

WHAT'S MY LINE

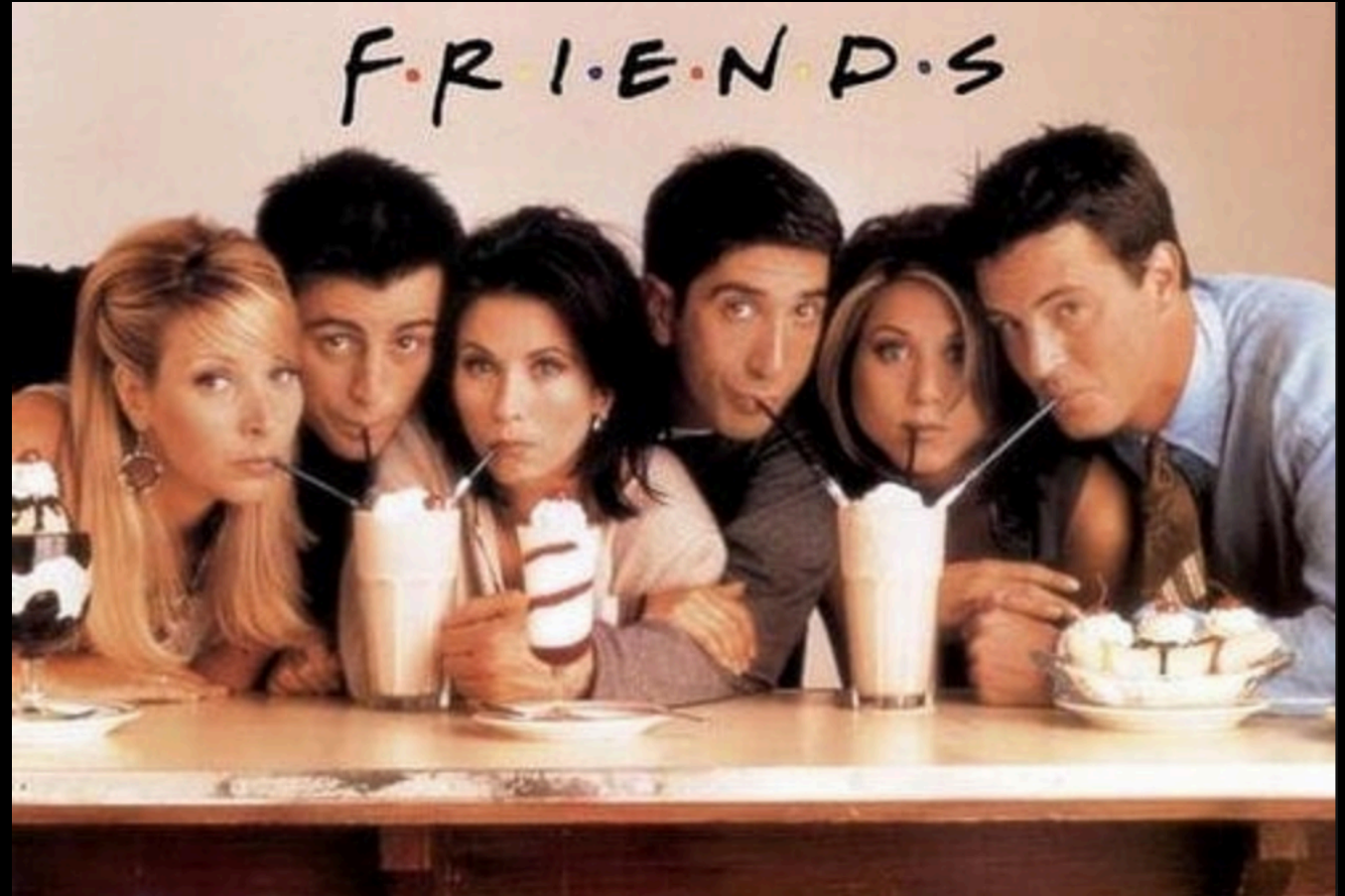
Xuan Qi

F.R.I.E.N.D.S



Objective

- Use LDA to discover the abstract "topics" that occur in the collection of documents
- Text generator: generate dialog for the main characters.



DATA

- Scraping dialog from 224 episode (<https://fangj.github.io/friends/>)

1. The One Where Monica Gets a New Roommate (The Pilot-The Uncut Version)

Written by: Marta Kauffman & David Crane

Transcribed by: [guineapig](#)

Additional transcribing by: [Eric Aasen](#)

(Note: The previously unseen parts of this episode are shown in blue text.)

[Scene: Central Perk, Chandler, Joey, Phoebe, and Monica are there.]

Monica: There's nothing to tell! He's just some guy I work with!

Joey: C'mon, you're going out with the guy! There's gotta be something wrong with him!

Chandler: All right Joey, be nice. So does he have a hump? A hump and a hairpiece?

Phoebe: Wait, does he eat chalk?

(They all stare, bemused.)

Phoebe: Just, 'cause, I don't want her to go through what I went through with Carl- oh!

Data mining

1.

Monica: There's nothing to tell! He's just some guy I work with!

Joey: C'mon, you're going out with the guy! There's gotta be something wrong with him!

Chandler: All right Joey, be nice. So does he have a hump? A hump and a hairpiece?

Phoebe: Wait, does he eat chalk?

(They all stare, bemused.)

Separate staging directions with dialog.



2.

| | diag | expression | diag_filt |
|------------------|---|---|--|
| Phoebe: | Wait, does he eat chalk?, Just, 'cause, I don... | She starts to pluck at the air just in front o... | Wait, does he eat chalk?, Just, 'cause, I don... |
| Rachel: | Oh God Monica hi! Thank God! I just went to y... | to,the waitress that brought her coffee,to Mon... | Oh God Monica hi! Thank God! I just went to y... |
| Ross: | (mortified) Hi., I just feel like someone reac... | nortified,in a deep voice,choked voice,normal ... | Hi., I just feel like someone reached down my... |
| Monica: | There's nothing to tell! He's just some guy l... | explaining to the others,to Ross,pointing at R... | There's nothing to tell! He's just some guy l... |
| Chandler: | All right Joey, be nice. So does he have a hu... | They all stare at him.,He extends his hand hop... | All right Joey, be nice. So does he have a hu... |
| Joey: | C'mon, you're going out with the guy! There's... | comforting her,picking up a leftover part,look... | C'mon, you're going out with the guy! There's... |

Non Supervised LDA

Corpus

6 documents



Corpus

6 documents (e.g.
Phoebe)



LDA: work flow

"I was, fourteen. My mom had just killed herself and my step-dad was back in prison, and I got here, and I didn't know anybody."



**Tokenization, lemmatization,
non stop words, length ≥ 3**

[fourteen, mom, just, kill, step-dad, back, prison, know ...]



Bag-of-words

{fourteen:1, mom: 1, just: 1, kill: 1, step-dad: 1, back: 1,
prison: 1, know: 1 ...}



LDA: Gensim

file:///Users/xuanqi/Documents/Metis/project4/friend_lda.html

Text generator: work flow

"I was, fourteen. My mom had just killed herself and my step-dad was back in prison, and I got here, and I didn't know anybody. And I ended up living with this, albino guy who was, like, cleaning windshields outside port authority and then, he killed himself, and then I found aromatherapy. So believe me, I know exactly how you feel"



Character to integer

{ ' ': 0, ',': 1, '-': 2, '.': 3, 'A': 4, 'I': 5, 'M': 6, 'S': 7, 'a': 8, 'b': 9, 'c': 10, 'd': 11, 'e': 12, 'f': 13, 'g': 14, 'h': 15, 'i': 16, 'j': 17, 'k': 18, 'l': 19, 'm': 20, 'n': 21, 'o': 22, 'p': 23, 'r': 24, 's': 25, 't': 26, 'u': 27, 'v': 28, 'w': 29, 'x': 30, 'y': 31, '"': 32 }

Text generator

Seed:

" Wait, does he eat chalk?, Just, 'cause, I don't want her to go through what I went through with Caral "

Epoch 1: loss: 3.1725 - acc: 0.186276225/76225

ho ho ho ho ho ho ho ho ho ho ho ho ho ho ho ho
ho ho ho ho ho ho ho ho ho ho ho ho ho ho ho
ho ho ho ho ho ho ho

Epoch 50: loss: 1.8456 - acc: 0.458476160/76225

I mann tous feen bod I was juit to toe tooeoni yo hnow
a''s not ave you think to hete thet aeTraceback

Rachel model

tjnn. ccf kkgg Tqtngt0. Jmcj uqqjv jq jgdqp qpq vqtj qqt
vq vqug vqtj vh vqg vqt qhtkg0 Kpf c egvpq vg0 K yccf K
jcf fv vjtgcuj c fcp ygct iqij K jcx



Model

LSTM

Dropout 0.2

Dense

Compile

Model



FUTURE WORK

- Text mining
- Use word based text generator.



THANK YOU

QUESTIONS?

Email: sherryqixuan@gmail.com

Linkedin:

<https://www.linkedin.com/in/xuanqi-phd/>

LDA

- Six Chatracter
- file:///Users/xuanqi/Documents/Metis/project4/friend_Ida.html
- file:///Users/xuanqi/Documents/Metis/project4/rachel_Ida.html

Text generator

“I was, fourteen. My mom had just killed herself and my step-dad was back in prison, and I got here, and I didn't know anybody.

Sequence length: 100



[19 0 29 8 25 1 ...]



Reshape

M samples

Time Steps

| | | | | | | | | | | | | | | | |
|---|---|----|-----|----|---|---|----|-----|----|---|---|----|-----|----|---|
| [| [| 19 | ... | 25 |] | [| 0 | ... | 27 |] | [| 29 | ... | 26 |] |
| | | ⋮ | ⋱ | ⋮ | | | ⋮ | ⋱ | ⋮ | | | | | | |
| | | 36 | ... | 0 | | | 33 | ... | 6 | | | 31 | ... | 4 | |

Features (Sequence length)

- <https://machinelearningmastery.com/how-to-develop-a-word-level-neural-language-model-in-keras/>
- <http://nbviewer.jupyter.org/github/bmabey/pyLDavis/blob/master/notebooks/Gensim%20Newsgroup.ipynb>
- <https://github.com/bmabey/pyLDavis>