

The **Best** Words

Leon Zhou

lzhou95@gmail.com

LinkedIn: [zhouleon](#)

GitHub: [tetrahydrofuran](#)



**“I know words. I have the best words.”
- Donald J. Trump**

Opening slide to introduce the project

A Unique Style



Donald J. Trump ✓

@realDonaldTrump

Follow



Sorry losers and haters, but my I.Q. is one of the highest -and you all know it! Please don't feel so stupid or insecure, it's not your fault

6:37 PM - 8 May 2013



Donald J. Trump ✓

@realDonaldTrump

Follow



A TOTAL WITCH HUNT!!!

4:08 AM - 10 Apr 2018



Donald J. Trump ✓

@realDonaldTrump

Follow



....to President of the United States (on my first try). I think that would qu smart, but genius....and a ver at that!

4:30 AM - 6 Jan 2018

Slide to say that Trump has unique style of speaking, hopefully enough so for a neural net to pick up on

Tweets will appear one at a time

The Data



via precompiled
database



officially released
transcripts

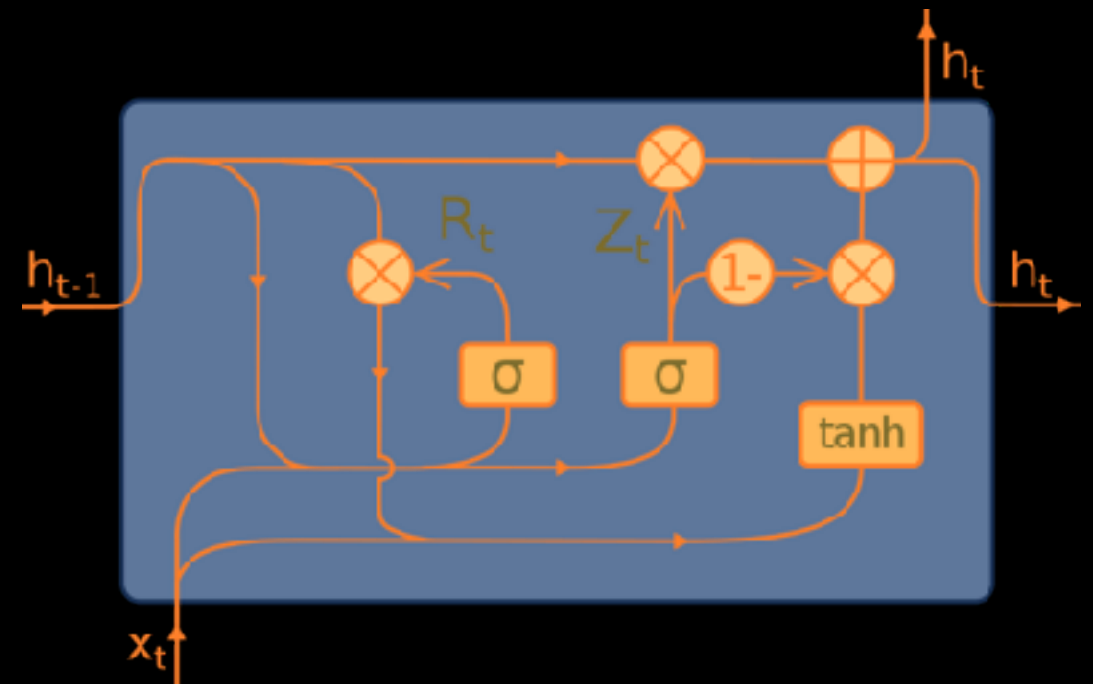
Description of data source:

Database of Donald's tweets
Official White House transcripts

Generation Methods

Current State	Next Word (p)
...cut taxes	bigly (0.70)
	. (0.23)
	soon (0.07)

Markov
Chain



Recurrent
Neural Network

Description of recurrent models
RNN and Markov

Sample Text

Quick slide with sample text from
both methods

Quirks

Tweet-trained models deal with uncertainty with excessive username mentions:

This is too good to be true for Republicans are a total job to community and endorsement is a complete course in the country is a complete to see you can t know it will be a careful course in the country is a complete course in chicago. --
@realdonaldtrump @realdonaldtrump @foxnews @realdonaldtrump @realdonaldtrump
@mittromney @celebapprentice @realdonaldtrump @realdonaldtrump @realdonaldtrump
@realdonaldtrump @realdonaldtrump @realdonaldtrump @cnn @cnn @cnn @cnn @cnn @cnn @cnn
@cnn @jerrysalmond @realdonaldtrump @realdonaldtrump @cnn @cnn @cnn @cnn @cnn @cnn
@cnn @cnn @jack_wallown @trumpchicago @realdonaldtrump @celebapprentice
@realdonaldtrump @foxnews @realdonaldtrump @realdonaldtrump @macys @realdonaldtrump
@celebapprentice @realdonaldtrump @realdonaldtrump @macys @realdonaldtrump
@realdonaldtrump @realdonaldtrump @apprenticenbc @realdonaldtrump @realdonaldtrump
@trumpchicago @realdonaldtrump @realdonaldtrump @realdonaldtrump @realdonaldtrump
@mittromney @jerrysalor @realdonaldtrump @realdonaldtrump @realdonaldtrump @re

Tweet “Hidden-ELU” Unidirectional, Temperature=0.2

Quirks

... or hyperlinks:

Why didn't President Obama do something it was a complete course in the country is a complete course in the country is a complete security and we will continue to care of the country will solve the special country and the country is a complete course in the country is a complete course in chicago - and will be a great country great again!
<https://t.co/usyaccclw> <https://t.co/syukcbycck> <https://t.co/cuygpyccecy> <https://t.co/aqasyvice0> <https://t.co/qcyccsyccps> <https://t.co/qxckcyccc4> <https://t.co/agec2yhccc> <https://t.co/lswxxxsccc> <https://t.co/jcsyywwyoxx> <https://t.co/oc3uycocpcc> <https://t.co/qysygcccol> <https://t.co/cssyacywcc> <https://t.co/ecymsycccccw> <https://t.co/icycccyccc> <https://t.co/ycbuyccecc> <https://t.co/ysecxacco4ch> <https://t.co/ec6rcs4crt> <https://t.co/as0cxckyco> <https://t.co/s4xocwpycc> <https://t.co/xcsc3wyyc> <https://t.co/c7yschcooqy> <https://t.co/s4csmgrcccc> <https://t.co/c4yxcpcccws> <https://t.co/syskcxxsxxh> <https://t.co/csyp4iiccc> <https://t.co/aasyscrccc> <https://t.co/s4wawc5ycc> <https://t.co/ycyjpqswys> <https://t.co/csyj>

Tweet “Hidden-ELU” Unidirectional, Temperature=0.2

Quirks

PLACEHOLDER; brief example of remarks-corpus trained model

Why didn't President Obama do something it was a complete course in the country is a complete course in the country is a complete security and we will continue to care of the country will solve the special country and the country is a complete course in the country is a complete course in chicago - and will be a great country great again!
<https://t.co/usyaccclw> <https://t.co/syukcbycck> <https://t.co/cuygpyccecy> <https://t.co/aqasyvice0> <https://t.co/qcyccsyccps> <https://t.co/qxckcyccc4> <https://t.co/agec2yhccc> <https://t.co/lswxxxsccc> <https://t.co/jcsyywwyoxx> <https://t.co/oc3uycocpcc> <https://t.co/qysygcccol> <https://t.co/cssyacywcc> <https://t.co/ecymsycccccw> <https://t.co/icycccyccc> <https://t.co/ycbuyccecc> <https://t.co/ysecxacco4ch> <https://t.co/ec6rcs4crt> <https://t.co/as0cxckyco> <https://t.co/s4xocwpycc> <https://t.co/xcsc3wyyc> <https://t.co/c7yschcoogy> <https://t.co/s4csmgrcccc> <https://t.co/c4yxcpcccws> <https://t.co/syskcxxsxxh> <https://t.co/csyp4iiccc> <https://t.co/aasyscrccc> <https://t.co/s4wawc5ycc> <https://t.co/ycyjpgswys> <https://t.co/csyj>

Tweet “Hidden-ELU” Unidirectional, Temperature=0.2

- Which is real

If I get good results from my RNN, I would like to have a slide placing trump quotes and generated quotes side by side for comparison; audience interaction in determining which is real

Thank You



Leon Zhou
lzhou95@gmail.com
LinkedIn: zhouleon
GitHub: tetrahydrofuran

Appendix

Recurrent Model Architecture

6 variants (3 base + uni/bi-directional)

```
def rnn_basic(X, y, bidirectional=False):
    seq_input, x_ohe, output = __init_shared_layers(X, y)

    # Network architecture
    x = x_ohe(seq_input)
    gru = GRU(units=256, dropout=0.2, recurrent_dropout=0.2)
    if bidirectional:
        x = Bidirectional(gru(x))
    else:
        x = gru(x)

    x = output(x) # Dense
```

```
def rnn_flanked_elu(X, y, bidirectional=False):
    seq_input, x_ohe, output = __init_shared_layers(X, y)

    # Network architecture
    x = x_ohe(seq_input)
    x = Dense(units=256, activation='elu')(x)
    gru = GRU(units=256, dropout=0.2, recurrent_dropout=0.2)
    if bidirectional:
        x = Bidirectional(gru(x))
    else:
        x = gru(x)

    x = Dense(units=512, activation='elu')(x)
    x = output(x) # Dense
```

```
def rnn_hidden_elu(X, y, bidirectional=False):
    seq_input, x_ohe, output = __init_shared_layers(X, y)

    # Network architecture
    x = x_ohe(seq_input)
    gru = GRU(units=256, dropout=0.2, recurrent_dropout=0.2)
    if bidirectional:
        x = Bidirectional(gru(x))
    else:
        x = gru(x)

    x = Dense(units=512, activation='elu')(x)
    x = output(x) # Dense
```

Adam Optimizer, 0.01 LR
Categorical Cross-Entropy Loss
200 Epochs
4096 Batch Size
40 Character Window Length

Markov Chain Text Samples

America the Emir of steel.

Democrats seem desperate.

FBI are wearing Revlon eyeliner today.

Russia thinks that I really cares about my presentation.

Immigration is coming.

Loser newspaper MUST READ.

RNN Text Samples

America the Emir of steel.

Democrats seem desperate

Placeholder

FBI are wearing Revlo

Russia thinks that I really cares about my presentation.

Immigration is coming.

Loser newspaper MUST READ

REGULAR EXPRESSION

65 matches, 156164 steps (~46ms)

```
:/ ([PRESIDENT TRUMP|THE PRESIDENT]): (.*) ([A-Z']+) ?[A-Z']+?:|Q ?|END)
```

TEST STRING

SWITCH TO UNIT TESTS ▶

and it's just an endless honor to be a part of it, and even in any a small way. So thank you, sir. **THE PRESIDENT:** Thank you, Stephen. That's great. A great job you do, too. **GASPAR:** Good afternoon, Mr. President. It's an honor to be here. I'm Kristin Gaspar representing the largest county here today, San Diego County. I have 3.5 million constituents that I'm responsible for their public safety. If you look around this room, your tiny but mighty team, this is what Governor Brown classifies as low-life politicians. Well, here we are. You've heard about the problems. You've heard about the statistics. And I could have thought of a million things to say to you. I have a stack of 3,000 emails in my office. These are the emails that have come in – thank yous, people supporting what we're doing. And I have a tiny little stack of less than 50 where people are very upset with what we're doing in San Diego County. **THE PRESIDENT:** How is the wall going? How is the wall? (Laughter.) **GASPAR:** It's going. It's going. **THE PRESIDENT:** We're getting it built, right? **GASPAR:** It is being built. **THE PRESIDENT:** They wanted it so badly – San Diego. They wanted it so badly. And I said, you know, if we build it, we will lose a big constituency, because there won't be anybody saying, "We want the wall." But we had to build it. So I know they're very happy about it. **GASPAR:** And I'd like to share with you a story, because sometimes humanizing the issue is really important. And a family reached out to me, and I brought with me one single photo on that plane, since the stack of 3,000 emails is a little difficult to carry. But that photo was the last photo taken

SUBSTITUTION

Example of complex regex extractions

Topic Modeling via LDA

WIP

Potential quotes to use if “can you tell which quote is real” slide is made

- “unattractive, both inside and out”
- “the beauty of me is that I’m very rich”
- “my fingers are very beautiful, as, it has been well documented”
- “i have never seen a thin person drinking diet coke”
- “eventually we’re going to get something done and it’s going to be really, really good”