

GenAI Project Unit 1

Name: Maya Nithyanand Bhagath

SRN: PES2UG23CS331

Section: F

Problem Statement:

Travel Destination Extractor: Parses the blog and lists all the Locations mentioned.

Abstract:

This project builds a Travel Destination Extractor using Named Entity Recognition. The system takes the text and parses through it to identify the locations mentioned in it. It helps in automated travel analysis and information extraction.

Documentation:

Travel blogs are a form of unstructured data. Named Entity Recognition (NER) categorizes each word in a sentence as a different entity, such as person, location, etc. Using this, we can automatically detect the locations mentioned in the blog.

The Travel Destination Extractor was build using the Hugging Face Transformers library. It accepts a travel blog as an input. A pre-trained NER model, called **bert-base-NER** is used. The model identifies all the names entities in the text, and filters only the location related entities. It also removes duplicate locations.

Screenshots:

```
[4]: blog = """
Last summer I traveled across Europe starting in Paris.
From there, I visited Rome, Florence, and Venice in Italy.
Later, I explored Barcelona in Spain and ended my journey in Amsterdam.
"""

places = extract_locations(blog)

print(" Locations mentioned:")
for place in places:
    print("-", place)
```

Locations mentioned:

- Rome
- Venice
- Florence
- Barcelona
- Spain
- Europe
- Paris
- Amsterdam
- Italy