Stance detection on Twitter

Lixing Zhu

Introduction

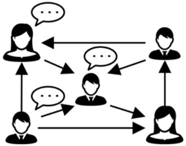
- Stance detection aims at classifying users' stances towards topics

Tweets

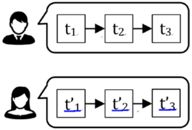
Users

Timestamps

User impacts



User behaviors



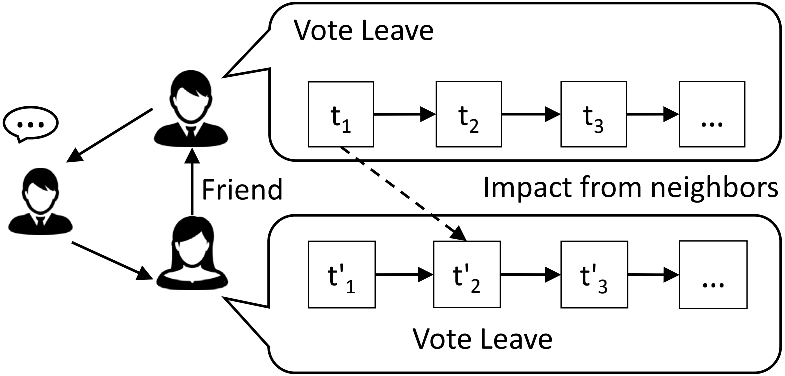
Public Opinion Detection

Market Analysis

Wide range of applications

Introduction

- Dynamic Stance Prediction: To predict a user's stance label of the next tweet within a short period of time.



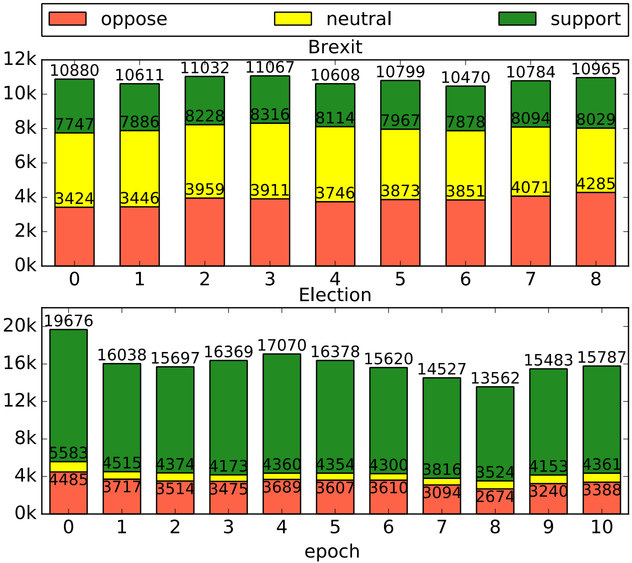
- Users are influenced by others.
- Users pay more attention to their interested topics.
- Users opinions are changing dynamically.

Tweet Datasets

- Brexit
  - Tweets were crawled in June 2016 using hashtags: #EURef, #EU, #Referendum, #Brexit, #VoteRemain, #VoteLeave, etc.
  - Tweets are split into epochs every other day, which resulted in a total of 9 epochs with each epoch consisting of 40,440 tweets on average
- US General Election 2016
  - Tweets were crawled in November 2016 using keywords: Trump, Clinton and Hillary, etc.
  - A total of 452,128 tweets split into 11 count-based epochs with the epoch size set to 40,000 and on average 16,019 users per epoch

https://developer.twitter.com/en/docs/basics/getting-started

Tweet Datasets



- The Brexit dataset has on average 10,802 users per epoch. Election dataset has 16,019 users in each epoch.
- In the Brexit dataset, each user on average posted 3.7 tweets. In the Election dataset the number is 2.5.

Tweet Datasets

WARWICK
THE UNIVERSITY OF WARWICK

```
>>>>>>>>
1
--------
its fantastic , at last people ignoring their party loyalties : it
kill the traitorous scots ! just kidding ? <hashtag> brexit </hash
cracking work mr carswell that is bloody great ! you shared our se
<hashtag> eu ref </hashtag> watch bbc2 britain europe the immigrat
well done the <hashtag> rmt </hashtag> union recommending <hashtag
eu remainie has informed me we can control our borders we know but
--------
<hashtag> brexit </hashtag> save the uk . the dawn o / t revival o
i hope and pray that the people of the <hashtag> uk </hashtag> wil
lord bamford ' s voting leave eu . he runs a global business knows
bigot with no respect for the dead / family or indeed integrity <h
still too close to call , all to play for ! <hashtag> brexit </has
<hashtag> brexit </hashtag> is necessary to protect <hashtag> nhs
--------
this is the reason our pm has cut our military to pieces . he was
sorry but fullfact had to acknowledge eu makes over <percent> of u
an open letter on why we should vote leave on <date> __full_zero_w
<hashtag> remain </hashtag> what ' s with the xenophobic comments
looking ahead to <hashtag> eu ref </hashtag> night : where are the
vote_leave : iankatz1000 : watch senior erdogan advisor ilnur chev
```

`+0000 2016-06-14 20:15:40`

- Each line consists of serial of user's tweets.

`+0000 2016-06-20 20:56:47`

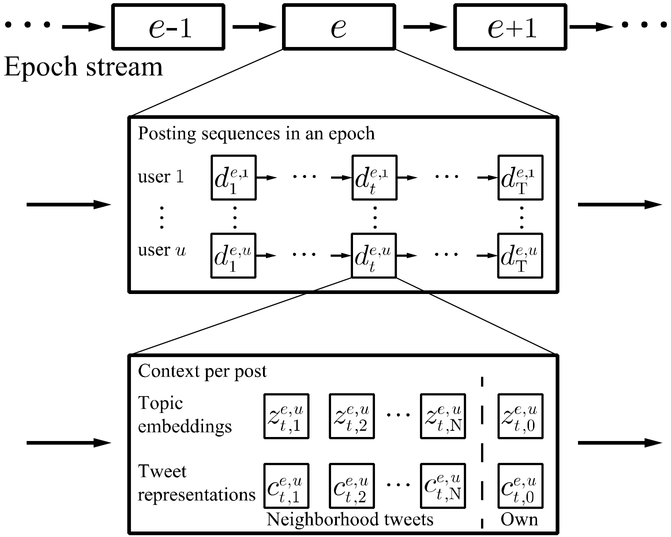
- Each user's tweet comprises neighbors.

`+0000 2016-06-21 22:59:54`

Neural Opinion Dynamics (NOD) model

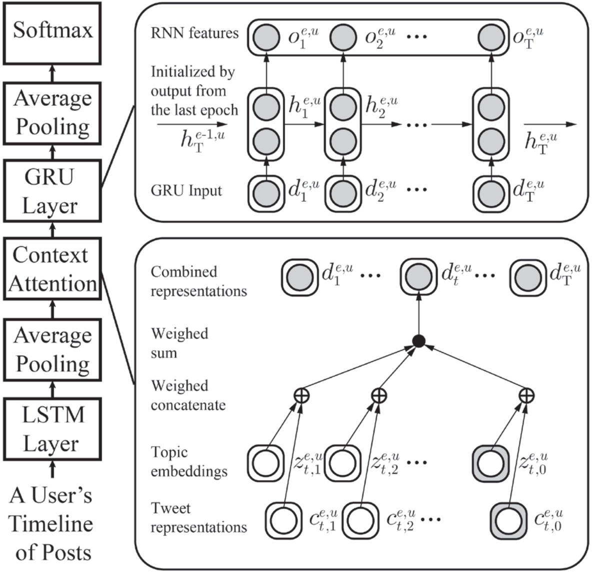
- NOD leverages content and context information for the tracking of user-level stance dynamics over time
- We assume tweets arrived in a temporal order and can be split into epochs
- In each epoch, a user posted a sequence of tweets
  - Each tweet has a representation derived from its content and the associated topic embedding
  - In addition, we assume that when user posts a tweet at epoch *e*, their opinion is also influenced by the most recent N tweets posted by their neighbours in their social network

Problem Setup



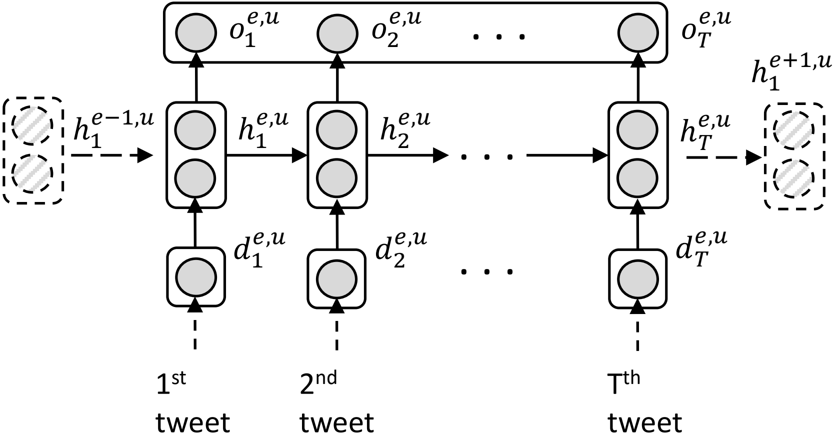
- Dataset has been split into epochs.
- Each epoch consists of user timelines where each step is a tweet.
- Each tweet is represented by its own content and neighbourhood context.

Model Architecture



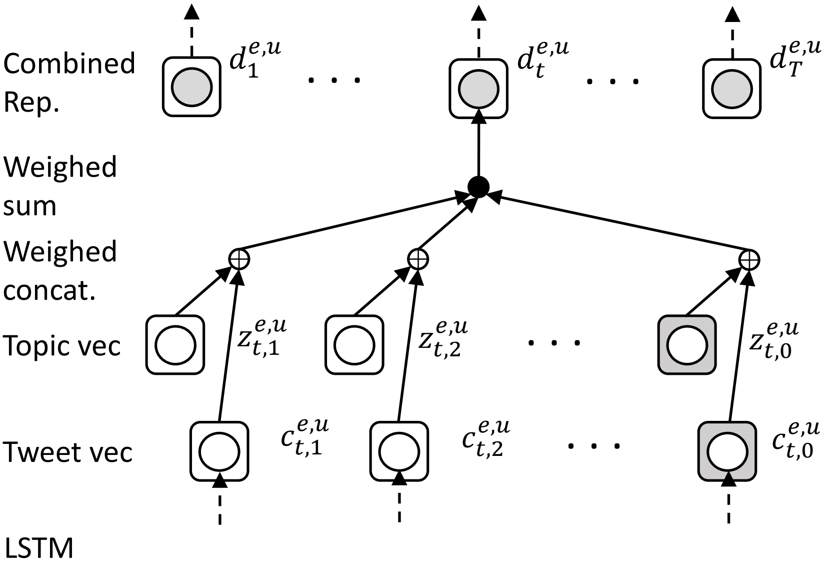
- Topic embeddings have already been discovered by an independent topic model.
- Tweet representation is a vector pooled over an LSTM.

Model Components



- Users sequential behaviors are modelled by a GRU.
- The GRU state are initialized by the last GRU state from the previous epoch.

Model Components



- Features are extracted from the user's tweet, neighbors' tweets and tweet-associated topics.

Model Components

- Topic embedding
  - Extract hierarchical topics from tweets using Hierarchical Latent Dirichllet Allocation (HLDA)
  - Convert each topic into topic embeddings
- Context attention - when a user posts a tweet, they would pay more attention to their neighbours' tweets carrying topics of their interests
  - For each tweet, its final representation is generated by combining its content representation with its corresponding topic embedding by:

$$g_{t,n}^{e,u} = \frac{\alpha_1 z_{t,n}^{e,u} \oplus \alpha_2 c_{t,n}^{e,u}}{\alpha_1 + \alpha_2}$$

Model Components

- Use β_n to measure the degree of influence from the nth neighbourhood tweet and β_0 is the attention signal on the user's current tweet

$$d_t^{e,u} = \sum_{n=0}^{\mathrm{N}} \beta_n g_{t,n}^{e,u}$$

$$\beta_n = \frac{\exp(\mathbf{v}^{\mathrm{T}} u_{t,n}^{e,u})}{\sum_{n=0}^{\mathrm{N}} \exp(\mathbf{v}^{\mathrm{T}} u_{t,n}^{e,u})}$$

$$u_{t,n}^{e,u} = \tanh(\mathbf{W} g_{t,n}^{e,u} + \mathbf{b})$$

- Self attention – The weight of g_n, that is β_n, is determined by g_n themselves

$$\mathbf{d} = \mathrm{softmax}\left(\mathbf{v}^{\mathrm{T}} \tanh(\mathrm{W}\quad)\right)\quad{}^{\mathrm{T}},$$

where $\mathrm{G}=[g_1, g_2, \ldots, g_N]$

Model Components

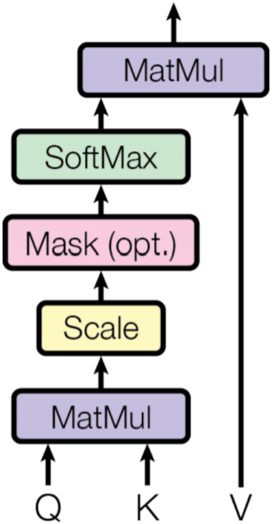
- Compared to Query-Key-Value Attention

$$\mathbf{d} = \text{softmax}\left(\cancel{W_Q Q}\,\tanh(W_K \cancel{K})\right)\cancel{W_V} V^T$$

$$\Downarrow$$

$$\mathbf{d} = \text{softmax}\left(\tanh\left(\mathbf{v}^T \quad\right)\right) \quad {}^T$$

or $$\mathbf{d} = \text{softmax}\left(\mathbf{v}^T \tanh(W \quad)\right) \quad {}^T$$



This figure is from paper: https://arxiv.org/pdf/1706.03762.pdf

User-Level Topic-Stance Prediction

- The integrated representations of user u's post sequence in epoch e is fed to a GRU layer for user-level topic-stance prediction, output is a $\mathrm{C} = \mathrm{K} \times \mathrm{S}$ vector $y^{e,u}$

- Objective function is to minimise the KL divergence between the predicted distribution over topic-stance categories and the ground truth distribution over topic-stance categories

$$\mathcal{L} = \sum_{u=1}^{\mathrm{U}} \sum_{i=1}^{\mathrm{C}} \mathbf{KL}(y_i^{e,u} || \mathrm{g}_i^{e,u})$$

- Alternatively, we can perform coarse-level stance classfication in which the output is a three-class stance label (\`oppose', \`neutral' or \`support') with topics ignored
  - Use cross-entropy loss instead
  - Output is an S dimensional vector, ground truth is a label

Experimental Setup

- Ground truth stance label acquisition – trained supervised classifier with distant supervision
  - Brexit – collect over 4 million tweets between May 16th and June 2nd 2016 with hashtags clearly indicating stances as training data
  - Election – collect over 17 million tweets in the first week of November 2016 with stance-indicative hashtags
- Results on manually annotated 1,000 tweets:

Table 3: Accuracy of the ground truth acquisition methods.

	Brexit	Election
DataStories	0.906	0.895
Sentiment140Lex	0.579	0.562
Vader	0.538	0.481

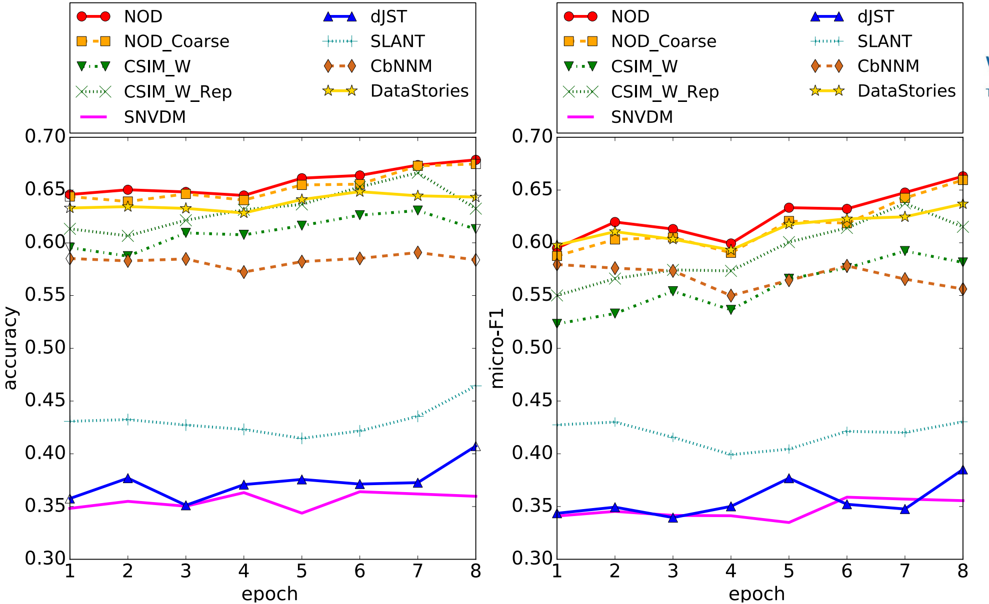
Experimental Setup

- Obtain the user-level topic-stance distributions
  - For each user, calculate the number of tweets under each topic with different stance labels and normalize the counts to obtain the topic-stance distributions
- Update the model with the data in the current epoch and use the trained model to predict the stance labels in the next epoch

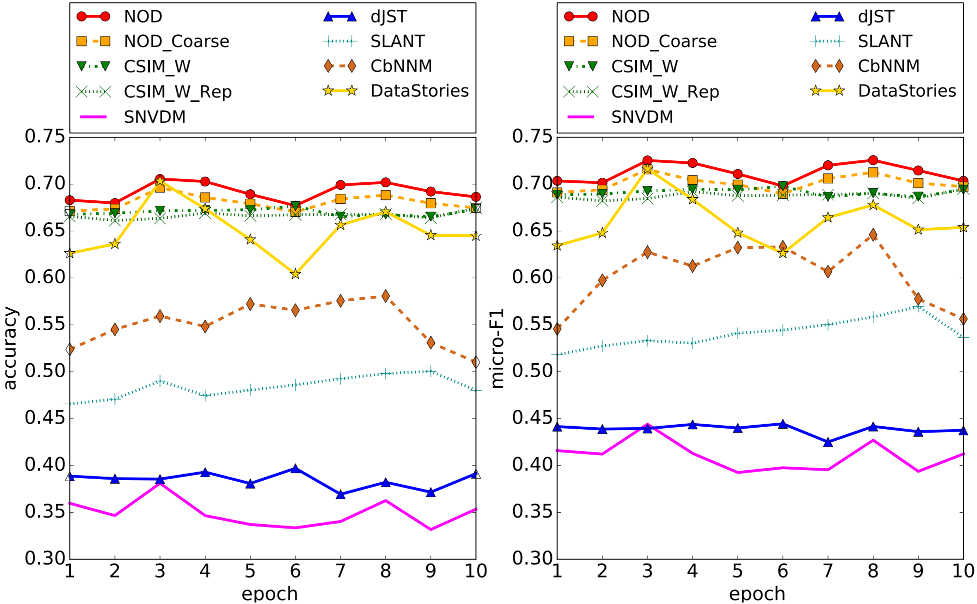
Baselines

- dJST (He et al., 2013) – weakly-supervised LDA-based generative model for dynamic sentiment-topic detection.
- SLANT (De et al., 2016) – a supervised probabilistic generative model which models each user's latent opinions over time as a multidimensional stochastic process.
- CSIM-W (Chen et al., 2018) – used an attention layer to weigh the importance of a given user's previously published tweets, their current tweet and their neighbors' tweets and employed an LSTM layer to capture the influence in the previous epochs.
- SNVDM (Thonet et al., 2017) – an unsupervised LDA-based generative model where the sender/receiver information is regarded as observed variables, which is generated by a hidden viewpoint variable
- CbNNM (Ren et al., 2016) – the contextual information of a tweet (i.e., the neighbours' tweets sharing the same hashtag) serve as features for tweet-level stance classification.
- DataStories (Baziotis et al., 2017) – the state-of-art method in tweet-level stance classfication.

Experiment Results – Brexit



Experiment Results – US Election



Experiment Results – Topical Stance

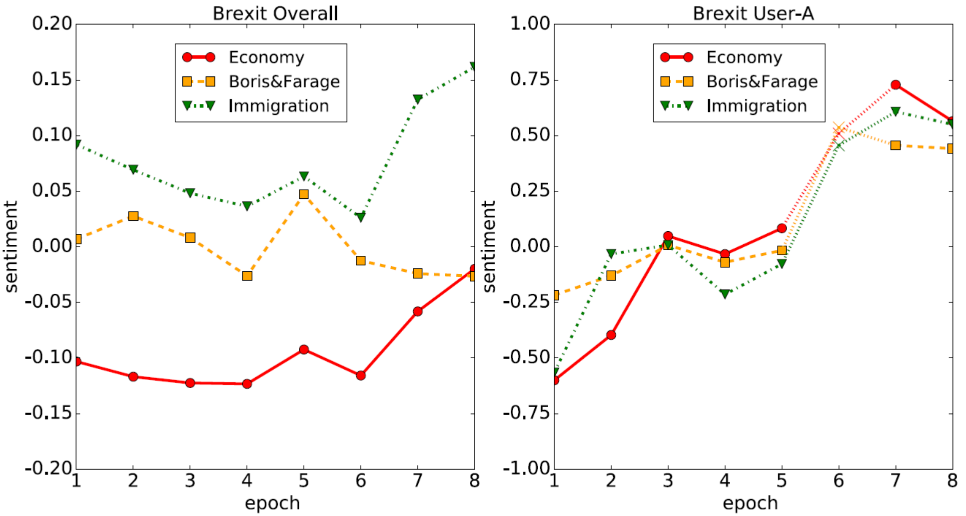


Figure 6: Global topic-stance and user-level topic-stance on BREXIT.

Conclusion

- Neural Opinion Dynamics (NOD) model for user-level topic-stance prediction
  - take into account tweet content, topical and neighborhood context

- Future work:
  - The neighborhood information was derived from the following-follower relations. It is also possible to construct the social networks using the re-tweeting or mentioning relations.
  - Consider each neighbour's social influence score that opinions from more influential users should carry higher weight.
  - Investigate a unified model for joint topic-stance detection over streaming data.