

FinNum-2: Numeral Attachment in Financial Tweets

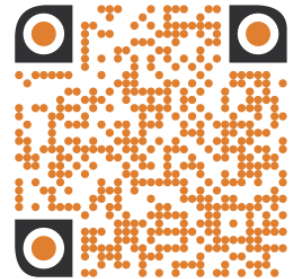


Chung-Chi Chen (National Taiwan University, Taiwan)

Hsin-Hsi Chen (National Taiwan University, Taiwan)

Hiroya Takamura (AIRC, AIST, Japan)

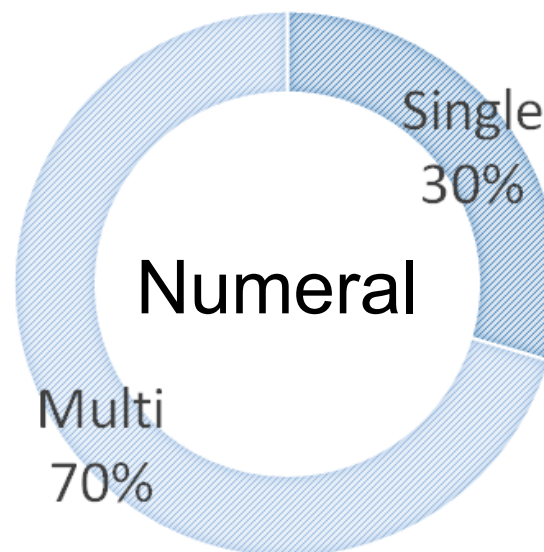
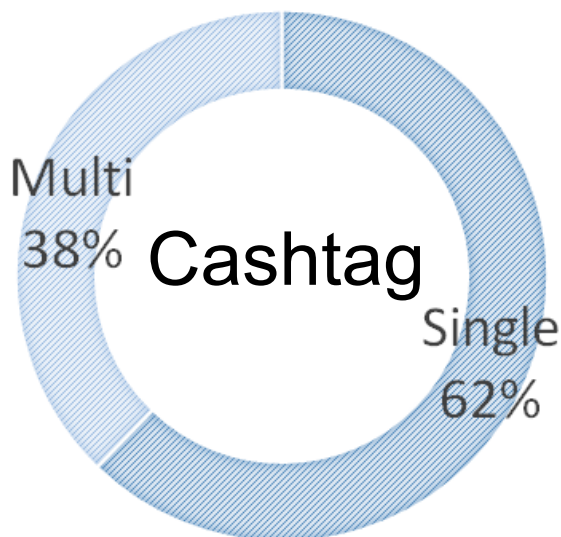
Hen-Hsen Huang (National Chengchi University, Taiwan)



<http://finnum.nlpfin.com>

Introduction

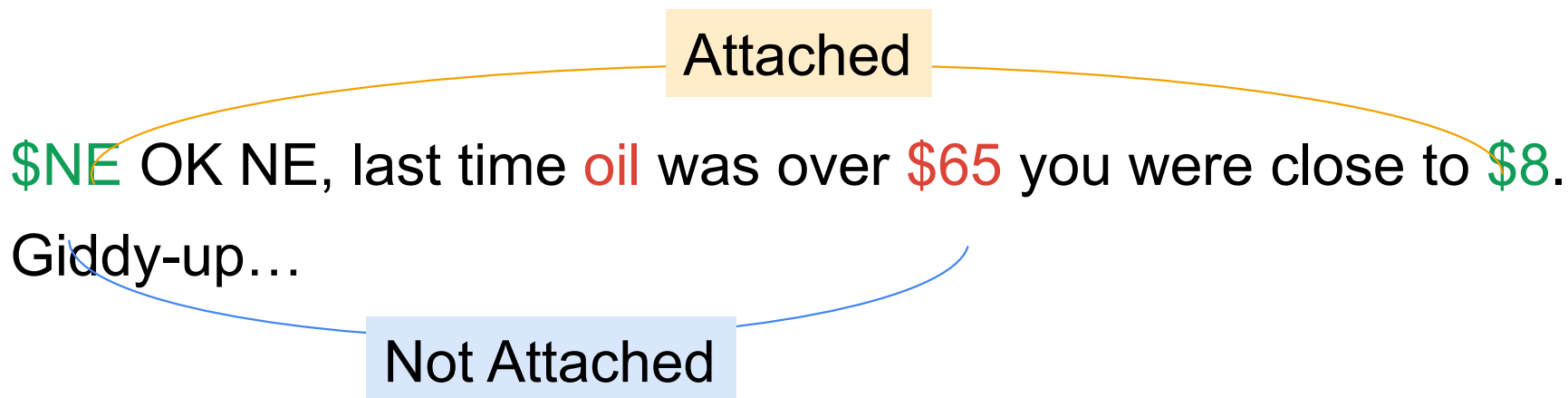
- FinNum-1: Fine-Grained Numeral Understanding in Financial Tweets
 - We focus on disambiguating the **meaning** of the numerals in financial tweets.
- However, in a financial tweet, there may have more than one cashtag and more than one numeral.
- **FinNum-2: Numeral Attachment in Financial Tweets**



Task Definition

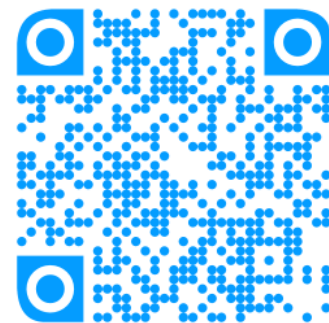
Numeral Attachment:

- Detect the attached target of the numeral
- Target = Cashtag
- Source: Financial Tweets



Dataset and Baseline

- Labeled by Experts
- Pilot Dataset
 - 4,847 unique tweets
 - 7,984 instances
 - Macro-F1: 67.14%
- Full Dataset (Pilot + Train + Test)
 - **In Progress: Checking the agreement**
 - About 10,000 unique tweets
 - 12,000-15,000 instances



<http://numattach.nlpfin.com>

Numeral Attachment with Auxiliary Tasks

Chung-Chi Chen¹, Hen-Hsen Huang^{2,3}, Hsin-Hsi Chen^{1,3}

¹Department of Computer Science and Information Engineering, National Taiwan University, Taiwan

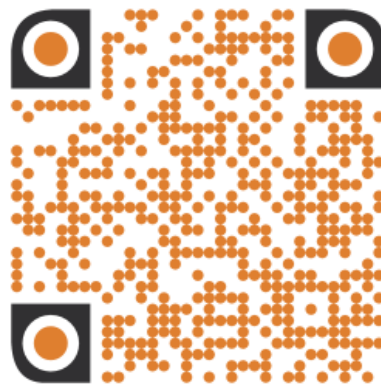
²Department of Computer Science, National Chengchi University, Taiwan

³MOST Joint Research Center for AI Technology and All Vista Healthcare, Taiwan

Important Dates (Tentative)

- 2019
 - Jun 15th : Pilot Dataset Release
 - Oct 20th : Full Training Set Release
 - Dec 15th : Registration Start

- 2020
 - Jun 30th : Registration Due
 - Jul 1st : Formal Run – Test Set Release
 - Jul 31st : Formal Run – Submission Due
 - Aug 15th : Formal Run – Evaluation Result Release
 - Sep 15th : Submission Due of Participants Papers
 - Oct 15th : Acceptance Notification & Task Overview Paper Release
 - Nov 1st : Camera-ready Participant Paper Due
 - Dec 5th-8th: NTCIR-15 Conference



<http://finnum.nlpfin.com>

- 7 teams have interest in FinNum-2
- 3 teams already signed the agreement for the pilot dataset

