TECHNISCHE UNIVERSITÄT WIEN

News Sentiment and Equity Returns

Institute of Management Science
Financial Management & Controlling

Ao.Univ.-Prof. Mag. DDr. Thomas Dangl
Dipl. Ing. Stefan Salbrechter

Overview

- We study the influence of financial news on equity returns.
- Therefore we perform a Sentiment Analysis of company related news articles.
- We train a state-of-the art Natural Language Model (BERT) to capture the news sentiment.
- Fundamental Hypothesis:
  - Positive Sentiment is associated with positive future asset returns
  - Negative Sentiment is associated with negative future asset returns
- News sentiment trading strategy
- 100% data driven approach

What we find

- Our sentiment measure shows predictive quality
- It takes up to two days until financial news are fully incorporated into asset prices
- Markets are very efficient

Fama-French Reg. (FF5+MOM)	Long Portfolio	Long/Short Portfolio
Alpha /Month	2.25%	6.46%

Natural Language Processing (NLP)

NLP is a growing subfield of computer science and artificial intelligence that gives computers the ability to process, analyze and understand human language.

Methods for generating features from text

- Dictionary approaches
- Bag-of-words
- Word embeddings (Word2Vec)
- Contextual word embeddings (BERT)

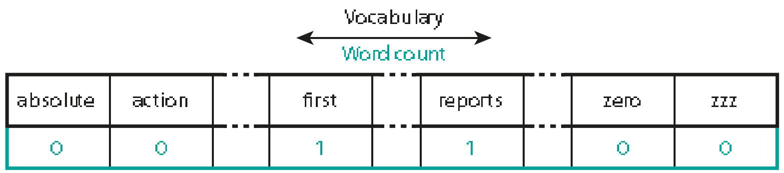


Fig 1: Bag-of-words approach

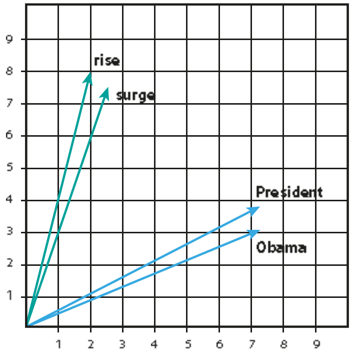


Fig 2: Word embeddings approach

Use cases of NLP

- Machine translation
- Spam detection
- Text summarization
- Virtual agents and chatbots
- <u>Sentiment analysis</u>

Data

News Data

- Thomson Reuters news dataset
- Time period: 1996 to 2020
- Selection of English news, related to S&P 500 companies
- 372,438 cleaned news articles
- We further consider:
  - News novelty („fresh“ vs. „stale“)
  - News topics (analyst forecast, earnings reports)

Price data[1]

- In Total 1130 S&P 500 Companies
- Adjusted Open/Close Prices
- Frequency: Daily

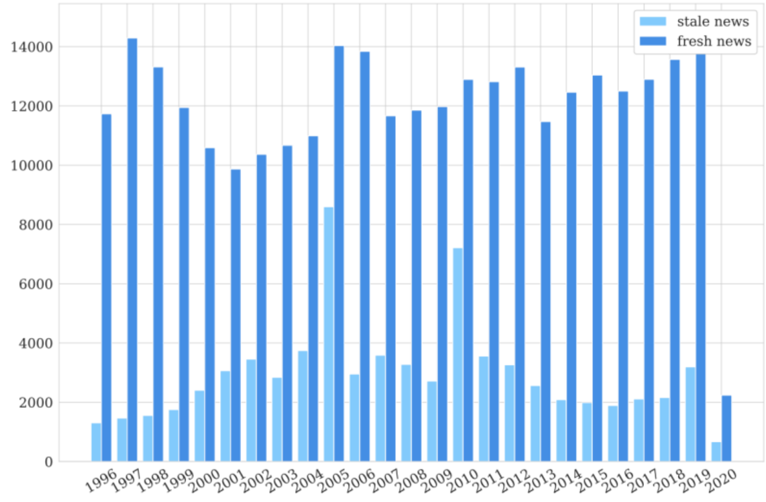


Fig. 3: Annual amount of „fresh“ und „stale“ news

[1] Data source: Thomson Reuters DataStream

BERT Model (**B**idirectional **E**ncoder **R**epresentations for **T**ransformers)

- Developed by Google in 2018
- Pre-trained on a book corpus with 800 million words and the English Wikipedia with 2,500 million words in 2018.[1]
- Uses contextual word embeddings
- BERT is part of the Google search algorithm [2]
- Multiple use cases:
  - Question Answering
  - Named Entity Recognition
  - Text Similarity
  - Classification

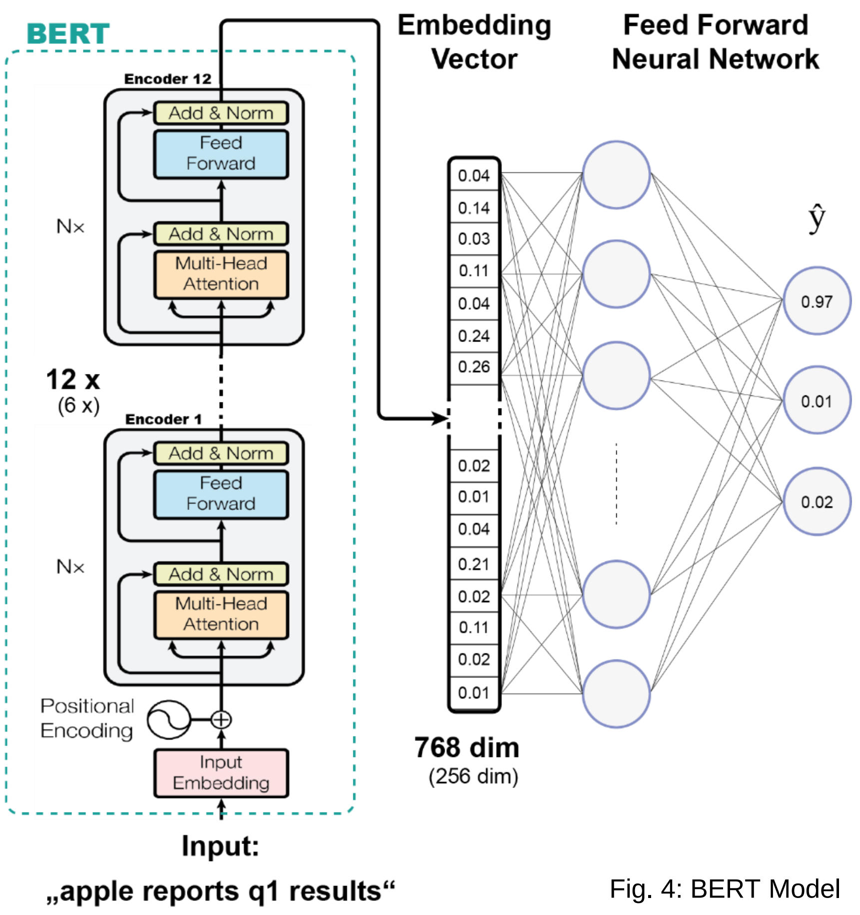


Fig. 4: BERT Model

[1] Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina (2018): BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding
[2] Schwartz, Barry (2020): Google: BERT now used on almost every English query. In: *Search Engine Land*, 15.10.2020.

$\text{BERT}_{\text{MINI}}$ vs. $\text{BERT}_{\text{BASE}}$

	$\text{BERT}_{\text{MINI}}$ [1]	$\text{BERT}_{\text{BASE}}$
Transformer layers	6	12
Hidden embedding size	256	768
Max. Input Sequence	256	512
# Parameters	**18.95 Mio.**	110 Mio.
Pre-training data	Thomson Reuters News (4 GB)	Book corpus & Wikipedia (16 GB) [2]

Advantages of BERT-Mini

- Focus on financial domain
- Less computation power required
- Training cheaper compared to $\text{BERT}_{\text{BASE}}$
- No look-ahead bias

Disadvantages of BERT-Mini

- Limited input sequence length of 256 words

[1] Turc, Iulia; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina (2019): Well-Read Students Learn Better: On the Importance of Pre-training Compact Models.
[2] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach.

Model architecture

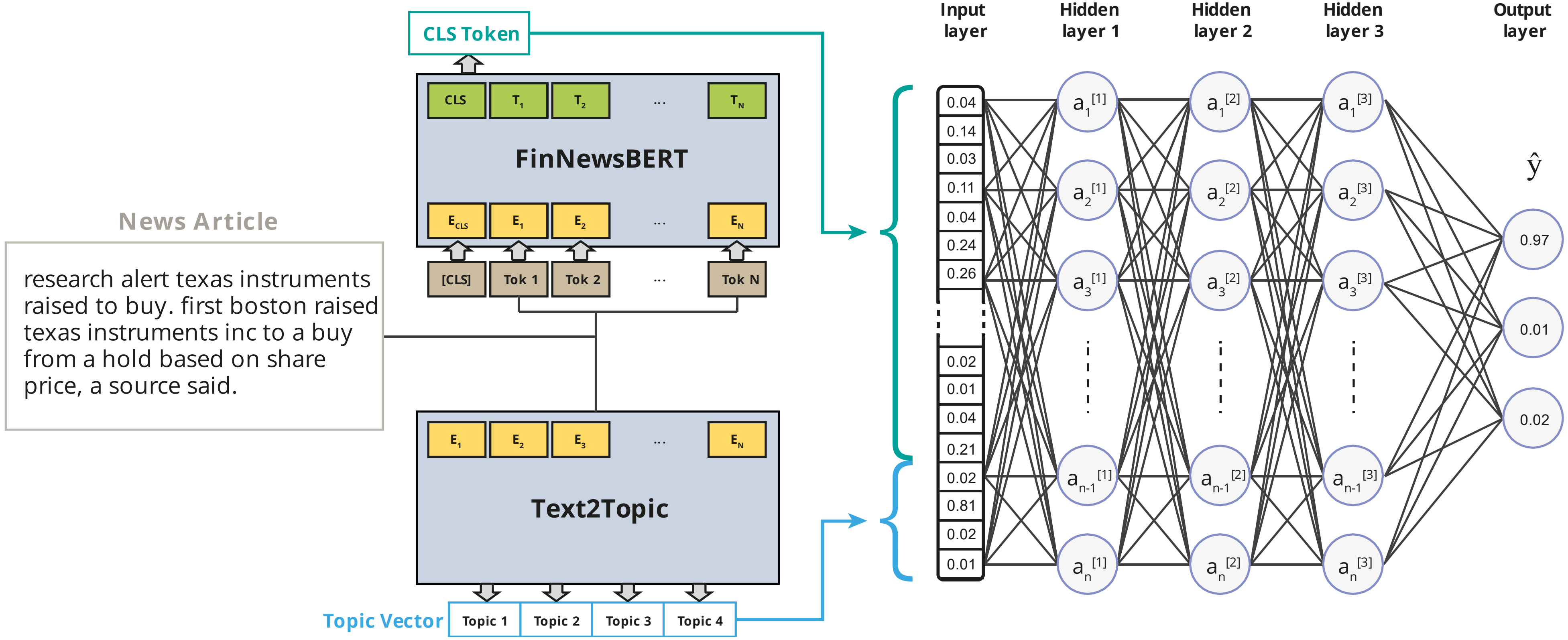


Fig. 5: Model architecture

Annotation of Financial News

- Artificial neural networks are trained on datasets that contain large numbers of example input-output pairs (Supervised Machine Learning).
- We use the joint behaviour of news articles and stock returns to derive the sentiment labels [1]

1. Calculation of daily idiosyncratic returns
2. Calculation of the arithmetic mean values between t-1, t, and t+1
3. Calculation of z-values over a 2-year rolling window
4. Annotation:
   - z-value > 1.4: positive label
   - z-value < -1.4: negative label
   - -1 < z-value < 1: neutral label

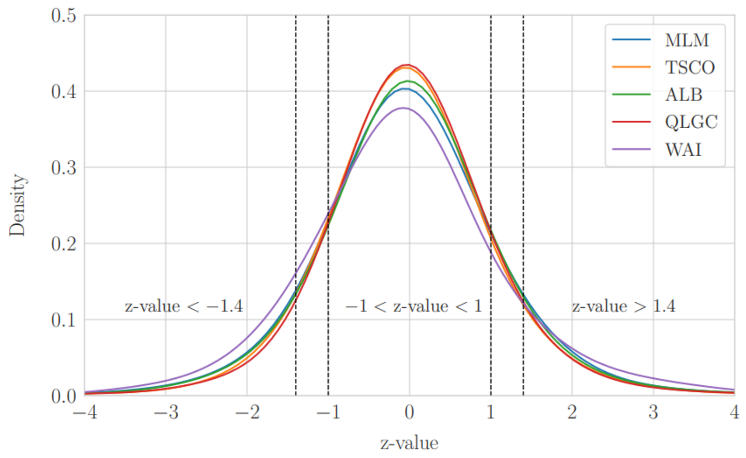


Fig. 6: Density distribution of z-values

[1] Kelly, Bryan T.; Ke, Zheng Tracy; Xiu, Dacheng (2019): Predicting Returns with Text Data (2019-10), S. 54.

Training of FinNewsBERT

1. Pre-training

- Unsupervised learning
- Masked Language Modelling (MLM)

"lincoln national corp deutsche bank starts with <mask> rating price target $84."

Predictions for the <mask> Token	Score
buy	**0.56**
hold	0.24
neutral	0.08
outperform	0.05
sell	0.03

2. Fine-tuning

- Supervised learning
- Fine-tuning on the training dataset with annotated financial news
- Iterative Training
- Training over 3 epochs
- Loss function: Symmetric Cross Entropy Loss [1]

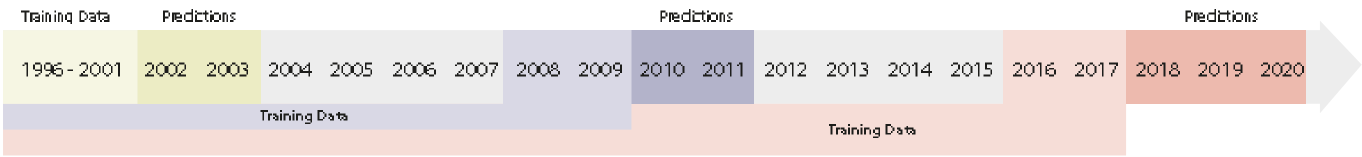


Fig. 7: Iterative pre-training and fine-tuning. The models are re-trained every two years.

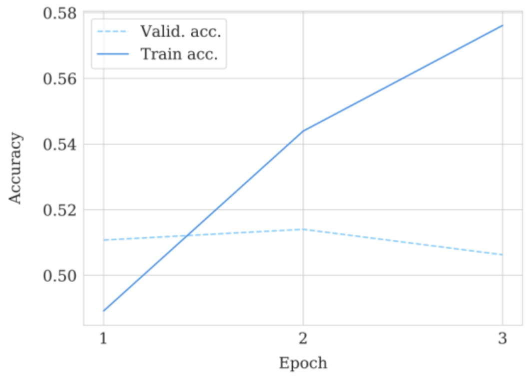
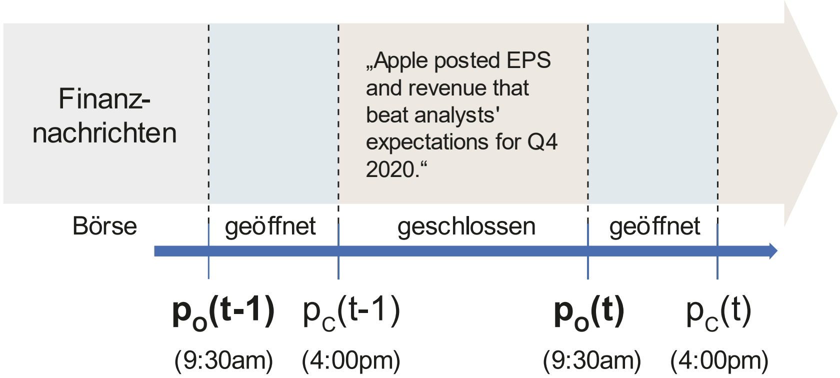


Fig. 8: Training and validation accuracy (The model is trained with data from 1996 to 2017)

[1] Wang, Y., Ma, X., Chen, Z., Luo, Y., Yi, J., and Bailey, J. (2019). Symmetric crossentropy for robust learning with noisy labels.

Event Study

Fig. 9: Abnormal CAPM-Returns (in %) of different sectors.
News are released between market close at 4:00pm on day t-1 and market opening at 9.30am on day t. Time period: 01-2002 to 01-2020.



p_O … Opening price
p_C … Closing price

Forecast of the price movement

Prediction whether the share price of company XY will rise or fall on the following trading day in response to financial news.

- News article classified as positive & r(t+1) > 0 → True Positive
- News article classified as negative & r(t+1) < 0 → True Negative

	All news		Filter: Topic: Analyst Forecast		Filter: Topic: Analyst Forecast & z-value(t) < -1.96	
	Neg. pred.	Pos. pred.	Neg. pred.	Pos. pred.	Neg. pred.	Pos. pred.
Precision	51.40 %	52.03 %	52.36 %	52.78 %	**56.05 %**	52.78 %
Avg. return p.a. (arithm.)	-13.77 %	33.03 %	-31.16 %	47.15 %	**-96.19 %**	47.15 %
Support	61916 (48.44 %)	65878 (51.55 %)	19746 (49.09 %)	20480 (50.91 %)	9680 (48.74 %)	10179 (51.26 %)

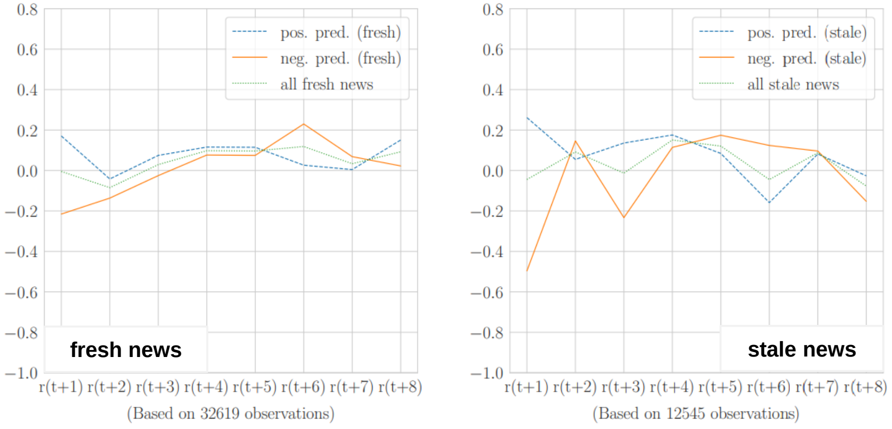
Table 1: Precision and average returns of the predictions at time t+1

Fresh & stale news published during market closing hours, 4:00pm (day t-1) to 9:30am (day t)

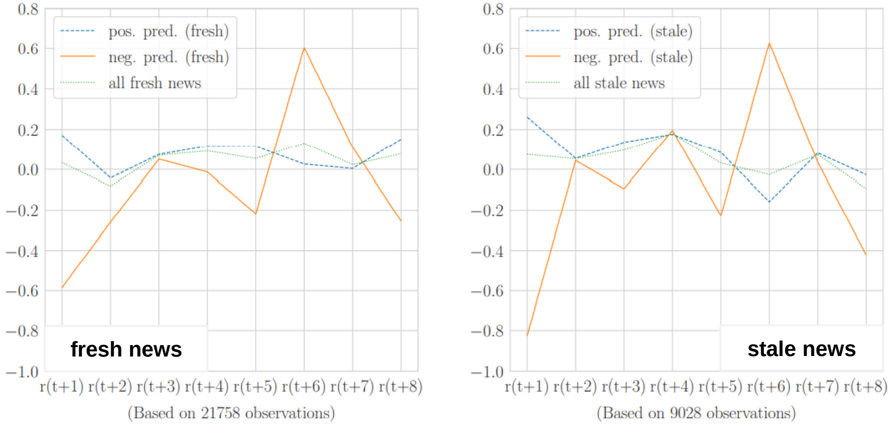
Short-term Momentum Effect? (1)

- Stale news show larger abnormal returns than fresh news at time t+1
- Strongly negative returns at time t (z-value < -1.96) tend to be followed by large negative returns at time t+1 (and also at t+2 in the case of fresh news)
- Tetlock, P. C. (2011). Observes overreactions of market participants to stale news.

Fig. 10: Daily abnormal returns following a news event



Filter: z-value(t, neg.news) < -1.96



[1] Tetlock, P. C. (2011). All the news that’s fit to reprint: Do investors react to stale information? The Review of Financial Studies, 24(5):1481–1512

Short-term Momentum Effect? (2)

- This effect can only be observed in combination with negative news
- The effect can always be observed except for the period 2008-2009 (financial crisis).
- Without taking the news into account, a short term reversal can be observed in the markets.

→ Effect is not driven by recessions

Daily abnormal returns in multiple subperiods following a news event

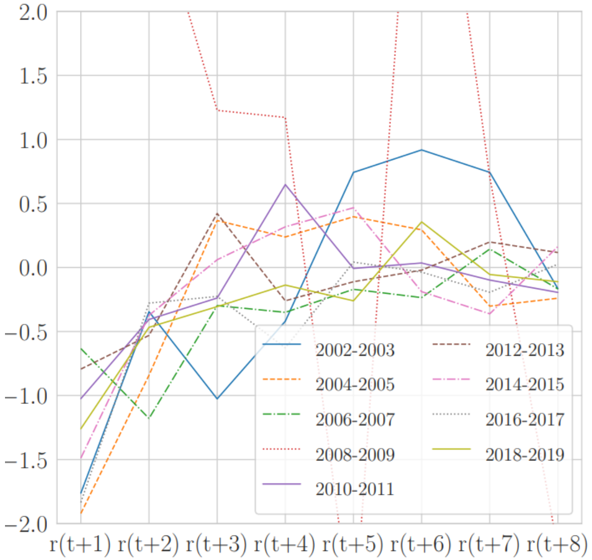


Fig. 11: fresh & stale News
z-value(t) < -1.96

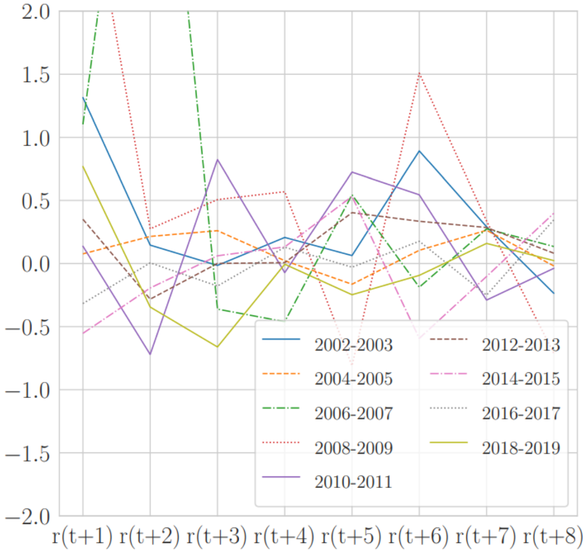


Fig. 12: No consideration of news sentiment
z-value(t) < -1.96

Backtest (Settings B)

Trading strategy: Positive sentiment → LONG; Negative sentiment → SHORT | Buy: MOO(t); Sell: MOO(t+1)

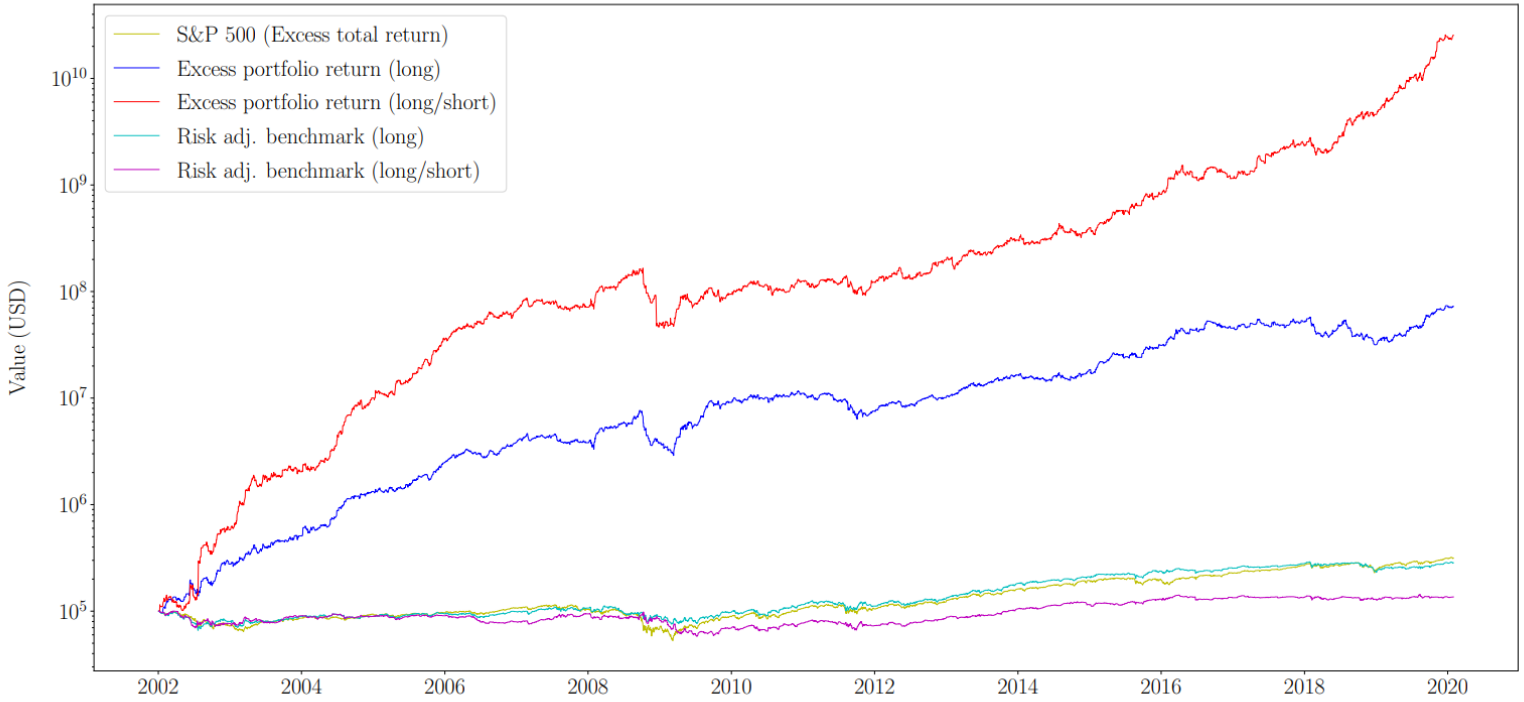


Fig. 16: Backtest
Fresh news, analyst forecast, time-window: 4pm - 9:30am (17.5h), z-value(t, neg. news) < -1.96

	Settings B
Max. weight	100%
News	**Fresh news, Topic: Analyst Forecast**
z-value(t) neg. /pos. news	(−∞, **-1,96**)/(∞/∞)

	Long portfolio	Long/short portfolio	Market portfolio
Sharpe ratio	1.44	2.26	0.36
CAGR	44.04% p.a.	61.22% p.a.	6.56% p.a.
Std. dev.	30.58% p.a.	44.00% p.a.	18.33% p.a.
Beta	0.84	0.49	1
Alpha	13.43 bps/day	29.94 bps/day	
Avg. Portfolio-size	1.93	1.92 (long) 0.42 (short)	
Trades p.a.	454	553	
Return/trade	18.31 bps	**24.06 bps** (50.28 bps im Short leg)	

Table 3: Performance metrics of the long and the long/short portfolio (Settings B)

Backtest (Settings B)

Investment ratio

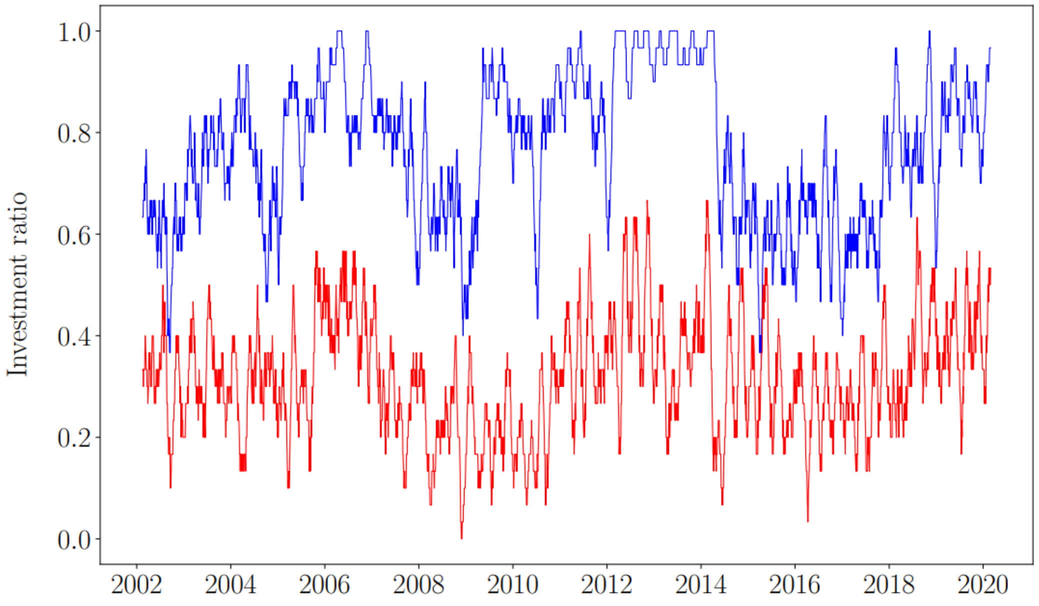


Fig. 17: Investment ratio of the long and the long/short portfolio

Beta (ex-ante)

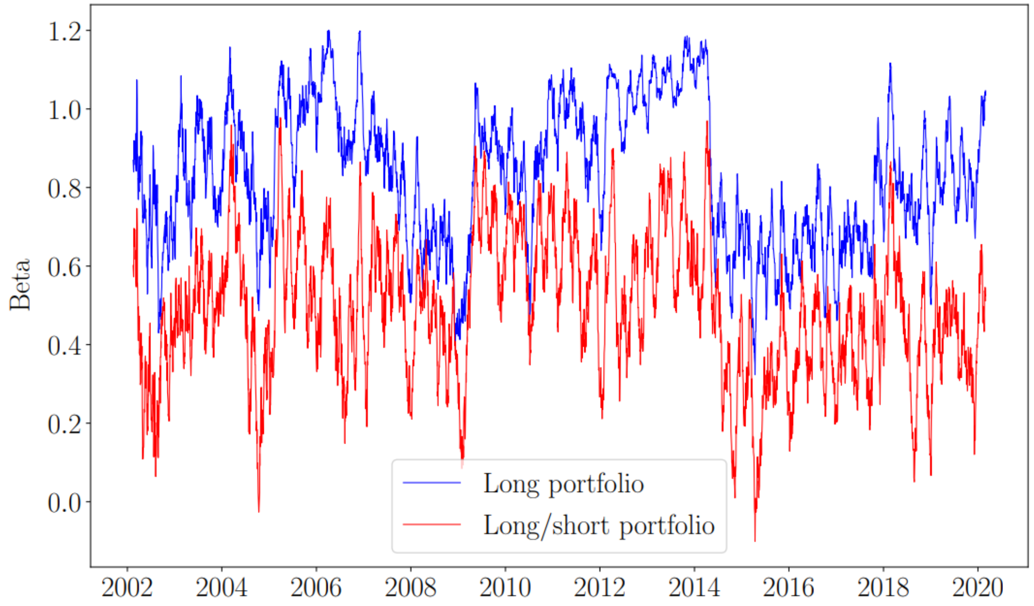


Fig. 18: Ex-ante beta of the long and the long/short portfolio

Fama-French Factor Analysis

	Settings A					
	FF3		FF5		FF5+MOM	
Portfolio	α (bps)	R^2	α (bps)	R^2	α (bps)	R^2
Long	178.19***	18.69%	160.94**	19.59%	160.18**	19.81%
Long/short	529.52***	9.07%	500.20***	11.01%	$\boxed{500.67***}$	11.06%
	Settings B					
Long	262.33***	16.64%	226.72***	17.40%	225.17***	18.10%
Long/short	641.52***	4.55%	648.12***	5.19%	$\boxed{646.36***}$	5.56%

	Long Portfolio					
	FF3		FF5		FF5+MOM	
Factor	Loading	t-value	Loading	t-value	Loading	t-value
Alpha	0.018	2.667	0.016	2.242	0.016	2.227
Beta	1.016	5.630	1.081	5.508	1.057	5.288
SMB	-0.517	-1.585	-0.434	-1.302	-0.409	-1.218
HML	0.463	1.641	0.463	1.241	0.409	1.066
RMW			0.503	1.262	0.574	1.385
CMA			-0.024	-0.047	0.002	0.005
MOM					-0.113	-0.642
	Long/short Portfolio					
Alpha	0.053	5.165	0.050	4.565	0.050	4.554
Beta	-0.872	-3.147	-0.754	-2.517	-0.739	-2.419
SMB	-0.268	-0.535	-0.094	-0.186	-0.110	-0.214
HML	0.357	0.825	0.435	0.764	0.469	0.801
RMW			1.037	1.704	0.993	1.568
CMA			-0.215	-0.275	-0.232	-0.295
MOM					0.071	0.264

Table 4: Fama-French factor analysis of the two backtests with monthly alphas in basis points. **, *** corresponds to significance levels of 5% and 1%.

Conclusion

- The presented models are able to extract a sentiment measure from financial news.
- Financial news contains predictive information about stock returns. Positive correlation between news classification (sentiment) and stock returns.
- Information from financial news is priced in very quickly, usually within a day.
- We observe abnormal returns of up to 2 days after strongly negative news events.
- The additional news topics are valuable features that significantly improve the performance of the model.
- Analyst forecast news contains more predictive power than general financial news.
- Training of BERT on financial text corpora is beneficial

Outlook

- By considering companies with lower market capitalization, higher achievable returns can be expected. According to Kelly et. al. (2019), the price reaction is about four times as large for small companies as for companies with high market capitalization.
- By combining news sentiment with other signals (e.g. historical price data, volatility, market trends, etc.) the return per trade could be further improved.
- The news sentiment could be used as an additional signal for other (longer-term) trading strategies.
- A higher number of additional topic features might further improve prediction accuracy.

Kelly, Bryan T.; Ke, Zheng Tracy; Xiu, Dacheng (2019): Predicting Returns with Text Data (2019-10), S. 54.

References

Araci, Dogu (2019): FinBERT: Financial Sentiment Analysis with Pre-trained Language Models

Cong, Lin; Liang, Tengyuan; Zhang, Xiao (2018): Textual Factors: A Scalable, Interpretable, and Data-driven Approach to Analyzing Unstructured Information.

Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina (2018): BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Gentzkow, Matthew; Kelly, Bryan T.; Taddy, Matt (2017): Text as Data

Kelly, Bryan T.; Ke, Zheng Tracy; Xiu, Dacheng (2019): Predicting Returns with Text Data (2019-10), S. 54.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach.

Mikolov, Tomas; Chen, Kai; Corrado, Greg; Dean, Jeffrey (2013): Efficient Estimation of Word Representations in Vector Space.

Schwartz, Barry (2020): Google: BERT now used on almost every English query. In: *Search Engine Land*, 15.10.2020. Online verfügbar unter https://searchengineland.com/google-bert-used-on-almost-every-english-query-342193, zuletzt geprüft am 16.12.2020.

Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information? The Review of Financial Studies, 24(5):1481–1512

Turc, Iulia; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina (2019): Well-Read Students Learn Better: On the Importance of Pre-training Compact Models.

Wang, Y., Ma, X., Chen, Z., Luo, Y., Yi, J., and Bailey, J. (2019). Symmetric crossentropy for robust learning with noisy labels. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 322–330.

Dangl, Thomas, Salbrechter, Stefan (2021): News Sentiment and Equity Returns. Available at SSRN: https://ssrn.com/abstract=3971880