PS-04: Stock Price Prediction + Algorithmic Trading (BTC/USDT)

Track / Theme	Quant / FinTech / Trading	Difficulty	Expert (2x)

Background / Context

Algorithmic trading in crypto markets relies on data-driven models to forecast price movement and convert forecasts into disciplined trading decisions. Compared to traditional markets, crypto is highly volatile and runs 24/7, making robustness, risk control, and realistic backtesting essential.

Core Challenge

- Build a predictive model using historical BTC/USDT market data (2017–2021) and evaluate generalization on a large held-out period.
- Design a trading strategy that converts model outputs into buy/sell/hold decisions and demonstrate profitability with proper risk controls.
- Ensure the full pipeline is reproducible: data loading → feature engineering → training → evaluation → backtesting → reporting.

Task 1: Price Prediction (Forecasting)

- Use historical data from 2017–2021. **Split the dataset so that more than 50% of the data (2021) is used for testing.**
- Train on the smaller portion (2017–2020 + part of 2021 if you choose), then evaluate on the larger 2021 test set.
- Primary metric: Mean Squared Error (MSE) on the test set price predictions.

Task 2: Algorithmic Trading Strategy (Execution)

- Build an algorithmic trading model for the cryptocurrency market using a **separate dataset from Task 1**, based on **daily or hourly price data**.
- Deploy machine learning models and computer programs by integrating financial market expertise, data analysis, statistical modeling, and programming skills.
- Generate trading signals and design an executable trading strategy based on the higher-frequency dataset.
- The objective is to unlock the potential of ML-based algorithmic trading to generate consistent and robust returns, outperform relevant benchmarks, and safeguard capital in the dynamic cryptocurrency market.

Must-Have Requirements

- NO UI/UX required, only a trained ML model with results is sufficient.

- Clear train/test split with test set dominated by 2021 (≥50% of total samples).
- Reproducible evaluation code for MSE (Task 1).
- Backtesting engine or framework (custom or library) that reports at least: total return, max drawdown, and a risk-adjusted metric (e.g., Sharpe).
- Benchmark comparison: Buy & Hold BTC/USDT over the same test period (and optionally a simple moving-average strategy).
- Visualizations: equity curve, drawdown curve, and at least one plot explaining signals/entries.

Nice-to-Have

- Walk-forward validation or rolling retrain (to mimic real deployment).
- Multi-timeframe features (hourly + daily) and regime detection (volatility, trend/range).
- Uncertainty estimation and confidence-based position sizing.
- Explainability: top drivers of signals or an interpretable policy.

Deliverables

10. Working code (repo / zip) + a README with setup + run instructions.
11. A short report (slides or doc) that explains: data used, features, model choice, training procedure, and results for Task 1 & Task 2.
12. Backtest summary table + plots (equity curve, drawdown, key trades).
13. A short demo (live or recorded) showing the pipeline end-to-end.

Dataset

Links:-

https://drive.google.com/file/d/1DmwNlpqGdhU8U_RsiEWXx-Wu1w0HTbUb/view

Judging Notes

This is the **highest-weighted** problem via the Expert difficulty tier. Judges will prioritize realistic evaluation (no leakage), risk management, and robustness over raw profit alone.