PYTHON PROGRAM TO BUILD AN RNN WITH KERAS

Aim:

To build a RNN(Recurrent Neural Networks) with keras in python.

Procedure:

1. Import NumPy, TensorFlow, IMDB dataset utilities, RNN layers, and Matplotlib for plotting.
2. Set a random seed for reproducibility in NumPy and TensorFlow.
3. Load the IMDB dataset, restricting to the top 10,000 words and specifying a maximum sequence length of 200.
4. Pad sequences to ensure uniform input length across all samples.
5. Build a Sequential model with an Embedding layer, a SimpleRNN layer, and a Dense output layer for binary classification.
6. Compile the model using Adam optimizer and binary cross-entropy loss, with accuracy as the evaluation metric.
7. Print the model summary to check the architecture and layer details.
8. Train the model for 5 epochs with a batch size of 64, including validation data.
9. Evaluate the model on the test data and print the accuracy.
10. Plot training and validation accuracy and loss over epochs using Matplotlib.

Code:

```python
# Import necessary libraries
import numpy as np
import tensorflow as tf
from tensorflow.keras.datasets import imdb
from tensorflow.keras.preprocessing import sequence
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Embedding, SimpleRNN, Dense
import matplotlib.pyplot as plt

# Set random seed for reproducibility
np.random.seed(42)
tf.random.set_seed(42)

# Load the IMDB dataset
max_features = 10000  # Number of words to consider as features
maxlen = 200  # Maximum length of input sequences

(train_data, train_labels), (test_data, test_labels) =
imdb.load_data(num_words=max_features)

# Pad sequences to ensure consistent input size
train_data = sequence.pad_sequences(train_data, maxlen=maxlen)
test_data = sequence.pad_sequences(test_data, maxlen=maxlen)

# Build the RNN model
model = Sequential()
model.add(Embedding(input_dim=max_features, output_dim=128,
input_length=maxlen))
model.add(SimpleRNN(128, activation='relu'))
model.add(Dense(1, activation='sigmoid'))
```

```python
# Compile the model
model.compile(optimizer='adam',
        loss='binary_crossentropy',
        metrics=['accuracy'])

# Print model summary
model.summary()

# Train the model
history = model.fit(train_data, train_labels,
            epochs=5,
            batch_size=64,
            validation_split=0.2,
            verbose=1)

# Evaluate the model
test_loss, test_acc = model.evaluate(test_data, test_labels)
print(f'Test accuracy: {test_acc}')

# Plot training and validation accuracy and loss
plt.figure(figsize=(12, 4))

# Accuracy plot
plt.subplot(1, 2, 1)
plt.plot(history.history['accuracy'], label='Training Accuracy')
plt.plot(history.history['val_accuracy'], label='Validation Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()
```

```python
# Loss plot
plt.subplot(1, 2, 2)
plt.plot(history.history['loss'], label='Training Loss')
plt.plot(history.history['val_loss'], label='Validation Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

plt.show()
```

Output:

```
python -u "/Users/manoj/Documents/PYTHON/DLC/RNN.py"
Model: "sequential"

┌─────────────────────────────────┬───────────────────────┬─────────────────┐
│ Layer (type)                    │ Output Shape          │        Param #  │
├─────────────────────────────────┼───────────────────────┼─────────────────┤
│ embedding (Embedding)           │ ?                     │    0 (unbuilt)  │
├─────────────────────────────────┼───────────────────────┼─────────────────┤
│ simple_rnn (SimpleRNN)          │ ?                     │    0 (unbuilt)  │
├─────────────────────────────────┼───────────────────────┼─────────────────┤
│ dense (Dense)                   │ ?                     │    0 (unbuilt)  │
└─────────────────────────────────┴───────────────────────┴─────────────────┘

 Total params: 0 (0.00 B)
 Trainable params: 0 (0.00 B)
 Non-trainable params: 0 (0.00 B)
Epoch 1/5
313/313 ──────────────── 20s 64ms/step - accuracy: 0.5807 - loss: 0.6742 - val_accuracy: 0.7862 - val_loss: 0.4629
Epoch 2/5
313/313 ──────────────── 20s 65ms/step - accuracy: 0.7752 - loss: 0.4695 - val_accuracy: 0.6738 - val_loss: 0.5825
Epoch 3/5
313/313 ──────────────── 20s 63ms/step - accuracy: 0.7908 - loss: 0.4402 - val_accuracy: 0.8024 - val_loss: 0.4511
Epoch 4/5
313/313 ──────────────── 21s 68ms/step - accuracy: 0.8636 - loss: 0.3253 - val_accuracy: 0.7790 - val_loss: 0.5107
Epoch 5/5
313/313 ──────────────── 20s 64ms/step - accuracy: 0.9018 - loss: 0.2383 - val_accuracy: 0.7898 - val_loss: 0.5402
782/782 ──────────────── 7s 9ms/step - accuracy: 0.7872 - loss: 0.5587
Test accuracy: 0.787880003452301
2024-09-16 16:21:00.459 Python[2844:1632883] +[IMKClient subclass]: chose IMKClient_Legacy
2024-09-16 16:21:00.459 Python[2844:1632883] +[IMKInputSession subclass]: chose IMKInputSession_Legacy
```

Result:

Thus, to build a RNN using Keras in python has been completed successfully.