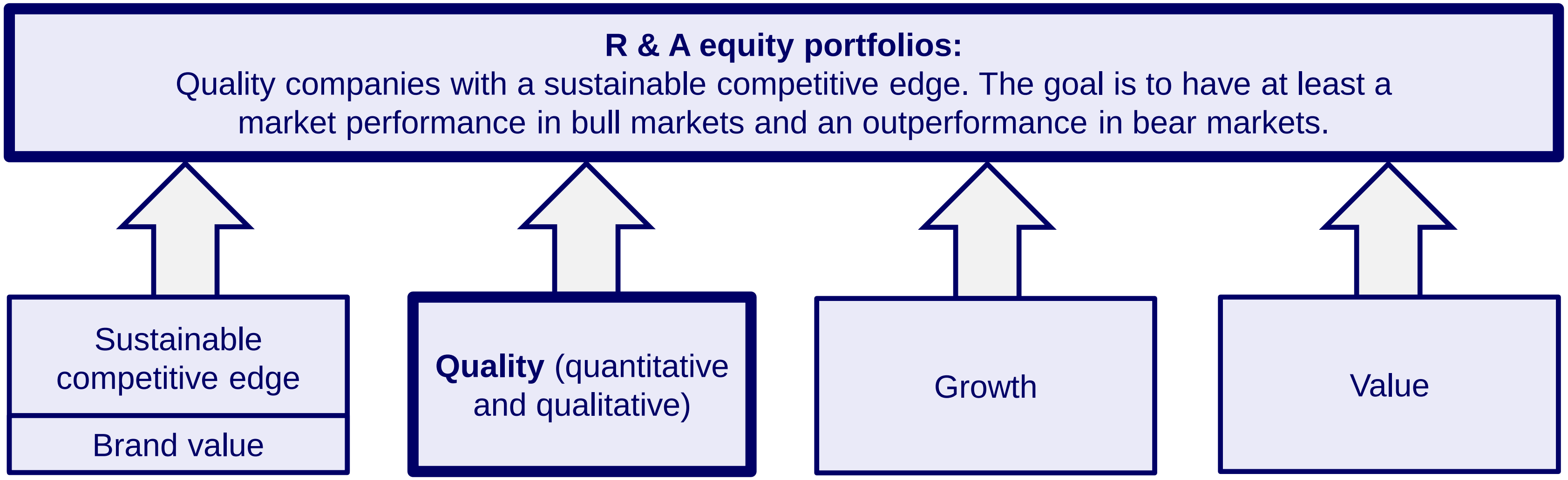
Stock selection with a neural network: Focus on bear market outperformance

© R & A Group Research & Asset Management AG · Bodmerstrasse 3 · CH-8002 Zürich
Phone +41-44-201 07 20 · e-Mail info@ragroup.ch · www.ragroup.ch

Introduction: Stock selection approach



General objective: We aim to **replicate our stock selection process using machine learning techniques.** We tell the machines how we select stocks (and then see what they come up with, even though we expect a significant overlap with our existing portfolios). We use a modular approach, modelling the pillars separately: Competitive edge, brand value, quality, growth, value.

This presentation: Model the input factor **"Quality**" with the aim to create portfolios which have the desired performance attributes. Hence, we **train a neural network to select stocks, with a focus on outperformance in bear markets and at least a market performance in bull markets.**

R & A GROUP
RESEARCH & ASSET MANAGEMENT

Agenda

1. Translate the study objective into a data requirement

- What is the output of the neural network (= the dependent variable(s)), and
- which input variables (explanatory variables) do we need/have.

2. Neural network used in this study

- Basic characteristics of the neural network
- Two-step-approach
  - Step 1: Train a Base neural network to select outperforming stocks.
  - **Step 2: Adjust the Base neural network to improve its bear-market performance.**

1. Data

To select stocks: We estimate 1-month forward returns of individual stocks. If the model works: The stocks with the highest model-estimate, e.g. the top 20%, should outperform on average.

Universe of companies: Members of the **Stoxx Europe 600 and the S&P 500** (non-financial companies).

Monthly data for the years 2005 to 2020 (= 181 months).

To estimate/explain forward returns **we need company data:**

1. **Financial statement data:** Basic financial ratios, such as the net-debt-to-equity ratio and operating margin, and their year-on-year change.
2. **Fama-French-Carhart factors:** Beta, the price-to-book ratio, market capitalization and momentum variables.
3. Categorical variables of sector membership (e.g. technology, consumer staples, etc.).

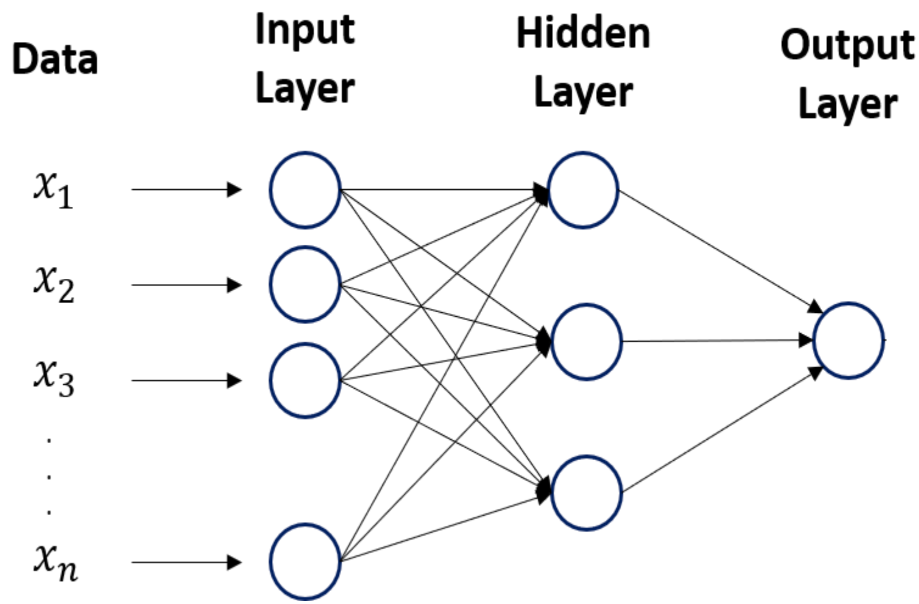
Size of the data set: 557 companies have a full data history multiplied by 181 months yields over **100,000 samples/records** to train and test the neural network (panel data is interpreted as cross section data, which is a typical set-up in neural network equity models, see the reference list).

2. Approach: Feed-forward neural network

Feed-forward neural network with 1 hidden layer

1. Output Layer: 1 variable/nod = equity forward returns.
2. Input layer: 30 variables/nods (n=30 in the graph, i.e. the variables explaining forward returns).
3. Hidden layer: 30 nods.
4. Neural network: Calculation/specifications
   - Python/Pytorch neural network functionalities.
   - Google Colab environment.
   - Loss function: MSE (mean squared error).
   - Activation functions: Softplus and Tanh.

Now, we have the data and the network structure:
→ let's train the neural network (let the optimiser do its job, i.e. how to best forecast 1-month forward returns) and let's look at the results.



STEP 1: The Base neural network

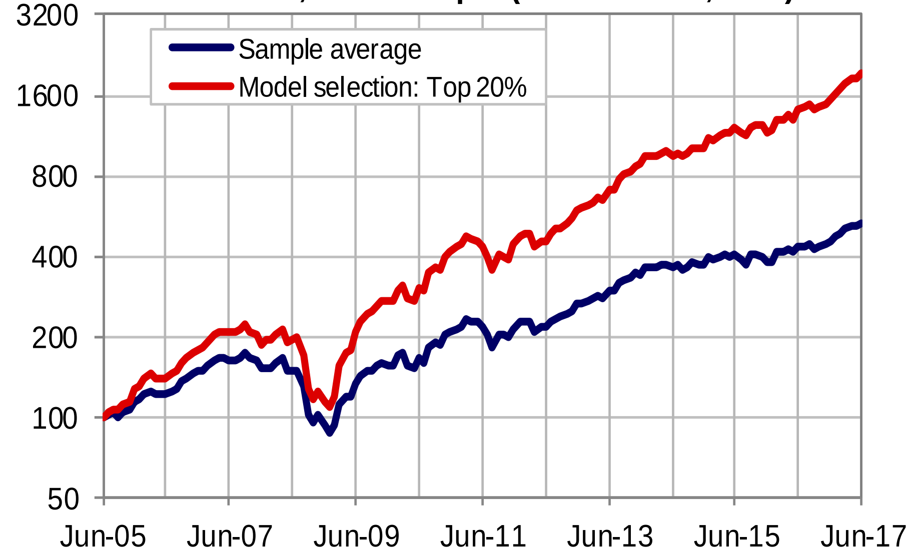
Visualising the model performance

1. Take the 20% stocks with the highest 1-month return forecast provided by the neural network (Top 20%) and calculate their average total return (in US dollar).
2. Similarly, calculate the average return of all the stocks in the sample for each monthly cohort, with this sample average serving as the benchmark.
3. Chain the monthly returns to an index (June 2005 = 100).

→ Conclusion: **The Top 20% outperform (red line vs. blue line in the chart).**

→ **But how are we doing in bear markets?**

Base NN, train sample (return index, USD)

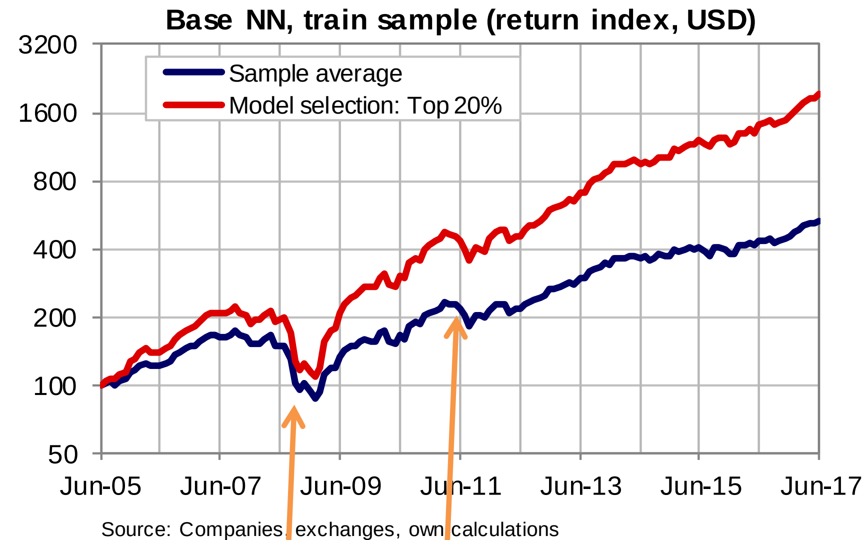


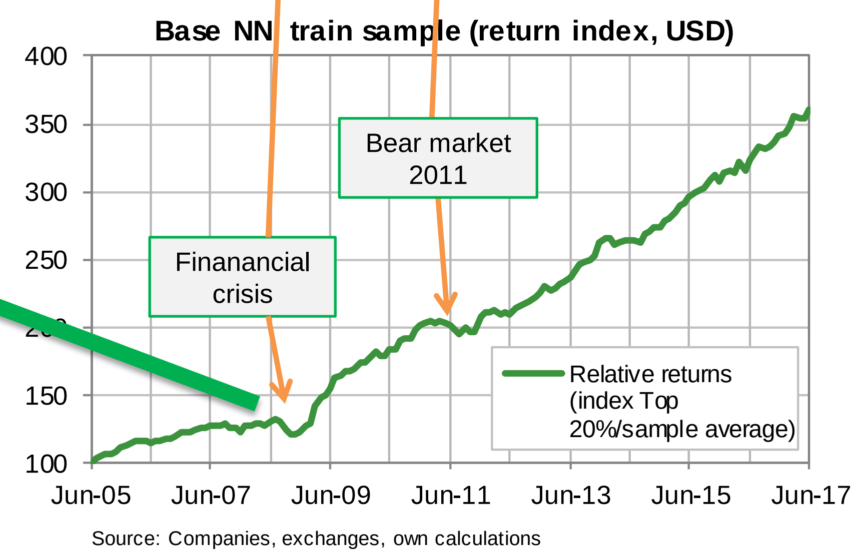
Source: Companies, exchanges, own calculations

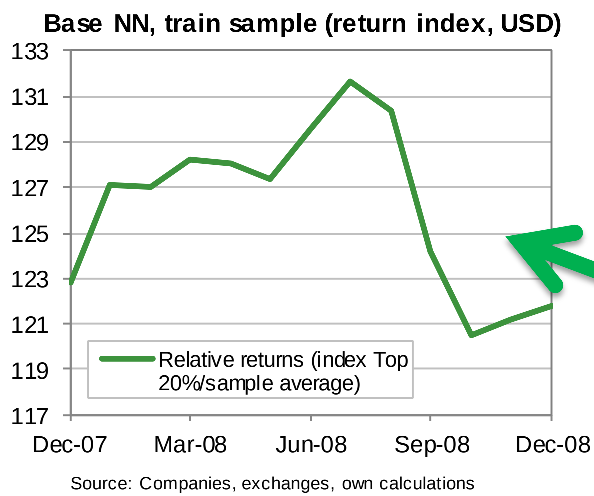
The Base neural network in bear markets

Bottom chart: Relative return index = Top 20%/index of sample average returns, i.e. the red line divided by the blue line of the top chart.

→ **The Top 20% underperform during the financial crisis** (Lehman bankruptcy Sept. 2008) **and the bear-market of early summer 2011.**







R & A GROUP
RESEARCH & ASSET MANAGEMENT

STEP 2: Improving bear-market performance

How can we train our neural network such that its bear-market performance improves?

Solution: **Reward desired outcomes and penalise undesired outcomes,** i.e. adjust the loss calculation**,** with the "loss" the key variable in the training process of the neural network.

Loss = ModelEstimate – TrueReturn

ModelEstimate = 1-month forward-return estimate (output of the neural network)

Adjusted loss: Conditional (linear) transformation

$\mathbf{Loss_{adj}}$ = ModelEstimate – $\mathbf{k_{cond}}$*TrueReturn

$\mathbf{k_{cond}}$ **> 1 whenever the average sample forward return is negative** (otherwise $k_{cond} = 1$).

The higher k is the more weight negative return periods get in the optimisation process. Starting point for k: the ratio of positive to negative return periods.

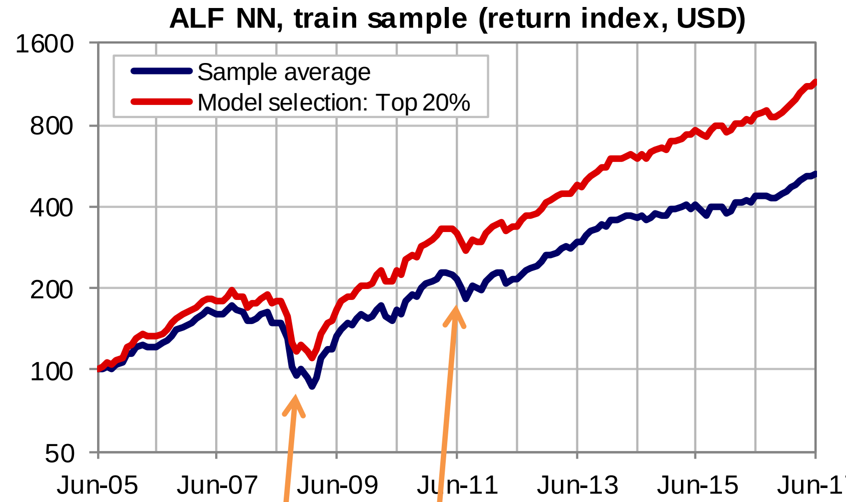
Variations to the adjustment: We know that companies with above-average financial ratios (notably high profit margins, low short-term debt) outperform in bear markets. **Add another adjustment for companies with good financial ratios** ($Loss_{adj}$ = ModelEstimate – [k_{cond}*TrueReturn + y_{cond}*f(FinancialRatios)]).

STEP 2: Adjusted-Loss-Function NN

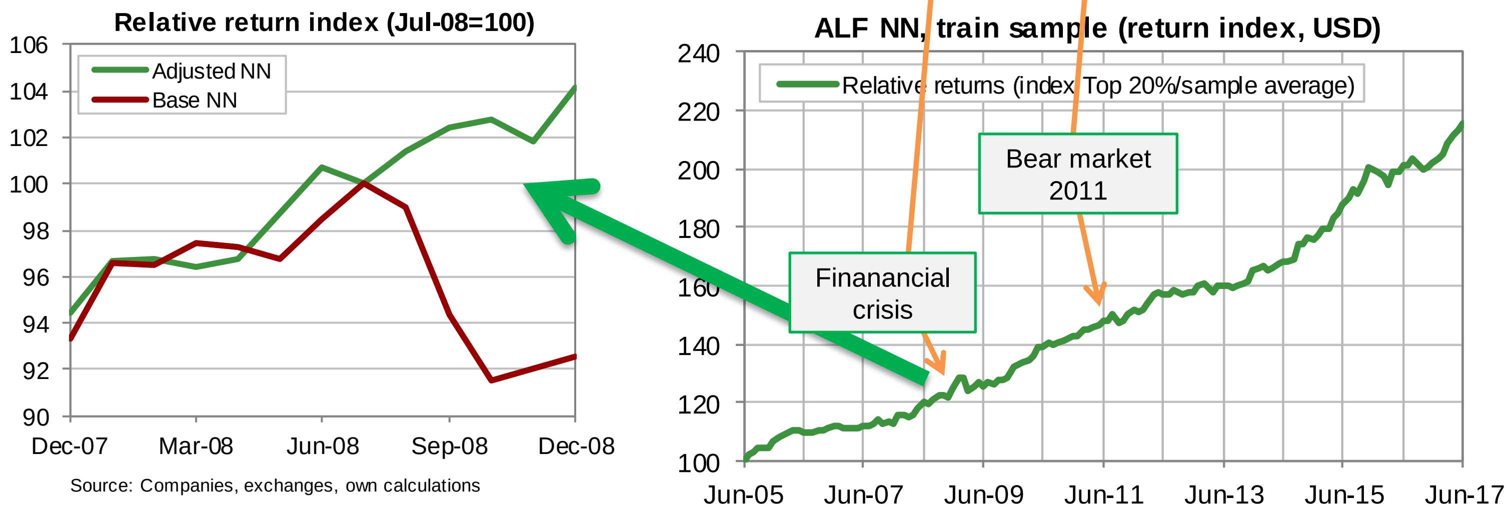
Adjusted-Loss-Function neural network (ALF NN): The only difference to the Base NN: The loss function is adjusted as described.

ALF NN Results: The chart of the ALF NN looks **similar** to the Base NN, i.e. a steady outperformance of the Top 20%. However: **Bear-market performance has clearly improved.**

→ Conclusion: **We are able to tweak a simple feed-forward neural network towards desired return characteristics by adjusting the loss function.**



Source: Companies, exchanges, own calculations



Source: Companies, exchanges, own calculations

R & A GROUP
RESEARCH & ASSET MANAGEMENT

The final test: Out-of-sample simulation

We apply the parameters obtained using the train data (2005-2017) to the test/out-of-sample data set (2017-2020) for both networks.

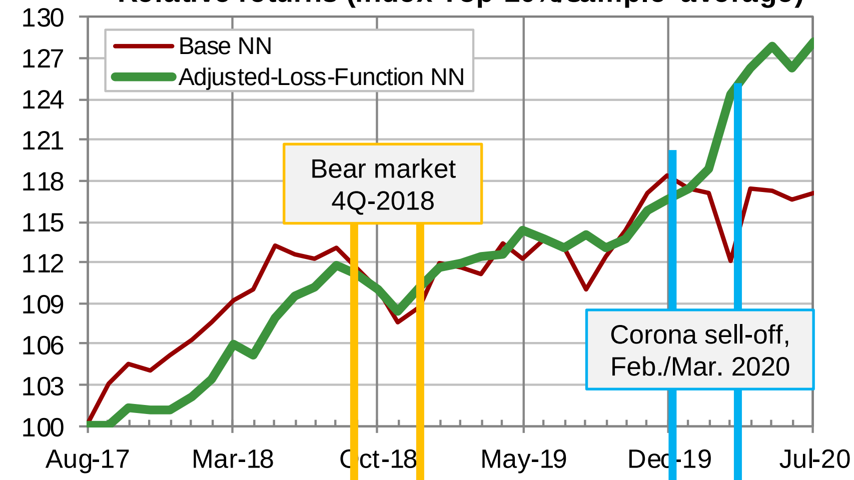
ALF: Much better return and risk characteristics than the Base NN (see the top chart and the table), with the ALF outperforming in the **Covid-19 stock market sell-off** (Feb./March 2020). Part of the explanation: The Base NN trains for a high beta (which is no surprise, given bull market prevalence in the train sample 2005 to 2017).

Base and ALF neural network (NN): Test data set, 2017 to 2020

	Base NN Top 20%	ALF NN Top 20%	All stocks
Alpha (annualised)	4.3%	9.4%	0.0%
Beta	1.22	0.95	1.00
Monthly returns			
CAGR	14.1%	17.6%	8.2%
Standard deviation (annualised)	23.0%	17.9%	18.4%
Maximum drawdown	-28.5%	-19.4%	-24.5%
Negative-period characteristics			
Mean when sample return negative (ann.)	-53.6%	-41.6%	-49.6%
Share of negative periods	34.3%	31.4%	28.6%

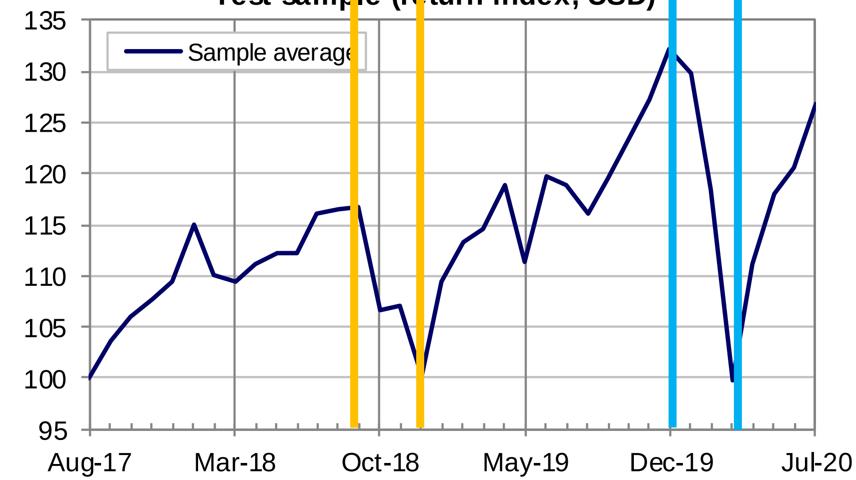
Source: Companies, exchanges, own calculations

Relative returns (index Top 20%/sample average)



Source: Companies, exchanges, own calculations

Test sample (return index, USD)



Source: Companies, exchanges, own calculations

R & A GROUP
RESEARCH & ASSET MANAGEMENT

3. Summary and conclusion

Step 1: Train a Base neural network to select outperforming stocks

Fundamental company characteristics and Fama-French-Carhart factors help to systematically identify outperforming stocks. This is a well-documented finding in the literature.

Step 2: Tweak the Base neural network by rewarding outperforming and penalising underperforming stocks in negative market environments (adjustment of the loss):

- “Tweak” via a simple correction of the loss calculation.
- The “tweak” works, and surprisingly well, both in-sample (train data set) and out-of-sample (test data set 2017-2020), **especially also in the corona-crisis in the spring of 2020.**
- The approach can be applied to tweak neural networks towards other properties as well, e.g. **high-quality attributes**, such as a focus on high profit margins and sound balance sheets.

Conclusion

Training models towards specific return characteristics:

- A way to overcome issues arising from the idiosyncrasies of data sets.
- Allows to incorporate a number of objectives (as long as we can identify these quantitatively) into a neural network’s stock selection. Examples: Investment style, factors, ESG criteria or a top-down view.

References

Arias Chao, Manuel, Bryan T. Adey and Olga Fink: “Knowledge-Induced Learning with Adaptive Sampling Variational Autoencoders for Open Set Fault Diagnostics”. December 2019. https://doi.org/10.3929/ethz-b-000430653

Huang, Yuxuan, Luiz Fernando Carpetz and Danny Ho, “Neural Network Models for Stock Selection Based on Fundamental Analysis”, May 2019. https://arxiv.org/abs/1906.05327.

Rasekhschaffe Keywan Christian and Robert C. Jones (2019), Machine Learning for Stock Selection, Financial Analysts Journal, 75:3, 70-88, DOI: 10.1080/0015198X.2019.1596678.

Annex: Return characteristics

The Base neural network (NN) trains for a high beta of 1.23, whereas the ALF NN trains for a moderate beta of 0.92 (see the top table).

Given its high beta, the standard deviation of the Base NN is higher than that of the sample (“All stocks” in the table), whereas the ALF NN has both a higher return and a lower standard deviation than the sample. This property makes the ALF NN unequivocally superior to the sample in terms of risk and return.

Base and ALF neural network (NN): Train sample, 2005 to 2017

	Base NN Top 20%	ALF NN Top 20%	All stocks
Alpha (annualised)	8.5%	7.8%	0.0%
Beta	1.23	0.92	1.00
Monthly returns			
CAGR	27.7%	22.4%	14.8%
Standard deviation (annualised)	21.6%	16.1%	17.1%
Maximum drawdown	-51.1%	-44.0%	-50.1%
Negative-period characteristics			
Mean when sample return negative (ann.)	-37.3%	-30.0%	-37.2%
Share of negative periods	30.6%	27.8%	33.3%

Source: Companies, exchanges, own calculations

Base and ALF neural network (NN): Test data set, 2017 to 2020

	Base NN Top 20%	ALF NN Top 20%	All stocks
Alpha (annualised)	4.3%	9.4%	0.0%
Beta	1.22	0.95	1.00
Monthly returns			
CAGR	14.1%	17.6%	8.2%
Standard deviation (annualised)	23.0%	17.9%	18.4%
Maximum drawdown	-28.5%	-19.4%	-24.5%
Negative-period characteristics			
Mean when sample return negative (ann.)	-53.6%	-41.6%	-49.6%
Share of negative periods	34.3%	31.4%	28.6%

Source: Companies, exchanges, own calculations

Annex: Structure of the data base

The table visualises the data set-up.

Treat the panel data as a single cross section (which is a typical set-up in neural network equity models, see the reference list).

Structure of the data set: "Cross-section view" of panel data, 557 companies, monthly data 2005-2020

Obser-			Input variables/features of the neural network						Output variable
vation	Month	Company	Variable 1	Variable 2	....	Variable 25	....	Variable 30	neural network
1	**Jun-2005**	Company 1	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
2	Jun-2005	Company 2	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
3	Jun-2005	Company 3	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
...	...	...	...	...	...	...	...	...	...
556	Jun-2005	Company 556	Net debt/assets	RoIC		Market capital.	...	...	1-month-fwd-ret.
557	Jun-2005	Company 557	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
558	**Jul-2005**	Company 1	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
559	Jul-2005	Company 2	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
560	Jul-2005	Company 3	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
...	...	...	...	...	...	...	...	...	...
1113	Jul-2005	Company 556	Net debt/assets	RoIC		Market capital.	...	...	1-month-fwd-ret.
1114	Jul-2005	Company 557	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
100261	**Jun-2020**	Company 1	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
100262	Jun-2020	Company 2	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
100263	Jun-2020	Company 3	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.
...	...	...	...	...	...	...	...	...	...
100816	Jun-2020	Company 556	Net debt/assets	RoIC		Market capital.	...	...	1-month-fwd-ret.
100817	Jun-2020	Company 557	Net debt/assets	RoIC	...	Market capital.	...	...	1-month-fwd-ret.

Note: See the text for further details.

R & A Group is an independent investment research and asset management company. R & A Group has a comprehensive investment research offering with a focus on investment strategies and global stock selection.

Founded in 2001, R A & Group is a joint stock company domiciled in Zurich and a member of the leading independent asset manager organization VSV (Swiss Association of Asset Managers).

Investment-
Research at Work ™

Disclaimer: We do not make any guaranties for the validity of the information herein even though the information is based on public sources believed to be reliable. We are independent and do not have business relationships with any company possibly mentioned in this report. Opinions, estimates and projections in this report reflect our judgment at the date of writing and are subject to change without notice. We have no obligation to update, modify or amend this report or to notify a reader hereof if any matter, opinion, estimate or projection subsequently changes or becomes inaccurate. This report is provided for information purposes only and does not constitute an offer to buy or sell company stock or related securities or take any other investment decision. This report must not be made available to US or UK residents or to any other person to whom governing law may prohibit distribution of this report.

Data sources in charts or tables that are not specifically mentioned are either companies or index provider.

R & A GROUP
RESEARCH & ASSET MANAGEMENT