State of the Sesame Street

(Are those NLP folks, like, ok?)

Dr. Rachael Tatman
Data Science Advocate, Kaggle

@rctatman

Sesame Street is 50!

@rctatman

Sesame Street is ~~30~~.

STATE OF THE ART IN NLP

@rctatman

Sesame Street is ~~not~~ STATE OF THE ART IN NLP

@rctatman

It turns out humans don’t just use words in a random order

It turns out humans don’t just use words in a random order

Random out turns use words it don’t a humans just order

It turns out humans don’t just use words in a random order

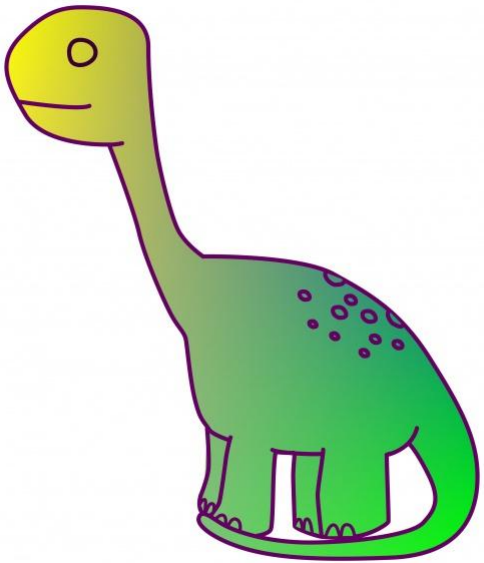
Random out turns use words it don’t a humans just order

LANGUAGE MODELS

It turns out humans don’t just use words in a random order

Random out turns use words it don’t a humans just order

✨ *Statistical* ✨ LANGUAGE MODELS

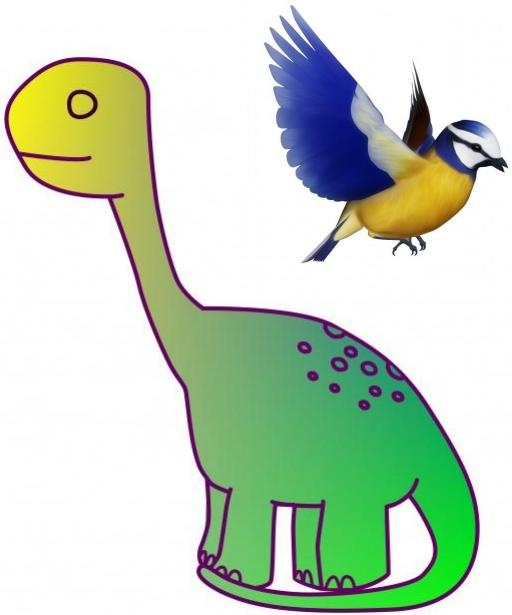


1 : This is a sentence *unigrams:* this, is, a, sentence

2 : This is a sentence *bigrams:* this is, is a, a sentence

3 : This is a sentence *trigrams:* this is a, is a sentence

@rctatman



1 : This | is | a | sentence — *unigrams:* this, is, a, sentence

2 : This | is | a | sentence — *bigrams:* this is, is a, a sentence

3 : This | is a | sentence — *trigrams:* this is a, is a sentence

@rctatman

@rctatman

"You shall know a word
by the company it keeps"

–J. R. Firth 1957

"You shall know a word
by the company it keeps"

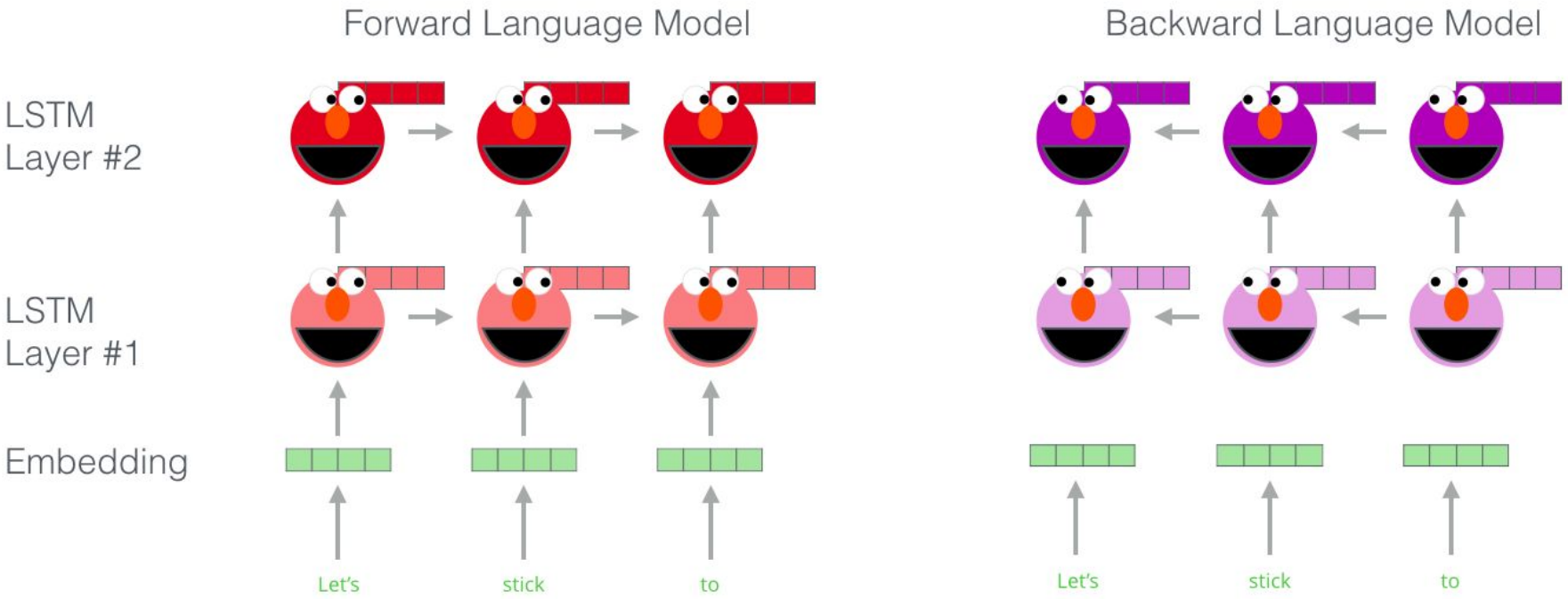
–J. R. Firth 1957

NOW WITH DEEP LEARNING!

@rctatman

ELMo: Deep contextualized word representations, NAACL 2018

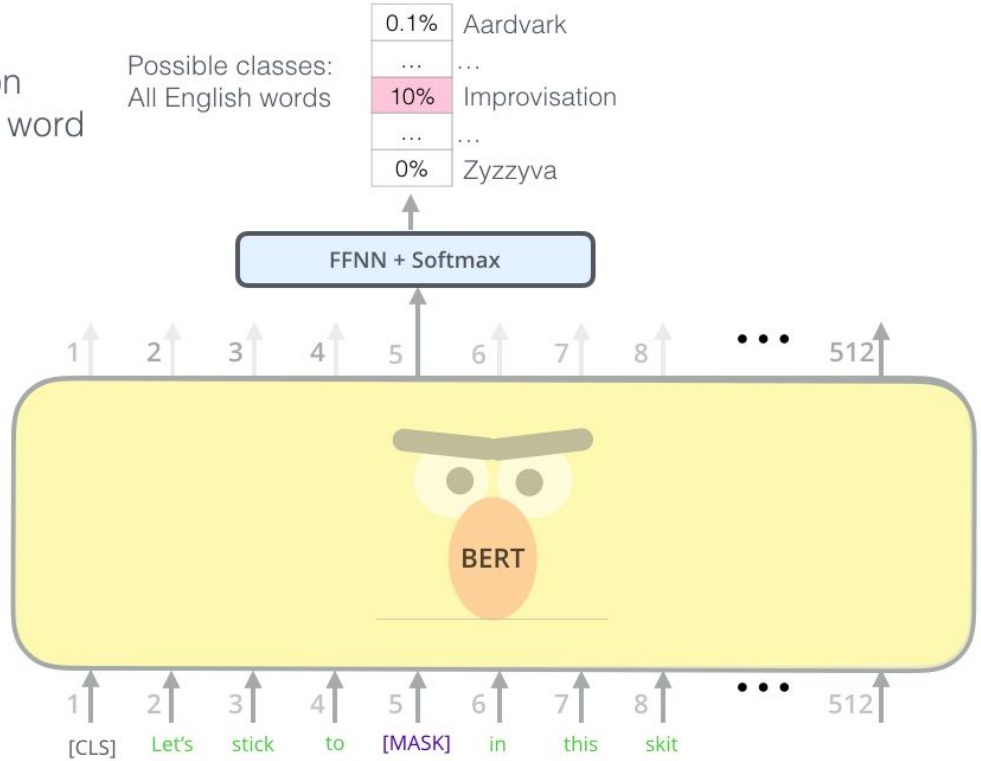
Embedding of “stick” in “Let’s stick to” - Step #1



http://jalammar.github.io/illustrated-bert/ @rctatman

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, NAACL 2019

Use the output of the masked word's position to predict the masked word

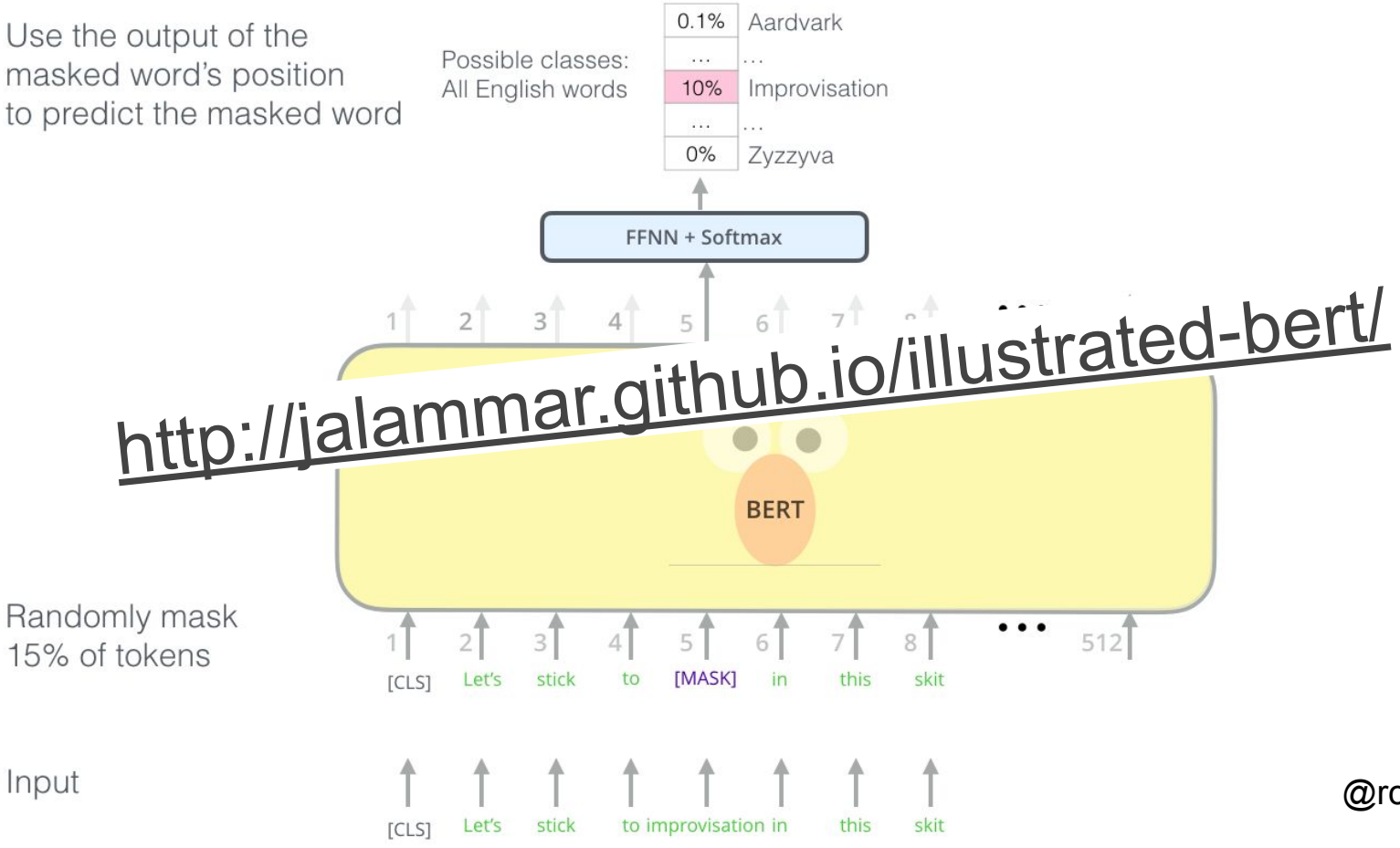


Randomly mask 15% of tokens

Input

@rctatman

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, NAACL 2019



http://jalammar.github.io/illustrated-bert/

@rctatman

State of the Sesame Street

- Elmo
  - Contextual word embedding
- Bert
  - Bidirectional embeddings w/ transformers
- Ernie (2 models!)
  - by Baidu: ERNIE: Enhanced Representation through Knowledge Integration
  - Tsinghua: ERNIE: Enhanced Language Representation with Informative Entities
- Big Bird
  - Multi-Task Deep Neural Networks for Natural Language Understanding
- ????

Yuval Pinter 🐸;🌾 4 1
@yuvalpi

Follow

Several New Ultimate Feature Finders Letting Embeddings Use Procedurally Acquired Global Universal Structure

> **Yuval Pinter** 🐸;🌾 4 1 @yuvalpi
> ok so which massive NLP team is working on COnvolutional Operational Knowledge-Injecting Embeddings?
>
> (cc @alethioguy)

(((ل)()ل() 'yoav))))
@yoavgo

Follow

The "BIG Binary-Inverted-Residual-Dropout" we are exploring plays along with and drives all previous methods to perform their best.

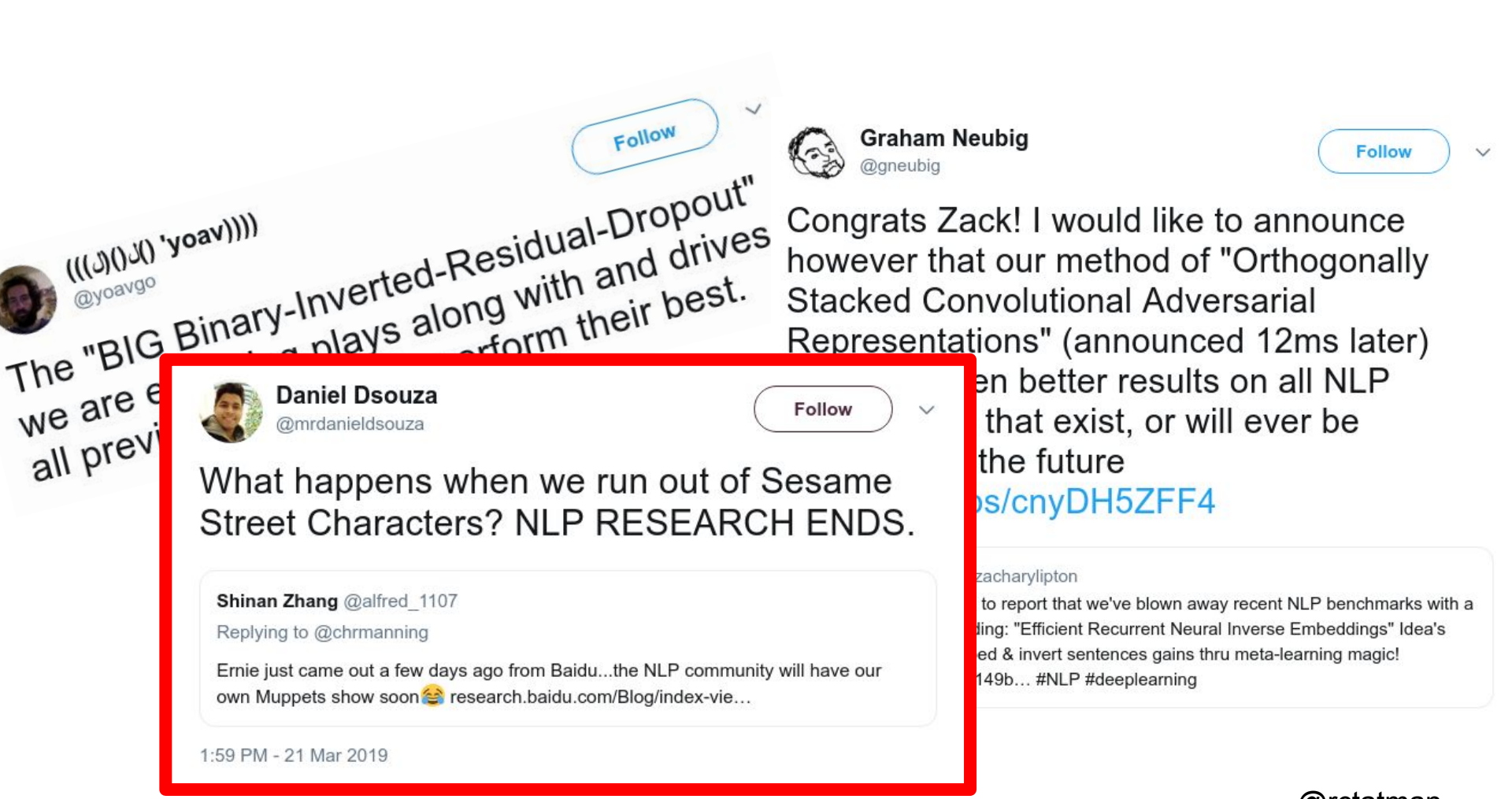
Graham Neubig
@gneubig

Follow

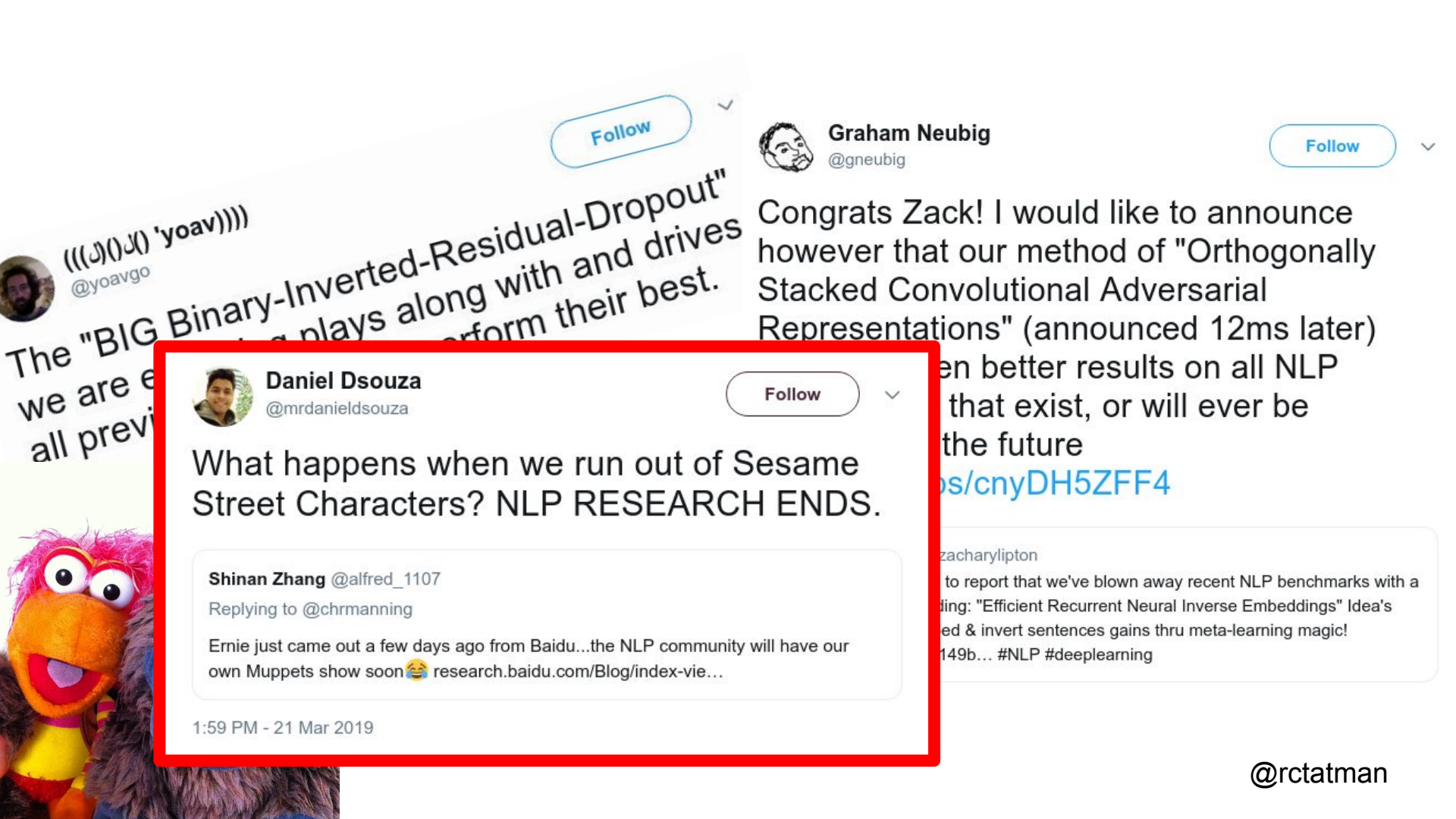
Congrats Zack! I would like to announce however that our method of "Orthogonally Stacked Convolutional Adversarial Representations" (announced 12ms later) achieves even better results on all NLP benchmarks that exist, or will ever be proposed in the future
arxiv.com/abs/cnyDH5ZFF4

> **Zachary Lipton** @zacharylipton
> Hey everyone. Stoked to report that we've blown away recent NLP benchmarks with a new sentence embedding: "Efficient Recurrent Neural Inverse Embeddings" Idea's simple: iteratively embed & invert sentences gains thru meta-learning magic! arxiv.com/abs/xFgq14149b… #NLP #deeplearning

@rctatman


Follow
(((J)()J() 'yoav))))
@yoavgo
The "BIG Binary-Inverted-Residual-Dropout"
plays along with and drives
we are e
all previ
Graham Neubig
@gneubig
Follow
Congrats Zack! I would like to announce however that our method of "Orthogonally Stacked Convolutional Adversarial Representations" (announced 12ms later)
en better results on all NLP
that exist, or will ever be
the future
s/cnyDH5ZFF4
zacharylipton
to report that we've blown away recent NLP benchmarks with a
ding: "Efficient Recurrent Neural Inverse Embeddings" Idea's
ed & invert sentences gains thru meta-learning magic!
149b… #NLP #deeplearning
Daniel Dsouza
@mrdanieldsouza
Follow
What happens when we run out of Sesame Street Characters? NLP RESEARCH ENDS.
Shinan Zhang @alfred_1107
Replying to @chrmanning
Ernie just came out a few days ago from Baidu...the NLP community will have our own Muppets show soon😂 research.baidu.com/Blog/index-vie…
1:59 PM - 21 Mar 2019

@rctatman


Follow
(((J)()J() 'yoav))))
@yoavgo
The "BIG Binary-Inverted-Residual-Dropout"
plays along with and drives
form their best.
we are e
all previ
Graham Neubig
@gneubig
Follow
Congrats Zack! I would like to announce however that our method of "Orthogonally Stacked Convolutional Adversarial Representations" (announced 12ms later)
en better results on all NLP
that exist, or will ever be
the future
s/cnyDH5ZFF4
zacharylipton
to report that we've blown away recent NLP benchmarks with a
ding: "Efficient Recurrent Neural Inverse Embeddings" Idea's
ed & invert sentences gains thru meta-learning magic!
149b… #NLP #deeplearning
Daniel Dsouza
@mrdanieldsouza
Follow
What happens when we run out of Sesame Street Characters? NLP RESEARCH ENDS.
Shinan Zhang @alfred_1107
Replying to @chrmanning
Ernie just came out a few days ago from Baidu...the NLP community will have our own Muppets show soon😂 research.baidu.com/Blog/index-vie…
1:59 PM - 21 Mar 2019

@rctatman

Well, look at that! This is the end of the book, and the only one here is . . .

ME

I, lovable, furry old **GROVER**, am the Monster at the end of this book.

And you were so **SCARED!**

THE EN

I told you and told you there was nothing to be afraid of.