

## BAB V PENUTUP

### 5.1 Kesimpulan

Berdasarkan tahapan analisis yang dilakukan dan hasil yang didapatkan, maka diperoleh beberapa kesimpulan berikut :

1. Penelitian menggunakan data ulasan aplikasi sebanyak 25.000 data awal setelah didrop data null yang berjumlah 11. Maka, data yang digunakan berjumlah 24989 dengan menggunakan atribut content dan score.
2. Metode validasi algoritma SVM Rbf dan Random Forest yang merupakan pokok dari penelitian ini. Pada metode SVM Rbf dengan parameter menggunakan *SVC* (*kernel = 'rbf'*) sedangkan pada Random forest menggunakan parameter *RandomForestClassifier* yaitu *n\_estimators = 100*. Melalui permodelan split data yang berbeda yaitu dengan split data menggunakan *K-fold Cross Validation* dan *Random state 80 : 20* untuk data train dan test.
3. Didapatkan hasil dari masing-masing pemodelan split data yang dilakukan dan parameter tersebut. Pada SVM Rbf dengan pemodelan split data K-fold mendapatkan akurasi terbaik dari fold 74% dan pada split data rasio 80 : 20 mendapatkan 94% . Sedangkan hasil dari Random Forest menggunakan K-fold 75% dan split data rasio 80:20 mendapatkan 95%. Oleh itu dari kedua metode yang telah dilalui bahwa hasil dari Random Fores dari pemodelan split data tersebut mendapatkan hasil lebih baik.
4. Dalam perbandingan metode SVM Rbf, *Random Forest*, *SVM Linear*, *SVM Polynomial*, *Adaboost* dan *Bagging* dalam data ulasan Dana, pada pemodelan split data K-fold yang memberikan performa yang paling terbaik Random Forest dengan rata-rata akurasi fold akurasi 75%, Sedangkan untuk split data rasio 80:20 mendapatkan bahwa SVM Polynomial memberikan performa yang paling baik dengan akurasi 98% ,*precision* 98% ,*recall* 98% dan *F1-Score* 98% . Untuk metode yang mempunyai performa yang tidak

bergitu baik dari kedua pemodelan split data yaitu metode Adaboost dengan akurasi pada K-fold hanya 70% dan pada split data biasa rasio 80:20 sebesar 68%.

## 5.2 Saran

Berdasarkan hasil dan kesimpulan yang telah dicapai dalam penelitian ini, peneliti memberikan saran untuk meningkatkan kualitas penelitian sebagai berikut:

1. Menggunakan metode yang lebih baik pada pelabelan data untuk mendapatkan data yang lebih kompleks untuk meningkatkan akurasi dari metode yang digunakan.
2. Melakukan percobaan dengan metode lainnya selain dari metode yang digunakan dan sebagai perbandingan dari penelitian ini untuk memperoleh akurasi yang baik pada metode split data dengan *K-fold cross validation*.
3. Menambahkan metode *preprocessing* lainnya dapat berupa *word normalization* untuk mendapatkan representasi kata yang lebih baik dalam analisis sentiment.

