Rappi Machine Learning Engineer Challenge

The Challenge

Based on the well known Kagle problem "Titanic: Machine Learning from Disaster", we would ask you to develop two process:

- Train a classifier
- Create an API

Train a classifier

Create a pipeline for training a binary classifier. The input for this training can be found in the link from above. The output should be a binary classifier model and exported somewhere so it can be used from the API.

It would be great if your solution covers the following points:

- Logs configuration
- Automatized pipeline execution
- Pipeline profiling (CPU and RAM usage)
- Virtual environment configuration
- Dockerized solution

Create an API

After exporting the classifier as a binary, it is required to use it for real time predictions. With the mentioned objective in mind, you would need to create an API that receives a list of passengers and predicts for each of them if they're likely to survive or not.

It would be great if your solution covers the following points:

- Download model and load it on API deployment
- Proper error handling (http status code, message, unhandled exceptions, etc)
- Logs configuration
- Unit testing of model's predict function
- Binary model issues tracking
- API profiling (CPU, RAM usage, maximum requests handled) and how it would scale
- API monitoring and alerts
- A/B Testing so two models can be deployed and compete between them
- Dockerized solution

Results delivery

1. **Implementation code folder/repository and instructions to run it.**
2. **You may use any language, tools, and cloud services you want.**
3. **You will have a 30 minutes presentation to show your work.**
4. **We expect to receive your work within 15 days since we send you the challenge.**

Documentation

We encourage you to write down every step down the road. Design decisions, references you used, difficulties and insights, and partial results.

If you have any doubts, feel free to contact us at ariel.wolfmann@rappi.com