THIRD YEAR PROJECT REPORT

SUBJECT

IMPLEMENTATION OF A DEEP LEARNING ARCHITECTURE TO CREATE FINANCIAL FACTORS

Realized by:
Chady RAACH
Djamel CHAIBI
Mehdi ZEMNI
Mohamed Alaa MALLOULI

Supervised by:
Damien CHALLET
Jérémi ASSAEL

April 13, 2022

Abstract

The development of machine learning and deep learning methods has enabled the emergence of new techniques that can be used in asset pricing and particularly in characteristics-sorted factor models.
Through this project, we aim to highlight the contribution of deep learning frameworks in the generation of factors that are based on company characteristics. The main objective is to reduce pricing errors by training neural networks to generate deep factors (intermediate features) based on firm characteristics (inputs) in order to predict portfolio returns (outputs).
Another objective is to reproduce some benchmark sorted-factor models (CAPM, Fama French...) to test the robustness of the architecture.

Key words: Deep Learning, , Characteristics-Sorted Factor Models, Firm Characteristics, Deep Factors.

Introduction

We use multi-factor asset pricing models to describe the relationship between the expected return of assets and a set of common factors that reflect the market structure. They can be used to explain the return of one asset or a portfolio of multiple assets.

There are many categories of multi-factor models such as: macroeconomic models, fundamental models and statistical models. Besides, there are multiple methods to build those models (combination model, sequential model and intersectional model).

We will focus more on fundamental models that capture the relationship between the security's return and the firm characteristics (revenues, assets, debt level, market capitalization...) using the intersectional method which is based on sorting stocks using the intersection of multiple factors. This specific type of model is called :**Characteristics-Sorted Factor Model** and the main goal is to generate common factors to minimize pricing errors.

More specifically, we want to study the implementation of a deep learning framework that replicates a characteristics-sorted factor model. The major challenge consists in defining accurately the architecture and the loss function after understanding the mechanism of factor models.

The paper of *Feng, G., Polson, N. G., Xu, J. (2018)* [1] provides a complete deep learning architecture for asset pricing, which will be the basis of our work.

We are aware that deep learning is well known for complex architectures and “black box” models. However, we can reach a good level of transparency in this context by linking each part in the asset pricing mechanism with the deep learning elements.

Contents

1 Asset Pricing Factor Models

1.1 CAPM

Historically, the first asset pricing model was developed by Jack Treynor, William Sharpe, John Lintner and Jan Mossin in the 1960s. It is the CAPM (Capital asset pricing model). The CAPM is based on the following idea: investors are remunerated by the time value of money and by the risk. The time value of money is represented by the risk-free rate. The risk is represented by the beta, which is a historical ratio of the volatility of the price of an asset to the volatility of market prices in general (e.g. SP500 index).

$$\beta = \frac{\operatorname{Cov}(r_p, r_m)}{\operatorname{Var}(r_m)} \tag{1}$$

such that r_p is the implicit portfolio return and r_m is the market return. We obtain that the expected return on financial assets is equal to:

$$\mathbb{E}(r_{\mathrm{p}}) = r_f + \beta \cdot [\mathbb{E}(r_m) - r_f] \tag{2}$$

This model is based on strong assumptions:

- No transaction costs or taxes
- Short selling is allowed and has no impact on price
- Investors are risk averse and rational
- Investors control their portfolio risk through diversification
- The market is completely free and all assets can be traded.
- Single rate loan r constant
- The market is efficient
- Divisibility of assets

This model is very restrictive and it not seems to explain average returns on stocks and bonds.

1.2 Fama-French 3 factors

The Fama-French three-factor model is an empirical explanation of the expected return on a financial asset. Fama and French argued that the size and value factors capture a dimension of systematic risk that is not captured by market beta in the Capital Asset Pricing Model (CAPM). In fact, they observed that value stocks outperformed growth stocks. Besides, stocks with small capitalization outperformed stocks with high capitalization. In addition, small-cap stocks outperformed large-cap stocks. They therefore proposed to extend the CAPM, which resulted in the 3-factor model.

The cross-section of average returns on U.S. common stocks shows little relation to the market β of the asset pricing model developed before and variables that have no special standing in asset-pricing theory. These variables show reliable power to explain the cross-section of average returns. The list of empirically determined average return variables includes size, leverage, earnings/price (E/P), and book-to-market equity (the ratio of the book value of a firm's common stock, BE, to its market value, ME).

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1):3–56, 1993[2] identifies five common risk factors in the returns on stocks and bonds. Three stock-market factors: an overall market factor and factors related to firm size (ME, total number of a company's outstanding shares multiplied by the current market price of one share) and book-to-market equity (the ratio of the book value of a firm's common stock, BE, to its market value, ME). Two bond-market factors, related to maturity and default risks. Equity returns have a common variation due to equity market factors, and they are linked to bond returns by a common variation in bond market factors. With the exception of non-investment grade companies, bond market factors capture the common variation in bond returns. All five factors appear to explain average stock and bond returns.

[2] shows that the final result is that two empirically determined variables, size and book-to-market equity, explain well the cross-section of average returns of NYSE, Amex and NASDAQ stocks for the period 1963-1990. The size effect corresponds to the fact that small-cap stocks obtain higher returns than large-cap stocks. The value effect is the superior performance of stocks with low price-to-book ratios compared to those with high price-to-book ratios.

The proposed formula to describe the expected return Fama-French 3 factors is:

$$\mathbb{E}\left(R_p\right) - R_F = \beta\left[\mathbb{E}\left(r_m\right) - r_f\right] + s\mathbb{E}(SMB) + h\mathbb{E}(HML), \tag{3}$$

where $\mathbb{E}[SMB]$ is difference between the expected return of a small-cap portfolio and a large-cap portfolio and $\mathbb{E}[HML]$ difference between the expected return of a portfolio of securities with a high book-to-market ratio and one with a low book-to-market ratio.

1.3 Fama-French 5 factors

In 2015, Fama and French *Eugene F. Fama a , Kenneth R. French, A five-factor asset pricing model, Journal of Financial Economics Volume 116, Issue 1, April 2015, Pages 1-22*[3] adds two additional factors to their 3-factor model, namely profitability (stocks of companies with a high operating profitability perform better) and investment (stocks of companies with high total asset growth have below average returns).

This 5-factor model will certainly become the new standard in asset valuation.

$$\mathbb{E}\left(R_p\right) - R_F = \beta\left[\mathbb{E}\left(r_m\right) - r_f\right] + s\mathbb{E}(SMB) + h\mathbb{E}(HML) + r\mathbb{E}(RMW) + c\mathbb{E}(CMA) + e, \tag{4}$$

where RMW (Robust Minus Weak) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios, CMA (Conservative Minus Aggressive) is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios and e is a zero-mean residual.

The Fama/French 5 factors are constructed using the 6 value-weight portfolios formed on size and book-to-market, the 6 value-weight portfolios formed on size and operating profitability, and the 6 value-weight portfolios formed on size and investment.

According to [3] the five-factor model performs better than the three-factor. The five-factor model's main problem is its failure to capture the low average returns on small stocks whose returns behave like those of firms that invest a lot despite low profitability.

We want to improve these models, with machine learning to answer the problem: How can we construct a factor model which minimizes pricing errors or alpha?

2 Deep Learning for Asset Pricing

The main objective of introducing a deep model is to generate factors from firm characteristics in order to reduce the pricing errors in benchmark asset pricing models presented in the previous sections. By following the standard literature, the goal is to explain the excess return of a portfolio i by characteristic related factors. $R_{i,t}$ denotes the excess return of the portfolio i at time t that is explained by the following equation:

$$E(R_{i,t}) = \alpha_i + \beta_i^T E(f_t) + \gamma_i^T E(g_t) \tag{5}$$

The time series expectation difference α_i represents the pricing error of the portfolio to be reduced. g_t represents the benchmark factors vector while f_t represents the vector of generated factors that aims to better explain the excess return of market portfolios.

Therefore, the deep model has two different families of parameters to learn:

- Layers to build deep factors from firms characteristics. These parameters are specific to the market firms and do not depend on the portfolios to explain.
- Layers to fit a regression model that predicts portfolios excess returns from the benchmark factors and the generated factors. These parameters are specific to the portfolios and are tuned according to the portfolios to explain.

Even though this deep learning framework has a generative task, the end-to-end training allows to have a simple objective which is the minimization of the excess returns of market portfolios. The generation task is accomplished by the feedback of the pricing error task: reducing the alphas of the studied portfolios is done through the generation of pertinent factors from the market firms characteristics.

3 Deep Learning Architecture

3.1 Deep Learning Model

In this section, we introduce in details our deep learning model. The process can be devided into 4 main steps. First, we build deep characteristics for firms which are non linear transformations of the input firm characteristics (sec. 3.2. Then, we compute the deep factors (sec. 3.4) based on long-short portfolio weights W presented in section 3.3. In section 3.5, we discuss the optimization process, objective functions and metrics used for evaluation.

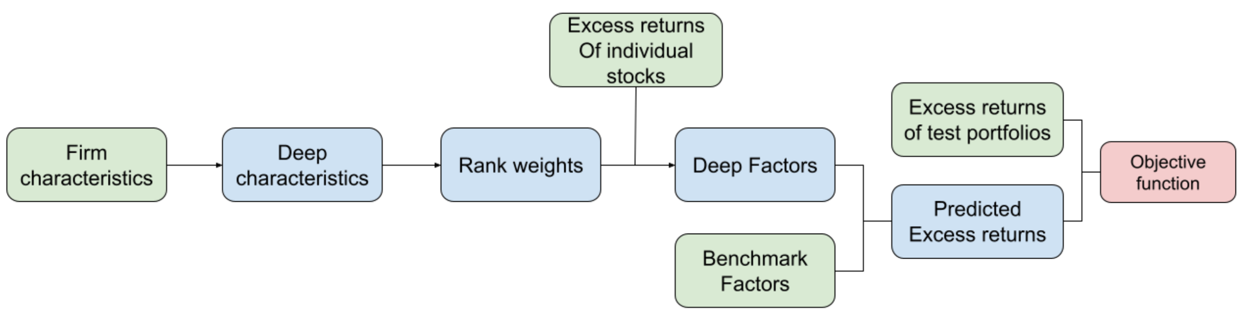

Figure 1: Pipeline of the deep learning model.

In fig. 2, we present the pipeline of the deep learning model. Cells in green denote data used during training either as an input variable or as a target variable. Blue cells denote variables computed at different parts of our deep learning model.

Following the same notation as the paper, we will use these 4 types of observations:

- $\{R_{i,t}\}_{i=1}^{N}$: excess returns of N training portfolios
- $\{j_{j,t}\}_{j=1}^{M}$: excess returns of M individual stocks
- $\{z_{k,j,t-1} : 1 \leq k \leq K\}_{j=1}^{M}$: K lagged characteristics of M firms
- $\{g_{d,t}\}_{d=1}^{D}$: D benchmark factors

All these variables are indexed by time t

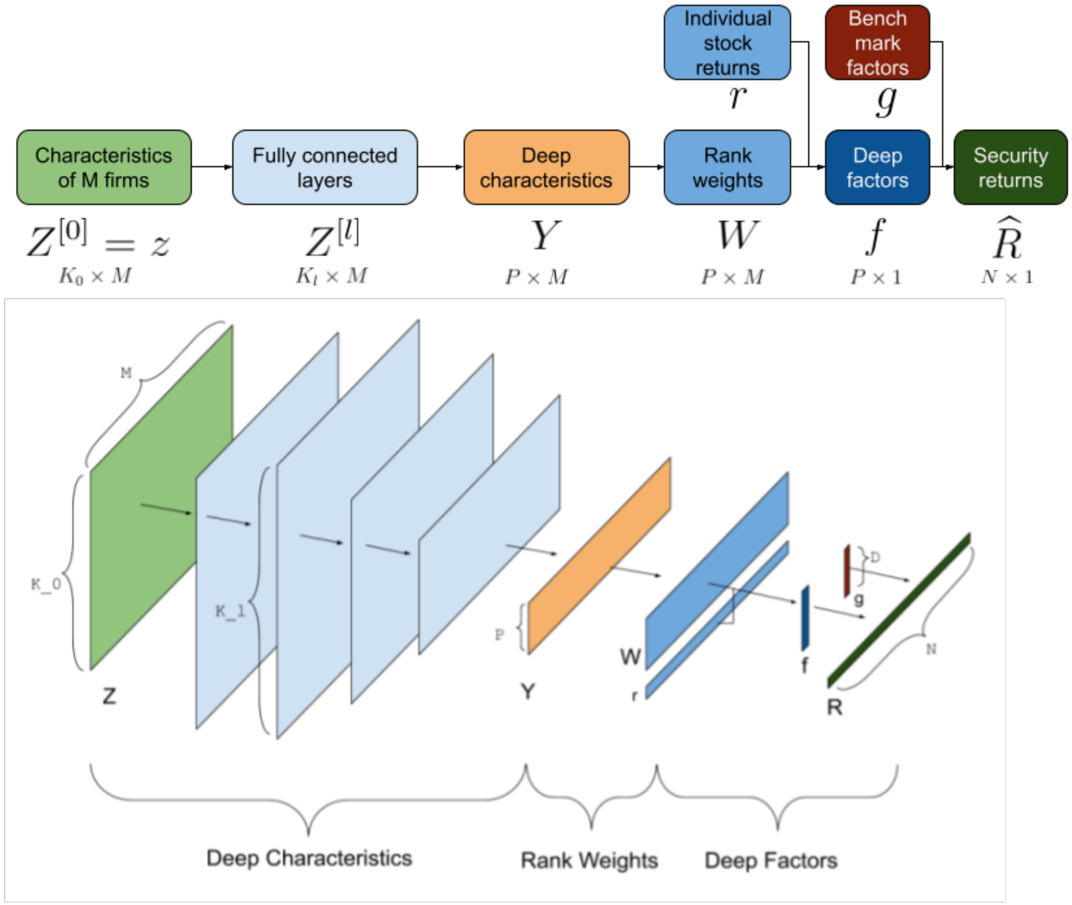

Figure 2: Deep Learning Network Architecture.

3.2 Deep Characteristics

The purpose of the stacked layers is to generate P deep characteristics from K raw characteristics. This is achieved using a fully connected network. This network is designed by authors to perform the same operation on all market firms: given two firms, the network encodes new characteristics for each, without interfering with other firms. Therefore, during training, batches will consist in a number of firms that are forwarded independently. The advantage of the fully connected layers over the standard asset pricing method is that we add trainable non linear transformations to encode low-dimensional firm characteristics. The model builds consequently more rich characteristics Y that have the advantage of encoding all the usual characteristics by non linear functions. They are also built by backpropagating information from layers that predict portfolios excess returns. This means that the selection of characteristics is oriented to better explain the training portfolio returns. So, if we feed the same characteristics that are used in Fama-French model to the deep structure we expect to reduce the portfolio alphas or at least obtain the same level of performance. This intuition is detailed by numerical experiments in the next sections of this report.

3.3 Nonlinear Rank Weights

Frazzini, A. and L. H. Pedersen (2014). Betting against beta. Journal of Financial Economics 111(1),1–25 [4] has developed their factor with a ”rank weighting”. Each stock is assigned to a ”high” or ”low” portfolio with a weight proportional to the cross-sectional rank of the stock’s estimate beta. This follows the idea developed by Fama and

French in their model with the separation of ME into small and big stocks.
Let $h^{[0]} : \mathbb{R}^M \to [-1, 1]^M$ be the function which computes the portfolio weights using the rankings of deep characteristics.

We introduce the softmax function:

$$\mathrm{softmax}\,(y_j) = \frac{e^{y_j}}{\sum_{j'=1}^{M} e^{y_{j'}}}, \sum_{j=1}^{M} \mathrm{softmax}\,(y_j) = 1$$

Given that y is a vector of a deep characteristic, we have:

$$h^{[0]}(y) = \underbrace{\begin{bmatrix} \mathrm{softmax}\left(y_1^+\right) \\ \mathrm{softmax}\left(y_2^+\right) \\ \vdots \\ \mathrm{softmax}\left(y_M^+\right) \end{bmatrix}}_{\text{long portfolio}} - \underbrace{\begin{bmatrix} \mathrm{softmax}\left(y_1^-\right) \\ \mathrm{softmax}\left(y_2^-\right) \\ \vdots \\ \mathrm{softmax}\left(y_M^-\right) \end{bmatrix}}_{\text{short portfolio}} \tag{6}$$

where $y^+ := y, y^- := -y$.
The first vector of $h^{[0]}$ reflects the weights of firms in the long portfolio and the second one represents those in the short portfolio.

3.4 Deep Factors

After the step of nonlinear rank weighting, we are left with long-short portfolio weights W. Deep factors step consists in creating the long-short factors f using individual stock returns denoted r and corresponding weights W.

3.5 Objective Function

As explained in section 2, the deep model is trained end-to-end to generate deep factors through the minimization of portfolios pricing errors. On one hand, the pricing error is the time series expectation. Using the same notations, we can express the pricing error of a portfolio i as an approximation of the expectation on a finite time series of length T:

$$\alpha_i = \frac{1}{T}\sum_{t=1}^{T}(R_{i,t} - \hat{R}_{i,t}) \tag{7}$$

where $\hat{R}_{i,t}$ is the predicted excess return for portfolio i at time step t.

On the other hand, the predicted excess return at a time step is a function of lag characteristics of market firms. While the alphas stands for the residual error that the factors are unable to explain, the time-serie error stands for the punctual error that the pricing model commits when predicting the excess return at time step t. Hence, this error should be penalized by the loss function in order to train the model on predicting accurate excess return as a function of lag characteristics, instead of only reducing the pricing error on the overall time serie. To formalize this, $\epsilon_{i,t}$ denotes the pricing error at time t for a portfolio i.

$$R_{i,t} = \hat{R}_{i,t} + \epsilon_{i,t} + \alpha_i \tag{8}$$

The first type error is called **cross-sectional error** while the second type is called **time-series error**. The loss function of the training combines both of them by an aggregation parameter λ. It is computed simultaneously on the set of N training portfolios. Finally, the objective functions of the training is the following:

$$\begin{aligned} \mathcal{L}_\lambda &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\alpha_i + \epsilon_{i,t})^2 + \lambda \times \frac{1}{N}\sum_{i=1}^{N}\alpha_i^2 \\ &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(R_{i,t} - \hat{R}_{i,t})^2 + \lambda \times \frac{1}{N}\sum_{i=1}^{N}[\frac{1}{T}\sum_{t=1}^{T}(R_{i,t} - \hat{R}_{i,t})]^2 \end{aligned} \tag{9}$$

The quality of trainings and the pertinence of the generated factors are sensitive to the diversity of the training portfolios. For example, if the model target is to generate one of the Fama-French factors, Fama-French portfolios are well adapted to the task. The backpropagation mechanism promotes the contribution of firm characteristics related to the economic context of the portfolios. However, the generated factors are very unlikely to correctly explain the returns of other industry portfolios, different from the Fama-French portfolios.

3.6 Evaluation metrics

In order to report the empirical results, three metrics are defined to assess the model capability of reducing pricing error:

- Time series $\mathcal{R}^2_{TS}$

$$\mathcal{R}^2_{TS} = 1 - \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(R_{i,t} - \hat{R}_{i,t})^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}(R_{i,t} - \overline{R}_i)^2} \tag{10}$$

  where $\overline{R}_i = \frac{1}{T}\sum_{t=1}^{T} R_{i,t}$. This metric evaluates how better is the time series prediction with respect to predicting a constant portfolio return equal to ground truth time series mean. Ideally, the fractional term tends to zero and the $\mathcal{R}^2_{TS}$ tends to one.

- Pricing error PE

$$PE = \frac{1}{N}\sum_{i=1}^{N}\alpha_i^2 \tag{11}$$

  where α_i is given by equation 7. Reducing the portfolio alphas means that the factors are able to better explain the excess return of portfolios by only fitting a linear regression.

- cross sectional $\mathcal{R}^2_{CS}$

$$\mathcal{R}^2_{CS} = 1 - \frac{Q}{Q_0} \tag{12}$$

  where $Q = \min_{\lambda}(X\lambda - \overline{R})^T(X\lambda - \overline{R})$, $Q_0 = \min_{\lambda}(\mathbf{1}_N\lambda - \overline{R})^T(\mathbf{1}_N\lambda - \overline{R})$, $X = [\mathbf{1}_N, \hat{\beta}]$, with $\hat{\beta}$ is the multivariate betas for factors and $\overline{R} = [\overline{R}_1, \ldots, \overline{R}_N]$. This metric evaluates the pertinence of introducing new factors to explain the portfolio excess returns compared to basic mono-factor model that predicts a constant value that aims to reduce the alpha for each portfolio.

4 Data

The implementation of the deep learning framework requires multiple types of data that cover a long period such as:

- Returns of training portfolios.
- Returns of individual stocks.
- Firm characteristics that can be reported with different frequencies (daily, monthly, quarterly…)
- Benchmark factors like the Fama French factors (SMB, HML, RMW…)

4.1 Data Source

We had difficulties in collecting reliable data from a centralized source given that information about financial characteristics of firms is not publically available. At the beginning , we had access to a pile of data which had the following description:

- **Region:** United States and Europe.

- **Number of firms:** 619 in the US market / 859 in the European market.
- **Time range:** 2016-2021.
- **Variables:** Total Assets / Total liabilities / Market Capitalization / Closing prices.

The data described above had some limitations which affected the empirical results after the training process:

- The insufficient number of companies that don't reflect the market behavior in terms of market capitalization.
- The short time range that does not allow for a robust statistical analysis.
- The limited number of characteristics that does not allow for the generation of multiple deep factors.

Eventually, we managed to find a reliable data source that allowed us to extract a great part of the data used in training the models: **Wharton Research Data Services (WRDS)**

WRDS provides a centralized tool to extract and analyze financial data in the US market for academic purposes by giving access to several certified data sources :

- **Center for Research in Security Prices (CRSP):** CRSP maintains a comprehensive collection of security price, return and volume data for all NYSE, NASDAQ and AMEX markets.
  Target variables (returns , volumes..) were extracted with a monthly frequency between 1985 and 2021.
- **Compustat:** Compustat provides financial statement and market data for over 80,000 publicly traded companies. The data covers income statements, balance sheets and cash flows on a quarterly and annual basis.
  Target variables (revenues, cash flows, debt levels, expenses...) were extracted with a quarterly frequency between 1970 and 2021.
- **Fama-French Portfolios Factors:** We used this table and another open source data to download historical data that includes: Fama French monthly factors (SMB, HML, Market Excess Return…) and excess returns of multiple portfolios (Fama French portfolios, industry portfolios…)

Remark: WRDS provides an efficient method to combine **CRSP** and **Compustat** databases using **Link History Tables** which contain the link-history references between the two sources.

4.2 Merging monthly and quarterly data

We merge the monthly and quarterly data to construct a unique dataset containing all characteristics with a monthly frequency. For each month, we assign the data of the previous quarter to the quarterly characteristics and we keep the data of the current month for monthly characteristics. In fact, quarterly data is often published lately and we need to take this delay into account to construct our dataset in a consistent way.

4.3 Data description

The data sample that we used is from January 2000 to December 2021. This choice was done to avoid the heavy computation time which was required to merge the monthly and the quarterly data. Besides, we think that a time range covering 21 years is sufficient for the reliability of the experiments.

Each month, the data is sorted according to the Market Equity and we take a specific number n_{sort} of firms that cover the maximum of the total market capitalization in the given stocks universe. This number n_{sort} depends on the availability and the frequency of characteristics data that we are going to use to train our model.

The monthly-updated data includes: share volume, Return without dividends, Number of Shares Outstanding, Adjusted Price, Market capitalization.

The quarterly-updated data includes Total Assets, Total Liabilities, Book Equity, Total Revenues, Interest Expense, Selling, General and Administrative Expenses ,Deferred Taxes and Investment Tax Credit , Depreciation and Amortization , Non-Operating Income, Invested Capital, Operating Expense, Investment and Operating Profitability.

Remark:

- **Operating Profitability :** represents Total Revenues minus Cost of Goods Sold, Interest Expense, and Selling, General, and Administrative Expenses divided by Book Equity for the last fiscal year end in t-1.
- **Investment:** represents the change in Total Assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 Total Assets at the end of each June.

These 2 characteristics are used to generate the factors SMB, RMW and CMA in Fama-French 5 factor model.

The portfolios include: 6 Fama French portfolios formed on Size and Book-to-Market, 25 portfolios Fama French formed on Size and Book-to-Market, 5 Industry portfolios and 10 Industry portfolios.

5 Experiments and results

5.1 Validating the model with Fama-French 3-factor model

5.1.1 Experience settings and training results

In order to test the capability of the model to produce economically meaningful factors, we conduct a first experiment where we try to reproduce one of the 3 Fama-French factors described in section 1.2. We feed the model with firm characteristics that are used to produce the two factors SMB and HML. We also feed the model with 2 among 3 benchchmark factors computed using the classical Fama-French method.

In this experiment, we feed the model characteristics of 2300 with highest market equity. The model is only fed with firms Market Equity (ME) and Book Equity (BE) since they are the only characteristics that interfere in producing SMB and HML factors. The time range for training and validation do not overlap: training uses data from January 2000 to August 2017 while the validation is performed on data from September 2017 to December 2021. The length of training time series is equal to 30, which allows to give more sense to the time series error in the objective function given by equation 9. The number of layers in the deep structure that builds factors is equal to 1 or 2. The model is fed with 2 Fama-French benchmark factors which are HML or SMB and market free-risk rate.

In Figure 3, we show the evolution of training loss and validation loss. The loss decreases steadily until conver-

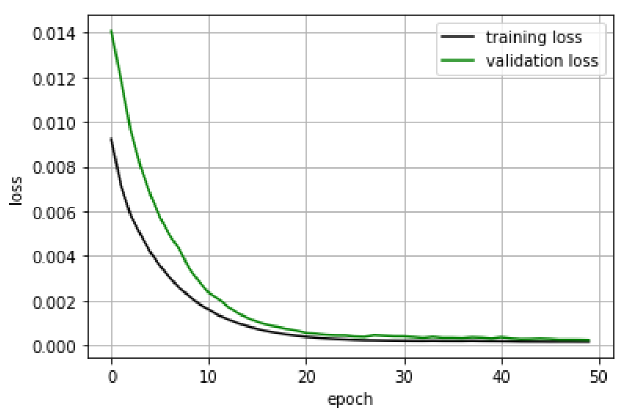

Figure 3: Training loss and validation loss evolution.

gence. The asset pricing error on validation time series is reduced when involving the newly generated deep factor.

5.1.2 Deep Learning Model Improvement

To test in more realistic conditions, the train and test portfolios are different in this section. In the training stage, the model learns how to predict a deep factor using the fed firms characteristics through the backpropagation of loss value computed on training portfolios. In the testing stage, only the generative part of the model is used to predict the value of the deep factor. On head of the inferred factor, we fit a regression model to new portfolios returns. To assess the added explainability of the deep factor, we compare the performance of the model using 2 Fama-French factors (2FF) with the trained model using the same 2 Fama-French factors in addition to the generated deep factor (2FF $+$ 1 deep factor).

In tables 1-2, we train the model using the 25 Portfolios Formed on Size and Book-to-Market (5×5) and test on the 6 Portfolios Formed on Size and Book-to-Market (3×2)[1]. In table 1, we report the result when the target factor to rebuild is SMB. In table 2, we report the result when the target factor to rebuild is HML. In both case, the generated factor improves the statistical evidence, the economic evidence and reduces the pricing error. This proves that the model is able to encode information from the characteristics of the market firms, that is not expressed by the 2 benchmark factors. In figure 4, we show a scatter plot of an unseen portfolio on the unseen time range during training. We can see that the pricing error is slightly improved with respect to the benchmark Fama-French 3-factor model.

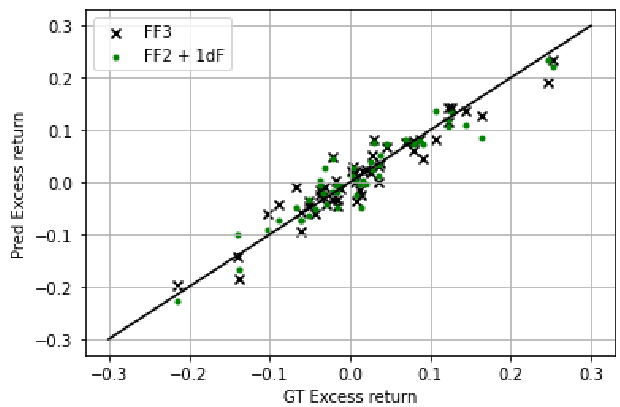

Figure 4: Ground truth excess return vs. predicted excess return for an unseen Fama-French portfolio.

5.1.3 Interpreting Deep Factors

To better understand the generated deep factor, we compute its correlation with the ground truth Fama French 3^{rd} factor that was not seen by the model. We report the result in the last row of tables 1-4. In the first case, when we train the model using the 25 Portfolios Formed on Size and Book-to-Market (5×5), we obtain a strong correlation with the Fama-French factor. In the second case, when training on the other set of portfolios, correlation with HML decreases, but the deep factor keeps improving the performance of excess return prediction in the test stage. This proves that the structure of the training portfolios has an impact on the quality of the generated deep factor. In figure 5, we illustrate the correlation between SMB factor and the factor generated from Market Equity (ME) and Book Equity (BE) of market firms. Note that the sign of the factor generated by the model is arbitrary and the generated factor may have the opposite sign of the Fama-French factor, but they are still strongly correlated.

With same trained model, we generate a factor then fit the regression coefficients to 10 industry portfolios. We have observed that that the (2FF $+$ 1 deep factor) model does not improve any evaluation metric (Statistical evidence levels off at 0.74 and economic evidence levels off at 0.85). Worse, pricing error are increased by 0.4 points. This

[1] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/tw_5_ports.html

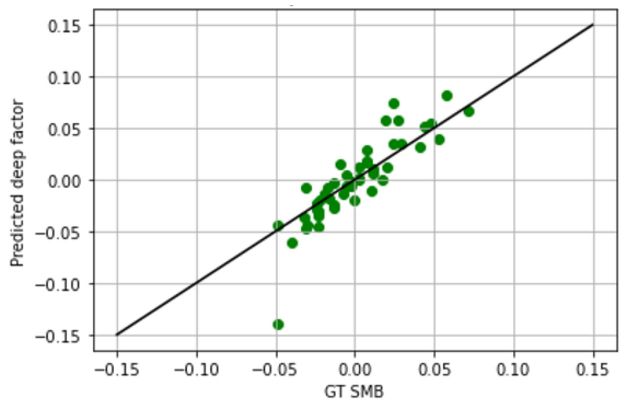

Figure 5: Scatter plot: x-axis contains SMB Fama-french factor, y-axis contains generated deep factor when the model is fed with HML and Market Risk-free rate. Correlation is equal to 0.88

last experiment shows the limits of deep frameworks: the generated factor is overfitted to Fama-French portfolios and poorly generalizes to other portfolios.

5.2 Deep learning model with Fama-French 3 factors as a benchmark model

In this experiment, we try to analyze the effect introduced by deep factors generated by the deep model, compared to the Fama-French 3 factors model.
The train and validation sets are the same as the previous experiment.

First, we use the 6 Portfolios Formed on Size and Book-to-Market (3×2). We have independently trained $2 * 2$ deep learning models with 1 or 2 layers and 1 or 2 deep factors.
The results are reported in Table 5. We can see that the FF3 model performed better than the deep model which is predictable given that we used the Fama-French portfolios that are the basis of construction of the Fama-French factors. Besides, we can see that the addition of deep factors has improved the performance of the deep model.

Second, we use the 5 industry portfolios. The results are reported in Table 6.
The analsysis $\mathcal{R}^2_{TS}$ and PE values shows that the use of these portfolios has impaired the overall performrance compared to the Fama-French portfolios.
By comparing the deep model and Fama-French, we can see that the values of $\mathcal{R}^2_{TS}$ and $\mathcal{R}^2_{CS}$ are very similar. However, the pricing errors PE of the deep model are relatively higher than the Fama-French 3 factors model. This can be due to the nature or the lack of characteristics that were fed to the model.

5.3 Deep learning model and Fama-French 5 factor model

In this part, we add more firm characteristics to the dataset including investment and operating profitability which are used in the additional 2 Fama-French factors in the 5 factor model. In this experiment we train our model using the 6 portfolios formed on size and book to market and we try each time to learn a missing deep factor (CMA (Conservative Minus Aggressive) or RMW (Robust Minus Weak)).
Unlike previous experiments, this time we didn't limit the input characteristics to the ones that are needed to build the missing Fama-French factor but we added many other characteristics (21 in total) to see the impact on the performance and on the correlation between the predicted deep factor and the removed benchmark factor.

The characteristics :Investment and Operating Profitability were calculated based on other characteristics based on the formulas given in section 4.3.

In table 7 and 8, we summarize the results of the experiments. We notice that there is no improvement compared to the model with 4 Fama-French factors. This suggests that the better the benchmark model (relevant benchmark factors) the more difficult it becomes to learn additional factors to improve the model. Additionally, we have observed very low correlation between the missing benchmark factor and the predicted one (between 0.13 and 0.2 depending on the experiment).

6 Conclusion

In this project we managed to implement a deep learning architecture to first reproduce the Fama-French factors to test it and then to try to find new factors to explain the cross section of expected returns not explained by the Fama-French factors. We have proved that it is possible to generate significant factors using deep structures. Nonetheless, we have proved the vulnerability of these models to poor generalization capacity and overfitting to training data. One main solution to encounter this problem is to feed the model with an important number of firm characteristics and to train the model using diversified portfolios.

Our main problem has been the accessibility to financial data which is not freely available and can be very expensive.
It would be appropriate to apply our model to other types of data, such as ESG data, at a later stage. We would have liked to test this but we could not have access to these types of data because of their price and their recent creation (not available for all companies and not available over a long time range).

References

[1] Guanhao Feng, Nicholas G. Polson, and Jianeng Xu. Deep learning in characteristics-sorted factor models. 2018.
[2] Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1):3–56, 1993.
[3] Eugene F. Fama a , Kenneth R. French, A five-factor asset pricing model, Journal of Financial Economics Volume 116, Issue 1, April 2015, Pages 1-22.
[4] Frazzini, A. and L. H. Pedersen (2014). Betting against beta. Journal of Financial Economics 111(1),1–25
[5] Yuriy Gorodnichenko and Michael Weber. Are sticky prices costly? evidence from the stock market. American Economic Review, 106(1):165–99, January 2016. 11

	L = 1			L = 2		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF2	0.86	3.28	0.93	0.86	3.28	0.93
FF2 + 1 deep factor	0.96	0.28	0.94	0.96	0.27	0.95
Improvement	11,63 %	-91,46 %	1,08 %	11,63 %	-91,77 %	2,15 %
Correlation DF% SMB factor	0.88			0.81		

Table 1: Predict SMB with Deep factor (6 Fama-French portfolios)

	L = 1			L = 2		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF2	0.86	4.22	0.87	0.86	4.22	0.87
FF2 + 1 deep factor	0.91	0.42	0.98	0.91	0.25	0.98
Improvement	11,63 %	-91,46 %	1,08 %	11,63 %	-91,77 %	2,15 %
Correlation DF% HML factor	0.88			0.81		

Table 2: Predict HML with Deep factor (6 Fama-French portfolios)

	L = 1			L = 2		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF2	0.8	3.12	0.86	0.8	3.12	0.86
FF2 + 1 deep factor	0.9	0.82	0.85	0.89	1.31	0.86
Improvement	12.50%	-73.72%	-1.16%	11.25%	-58.01%	0.00%
Correlation DF% SMB factor	0.88			0.81		

Table 3: Predict SMB with Deep factor (25 Fama-French portfolios)

	L = 1			L = 2		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF2	0.81	4.44	0.77	0.81	4.44	0.77
FF2 + 1 deep factor	0.85	0.85	0.84	0.85	0.82	0.85
Improvement	4,94 %	-80,86 %	9,09 %	4,94 %	-81,53 %	10,39 %
Correlation DF% HML factor	0.41			0.43		

Table 4: Predict HML with Deep factor (25 Fama-French portfolios)

	L = 1			L = 2		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF3	0.97	0.2	0.97	0.97	0.2	0.97
1 deep factor	0.94	0.91	0.99	0.96	0.39	0.93
2 deep factors	0.96	0.35	0.99	0.97	0.44	0.99

Table 5: Deep Learning model with a Benchmark model: FF3 (6 Fama-French portfolios)

	L = 1			L = 2		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF3	0.87	0.48	0.99	0.87	0.48	0.99
1 deep factor	0.86	1	0.99	0.86	1.6	0.99
2 deep factors	0.84	1.03	0.99	0.86	1.2	0.99

Table 6: Deep Learning model with a Benchmark model: FF3 (5 industry portfolios)

	L = 2			L = 3		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF4	0.97	0.29	0.99	0.97	0.29	0.99
FF4 + 1 deep factor	0.97	0.62	0.98	0.97	0.63	0.98
FF4 + 2 deep factors	0.97	0.74	0.96	0.97	0.49	0.95

Table 7: Predict CMA with Deep factor (6 Fama-French portfolios)

	L = 2			L = 3		
	R^2_{TS}	PE(10^{-5})	R^2_{CS}	R^2_{TS}	PE(10^{-5})	R^2_{CS}
FF4	0.97	0.29	0.99	0.97	0.29	0.99
FF4 + 1 deep factor	0.97	0.62	0.98	0.97	0.63	0.98
FF4 + 2 deep factors	0.97	0.74	0.96	0.97	0.49	0.95

Table 8: Predict CMA with Deep factor (6 Fama-French portfolios)