Machine Learning Model Deployment for Real-time Predictions
PHASE 5

Objective:

- The objective of this project is to deploy a machine learning model as a web service for real-time predictions. The model is designed to make predictions on a selected use case and should be accessible via an API. This project aims to provide a scalable and user-friendly solution for utilizing machine learning models in real-time applications.
-
- **Design Thinking Process:**
- 1. **Identify Problem:** The initial phase involved identifying a specific use case that could benefit from machine learning predictions in real-time. This could be anything from fraud detection to image classification.
-
- 2. **Data Collection:** The next step was to select and gather the dataset relevant to the chosen use case. Data quality and quantity are crucial for model development.
-
- 3. **Model Selection:** Once the data was collected and cleaned, the appropriate machine learning model was chosen based on the specific problem and dataset characteristics.
-
- 4. **Model Development:** The model was then developed using appropriate libraries and tools, and the performance was evaluated through various metrics and techniques.
-
- 5. **Deployment Planning:** The deployment process was planned, considering factors like scalability, accessibility, and integration with other systems.
-
- 6. **Deployment and Integration:** The model was deployed as a web service, and integration steps were taken to make it accessible via an API.
-
- 7. **Testing and Optimization:** Extensive testing and optimization were carried out to ensure that the deployed model performed efficiently and accurately.
-
- 8. **Documentation and Submission:** Finally, the project was documented and prepared for submission.
-
- **Predictive Use Case:**
- The chosen use case for this project is sentiment analysis. We aim to develop a model that can predict the sentiment (positive, negative, or neutral) of text data in real-time. This could be used for analyzing customer reviews, social media comments, or any textual data source.
-
- **Dataset Selection:**
- For sentiment analysis, a dataset containing labeled text data with sentiment labels (positive, negative, neutral) was selected. The dataset was preprocessed to remove noise, clean the text, and ensure uniformity.
-

- **Model Training:**
- The model for sentiment analysis was trained using natural language processing (NLP) techniques. This involved tokenization, embedding layers, recurrent neural networks (RNNs), or transformer-based architectures, depending on the choice of model.
-
- **Deployment Process:**
- The deployment process involved the following steps:
- 1. Containerization: The trained model was containerized using Docker for easy deployment.
- 2. Hosting: The Docker container was hosted on a cloud server or platform like AWS, Google Cloud, or Heroku.
- 3. API Development: An API was developed to enable real-time predictions. Flask or FastAPI could be used for building the API.
- 4. Integration: The API was integrated with the deployed model to make predictions accessible.
-
- **Accessing and Utilizing the Deployed Model:**
- The deployed model can be accessed for real-time predictions through API requests. Users can make POST requests to the API endpoint with their input text, and the API will respond with the predicted sentiment label (positive, negative, or neutral).
-
- **Submission:**
- 1. GitHub Repository Link: [Provide the link to the GitHub repository containing the project's code and files.]
- 2. Deployment Instructions:
- - Clone the GitHub repository.
- - Install the required dependencies (e.g., Flask, Docker).
- - Build the Docker container using the provided Dockerfile.
- - Run the container on a cloud server or platform.
- - Access the API endpoint for real-time predictions.
-
- 3. Example API Requests:
- - Request: `POST /predict_sentiment`
- ```
- {
- "text": "I love this product! It's amazing."
- }
- ```
- Response:
- ```
- {
- "sentiment": "positive"
- }
- ```
-
- - Request: `POST /predict_sentiment`
- ```
- {
- "text": "This is terrible, I hate it."
- }
- ```

- Response:
- ```
- {
- "sentiment": "negative"
- }
- ```
-
- By following the provided instructions, users can deploy the machine learning model and make real-time predictions using the API
-

NAME	E-MAIL	NM REGISTER NO:
D. YUGESH	Yugeshfarah123@gmail.com	AU820321104059
M. SUNDHARLIGAM	Sundharlinga00@gmail.com	AU820321104050
J.LOKESH	lokeshchithralokesh@gmail.com	AU820321104024
P. Praveen	Priveenroman@gmail.com	AU820321104032