How much time did you spend on this exercise? On each part?

- In total about 8 hours.
- 5 hours for testing differents models with different data formats
  - In `notebook1` I tried several models and lightgbm gives the best results.
  - In `notebook2` I added the data of the match.
  - Finally in `notebook3` I remove the data of the match because a priori it is not known, besides I increase the data to get a final accuracy of 94%.
- 3 hours to create the API and the Docker ready to be deployed in the cloud or in a local machine.

If you had more time, what would you do?

- I will have a data-centric approach:
  - I would search for more data and continue to increase the data.
  - Train the model with vectors which have some Null information among the 32 data points and test again the accuracy when information is missed.

To train model and get model and pkls:

```
python app/train_model.py
```

Build docker

```
docker build -t atp .
```

Run in development mode:

```
sudo docker run -it --mount type=bind,source="$(pwd)/app",target=/app/app -
p 4004:4004 atp uvicorn app.main:app --host 0.0.0.0 --port 4004 --reload
```

Lauch docker

```
docker run --name atpR {image id}
```

Test

```
curl -X POST "http://localhost:4004/predict/" \
 -d '{"player1": "30.0, 17.0, 7.0, 8.0, 3.0, 6.0, 2.0, 0.0, 46.0, 20.64, R,
175.0, 101746, ITA, 78.0, 459.0", "player2": "37.0, 30.0, 7.0, 9.0, 1.0,
6.0, 5.0, 0.0, 53.0, 25.61, R, 180.0, 101142, ESP, 9.0, 1487.0"}' \
 -H "Content-Type: application/json"
```