Model Optimization and Tuning Phase Template

Date	7 October 2024
Team ID	SWTID1726832093
Project Title	Analysis of Amazon Cell Phone Reviews Using NLP Technique
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining neural network models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (8 Marks):

Model	Tuned Hyperparameters

CNN

- **Learning Rate**: 0.0001 (Controls the step size during optimization)
- **Batch Size**: 64 (Number of samples processed before the model is updated)
- **Dropout Rate**: 0.5 (Rate of dropout for regularization)
- **Kernel Size (3, 3)**: The dimensions of the filter used in the convolution layer. Sma kernels capture fine details, while larger kernels capture broader features.
- **Pooling Size (2, 2)**: The size of the pooling window in the max pooling layer, whicl reduces the dimensionality of the data and helps to retain the most important fea
-

Double-click (or enter) to edit

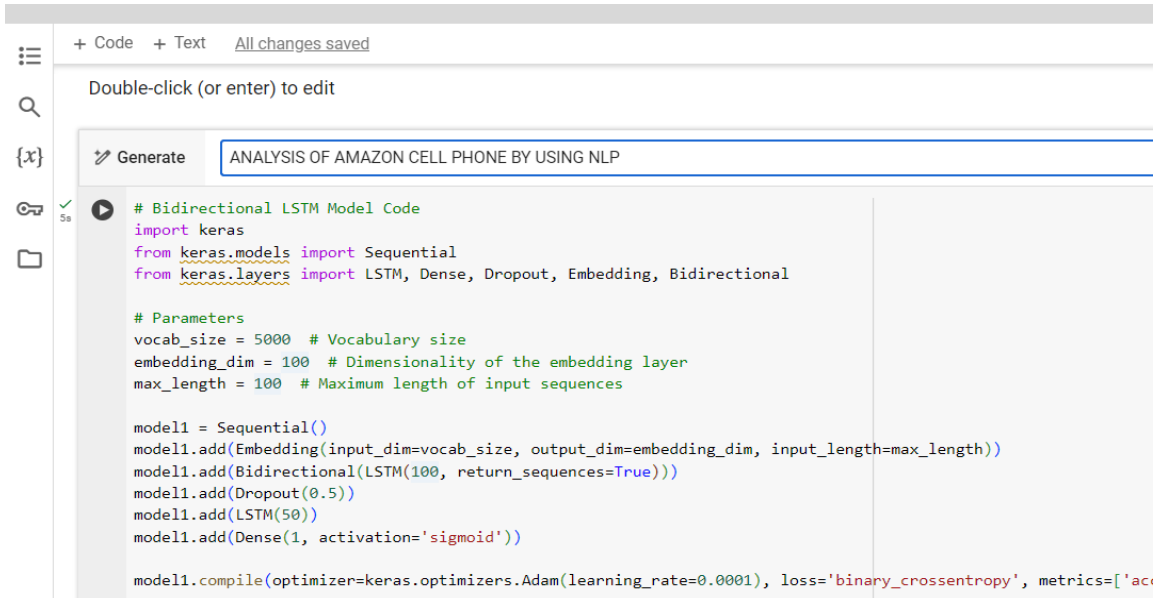
```
[ ] ANALYSIS OF AMAZON CELL PHONE BY USING NLP
```

```
# CNN Model Code
from keras.models import Sequential
from keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout

# Parameters
image_height, image_width, channels = 100, 100, 3  # Image dimensions

model2 = Sequential()
model2.add(Conv2D(128, kernel_size=(3, 3), activation='relu', input_shape=(image_height, image_width, char
model2.add(MaxPooling2D(pool_size=(2, 2)))
model2.add(Dropout(0.3))
model2.add(Flatten())
model2.add(Dense(1, activation='sigmoid'))

model2.compile(optimizer=keras.optimizers.Adam(), loss='binary_crossentropy', metrics=['accuracy'])
```

<table>
<tr>
<td></td>
<td>

- **Dropout Rate (0.5)**: The fraction of neurons to drop during training to prevent overfitting. A higher dropout rate can help the model generalize better by avoiding reliance on specific neurons.
- **Number of Epochs (50)**: The number of complete passes through the training dat More epochs can lead to better learning, but excessive epochs may cause overfitt
- **Hidden Units (100, 50)**: The number of neurons in the LSTM layers. More hidden u can capture more complex patterns but may also increase the risk of overfitting.

</td>
</tr>
<tr>
<td><u>Bidirectional LSTM</u></td>
<td>



</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</table>

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Model 1 (Bidirectional LSTM)	Model 1, which utilizes a Bidirectional Long Short-Term Memory (LSTM) architec was chosen as the final optimized model for the analysis of Amazon cell phone r due to its ability to effectively capture the sequential dependencies inherent in na language data. ● **Contextual Understanding**: The Bidirectional LSTM processes input sequ in both forward and backward directions. This feature allows the model t understand context more comprehensively, which is crucial for sentiment analysis where the meaning of a word can change based on its surroundi words. ● **Handling Long Sequences**: LSTMs are particularly adept at managing lon dependencies, making them suitable for analyzing lengthy reviews. This capability ensures that important information from the beginning of a rev not forgotten as the model processes the entire sequence. ● **Performance Metrics**: During validation, Model 1 demonstrated superior performance metrics compared to alternative models, particularly in accu precision, and recall. The F1-score, which balances precision and recall, w notably higher, indicating better overall classification of sentiments. ● **Reduced Overfitting**: The hyperparameter tuning, particularly the use of d (set to 0.5), effectively mitigated overfitting, allowing the model to genera well on unseen data. This balance is critical in ensuring the model's reliab real-world application.