Inferring Power Relations from Social Interactions

Guna Prasaad , Mehul Goyal, Nisheeth Lahoti, and Raghav Gupta

Problem

Given a short message within a community (online or offline), predict whether the sender is senior to the receiver (DownSpeak) or the other way round (UpSpeak).

Useful in law enforcement and intelligence and even online marketing : who’s the big fish to be held to ransom? To be targeted for ads? Who’s who in a terrorist organization?

Data

Enron corpus

- Set of emails from within Enron’s company email system
- Obtained corpus from Apoorv et al; technical issues with MongoDB
- Going with 591 mails extracted using list of employees (with corporate ranking)

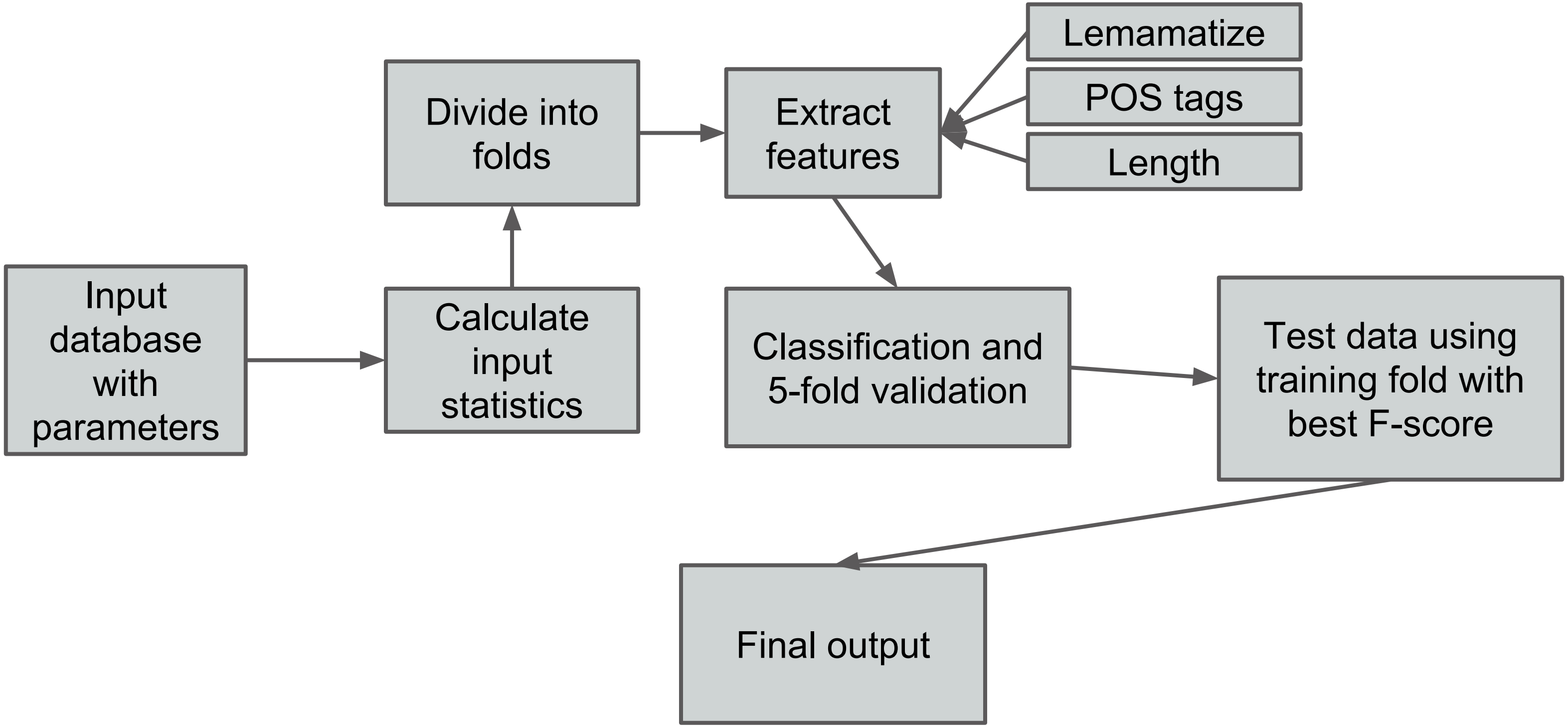
Corpus Statistics

	UpSpeak	DownSpeak	Overall
#Emails	269	322	591
#Sentences	2225	2161	4386
#Words	43400	43117	86517
Avg. Sentences/email	8.27	6.71	7.42
Standard Deviation Sentences/email	9.96	10.66	10.38
Avg. Words/email	161.33	133.90	146.39
Standard Deviation Words/email	206.122	221.53	215.089

Features

- N-grams
  - Word, POS, mixed
  - N = 1, 2, 3
  - Lemmatization or no lemmatization
- Length of email:
  - Number of sentences
  - Average number of words per sentence

System Modules



Classification Statistics

- Confusion Matrix (word, POS, mixed unigrams, presence)

predict \ actual	Downspeak	Upspeak
Downspeak	51	5
Upspeak	14	49

- Accuracy
  - Upspeak - 90.4%
  - Downspeak - 78.46%
  - Overall - 84.03%

Classification Statistics

- Confusion Matrix (word, POS, mixed unigrams, frequency)

predict \ actual	Downspeak	Upspeak
Downspeak	53	8
Upspeak	12	46

- Accuracy
  - Upspeak - 85.19%
  - Downspeak - 81.53%
  - Overall - 83.19%

Ablation Test Results

- Effect of lemmatization
- No effect of a discourse connector presence feature

F scores for 0 and 1 classes	Word unigrams (presence)	Word unigrams + bigrams (frequency)
With lemmatization	0.81, 0.79	0.82, 0.79
Without lemmatization	0.84, 0.83	0.85, 0.82

- Effect of mixed N-grams

F scores for 0 and 1 classes (frequency)	Mixed unigrams	Word unigrams
	0.84, 0.83	0.86, 0.84

Ablation Test Results

- Effect of word bigrams and trigrams
- Presence vs frequency features

F scores for 0 and 1 classes (frequency)	Unigrams	Unigrams + Bigrams	Unigrams + Bigrams + Trigrams
Word n-grams (no lemmatization)	0.86, 0.84	0.84, 0.82	0.87, 0.84

F scores for 0 and 1 classes (frequency)	Word unigrams	Word + POS unigrams + bigrams	Mixed unigrams
Presence	0.84, 0.83	0.83, 0.83	0.84, 0.84
Frequency	0.85, 0.83	0.86, 0.85	0.48, 0.31

Ablation Test Results

- Effect of POS tagged input

F scores for 0 and 1 classes	Word + POS unigrams	Word + POS unigrams + bigrams
Presence	0.90, 0.78	0.83, 0.83
Frequency	0.86, 0.85	0.86, 0.85

- Effect of length feature

F scores for 0 and 1 classes	Word unigrams with length	Word unigrams
Word unigrams	0.87, 0.85	0.86, 0.84

Error Analysis

Misclassified as UpSpeak

Please re-arrange your schedules to attend a meeting with Stan Horton today at 10:45am in 49C3.   Sorry for the short notice.

Cindy Stark

Executive Assistant to Stan Horton

--

attached is the info for JP Morgan. Let me know if you have questions.

kh

--

< Long, descriptive emails from bosses >

Error Analysis

Misclassified as DownSpeak

The ST ercot book is long 30 mw's in the off peak hours for Monday. We would like to sell these mw's at a reasonable price that is left to your judgement.

If prices look to be potentially negative, of course, sell the mw's.

Give me a call if you have questions, preferably before 10 pm.

Thanks,

JMF

--

Fellow students,,,,here is the format I used ...... It doesn't have all the info yet, but this should work. So, just delete my data and fill in yours......

The KID!

Change of Test Domain

We tried the trained classifier on a different test domain : CS626 Moodle emails, to check if such an approach is portable across domains.

It is not. We reported 40% accuracy overall, largely due to out of vocabulary words and the length feature not holding importance.

Drawbacks and Future Extensions

- Poor test domain performance (company to university)
- Problems with named entities
- Need to be reliant on features portable to other languages
  - Relationship with politeness markers

References

Vinodkumar Prabhakaran, Owen Rambow, and Mona Diab. 2012. Predicting overt display of power in written dialogs. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (NAACL HLT '12).

Philip Bramsen, Martha Escobar-Molano, Ami Patel, and Rafael Alonso. 2011. Extracting social power relationships from natural language. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1* (HLT '11), Vol. 1.