BAB I
PENDAHULUAN

1.1 Latar Belakang

Cryptocurrency adalah mata uang digital yang berjalan di atas teknologi *blockchain* dan menggunakan kriptografi untuk mengamankan pencatatan transaksi secara terdesentralisasi [1]. Pasarnya aktif 24 jam sehari tanpa terikat pada sistem perbankan, sehingga menarik investor dan peneliti [2]. Harga aset *cryptocurrency* berfluktuasi tajam (*volatile*) mengikuti sentimen pasar [3]. Pergerakan cepat ini membuka peluang keuntungan besar, tetapi juga membawa risiko yang besar juga. Karena itu, *trader* membutuhkan sistem peramalan harga yang akurat [2].

Untuk memperkirakan arah harga, *trader* biasanya menggunakan analisis teknikal, dengan menghitung indikator teknikal dari data historis, seperti *Relative Strength Index* (RSI), *Moving Average Convergence Divergence* (MACD), dan *Moving Average* (MA), untuk membaca momentum dan menghasilkan sinyal *buy* atau *sell* [4]. Model yang dilatih menggunakan data mentah mencapai rata-rata akurasi 52,11%, dan angkanya naik menjadi 72,15% ketika indikator teknikal dimasukkan sebagai fitur [5]. Namun, indikator teknikal memiliki kelemahan, yaitu terlambat merespons pembalikan arah pasar [6]. Indikator ini juga memperlakukan semua data secara sama tanpa memperhatikan fase pergerakan harga, sehingga menghasilkan sinyal palsu saat pasar dipenuhi gangguan acak (*noise*) [7].

Untuk mengatasi keterbatasan tersebut, *Hidden Markov Model* (HMM) mendeteksi rezim pasar yang tidak terlihat pada grafik harga [8]. Pada data *cryptocurrency*, HMM mengelompokkan kondisi data ke dalam tiga status tersembunyi, yaitu rezim pasar *bullish*, *sideways*, dan *bearish* [7]. Dengan menghitung peluang, HMM memetakan kecenderungan pasar untuk bertahan pada tren tertentu dan kemungkinan berpindah antarrezim berdasarkan pergerakan harga sebelumnya [8]. Status rezim pasar memberikan konteks tentang kondisi pasar sebelum *trader* mengambil keputusan transaksi [8].

Seiring perkembangan *machine learning* untuk data keuangan, *Extreme Gradient Boosting* (XGBoost) menunjukkan kinerja unggul dalam berbagai uji coba. Pada klasifikasi sinyal perdagangan tiga kelas (*Buy*, *Hold*, dan *Sell*), XGBoost mencatat akurasi 98,5%, lebih tinggi daripada *Random Forest* (97,5%) dan

Decision Tree (97,4%) [4]. Dalam penentuan arah harga, XGBoost mencapai tingkat keberhasilan 81,11%, sedangkan arsitektur *deep learning* LSTM hanya 46,81% [9]. Pada simulasi perdagangan *cryptocurrency*, XGBoost menghasilkan *Return on Investment* (ROI) 141,4% pada Bitcoin dan 246,6% pada Ripple, dengan *Sharpe ratio* masing-masing 1,78 dan 1,59 [10]. Struktur pohon keputusan bertahap menangkap hubungan rumit antarfitur, dan penalti regularisasi menjaga model dari *overfitting* [4].

Integrasi HMM berpotensi meningkatkan kinerja peramalan. Salah satu studi memangkas *Mean Absolute Error* sekitar 56% [8]. Namun, penelitian terdahulu tidak menggunakan HMM sebagai fitur masukan (*feature augmentation*) secara langsung, melainkan menggunakan HMM sebagai mekanisme pergantian model atau penyaring data sebelum pelatihan [8]. Penggabungan probabilitas rezim pasar HMM (*bullish*, *sideways*, dan *bearish*) secara langsung dengan indikator teknikal sebagai fitur dalam model klasifikasi jarang dieksplorasi [8]. Karena itu, belum jelas apakah penambahan fitur probabilitas rezim pasar dari HMM meningkatkan kinerja klasifikasi sinyal perdagangan (*Buy*, *Hold*, dan *Sell*) pada XGBoost dibandingkan dengan model *baseline* yang hanya menggunakan indikator teknikal sebagai fitur [4].

Celah berikutnya adalah minimnya pemahaman tentang cara model membuat keputusan prediksi. Studi terdahulu menggunakan algoritma *machine learning* tanpa menjelaskan bagaimana model menghasilkan prediksi dan hanya berfokus pada skor akurasi akhir [11]. Dengan probabilitas rezim pasar dan indikator teknikal dalam satu model, sulit untuk memastikan apakah fitur rezim pasar berpengaruh signifikan atau hanya menambah beban variabel [12]. Kontribusi fitur probabilitas rezim pasar perlu dibedah dan dibandingkan secara langsung dengan indikator teknikal melalui analisis tingkat kepentingan fitur (*feature importance*) [13].

Celah lain terletak pada ketidaksesuaian antara hasil evaluasi klasifikasi dan keuntungan dalam perdagangan langsung. Banyak penelitian menilai keberhasilan model hanya berdasarkan angka akurasi atau skor F1 pada data uji [14]. Metrik klasifikasi yang tinggi tidak selalu sejalan dengan keuntungan dalam simulasi perdagangan (*backtesting*) [15]. Model yang akurat dalam klasifikasi bisa merugi

bila terlambat mengantisipasi pembalikan tren atau terlalu sering membuka posisi, sehingga biaya transaksi menggerus keuntungannya [16]. Maka dari itu, perlu diuji apakah fitur rezim pasar menghasilkan keuntungan yang lebih baik dibandingkan dengan hanya menggunakan indikator teknikal dalam simulasi *backtesting*.

Berdasarkan uraian tersebut, penelitian ini membandingkan dua model dengan fitur masukan yang berbeda. Model XGBoost *baseline* hanya menggunakan fitur masukan berupa indikator teknikal, sementara model hibrida HMM-XGBoost menambahkan fitur masukan (*feature augmentation*) berupa nilai probabilitas rezim pasar dari HMM, yaitu *bearish*, *sideways*, dan *bullish*. Perbandingan kedua model mencakup evaluasi metrik klasifikasi, analisis *feature importance*, dan simulasi *backtesting* dengan memperhitungkan biaya transaksi. Tujuannya adalah menguji apakah penambahan probabilitas rezim pasar sebagai fitur masukan mampu meningkatkan kinerja klasifikasi, menganalisis besarnya kontribusi fitur melalui *feature importance*, serta menilai hasil perdagangan yang didapatkan.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah penelitian ini adalah:

1. Apakah penambahan probabilitas rezim pasar sebagai fitur meningkatkan kinerja klasifikasi sinyal perdagangan (*Buy*, *Hold*, dan *Sell*) pada XGBoost dibandingkan dengan model *baseline* yang hanya menggunakan indikator teknikal?
2. Seberapa besar kontribusi fitur rezim pasar HMM dibandingkan dengan indikator teknikal pada model XGBoost berdasarkan analisis tingkat kepentingan fitur (*feature importance*)?
3. Apakah strategi perdagangan model hibrida HMM-XGBoost menghasilkan hasil yang lebih tinggi dalam simulasi *backtesting* dibandingkan dengan model *baseline* yang hanya menggunakan indikator teknikal?

1.3 Tujuan

Berdasarkan rumusan masalah di atas, tujuan penelitian ini adalah:

1. Menguji apakah penambahan probabilitas rezim pasar dari HMM sebagai fitur dapat meningkatkan kinerja klasifikasi sinyal perdagangan (*Buy*, *Hold*, dan *Sell*) pada model XGBoost dibandingkan dengan model *baseline* yang hanya menggunakan indikator teknikal.

2. Menganalisis besarnya kontribusi fitur rezim pasar HMM dibandingkan dengan indikator teknikal pada model XGBoost berdasarkan analisis tingkat kepentingan fitur (*feature importance*).
3. Menguji apakah strategi perdagangan berbasis model hibrida HMM-XGBoost menghasilkan hasil yang lebih tinggi dalam simulasi *backtesting* dibandingkan dengan model *baseline* yang hanya menggunakan indikator teknikal setelah memperhitungkan biaya transaksi.

1.4 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat dari segi akademis dan praktis. Secara akademis, penelitian ini diharapkan dapat memberi gambaran mengenai pengaruh penambahan fitur probabilitas rezim pasar dari HMM terhadap kinerja model XGBoost, serta membantu memahami kontribusi fiturnya melalui analisis *feature importance*. Sementara untuk sisi praktis, penelitian ini dapat menjadi bahan pertimbangan bagi *trader* dalam menentukan posisi *buy* atau *sell* yang sesuai dengan tren pasar. Selain itu, adanya simulasi dengan biaya transaksi diharapkan bisa memberi perkiraan hasil yang lebih realistis pada perdagangan *cryptocurrency*.

1.5 Batasan Masalah

Untuk menjaga fokus pada pengaruh fitur rezim HMM terhadap kinerja XGBoost, penelitian ini menetapkan empat batasan berikut.

1. Data yang digunakan hanya pasangan SOL/USDT dan XRP/USDT dari bursa Tokocrypto pada interval 4 jam dengan rentang 1 Januari 2021 sampai 30 Juni 2026.
2. Fitur masukan hanya berasal dari data OHLC (tanpa volume, sentimen, atau data *on-chain*), yaitu 17 indikator teknikal dan 3 probabilitas rezim pasar HMM dengan tiga rezim (*bullish*, *sideways*, *bearish*).
3. Algoritma klasifikasi yang digunakan hanya XGBoost, dengan label *Buy*, *Hold*, dan *Sell* berbasis ATR (n = 14, k = 0,5) dan pembagian data kronologis 80:20.
4. *Backtesting* hanya menerapkan *spot trading long-only* dengan parameter tetap (alokasi 15%, Take Profit 3%, Stop Loss 4%, ambang keyakinan beli 0,38, biaya 0,4044%) tanpa memperhitungkan *slippage*.