PHASE 2- CREATE A CHATBOT IN PYTHON

Introduction:

In this section, we will explore the innovative solution that has been developed to address the problem statement established in the earlier phase of this project. We will discuss the design thinking process that led to this innovation and the significance of innovation in problem-solving.

Background:

Recap the problem statement and the challenges it poses. Briefly introduce the context in which the innovation is taking place.

Innovative Solution:

Describe the innovative solution that has been designed to tackle the identified problem. Explain how this solution represents a creative and forward-thinking approach to addressing the issue.

Rationale:

Explain the reasoning behind choosing this specific solution. How does it align with the insights gained during the design thinking phase?

Key Features:

Enumerate the key features and components of the innovative solution. Elaborate on how each of these features contributes to solving the problem and enhancing the user experience.

Feature 1

- Describe the first key feature.

- Explain its role in the solution.

Feature 2

- Describe the second key feature.
- Explain its role in the solution.

Technology and Tools:

Discuss the technologies and tools that will be leveraged to implement the innovative solution. Justify the choice of these technologies and tools and how they promote innovation.

Technology 1

- Describe the first technology.
- Explain why it's essential for the solution.

Technology 2

- Describe the second technology.
- Explain its relevance to the solution.

User Experience

Explain how the innovative solution enhances the overall user experience. Highlight the benefits and improvements compared to existing solutions.

User-Centered Design

- Describe how user feedback and needs were incorporated into the solution.
- Explain how this ensures a user-friendly experience.

Testing and Validation

Detail the testing process that the innovative solution underwent. Share insights from user feedback and validation results. Discuss any refinements or adjustments made based on this feedback.

Testing Phase

- Describe how the testing with users was conducted.

- Share user feedback and initial reactions.

Validation Results

- Explain the results of the validation phase.

- Discuss any changes or improvements made following the validation.

Scalability and Sustainability

Explain the scalability of the solution to meet future demands. Discuss its long-term sustainability and potential for growth, ensuring it remains effective over time.

Scalability Strategy

- Outline the plan for scaling the solution.

- Identify potential challenges and how they will be addressed.

Long-Term Sustainability

- Discuss how the solution will remain relevant and effective in the long term.

- Address any potential obstacles to sustainability.

Impact Assessment

Outline the expected impact of the innovative solution on the problem and its stakeholders. Include measurable Key Performance Indicators (KPIs) to track progress.

KPIs

- List the specific KPIs that will be used to assess the solution's impact.

- Explain their significance and how they will be measured.

Challenges and Mitigations

Identify potential challenges or risks associated with the innovation and discuss strategies or plans to mitigate these challenges.

Challenge 1

- Describe the first identified challenge.

- Explain how it may impact the innovation.

Mitigation Plan

- Outline the strategy to mitigate the challenge.

- Discuss how the mitigation plan will address the challenge effectively.

.

```
import pandas as pd
import numpy as np
from functools import reduce
```

Load the data raw = pd.read_csv('/kaggle/input/argentina-car-prices/argentina_cars.csv')

Setup

ratio_usd_ars_official = 1/172.41408
ratio_usd_ars_blue = 1/319

Adding column for English variables

colours_dict = {'Plateado': 'Silver', 'Blanco': 'White', 'Gris oscuro': 'Dark Grey', 'Gris': 'Grey', 'Negro': 'Black', 'Naranja': 'Orange', 'Beige': 'Beige', 'Rojo': 'Red', 'Azul': 'Blue', 'Dorado': 'Gold', 'Marrón': 'Purple', 'Verde': 'Green', 'Celeste': 'Light Blue', 'Violeta': 'Violet'}
fuel_type_dict = {'Nafta': 'Petrol', 'Diésel': 'Diesel', 'Nafta/GNC': 'Petrol/CNG', 'Híbrido/Nafta': 'Petrol/Hybrid'}
gears_dict = {'Automática': 'Auto', 'Manual': 'Manual'}

```
raw_anglicised = raw.assign(colour_eng = raw.color.replace(colours_dict),
                            fuel_type_eng = raw.fuel_type.replace(fuel_type_dict),
                            gears_eng = raw.gear.replace(gears_dict),
                            price_usd_official = np.where(raw.currency == 'pesos',
raw.money * ratio_usd_ars_official, raw.money),
                            price_usd_blue = np.where(raw.currency == 'pesos',
raw.money * ratio_usd_ars_blue, raw.money))
```

add Codeadd Markdown

Chunk two was also fine

add Codeadd Markdown

```
import seaborn as sns
import matplotlib.pyplot as plt
```

Count the number of cars by brand

```
cars_by_brand =
raw_anglicised.groupby('brand').size().reset_index(name='Cars')
```

Plot the number of cars by brand

```
sns.catplot(x='brand', y='Cars', data=cars_by_brand, kind='bar')
plt.show()
```

Filter for the top 6 brands by number of listings

top_brands = ['Chevrolet', 'Ford', 'Peugeot', 'Renault', 'Toyota', 'Volkswagen']

```
top_brands_data = raw_anglicised[raw_anglicised.brand.isin(top_brands)]
```

Plot the price of cars by kilometers and brand for the top brands

```
sns.relplot(x='kilometres', y='price_usd_blue', col='brand', alpha=.6,
kind='scatter', data=top_brands_data, facet_kws={'sharey': False})
plt.subplots_adjust(top=0.9)
plt.suptitle('Argentine Car Price by Kms and Brand')
plt.show()
```

add Codeadd Markdown

The graphing code is quite different between ggplot 2 and pyhthon so I had to edit down the code till it could make a boxplot, then I could build it back up again with the other plot attributes.

add Codeadd Markdown

Filter for Ford cars with price over 100,000 USD

```
out_1 = raw_anglicised.loc[(raw_anglicised['brand'] == 'Ford') &
                    (raw_anglicised['price_usd_blue'] > 100000)]
```

Filter for cars with over 300,000 km

```
out_2 = raw_anglicised.loc[raw_anglicised['kilometres'] > 300000]
```

Filter for selected brands and calculate age

```
selected_brands = ['Chevrolet',
                   'Ford',
                   'Peugeot',
                   'Renault',
                   'Toyota',
                   'Volkswagen']

filtered =
raw_anglicised.loc[raw_anglicised['brand'].isin(selected_brands)].copy()
filtered['age'] = 2022 - filtered['year']
```

The graph didn't work first time out of gpt3 we had to have a bit of back and forth to get a basic graph.

```
fig, ax = plt.subplots()

sns.boxplot(x = 'brand', y = 'age', data = filtered)
sns.stripplot(x="brand", y="age", data=filtered, jitter=2, size=4, alpha = .5)
```

add titles and labels

```
plt.title('Car Age distribution by Brand', fontsize=14)
plt.xlabel('Brand', fontsize=12)
plt.ylabel('Age in years', fontsize=12)

plt.show()

print(out_1)
print(out_2)
```

Conclusion

Summarize the innovative solution, its key features, and the anticipated impact it will have on solving the problem. Reiterate the importance of innovation in achieving the project's goals.

References

Include a list of sources, studies, or materials used for research and development to provide credibility to the innovative solution.