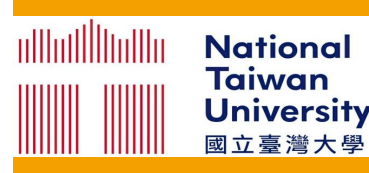


Distilling Numeral Information for Volatility Forecasting



Chung-Chi Chen

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

Hen-Hsen Huang

Institute of Information Science, Academia Sinica, Taiwan
MOST Joint Research Center for AI Technology and All Vista Healthcare, Taiwan
hhhuang@iis.sinica.edu.tw

Yu-Lieh Huang

Department of Quantitative Finance,
National Tsing Hua University, Taiwan
Center for Research in Econometric Theory and Applications, National Taiwan University, Taiwan
ylihuang@mx.nthu.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

Earnings Conference Call

- Modeling Financial Analysts' Decision Making
 - Keith, Katherine, and Amanda Stent. "Modeling Financial Analysts' Decision Making via the Pragmatics and Semantics of Earnings Calls." *ACL-2019*.
- Risk Forecasting
 - **Chung-Chi Chen, Hen-Hsen Huang, Yu-Lieh Huang, and Hsin-Hsi Chen. "Distilling Numeral Information for Volatility Forecasting" *CIKM-2021***
 - Sawhney, Ramit, et al. "Multimodal Multi-Speaker Merger & Acquisition Financial Modeling: A New Task, Dataset, and Neural Baselines" *ACL-2021*
 - Ye, Zhen, Yu Qin, and Wei Xu. "Financial Risk Prediction with Multi-Round Q&A Attention Network." *IJCAI-2020*.
 - Sawhney, Ramit, et al. "VoTAGE: Volatility Forecasting via Text-Audio Fusion with Graph Convolution Networks for Earnings Calls." *EMNLP-2020*.
 - Qin, Yu, and Yi Yang. "What you say and how you say it matters: Predicting financial risk using verbal and vocal cues." *ACL-2019*
 - Yang, Linyi, et al. "HTML: Hierarchical Transformer-based Multi-task Learning for Volatility Prediction." *Proceedings of The Web Conference 2020*.
 - Li, Jiazheng, et al. "MAEC: A Multimodal Aligned Earnings Conference Call Dataset for Financial Risk Prediction." *CIKM-2020*.

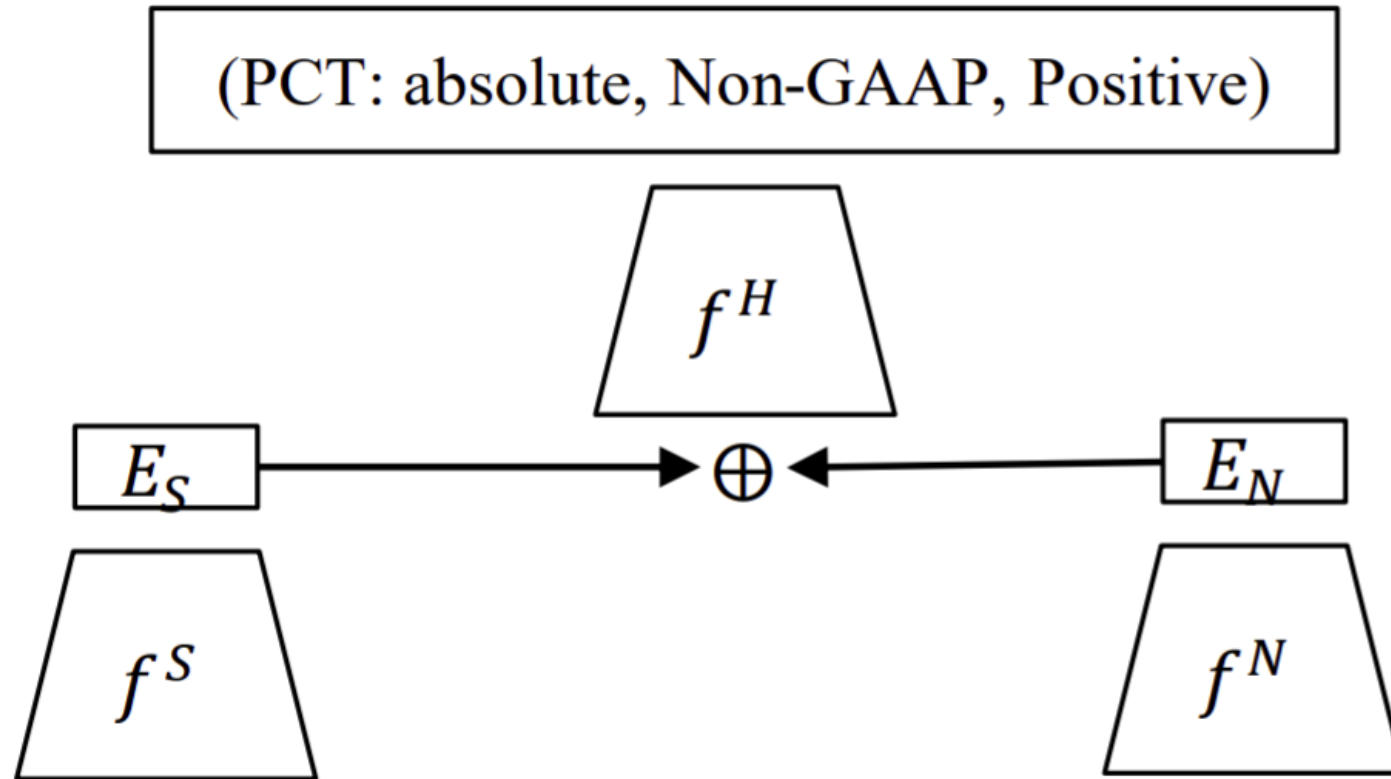


Numerals in Financial Documents

	Earnings call	Analysis report	Financial tweet	Financial news	
Unit of measurement	Sentence	Sentence	Tweet	Headline	Article
Instances	13,574	4,952	2,028	75,448	75,448
Instances with numerals	7,499	2,938	1,395	45,073	75,297
Proportion with numerals	55.3%	59.3%	68.8%	59.7%	99.8%



Proposed Idea



[CLS] Our adjusted tax rate is expected to be **20.5**.

- **Numeral Understanding**

Category	Earnings Call		Analysis Report		Tweet	
	# of instance	Ratio	# of instance	Ratio	# of instance	%
Monetary - money	2,656	19.72%	874	16.99%	736	8.30%
Monetary - quote	-	-	75	1.46%	1,033	11.65%
Monetary - change	753	5.59%	18	0.35%	176	1.98%
Monetary - buy price	-	-	-	-	415	4.68%
Monetary - sell price	-	-	-	-	135	1.52%
Monetary - forecast	-	-	-	-	355	4.00%
Monetary - stop loss	-	-	-	-	35	0.39%
Monetary - support or resistance	-	-	-	-	302	3.41%
Percentage - relative	3,040	22.57%	708	13.76%	767	8.65%
Percentage - absolute	969	7.19%	810	15.75%	346	3.90%
Temporal - date	2,647	19.65%	2,134	41.49%	2,653	29.92%
Temporal - time	8	0.06%	3	0.06%	365	4.12%
Option - exercise price	-	-	-	-	132	1.49%
Option - maturity date	-	-	-	-	70	0.79%
Indicator	-	-	-	-	216	2.44%
Quantity	2,199	16.33%	278	5.40%	982	11.07%
Product/Version	349	2.59%	136	2.64%	150	1.69%
Ranking	50	0.37%	3	0.06%	-	-
Other	798	5.92%	105	2.04%	-	-
	13,469	100.00%	5,144	100%	8,868	100%

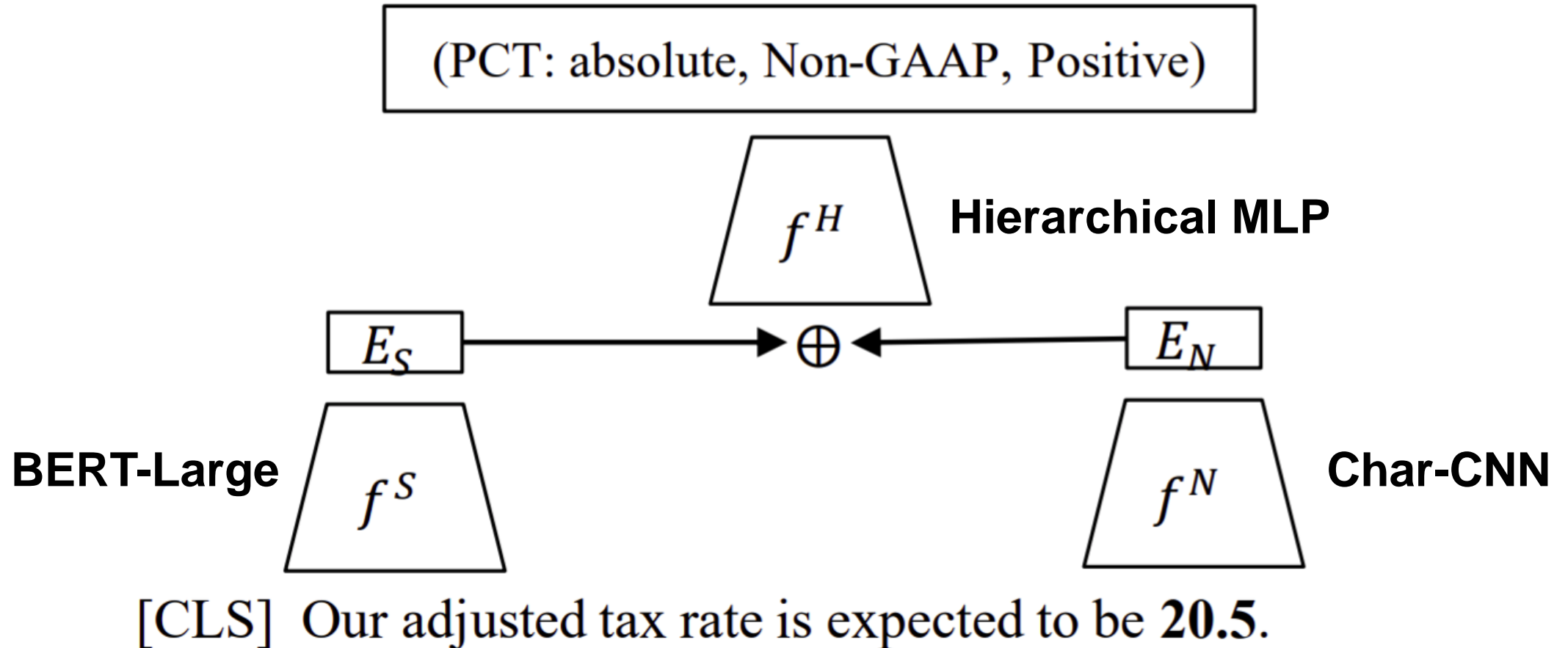
- **Account Differentiation**

- Generally Accepted Accounting Principles (GAAP)
- Non-GAAP: Adjust Term for GAAP
- Other

- **Influence Analysis**

Account Differentiation			Influence Analysis		
Class	# labels	%	Class	# labels	%
GAAP	3,675	40.68%	Positive	5,467	60.52%
Non-GAAP	432	4.78%	Negative	669	7.41%
Other	4,927	54.54%	Neutral	2,898	32.08%

Numeral-Aware Model (NAM)



Distilling Numeral Information

	NU	Account	Influence
BERT-Large	36.83	75.34	62.44
Multi-task BERT	48.61	75.13	72.77
T5-Base (220M)	74.60	85.28	74.38
NAM-Base (110M)	86.45	86.05	87.39
NAM (340M)	88.42	88.15	89.87

	NU	Account	Influence
NAM-Base	86.45	86.05	87.39
w/o f^N	48.37	77.50	67.57
w/o f^N & AT	36.98	77.70	64.20



Volatility Forecasting

	3-day	7-day	15-day	30-day
MDRM (Text Only)	1.431	0.439	0.309	0.219
MDRM (Text + Audio)	1.371	0.420	0.300	0.217
HTML (Text Only)	1.175	0.372	0.153	0.133
HTML (Text + Audio)	0.845	0.349	0.251	0.158
Proposed	0.745	0.300	0.232	0.187



Conclusion

- We present a dataset, **ECNum**, for understanding the numeral information in earnings conference calls.
- We propose a simple but efficient model for distilling the numeral information, and show that the proposed model performs better than other previous models.
- We also experiment on a real-world application scenario, volatility forecasting, and find that our method performs well in short-term volatility forecasting tasks.



Related Works and Events

• Related Works

- **From Opinion Mining to Financial Argument Mining.** (Springer Nature – Open Access)
 - <http://springer.nlpfin.com/>
- Constructing Noise Free Economic Policy Uncertainty Index (CIKM'21)
- NQuAD: 70,000+ Questions for Machine Comprehension of the Numerals in Text (CIKM'21)

• Related Events

- **Call for Paper:** Financial Technology on the Web @ ACM TWEB
 - <https://acmfinweb.nlpfin.com/>
- **FinNum-3 Shared Task @ NTCIR-2022 – Investor's and Manager's Fine-grained Claim Detection**
- **EMNLP-2021 Tutorial:** Financial Opinion Mining
- The **Workshop of Financial Technology** and Natural Language Processing (FinNLP @ **IJCAI**)
- The Workshop on Financial Technology on the Web (FinWeb @ **WWW**)

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



Thank you for your attention!

Feel free to contact us if you have any questions.

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

<http://cjchen.nlpfin.com/>

Join Mail List for Latest News

➔ <https://forms.gle/RB9Qq9ok6z5exu1G6>



**National
Taiwan
University**

國立臺灣大學

