



ANALISIS SENTIMEN MASYARAKAT TERHADAP KEBIJAKAN PENERAPAN PPKM DI MEDIA SOSIAL TWITTER DENGAN MENGUNAKAN METODE XGBOOST

Dosen Pembimbing: Ramaditia Dwiyanaputra, S.T., M.Eng. & Arik Aranta, S.Kom., M.Kom.
I Putu Angga Purnama Widiarta (FID018024)



ABSTRAK

Covid-19 adalah virus yang menyebabkan infeksi saluran pernafasan dan merebak di Indonesia, PPKM diterapkan pemerintah demi menekan angka persebaran. Pro & kontra muncul akibat dampak yang ditimbulkan kebijakan tersebut. Penelitian ini bertujuan untuk mengimplementasikan algoritma XGBoost pada proses klasifikasi sentimen guna menilai opini masyarakat Indonesia terhadap kebijakan tersebut. Sentimen analisis menargetkan opini masyarakat di twitter. Dataset yang digunakan sejumlah 1958 tweets positif & 3980 tweets negatif. Pada preprocessing, dilakukan case folding, stopwords removal, tokenizing, dan stemming. Pemberian bobot pada term menggunakan metode TF-RF untuk mengubah setiap term menjadi angka. Pada tahapan akhir, klasifikasi dilakukan dengan mengimplementasikan metode XGBoost dengan score hyperparameter yang optimal. K-fold cross validation digunakan untuk mengevaluasi kinerja model. Hasil evaluasi performa terbaik diperoleh model dengan nilai hyperparameter dengan $n_estimator$ sebesar 1000, $learning_rate$ sebesar 0.1, max_depth sebesar 6, $subsample$ sebesar 1, $gamma$ sebesar 0 dan memanfaatkan proses stemming dalam preprocessing dengan nilai akurasi sebesar 85.27%, presisi sebesar 86.07%, dan recall sebesar 85.23%.

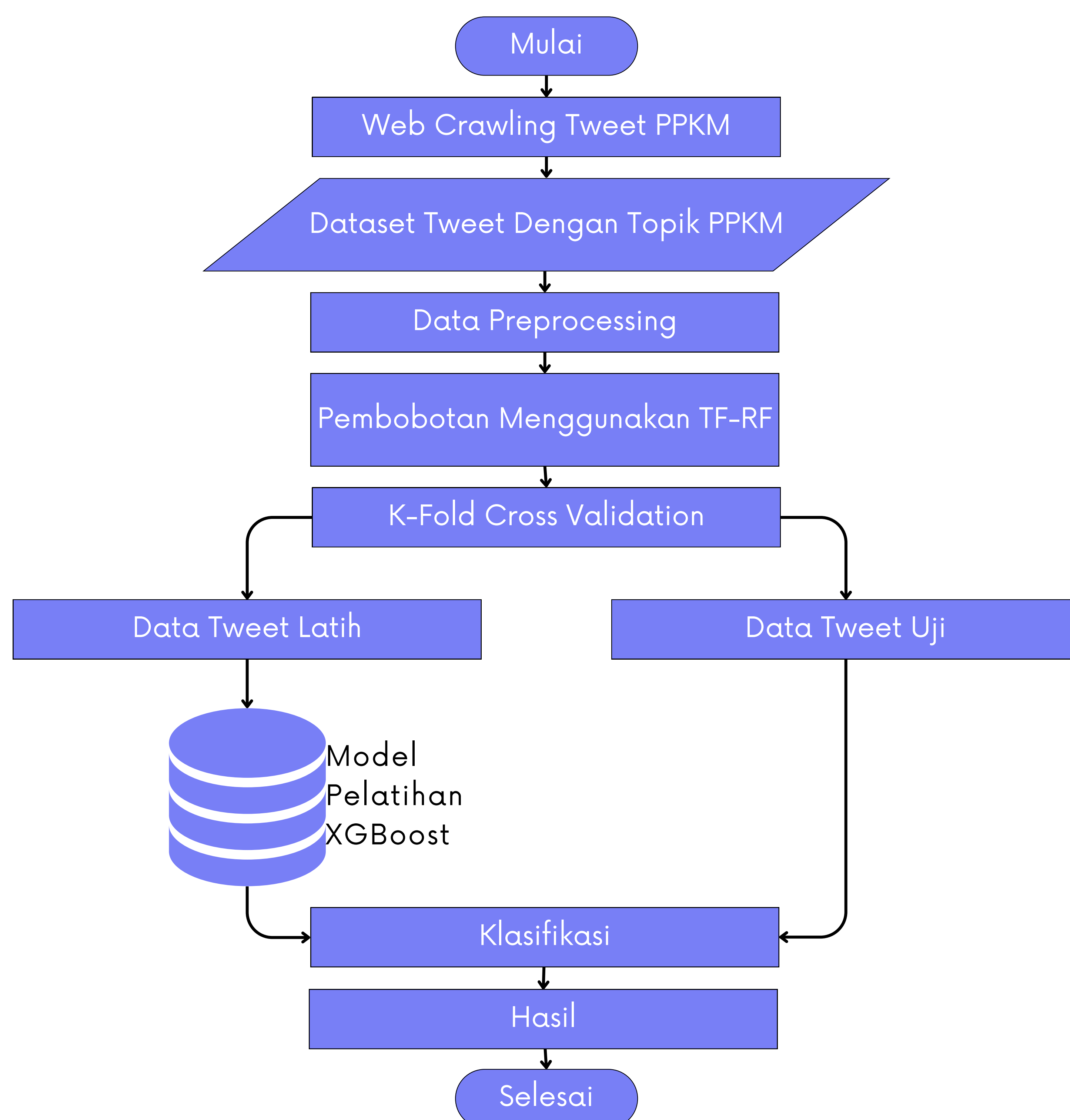
TUJUAN

1. Mengimplementasikan metode eXtreme Gradient Boosting (XGBoost) dalam melakukan analisis sentimen masyarakat terhadap penerapan kebijakan PPKM di media sosial Twitter.
2. Mengetahui performa pengujian pada analisis sentimen masyarakat terhadap penerapan kebijakan PPKM di media sosial Twitter.
3. Mengetahui sentimen mayoritas masyarakat Indonesia terhadap penerapan kebijakan PPKM di media sosial Twitter.
4. Mengetahui jumlah klusterisasi kata yang akan dicari yaitu positif dan negatif terhadap penerapan kebijakan PPKM di media sosial Twitter.

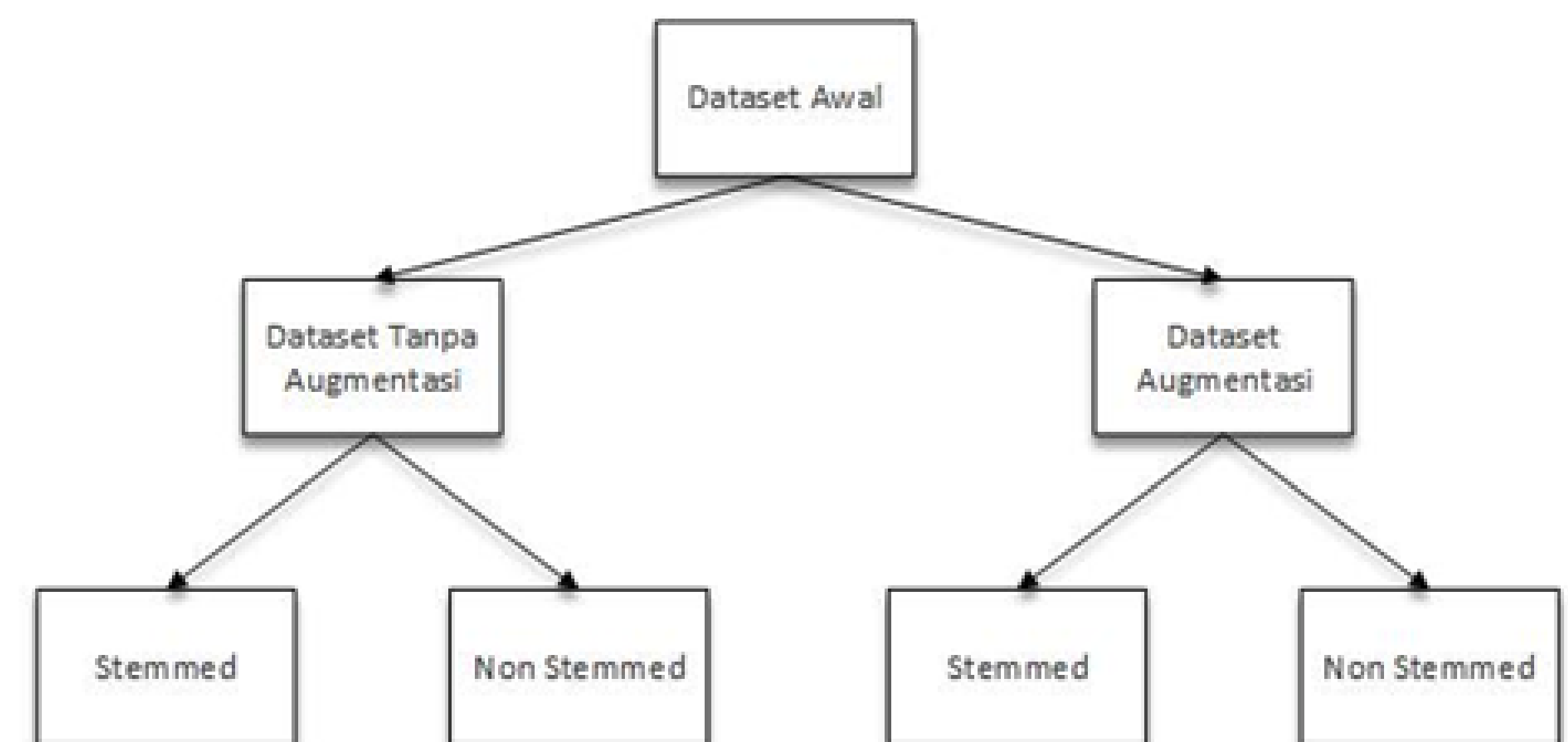
PENYELESAIAN

Mengimplementasikan algoritma klasifikasi XGBoost, dan metode Term Frequency - Relevance Frequency untuk menentukan bobot suatu term pada teks. XGBoost menghemat waktu, mengoptimalkan sumber daya memori dan dapat diterapkan secara paralel selama proses implementasi untuk mengelola sentimen.

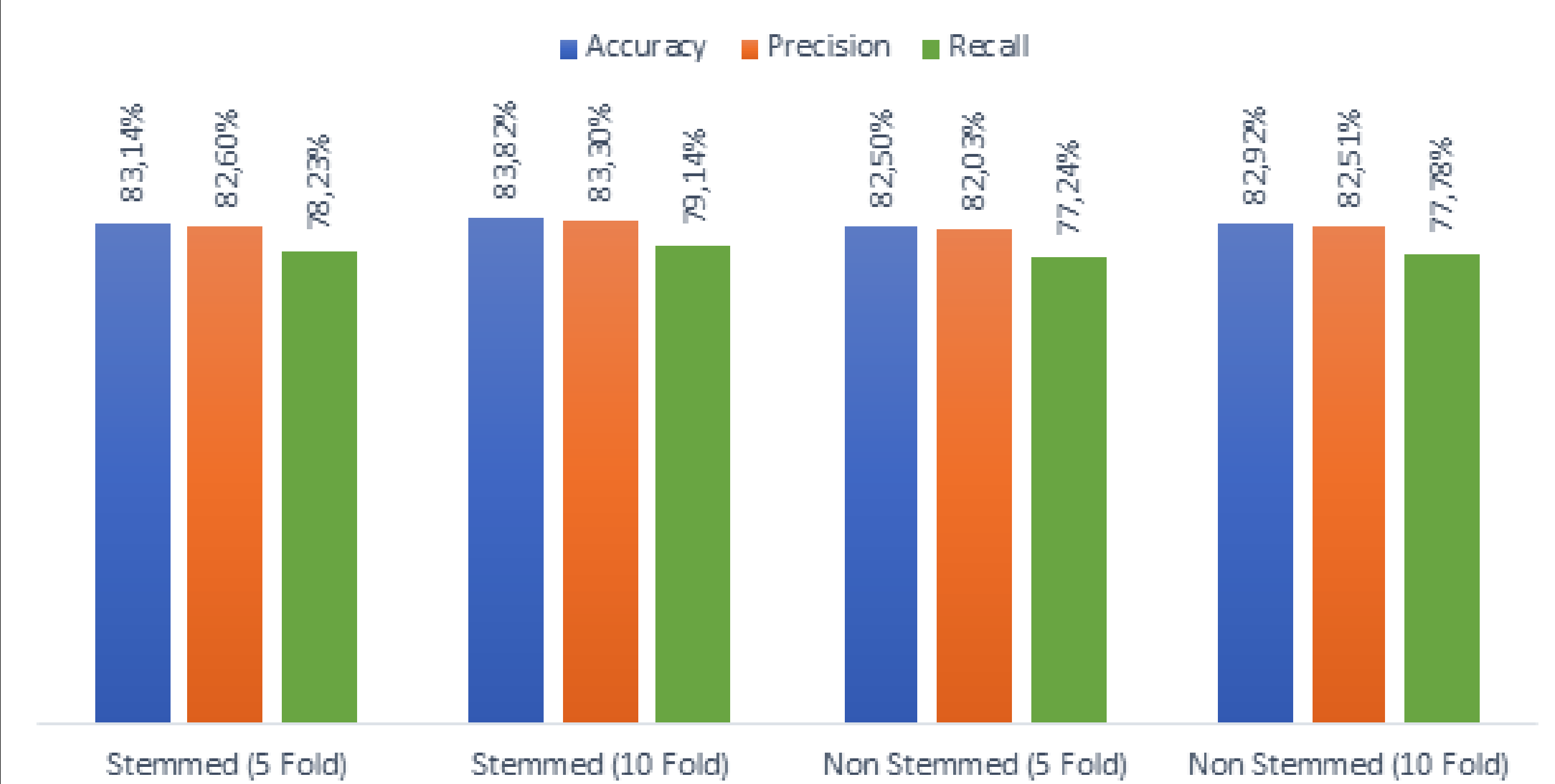
RANCANGAN KLASIFIKASI SISTEM



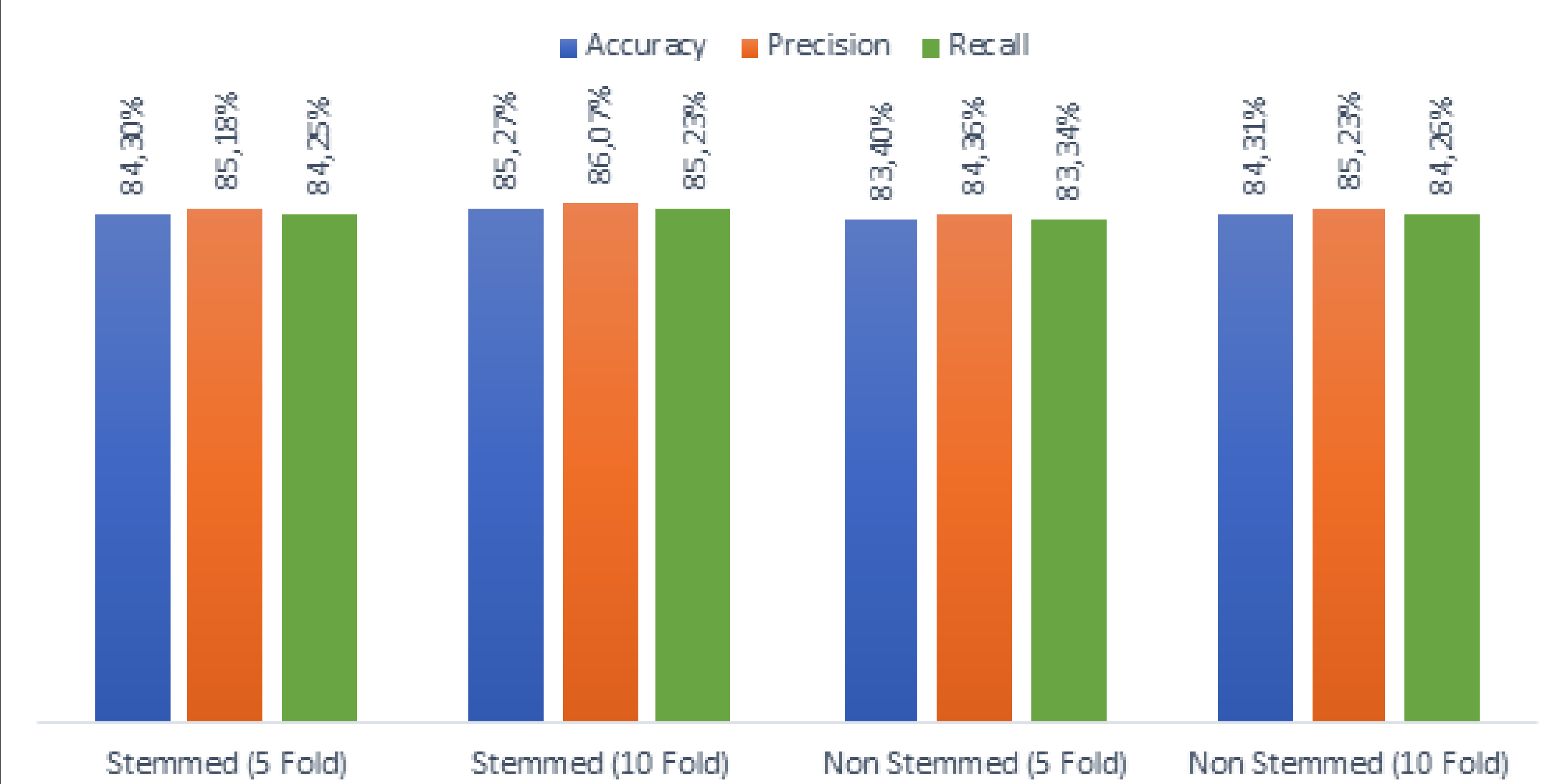
PENGUJIAN & HASIL



Dataset Tanpa Augmentasi



Dataset Dengan Augmentasi



KESIMPULAN

- Implementasi metode XGBoost dalam melakukan klasifikasi sentimen pada tweets dengan menggunakan dataset yang telah diaugmentasi dan telah melewati proses stemming, menghasilkan nilai akurasi sebesar 85.27%, presisi sebesar 86.07%, dan recall sebesar 85.23% setelah diuji dengan menggunakan teknik cross validation dengan nilai k-fold yaitu sama dengan 10.
- Dataset yang mengalami augmentasi cenderung menghasilkan nilai akurasi, presisi, dan recall yang lebih tinggi dibandingkan dengan dataset yang tidak mengalami augmentasi, hal ini dikarenakan augmentasi dataset merujuk pada teknik penambahan variasi atau modifikasi pada dataset sehingga meningkatkan jumlah sampel yang tersedia yang menyebabkan model memiliki lebih banyak variasi dan informasi yang dapat dipelajari.
- Skenario pengujian yang menggunakan tahapan stemming cenderung menghasilkan nilai accuracy, precision, dan recall yang lebih tinggi dibandingkan dengan pengujian tanpa tahapan stemming. Hal ini terjadi karena imbuhan berupa prefiks maupun sufiks pada kata yang memiliki kata dasar yang sama malah menambah variasi fitur saat melakukan pembobotan, sehingga memberikan nilai bobot yang berbeda - beda.

