

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model prediksi penjualan untuk Big Mart dengan menggunakan teknik pembelajaran mesin. Data penjualan yang digunakan mencakup 1.559 produk di 10 toko pada tahun 2013 yang diperoleh dari Analytics Vidhya. Permasalahan yang dihadapi adalah cara optimasi *hyperparameter* dalam meningkatkan akurasi prediksi. Sebagian besar penelitian sebelumnya menggunakan metode konvensional atau optimasi *hyperparameter* yang sederhana, sementara penelitian ini mencoba mengisi kekosongan tersebut dengan membandingkan dua metode optimasi *hyperparameter*, *Grid Search Optimization* (GSO) dan *Bayesian Optimization* (BO), dalam konteks model XGBoost untuk prediksi penjualan. Tujuan penelitian ini adalah menilai performa model XGBoost dengan pengaturan *hyperparameter* bawaan, menganalisis peningkatan akurasi dari optimasi *hyperparameter* menggunakan GSO, serta membandingkan efisiensi waktu dan akurasi antara GSO dan BO. Data penjualan yang digunakan berasal dari *dataset* penjualan Big Mart yang mencakup berbagai atribut produk dan toko, dengan evaluasi performa model dilakukan menggunakan metrik *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE). XGBoost *default* menggunakan *max_depth* = 6, *learning_rate* = 0,3, *n_estimators* = 100, dan *subsample* = 1 menghasilkan nilai RMSE sebesar 1213 dan MAE sebesar 850 dan digunakan sebagai *baseline*. GSO pada model XGBoost menghasilkan *hyperparameter* terbaik dengan *max_depth* = 3, *learning_rate* = 0,01, *n_estimators* = 50, dan *subsample* = 0,9 memberikan nilai RMSE sebesar 1070 dan MAE sebesar 754. Sedangkan, BO yang difokuskan pada minimisasi RMSE menghasilkan *hyperparameter* *max_depth* = 3, *learning_rate* = 0,0126, *n_estimators* = 357, dan *subsample* = 0,6 menghasilkan nilai RMSE terbaik sebesar 1069 dan MAE sebesar 755. BO yang difokuskan pada minimisasi MAE menghasilkan *hyperparameter* *max_depth* = 3, *learning_rate* = 0,0668, *n_estimators* = 100, dan *subsample* = 0,9 mencapai nilai RMSE sebesar 1071 dan MAE sebesar 753. Hasil penelitian menunjukkan bahwa BO menghasilkan performa yang sedikit lebih baik dibandingkan dengan GSO dalam hal akurasi prediksi, tetapi BO jauh lebih efisien dalam hal total waktu komputasi sebesar 530 detik dibanding GSO sebesar 1481 detik. Penelitian ini memberikan kontribusi dalam pemilihan metode optimasi *hyperparameter* yang lebih efisien dan efektif untuk model prediktif di industri ritel.

Kata kunci: XGBoost, optimasi *hyperparameter*, *Grid Search Optimization*, *Bayesian Optimization*, prediksi penjualan