

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Perkembangan teknologi informasi yang terus meningkat seiring dengan kemajuan peradaban manusia juga berdampak pada dunia bisnis. Dalam menghadapi persaingan, pihak usaha perlu untuk memaksimalkan kemampuannya agar tetap bertahan dan bersaing dengan kompetitor lainnya. Dengan persaingan yang semakin ketat, pihak usaha tidak hanya harus berfokus pada produk yang ditawarkan, tetapi juga perlu mengarahkan perhatian mereka pada prediksi penjualan. Pemahaman terhadap analisis penjualan juga menjadi kunci penting dalam menciptakan strategi pemasaran yang lebih optimal agar dapat tetap bersaing di pasar yang kompetitif [1].

Seiring dengan berkembangnya pertumbuhan *e-commerce*, banyak pihak usaha mulai mengadopsi teknologi digital untuk meningkatkan efektivitas penjualan dan pemasaran mereka. Teknologi ini memungkinkan bisnis untuk mengumpulkan, menyimpan, serta menganalisis data historis untuk memahami analisis penjualan dengan lebih akurat. Berdasarkan studi terbaru, prediksi berbasis data historis telah terbukti dapat membantu pihak usaha dalam mengenali prediksi penjualan dan membantu menyusun strategi pemasaran yang lebih optimal [2].

Infinity Jar sebagai salah satu usaha yang berlokasi di Surabaya, merupakan toko online yang menjual produk dengan fokus pada produk-produk inovatif yang dirancang untuk menarik minat konsumen. Dalam upaya memperluas pangsa pasar dan memahami prediksi penjualan produk, pihak usaha perlu menganalisis produk-produk dengan permintaan tinggi serta mengidentifikasi prediksi pembelian konsumen. Dengan memanfaatkan data historis yang ada, Infinity Jar dapat memahami produk-produk dengan permintaan tertinggi, sehingga dapat membuat rencana pemasaran yang lebih optimal dan tepat sasaran [3].

Dalam konteks pemasaran, analisis data historis sangat penting untuk merancang strategi yang optimal. Dengan menggunakan teknik analisis modern, pihak usaha dapat mengidentifikasi produk-produk dengan permintaan tertinggi

yang sering tidak terlihat dengan metode tradisional. Misalnya, analisis ini dapat mengungkap produk dengan peluang penjualan tertinggi, dengan demikian dapat mengoptimalkan penggunaan sumber daya secara lebih efisien dan meningkatkan ROI (*Return on Investment*) dari kampanye pemasaran [4].

Di bidang analisis data, algoritma machine learning seperti XGBoost dan Random Forest telah terbukti efektif dalam memprediksi hasil berdasarkan data historis. XGBoost (Extreme Gradient Boosting) menggunakan teknik boosting untuk meningkatkan akurasi prediksi [5], sementara Random Forest menggunakan pendekatan *ensemble* dengan mengombinasikan banyak pohon keputusan untuk memperoleh akurasi yang lebih tinggi dan menekan kemungkinan overfitting. Karena kemampuannya menangani beragam variabel yang mempengaruhi hasil akhir, kedua algoritma ini sangat ideal untuk analisis data transaksi penjualan yang kompleks [6].

XGBoost (Extreme Gradient Boosting) merupakan algoritma machine learning berbasis boosting yang populer untuk analisis prediktif, terutama karena kemampuannya yang unggul dalam menangani data besar, tidak seimbang, dan dengan fitur berkolinear [7]. Dalam konteks analisis penjualan, XGBoost sangat relevan karena mampu mengolah data transaksi dengan jenis produk dan pola pembelian yang beragam, mengidentifikasi pola tersembunyi yang tidak terdeteksi oleh metode tradisional [8]. Selain karena hasil akhirnya, XGBoost menawarkan kecepatan dan efisiensi dalam proses pelatihan model, menjadikannya alat yang tepat untuk analisis penjualan dan optimalisasi strategi pemasaran di Infinity Jar.

Random Forest menawarkan keunggulan dalam hal interpretabilitas, karena algoritma ini membentuk sejumlah pohon keputusan secara paralel dan menggabungkan hasilnya guna menghasilkan prediksi yang lebih konsisten dan tepat [9]. Setiap pohon dalam Random Forest menganalisis peran variabel-variabel yang berbeda, sehingga memungkinkan penilaian terhadap pentingnya setiap variabel dalam mempengaruhi hasil akhir [10]. Kemampuan ini sangat bermanfaat bagi Infinity Jar karena memungkinkan pihak usaha untuk mengetahui prediksi penjualan produk, seperti jenis produk atau frekuensi pembelian. Dengan mengetahui hal-hal ini, mereka dapat menyesuaikan strategi pemasaran untuk mengoptimalkan penjualan mereka.

Dengan menerapkan kedua algoritma ini, Infinity Jar dapat memperoleh wawasan yang lebih mendalam terkait analisis penjualan. Misalnya, analisis data dapat mengungkap prediksi pembelian konsumen pada waktu tertentu. Informasi ini memungkinkan pihak usaha untuk merancang strategi pemasaran yang lebih optimal dan sesuai [11].

Meskipun ada banyak keuntungan, implementasi algoritma machine learning juga menghadapi tantangan. Salah satu tantangan utama adalah kebutuhan akan data yang berkualitas tinggi. Data yang tidak bersih atau tidak representatif dapat menghasilkan prediksi yang tidak akurat [12]. Oleh karena itu, data historis yang digunakan dalam analisis adalah data yang bersih dan relevan untuk menghasilkan wawasan yang akurat.

Dengan memanfaatkan algoritma XGBoost dan Random Forest dalam analisis data transaksi penjualan, Infinity Jar dapat mengoptimalkan strategi pemasaran mereka. Harapannya, penggunaan data dan teknologi ini tidak hanya dapat membantu menganalisis penjualan, tetapi juga memberikan pengalaman yang lebih baik bagi konsumen. Dengan menganalisis data historis, pihak usaha dapat memprediksi pembelian produk, sehingga strategi pemasaran yang mengoptimalkan pengelolaan stok dapat disusun dengan lebih terarah, seperti dengan menggabungkan produk permintaan tinggi dan permintaan rendah berdasarkan prediksi penjualan yang ada.

## **1.2. Rumusan Masalah**

Berdasarkan uraian latar belakang di atas, permasalahan yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana penerapan model hybrid XGBoost dan Random Forest untuk memprediksi penjualan produk?
2. Bagaimana performa model hybrid XGBoost dan Random Forest berdasarkan metrik evaluasi MAE, MSE, dan MAPE?
3. Bagaimana hasil prediksi dari model hybrid XGBoost dan Random Forest dapat diinterpretasikan untuk merancang rekomendasi paket produk sebagai strategi pemasaran?

### **1.3. Tujuan Penelitian**

Adapun tujuan dari penelitian ini dirumuskan berdasarkan rumusan masalah sebagai berikut:

1. Menerapkan model hybrid XGBoost dan Random Forest dalam memprediksi penjualan produk.
2. Mengevaluasi performa model hybrid XGBoost dan Random Forest dengan menggunakan metrik evaluasi seperti MAE, MSE, dan MAPE untuk mengetahui akurasi dan kehandalan model dalam memprediksi penjualan.
3. Menginterpretasikan hasil prediksi model hybrid XGBoost dan Random Forest untuk merancang kombinasi paket produk sebagai strategi pemasaran.

### **1.4. Manfaat Penelitian**

Penelitian ini memberikan manfaat bagi sejumlah pihak. Adapun manfaat penelitian ini bagi berbagai pihak terkait adalah sebagai berikut:

1. Bagi Akademisi
  - a. Menambah wawasan tentang penerapan hybrid algoritma XGBoost dan Random Forest dalam memprediksi penjualan produk untuk dapat merumuskan kombinasi paket produk sebagai strategi pemasaran berbasis data.
  - b. Menyediakan referensi untuk penelitian lebih lanjut di bidang prediksi penjualan, strategi pemasaran, dan penerapan machine learning dalam bisnis, khususnya dengan hybrid algoritma XGBoost dan Random Forest.
2. Bagi Praktisi
  - a. Menambah wawasan tentang prediksi penjualan produk.
  - b. Mengoptimalkan perancangan kombinasi paket produk sebagai strategi pemasaran dengan dukungan prediksi yang lebih akurat.

### **1.5. Batasan Masalah**

Adapun batasan masalah ini ditetapkan untuk memperjelas fokus dan arah penelitian sebagai berikut:

1. Data yang dikumpulkan merupakan data transaksi penjualan selama rentang waktu Januari 2022 hingga Desember 2024. Data pada bulan Desember 2024 akan digunakan sebagai periode validasi untuk mengukur akurasi prediksi penjualan yang dilakukan oleh model di periode tersebut.
2. Penelitian ini memprediksi penjualan produk untuk satu bulan ke depan, yaitu pada bulan Desember 2024.
3. Penelitian ini menggunakan algoritma XGBoost dan Random Forest untuk memprediksi penjualan produk, sehingga tidak mencakup dengan penggunaan algoritma lain.
4. Analisis akan dibatasi pada variabel-variabel tertentu yang relevan dengan prediksi penjualan seperti jumlah pembelian, jenis produk, dan tanggal pembelian.
5. Penelitian ini akan berfokus pada hasil prediksi dari model hybrid dan perumusan strategi pemasaran, tanpa mencakup implementasi langsung atau evaluasi strategi pemasaran dalam praktiknya.

*Halaman ini sengaja dikosongkan*