

# Data-driven Methods in Finance: Final Presentation

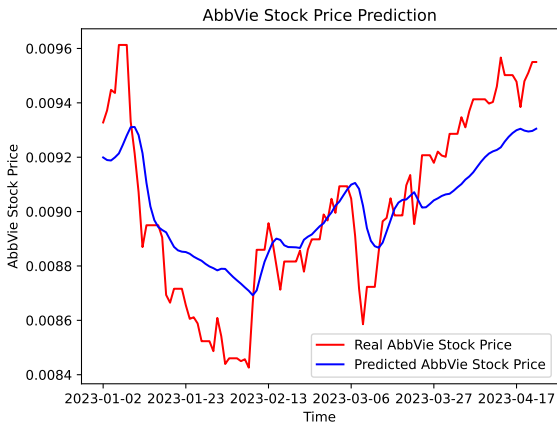
April 24, 2023

Use forecasting models to predict expected return and determine the weights for each asset by solving a constrained optimization problem.

Main models used are:

- 1 Prophet
  - developed by Facebook
  - can capture daily/weekly/yearly seasonality along with holiday effects
  - not in our data 😊
- 2 LSTM
- 3 ARIMA

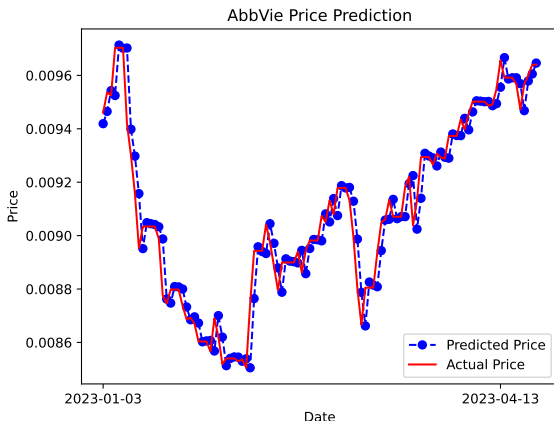
An example of using LSTM to predict stock price:



- 1 Normalize the data to  $[0, 1]$  through the min-max scaler
- 2 Time lag of 1 day and 60 time steps
- 3 4 hidden layers each with 100 neurons
- 4 1 neuron in the output layer for predicting the normalized stock price
- 5 It is better to fit the price than to fit the lead return

# ARIMA

Using the ARIMA model with parameters  $p = 4, d = 1, q = 0$ , and the rolling forecasting procedure: the model is recreated after each new observation is received.



# Portfolio weights

Data is obtained using python yfinance.

Risk-adjusted return

$$\begin{aligned} \min_w \quad & Aw^T \Sigma w - w^T \mu \\ \text{s.t.} \quad & \|w\|_1 = 1 \end{aligned}$$

where the first term in the objective is the risk, and the second term can be seen as the expected return.

$\mu$  is predicted using the average of the outputs from ARIMA model and LSTM model.

Risk-adjusted return

$$\begin{aligned} \min_w \quad & Aw^T \Sigma w - w^T \mu \\ \text{s.t.} \quad & \|w\|_1 = 1 \end{aligned}$$

- 1 diversification constraints

$$\underline{w} \leq w \leq \overline{w}$$

- 2 dollar neutrality  $w$  is rescaled so that

$$\|\max(w, 0)\|_1 = \|\min(w, 0)\|_1$$

# Performance analysis

For the past week (April 17 - April 21), the forecast performance is 0.159961 and the decision performance is  $-8.6$ . 51 out of 110 contribute positive return, and the **worse** 8 picks are:

symbol	name	weights	rank in submission
UNH	UnitedHealth Group	0.203	1
IEF	ETF	-0.099	110
XLK	ETF	$8 \times 10^{-4}$	26
CZR	Casears Entertainment	0.029	4
PRU	Prudential Financial	$2.8 \times 10^{-5}$	49
COP	ConocoPhillips	0.007	12
REET	ETF	$-2 \times 10^{-4}$	78
AXP	American Express	-0.014	99

If remove the worst one (UNH), the decision performance is 0.701.



# Performance analysis

For the past week (April 17 - April 21), the **best** 8 picks are:

symbol	name	weights	rank in submission
DPZ	Domino's Pizza	-0.043	108
CNC	Centene	0.106	2
XLU	ETF	-0.036	97
IAU	ETF	0.011	8
ICLN	ETF	-0.010	96
META	Meta Platforms	0.010	11
XLP	ETF	0.010	10
PG	Procter & Gamble	-0.029	103

# Future steps

- 1 Impose stronger diversification constraints
- 2 Collect data and extract factors/features. For instance, sentiment analysis using Refinitiv.