



**Artificial Intelligence (Machine Learning & Deep Learning)  
[Course]**

**Week 10 - AI Deployment**

**[See examples / code in GitHub code repository]**

**It is not about Theory, it is 20% Theory and 80% Practical –  
Technical/Development/Programming [Mostly Python based]**

# AI Deployment

Deploying machine learning models enables your applications to make real-time predictions and decisions.



## Reference:

<https://medium.com/@emyasenc/deploying-machine-learning-models-with-flask-fastapi-or-streamlit-an-in-depth-guide-30c2e1f2ee44>

# AI Deployment Via Flask

**Flask is a web application framework written in Python. It was developed by Armin Ronacher, who led a team of international Python enthusiasts called Pocco. Flask is based on the Werkzeug WSGI toolkit and the Jinja2 template engine. Both are Pocco projects.**

## Reference:

<https://www.geeksforgeeks.org/python/flask-tutorial/>

<https://pythonbasics.org/what-is-flask-python/>

<https://flask.palletsprojects.com/en/stable/>

# Practical Development Case Study

## Reference:

<https://medium.com/@emyasenc/deploying-machine-learning-models-with-flask-fastapi-or-streamlit-an-in-depth-guide-30c2e1f2ee44>

## Model Saving:

<https://github.com/ShahzadSarwar10/FULLSTACK-WITH-AI-BOOTCAMP-B1-MonToFri-2.5Month-Explorer/blob/main/Week9/Case9-5-LinearRegressionViaSciKitLearn-SaveModel.py>

## Model Calling via Flask

<https://github.com/ShahzadSarwar10/FULLSTACK-WITH-AI-BOOTCAMP-B1-MonToFri-2.5Month-Explorer/blob/main/Week9/Case9-5-LinearRegressionViaSciKitLearn-ModelDeploymentWithFlask.py>

# AI Deployment Via FastAPI

FastAPI is a modern, high-performance web framework for building APIs with Python based on standard type hints. It has the following key features:

**Fast to run:** It offers very high performance, on par with NodeJS and Go, thanks to Starlette and pydantic.

**Fast to code:** It allows for significant increases in development speed.

**Reduced number of bugs:** It reduces the possibility for human-induced errors.

**Intuitive:** It offers great editor support, with completion everywhere and less time debugging.

**Straightforward:** It's designed to be uncomplicated to use and learn, so you can spend less time reading documentation.

**Short:** It minimizes code duplication.

**Robust:** It provides production-ready code with automatic interactive documentation.

**Standards-based:** It's based on the open standards for APIs, OpenAPI and JSON Schema.

## Practical Development Case Study

### Reference:

<https://medium.com/@emyasenc/deploying-machine-learning-models-with-flask-fastapi-or-streamlit-an-in-depth-guide-30c2e1f2ee44>

### Model Saving:

<https://github.com/ShahzadSarwar10/FULLSTACK-WITH-AI-BOOTCAMP-B1-MonToFri-2.5Month-Explorer/blob/main/Week9/Case9-5-LinearRegressionViaSciKitLearn-SaveModel.py>

### Model Calling via FASTAPI

<https://github.com/ShahzadSarwar10/FULLSTACK-WITH-AI-BOOTCAMP-B1-MonToFri-2.5Month-Explorer/blob/main/Week9/Case9-5-LinearRegressionViaSciKitLearn-ModelDeploymentWithFASTAPI.py>



Thank you - for listening and participating

- Questions / Queries
- Suggestions/Recommendation
- Ideas.....?

Shahzad Sarwar  
Cognitive Convergence

<https://cognitiveconvergence.com>

[shahzad@cognitiveconvergence.com](mailto:shahzad@cognitiveconvergence.com)

voice: +1 4242530744 (USA) +92-3004762901 (Pak)