ANALYSIS OF TEXT GENERATION MODELS TO DEPLOY A GENRE-BASED MOVIE DIALOGUE GENERATOR

INTRODUCTION

Movies have a way of capturing our imaginations, and dialogues play a big part in that. Recognizing their significance, our project aims to transform scriptwriting.

Our approach uses a sequence-to-sequence model, tailored to produce dialogues fitting specific movie genres. We experimented with three architectures: Bi-directional LSTM, Bi-directional GRU, and the GPT-2 Medium Transformer, ensuring robustness through cross-validation.

Crafting unique movie dialogues can help writers in script development and open new avenues in entertainment through automated content generation.

LITERATURE REVIEW

- **Seq2Seq with Attention for Dialogue Context:** Johnson and Lee (2021) used LSTM-based sequence-to-sequence architecture with multi-head attention, effectively capturing dialogue context in longer exchanges, resulting in coherent outputs, notably excelling in action and drama genres.

- **Genre-Embedded Bidirectional GRUs:** Fernandez and Gupta (2022) introduced Bidirectional GRUs with genre embeddings, enhancing dialogue authenticity by recognizing and encoding movie genres. Notably successful in thrillers and historical dramas, evaluated highly by human judges.

- **GPT-3's Genre Adaptability:** Davis et al. (2021) fine-tuned GPT-3 with genre tags using diverse movie scripts, demonstrating its adaptability across genres. Particularly effective in fantasy and science fiction, aligning well with intended narrative styles.

- **AI's Expanded Cinema Role:** Matthews and Rahman (2022) explored a transformer-based architecture beyond dialogues, generating plot twists and character arcs. Integrated sentiment analysis tailored emotional content, well-received for mystery and thriller genres, indicating potential for further exploration.

DATA SET

- **Total Entries: 304,713 rows**
- **Key Columns: 'text' (dialogue content) and 'genres'**

	Unnamed: 0	lineID	characterID	movieID	character_x	text	character_y	movie_x	gender	position	movie_y	year	rating	votes	genres
0	0	L1045	u0	m0	BIANCA	they do not!	BIANCA	10 things i hate about you	f	4	10 things i hate about you	1999	6.9	62847	['comedy', 'romance']
1	1	L1044	u2	m0	CAMERON	they do to!	CAMERON	10 things i hate about you	m	3	10 things i hate about you	1999	6.9	62847	['comedy', 'romance']
2	2	L985	u0	m0	BIANCA	i hope so.	BIANCA	10 things i hate about you	f	4	10 things i hate about you	1999	6.9	62847	['comedy', 'romance']
3	3	L984	u2	m0	CAMERON	she okay?	CAMERON	10 things i hate about you	m	3	10 things i hate about you	1999	6.9	62847	['comedy', 'romance']
4	4	L925	u0	m0	BIANCA	let's go.	BIANCA	10 things i hate about you	f	4	10 things i hate about you	1999	6.9	62847	['comedy', 'romance']

MODEL SELECTION

Bidirectional LSTMs:

- Bidirectional LSTMs combined with attention mechanisms and teacher forcing enhance contextual understanding. Bidirectional processing ensures comprehensive context grasp, attention focuses on relevant data, and teacher forcing accelerates convergence and stability.

Bidirectional GRUs:

- Bidirectional GRUs, efficient in capturing long-term dependencies, are used with pre-trained word embeddings (glove) for understanding context. This combination contributes to a holistic context comprehension.

GPT2LMHeadModel:

- GPT2LMHeadModel, a transformer-based architecture, excels in text generation with multi-head self-attention. Its extensive training data and 'Language Model' capabilities make it suitable for dialogue generation.

PRE-PROCESSING-1

Data Transformation:

- Utilized the Cornell Movie Dialogs Corpus, integrating character, movie, genre, and dialogue data from four different files into a cohesive genres.csv file.
- Removed nan values and duplicates from the genre column, ensuring data integrity.
- Grouped dialogues by movieID, crafting sequences where each input dialogue was paired with the subsequent dialogue as its target.
- Linked each dialogue with its corresponding movie genre.

PRE-PROCESSING-2

Tokenization and Sequencing:

- Initialized Keras's Tokenizer on both input and target dialogues, converting them into sequences of integers.
- Set a maximum sequence length for padding based on the 90th percentile of sequence lengths in our dialogues.
- Designed teacher input sequences by appending a zero at the start of target dialogues and eliminating their last element.
- One-hot encoded the genre labels using MultiLabelBinarizer.

Embedding Matrix Creation :

- Leveraged pre-trained GloVe embeddings.
- Built an embedding matrix tailored to the vocabulary of our dataset.

BI-DIRECTIONAL GRU

ARCHITECTURE

- Input Layer for Dialogues
- Input Layer for Genres
- Embedding Layer
- Bidirectional GRU Layer
- Dropout Layer
- Concatenation Layer
- GRU Decoder Layer
- Dropout Layer
- Output Dense Layer

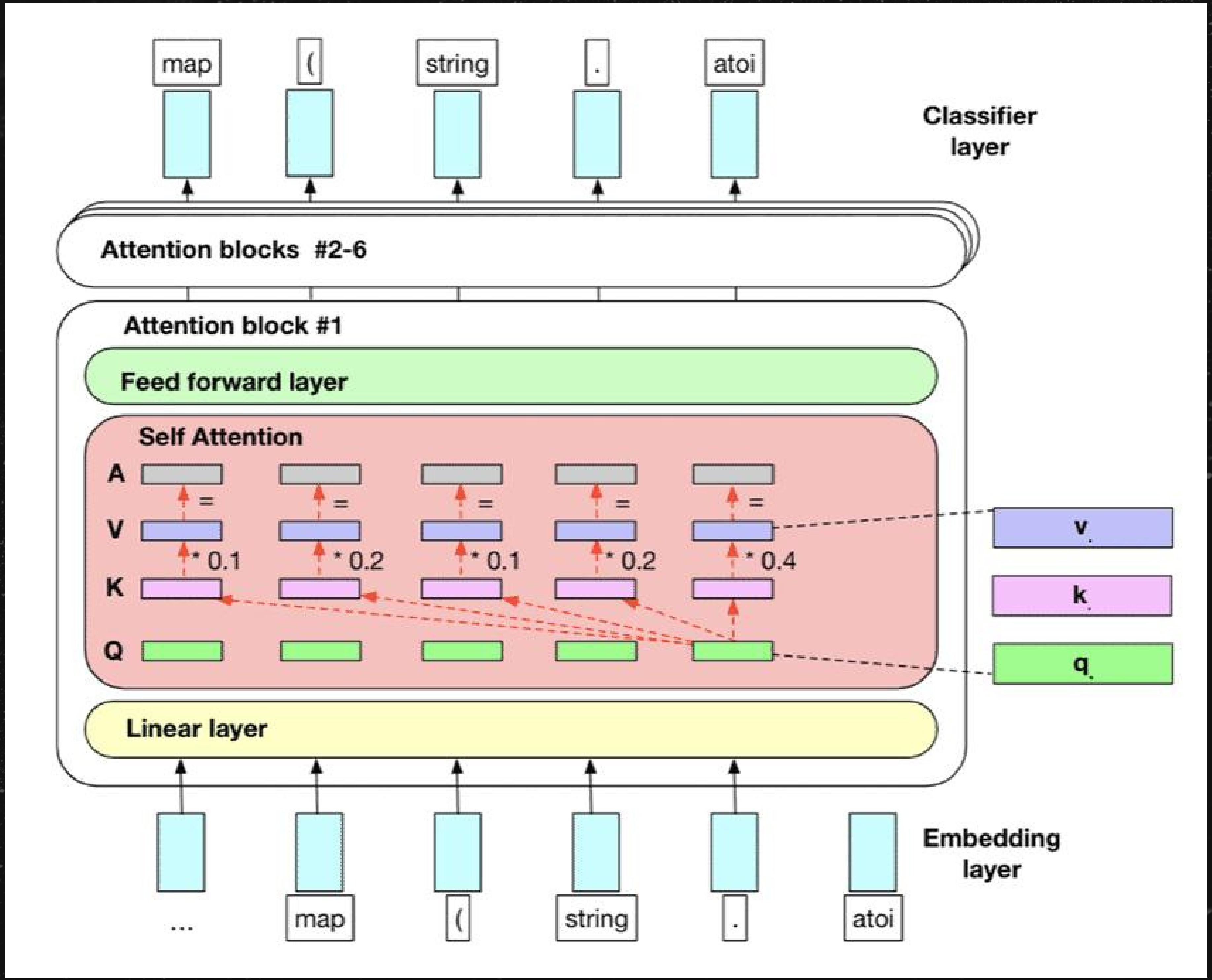
BI-DIRECTIONAL LSTM

ARCHITECTURE

- Input Layer for Dialogues
- Input Layer for Genres
- Input Layer for Teacher Forcing
- Embedding Layer
- Bidirectional LSTM
- Dropout Layer
- Attention Mechanism
- Embedding Layer for Teacher Forcing
- Concatenation Layer
- LSTM Decoder Layer
- Dropout Layer
- Output Dense Layer

GPT2LMHEAD MODEL

ARCHITECTURE



PERFORMANCE COMPARISSION

MODEL	SPARSE CATEGORICAL CROSS ENTROPY LOSS	BLEU SCORE	HUMAN EVALUATION
BIDIRECTIONAL LSTM WITH ATTENTION	2.772	0.09	BAD
BIDIRECTIONAL GRU	2.371	0.7	OKAY
GPT2LMHEADMODEL	8.351	0.992	GREAT

RESULT

BI-DIRECTIONAL LSTM

```
I thought you were cool
Generated Dialogue: i
```

BI-DIRECTIONAL GRU

```
You: understand it
Bot: Understand it
```

GPT2LMHead Model

```
You: fact
Bot: Fact that the government has been able to k
You: back
Bot: Back to the beginning of the game
```

ABLATION

- For the ablation study, we first split the data set into training, validation, testing and teacher input.We used a 70-15-15 split. Next, we used K-Fold cross validation to tune hyper parameters, taking cross validation over average validation loss on the validation data set and training each model for 2 epochs due to time and GPU constraints.

- For the LSTM model, the best parameters were found to be: Dropout of 0.5, L1 and L2 regularization of 0.001, Embedding dimension of 32, and 32 LSTM units.

- For the GRU model, the best parameters were found to be: Dropout of 0.1, L1 and L2 regularization of 0.0001, Embedding dimension of 32, and 64 GRU units.

CONCLUSION

- **Comparison of Models:** Three distinct models were evaluated for genre-based movie dialogue generation, each showcasing its strengths and weaknesses.
- **Bidirectional LSTM with Attention:** Despite its theoretical promise, this model faced challenges in practical application for the given task.
- **Bidirectional GRU's Effectiveness:** Simplifying the architecture by using Bidirectional GRU resulted in better and more interpretable outcomes than the LSTM with Attention.
- **Superiority of GPT2LMHeadModel:** The GPT2LMHeadModel outperformed others, highlighting the advantages of fine-tuning large-scale pre-trained models. This model excelled in generating technically accurate dialogues and capturing genre-specific nuances.

FUTURE WORK

Future research will focus on refining model architecture, incorporating data augmentation, enhancing generalization techniques, and a deeper exploration of genre-specific performance metrics.

THANK YOU