Tweet Argument Summarization

Marcus Hughes
5/11/17 NLP CS 375
Williams College

Can we improve tweet summarization using argument mining ideas for classification?

Looks like our sense of conscience is somewhat twistet... #abortion #PlannedParenthood #ProFamilia #defundPP https://t.co/KOZO7gLv9v

Good Morning Patriots!! Happy #TGIF! If you haven't heard, we are WINNING the fight for #Prolife. Keep Praying for the unborn & to #DefundPP

"WTF? These monsters should welcome an investigation (of nothing is being done wrong) #DefundPP NOW! https://t.co/qU4XRyLEnX

My favorite #prolife follows: @Liberty_Pike @LifeNewsHQ @LilaGraceRose @SBAList #DefundPP #AllLivesMatter #HandsUpDontCrush"

"@DrottM: #DefundPP #Life begins at conception. https://t.co/UoXIMerIM7\\u2026 https://t.co/a329TM5jok\
" #NoHillary2016 #ChooseLife #UrMotherDid!

This girl didn't #ShoutYourAbortion and a lullaby to her baby gave her the career she always wanted https://t.co/ZveevxWw81 #DefundPP

It's meant to stop forcing taxpayers to fund an organization that kills babies. #DefundPP https://t.co/8RNPByoq0N

"#PepsiPerfect Is this the one made from aborted human cells used in the development of artificial flavor enhancers? #defundPP

stevenlaststraw: RT catwahler: RT LeahR77 Abortion Has NEVER Been About CHOICE #DefundPP #PJNET #TCOT https://t.co/4BS4ni0bfi

Tell your Senators to Defund Planned Parenthood SIGN & RT #DefundPP https://t.co/aKvL3tuSEP

Planned Parenthood Supporters Funnel Cash to Local School Elections https://t.co/kw1Mwsks5e Baby killers pay for protxn w/our $. #DefundPP

#DefundPP better yet close it down. Plenty of other clinics for women. https://t.co/RopOXtfgHX

Looks like our sense of conscience is somewhat twistet... #abortion #PlannedParenthood #ProFamilia #defundPP https://t.co/KOZO7gLv9v

Good Morning Patriots!! Happy #TGIF! If you haven't heard, we are WINNING the fight for #Prolife. Keep Praying for the unborn & to #DefundPP

"WTF? These monsters should welcome an investigation (of nothing is being done wrong) #DefundPP NOW! https://t.co/qU4XRyLEnX

My favorite #prolife follows: @Liberty_Pike @LifeNewsHQ @LilaGraceRose @SBAList #DefundPP #AllLivesMatter #HandsUpDontCrush"

"@DrottM: #DefundPP #Life begins at conception. https://t.co/UoXlMerlM7\\u2026 https://t.co/a329TM5jok\ " #NoHillary2016 #ChooseLife #UrMotherDid!

This girl didn't #ShoutYourAbortion and a lullaby to her baby gave her the career she always wanted https://t.co/ZveevxWw81 #DefundPP

It's meant to stop forcing taxpayers to fund an organization that kills babies. #DefundPP https://t.co/8RNPByoq0N

"#PepsiPerfect Is this the one made from aborted human cells used in the development of artificial flavor enhancers? #defundPP

stevenlaststraw: RT catwahler: RT LeahR77 Abortion Has NEVER Been About CHOICE #DefundPP #PJNET #TCOT https://t.co/4BS4ni0bfi

Tell your Senators to Defund Planned Parenthood SIGN & RT #DefundPP https://t.co/aKvL3tuSEP

Planned Parenthood Supporters Funnel Cash to Local School Elections https://t.co/kw1Mwsks5e Baby killers pay for protxn w/our $. #DefundPP

#DefundPP better yet close it down. Plenty of other clinics for women. https://t.co/RopOXtfgHX

Looks like our sense of conscience is somewhat twistet... #abortion #PlannedParenthood #ProFamilia #defundPP https://t.co/KOZO7gLv9v

Good Morning Patriots!! Happy #TGIF! If you haven't heard, we are WINNING the fight for #Prolife. Keep Praying for the unborn & to #DefundPP

"WTF? These monsters should welcome an investigation (of nothing is being done wrong) #DefundPP NOW! https://t.co/qU4XRyLEnX

My favorite #prolife follows: @Liberty_Pike @LifeNewsHQ @LilaGraceRose @SBAList #DefundPP #AllLivesMatter #HandsUpDontCrush"

"@DrottM: #DefundPP #Life begins at conception. https://t.co/UoXlMerlM7\\u2026 https://t.co/a329TM5jok\
" #NoHillary2016 #ChooseLife #UrMotherDid!

This girl didn't #ShoutYourAbortion and a lullaby to her baby gave her the career she always wanted https://t.co/ZyeevxWw81 #DefundPP

It's meant to stop forcing taxpayers to fund an organization that kills babies. #DefundPP https://t.co/8RNPByoq0N

"#PepsiPerfect Is this the one made from aborted human cells used in the development of artificial flavor enhancers? #defundPP

stevenlaststraw: RT catwahler: RT LeahR77 Abortion Has NEVER Been About CHOICE #DefundPP #PJNET #TCOT https://t.co/4BS4pi0tfi

Tell your Senators to Defund Planned Parenthood SIGN & RT #DefundPP https://t.co/aKvL3tuSEP

Planned Parenthood Supporters Funnel Cash to Local School Elections https://t.co/kw1Mwsks5e Baby killers pay for protxn w/our $. #DefundPP

#DefundPP better yet close it down. Plenty of other clinics for women. https://t.co/RopOXtfgHX

Very short statements
Spelling errors
Usage of *nontraditional* POS
Usage of abbreviations
Interlinked meaning with other hashtags
Way too many to look at by hand

Looks like our sense of conscience is somewhat twistet... #abortion #PlannedParenthood #ProFamilia #defundPP https://t.co/KOZO7gLv9v

Good Morning Patriots!! Happy #TGIF! If you haven't heard, we are WINNING the fight for #Prolife. Keep Praying for the unborn & to #DefundPP

"WTF? These monsters should welcome an investigation (of nothing is being done wrong) #DefundPP NOW! https://t.co/qU4XRyLEnX

My favorite #prolife follows: @Liberty_Pike @LifeNewsHQ @LilaGraceRose @SBAList #DefundPP #AllLivesMatter #HandsUpDontCrush"

"@DrottM: #DefundPP #Life begins at conception...
CoXiMgnM7Ad2928

" #NoHillary2016 #ChooseLife

This girl didn't #ShoutYourAbortion and a lullaby to her baby gave her the career she always wanted https://t.co/ZyeevxWw81 #DefundPP

It's meant to stop forcing taxpayers to fund an organization babies. #DefundPP https://t.co/8RNPByoq0N

made from aborted human cells used in the development of artificial flavor enhancers? #defundPP

stevenlaststraw: RT catwahler: RT abortion Has NEVER Been About CHOICE #DefundPP #PJNET #TCOT https://t.co/tBS4p0t.f

Tell your Senators to Defund Planned Parenthood SIGN & RT #DefundPP https://t.co/aKvL3tuSEP

Planned Parenthood Supporters Funnel Cash to Local School Elections https://t.co/kw1Mwsks5e Baby killers pay for protxn w/our $. #DefundPP

#DefundPP better yet close it down. Plenty of other clinics for women. https://t.co/RopOXtfgHX

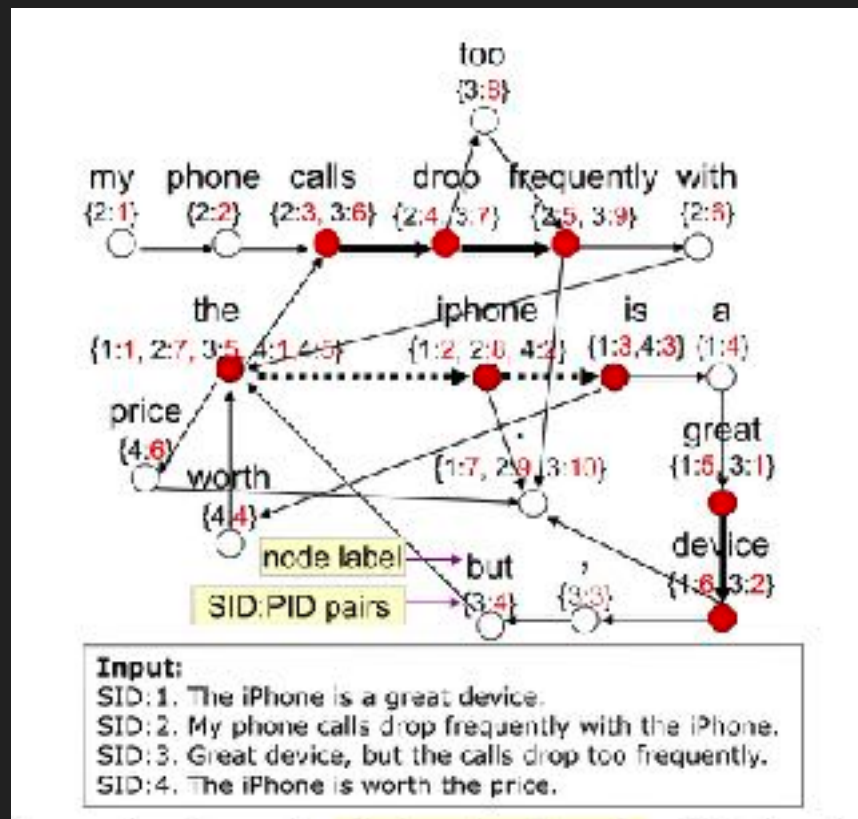# Summarization Approaches

1.Extractive

- "Abortion is the intentional killing of an unborn baby. And it doesn't take religion to know that killing an unborn baby is wrong. #DefundPP" EV_Black: RT jstines3: Nothing in MY Constitution says FED GOVT can fund infant killing! \n\n#DefundPP #PJNET #TCOT \\u2026 https://t.co/c9ou7qELVi" BC you R a Heartless Baby Killing Bitch! Go get a Job at #DefundPP https://t.co/PFlEw79ssN"

2.Abstractive

- Nothing in my constitution says fed govt can fund infant killing. Funding death of unborn is not role of fed govt. Birth control is free and readily available. Cruz attacked for stating truth & attempting to t_url #cruzcrew #pjnet #trusted .

Summarization algorithms

- **Baseline**: random, most recent

- **Frequency based**: SumBasic, Hybrid TF-IDF

- **Graph based**: Opinosis, LexRank, TextRank, Phrase reinforcement

- **Cluster based**: k-means clustering

- **Other**: decay models, neural networks

Ganesan, K., Zhai, C., & Han, J. (2010).

Evaluation

- Human evaluation
  - Readability, Informativeness, Redundancy, Grammaticality
- Automatic evaluation
  - ROUGE-N, ROUGE-L, ROUGE-SU4
  - BLEU
  - Pyramid method
  - Precision, recall, and F-measure

Summary One
So Instagram can sell your pictures to advertisers without u knowing starting January 16th I'm bout to delete my instagram! Instagram debuts new privacy policy , set to share user data with Facebook

Summary Two
Instagram changed their policies to state that they will have the right to sell your photos to advertisers beginning January 16th. You can only opt out by deleting your account.

Summary Three
Instagram will have the rights to sell your photos to Advertisers as of jan 16. Over for Instagram on January 16th - Instagram says it now has the right to sell your photos unless you delete your account by January 16th http://t.co/tsjic6yA
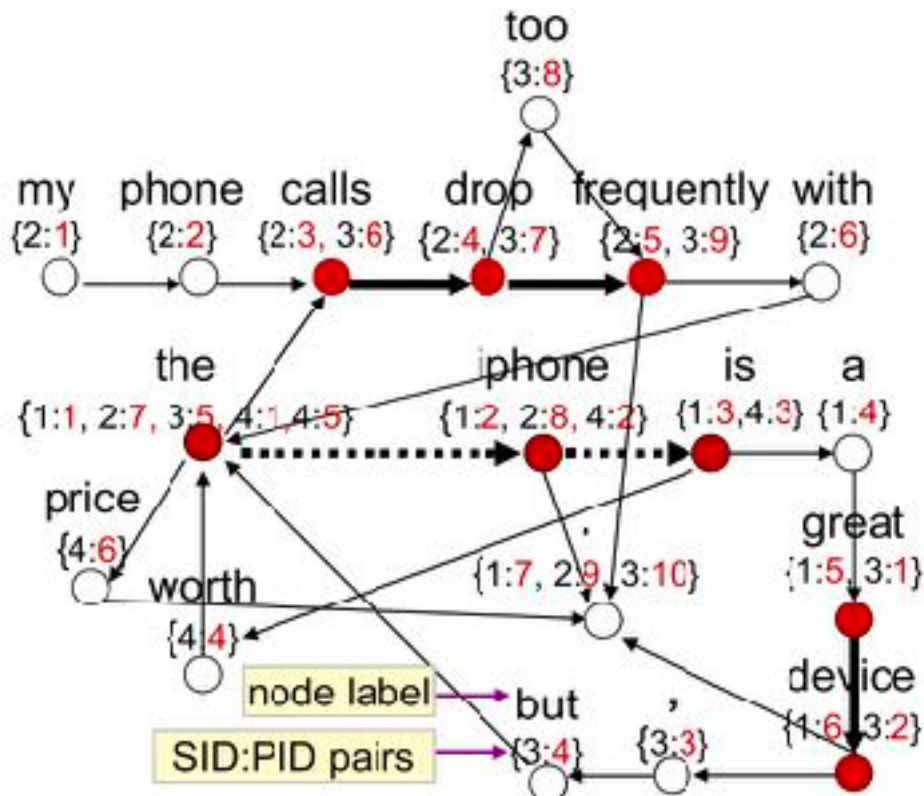
Summary 1 and 3 from Xu, Grishman, Meyers, and Ritter (2013)

Apply to Tweets?

TABLE III
AVERAGE VALUES OF F-MEASURE, RECALL AND PRECISION ORDERED BY F-MEASURE.

	F-measure	Recall	Precision
LexRank	0.2027	0.1894	0.2333
Random	0.2071	0.2283	0.1967
Mead	0.2204	0.3050	0.1771
Manual	0.2252	0.2320	0.2320
Cluster	0.2310	0.2554	0.2180
TextRank	0.2328	0.3053	0.1954
MostRecent	0.2329	0.2463	0.2253
Hybrid TF-IDF	0.2524	0.2666	0.2499
SumBasic	0.2544	0.3274	0.2127

Inouye, D., & Kalita+, J. K. (2011). Comparing Twitter Summarization Algorithms for Multiple Post Summaries. 2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing, 298–306. https://doi.org/10.1109/PASSAT/SocialCom.2011.31
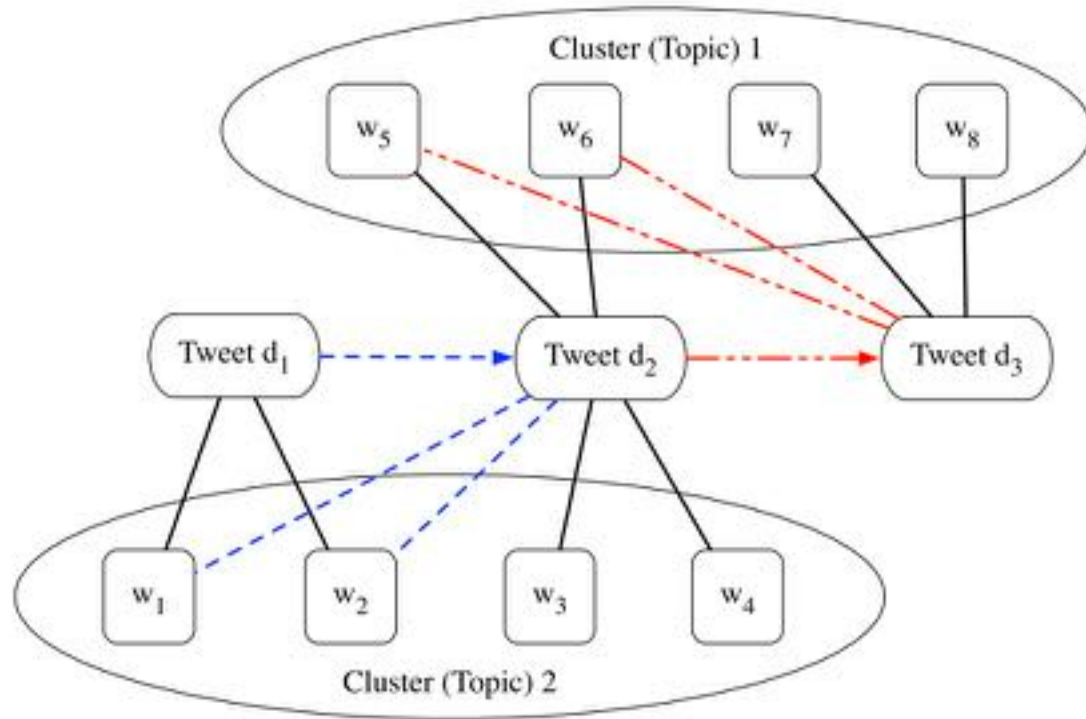
Input:
SID:1. The iPhone is a great device.
SID:2. My phone calls drop frequently with the iPhone.
SID:3. Great device, but the calls drop too frequently.
SID:4. The iPhone is worth the price.

Freddy Chong, Tat Chua, S. A. (2013). Automatic Summarization of Events From Social Media. International AAAI Conference on Weblogs and Social Media, 81–90. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.309.517

Nenkova, A., & Vanderwende, L. (2005). The impact of frequency on summarization. Microsoft Research Redmond Washington Tech Rep MSRTR2005101. Retrieved from http://www.cs.bgu.ac.il/~elhadad/nlp09/sumbasic.pdf

Step 1 Compute the probability distribution over the words w_i appearing in the input, $p(w_i)$ for every i; $p(w_i) = \frac{n}{N}$, where n is the number of times the word appeared in the input, and N is the total number of content word tokens in the input.

Step 2 For each sentence S_j in the input, assign a weight equal to the average probability of the words in the sentence, i.e.

$$weight(S_j) = \sum_{w_i \in S_j} \frac{p(w_i)}{|\{w_i | w_i \in S_j\}|}$$

Step 3 Pick the best scoring sentence that contains the highest probability word.

Step 4 For each word w_i in the sentence chosen at step 3, update their probability

$$p_{new}(w_i) = p_{old}(w_i) \cdot p_{old}(w_i)$$

Step 5 If the desired summary length has not been reached, go back to Step 2.