Hedging via Opinion-based Pair Trading Strategy

Ting-Wei Hsu
Graduate Program of Data Science, National Taiwan University and Academia Sinica, Taiwan
twhsu@nlg.csie.ntu.edu.tw

Chung-Chi Chen
Department of Computer Science and Information Engineering, National Taiwan University, Taiwan
cjchen@nlg.csie.ntu.edu.tw

Hen-Hsen Huang
Department of Computer Science, National Chengchi University, Taiwan
MOST Joint Research Center for AI Technology and All Vista Healthcare, Taiwan
hhhuang@nccu.edu.tw

Meng Chang Chen
Graduate Program of Data Science, National Taiwan University and Academia Sinica, Taiwan
Institute of Information Science, Academia Sinica, Taiwan
mcc@iis.sinica.edu.tw

Hsin-Hsi Chen
Department of Computer Science and Information Engineering, National Taiwan University, Taiwan
MOST Joint Research Center for AI Technology and All Vista Healthcare, Taiwan
hhchen@ntu.edu.tw

ABSTRACT

Risk is an important component when constructing a trading strategy. However, most of the previous works that make the price movement prediction on the basis of the opinions on social media platforms do not take the risk into consideration. In order to hedge the market-risk, we propose an idea of an opinion-based pair trading strategy. Comparing with the task setting of the previous works, our experimental results show that the neural network models with the pair-wise task setting perform better in both accuracy and profitability metrics. That introduces a new research direction for future researches on opinion-based price movement predictions.

CCS CONCEPTS

• **Information systems → Decision support systems**.

KEYWORDS

Pair trading, financial social media, text mining

ACM Reference Format:
Ting-Wei Hsu, Chung-Chi Chen, Hen-Hsen Huang, Meng Chang Chen, and Hsin-Hsi Chen. 2020. Hedging via Opinion-based Pair Trading Strategy. In *Companion Proceedings of the Web Conference 2020 (WWW '20 Companion), April 20–24, 2020, Taipei, Taiwan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3366424.3382701

1 INTRODUCTION

In order to eliminate the influence of systematic risks, hedging is a common method in the financial industry. In hedging, investors long a financial instrument and short the other related financial instrument simultaneously. Although systematic risks influence every financial instrument in the same sector or the same market, few opinion-based price movement prediction works deal with this issue when constructing models. In this paper, we introduce an idea of pair trading to the community and experiment on a benchmark dataset with several neural network (NN) models.

The price movement of a certain financial instrument may be influenced by the information of both the market and the instrument itself. However, previous work [2] omits the market influence toward the price movement of financial instrument, and evaluate the price movement prediction task directly. This kind of evaluation results may be noisy because the market risk is not controlled. The basic idea of pair trading is to make a profit from a market-neutral position. That means we only need to consider the convergence and divergence of the traded pair, and the market risk is already smoothed away. Therefore, the pair-wise setting is more suitable for reflecting the performance of the models on capturing the information of the target financial instrument.

Our experimental results show that, comparing with making predictions one-by-one, considering the stock pair in the same sector at the same time can improve the performance of price movement prediction. Furthermore, we also show the robustness of the pair trading strategies when market drops down.

2 MODELS

In this paper, we experiment on the benchmark dataset, StockNet [2], which is collected from Twitter. We select the stock pairs (s_1,s_2) from the same sector, and aim at predicting the price convergence and divergence of the two target stocks s_1 and s_2 on the trading day d. That is, the models deal with a binary decision, (long s_1 and short s_2) or (short s_1 and long s_2). We experiment on two settings shown as follows.

Independent (IND): As the setting of the previous works, we consider the price movement prediction independently, and trading is based on the following condition.

$$y = \mathbb{1}((p1_d^c > p1_d^o) \ \& \ (p2_d^c < p2_d^o)) \quad (1)$$

WWW '20 Companion, April 20–24, 2020, Taipei, Taiwan

ACM ISBN 978-1-4503-7024-0/20/04.
https://doi.org/10.1145/3366424.3382701

Table 1: Accuracy of Pair Trading and Stock Movement Prediction

Method	Setting	Tech	Finance	BM	CG	Health	Services	Utilities	IG	Avg
Random		51.07	50.29	51.85	48.32	50.41	50.48	49.29	50.79	50.31
CNN	IND	50.13	52.71	53.87	51.96	47.05	53.16	49.83	50.44	50.85
	PW	52.63	51.79	55.28	54.65	55.87	54.92	49.76	48.57	52.94
RNN	IND	48.80	49.70	47.92	50.75	48.12	46.52	48.17	47.91	48.49
	PW	52.55	50.12	53.22	**56.44**	52.98	55.06	**58.43**	55.36	**54.27**
HAN	IND	50.66	48.09	**57.44**	52.26	51.69	51.00	52.49	**57.44**	52.63
	PW	50.96	**53.70**	49.79	55.13	53.81	54.45	55.11	57.37	53.79
SANN	IND	52.78	46.29	55.96	52.26	52.23	50.23	47.84	55.95	51.69
	PW	**53.95**	51.54	53.50	56.32	**55.88**	**55.85**	49.28	56.51	54.10
ABCNN	PW	49.76	51.79	56.79	54.77	55.14	55.38	50.00	52.22	53.23

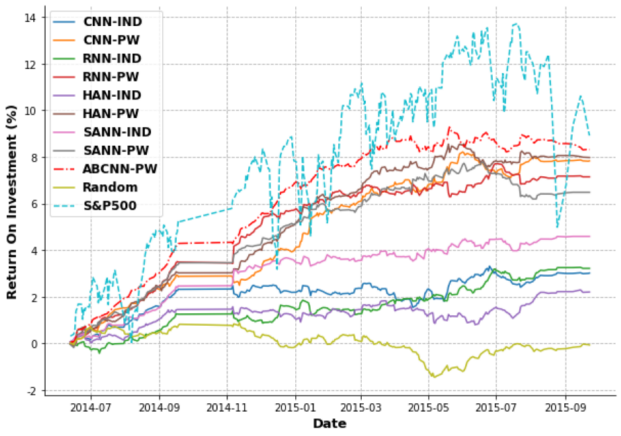

Figure 1: The cumulative profit curve.

where $p1^c_d$ ($p2^c_d$) denotes the closing price for s_1 (s_2), and $p1^o_d$ ($p2^o_d$) denotes the opening price for s_1 (s_2). Note that, 1 denotes long s_1 and short s_2; 0 denotes short s_1 and long s_2.

Pair-wise (PW): In the pair-wise setting, the models consider the tweets of both s_1 and s_2 on day $d-1$ at the same time, and make the decision based on the following condition.

$$y = \mathbb{1}\left(\frac{p1^c_d - p1^o_d}{p1^o_d} > \frac{p2^c_d - p2^o_d}{p2^o_d}\right) \quad (2)$$

We compare several different NN architectures, including CNN, RNN, HAN [3], SANN [1], and ABCNN [4], which are used to do text pair classification, under IND and PW settings. We release the codes for reproducing the results.[1]

3 RESULTS AND DISCUSSION

Table 1 shows the experimental results. Although there does not exist a model performing the best in every sector, we find that considering the information of both target stocks at the same time performs better than the approaches making prediction independently, i.e., PW always performs better than IND no matter which model is adopted. PW performs relatively worse in BM and IG, which are majorly influenced by the macro-economic information. That means PW cannot remove the risk in the industries where stock prices rise and fall at the same time.

After further backtesting experiments with the historical price data, the cumulative profit curve on Services sector is shown in Figure 1. ABCNN-PW performs the best in this experiment. Comparing with the market index (S&P 500), we find that the pair trading methods do help reduce the risk when market drops down under both IND and PW settings. Besides, the PW setting also performs the best in the profitability test. These results show the usefulness of making pair-wise consideration when constructing trading strategies.

4 CONCLUSION

In this paper, we explore the opinion-based pair trading strategy with several recent classification models, and find that the task setting of previous works, which consider the financial instruments independently, performs worse than the pair-wise setting proposed in this paper under both accuracy and profitability metrics. Furthermore, we also evidence that the opinion-based pair trading strategy can reduce the downside risk when the market crash.

ACKNOWLEDGMENTS

This research was partially supported by Ministry of Science and Technology, Taiwan, under grants MOST-106-2923-E-002-012-MY3, MOST 108-2218-E-009-051, MOST 108-2634-F-002-017, and MOST 109-2634-F-002-034, and by Academia Sinica, Taiwan, under grant AS-TP-107-M05.

REFERENCES

[1] Zhouhan Lin, Minwei Feng, Cicero Nogueira dos Santos, Mo Yu, Bing Xiang, Bowen Zhou, and Yoshua Bengio. 2017. A structured self-attentive sentence embedding. *arXiv preprint arXiv:1703.03130* (2017).
[2] Yumo Xu and Shay B. Cohen. 2018. Stock Movement Prediction from Tweets and Historical Prices. In *ACL*.
[3] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *NAACL*. 1480–1489.
[4] Wenpeng Yin, Hinrich Schütze, Bing Xiang, and Bowen Zhou. 2016. Abcnn: Attention-based convolutional neural network for modeling sentence pairs. *Transactions of the Association for Computational Linguistics* (2016).

[1] https://github.com/quanthsu/PairTrading