Model Details

Overview

This model tries to use neural networks architecture applied to financial time series for forecasting prices. The aim is to construct an optimal portfolio allocation according to forecasted prices from the different neural networks. The allocation is made while using forecasted price returns and trying to maximize the overall return of the portfolio.

Input: time series of stock (or index or portfolio) price
Output: scalar of predicted stock price value

Model architecture: CNN or LSTM fine-tuned with the Kenneth French 49-industry monthly portfolios (ex-div). The portfolio allocation is next made with a sequential least squares programming optimizer according to forecasted returns.

Considerations

Use cases

The model has been trained over the Kenneth French 49-industry portfolios. The model has been tested with the monthly industry price returns ex-dividends from 1st January 2020 to 1st May 2024. The forecasted returns by neural networks models allowed to compute portfolio weights. Predicted returns were computed for a period of 48 months ahead. (because of the 10-month time step).

Limitations

The stock market is highly unpredictable so this model cannot fully predict the price of industry stock market.
The validation phase shows that we can experience very large error rate on some cases.

Ethical considerations

No filtering is made for controversial sectors so sectors like guns, oil, coal can be invested.

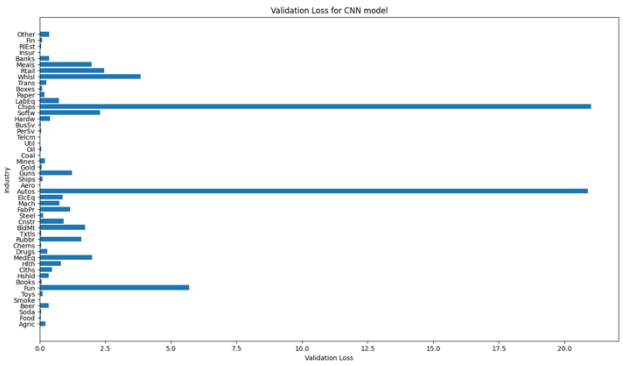
Train set

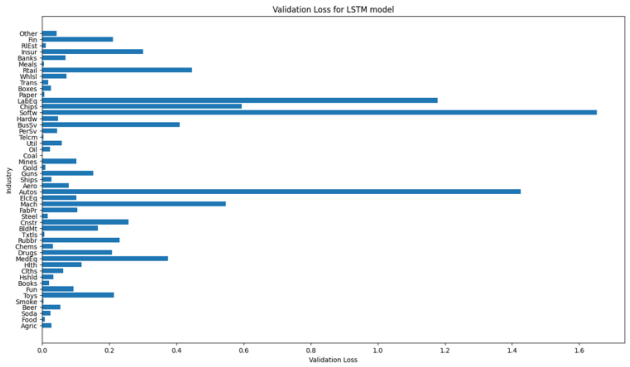
The training data was the Kenneth French 49-industry portfolios ex-dividend on monthly basis. The training set is composed of 80% of the total data available so returns from 1st January 2000 to 1st July 2019.

Validation set

The validation set is composed of the remaining 20% of the total data available. So from 1st august 2019 to 1st May 2024.

We can see that with LSTM the validation loss depending on sectors is smaller than the CNN model. However there is no uniform law according to the industry and so the path of the industry stock performance.

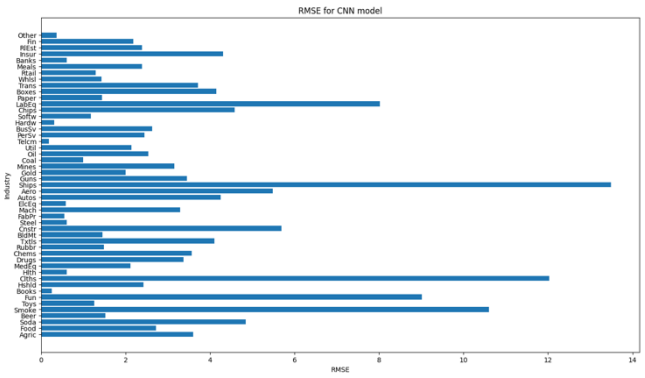


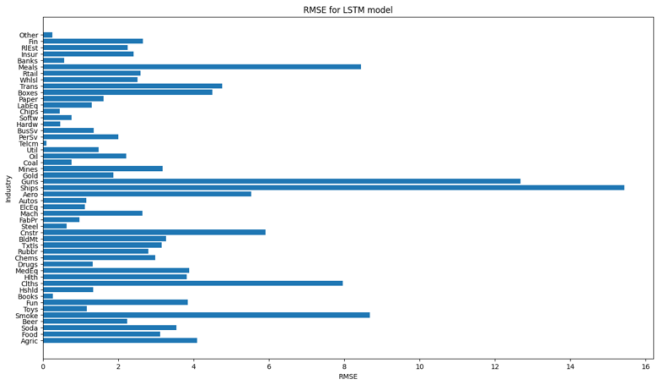


Quantitative Analysis

Performance highly depends on the industry time series tested.
Some of them experience a large root mean squared error in both models like Ships sector.
However, we can see that we still experience large RMSE for both models. This is because stock market timeseries are highly unpredictable due to their multivariate property.

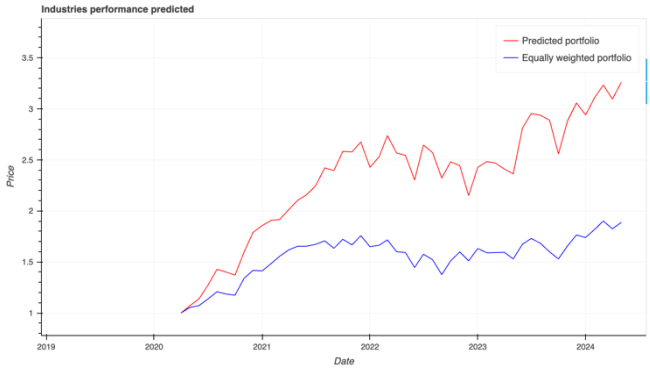
RMSE :

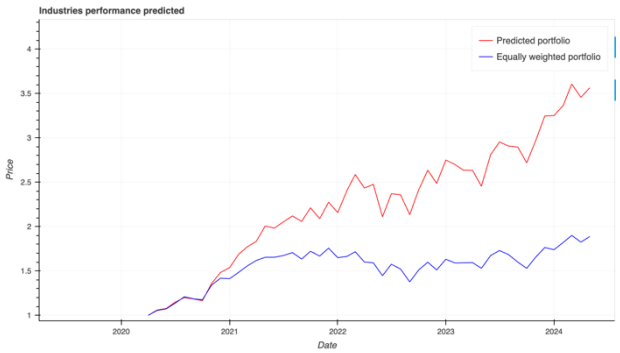




Portfolio price performance:

We can see that the LSTM model gives a more linear steady performance with less volatile that could be quantified by a Sharpe ratio.





Portfolio allocation

The optimizer is bound to give weights up to 15% of our portfolio.

We can see that the allocation is slightly different according to the selection of the CNN or LSTM model.
However the two models are very drastic in the selection because weights are maximized on a few industries instead of selecting multiple one with less than 15% portfolio allocation.