FiQA - 2018

Financial Opinion Mining and Question Answering

Open Challenge - WWW 2018 LYON, FRANCE (23 - 27 April 2018)

FiQA - 2018

Summary

The growing maturity of Natural Language Processing (NLP) techniques and resources is drastically changing the landscape of many application domains which are dependent on the analysis of unstructured data at scale. The financial domain, with its dependency on the interpretation of multiple unstructured and structured data sources and with its demand for fast and comprehensive decision making is already emerging as a primary ground for the experimentation of NLP, Web Mining and Information Retrieval (IR) techniques. This challenge focuses on advancing the state-of-the-art of aspect-based sentiment analysis and opinion-based Question Answering for the financial domain.

Motivation

The increasing interest and investment around technologies which can support better financial analysis and decision making creates the demand for an increasing dialog between academia and industry. The specificity of the language use and its underlying conceptualizations in the financial and economic domains requires the creation of new fine-grained models and techniques which capture the particular semantic phenomena of this field.

This challenge aims to provide an experimentation and discussion ground for novel NLP approaches targeting the interpretation of financial data using the tasks of aspect-based sentiment analysis and opinionated Question Answering(QA) as motivational scenarios. The challenge aims at catalyzing theoretical and empirical discussions around principles, methods and resources focused on financial data. A special emphasis is given to multi-lingual and multiple data sources.

While previous tasks and challenges have focused on, multilingual document, message sentence or even entity level sentiment classification, no challenge that we are aware of attempts to analyse to the aspect level. Moreover such tasks often do not focus beyond the English language for such fine grained sentiment analysis. In addition, research in Question Answering(QA) from opinionated datasets is also under-explored.

- Topics of particular interest to be discussed and developed within the task include (but are not limited to):
- Aspect-oriented sentiment analysis and opinion mining.
- Aspect-identification extraction/classification for finance for opinion mining.
- Question Answering and opinion-based Question Answering over financial text.
- Multi-lingual sentiment analysis.
- Linguistic analysis tools for the financial domain, in particular financial social media (e.g. tokenisation, part-of-speech tagging, normalization, parsing)
- Sentiment classification on financial texts;

- Sentiment classification on financial texts;
- Analysing and understanding linguistic phenomena associated with financial text corpora (including the sub-language of financial microblogs);
- New semantic and ontological models for finance;
- Construction and application of distributional semantic models on finance;
- Lexical resources for the financial domain;

Challenge Timeline & Evaluation Criteria

The two tasks will be available to participating systems (participants can be involved in one of the tasks or both):

Task 1: Aspect-based financial sentiment analysis

Given a text instance in the financial domain (microblog message, news statement or headline) in English, detect the target aspects which are mentioned in the text (from a pre-defined list of aspect classes) and predict the sentiment score for each of the mentioned targets. Sentiment scores will be defined using continuous numeric values ranged from -1(negative) to 1(positive).

Systems will be evaluated with regard to aspect classification, sentiment classification and aspect-sentiment attachment. Participating systems will be evaluated with regard to precision, recall and F1-score for aspect classification approaches and regard to MSE and R Squared(R^2) metrics for sentiment prediction approaches.

An example of the input/output of the task is defined below:

"55": {

"sentence": "Tesco Abandons Video-Streaming Ambitions in Blinkbox Sale",

"info": [

{

"snippets": "['Video-Streaming Ambitions']",

"target": "Blinkbox",

"sentiment_score": "-0.195",

"aspects": "['Corporate/Stategy']"

},

{

"snippets": "['Tesco Abandons Video-Streaming Ambitions ']",

arget": "Tesco",

"sentiment_score": "-0.225"

"sentiment_score": "-0.335",
"aspects": "['Corporate/Stategy']"
}
]
}

Task 2: Opinion-based QA over financial data

Given a corpus of structured and unstructured text documents from different financial data sources in English (microblogs, reports, news) build a Question Answering system that answers natural language questions. For this challenge, part of the questions will be opinionated, targeting mined opinions and their respective entities, aspects, sentiment polarity and opinion holder.

The challenge takes both an Information Retrieval (IR) and a Question Answering (QA) perspective. Systems can rank relevant documents from the reference knowledge base with regard to a natural language question or generate their own answer. The document relevance is evaluated using 0 as non-relevant document score and 1, otherwise. The relevant score information is implicit if you consider the question-doc matches information contained in the training FiQA_question_doc data source. Participating systems will be evaluated using **F-score,** Normalized Discounted Cumulative Gain (NDCG) and Mean Reciprocal Rank(MRR) for the top 10 answers . The goal is to evaluate the ranking accuracy and the most-ranked answer using these two measures.

The QA test collection is built by crawling Stackexchange, Reddit and StockTwits.

An example about opinion question-answering matching is showed below:

"question": "Why are big companies like Apple or Google not included in the Dow Jones Industrial Average (DJIA) index?",

"answers":{

"290156": { "text":" That is a pretty exclusive club and for the most part they are not interested in highly volatile companies like Apple and Google. Sure, IBM is part of the DJIA, but that is about as stalwart as you can get these days. The typical profile for a DJIA stock would be one that pays fairly predictable dividends, has been around since money was invented, and are not going anywhere unless the apocalypse really happens this year. In summary, DJIA is the boring reliable company index." ," timestamp": "Sep 11 '12 at 0:53"}

}

Challenge Timeline

- lease of the training data : ~~December, 5th 2017~~
- Challenge test data published : ~~February 2018, 17th~~

- Submission of the results: February, 23th

FiQA - 2018

- Challenge papers paper submission deadline : ~~4 February 2018~~ February, 25th
- Challenge papers acceptance notification and gold standard release : February, 28th

How to participate?

Instructions on how to submit your results will be available soon. Participants should subscribe to the challenge's mailing list: https://groups.google.com/forum/#!forum/fiqa-2018 .

Participating systems should send their results as an attached JSON file via email to sousam02@gw.uni-passau.de.

The test data will be made available on February 12th, 2018.

The submissions should adhere to the format below:

- Task 1:

```
{
  'team': "name of the team",
 'paper': 'paper title',

 'results': [
    { 'id':'Num1',
     'snippet': 'snippet1'
      'aspect_categories': 'aspect11'
      'sentiment_scores': '0.2'
    }
   { 'id':'Num2',
     'snippet':'snippet2'
      'aspect_categories': 'aspect21'
      'sentiment_scores': '0.3'
   }
 ]
```

FiQA - 2018

- Task 2:

```
{
  'team': "name of the team",
  'paper': 'paper title',
  'results': [
      { 'question_id':'Num1',
        'ranked_answers_id': [ 'answerid10','answerid11',...,'answerid19']
      }
     { 'question_id':'Num2',
        'ranked_answers_id': [ 'answerid20','answerid21',..,'answerid29']
     }
     ...
  ]
}
```

Where the most ranked answer most be in descending order

Papers Submissions

Paper submissions are invited independently of having an associated participating system in the challenge. Submissions should be formated using the ACM Sigconf template (as for the main conference track). Long papers have a page limit of 6 pages while short papers have a 2 page limit.

Submissions should be sent via EasyChair. CEUR open online proceedings will be published after the challenge.

Submissions will be closed on February 14th midnight, Hawaii time. A confirmation email will be send to the participants.

Train Data

Task 1: Aspect-based financial sentiment analysis

~~The training data source to Aspect-based financial sentiment analysis is available here.~~

The final Aspect-based financial sentiment analysis train set version for post and headlines is available

here

FiQA - 2018

Task 2: Opinion-based QA over financial data

The training data source to Opinion-based QA task is available here.

The training data is available only for non-commercial use.

Test Data

Task 1: Aspect-based financial sentiment analysis

The final Aspect-based financial sentiment analysis test set version for post and headlines is available here.

Task 2: Opinion-based QA over financial data

The testing data source to Opinion-based QA task is available here.

The testing data is available only for non-commercial use.

Organizers

Macedo Maia, Department of Computer Science and Mathematics, University of Passau.

André Freitas, School of Computer Science, University of Manchester.

Alexandra Balahur, Joint Research Centre , European Commission(EC)

Siegfried Handschuh, Department of Computer Science and Mathematics, University of Passau.

Manel Zarrouk, Insight Centre for Data Analytics, National University of Ireland, Galway.

Ross McDermott, Insight Centre for Data Analytics, National University of Ireland, Galway.

Brian Davis, Department of Computer Science, Maynooth University.

Contact

sousam02 - at - gw.uni-passau.de

Acknowledgements

This challenge is sponsored by the SSIX Horizon 2020 project (grant agreement No 645425) (https://ssix-project.eu/).


FiQA - 2018
European
Commission