Stock Forecasting With Machine Learning

By David Jia

Quantitative/Algorithmic Trading Overview

Quantitative:

1. Strategy Identification:

   -Come up with a hypothesis (usually statistical or mathematical) in predicting future prices

2. Back-Testing:

   -Test and study results on other similar types of data

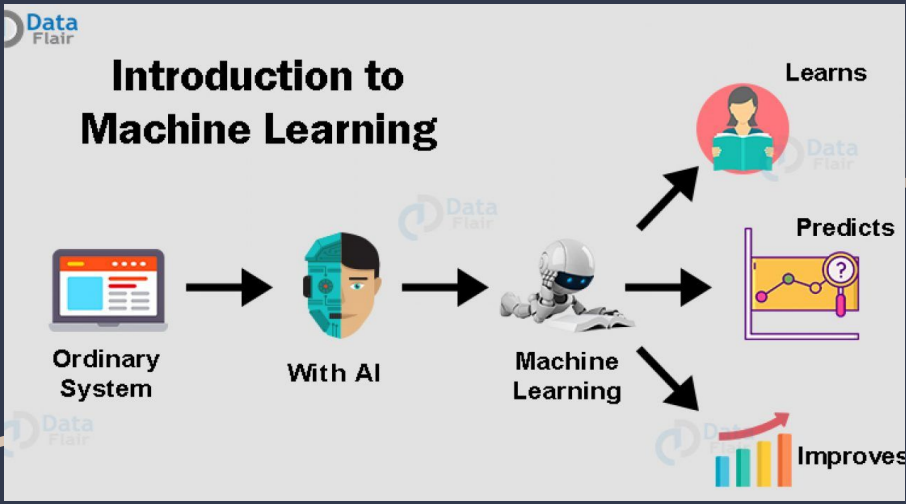
3. Risk Management:

   -Make sure trade sizes are appropriate, won't get blown up due to bad luck

Algorithmic: Strategies are sometimes automated and executed without human intervention (not necessary but can help with drawbacks of human emotion)

Quantitative Approach:

Why LSTM? ARIMA?



ARIMA Time Series:

Trains on gradient descent, mean-average

-Weighted-Moving Average

LSTM Neural Networks (Multi-Variate)

1. better “long-term memory” than traditional neural networks
2. multiple inputs

The Dataset

Taken from **Yahoo Finance** API

-2400 rows with data standardized

Stock List:

1. SPY: Top 500 US Companies
2. UVXY: Volatility Index
3. TLT: Bond-yield Derivative
4. GLD: Price of Gold

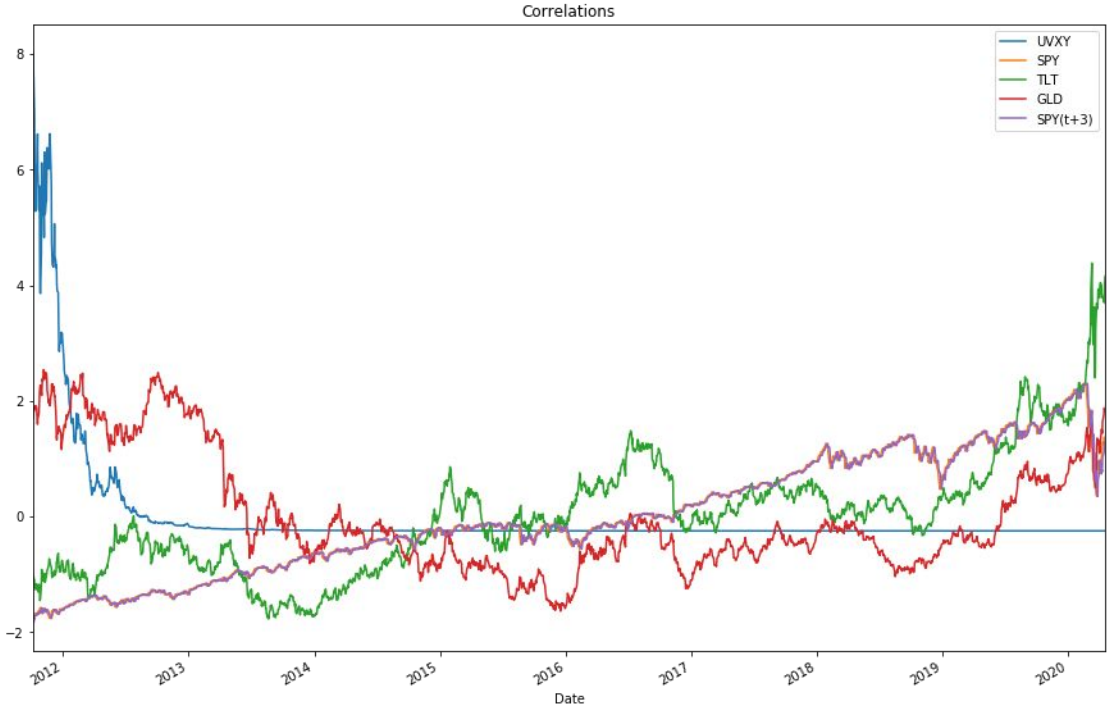
Fundamental Relationships

Quick overview of fundamental relationships between our selected stocks and commodities accordingly to economic theory.

Spy: generally considered a US economic indicator

1. **GLD**: should be inversely correlated with gold prices (gold is classified as a “safe haven” commodity)
2. **TLT**: should be inversely correlated with bond yields because risk-on environment = less money in bonds (safe), more in stocks (risk)
3. **UVXY**: inversely correlated to UVXY (more volatility and fear = greater stock sell-of)

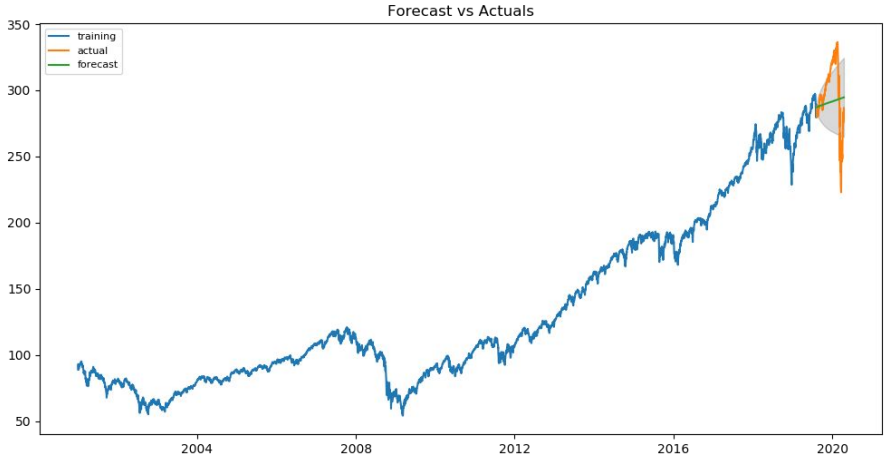
Explicit Relationships Recent

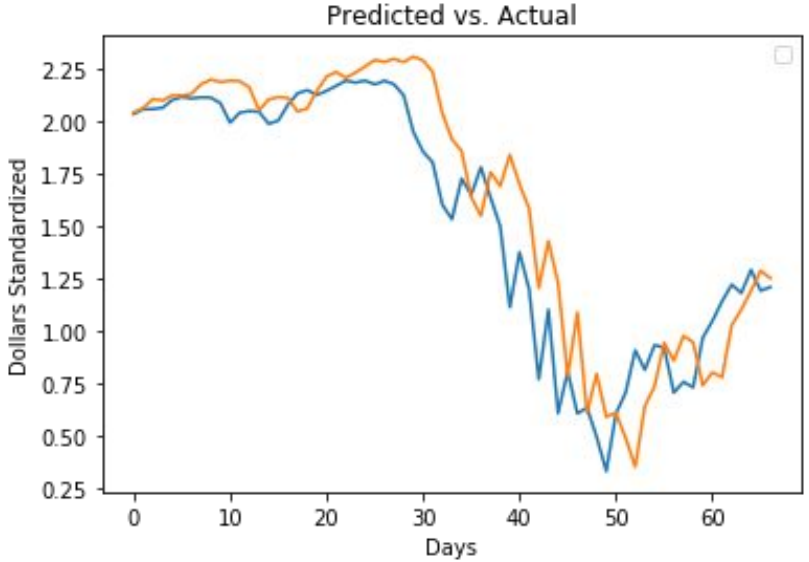


-hard to identify strict correlations

-alot of fundamental context in last ten years

Forecasting & Prediction





Possible Follow-up Projects

Dataset generally isn't the most robust,

Preliminary model and data to test whether there is an edge in LSTM neural networks in prediction

Next Dataset Possibilities:

More detailed dataset with smaller timeframes (1H, 5m)

-reasoning: known that it is harder to predict price further ahead of time than recent

-use different types of parameters instead of just stock/commodity prices such as volume or options profile.

Thanks for Listening!