

A decorative graphic in the top-left corner consisting of a network of interconnected nodes and lines. Some nodes are highlighted with blue circles, and others with blue dots.

Dialogue Generation And Classification with GPT-2 and BERT

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



The new season is coming!



matthew perry ✓

@MatthewPerry



Friends reunion being rescheduled for the beginning of March. Looks like we have a busy year coming up. And that's the way I like it!

11:44 AM · Nov 12, 2020



148.9K



25K people are Tweeting about this

HEY!!

**A robot can write
better story!**



Methods



Text Generation:

Finetune GPT-2 with scripts of F.R.I.E.N.D.S

But how to evaluate?

Measure the performance on a classification task!



Multi-Class Classification:

Labels: 6 main characters

Text: a single line

Model: BERT



**More
Details**

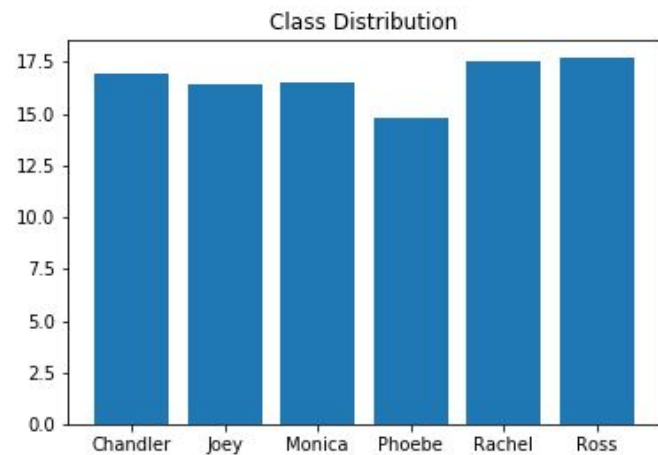
Data Augmentation

Insert/substitute words by feeding surrounding words to BERT to find out the most suitable ones for augmentation

Generator Evaluation

Compare the test accuracy of original lines and the generated lines.

DataSet



Quantile	Length of the Line
10%	4
25%	6
50%	10
75%	17
90%	28
100%	248

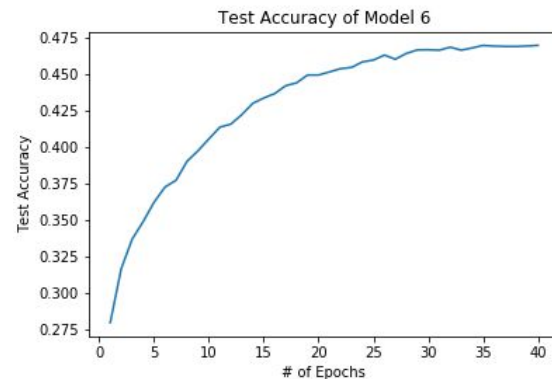
Table 1: Length of the Lines

“

1. Data Augmentation is powerful. Increasing the number of samples improves model performance.
2. Test accuracy of the best model: 47%

Models	1	2	3	4	5	6	7
Augmentation	F	F	T	T	T	T	T
# of Samples	42K	42K	77K	77K	84K	84K	84K
Max Input length	32	64	32	64	32	64	128
# of Epochs	10	10	10	10	10	10	10
Batch Size	128	128	512	128	512	128	64
Accuracy	0.328	0.334	0.340	0.363	0.351	0.406	0.410

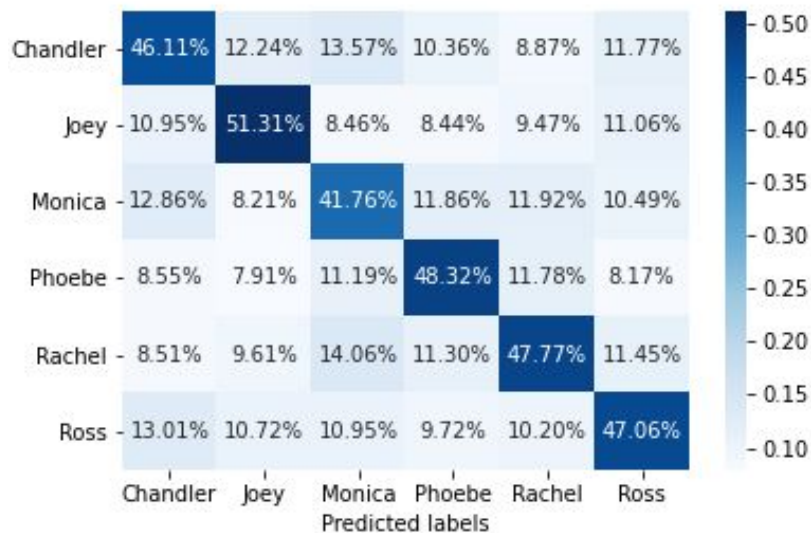
Table 2: BERT classification Results



“

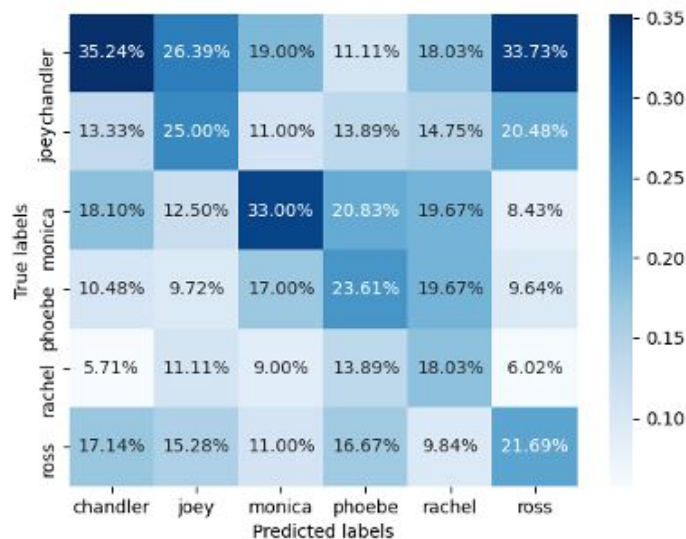
Confusion Matrix of Original Test DataSet

Accuracy: **42.78%**



Confusion Matrix of Dialogues Generated by GPT-2

Accuracy: **26.17%**





Defamation!

Who Said This In *F.R.I.E.N.D.S*

Bert Multi-Classifer Prediction: Finetuned with F.R.I.E.N.D.S. dialogue

Enter the text for prediction here!

Submit

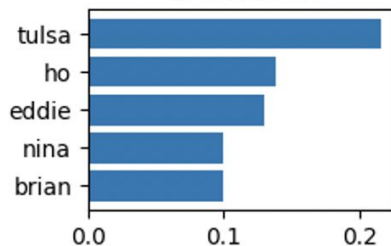
is most likely to say this line! With a confidence of

Demo

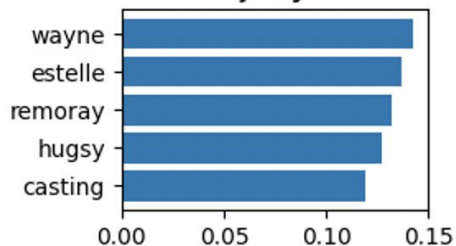
“

Words that matters!

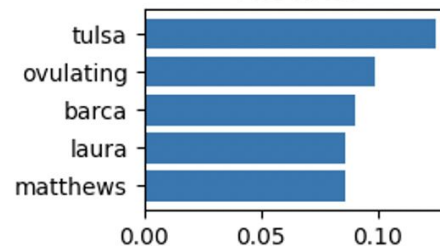
chandler



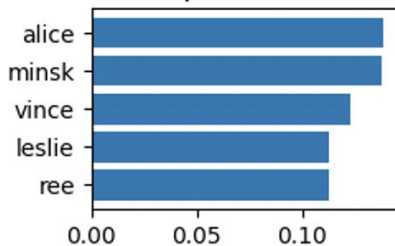
joey



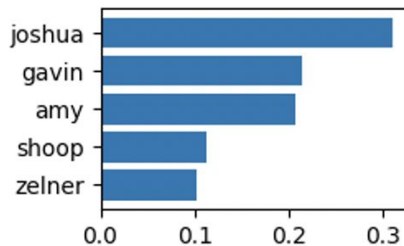
monica



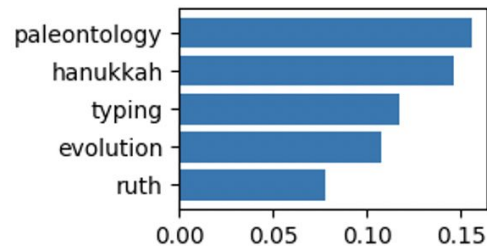
phoebe



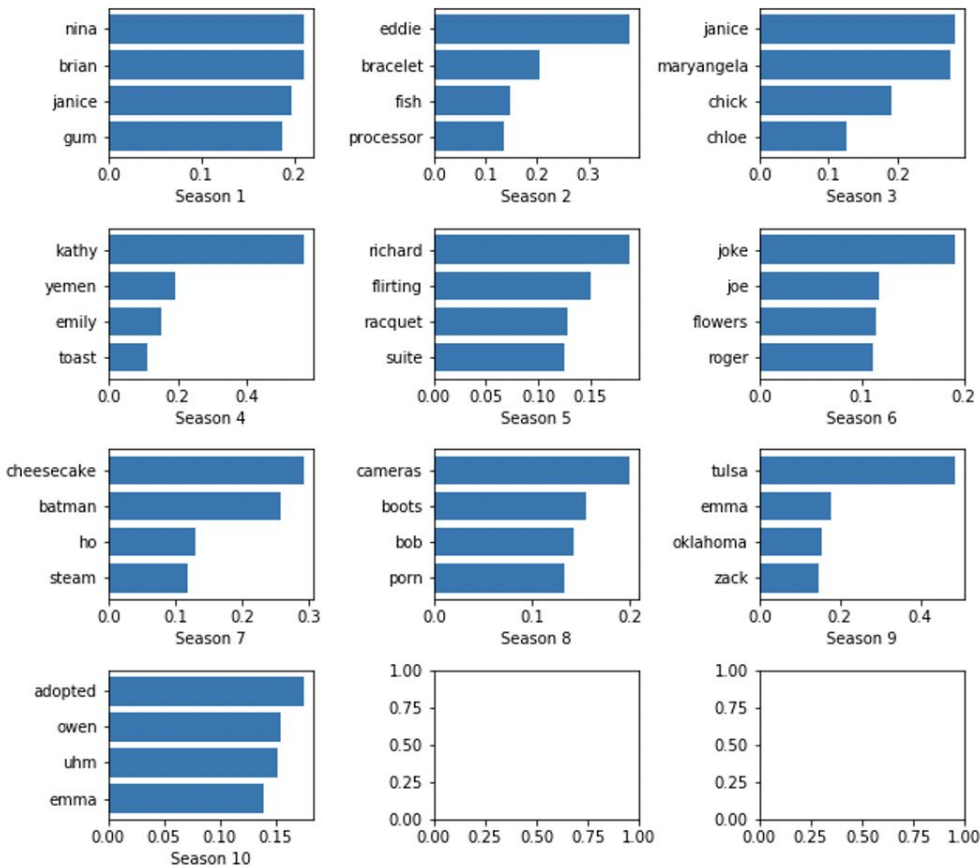
rachel



ross



Chandler's Life





Thanks!

Any questions?

