

nocodeML Workflow

Introduction

The primary philosophy behind nocodeML is to extract meaningful insights from trading data through the use of Machine Learning (ML), Deep Learning (DL), Statistics, and Mathematics. Our goal is to uncover hidden patterns and trends that are not readily visible to the human eye. This process involves several stages, each meticulously designed to ensure the accuracy and effectiveness of the analysis.

Workflow Overview

The workflow of nocodeML is divided into several stages, from data ingestion to advanced model exploration and performance analysis. Below is a comprehensive overview of each stage and its associated scripts:

1. Data Ingestion Preparation

Objective: The first stage is to retrieve and prepare data for analysis. We extract data from our NT8 market replay backtest, capturing trade data, indicator values (over 50 indicators), signal data, and event data (e.g., long or short positions).

Process:

- **Data Retrieval:**
 - Extract data from NT8 market replay backtest.
 - Record trade data, indicator values, signal data, and event data.
- **Data Storage:**
 - Store the retrieved data in a PostgreSQL database using our data ingestion preparation script.

Script:

- **Data Ingestion Preparation Script:**
 - This script performs extensive data cleaning and preprocessing.
 - Calculates additional indicators and metrics as required.
 - Generates a final dataframe called `merged_trade_event_data`, which contains all the indicators and the newly calculated metrics (e.g., % away, binary indicators).

2. Trading Dashboard

Objective: To create an interactive dashboard for analyzing and visualizing the performance of various machine learning models on the trading data.

Process:

- **Data Loading:**
 - Load data from the PostgreSQL database.
- **Data Preprocessing:**
 - Clean and preprocess the data for analysis.
- **Model Training:**
 - Train multiple classifiers using the preprocessed data.
- **Result Visualization:**
 - Display the results using Streamlit and Plotly for interactive exploration.

Script:

- **Trading Dashboard Script:**
 - This script integrates data loading, preprocessing, model training, and visualization into an interactive dashboard.
 - Users can explore the performance of different models on the trading data through an intuitive interface.

3. Model Exploration

Objective: To perform advanced model exploration, including detailed exploratory data analysis (EDA) and the training of various machine learning models.

Process:

- **Data Loading:**
 - Load the merged_trade_event_data dataframe from the PostgreSQL database.
- **Data Preprocessing:**
 - Clean and preprocess the data.
- **Model Training:**
 - Train a variety of machine learning models (e.g., Random Forest, Gradient Boosting, XGBoost, LightGBM, CatBoost, Neural Networks).
- **Exploratory Data Analysis:**
 - Perform detailed EDA to uncover patterns and trends.

- Calculate optimal win ranges using KDE distribution plots or histograms to identify the most interesting indicator values.

Script:

- **Model Exploration Script:**

- This script performs comprehensive model exploration and EDA.
- Includes advanced techniques for identifying optimal indicator values for trading strategies.

4. Strategy Performance

Objective: To analyze the performance of a trading strategy by calculating various performance metrics and visualizing the results.

Process:

- **Data Loading:**

- Load data from the PostgreSQL database.

- **Data Processing:**

- Process the data to prepare it for analysis.

- **Performance Metrics Calculation:**

- Calculate various performance metrics (e.g., net profit/loss, max drawdown, Sharpe ratio, Sortino ratio).

- **Result Visualization:**

- Visualize the results using Streamlit and Plotly for a comprehensive performance analysis.

Script:

- **Strategy Performance Script:**

- This script focuses on the detailed analysis of trading strategy performance.
- It provides insights into the effectiveness and risk associated with the trading strategies.

5. Trading Hours

Objective: To analyze trading activity and performance during different hours of the day, providing insights into the best times for trading.

Process:

- **Data Loading:**
 - Load event and merged data from the PostgreSQL database.
- **Data Processing:**
 - Convert timestamps to appropriate time zones (e.g., America/New_York).
 - Process data to calculate profit and loss for each hour.
- **Performance Analysis:**
 - Analyze trading performance by hour, including total trades, total profit, total loss, net profit, and win rate.
- **Result Visualization:**
 - Display the results using interactive plots (e.g., bar charts, scatter plots, box plots) to identify the best trading hours.

Script:

- **Trading Hours Script:**
 - This script analyzes trading activity by hour and visualizes the performance using various interactive plots.
 - Provides detailed insights into the trading performance during different hours of the day.

Conclusion

The nocodeML project is designed to harness the power of machine learning and data analysis to provide deep insights into trading data. By following a structured workflow from data ingestion to advanced model exploration and performance analysis, we aim to uncover hidden patterns and optimize trading strategies effectively. Each script is meticulously crafted to ensure robust data processing, comprehensive analysis, and intuitive visualization, making machine learning accessible and actionable for trading strategy optimization.