

Evaluation of Machine Learning and Deep Learning Investment Strategies

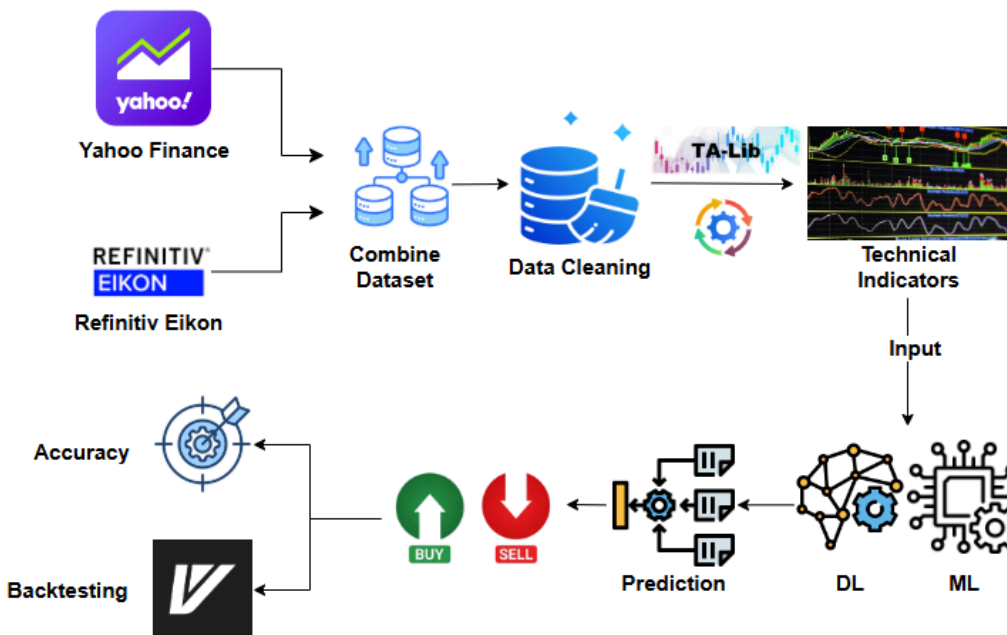
1. Overview

This project has a goal to analyze historical market data and technical indicators to create ML and DL-based trading strategies. The project automates data collection, feature extraction, model development, strategy backtesting, and portfolio performance evaluation and aims to maximize investment portfolio returns.

2. System Architecture

The basic structure is made up of these main parts:

- Data Acquisition: Used Yahoo Finance API to obtain historical stock market data for all types of market (Frontier, Emerging, Developed).
- Data Processing: Calculated technical features (RSI, MACD, Bollinger Bands, and Ichimoku Cloud) using the talib module and technical analysis library
- Model Training: Created ML models (Random Forest, XGBoost, LightGBM) and DL models (LSTM, GRU, CNN-Transformer) for strategy formulation and it's based on buy or sell signal.
- Backtesting Engine: Implemented Combinatorial Purged Cross Validation (CPCV) and risk metrics (Sharpe and Sortino ratios, Maximum Drawdown) for strategy evaluation.
- Deployment & Automation: Hosted on a version-controlled GitHub repository and packaged as a Jupyter Notebook.



3. Key Code Components

- `data_processing.ipynb`: This module is for data retrieval, data cleansing, and feature extraction using talib
- `frontier.ipynb`: Sets up and trains hyper-parameterized ML/DL models. Backtests strategies against performance metrics that are adjusted for risk.
- `Codebase`: Contains EDA and strategy performance visualization.

4. Expected Outcomes

- Attribution of optimization ML/DL algorithms for asset allocation and risk management resulted in automatic achievements.
- Accomplished better prediction results than with traditional approaches.
- Includes full coverage automated verification based on business rules and crosses over all markets.

5. Deployment Considerations

- This work can be replicated through the provided repository on [GitHub](#).
- Additional developments will include streaming market data, and RL for portfolio balancing.