Quantitative Trading with Machine Learning

René M. Glawion
rglawion@stanford.edu
Department of Computer Science, Stanford University

Stanford | ENGINEERING
Computer Science

Executive summary

We explore the application of Machine Learning for predicting the return of the VW stock by using the information of stock returns in its supply chain. Starting a horse race between Elastic Nets, Decision Trees, XGBoost, and LightGBM, we find that Elastic Nets generates the highest prediction accuracy across forecast horizons. A trading strategy based on this analysis leads to increased trading profits up to three times compared with a simple buy and hold strategy.

Dataset and features

We identified 70 companies in the supply chain of Volkswagen AG from the Volkswagen page [5] and several other sources for news about the supply chain of VW (e.g., Patzer [4]) and the VW Group Award for the best suppliers (e.g., Eisert [2]). Of those 70 companies, 36 are traded on a public stock exchange, and 30 are traded at the same stock exchange in Frankfurt. For those publicly traded firms, we downloaded the time series for the time period from 01-01-2005 to 01-05-2021 with the help of the yahoo finance API. Those time series include the ticker, open, high, low, and closing prices, as well as the stock exchange, currency, and other stock-specific data. We transform all level data to return data and use those as features. Figure 1 plots the trajectory of the VW stock over our observation period.

Additionally, to capture the shape of the overall economy, we collected a set of macroeconomic variables. Namely, the output gap, CPI, and the yield of the 10-year government bond of Germany, where the headquarter of VW is located. Further, to ascertain the overall structure of the yield curve, we included the 3-month Euribor rate and the 10-year Euirs rate as well as 1-month and 6-months relative differences of those. Last, we included the iTraxx Crossover together with 1- and 6-month relative differences to model changes in credit risk in the Euro Area. We obtained all time series from Refinitiv Eikon.

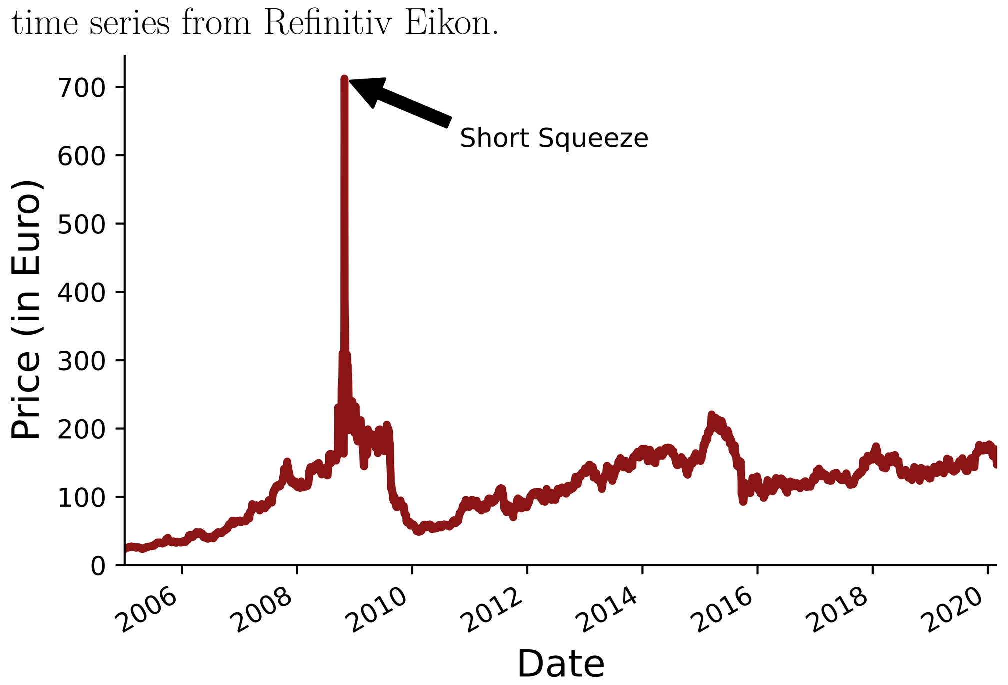

Figure 1: Trajectory of the VW Stock
The evolution of the VW stock price during our observation period from 01-01-2005 to 01-05-2020. The peak in 2008 was invoked by a failed takeover attempt of VW by Porsche [1].

Model evaluation criteria

The model we use for our analysis reads

$$\mathbb{E}[y_{t,t+h}|\mathcal{F}_t] \approx g(X_t, \theta),$$

where $y_{t,t+h} \in \mathbb{R}$ denotes the return of the VW stock over the period t to $t+h$, $h \in \{1, 5, 10, 15, 20, 25\}$, $\theta \in \mathbb{R}^{p+1}$ is a vector of weights (hyperparameters) we want to choose optimally in the sense of some given metric, $X_t^{(i)} \in \mathcal{F}_t$ is the matrix of features, which are the daily returns of all companies in the supply chain of VW over the past 28 days, $\mathcal{F}_t$ is the filtration. In our dataset we have $p = 850$ features. The function $g(\cdot)$ depends on the machine learning approach.

The model is a regression problem. Thus, we decided to use the root mean squared error (RMSE) metric to fit the models, which is defined by

$$RMSE(y_{t,t+h}, \hat{y}_{t,t+h}) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_{t,t+h}^{(i)} - \hat{y}_{t,t+h}^{(i)}\right)^2},$$

where y denotes the realized stock return of VW, $\hat{y}$ is the predicted return, and T is the number of samples used for training, validation, and testing.

Splitting the dataset

To mitigate the risk of overfitting, we divide the dataset into three subsamples. Regularization is controlled by the tuning of the aforementioned hyperparameters. First, using a given combination of hyperparameters, the parameter vector θ^T is estimated on the training sample. Second, the model performance gets evaluated, in terms of forecast RMSE, on the validation sample. A search across provided hyperparameter combinations points to the specification that delivers the lowest error on the validation sample. After that, we use a test sample that the model has never seen before to ultimately determine the performance in terms of the RMSE and several additional metrics.

Further, we do not split the data randomly but rather sequentially to account for our data's time series nature. In the beginning, we start with a five-year training period which, a one-year validation period and a one-month test period. Afterward, we roll over this window, including the test set into our training and validation set, and start with a fresh test month. Figure 2 visualizes this procedure.

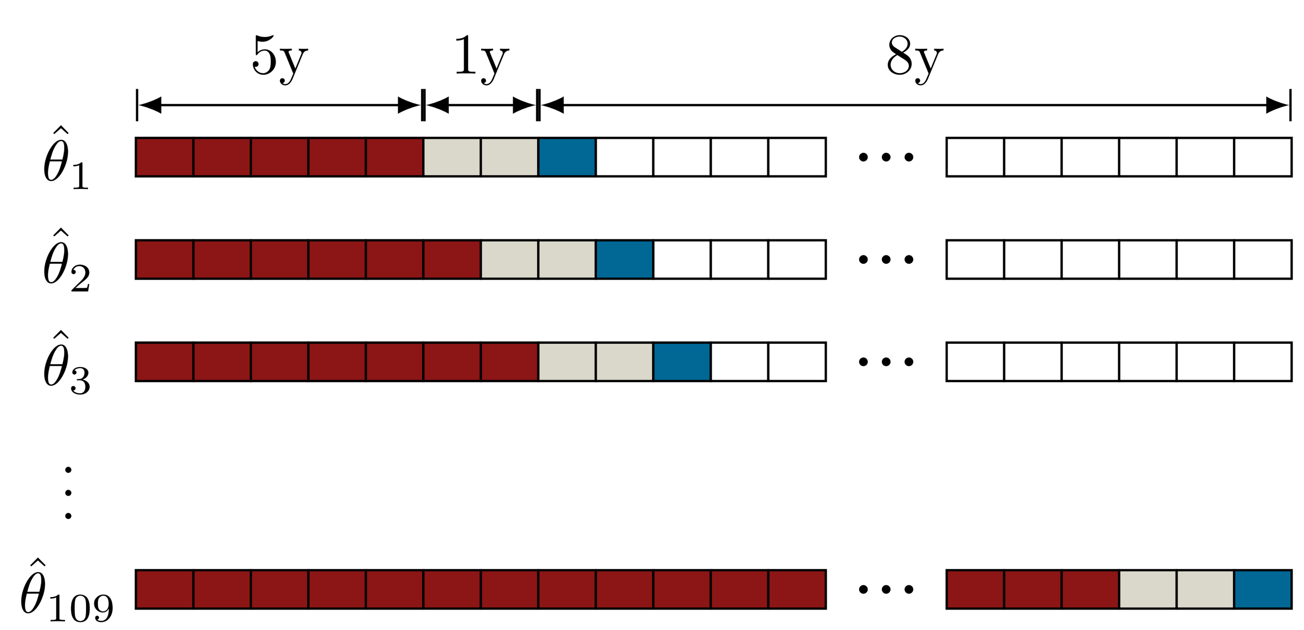

Figure 2: Training, validation and testing procedure
The data ranges from January 2005 to February 2020. The training period (red) initially spans 5 years and increases by one month after each validation step. Each of the 109 validation steps delivers a new set of parameter estimates. Each validation window (westar) covers 1 year and is rolled forward. After each validation step, there is one month of out-of-sample testing (blue). Note that the figure is for illustrative purposes only and not drawn to scale.

Results

Table 1: Prediction performances on the testset
The table on the left-hand side presents the root mean squared error, and the table on the right-hand side the estimation time for each model.

model \ h	1	5	20	model \ h	1	5	20
Elastic Net	.0174	.0401	.0751	Elastic Net	4.62	4.88	4.92
Decision Trees	.0190	.0422	.0952	Decision Trees	11.37	7.41	4.15
XGBoost	.0176	.0410	.0850	XGBoost	92.41	107.03	109.67
LightGBM	.0177	.0408	.0815	LightGBM	44.81	47.70	53.93

For all four models, the error metrics increase with increasing forecast horizon h. For example, using the Elastic Net the RMSE increases from 1.74% for $h = 1$ to 7.51% for $h = 20$ days. Second, for all forecast horizons, the Elastic Net results in the lowest RMSE on the test set, beating LightGBM with a tight margin.

Turning knowledge into profit

We set up our strategy as follows: We start with a bankroll of 1,000 EUR. If we do not own stock, we buy one position if our forecast predicts that the stock price will increase by more than the predetermined threshold c over the period h. Contrastingly, if we predict a decrease in the stock price below the threshold over the period h, we sell the stock and go short one position. If we predict an increase, we close out the short position and go long again. Here, we abstract from transaction costs. Thus, bid and ask prices are the same, which in our case is the closing price of each day at the Frankfurt Stock Exchange.

Table 2: The tables present the trading profit in percent for each model and forecast horizon. In brackets, we provide the number of trades for each strategy. For comparison, a buy and hold strategy would have resulted in a profit of 7.98%.

	Elastic Net				Decision Tree Regressor		
c \ h	1	5	20	c \ h	1	5	20
0%	-9.07	9.56	21.49	0%	-8.02	-3.56	7.97
	(283)	(127)	(59)		(319)	(237)	(399)
.25%	-6.33	-2.22	14.38	.25%	-5.66	-12.39	4.85
	(61)	(61)	(45)		(205)	(233)	(395)
.5%	-8.69	-5.24	13.93	.5%	-5.69	-7.85	7.10
	(25)	(53)	(57)		(165)	(217)	(395)
1%	-7.83	-7.20	-5.28	1%	-5.43	-7.33	3.11
	(9)	(13)	(25)		(125)	(201)	(349)

	XGBoost				LightGBM		
c \ h	1	5	20	c \ h	1	5	20
0%	0.63	-8.83	6.57	0%	9.85	5.39	2.91
	(191)	(439)	(459)		(1121)	(687)	(471)
.25%	1.41	-10.87	7.63	.25%	-5.09	-10.95	6.67
	(165)	(399)	(459)		(409)	(619)	(447)
.5%	-2.40	-9.54	0.75	.5%	-7.54	-3.27	17.04
	(85)	(339)	(467)		(169)	(517)	(439)
1%	-6.22	-11.72	4.37	1%	-5.98	-6.20	10.34
	(41)	(303)	(451)		(77)	(369)	(439)

Conclusion

This study provides evidence that trading strategies based on machine learning models and supply chains can outperform simple buy and hold strategies. The Elastic Net generates a trading profit three times larger than the buy and hold strategy in our best-case scenario. Further, using regression models instead of classification models, we can enhance the trading thresholds and only execute trades with an expected profit above a predetermined threshold.

Future work

Seeing the results, it will be fruitful to combine different ML methods. For example, Guo et al. [3] combine both LightGBM and LSTM models to predict the direction of the Apple stock. They only achieve an accuracy of 54.1%, but it would be interesting to enhance their approach to our regression problem with the information of the whole supply chain. For example, Taiwan Semiconductor Manufacturing Company (TSMC) is a supplier for both Infineon and Nvidia, which are both suppliers of Volkswagen. Motivated by the current capacity limit of TSMC and other chip manufacturers, we will also want to include those in our database. This analysis should help to predict stock price movements of Volkswagen AG itself. Additionally, here we only analyzed the return of a trading strategy. Due to personal risk tolerance and regulations in the banking industry, it will be necessary to measure the ML-based trading strategies' risk in terms of standard deviations, maximum drawdowns, and several other risk measures.

One last disclaimer: before implementing this strategy, one needs to perform a proper robustness analysis. This includes splitting the dataset into different subperiods and test if results carry over, e.g., before the financial crisis and the simultaneous short squeeze, after the financial crisis 2010-2020, and the last five years from 2016-2021. Furthermore, one needs to check if results carry over to the other exchanges. Further, looking at additional price information than closing prices. Maybe the Open, High, Low, and Volume Information also help to predict the stock prices. Finally, one can replicate this analysis to other stocks, for example, by invoking the FactSet Supply Chain Relationships dataset, which includes deeper information on 13 key company relationship types like important customers and also competitors.

References

[1] AP (2008). Squeezing the accelerator. The Economist `https://www.economist.com/business/2008/10/29/squeezing-the-accelerator`.

[2] Eisert, R. (2018). Group award 2018 für 19 top-lieferanten:das sind volkswagens beste zulieferer. `https://www.automobilwoche.de/article/20180524/NACHRICHTEN/180529953/group-award--fuer--top-lieferanten-das-sind-volkswagens-beste-zulieferer`. Accessed: 2021-05-05.

[3] Guo, Y., Li, Y., and Xu, Y. (2021). Study on the application of lstm-lightgbm model in stock rise and fall prediction. In *MATEC Web of Conferences*, volume 336, page 05011. EDP Sciences.

[4] Patzer, K.-H. (2015). 22 deutsche lieferanten dürfen mitspielen:fast, die vw-champions league der zulieferer. `https://www.automobilwoche.de/article/20150810/NACHRICHTEN/150819994/-deutsche-lieferanten-duerfen-mitspielen-fast-die-vw-champions-league-der-zulie` Accessed: 2021-05-05.

[5] Volkswagen AG (2021). Das fast programm ist der weg des volkswagen konzerns, die zusammenarbeit mit seinen wichtigsten lieferanten zu intensivieren. `https://www.vwgroupsupply.com/one-kbp-pub/de/kbp_public/fast_2/basicpage_for_general_pages__html_2.html`. Accessed: 2021-05-05.

Acknowledgements and replication files

We thank Pengda Liu for his help in developing the ideas for this project and helpful discussions throughout this study. We also thank Josie Oetjen and Gabriel Spil for helpful discussions and feedbacks. All replication files are on GitHub https://github.com/rglawion/cs229_project_report.