

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan dari analisis dan pengujian yang telah dilaksanakan menggunakan dataset yang telah diambil dan menggunakan algoritma *Random Forest*, *Decision Tree*, *KNN* mendapatkan kesimpulan bahwa Perbandingan algoritma *Random Forest*, *Decision Tree*, *KNN* telah mendapatkan hasil bahwa algoritma *Random Forest* merupakan algoritma yang terbaik pada skenario testing 60:40 dengan menggunakan dataset *review* aplikasi IDN. Dengan hasil akurasi 82%, *Precision* 82%, *Recall* 89%, *F1_Score* 86%. Algoritma *Decision Tree* mengalami penurunan performa akurasi pada skenario 80:20 mendapatkan nilai akurasi 80% kemudian pada skenario 70:30 mendapatkan hasil 78%, hal tersebut bisa terjadi dikarenakan kemampuan dalam menangani data yang lebih besar. Algoritma *KNN* mendapatkan hasil akurasi yang buruk dengan nilai akurasi tertinggi 65% pada nilai *K* sebesar 5. Pengujian dengan nilai *K* yang lebih tinggi mengurangi hasil akurasi yang didapat pada algoritma *KNN*.

Pada penelitian yang telah dilakukan menggunakan *framework* CRISP-DM. dapat disimpulkan menjadi beberapa tahapan. Pada tahapan *business understanding* melakukan pemahaman terkait penelitian seperti pemilihan algoritma *Random Forest*, *Decision Tree*, *KNN* serta pemilihan aplikasi IDN sebagai aplikasi berita paling populer pada *Google Playstore*. Tahapan kedua adalah *data understanding* dimana pada penelitian ini melakukan proses *Scraping* data ulasan aplikasi IDN pada *Google Playstore*. Jumlah data yang diperoleh sebesar 995 data. Tahapan selanjutnya yaitu data *preparation* dengan melakukan pembersihan data. tahapan selanjutnya adalah tahapan *modeling* untuk melakukan pemodelan klasifikasi dengan tiga skenario yang berbeda. Pada tahap *evaluation* melakukan proses perbandingan dari seluruh pemodelan yang telah dilakukan. Tahapan *deployment* adalah melakukan *deploy* model terbaik kedalam bentuk *website*.

5.2 Saran

Berdasarkan penelitian yang telah dilaksanakan terdapat saran yang dapat di berikan pada penelitian selanjutnya yaitu:

1. Menggunakan teknik scraping selain *Google Play Scrap* karena teknik ini telah membatasi jumlah data yang bisa di ambil. Teknik scraping yang bisa digunakan seperti *Beautiful soup*.
2. Pengujian Menggunakan algoritma SVM, Naïve bayes, C.45 untuk dapat dibandingkan.
3. Menggunakan teknik pembobotan kata lain seperti *word2vec*.
4. Menggunakan *feature selection* lain seperti *Information Gain*.
5. Penambahan fitur *website* seperti mengupload *file* untuk deteksi sentimen.
6. Menggunakan *library textblob* untuk pelabelan sentimen positif dan negatif.

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA