

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penjualan merupakan indikator utama dalam keberhasilan bisnis dan memengaruhi strategi operasional maupun manajerial perusahaan. Dengan kemajuan teknologi, data transaksi penjualan kini dapat dianalisis secara lebih mendalam untuk memahami pola pembelian pelanggan, tren produk, serta memprediksi penjualan di masa depan. Prediksi atau peramalan adalah proses aktivitas meramalkan suatu kejadian yang mungkin terjadi di masa mendatang dengan cara mengkaji data yang sudah ada [1]. Prediksi penjualan yang akurat dapat membantu perusahaan dalam pengelolaan stok, perencanaan produksi, dan pengambilan keputusan strategis.

Salah satu platform yang saat ini berkembang pesat dalam penjualan produk fashion adalah TikTok Shop. Marketplace ini menawarkan model penjualan berbasis video pendek dan *live shopping* yang memungkinkan interaksi langsung antara penjual dan pembeli, berbeda dengan marketplace tradisional yang lebih berfokus pada katalog produk. Karakteristik unik TikTok Shop, seperti fitur promosi berbasis algoritma, keterlibatan pelanggan dalam bentuk komentar dan *like*, serta pengaruh influencer dalam strategi pemasaran, menjadikannya ekosistem yang dinamis dengan pola penjualan yang tidak selalu linier. Oleh karena itu, metode prediksi konvensional yang hanya mengandalkan data historis sederhana kurang efektif dalam menangkap kompleksitas pola transaksi di TikTok Shop [2].

Metode tradisional dalam memprediksi penjualan sering kali menggunakan pendekatan manual atau teknik statistik sederhana yang memiliki keterbatasan dalam memahami hubungan kompleks antar variabel yang memengaruhi penjualan [3]. Selain itu, metode ini membutuhkan waktu yang lama untuk menganalisis data dalam jumlah besar, sementara proses manual tidak dirancang untuk menangani skala data yang luas. Ketergantungan pada keahlian individu juga menyebabkan hasil prediksi menjadi tidak konsisten dan rentan terhadap subjektivitas, sehingga menurunkan akurasi dan keandalan prediksi[4]. Dalam konteks bisnis yang kompetitif, terutama di platform digital seperti TikTok Shop yang terus berkembang, keterbatasan ini menjadi hambatan besar bagi pelaku usaha yang harus mengambil keputusan strategis dalam waktu singkat dan berbasis data yang akurat. Oleh karena itu, diperlukan pendekatan yang lebih efisien, otomatis, dan

berbasis teknologi untuk memprediksi penjualan secara lebih akurat.

Seiring berkembangnya teknologi, pendekatan berbasis machine learning mulai banyak diterapkan dalam analisis prediksi penjualan. Salah satu algoritma yang banyak digunakan dalam prediksi penjualan adalah Extreme Gradient Boosting (XGBoost). XGBoost memiliki keunggulan dalam menangani masalah *overfitting* melalui mekanisme regularisasi yang lebih baik dibandingkan metode boosting lainnya[5]. Algoritma ini dikenal sebagai metode yang andal untuk memproses data dalam jumlah besar dengan performa tinggi. Selain itu, model XGBoost juga memiliki fleksibilitas dalam menangani data dengan skala besar serta variabel dengan distribusi yang tidak merata. Kemampuan algoritma ini dalam melakukan pemangkasan pohon keputusan yang tidak relevan membantu meningkatkan efisiensi pemrosesan data dan mempercepat waktu prediksi. Dalam konteks bisnis ritel yang memiliki data transaksi dengan variasi tinggi, seperti jumlah pesanan yang fluktuatif dan tren musiman, XGBoost dapat mengidentifikasi pola yang sulit ditangkap oleh metode konvensional. Dengan demikian, algoritma ini menjadi pilihan yang tepat dalam meningkatkan akurasi prediksi penjualan, terutama di platform dinamis seperti TikTok Shop[6].

Salah satu kelebihan utama XGBoost adalah kemampuannya dalam mengoptimalkan kecepatan dan efisiensi komputasi. Algoritma ini menggunakan komputasi paralel dan pengolahan berbasis cache, sehingga lebih cepat dibandingkan model *boosting* lainnya. Selain itu, XGBoost memiliki fitur *handling missing values* secara otomatis, yang memungkinkan model tetap bekerja meskipun terdapat data yang hilang dalam dataset[7]. Keunggulan lainnya adalah kemampuan menangani dataset besar dengan variabel yang saling berkorelasi, menjadikannya sangat efektif dalam analisis prediksi penjualan yang melibatkan banyak faktor. Selain itu, algoritma ini memiliki kemampuan *feature importance scoring*, yang memungkinkan pemilihan fitur yang paling berkontribusi terhadap hasil prediksi tanpa memerlukan metode seleksi fitur tambahan. Keunggulan lainnya adalah fleksibilitas dalam menangani berbagai jenis data, baik untuk klasifikasi maupun regresi, serta mendukung *hyperparameter tuning* yang memungkinkan optimasi model secara lebih efektif[8].

Meskipun memiliki banyak kelebihan, XGBoost juga memiliki beberapa kekurangan. Salah satunya adalah sensitivitas terhadap hyperparameter, di mana pemilihan parameter yang kurang tepat dapat menyebabkan model mengalami *overfitting* atau *underfitting*. Proses *hyperparameter tuning* yang kompleks sering kali memerlukan teknik optimasi

seperti Grid Search atau Bayesian Optimization agar mendapatkan performa terbaik. Selain itu, meskipun XGBoost lebih cepat dibandingkan metode *boosting* lainnya, algoritma ini masih memerlukan sumber daya komputasi yang tinggi, terutama saat bekerja dengan dataset yang sangat besar[9]. Kekurangan lainnya adalah kurangnya interpretabilitas, karena XGBoost menggunakan banyak pohon keputusan yang membuat model sulit dipahami dibandingkan metode seperti regresi linear. Selain itu, XGBoost tidak selalu optimal untuk dataset kecil, di mana metode sederhana seperti Decision Tree atau Logistic Regression terkadang memberikan hasil yang lebih baik dengan waktu pemrosesan yang lebih singkat. Dalam penelitian ini, data transaksi penjualan di TikTok Shop akan dianalisis dengan mempertimbangkan variabel-variabel yang berpengaruh terhadap penjualan, seperti harga produk dan jumlah unit terjual. Dengan menggunakan metode machine learning yang mampu menangkap pola kompleks dalam data, perusahaan dapat meningkatkan akurasi prediksi penjualan dan mengoptimalkan strategi pemasaran serta pengelolaan stok[10].

Dalam pengembangan model prediksi penjualan berbasis machine learning, *feature selection* menjadi langkah penting untuk meningkatkan akurasi dan efisiensi model. Dataset transaksi penjualan sering kali memiliki banyak fitur, namun tidak semua fitur memberikan kontribusi signifikan terhadap hasil prediksi. Oleh karena itu, XGBoost memiliki mekanisme bawaan dalam melakukan *feature selection* dengan mengukur *importance score* dari setiap fitur berdasarkan kontribusinya terhadap hasil prediksi. Dengan menggunakan metode ini, fitur yang paling berpengaruh akan dipertahankan, sementara fitur yang kurang relevan dapat diabaikan. *Feature selection* bawaan XGBoost ini lebih efektif dibandingkan metode manual, karena secara otomatis menyesuaikan bobot fitur sesuai dengan pola yang ditemukan dalam data[11].

Selain *feature selection*, optimasi hyperparameter dilakukan untuk meningkatkan performa model XGBoost dengan memastikan kombinasi parameter yang digunakan dapat menghasilkan prediksi yang lebih akurat dan stabil. Metode Bayesian Optimization digunakan untuk mencari kombinasi learning rate, max depth, dan number of estimators secara lebih efisien dibandingkan metode konvensional seperti Grid Search dan Random Search. Bayesian Optimization bekerja dengan membangun model probabilistik dari fungsi objektif dan menggunakan fungsi akuisisi untuk menentukan titik parameter yang paling potensial untuk diuji, sehingga dapat mengurangi jumlah iterasi yang diperlukan dalam pencarian parameter optimal. Dengan pendekatan ini, proses *hyperparameter tuning* dapat dilakukan lebih cepat tanpa kehilangan efektivitas dalam eksplorasi ruang parameter. Selain

itu, Bayesian Optimization mampu menyeimbangkan antara eksplorasi dan eksploitasi, yang memungkinkan pencarian nilai hyperparameter tidak hanya bergantung pada hasil sebelumnya, tetapi juga mempertimbangkan kemungkinan kombinasi yang lebih baik. Hal ini menjadikannya lebih unggul dalam menangani kompleksitas model XGBoost, terutama ketika bekerja dengan dataset besar dan variabel yang saling berkorelasi.[12].

Penelitian ini bertujuan untuk mengembangkan sistem prediksi penjualan berbasis analisis data transaksi menggunakan algoritma XGBoost. Fokus utama penelitian adalah mengoptimalkan parameter pada XGBoost menggunakan Bayesian Optimization untuk meningkatkan akurasi prediksi secara efisien dan menghindari eksplorasi parameter yang tidak optimal. Dengan pendekatan ini, model prediksi dapat secara adaptif menyesuaikan kombinasi learning rate, max depth, dan number of estimators untuk mencapai performa terbaik. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam mendukung pengambilan keputusan strategis perusahaan secara lebih efektif dan efisien, terutama dalam aspek pengelolaan stok, perencanaan produksi, dan strategi pemasaran di marketplace TikTok Shop. Output dari penelitian ini mencakup model prediksi penjualan yang lebih akurat, analisis fitur-fitur yang paling berpengaruh dalam penjualan, serta rekomendasi penerapan model dalam sistem bisnis digital. Dengan demikian, penelitian ini tidak hanya berkontribusi pada peningkatan akurasi prediksi tetapi juga memberikan wawasan bisnis yang dapat membantu perusahaan dalam pengambilan keputusan berbasis data.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, terdapat beberapa permasalahan utama yang menjadi fokus penelitian ini:

1. Bagaimana penerapan Bayesian Optimization dalam optimasi hyperparameter XGBoost untuk meningkatkan akurasi prediksi penjualan?
2. Bagaimana performa model XGBoost sebelum dan sesudah dilakukan optimasi hyperparameter menggunakan Bayesian Optimization berdasarkan metrik evaluasi MSE, MAE, RMSE, dan R2?
3. Bagaimana penerapan model XGBoost setelah dilakukan *tuning* hyperparameter untuk implementasi model prediksi penjualan?

1.3 Tujuan Penelitian

Tujuan Penelitian ini memiliki beberapa tujuan utama, yaitu:

1. Menerapkan Bayesian Optimization untuk mengoptimalkan hyperparameter

XGBoost dalam rangka meningkatkan akurasi prediksi penjualan.

2. Mengevaluasi dan membandingkan performa model XGBoost sebelum dan sesudah optimasi hyperparameter menggunakan Bayesian Optimization dengan menggunakan metrik evaluasi MSE, MAE, RMSE, dan R^2 .
3. Mengimplementasikan model XGBoost yang telah dioptimasi sebagai langkah untuk melakukan prediksi jumlah stok terjual bulan depan.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Akademis: Penelitian ini dapat menambah wawasan dan pengetahuan dalam bidang analisis data dan *machine learning*, khususnya dalam penerapan algoritma XGBoost untuk prediksi penjualan. Penelitian ini juga diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang ingin mengembangkan model serupa dengan data dan lingkungan yang berbeda.
2. Manfaat Praktis: Hasil dari penelitian ini dapat digunakan sebagai dasar pengembangan sistem prediksi penjualan yang lebih modern dan efisien. Sistem prediksi penjualan yang otomatis dapat membantu perusahaan dalam merencanakan stok, mengelola produksi, dan mengambil keputusan strategis secara lebih proaktif, sehingga dapat meningkatkan efisiensi operasional dan mengurangi biaya yang tidak perlu.
3. Manfaat Ekonomi dan Bisnis: Dengan adanya sistem yang dapat memprediksi penjualan secara cepat dan akurat, diharapkan dapat meningkatkan profitabilitas perusahaan melalui pengelolaan sumber daya yang lebih baik. Hal ini juga dapat membantu perusahaan untuk mengidentifikasi peluang pasar baru, mengurangi risiko kelebihan atau kekurangan stok, dan mendukung pertumbuhan bisnis secara berkelanjutan.

1.5 Batasan Masalah

Penelitian ini memiliki beberapa batasan sebagai berikut:

1. Dataset: Penelitian ini menggunakan dataset transaksi penjualan dari TikTok Shop yang mencakup informasi seperti tanggal transaksi, harga, jumlah produk terjual, dan kategori produk. Dataset dari platform lain selain TikTok Shop tidak dimasukkan ke dalam penelitian, sehingga hasil analisis hanya berlaku dalam konteks marketplace ini.

2. Variabel Penelitian: Penelitian ini berfokus pada analisis variabel utama yang secara langsung memengaruhi volume penjualan, seperti kuantitas transaksi, harga, dan waktu transaksi. Variabel eksternal, perubahan tren pasar, atau faktor ekonomi makro tidak dimasukkan dalam penelitian ini untuk menjaga fokus analisis pada aspek internal transaksi.
3. Parameter Model: Penelitian ini akan mengoptimalkan parameter utama dalam algoritma XGBoost guna meningkatkan akurasi prediksi penjualan. Namun, tidak semua parameter yang tersedia dalam XGBoost akan dioptimalkan, sehingga penelitian ini hanya mencakup parameter yang memiliki pengaruh signifikan terhadap performa model.
4. Evaluasi Model: Model prediksi yang dikembangkan akan dievaluasi menggunakan metrik MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) dan R^2 (Koefisien Determinasi) sebagai ukuran utama dalam menilai tingkat akurasi dan kesalahan prediksi model.
5. Optimasi hyperparameter: Proses optimasi hyperparameter dalam penelitian ini hanya menggunakan Bayesian Optimization, dengan fokus pada pencarian nilai optimal untuk parameter max depth, learning rate, n_estimators, subsample, dan colsample_bytree. Parameter ini dipilih berdasarkan dampaknya terhadap kinerja model XGBoost, dengan tujuan memperoleh kombinasi hyperparameter yang memberikan akurasi prediksi terbaik.