

Sub Task 3: Deployment of AI Model – API Development

Overview

In this subtask, we deployed a trained Titanic survival prediction model as an API using FastAPI. The API accepts user inputs (such as Age, Fare, Sex, Embarkation, etc.), processes the data, and returns the model's prediction (Survived: 0 or 1). This deployment allows for real-time inference and easy integration with frontend or third-party systems.

Technologies Used

- Python 3.13
- FastAPI
- Uvicorn
- Pickle (for model serialization)
- GitHub (for version control)
- Render (for deployment)

GitHub Repository

https://github.com/vivek-tolada/my_api_project

Live API Endpoint

<https://my-api-project-dwq8.onrender.com>

Swagger UI (API Docs)

<https://my-api-project-dwq8.onrender.com/docs>

How to Use the API

1. Visit the Swagger UI link above.
2. Click on the POST endpoint.
3. Enter values for Age, Fare, Sex_male, Embarked_Q, Embarked_S, Pclass, Parch, SibSp.
4. Click Execute.
5. The response will show whether the person is predicted to survive (1) or not (0).

Model Info

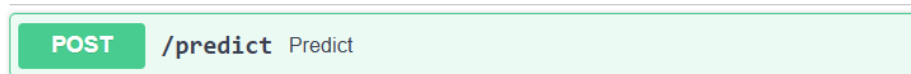
The deployed model is a Random Forest Classifier trained on the Titanic dataset. It was serialized using pickle (`model.pkl`) and loaded within the FastAPI backend. The app uses a POST method to accept input features in JSON format and returns predictions.

Output:

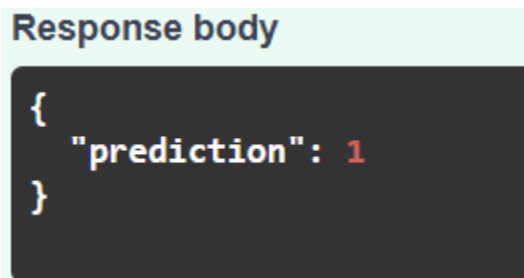
1. Swagger UI Opened



default



2. Sample Prediction Output



3. Render Deployment Logs (Success)

