CHARACTER LEVEL RNN MODEL

SHERLOCK HOLMES: A CHARACTER-BASED RNN ADVENTURE

Tricia S

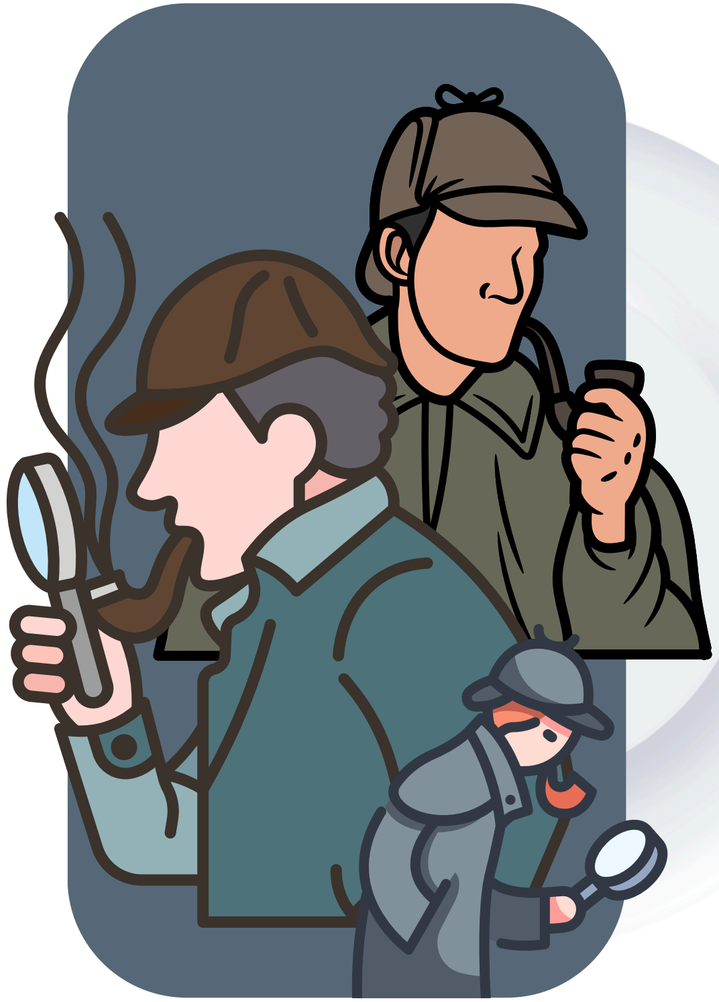

TABLE OF CONTENT

INTRODUCTION

In this project, I aim to develop an optimal recurrent neural network model capable of predicting and generating the next character in a sequence of words. And I would dive into the complexities of natural language generation, simulating how Recurrent Neural Networks (RNNs) might generate English text. Unlike human language generation, which naturally incorporates tone, emotion, and style, this project focuses on the computational challenges of modeling such a system.

PROBLEM STATEMENT

The goal is to create a character-level language model that predicts the next character in a sequence. This involves managing text sequences, long-term dependencies, and efficiency. The project explores RNN architectures like LSTMs and GRUs, with thorough hyperparameter tuning to capture English language patterns.

OVERALL WORKFLOW

Load Data

Data Processing

Build the model

Train the model

Generate Text

Analyse Output

BEST MODEL

```
acc: 0.6111 - loss: 1.2529 - val acc: 0.5464 - val loss: 1.5928
```

```python
# Build the Model
model13 = keras.models.Sequential()
model13.add(layers.LSTM(256, input_shape=(sequence_length, len(chars))))
model13.add(layers.Dense(len(chars), activation='softmax'))

optimizer = optimizers.Adam(learning_rate=0.01)
model13.compile(loss='categorical_crossentropy', optimizer=optimizer, metrics=['acc'])
model13.summary()
```

SAMPLE OUTPUT

```
Vectorization...
[[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0]]
20
The predicted next character for ['the game is afoo'] is  l
```

```
Vectorization...
[[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0]]
26
The predicted next character for ['he nodde'] is  r
```

FUTURE RECOMMENDATIONS

Moving forward, I plan to refine my approach by exploring hyperparameter tuning (batch size, L1/L2 regularization, dropout variations, increase epochs), investigating larger network architectures and data preprocessing adjustments (step parameter), conducting a more comprehensive model evaluation, experimenting with ensemble methods, and focusing on LSTM optimization.

01 COMPLEX MODELS

02 MORE TUNING

03 DATA AUGMENTATION

04 GPU

THANK YOU