

Introduction to Artificial Intelligence and Generative Learning

CPS 769

Segundo Trimestre de 2024

Professores:

Edmundo de Souza e Silva (PESC/COPPE/UFRJ)

Rosa M. Leão (PESC/COPPE/UFRJ)

Participação Especial: Gaspare Bruno (Diretor Inovação, ANLIX)

Lista de Exercícios 3

ATENÇÃO! Faça as listas de forma que TODAS AS RESPOSTAS sejam DEVIDAMENTE COMENTADAS (passos para se chegar a resposta). Ser claro e objetivo facilitará organizar as ideias para as discussões em classe.

Para facilitar escrever a lista de forma clara, é possível traduzir equações escritas a mão para LaTeX:

<https://mathpix.com/>, ver também

https://www.overleaf.com/learn/latex/Questions/Are_there_any_tools_to_help_transcribe_mathematical_formulae_into_LaTeX%3F.

Questão 1

O objetivo deste trabalho é treinar o que foi mostrado em classe nas últimas aulas, isto é, como usar o chatbots para chamada a outras ferramentas ou banco de dados. Especificamente, você trabalhará com as técnicas de chamada de função aprendidas em aula utilizando uma base de dados que fornece informações sobre o tempo, disponibilizada no Kagge pelo INMET.

Nesta tarefa, você deverá implementar (e rodar) os passos necessários para interpretar e responder perguntas em linguagem natural.

1. Implemente o que for necessário para responder perguntas como:
 - (a) “Qual o dia mais quente do ano de 2001”?
 - (b) “Qual o dia mais quente do ano de 2021”?
 - (c) “Qual a média da temperatura no ano de 2021”?
2. Tente também responder perguntas complexas como:
 - (a) “Qual a média do mês de janeiro de 2010”?
 - (b) “Janeiro de 2021 fez frio?”?
 - (c) “compare a média de temperatura do ano de 2001 com a de 2021”?

Base de dados:

<https://www.kaggle.com/datasets/gregoryoliveira/brazil-weather-information-by-inmet>

Aprenda a trabalhar com os dados em python. A base de dados é apresentada em CSV, onde cada linha representa um dia e uma hora. A base possui outras informações, além da temperatura. Para trabalhar de maneira prática com um arquivo csv em python, podemos utilizar o pandasql:

instalar:

```
pip install pandasql
```

exemplo de código:

```
from pandasql import sqldf
```

```
import pandas as pd

colunas = ['DATA (YYYY-MM-DD)', 'Hora UTC', 'TEMPERATURA DO PONTO DE ORVALHO (°C)']

df = pd.read_csv('datasets/weather_2024.csv', usecols=colunas)
df.columns = ['DATA', 'HORA', 'TEMP']
query = 'select DATA, max(TEMP) from df'
print(sqldf(query))
```

Documentação para uso do pandasql:

<https://www.kaggle.com/code/dataraj/sql-on-pandas-dataframe-using-pandasql>