Universidade Federal do Rio de Janeiro
Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia

Programa de Engenharia de Sistemas e Computação

CPS769 - Introdução à Inteligência Artificial e Aprendizagem Generativa

Prof. Dr. Edmundo de Souza e Silva (PESC/COPPE/UFRJ)
Profa. Dra. Rosa M. Leão (PESC/COPPE/UFRJ)
Participação Especial: Gaspare Bruno (Diretor Inovação, ANLIX)

Lista de Exercícios 1a

Luiz Henrique Souza Caldas
email: lhscaldas@cos.ufrj.br

10 de julho de 2024

Questão 1

Esse exemplo simples é para auxiliar a discussão do artigo "Serial Order A Parallel Distributed Processing Approach" que todos já devem ter lido. O objetivo é prever um padrão de figura, por exemplo um quadrado, usando uma Rede Neural Recorrente (RNN). Fornecemos o código em Python de um exemplo de geração do padrão 2-D de quadrados e treinamento de uma RNN para prever a sequência cíclica [0, 25, 0, 25], [0, 75, 0, 25], [0, 75, 0, 75], [0, 25, 0, 75], [0, 25, 0, 25].

1. Entenda o código e explique qual a RNN que ele modela (faça o desenho). Explique a parte do código que define a RNN.

   Resposta:

   O código pode ser explicado dividindo-o em 6 partes:

   (a) Definição do caminho quadrado: O código define um conjunto de coordenadas que formam um caminho quadrado na variável *square_path*.

```
square_path = np.array([
    [0.25, 0.25],
    [0.75, 0.25],
    [0.75, 0.75],
    [0.25, 0.75],
    [0.25, 0.25]
])
```

   (b) Geração dos dados de treinamento: O caminho quadrado é repetido várias vezes para formar os dados de treinamento.

```
num_repeats = 4
data = np.tile(square_path, (num_repeats, 1))
x_train = data[:-1].reshape(-1, 1, 2)
y_train = data[1:].reshape(-1, 2)
```

   (c) Definição e compilação do modelo RNN: O modelo RNN é definido usando uma camada LSTM (*long short-term memory*) seguida de uma camada densa e depois é compilado configurando o algoritmo ADAM como otimizador e o Erro Médio Quadrático como função de perda.

```
model = models.Sequential([
    layers.LSTM(50, activation='relu', input_shape=(num_repeats, 2)),
    layers.Dense(2)
])
model.compile(optimizer='adam', loss='mse')
```

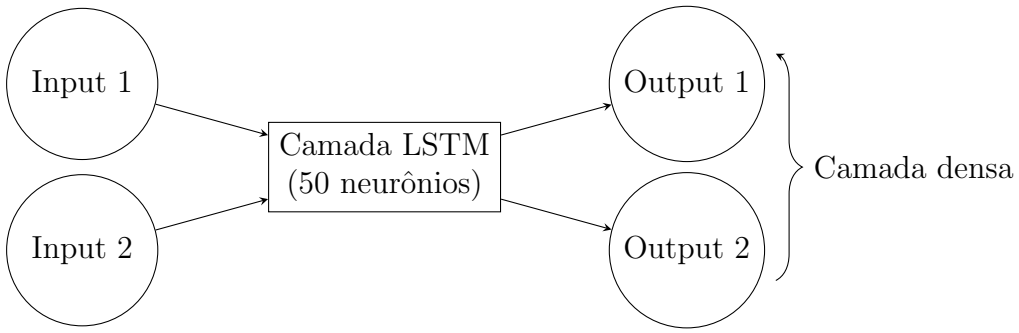


Figura 1: Diagrama de uma Rede Neural Recorrente (RNN) com dois neurônios de entrada e dois neurônios de saída (camada densa) e uma camada LSTM modelada no código fornecido.

(d) Treinamento do modelo: O modelo é treinado com os dados gerados, utilizando inicialmente 300 épocas.

```
model.fit(x_train, y_train, epochs=300, verbose=0)
```

(e) Geração das previsões: As previsões são geradas.

```
predictions = model.predict(x_train[:5])
```

(f) Plotagem dos resultados: As previsões são plotadas e comparadas com o caminho original.

```
plt.plot(data[:, 0], data[:, 1], label='Original Path', linestyle='
    dashed', color='gray')
plt.plot(predictions[:, 0], predictions[:, 1], label='Predicted Path',
     color='blue')
plt.scatter(square_path[:, 0], square_path[:, 1], color='red')
plt.legend()
plt.show()
```

2. Treine a rede. Aprenda como fazer, e explique.

   Resposta:

   Como dito no passo (d) do item anterior, o treinamento é ralizado utilizando a função "fit"do modelo. Utilizando a configuração inicial, com 300 épocas, o treinamento demorou cerca de 11 segundos.

3. Faça a previsão de algumas trajetórias, quando o ponto inicial varia. O que você conclui?

   Resposta:

   Foi feita a previsão para o ponto inicial original do código dado, ($x_1 = 0.25$ e $x_2 = 0.25$) e depois foram testados os pontos ($x_1 = 0.00$ e $x_2 = 0.00$), ($x_1 = 0.50$ e $x_2 = 0.50$) e ($x_1 = 0.75$ e $x_2 = 0.25$).

Figura 2: Previsão para ($x_1 = 0.25$ e $x_2 = 0.25$)

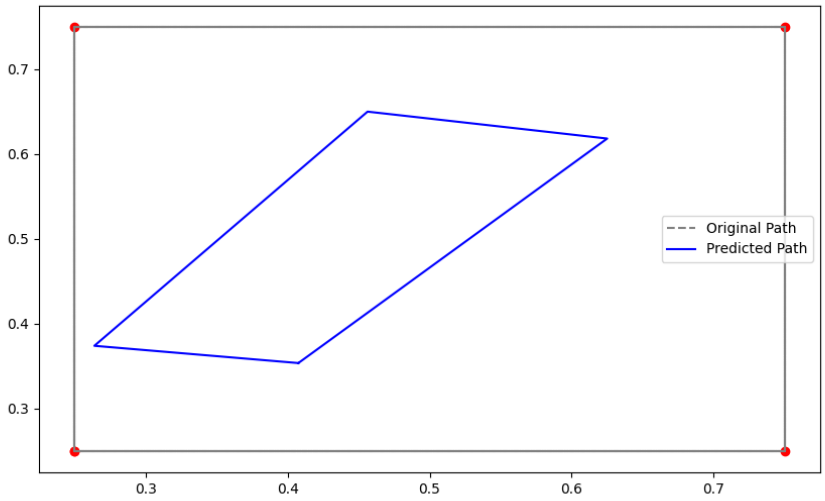


Figura 3: Previsão para ($x_1 = 0.00$ e $x_2 = 0.00$)

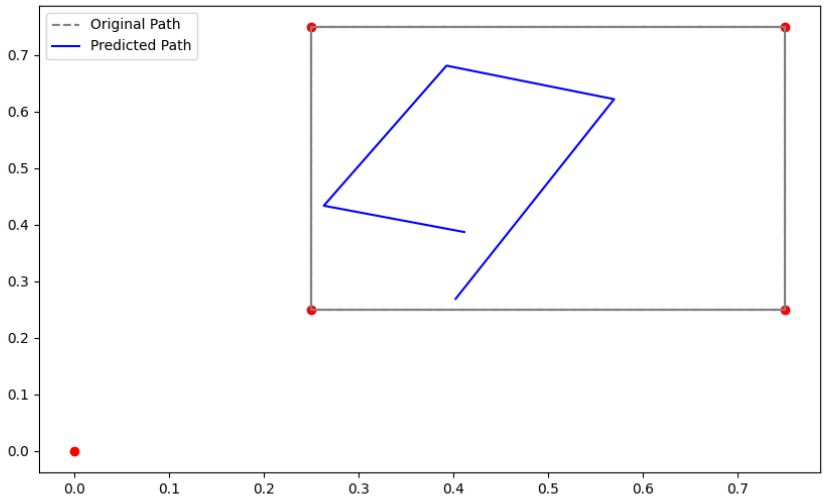


Figura 4: Previsão para ($x_1 = 0.50$ e $x_2 = 0.50$)

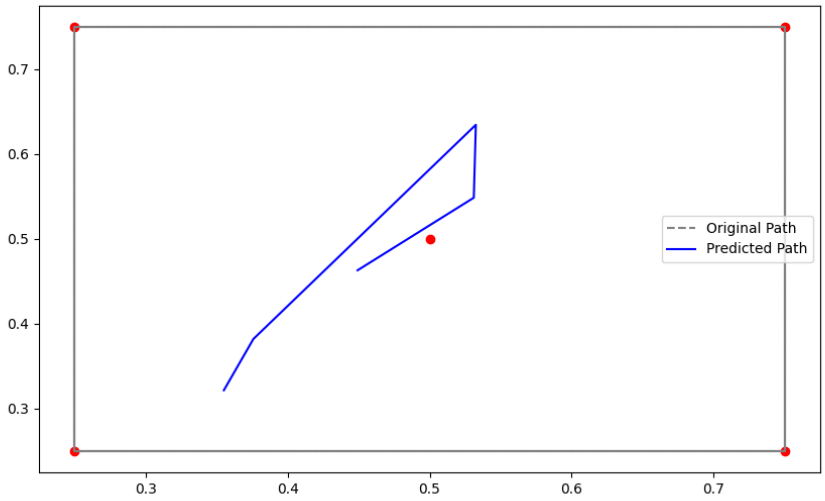
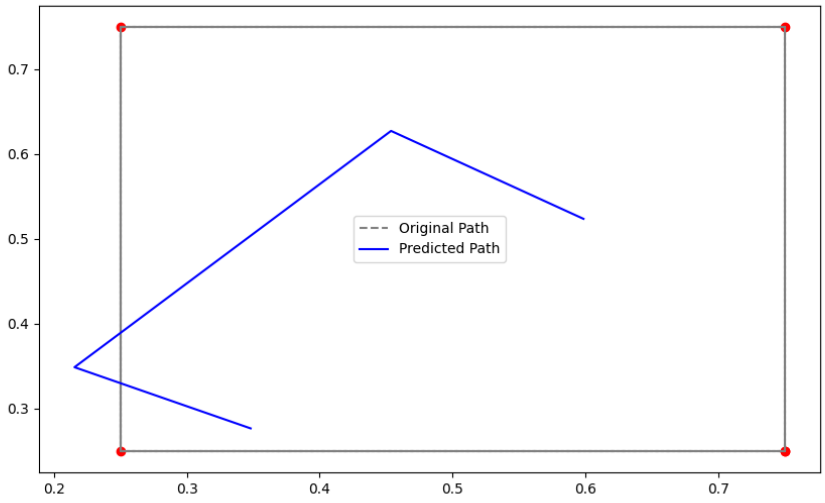


Figura 5: Previsão para ($x_1 = 0.75$ e $x_2 = 0.25$)



O tempo de execução foi cerca de 11,4 segundos, o que demonstra que a maior parte do custo computacional é para o treinamento da rede, sendo o tempo para a previsão deste problema específico irrelevante. O valor inicial ($x_1 = 0.25$ e $x_2 = 0.25$) foi o que apresentou o melhor resultado, uma vez que o polígono formado pelo menos foi fechado.

4. Modifique a RNN usada e observe o que acontece.

   Resposta:

Para este teste, o ponto inicial foi retornado para a configuração original ($x_1 = 0.25$ e $x_2 = 0.25$) e foram testadas diferentes combinações de épocas e número de repetições.

Tabela 1: Resultados das modificações

Nº Repetições	Épocas	Tempo Treinamento (s)	Tempo Total (s)
4	300	12	12.2
40	300	13.5	13.7
400	300	26.4	26.6
40	600	19.6	19.8
40	900	29.2	29.1

Pela tabela 1 é possível observar o aumento do tempo de execução, mais específicamente do tempo de treinamento, tanto com o aumento do número de repetições quanto com o aumento do número de épocas. Os resultados das previsões podem ser visualizados nas figuras abaixo.

Figura 6: Previsão para 4 repetições e 300 épocas

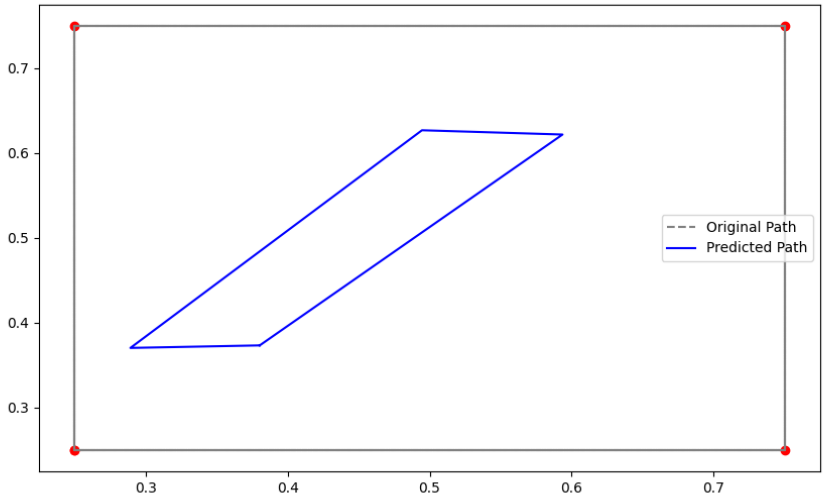


Figura 7: Previsão para 40 repetições e 300 épocas

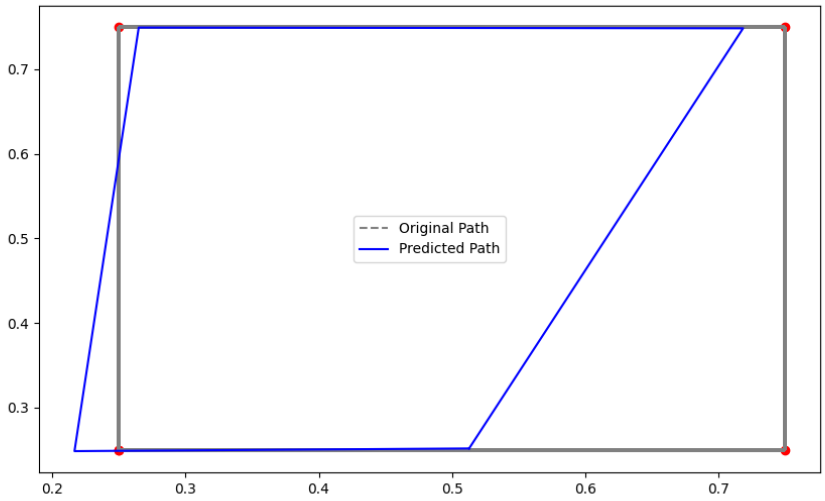


Figura 8: Previsão para 400 repetições e 300 épocas

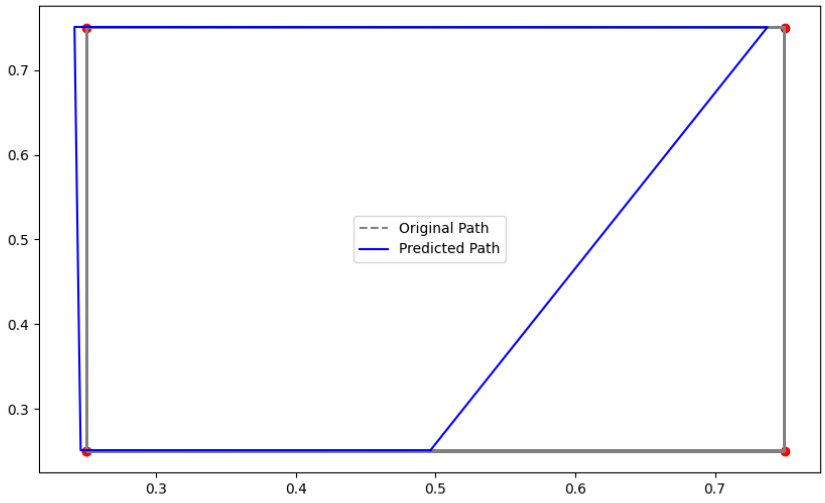


Figura 9: Previsão para 40 repetições e 600 épocas

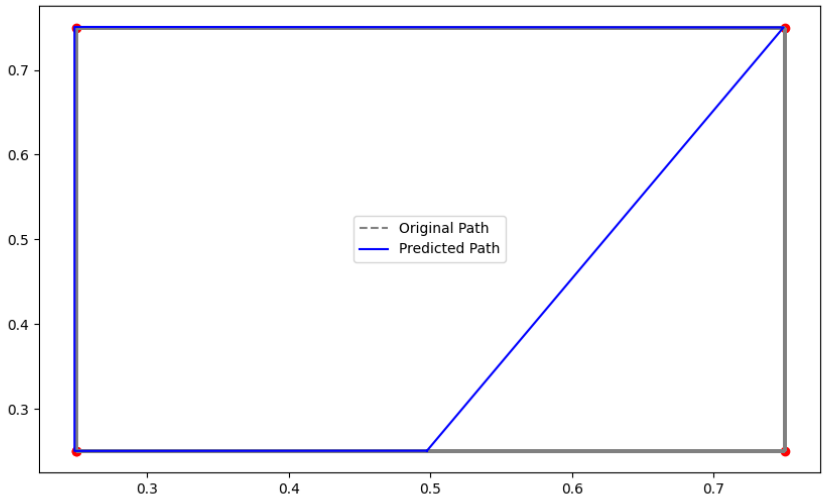
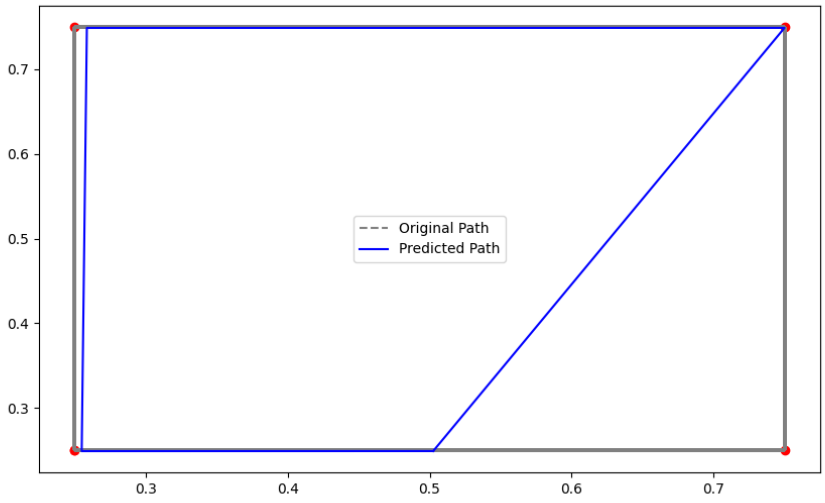


Figura 10: Previsão para 40 repetições e 900 épocas



Pelas figuras foi possível observar que o impacto do aumento do número de repetições na precisão da previsão foi muito maior que o impacto do aumento do número de épocas, o que demonstra a importância do tamanho do dataset de treinamento para o resultado de suas previsões.

5. Quais os pontos principais que você concluiu do artigo "Serial Order A Parallel Distributed Processing Approach"?

Resposta:

A teoria de Michael I. Jordan sobre ordem serial em sequências de ações usa redes neurais para entender e reproduzir a ordem das ações ao longo do tempo. Essas redes mantêm uma "memória"do que já aconteceu, usando conexões que alimentam as saídas de volta para as entradas, ajudando a lembrar das ações passadas. A rede aprende ajustando seus parâmetros para reduzir erros entre o que foi previsto e o que realmente aconteceu.

Isso faz com que a rede consiga generalizar a partir de sequências aprendidas e continuar funcionando bem, mesmo com pequenas perturbações. Essencialmente, a rede se torna uma memória dinâmica que pode voltar às suas trajetórias aprendidas, garantindo que as sequências de ações sejam produzidas corretamente, mesmo começando de pontos diferentes.

Código

O código abaixo encontra-se no repositório https://github.com/lhscaldas/cps769-ai-gen, bem como o arquivo LaTex com o relatório.

Código 1: código fornecido completo com algumas modificações

```
import time
start_time = time.time()

# Modify the Python Script to Disable GPU
import os
os.environ['CUDA_VISIBLE_DEVICES'] = '-1'

# to run the python script: python3 rnn_cyclic_sequence.py

import numpy as np
import tensorflow as tf
from tensorflow.keras import layers, models
import matplotlib.pyplot as plt

# Define the square path coordinates
square_path = np.array([
    [0.25, 0.25],
    [0.75, 0.25],
    [0.75, 0.75],
    [0.25, 0.75],
    [0.25, 0.25]
])

# Generate the training data by repeating the square path
num_repeats = 40
data = np.tile(square_path, (num_repeats, 1))

# Initial point modification
mod_square_path = square_path
mod_square_path[0] = [0.25, 0.25]
mod_data = np.tile(mod_square_path, (num_repeats, 1))

```

```
# Prepare training data
x_train = mod_data[:-1].reshape(-1, 1, 2)
y_train = data[1:].reshape(-1, 2)

# Define the RNN model
model = models.Sequential([
    layers.LSTM(50, activation='relu', input_shape=(num_repeats, 2)),
    layers.Dense(2)
])

# Compile the model
model.compile(optimizer='adam', loss='mse')

# Train the model
model.fit(x_train, y_train, epochs=900, verbose=0)

fit_time = time.time()
print(f"Tempo de execução até o treinemento: {fit_time - start_time} segundos")

# Generate predictions
predictions = model.predict(x_train[:5])

predict_time = time.time()
print(f"Tempo de execução até a predição: {predict_time - start_time} segundos"
    )

# Plot the results
plt.figure(figsize=(10, 6))
plt.plot(data[:, 0], data[:, 1], label='Original Path', linestyle='dashed',
    color='gray')
plt.plot(predictions[:, 0], predictions[:, 1], label='Predicted Path', color='
    blue')
plt.scatter(square_path[:, 0], square_path[:, 1], color='red')
plt.legend()
plt.show()
```