Module 10 exercises – Deploying Models and Insights

Part 1 – Deploying a ML model

1. In JupyterLab, navigate the file browser to 10_Deploying Models & Insights.

2. Click the blue + icon.

3. Open a new Terminal.

4. Paste the following command: `uvicorn main:app --reload --no-use-colors`

5. In a new browser tab, enter the following URL: http://127.0.0.1:8000/predict/5?ds=BARCELONA&vc=Turista&fr=Promo&dur=160

6. Inspect the output. What format is it in?

Extensions

7. In the url, change BARCELONA to SEVILLA. How did that affect the prediction?

8. Change 160 to 250. How did that affect the prediction?

Hard extension

1. Using the requests library, display the data returned by the API in Python

```python
import requests

result = requests.get(r"http://127.0.0.1:8000/predict/5?ds=BARCELONA&vc=Turista&fr=Promo&dur=160")

eval(result.content.decode())['model_out']

74.33
```

Part 2 – Model deployment considerations

1. Consider the following ML powered applications. How might it be best to deploy each model? What would influence your decision?

- Mobile text autocompletion model

- Mortgage eligibility predictor

- Sensor based personal health prediction

Part 3 – Checks

1. What checks might you put in place for the model we deployed in part 1?