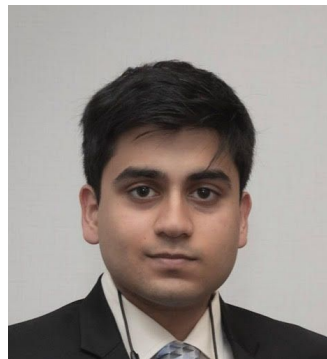




Biased Abstractive Summarization

By Vikrant Sharma and Zachary Capone

TEAM 14



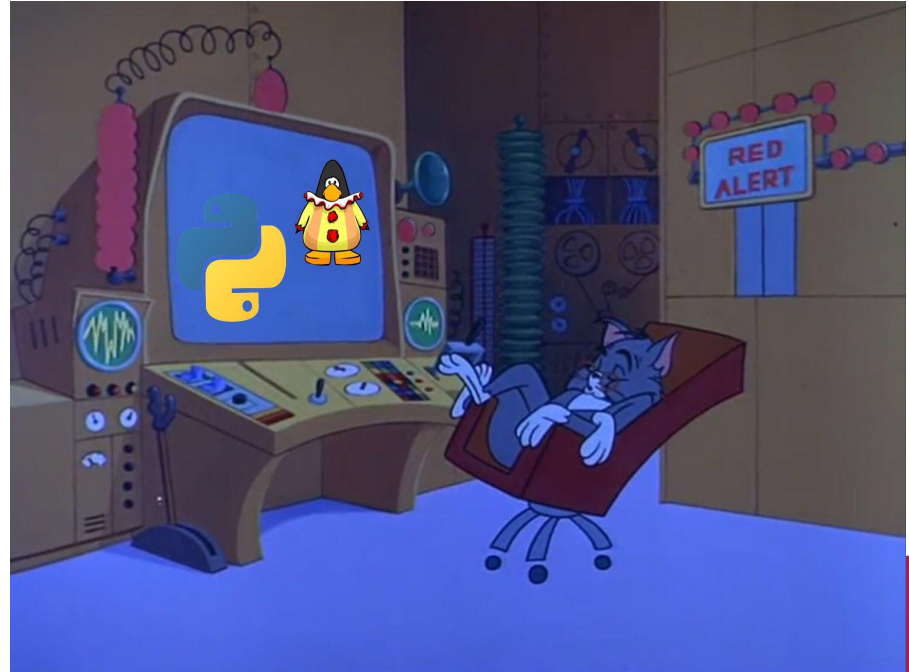
The Problem We Wanted to Solve

- Finding a good product, good movie, good car, and a good restaurant can often be tedious. Digging through reviews takes time, and may even be in a foreign language given you are on vacation in a foreign country.
- Our algorithm works to lessen the burden of these issues by sifting through reviews for the user and also has the capability to translate reviews to different languages.



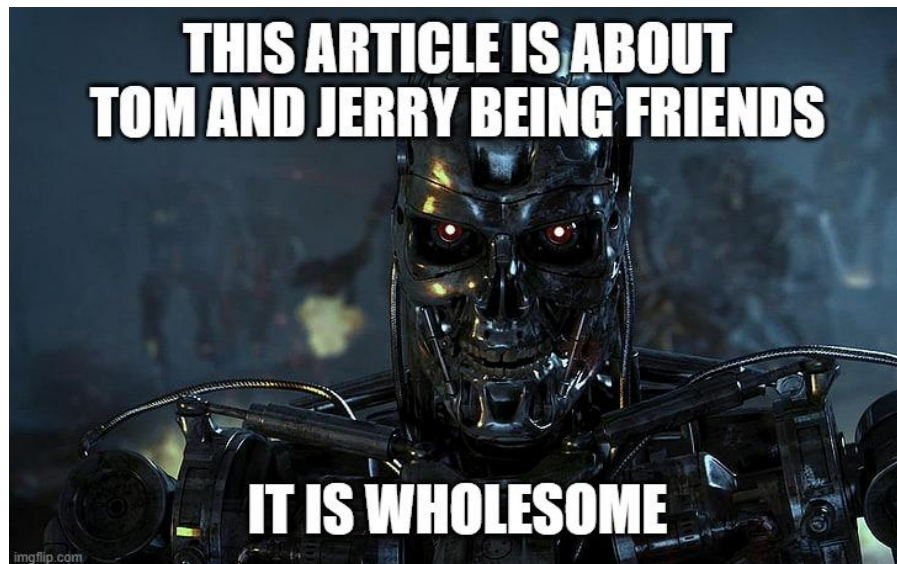
How did we make this?

- Our program was created in Python on a linux OS (Library Reasons)
- We started by initially training off of tweets, and once this became functional, expanded this into full articles.
- Algorithms were trained thanks to Vikrant's 2080 Ti and a lot of random angry people on twitter.



What does it do?

- Our first algorithm goes through articles and reviews fed to it, summarizes them, and then states the perceived mood of the media. (Positive, Negative, etc.)
- This was done using the proof of concept from tweet analysis, and adapting training from tweets to articles instead.
- Enables lazy readers and terminators to enjoy movie reviews.



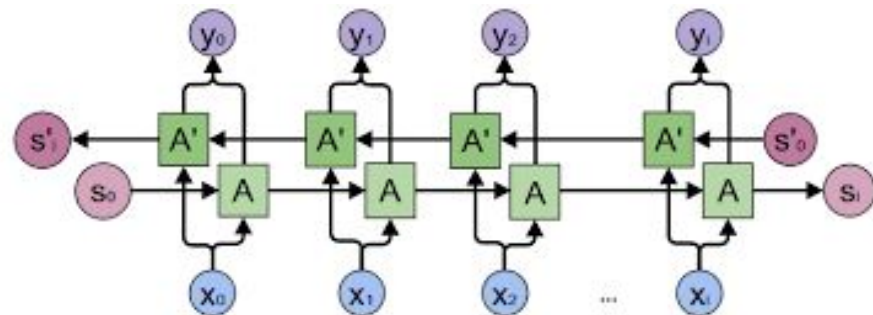
The methods behind our project.

- In order to make this possible, we essentially combined multiple smaller deep learning architectures, including: an attention mechanism, a Bi-Directional LSTM, and encoder/decoder framework.
- The program was built up slowly from when projects were assigned. We gradually added on more and more complexity in the time allowed, using the foundational framework of the software and continually programming it to be capable of more and more.



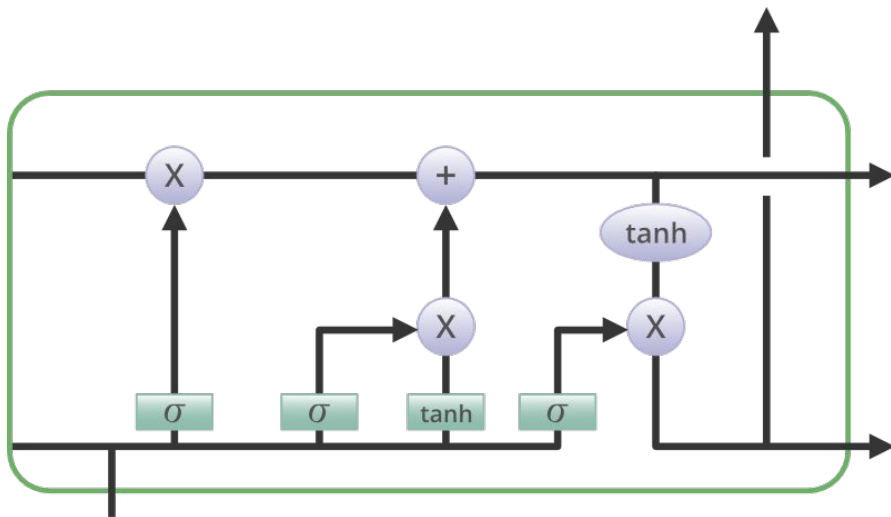
Bi-Directional LSTM

- Bi-Directional LSTM's are great in that unlike LSTM's and Unidirectional LSTM's, Bi-Directional LSTM's can communicate data both forwards and backwards (essentially acting for past and future considerations). This gives it the ability to consider more data at once, creating a better sense of context for the algorithm.
- For context, say the computer was fed the phrase 'Ps3 made me...' and told to finish it, a LSTM would only see that phrase, and a Unidirectional LSTM would see just the predicted second half, for example '...cry at 2 in the morning'. A Bi-Directional LSTM enables both to be suspended and analyzed at once.



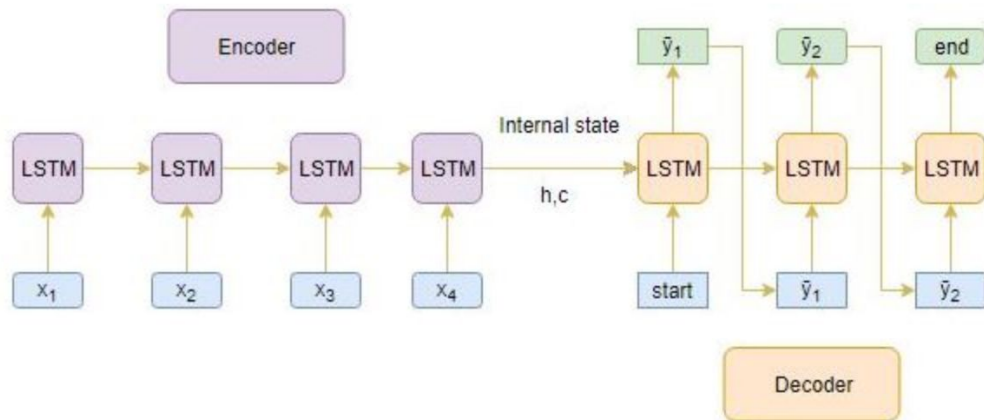
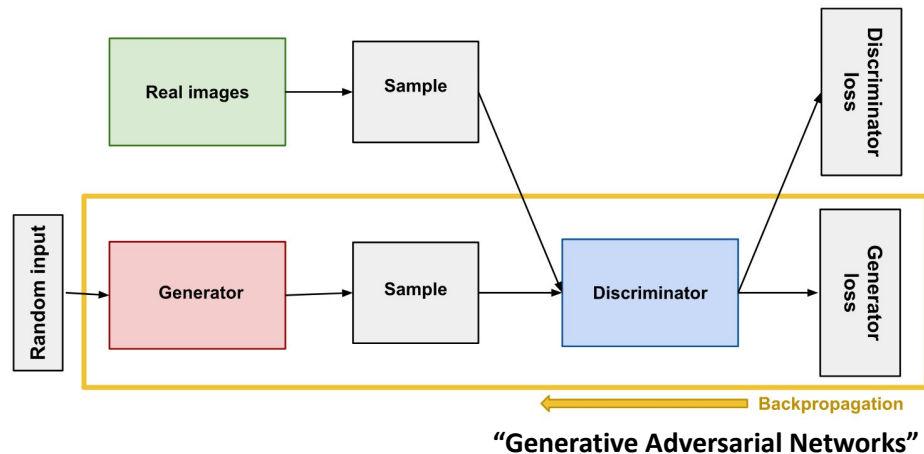
A bit more info about LSTM's

- 'LSTM' stands for long short-term memory.
- Because LSTM's can compare past and present variables with each other it makes them superior to RNN's in that they learn better with time lags between data sets.
- This greater versatility in memory allows for tackling more complex problems, such as our project.



Text Abstraction

- Synthesis of new sentences
- GANs
- Seq2Seq



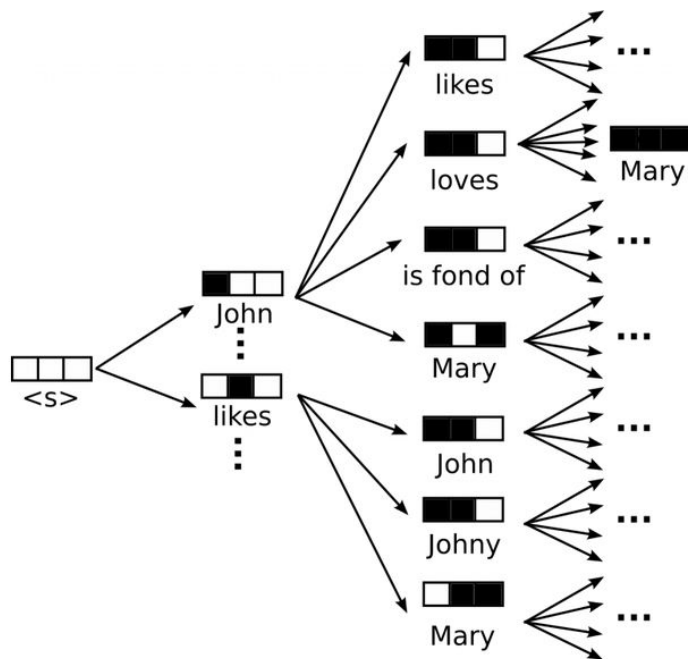
Pai, 2019

Data preprocessing

- Remove unintelligible/empty entries
- “Translate” contractions
- Remove punctuation
- Remove short words
- START and END words for summary decoder

```
Review: confection around centuries light pillowy citrus gelatin nuts case filberts cut tiny squares liberally coated powdered suga  
r tiny mouthful heaven chewy flavorful highly recommend yummy treat familiar story lewis lion witch wardrobe treat seduces edmund s  
elling brother sisters witch  
Summary: _START_ delight says it all _END_
```

Encoder Decoder Architecture

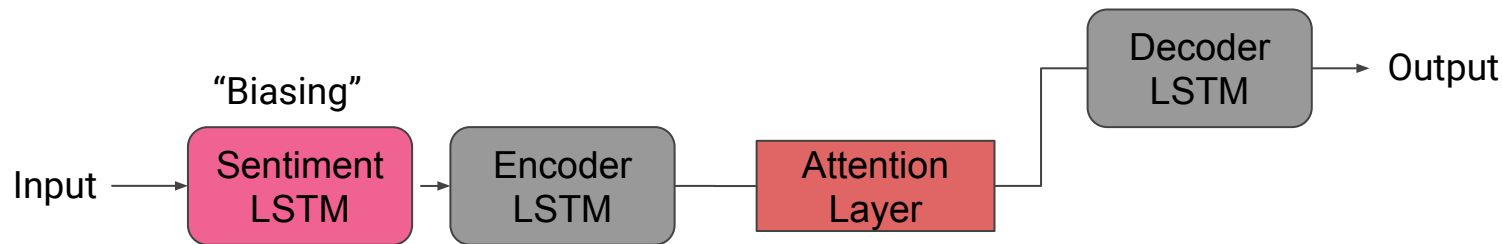


Shaikh, 2018

Too many arbitrary possibilities!

SOLUTIONS

- Beam Search
- Attention Mechanism

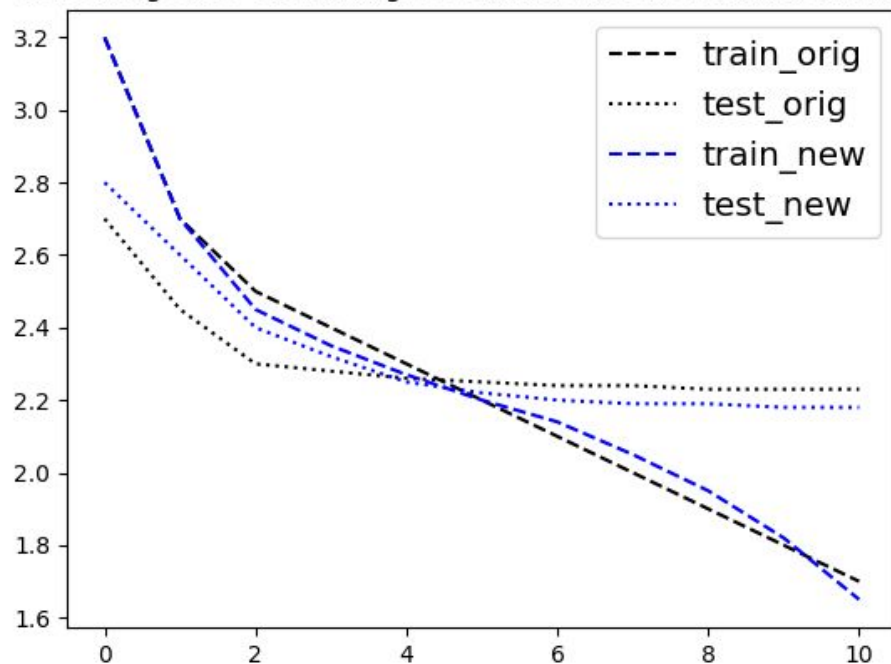


- Kaggle Amazon reviews
- 5 Stacked LSTM encoder
- Bahdanau AM
- 1x LSTM using encoder states
- Simple sentiment analysis
- Low batch size (memory issue)
- 0.9 train/test ratio

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 200)]	0	
embedding (Embedding)	(None, 200, 200)	8804800	input_1[0][0]
lstm (LSTM)	[(None, 200, 200), (320800		embedding[0][0]
lstm_1 (LSTM)	[(None, 200, 200), (320800		lstm[0][0]
lstm_2 (LSTM)	[(None, 200, 200), (320800		lstm_1[0][0]
lstm_3 (LSTM)	[(None, 200, 200), (320800		lstm_2[0][0]
input_2 (InputLayer)	[(None, None)]	0	
lstm_4 (LSTM)	[(None, 200, 200), (320800		lstm_3[0][0]
embedding_1 (Embedding)	(None, None, 200)	2337600	input_2[0][0]
tf.convert_to_tensor (TFOpLambd	(None, 200)	0	lstm_4[0][1]
tf.convert_to_tensor_1 (TFOpLam	(None, 200)	0	lstm_4[0][2]
lstm_5 (LSTM)	[(None, None, 200), 320800		embedding_1[0][0] tf.convert_to_tensor[0][0] tf.convert_to_tensor_1[0][0]
scanner (AttentionLayer)	((None, None, 200), 80200		lstm_4[0][0] lstm_5[0][0]
conc_layer (Concatenate)	(None, None, 400)	0	lstm_5[0][0] scanner[0][0]
time_distributed (TimeDistribut	(None, None, 11688)	4686888	conc_layer[0][0]
Total params: 17,834,288			
Trainable params: 17,834,288			
Non-trainable params: 0			

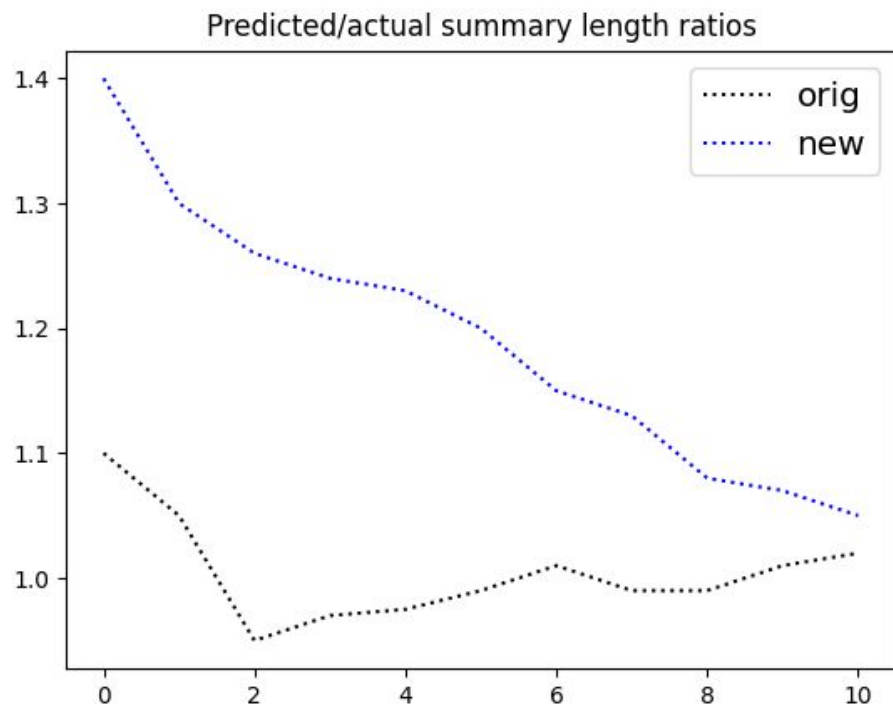
Results

Percentage loss on training/test data between old and new model



NOTE: better than previous LSTM method, but not improving on state of the art

Results



Results

Original string: This fire tv stick is horrible when it comes to streaming, it constantly goes out constantly buffering, i can't enjoy steaming my cable channels, apparently it's mainly designed just to watch your Amazon prime shows. I literally tried everything from restarting to uninstalling to reinstalling app, clearing data, clearing cache and it still doesn't let me steam, it only works for Amazon prime shows. What a waste of money.

Actual Summary: Fire stick having issues

Predicted Summary (old): stick again

Predicted Summary (new): bad television product

Original string: I have always wanted an ice cream maker but didn't want to mess around with preplanning to have ice and salt on hand, or prefreezing a bowl. With a built in compressor to do the freezing and a reasonable price, this Whynter Ice Cream Maker was made for me! I also needed a smaller footprint so it's nice that this maker comes in a vertical version, which is the one I got. It is so easy to use, I can have as much ice cream as I could ever want, any time I want! I have made ice cream every weekend: vanilla, blueberry green tea, orange sherbet, roasted plum honey ginger and even a boozy whiskey vanilla! Any recipe, custard or not, is accommodated. Although I haven't tried it for drinks, it's capable of creating frozen drinks, slushees, etc., which is a pleasant surprise. It runs quietly, too. It's just loud enough to hear when the ice cream is close to being finished (it has auto shut off if ice cream gets too thick to churn) so the motor won't burn out. There have been comments that the maker only freezes to soft serve state. This is NECESSARY because you can't get it out of the bowl or off the churn if it's frozen harder than that...I learned the hard way (pun intended). Cleanup is easy and fast, even though the parts must be hand washed. I LOVE IT! I don't think I will ever buy ice cream again!

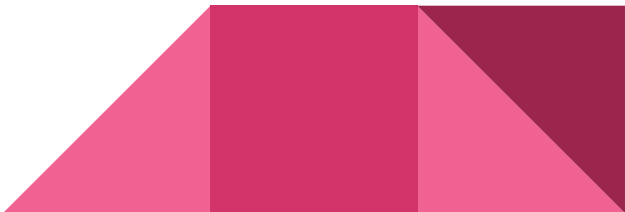
Actual Summary: Easy to use ice cream maker

Predicted Summary (old): Frozen cream maker

Predicted Summary (new): Best ice cream maker have ever tasted

Next steps

- Add precision, recall, bias/overfit analyses
- Feed result summaries into third party sentiment analyzer, see if analysis agrees with the labeled sentiment of the original data
- Run same analysis for more epochs, other NLP datasets
- Replace encoder LSTM with bi-LSTM, transformer
- Incorporate attention mechanism for full sentences
- Add different modalities, continuum to sentiment analysis output



Questions and Answers

Me: *uses machine learning*

Machine: *learns*

Me:

- Feel free to ask us any questions about our project!

