Northeastern University
Khoury College of Computer Sciences

TV Script Generation using Deep Learning

“The Office”

Rohit Awate
Rohan Devasthale
Shubham Bhagwat

TV Script Generation using Deep Learning

Rohit Awate
Rohan Devasthale
Shubham Bhagwat

Objectives/Goals

- Develop an NLP model for script generation for "The Office" TV show.
- Train RNNs and LSTMs on a large corpus of scripts to learn language patterns and structures.
- Explore NLP model capabilities for generating coherent and engaging text.
- Demonstrate the potential of these models to help writers with writer's block to generate ideas for their scripts.
- Generate coherent lines for a scene that capture the show's unique humor and style.

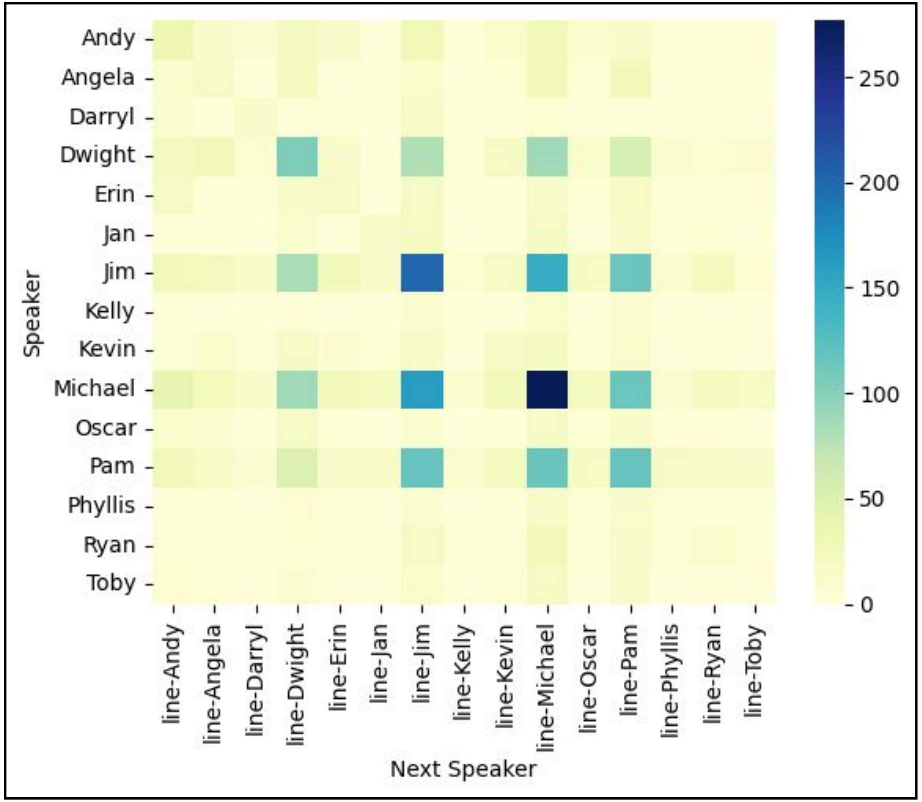
Tasks

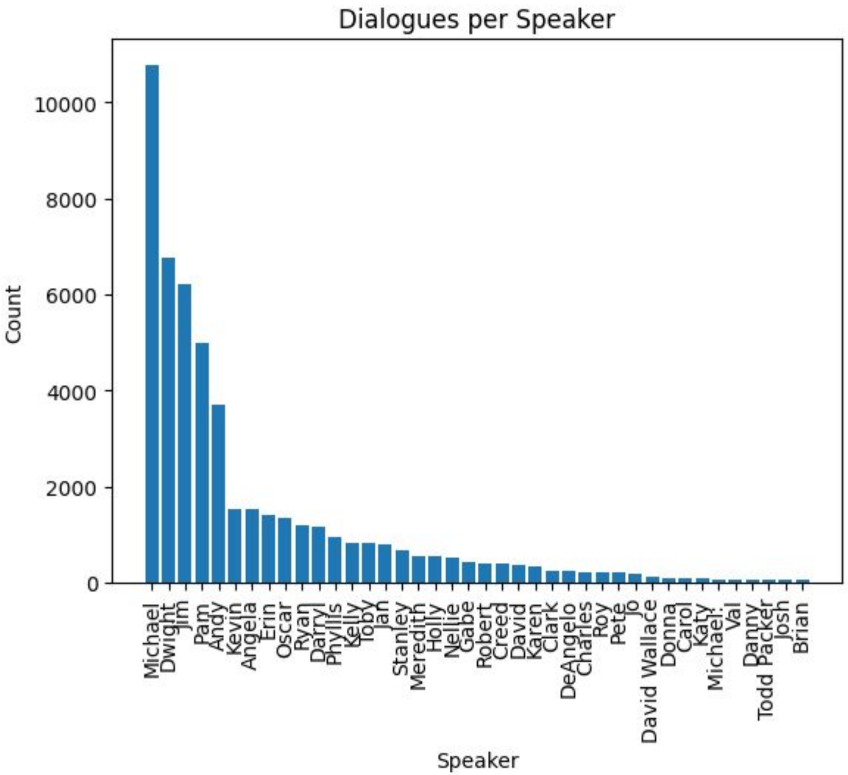
- Clean and preprocess the dataset of "The Office" scripts, including removing unwanted columns, tokenization, and vectorization.
- Train RNNs and LSTMs on the corpus of "The Office" scripts to learn language patterns and structures.
- Building a neural net classifier to classify the speaker of each line in the script.

Dataset

- Dataset consists of dialogues from all nine seasons of "The Office (US)"
- Preprocessing:
  - Tokenization of text
  - Conversion of tokens into integer sequences
  - Padding of sequences to have the same length
  - Filtering of stop words, rare words, and special characters
  - Removal of lines from deleted scenes
- Considered the top 40 speakers in the dataset; for the model to focus on the main characters of the show and their unique speech patterns.
- Focus on dialogues that were aired on TV

Dataset


Speaker
Andy
Angela
Darryl
Dwight
Erin
Jan
Jim
Kelly
Kevin
Michael
Oscar
Pam
Phyllis
Ryan
Toby
line-Andy
line-Angela
line-Darryl
line-Dwight
line-Erin
line-Jan
line-Jim
line-Kelly
line-Kevin
line-Michael
line-Oscar
line-Pam
line-Phyllis
line-Ryan
line-Toby
Next Speaker
250
200
150
100
50
0


Dialogues per Speaker
Count
10000
8000
6000
4000
2000
0
Michael
Dwight
Jim
Pam
Andy
Kevin
Angela
Erin
Oscar
Ryan
Darryl
Phyllis
Kelly
Toby
Jan
Stanley
Meredith
Holly
Nellie
Gabe
Robert
Creed
David
Karen
Clark
DeAngelo
Charles
Roy
Pete
Jo
David Wallace
Donna
Carol
Katy
Michael:
Val
Danny
Todd Packer
Josh
Brian
Speaker

Models

- LSTM and Bidirectional LSTM models were trained and compared for TV script generation.
- Both models were evaluated based on their performance in generating coherent lines of dialogue that "sound" like the characters from the show.
- A neural network classifier was used to measure the quality of the generated lines.
- Results showed that the Bidirectional LSTM model outperformed the LSTM model in terms of classification accuracy, indicating that it was better at capturing the nuances of the characters' speech patterns.
- The Bidirectional LSTM model demonstrated the potential of using advanced deep learning techniques for script generation tasks.

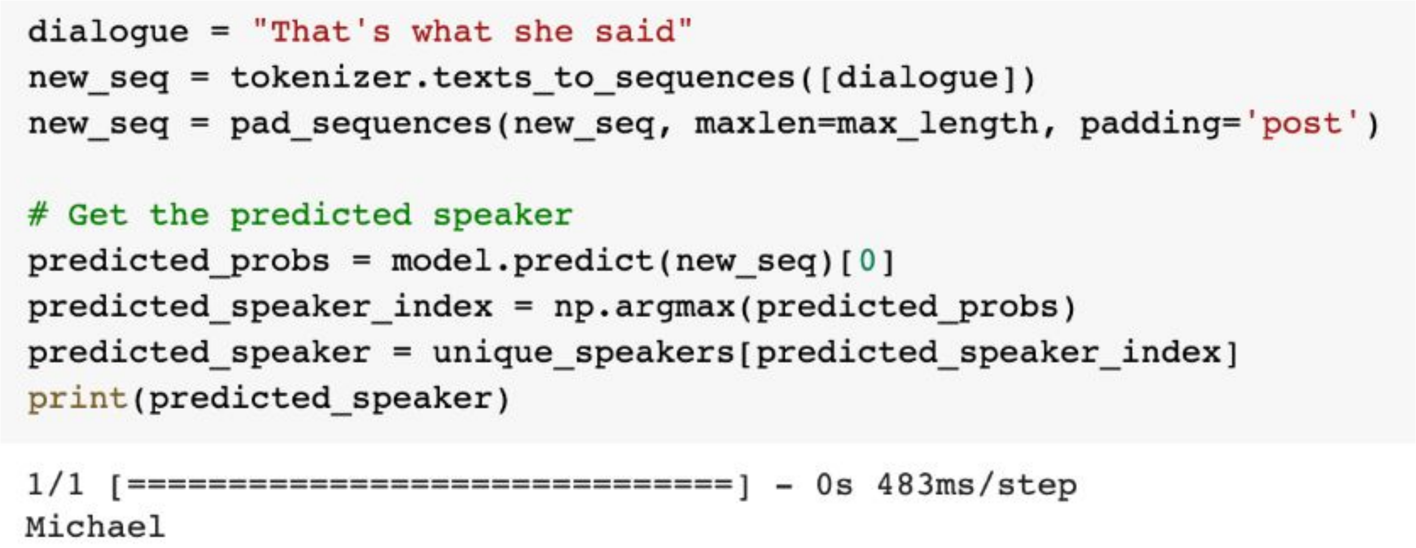
Evaluation

- Neural Network using LSTM is implemented which acts as classifier.
- It takes one dialogue at a time from generated script as input as tries to predicts the speaker.
- However, we do not entirely depend on this model for evaluation since the lines spoken by the characters are not distinctive enough to identify the speaker accurately.
- Therefore, we also check grammar, style, humor and coherence in general and determine whether the generated script is as close to the actual lines spoken in the show.

Results

```
dwight:  kids my deliciousity, uh, you did dwight, um, toby, um, jan, um, i don't know, i did, but,
thanks bye bye.
michael: what?
jim: no, actually, toby, did people, um, much, enjoy your baby can come and be baby
dwight: hey, how do you mean?
jim: you know, hooters? did you look... i just...
jim: so, thanks. i'm pretty cool.
michael: okay, let's do anything. i'm equipped-- i mean, we're just worried... i'm sorry i'm doing
it.
pam: oh, god.
michael: hey!
pam: michael, listen to me, it was just, incredibly late for tonight. but, i think she'd be great,
but it's coming to you, i promise it. i am so proud of that.
pam: yeah, but you know what? honestly, i blew the gifts. and i'm a regional manager. i think it's
really bad.
michael: oh, yeah.
pam: uh, you're the dad to make sure whatever you, " the company will still be a show of strengths,
right?
pam: hey!
michael: okay.
```

Results



```
dialogue = "That's what she said"
new_seq = tokenizer.texts_to_sequences([dialogue])
new_seq = pad_sequences(new_seq, maxlen=max_length, padding='post')

# Get the predicted speaker
predicted_probs = model.predict(new_seq)[0]
predicted_speaker_index = np.argmax(predicted_probs)
predicted_speaker = unique_speakers[predicted_speaker_index]
print(predicted_speaker)
```

```
1/1 [==============================] - 0s 483ms/step
Michael
```

Limitations

- The evaluation metrics that we used could be more better which would accurately judge the generated script.
- We believe that use of transformers could help in generation of more coherent and accurate scripts thereby depicting the actual humor and style as in actual show.

Thank You