

XGBoost 股票預測與回測系統說明文件

一、系統架構概述

本系統以 XGBoost 作為核心模型，建構一套自動化股票預測與回測流程。透過資料清洗、特徵工程、模型訓練、預測決策與回測模擬等步驟，評估策略在歷史資料上的表現，並可串接下單 API 實現自動交易。

二、預測流程說明

1. 資料蒐集 (Data Collection)

收集歷史股價與技術指標資料，常見欄位如下：- 時間序列欄位：日期、時間 - 價格類：開盤、收盤、最高、最低、成交量 - 技術指標：MA, RSI, MACD, KD 等 - 標籤欄位 (Label)：隔日上漲(1)或下跌(0)

2. 特徵工程 (Feature Engineering)

- 建立技術指標與延遲變數 (lag features)
- 標準化數值 (z-score 或 Min-Max)
- 清除無效值與極端值

3. 模型訓練 (Train)

使用 XGBoost 分類模型：

```
from xgboost import XGBClassifier
model = XGBClassifier(n_estimators=100, max_depth=3, learning_rate=0.05)
model.fit(X_train, y_train)
```

常見超參數：- `n_estimators`：樹數量 - `max_depth`：樹深度 - `learning_rate`：學習率 - `subsample`：隨機樣本比例

4. 預測 (Predict)

```
prob = model.predict_proba(X_test)[:, 1]
```

決策邏輯：- 機率 > 0.7 → 買進 - 機率 < 0.3 → 賣出

三、回測流程 (Backtesting)

1. 設計交易邏輯

依照預測結果決定買賣動作：

```
if prob > 0.7:
    action = 'buy'
elif prob < 0.3:
    action = 'sell'
```

2. 回測模擬交易行為

每日根據預測結果模擬下單，計算資產變化。

3. 評估績效

衡量指標：- 年化報酬率 - 最大回撤 (Max Drawdown) - 勝率 - Sharpe Ratio

四、自動下單整合

結合交易 API，例如：富邦 FubonAPI：

```
sdk.place_order(stock_id='2330', qty=100, side='buy')
```

自動化模組結構：1. Data Feed 模組 2. 預測與策略判斷模組 3. API 下單模組 4. 回測與績效追蹤模組

五、優勢與限制

優點：

- 準確率高、運算快速
- 支援非線性特徵
- 特徵重要性可視化

限制：

- 需嚴格交叉驗證防止過擬合
- 不適合即時高頻交易
- 過度依賴歷史資料可能失效

六、溝通簡報範例語句

我們的 AI 模型採用 XGBoost，能分析歷史數據中關鍵因素來預測股票漲跌，透過機率分類進行決策。整體系統可自動回測並接入下單 API，協助實現半自動或全自動交易策略。