# MMAI 5500 Assignment 3

This assignment continues where assignment 2 ended. The goal is to implement the second part of deep portfolio method presented in the article *Deep learning for finance: deep portfolios* by Heaton, Polson & Witte (2016) and covered on lecture 1 of MMAI 5500.

## Submission

The assignment should be submitted as Python 3 code and uploaded to Canvas as a single `PY` file (**not** a Jupyter Notebook) and the trained model. The due date is on July 13 at 8:30am.

## Data

Use the same data as for assignment 2. The daily closing prices of 118 stocks from the IBB biotechnology index are provided in the file `assing2_data.csv`. The prices have been normalized and span the period from 2016 to end of 2020.

Load the data using the following helper function.

```python
def load_data(data_fname, benchmark_fname):
    """
    """
    data = pd.read_csv(data_fname, index_col=0)
    X = data.values[:, 2:].astype('float64')
    years = data['year']
    X_train = X[years < 2020.]
    X_valid = X[years == 2020.]
    tmp = data.index[data['year'] == 2020.]
    tickers = np.array([ticker.rstrip('_2020') for ticker in tmp])
    benchmark = pd.read_csv(benchmark_fname, index_col=0)

    return X_train, X_valid, tickers, benchmark
```

Use `X_train` to train and `X_valid` for model selection. The array `tickers` holds the ticker names corresponding to the rows in `X_valid`. The array `benchmark` holds to IBB index for the period corresponding to `X_train` and `X_valid`.

## Task

1. Load the data.
2. Modify the benchmark.
3. Normalize data and benchmark.
4. Use the method developed in assignment 2 to find the portfolio that best reproduces the benchmark during the time period of the training data.
5. Evaluate the portfolio on the time period of the validation data.

## Deliverable

You need to submit a single Python file (`PY` **NOT** `IPYNB`) that does the following:

1. Reproduce the IBB: plot a comparision of the portfolio (from assignment 2) and the IBB on the validation data.
2. Create a modified IBB where all returns less than -5% are replaced by 5%.
3. Select a new portfolio based on the modified IBB.
4. Reproduce the modified IBB: plot a comparison of the new portfolio and the modified IBB on the validation data.

**Your code should follow the PEP 8 style guide.** See the original PEP 8 style guide, an easier to read version, or a short PEP 8 YouTube intro. Practically, adding a PEP 8 plugin to your text editor (e.g. Falke8) will make it easier to follow to style guide.

## Grading

For full marks the submitted code needs to be bug free, execute the **5 steps** described under **Deliverable**, **select the same stocks (with at least 80% overlap) as a reference solution (unseen by students)**, and follow the PEP 8 style guide.

## Help

See MMAI 5500 lecture 1 slides.

See the Keras blog about autoencoders for hints about the implementation and the Keras model API for ideas about how to train and get the losses for individual stocks.

**Good luck!**