

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan penelitian yang dilakukan dengan menggunakan algoritma *machine learning* terhadap *sentiment analysis cryptocurrency* melalui pendekatan *feature selection* berbasis *swarm intelligence*, dapat menarik beberapa kesimpulan penting. Tujuan dari penelitian adalah untuk mengetahui proses *scrapping data* dan pembersihan data yang dilakukan pada *sentiment analysis cryptocurrency*. Kemudian, mengetahui hasil perbandingan algoritma *machine learning* (SVM, KNN, dan Naïve Bayes) untuk *sentiment analysis* opini *cryptocurrency*. Kemudian, mengetahui algoritma *machine learning* terbaik dari SVM, KNN, dan Naïve Bayes menggunakan algoritma optimasi berbasis *swarm intelligence* (PSO, ACO, dan CSO) sebagai *feature selection*. Selain itu, penelitian ini juga berfokus pada mengukur kinerja hasil perbandingan dari algoritma SVM, KNN, dan Naïve Bayes terbaik sebelum dan sesudah penerapan *feature selection* berbasis *swarm intelligence* (PSO, ACO, dan CSO) berdasarkan akurasinya dan waktu pemrosesan.

Proses pengumpulan data dilakukan dengan menggunakan teknik *scrapping data*. Pada saat menerapkan *scrapping data*, diterapkan fungsi “*tweet-harvest*” untuk dapat mengambil data dari media sosial X dan diatur dalam bahasa Inggris. Hasil dari *scrapping data* tersebut menghasilkan total 9884 data yang diekstrak kedalam file CSV. Kemudian, untuk proses pembersihan data dilakukan dengan menerapkan *remove urls*, *remove hashtags*, *remove special characters and numbers*, *excluding punctuation*, *convert to lower case*, *remove stop words*, *lemmatization of words*, dan *remove words with length < 3*. Kemudian, semua kata akan disatukan kembali dengan fungsi “*join(words)*”. Hasil dari pembersihan data menghasilkan 9843 data yang terdiri dari hasil 3 kolom, yakni “*full_text*”, “*clean_text*”, dan “*vader_sentiment*”.

Berdasarkan hasil dari pemodelan dan evaluasi penelitian, didapatkan ketiga hasil akurasi dari algoritma *machine learning*, yaitu SVM sebesar 87,04%, KNN sebesar 72,27%, dan Naïve Bayes sebesar 79,18%. Dari ketiga algoritma tersebut SVM memiliki akurasi tertinggi. Selain itu, hasil *classification report* yang didapatkan juga memiliki akurasi yang cukup baik, yakni 87% untuk *precision*, *recall*, dan *f1-score*.

Selanjutnya, didapatkan juga hasil akurasi dan *classification report* terbaik untuk algoritma *machine learning* dengan menggunakan optimasi *feature selection* berbasis *swarm intelligence*, yaitu adalah ACO-Naïve Bayes dengan nilai akurasi sebesar 82,43%. Hasil *classification report* juga menunjukkan *output* yang baik, dimana skor untuk nilai *precision*, *recall*, dan *f1-score* adalah 83%, 82%, dan 82%. Hasil tersebut membuktikan bahwa algoritma *feature selection* berbasis *swarm intelligence* terbukti mampu dalam meningkatkan akurasi model *machine learning* sederhana terkait *sentiment analysis*.

5.2 Saran

Penelitian yang dilakukan telah memberikan kontribusi dan hasil yang baik pada optimasi kinerja model *machine learning* dalam *sentiment analysis cryptocurrency* menggunakan pendekatan *feature selection* berbasis *swarm intelligence*. Namun, terdapat beberapa saran yang dapat diberikan kepada peneliti selanjutnya untuk pengembangan penelitian lebih lanjut. Berikut merupakan beberapa saran yang dapat diberikan untuk penelitian selanjutnya.

1. Menggunakan data yang lebih terkini dan rentang waktu yang panjang. *Cryptocurrency* merupakan aset digital yang nilai jual mengalami pergerakan setiap saat dan tergantung oleh momentum. Oleh karena itu, penting untuk menggunakan data terkini dan rentang waktu yang panjang agar hasil prediksi dapat lebih akurat dan tervalidasi setiap saat.
2. Melakukan *deployment* terhadap hasil *sentiment analysis* yang telah dibuat, seperti *recommendation system*. Hal tersebut bertujuan agar model

dapat memberikan prediksi atau wawasan berdasarkan data nyata atau *real time*.

3. Menerapkan beberapa metode untuk mengatasi kendala *overfitting* pada model. Beberapa metode yang dapat diterapkan adalah implementasi *early stopping* (melibatkan *patience*, *threshold*, dan *metric to monitor*), pengurangan dimensi (*Principal Component Analysis*), dan optimasi *hyperparameter* berbasis *swarm intelligence*.



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA