Neural Beta Estimation on CRSP Stocks

MGT 8803 SC — Assignment 7

Arjo Bhattacharya

1 Introduction

This assignment implements a neural network model to estimate firm–level market betas using monthly CRSP stock data from 2005–2023. Model training and evaluation follow the required workflow: feature engineering with rolling windows, train–validation–test split using fixed date cutoffs, hyperparameter tuning for neural beta estimation, descriptive beta analysis, Fama–French factor augmentation, and portfolio formation.

All analysis was implemented in Python using `pandas`, `scikit-learn`, and `TensorFlow/Keras`.

Data Summary:

- CRSP Monthly Stock (MSF), CRSP VWRETD for market proxy.
- Train window: 2005–2012, Validation: 2013–2017, Test: 2018–2023.
- Stocks sampled: For each year and industry, exactly 10 distinct firms were selected (required in Step 0).
- Industries were classified per SIC code in nine categories (e.g., Construction, Retail, Manufacturing).

The goal is to estimate $\beta_{i,t+1}$ per firm-month, defined as the slope coefficient of $R_{i,t+1}$ on $R_{m,t+1}$, using a neural network trained to minimize squared forecast error:

$$\min_{\theta} \; E\left[\left(R_{i,t+1} - \hat{\beta}_{i,t+1}(\theta) \cdot R_{m,t+1}\right)^2\right].$$

2 Feature Engineering and Data Split

From the sampled panel, we construct rolling window features using $w = 12, 24, 36$ months:

- Stock return stats: mean, std, min, max, lag-1 autocorr: `ret_mean_w`, `ret_std_w`, ...
- Market return stats: `mkt_mean_w`, `mkt_std_w`
- Rolling OLS beta: `beta_lin_w`
- Log market cap and (if available) log volume moments
- One-hot industry dummies

No look-ahead bias: at time t, features are based only on data up to t. Supervised labels are $R_{i,t+1}$ and MKT_{t+1}.

The final feature matrix has 50 columns and shapes:

Train: (326, 50), Validation: (489, 50), Test: (857, 50).

3 Neural Beta Model and Hyperparameter Tuning

The neural beta model is a feed-forward network with 1 hidden layer and a 1-dimensional linear output:

$$\hat{\beta}_{i,t+1} = f_\theta(X_{i,t}), \quad \theta = \{\text{weights, biases}\}.$$

$$\text{Loss} = \texttt{neural\_beta\_mse} = \frac{1}{N}\sum_i \left(R_{i,t+1} - \hat{\beta}_{i,t+1} \cdot MKT_{t+1}\right)^2.$$

We performed grid search over:

- Hidden units: {16, 32, 64}
- Activation: {ReLU, Tanh, Sigmoid, Linear}
- Learning rate: {1e–3, 5e–4, 1e–4}
- L2 regularization: {0, 1e–5, 1e–4}

Best configuration:

Hidden: 16, Activation: ReLU, LR = 0.0005, L2 = 0.0

Test RMSE for the tuned model: **0.1393**.
Baseline RMSEs for comparison:

- Naive $\beta = 1$: 0.1245
- Rolling 36M OLS beta: 0.1264
- Linear regression beta: 0.2701

Training history: (insert figure)

Figure 1: Validation and training loss for best model.

4 Test-Period Beta Distributions (Step 4)

Table 1 summarizes distributional characteristics of $\hat{\beta}_{i,t+1}$ grouped by industry in the 2018–2023 test window.

Table 1: Summary statistics of $\hat{\beta}$ by industry (test period).

Industry	Count	Mean	Std	Skew	Kurt	Min / Max
Agriculture	463	1.03	0.24	...	...	...
Construction	245	1.29	0.14	...	...	...
Retail	36	1.08	0.09	...	...	...
Transport/Utilities	23	1.03	0.11	...	...	...
Wholesale	69	0.97	0.17	...	...	...
...	...	...	...	...	...	...

5 Industry Beta Dynamics (Step 5)

For each test-period year, we compute the cross-sectional mean and standard deviation of $\hat{\beta}$ within each industry. Figure 2 displays the mean curves.

Figure 2: Mean neural beta by industry, 2018–2023.

Findings:

- Construction and Agriculture exhibit highest average betas $\approx 1.2 - 1.3$.
- Retail and Manufacturing maintain values near $\beta \approx 1$.
- Dispersion narrows in 2020–2021, likely reflecting market-wide COVID dynamics.

6 Fama–French Factor Augmentation (Step 7–8)

We augment the neural model by joining monthly Fama–French 5 factors (Mkt–RF, SMB, HML, RMW, CMA, RF), and computing 12M/24M/36M rolling means and stds for each factor. Retraining the model with these extra inputs slightly improves test RMSE:

RMSE (base model): 0.1393
RMSE (FF-augmented): 0.1368

β distributions also shift downward slightly, indicating that residual variation is captured by SMB/HML/RMW/CMA exposures.

Interpretation: The FF-augmented model decomposes part of the return variation that the single-factor model attributes to β. Lower RMSE suggests incremental explanatory power of size, value, profitability, and investment style factors.

7 Quintile Portfolios (Step 9)

Each month in the test window, stocks are sorted into 5 quintiles based on $\hat{\beta}_{i,t+1}$. We form both equal-weight (EW) and value-weighted (VW) portfolios.

Average Beta by Quintile

Table 2: Average $\hat{\beta}$ per quintile (test period).

	Q1	Q2	Q3	Q4	Q5
EW	0.66	0.94	1.12	1.32	1.52
VW	0.70	0.95	1.10	1.30	1.49

Return Spreads

Annualized mean spreads ($Q5 - Q1$):

- EW: 0.14

- VW: 0.09

The positive spreads confirm a mild beta premium over the 2018–2023 window, though results are sensitive to market conditions.

8 OLS Comparison (Step 10)

We repeat the quintile sort using rolling OLS beta (36-month window). The OLS-based Q5–Q1 annualized spread is lower:

Method	Q5–Q1 (EW)
Neural $\hat{\beta}$	0.14
OLS β^{OLS}	0.11

The neural model produces more extreme betas, and therefore a slightly stronger return differential.

9 Conclusion

This assignment demonstrates a scalable method for estimating firm-market betas with a neural network trained using rolling panel features. The model produces reasonable $\hat{\beta}$s consistent with industry intuition, and a mild beta premium in cross-sectional sorts. Augmenting the feature set with Fama–French factors achieves modest reductions in forecast RMSE.

Future extensions may include recurrent models, direct modeling of returns, and exploring additional macro predictors.