FinNum-3 –
Investor's and Manager's Fine-grained Claim Detection

Hsin-Hsi Chen, Department of Computer Science and Information Engineering, National Taiwan University

Hiroya Takamura, AIRC, National Institute of Advanced Industrial Science and Technology

Yu-Lieh Huang, Department of Quantitative Finance, National Tsing Hua University

Hen-Hsen Huang, Department of Computer Science, National Chengchi University

Chung-Chi Chen, MOST Joint Research Center for AI Technology and All Vista Healthcare

FinNum-3

- Manager's and Investor's Fine-grained Claim Detection
    - **Chinese** financial analysis reports (**investor**)
    - **English** earnings conference call (**manager**)

- Information in NumClaim 2.0
    - Given
        - Target numeral
        - Context of Target numeral
    - Model Output
        - Category of target numeral (FinNum-1)
        - In-claim or out-of-claim (FinNum-3)
    - Metrics
        - Micro-F1 and Macro-F1 Scores
        - https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html

Motivation

- FinNum-1 & FinNum-2 Social Media

- FinNum-3 ➜ Formal Documents

- Argument mining in finance

- Over **58.47%** of sentences in analysis report contain at least one numeral

- Investors always make a claim **with an estimation**
  - **(X) We estimate that the sales may growth**
  - **(O) We estimate that the sales growth rate may exceed 40%**

- The importance of fine-grained claims and the numerals.
  - We estimate that the sales growth rate may exceed 5%
  - We estimate that the sales growth rate may exceed 40%

NumClaim

- **Chinese** financial **analysis reports (investor)**
- The annotators work in the financial industry (bank's treasury department and hedge fund)
- The Cohen's kappa agreements between the experts are 88.31%
- 5,144 instances: 23.78% "In-claim" and 76.22% "Out-of-claim"

Sentence	Label
We estimate that the sales growth rate may exceed **40**%.	In-claim
Professional audio/visual products account for **20**%.	Out-of-claim

In-claim		Out-of-claim	
estimate	2.86	lower/higher than	-1.37
price target	2.80	cause	-1.37
downgrade	2.58	last year	-1.26
upgrade	1.55	influence	-1.25

Chen, Chung-Chi, Hen-Hsen Huang, and Hsin-Hsi Chen. "NumClaim: Investor's Fine-grained Claim Detection." *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2020.

CIKM 2020
19-23 OCTOBER
Online

Auxiliary Task – Numeral Understanding (FinNum-1)

- The Cohen's kappa agreements between the experts are 89.55%

Category	Subcategory	In-claim	Out-of-claim	Sum
Monetary	price	42	33	75
	money	506	368	874
	change	3	15	18
Percentage	absolute	208	500	708
	relative	408	402	810
Temporal	date	0	2,134	2,134
	time	0	3	3
Quantity	absolute	55	219	274
	relative	0	4	4
Product Number		1	135	136
Ranking		0	3	3
Other		0	105	105
Total		1,223	3,921	5,144

Chung-Chi Chen, Hen-Hsen Huang, Yow-Ting Shiue, and Hsin-Hsi Chen. 2018. Numeral understanding in financial tweets for fine-grained crowd-based forecasting. In *IEEE/WIC/ACM International Conference on Web Intelligence*

Statistics

Dataset	NumClaim	CRC [13]	PE [12]
Language	Chinese	Chinese	English
Source	Analysis Report	Hotel Review	Persuasive Essay
# Word	42,594	21,848	97,420
# Numeral	5,144	67	111

	NumClaim	CRC
# hard words	**31.95**	18.28
# negative words	0.14	**0.60**
# synonym	0.28	**1.49**
Noun phrase modifier ratio	0.29	**0.38**
Noun phrase ratio	**31.79**	26.62
# transition words	**4.86**	1.62

[12] Steffen Eger, Johannes Daxenberger, and Iryna Gurevych. 2017. Neural End-to-End Learning for Computational Argumentation Mining. In ACL
[13] Steffen Eger, Johannes Daxenberger, Christian Stab, and Iryna Gurevych. 2018.Cross-lingual Argumentation Mining: Machine Translation (and a bit of Projection) is All You Need!. In COLING.

Experimental Results

- Encoding: BERT

- Baseline: CNN, BiGRU, CapsNet

- Metrics: Macro-F1

- Class Weight (CW)

- Numeral Encoder

  - Represent the digit (0-9) and the decimal point as a 11-dimension tensor, and concatenate it with a tensor for the inter-numeral position information.

- Joint Learning with Category Classification Task (CG)

Architecture	CNN	BiGRU	CapsNet
Baseline	76.15%	77.97%	77.93%
+ CW	77.26%	78.29%	78.68%
+ CW & NE (CNN)	78.19%	79.06%	80.91%
+ CW & NE (CNN) & CG	81.35%	81.65%	**82.62%**

7

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
  - Ye, Zhen, Yu Qin, and Wei Xu. "Financial Risk Prediction with Multi-Round Q&A Attention Network." IJCAI-2020.
  - Sawhney, Ramit, et al. "VolTAGE: 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.

Questions from Reviewers (1/3)

- The organisers ran a related task outside NTCIR that involves a cash prize. **If that task involves a cash prize, and the NTCIR task does not, what would be the motivation for researchers to sign up with the latter?**

  - Originally, we also plan to prepare a cash prize for FinNum-2 participants. However, after discussing with chairs and other organizers, we decide to cancel the cash prize. Thus, we did not show the cash prize in the proposal of FinNum-3.

- The size of the listed categories is highly **unbalanced**. Therefore, the authors should consider focusing on some major categories instead of using all of them in the second task.

  - Because category information is prepared as the auxiliary task for enhancing the performance of claim detection, our participants can explore different settings.

Questions from Reviewers (2/3)

- It was not very clear how beneficial it is to distinguish in-claim and out-claim. It should be better motivated. The lack of any end-user task (the proposal asserts that there will be beneficial effects on "**downstream tasks**", but no such tasks are named)
  - Argumentation Strategies
  - Argumentation Structures
  - **Chung-Chi Chen, Hen-Hsen Huang, and Hsin-Hsi Chen. 2021. Evaluating the Rationales of Amateur Investors. In Proceedings of The Web Conference 2021 (WWW'21)**

- It is not clear whether the macroaverage will be taken over types of subtypes in the type classification task, but the real puzzle is what the **macroaverage** will be computed over in the "in-claim" detection task.
  - Calculate metrics for each label, and find **their unweighted mean**. This does not take label imbalance into account.

- There is another concern for in- and out-claim classification: Is it a problem of **numerals**?
  - Based on the results of the pilot exploration, we find that numeral understanding can be used for better comprehending financial documents. Claim detection is the first step of argument mining. We hope to use FinNum-3 to link with the next shared task series we plan to propose in NTCIR-16-18 – **Argument Mining in Financial Narratives (FinArg)**

- Can Chinese Analyst's Report and English Earnings Conference Call **be released under the license**?
  - Chinese Analyst's Report ➔ Can be shared by request for academic usage
  - English Earnings Conference Call ➔ Publicly-available

Schedule

- 2021
  - Apr 30 Chinese Training and Development Set Release
  - Jul 31 English Training and Development Set Release
  - Dec 1 Test Set Release
- 2022
  - Jan 3 Task Registration Due
  - Jan 10 Participants' Results Submission Due
  - Feb 1 Evaluation Result Release & Draft Task Overview Paper Release
  - Mar 1 Participants' Papers Submission Due
  - May 1 Camera-ready Paper
  - Jun NTCIR-16 Conference in NII, Tokyo, Japan

Feel free to contact us if you have any questions.

March & April → Hen-Hsen Huang: hhhuang@nccu.edu.tw
From May → Chung-Chi Chen: cjchen@nlg.csie.ntu.edu.tw

NLP
Lab
National Taiwan University