Thirdlove - Machine Learning Engineer

Evaluation exam

Objective

We would like to screen your skills as part of the recruitment process, so we have designed a practical test that will help us evaluate it and give you an idea of the daily work we do at Thirdlove.

Section 1:

The goal is validated the general knowledge related to aspects like development practice and Software Engineer principles.

Section 2:

The goal is to measure how you design and implement a solution for a specific requirement.
The idea behind the test is for you to present all your knowledge and to do it as difficult as you want.
The process is very important to us. We don't want you to show just a couple of lines of code. We want you to explain your train of thought, how you did whatever you did, why you decided whatever you decided. You can do that either in this email thread or in the Readme file of the repository you'll be sending.

Presentation:

- The exercise is estimated to be completed in 9 days so the day after you should be sending us your work to cordoba-candidates@thirdlove.com. The deadline can be extended if you'd like to do additional things from those we ask you to do.
- If you have any questions or if you feel like you need some clarification, don’t hesitate to reach us.
- Create a private repository in Bitbucket to version the code and give access to cordoba-candidates@thirdlove.com.
- Create a running environment (AWS/Heroku/Google Cloud/DigitalOcean) with all the elements that are needed for running the microservice (database, computing, routing, etc.)
- The implementation needs to be production ready.
- Write a brief report explaining which technologies/libraries/patterns you used and why you chose them. If needed, include comments to the code to help us understand how it was implemented.

Section 1

1. What is the difference between a Python tuple, list, and dictionary? Are they sorted?
2. What are the benefits of writing multi-threaded programs?
3. What is a decorator?
4. What is PEP8?
5. What are the differences between iterator, generator, iterable, callable?
6. How would you test an API?. What testing tools have you used?
7. What are the differences between Flask and Django?
8. What common package have you used with Flask?
9. Where do you save sensitive data, i.e. secret keys?
10. Could you name some cloud technologies that you had the chance of use?
11. What are the differences between Docker and Docker Compose?
12. Did you have the chance to use CI/CD tool? If yes, name one example setup.

Section 2

1) Describe the design (overall big picture) in order to connect the API with the client side (in our case an app written in React JS (that communicates to Shopify)).

2) How would you design a pipeline to deploy Machine Learning algorithms in the cloud?. You can use any platform/tools

3) Design and implement an API service that can handle 1,000+ requests per minute that need to be evaluated by a machine learning classifier.

   The solution proposed should cover the following points:

   A. How would you think about authentication?

   B. How would you think about error handling from the classifier (ie, the classifier didn't get called for some reason or never returned an output)

   C. How would you think about error handling from the API service (ie, the host service such as Heroku doesn't have enough nodes and is failing, etc)?

   D. How would you log the data from the machine learning classifier?

   E. What languages would you support for ML models to be built in?

   F. How would you log and communicate out problems with the API service?

4) If you had all the time in the world and unlimited resources - what would you change?