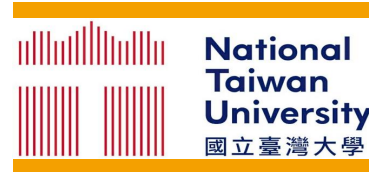


Retrieving Implicit Information for Stock Movement Prediction



Tsun-Hsien Tang

Department of Computer Science and Information Engineering,
National Taiwan University, Taiwan
thtang@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

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



國立臺灣大學
National Taiwan University

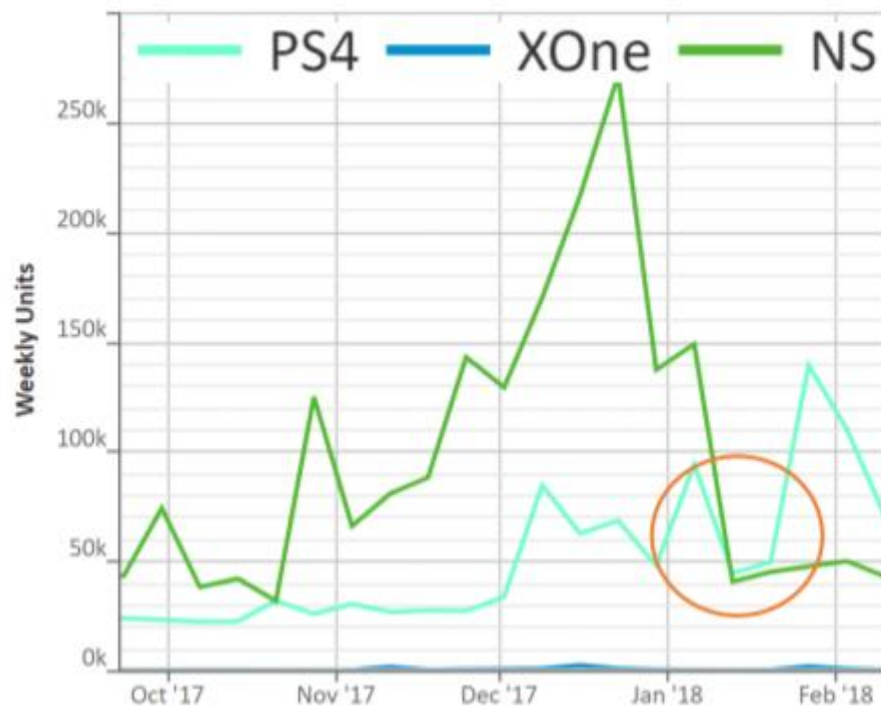


科技部人工智慧技術
暨全幅健康照護聯合研究中心
[Most Joint Research Center for AI Technology and All Vista Healthcare](#)

sigir21

Main Concept – Implicit Information

- Previous Work: Predict the Stock Movement of Sony (**Mentioned** in the News Article)
- **This Paper:** Predict the Stock Movement of Supply Chain Vendors (**Not Mentioned** in the News Article)



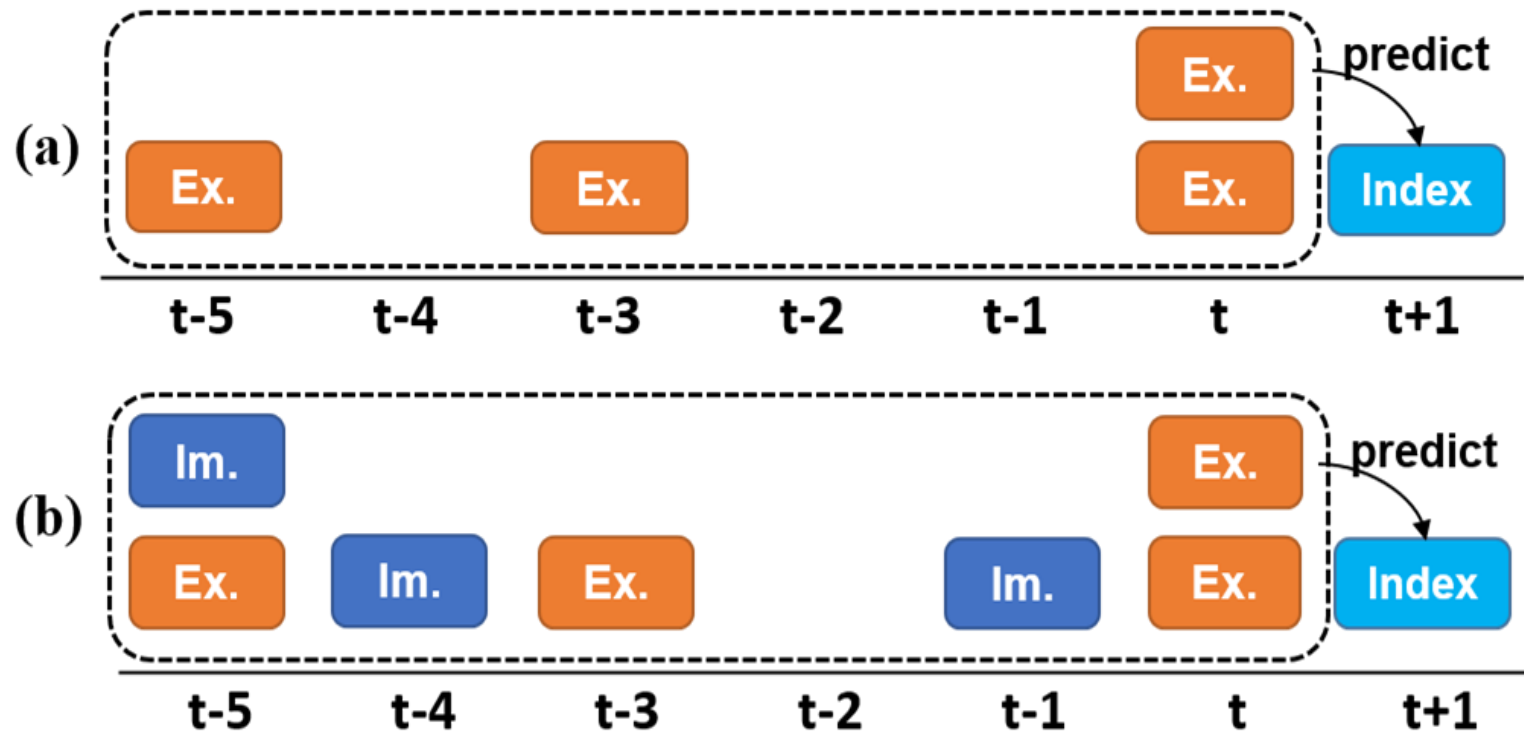
2018/01/19 news:

Sales of PS4 Breaks 6 Million and Beats Switch First Time

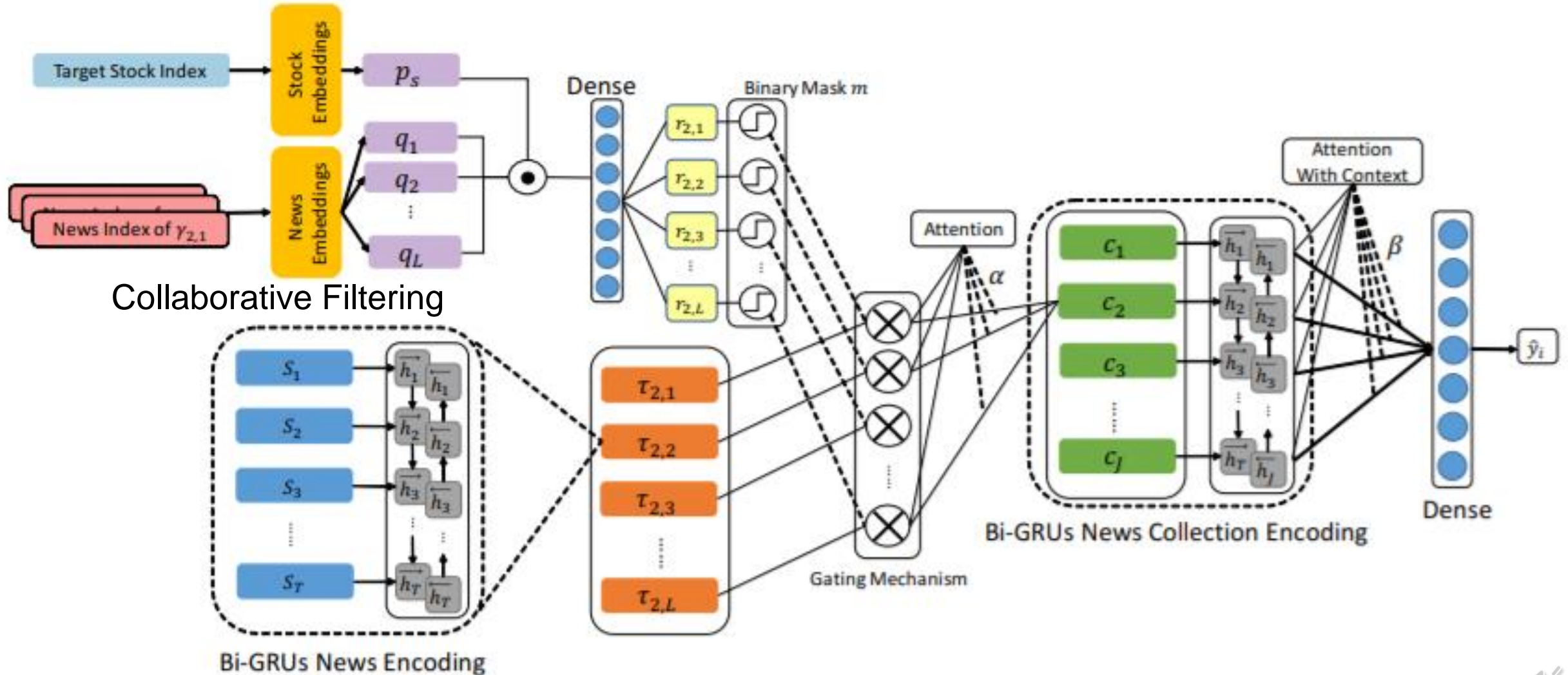
	Date	TAIEX	Foxconn	Delta Electronics
Price (NTD)	2018/1/17	11004.80	93.4	139.0
	2018/1/18	11071.57	93.8	143.0
	2018/1/19	11150.85	96.9	144.5
	2018/1/22	11231.46	96.6	149.0
Return (%)	2018/1/18	0.61%	0.43%	2.88%
	2018/1/19	0.72%	3.30%	1.05%
	2018/1/22	0.72%	-0.31%	3.11%
CAR ₃			1.38%	5.00%



Benefit – The Sparse Temporal News Sequence



News Distilling Network



Experiment

- Dataset

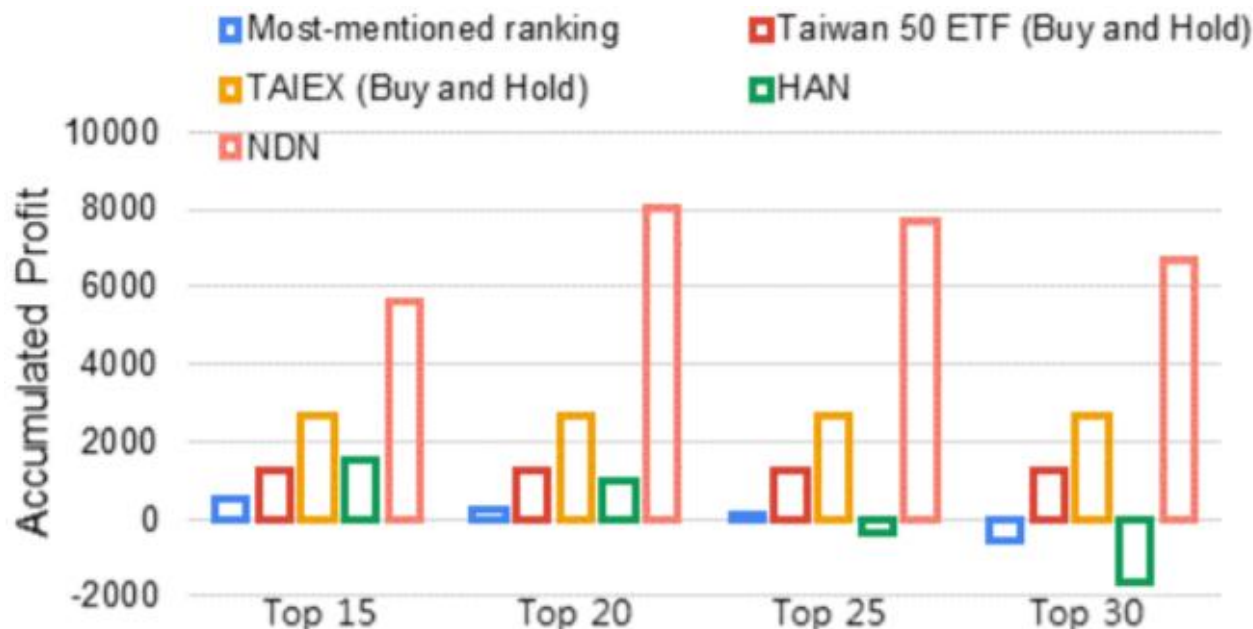
	Train	Test
Period	2013/06/22-2018/01/01	2018/01/01-2018/06/20
RISE	29,067	2,906
FALL	26,778	2,678
Total	55,845	5,584

- Cumulative Abnormal Return (CAR)

$$CAR_{s,n} = \sum_{d=1}^n (R_{s,d} - \hat{R}_d)$$

Results

Model	Acc. (%)	MCC
Random Guess	50.77	0.0147
BoW + Random Forest	50.97	0.0159
FastText + Random Forest	52.83	0.0485
HAN	54.37	0.0788
HAN _{s1}	53.85	0.0736
NDN _{s1} (Proposed)	56.75*	0.1302*
HAN _{s2}	56.77*	0.1300*
NDN _{s2} (Proposed)	57.89*	0.1536*



- Setting 1 (S_1): We **fused all news regardless of the existence of manual stock labels** to evaluate the robustness of the selection and attention mechanism.
- Setting 2 (S_2): **We integrate manually labeled implicitly related news with explicit news** to determine whether information from related news aids the prediction task. For NDN, all kinds of news were fed into the model. The results under this setting can be considered as the upper-bound of the models with implicit information



Conclusion

- We discuss **the problem of sparse temporal news sequences** as well as the shortcoming of previous works that are based on explicit news only.
- We develop **news distilling networks** for discovering implicitly related news given limited labels.
- The experiments on the real-world data show that the proposed methods **significantly improve the accuracy of stock movement prediction** and **yield greater profits** than the previous methods.



Related Works and Events

- **Related Works**

- Dynamic Graph Transformer for Implicit Tag Recognition. EACL'21
- FinSense: An Assistant System for Financial Journalists and Investors. WSDM'21
- Springer SpringerBriefs: **From Opinion Mining to Financial Argument Mining**



- **Related Events**

- FinNum-3 **Shared Task** @ NTCIR-2022 – Investor's and Manager's **Fine-grained Claim Detection**
- The Third **Workshop of Financial Technology** and Natural Language Processing (FinNLP-2021 @ **IJCAI-2021**)
- The 2nd Workshop on Financial Technology on the Web (FinWeb-2022) (May be in conjunction with WWW'22)
- **EMNLP-2021 Tutorial**: Financial Opinion Mining



Join Mail List for Latest News → <https://forms.gle/RB9Qq9ok6z5exu1G6>

Feel free to contact us if you have any questions.

Chung-Chi Chen: cjchen@nlg.csie.ntu.edu.tw

