat pada Tabel 6.1.

Tabel 6.1 Hasil Pengujian Akurasi Sistem

	Hasil Prediksi	Hasil Sebenarnya
Jumlah Benar		53
Jumlah Salah		27
Jumlah Akurasi		66%

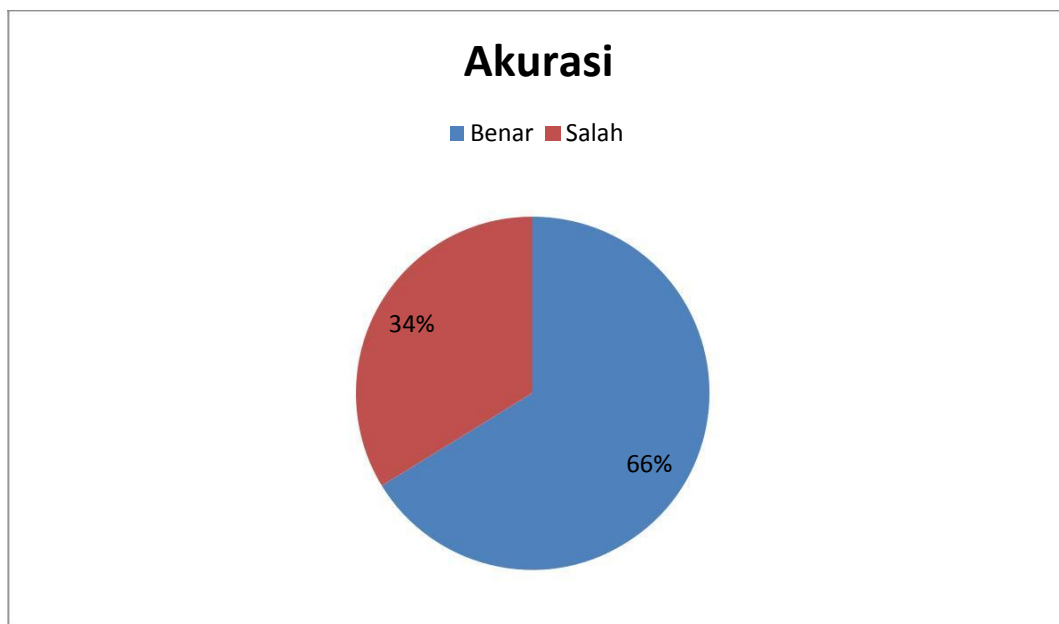
6.1.2 Analisis Hasil Pengujian Akurasi Sistem

Dari hasil pengujian akurasi sistem berdasarkan Tabel 6.1 menunjukkan bahwa tingkat akurasi pada sistem klasifikasi ini yaitu 66% jumlah data yang mengalami klasifikasi yang benar 53 data, sedangkan jumlah data yang mengalami klasifikasi yang salah terdapat 27 data. Hal ini karena data-data tweet calon karyawan yang digunakan sangat bervariasi sehingga untuk hasil

perhitungan nilai cosine similarity atau tingkat kemiripan antara data uji yang memiliki kelas yang sebenarnya sama dengan data latih tampak kurang, sedangkan tingkat kemiripan data uji dan data latih malah meningkat pada data-data yang memiliki perbedaan antara kelas sebenarnya yang dimiliki data uji dengan kelas pada data latih. Tingkat kemiripan data yang meningkat pada data uji yang memiliki perbedaan kelas yang sebenarnya dengan kelas dari

karena kata-kata yang mengandung dalam data-data tersebut terdapat banyak kesamaan walaupun sebenarnya memiliki makna yang berbeda. Sedangkan pada tingkat kemiripan yang rendah antara data uji yang memiliki kelas yang sebenarnya sama dengan kelas pada data karena terdapat banyak kata-kata yang tidak memiliki kesamaan antara keduanya namun sebenarnya memiliki makna yang berbeda atau data-data tersebut mewakili ciri-ciri kepribadian yang sama. Selain itu data yang digunakan adalah data Twitter yang memiliki isi data yang sedikit dan terbatas hingga 140 karakter sehingga hanya sedikit ciri-ciri suatu kepribadian yang sedikit dan sulit apabila dibandingkan kemiripannya dengan data lain yang memiliki kelas kepribadian yang sama.

Walaupun seperti itu, sistem masih bisa memberikan hasil klasifikasi dengan benar dengan tingkat akurasi sekitar 66% sehingga algoritme klasifikasi KNN telah diterapkan dengan baik untuk sistem ini.



Gambar 6.1 Hasil Pengujian Akurasi Sistem