

BAB I

PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang akan digunakan dalam dokumen skripsi ini.

1.1 Latar Belakang

Dalam studi mengenai dampak ekonomi dan sosial dari toko pengecer atau Big Mart, Haris (2006) menyoroti perilaku pengecer besar seperti Wal-Mart mengubah lanskap ritel baik bagi pengecer lokal hingga perusahaan ritel korporat nasional, serta mengubah kebiasaan berbelanja, harapan pelanggan, dan distribusi perdagangan ritel. Menurut Pragathi & Srinivasulu (2022), Big Mart menggunakan teknik pembelajaran mesin untuk memprediksi volume penjualan masa depan, mengoptimalkan tingkat inventaris, dan menyesuaikan strategi penetapan harga untuk meningkatkan penjualan dan profitabilitas. Kecenderungan Big Mart sangat penting, karena para ilmuwan data melihatnya untuk memahami fungsionalitas dari suatu produk dan lokasi yang dapat dimodelkan secara matematika menjadi hasil prediksi penjualan.

Dampak dari prediksi penjualan menunjukkan keberhasilan dan kinerja toko ritel. Oleh karena itu, prediksi yang tidak akurat dapat menyebabkan ketidaksesuaian stok, yang menyebabkan kerugian bagi perusahaan (Thomassey, 2010). Sangat penting untuk membuat prediksi yang akurat, terutama untuk industri yang berorientasi pada konsumen seperti Big Mart, di mana industri pengecer menghadapi sejumlah masalah. Dengan kata lain, mereka dapat menyimpan produk lebih awal karena mereka tahu permintaan pelanggan di masa depan. Selain itu, permintaan dapat dipengaruhi oleh acara publik, hari libur, lokasi, dan perubahan cuaca.

Pranav Kumar & Neethi (2022) menyajikan analisis prediktif lainnya menggunakan algoritma pembelajaran mesin, menyoroti keefektifan mereka dalam memprediksi volume transaksi untuk Big Mart, mendukung perencanaan pemasaran komersial. Demikian pula, (Ranjitha & Spandana, 2021) serta (Saranya, 2023)

menekankan penggunaan algoritma pembelajaran mesin, termasuk XGBoost, regresi linear, regresi polinomial, dan regresi Ridge, untuk melakukan prediksi penjualan di Big Mart. Dalam konteks analisis prediktif, (Behera & Nain, 2020) menyajikan studi perbandingan yang menunjukkan keefektifan teknik XGBoost dalam memprediksi penjualan untuk perusahaan seperti Big Mart, dengan kinerja yang lebih baik dibandingkan dengan model-model yang ada.

Perbedaan pendekatan antara pengoptimalan model menggunakan *Grid Search Optimization* (GSO) dan *Bayesian Optimization* belum sepenuhnya dipahami dalam konteks prediksi penjualan. Beberapa penelitian telah menyoroti keunggulan *Bayesian Optimization* dalam mengatasi kompleksitas model dan meningkatkan model secara signifikan dalam berbagai domain (Syed dkk., 2020) .

Penulis akan melakukan *Hyperparameter tuning* menggunakan kedua pendekatan tersebut dan mengukur kinerja model menggunakan metrik evaluasi yang umum digunakan seperti *Root Mean Squared Error* (RMSE) (Claypool dkk., 1999) dan *Mean Absolute Error* (MAE) (Smyth & Cotter, 2000). Dalam penelitian ini, penulis bertujuan untuk membandingkan kinerja model XGBoost yang dioptimalkan menggunakan GSO dan *Bayesian Optimization* dalam konteks prediksi penjualan di perusahaan ritel.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan dapat dirumuskan suatu rumusan masalah yaitu:

1. Bagaimana performa model XGBoost dengan pengaturan *Hyperparameter default* dalam memprediksi penjualan Big Mart?
2. Seberapa besar peningkatan akurasi yang diperoleh dari penggunaan *Grid Search Optimization* dalam optimasi *Hyperparameter* XGBoost?
3. Seberapa efisien *Bayesian Optimization* dalam menemukan kombinasi *Hyperparameter* optimal untuk XGBoost dibandingkan dengan *Grid Search Optimization*?

1.3 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah untuk menilai performa model XGBoost dengan pengaturan *Hyperparameter default*, menganalisis peningkatan akurasi yang diperoleh dari optimasi *Hyperparameter* menggunakan *Grid Search Optimization*, serta membandingkan efisiensi total waktu komputasi dan akurasi model antara *Grid Search Optimization* dan *Bayesian Optimization* dalam optimasi *Hyperparameter* XGBoost.

Manfaat dari penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik mengenai pentingnya optimasi *Hyperparameter* dalam meningkatkan kinerja model prediksi, menyediakan bukti empiris mengenai efektivitas dan efisiensi dari *Grid Search Optimization* dan *Bayesian Optimization* dalam optimasi *Hyperparameter*, serta menjadi referensi bagi peneliti dan praktisi dalam memilih metode optimasi *Hyperparameter* yang tepat untuk model prediksi penjualan.

1.4 Ruang Lingkup

Ruang lingkup mempunyai tujuan memberikan batasan dari skripsi ini agar tidak menyimpang dan sesuai dengan target yang diinginkan. Ruang lingkup dalam skripsi ini adalah sebagai berikut:

1. Melakukan pemodelan prediksi penjualan pada toko ritel Big Mart menggunakan model XGBoost dengan penentuan *hyperparameter* terbaik.
2. *Hyperparameter* terbaik yang ditentukan adalah *max depth*, *learning rate*, *n estimators*, dan *subsample*.
3. Data yang digunakan adalah dataset Penjualan Big Mart 2013 dengan pembagian data menjadi 5966 data latih dan 2557 data uji dari sumber Analytics Vidhya.
4. Model XGBoost dibangun memanfaatkan *library* XGBoost.
5. Pelatihan dan pengujian model XGBoost dilakukan di lingkungan *Jupyter Notebook* pada platform web *Google Colaboratory*.
6. Pengukuran skor yang digunakan untuk mengevaluasi kemampuan model XGBoost yaitu *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE).

1.5 Sistematika Penulisan

Sistematika penulisan memberikan gambaran laporan dari skripsi ini secara urut dan jelas. Berikut adalah sistematika penulisan dari skripsi ini:

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, manfaat dan tujuan, ruang lingkup, dan sistematika penulisan pada skripsi yang berjudul *Perbandingan Metode Grid Search dan Bayesian Optimization Pada Optimasi Hyperparameter Model XGBoost untuk Prediksi Penjualan Big Mart*.

BAB II TINJAUAN PUSTAKA

Bab ini membahas teori-teori yang berhubungan dengan topik atau masalah yang dibahas pada skripsi ini. Pada bab ini akan terdiri beberapa subbab antara lain mengenai big mart, penelitian terdahulu, *flowchart*, pembagian data, XGBoost, *Cross-Validation*, *Grid Search Optimization*, *Bayesian Optimization*, dan pengujian performa model.

BAB III METODE PENELITIAN

Bab ini menyajikan informasi mengenai metodologi penelitian yang digunakan pada penelitian yang menentukan *Hyperparameter* terbaik model XGBoost untuk segmentasi prediksi penjualan toko ritel Big Mart. Proses yang dilakukan dalam penyelesaian masalah ini antara lain adalah garis besar penyelesaian masalah, pengumpulan data, persiapan data, pelatihan model dengan memanfaatkan *library* XGBoost, optimasi *Hyperparameter*, pengujian performa model.

BAB IV HASIL DAN ANALISIS

Bab ini menyajikan informasi tentang lingkungan dan perangkat yang digunakan untuk penelitian ini serta menyajikan analisis hasil penelitian yang menentukan *Hyperparameter* terbaik model XGBoost untuk prediksi penjualan toko ritel Big Mart.

BAB V PENUTUP

Bab ini berisi tentang kesimpulan dari bab-bab yang telah dibahas dan berisi saran sebagai bahan masukan untuk pengembangan atau penelitian selanjutnya.