Overview of FinNum

Fine-Grained Numeral Understanding in Financial Social Media Data

Chung-Chi Chen, Hen-Hsen Huang, Hiroya Takamura and Hsin-Hsi Chen

NLP Lab National Taiwan University

National Taiwan University
國立臺灣大學

NATIONAL CHENGCHI UNIVERSITY
政治大學

AIRC

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

Motivation

Numerals on Social Trading Platforms

Bullish

Apr 26th, 1:39 am

$TSLA 256 Break-out thru 50 & 200- DMA (197-230) upper head res (274-279) Short squeeze in progress Nr term obj: 310 Stop loss:239

Bullish

Apr 12th, 7:02 pm

$AAPL support identified $198.8 ... next move to $215

5

Bearish

Apr 12th, 7:02 pm

$TVIX making a new 52 week low.

4

1

Introduction

$TSLA 256 Break-out thru 50 & 200- DMA (197-230) upper head res (274-279) Short squeeze in progress Nr term obj: 310 Stop loss:239. *25 tokens 9 numbers 6 meanings*

We

- propose **fine-grained numeral taxonomy** for financial social media data
- attempt to **leverage the numeral opinions made by the crowd** to mine **additional information** for trading

I will introduce the

- application of proposed tasks
- numeral taxonomy
- details of FinNum shared task
- empirical studies of extracted information
- further research direction of the numerals in financial data
- FinNum-2 proposal

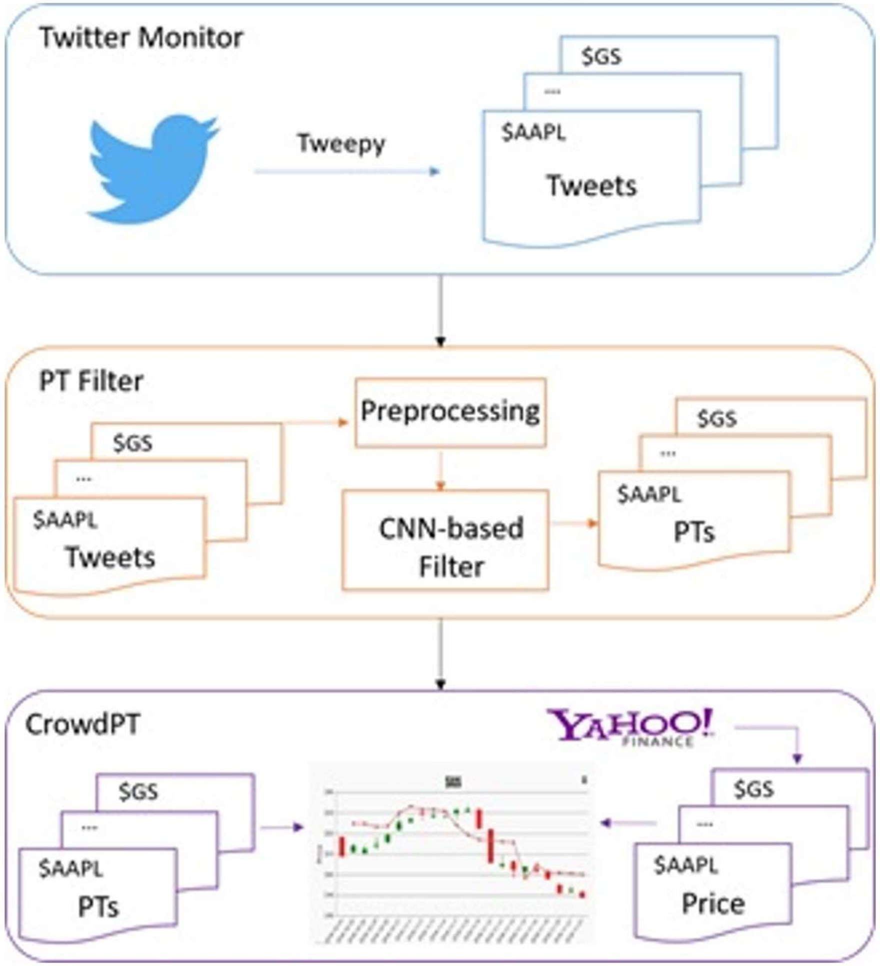
Application Scenario

National Taiwan University

Crowd View: Converting Investors' Opinions into Indicators

NLP Lab
National Taiwan University

System Flowchart



- Monetary
  - price target
  - support or resistance
  - buying
  - selling
  - other
- Percentage
- Temporal
- Option
- Indicator
- Quantity
- Product/ Version Number

IJCAI 2019 MACAO

Numeral Taxonomy

Numeral Taxonomy

NLP Lab
National Taiwan University

Category	Subcategory	Train	Dev.	Test	Total	Ratio
Monetary		2467	261	459	3187	35.94%
	money	589	52	95	736	8.30%
	quote	792	89	152	1033	11.65%
	change	143	8	25	176	1.98%
	buy price	319	36	60	415	4.68%
	sell price	103	10	22	135	1.52%
	forecast	270	33	52	355	4.00%
	stop loss	25	4	6	35	0.39%
	support or resistance	226	29	47	302	3.41%
Percentage		838	105	170	1113	12.55%
	relative	585	70	112	767	8.65%
	absolute	253	35	58	346	3.90%
Option		169	11	22	202	2.28%
	exercise price	113	5	14	132	1.49%
	maturity date	56	6	8	70	0.79%
Indicator		167	22	27	216	2.44%
Temporal		2364	253	401	3018	34.03%
	date	2079	223	351	2653	29.92%
	time	285	30	50	365	4.12%
Quantity		741	87	154	982	11.07%
Product/Version		114	14	22	150	1.69%
		6860	753	1255	8868	100.00%

Monetary

NLP Lab
National Taiwan University

- The Monetary category contains the following 8 subcategories:
  - "money", "quote" and "change"
  - **"buy price", "sell price", "forecast", "stop loss" and "support or resistance"**

- The identification of "buy price" and "sell price" can help us understand the performance of the writer.
  - $SPY Long 1/2 position 137.89

- Some investors "forecast" the price of the instruments depending on their analysis results.
- The concepts of support and resistance are always discussed in technical analysis.

Percentage

- The numeral that indicates the proportion of a certain amount is classified into “absolute”.

- The numeral that stands for the change relative to original amount is classified into “relative”.

- ¢Den up almost **10**% since Q1 and £áuro up around 7.5%, much more $ for $AAPL pocket. Remember **23**% of Apple revenues comes from this two @jimcramer
  - 10% and 7.5% are annotated as “relative”
  - 23% stands for “absolute”.

Option

- Option is a popular instrument frequently discussed.
- To capture the implications of investors' opinions, we propose two subcategories for Option category, "exercise price" and "maturity date".
- $XLU long April $**44** calls
- $MSFT those APR.**22** CALLS were getting hot.

Indicator

NLP Lab
National Taiwan University

- This category captures the parameters of the technical indicators.

- Different investors may use dissimilar parameters for the same indicator. In order to capture the price most investors pay attention to, we should identify the parameters being used.

- $ATHX riding **5**dma higher, dropping to **13**dma at the dips, sign of a healthy advancing stock that stays above **20**dma

Temporal

- Temporal information is also important in financial domain.
- The day most investor focusing on is the one with high volatility.
- We classify Temporal category into two subcategories, “date” and “time”

Quantity

- Quantity information can help us know the position of an investor, and we can give the large weighting to the opinions held by persons who have large positions.

Product/Version Number

- The version of products may contain numerals. We can use the product information to compare importance of different tweets.

- For example, the tweets discuss of iPhone 7 may be more important than the tweets that discuss iPhone 4.

Dataset

Corpus Creation

NLP Lab
National Taiwan University

- We collected the data from StockTwits.
- Two experts were involved in the annotating process.
- FinNum dataset contains only the numerals in full agreement.

Distribution

NLP Lab
National Taiwan University

Category	Subcategory	Train	Dev.	Test	Total	Ratio
Monetary		2467	261	459	3187	35.94%
	money	589	52	95	736	8.30%
	quote	792	89	152	1033	11.65%
	change	143	8	25	176	1.98%
	buy price	319	36	60	415	4.68%
	sell price	103	10	22	135	1.52%
	forecast	270	33	52	355	4.00%
	stop loss	25	4	6	35	0.39%
	support or resistance	226	29	47	302	3.41%
Percentage		838	105	170	1113	12.55%
	relative	585	70	112	767	8.65%
	absolute	253	35	58	346	3.90%
Option		169	11	22	202	2.28%
	exercise price	113	5	14	132	1.49%
	maturity date	56	6	8	70	0.79%
Indicator		167	22	27	216	2.44%
Temporal		2364	253	401	3018	34.03%
	date	2079	223	351	2653	29.92%
	time	285	30	50	365	4.12%
Quantity		741	87	154	982	11.07%
Product/Version		114	14	22	150	1.69%
		6860	753	1255	8868	100.00%

Task Setting

Task Formulation & Evaluation

NLP Lab
National Taiwan University

- The position of a numeral in a tweet is given in advance.
- Participants are asked to disambiguate its category.
- This task is further separated into two subtasks:
  - Classify a numeral into 7 categories, i.e., Monetary, Percentage, Option, Indicator, Temporal, Quantity and Product/Version Number.
  - Extend the classification task to the subcategory level, and classify numerals into 17 classes, including Indicator, Quantity, Product/Version Number, and all subcategories
- Micro-averaged F-score and macro-averaged F-scores are adopted for evaluating the classification performance of participants' runs.

Participants

12 Teams including 15 Institutions from 6 Countries

NLP Lab
National Taiwan University

Participants

Microsoft
Fortia Financial Solutions
中央研究院 ACADEMIA SINICA
ORACLE
opt
University of Zurich[UZH]
Lenovo
AUTO1
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL
南開大學 Nankai University
國立臺北大學 National Taipei University
清華大學 Tsinghua University
重慶郵電大學 CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS
哈爾濱工業大學 HARBIN INSTITUTE OF TECHNOLOGY
武漢科技大學

Methods

National Taiwan University

Models

NLP Lab
National Taiwan University

Features

Topic
Format
Position
Keywords
Named Entity
Brown Cluster
Part-of-speech
Term Frequency
Prefixes/ Suffixes
Bag-of-Characters
Numeral Information
Recognizers-Text Type

Task Setting

Classification
Sequential Labeling

Representation

Skip-Gram
GloVe
ELMo
BERT

Models

SVM
MLP
CNN
RNN
RNN + CNN
Attention-based LSTM

6/12 10:15-11:45 Session B-2

Results

Participants Results

NLP Lab
National Taiwan University

Subtask 1			Subtask 2		
Submission ID	Micro F1 (%)	Macro F1 (%)	Submission ID	Micro F1 (%)	Macro F1 (%)
Fortia1 - 1	93.94	90.05	Fortia1 - 2	87.17	82.40
Fortia1 - 2	93.70	88.98	Fortia1 - 1	86.53	80.49
DeepMRT - 1	91.87	87.94	DeepMRT - 1	83.03	77.90
DeepMRT - 2	91.16	84.72	DeepMRT - 2	81.27	75.59
ASNLU - 2	89.72	80.93	aiai - 1	80.24	74.11
ASNLU - 1	89.40	79.96	aiai - 2	80.64	73.43
ZHAW - 2	86.45	79.27	ASNLU - 1	79.12	72.51
Fortia2 - 1	89.88	79.26	ASNLU - 2	77.37	70.09
Fortia2 - 2	87.73	78.59	Fortia2 - 2	77.05	68.86
aiai - 1	86.45	78.09	Fortia2 - 1	79.28	68.33
aiai - 2	87.41	78.04	ZHAW - 2	75.54	66.44
ZHAW - 1	84.78	75.40	ZHAW - 1	72.67	64.84
WUST	74.02	63.71	Stark - 1	69.08	56.83
BRNIR - 1	74.27	63.53	WUST	60.88	52.93
Stark - 1	78.01	61.75	BRNIR - 1	63.67	51.90
BRNIR - 2	72.91	58.54	BRNIR - 2	61.99	47.14
word-based CNN	55.90	51.67	char-based CNN	43.75	31.12

6/12 10:15-11:45 Session B-2

Error Analysis

NLP Lab
National Taiwan University

Confusion matrix

True label \ Predicted label	mon	qut	cha	buy	sel	for	sto	sup	rel	abs	exe	mat	Ind	dat	tim	Qua	Pro
mon	0.75	0.08	0.02	0.01	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.07	0.00
qut	0.04	0.67	0.01	0.04	0.01	0.02	0.00	0.02	0.02	0.01	0.00	0.00	0.01	0.09	0.01	0.04	0.01
cha	0.07	0.16	0.43	0.02	0.00	0.03	0.00	0.02	0.03	0.01	0.00	0.00	0.01	0.07	0.03	0.11	0.01
buy	0.01	0.17	0.00	0.71	0.02	0.02	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.00
sel	0.02	0.24	0.00	0.19	0.37	0.02	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.05	0.00	0.10	0.00
for	0.02	0.08	0.02	0.01	0.00	0.77	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.00
sto	0.01	0.09	0.00	0.10	0.07	0.17	0.46	0.08	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00
sup	0.01	0.24	0.00	0.02	0.01	0.02	0.00	0.64	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.02	0.00
rel	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.86	0.10	0.00	0.00	0.00	0.01	0.00	0.01	0.00
abs	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.75	0.00	0.00	0.00	0.00	0.00	0.01	0.00
exe	0.02	0.09	0.00	0.05	0.00	0.04	0.00	0.03	0.00	0.00	0.58	0.06	0.00	0.09	0.00	0.04	0.00
mat	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.57	0.00	0.27	0.00	0.03	0.00
Ind	0.01	0.06	0.01	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.68	0.11	0.06	0.02	0.00
dat	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00	0.02	0.00
tim	0.01	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.83	0.03	0.00
Qua	0.10	0.05	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.13	0.01	0.66	0.02
Pro	0.04	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.02	0.23	0.01	0.19	0.44

Colour scale: 0.2, 0.4, 0.6, 0.8

Sell

Stop
Sup.

Option

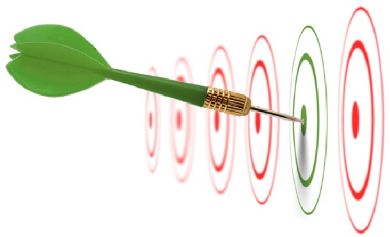
Ind.

Pro.

Empirical Study

Numeral Understanding in Financial Tweets for Fine-grained Crowd-based Forecasting

Comparable to Professional Analysts



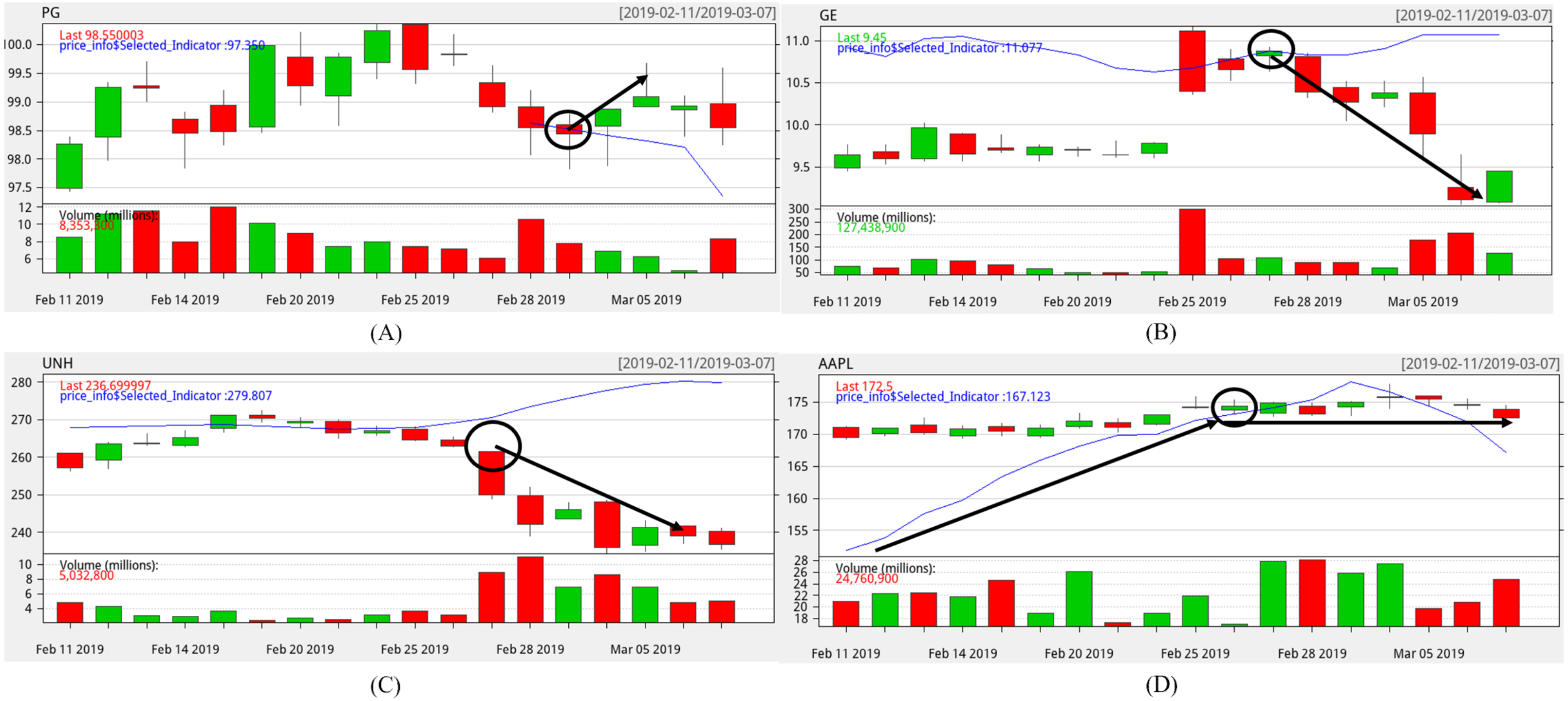
	Crowd	Analyst
Average difference	13.17%	6.75%
Achieving rate	67.03%	74.73%
Achieving duration	3.38 months	2.46 months
Average return	4.86%	2.93%

Progressive

	Crowd	Analyst
Winning ratio	68.13%	71.43%
Max profit	52.17%	17.23%
Max drawdown	-11.82%	-14.10%
Average profit	11.08%	6.42%
Average loss	-8.43%	-8.40%

Profitable

Crowd View: Converting Investors' Opinions into Indicators



- The indicators related to the analysis results of crowd investors (**support and resistance** price level) provide the incremental information for short-term (**3- and 5-day**) trading.
- The indicator constructed by **the cost of crowd investors** (buy-side and sell-side cost) furnish trader with additional long-term (**10-day**) information.

Further Research Directions

Numeracy-600K: Learning Numeracy for Detecting Exaggerated Information in Market Comments

- S&P 500 <.SPX> UP 1.53 POINTS, OR 0.08 PERCENT, AT ___ AFTER MARKET OPEN

- DOW JONES <.DJI> UP 8.70 POINTS, OR 0.05 PERCENT, AT ___ AFTER MARKET OPEN

- U.S. Q3 GDP rises _____ pct

Multilingual & Different Domain & Document Level

NLP Lab
National Taiwan University

近 5 年來京城銀行每股盈餘分別為 2.51、4.89、4.17、3.09、4.33 元，合併總損益分別為 28、56、47、36、51 億。除去年外，整體來說獲利向上，本年度光第一季就已經獲利 19 億，表現十分出色。

去年底，京城銀提列華映呆帳 16.4 億元，造成每股損失 1.42 元，也導致去年盈餘只剩 2.51 元，是近年來新低。若沒此呆帳，京城銀去年獲利應有 44 億。但據悉，京城銀行對華映六代廠有最高限額抵押權 18 億，透過該抵押 16 億債權可望全額受償，且已經獲得抵押物拍賣裁定，即將對華映六代廠做強制執行。也因此去年認列的呆帳損失可能回沖，進一步推升本年度獲利。若獲利與去年水準相同，有 44 億，在加上呆帳回沖的 16 億，一年就可獲利 60 億，每股盈餘突破 5 元。

【NQNロンドン】25日のフランクフルト株式市場で、ドイツ株式指数（DAX）は10営業日ぶりに反落した。終値は前日と比べて30.56ポイント（0.25%）安の12282.60だった。

オンライン決済サービスのワイヤーカードは、前日に大幅上昇した反動で、3%超下がった。ドイツ銀行の値下がりも目立った。同行とコメルツ銀行が3月から続けていた統合交渉を打ち切る見通しとなった。タイヤのコンチネンタルも売られた。一方で、第1四半期の決算を発表した医薬・農薬大手のバイエルは上昇した。鉄鋼のティッセン・クルップも買われた。

Cette révision à la baisse de la croissance fait suite au brusque ralentissement de l'économie canadienne à la fin de 2018, avec une progression de 0,4 % en rythme annuel au quatrième trimestre, et de chiffres décevants pour le début 2019.

Clinical

Geography

Cooperation

Next Step

FinNum-2: Numeral Attachment

NLP Lab
National Taiwan University

- $NE OK NE, last time oil was over $65 you were close to $8. Giddy-up…
- Given a target numeral and a cashtag, and we formulate the problem as a binary classification to tell if the given numeral is related to the given cashtag.
- Macro-F1 score is adopted for evaluating the experimental results.
- Baseline: CapsNet → Macro-F1 score: 67.14%

SIGIR 2019

NLP Lab

National Taiwan University