Challenge

This is a test, in which you will be provided a training file (.csv format) with English first names (uncleaned with non-alphabet characters and you will need to clean it up) and the respective gender. Your aim is to write code to train and serve a machine learning model to assign a gender (male, female) to first names. The machine learning model must be able to determine the gender of names which have not been used to train the model. Please attempt this task to the best of your ability within 72 hours and send back. If you take longer than 72 hours your attempt may be void.

Estimated time to complete: *72* **Hours**

Submission Guidelines

Please submit your code by sending a zip of all the files (no including the trained model) or Github link back to the email address which you received the test from.

Requirements

- The code base should be in python.
- The machine learning package can be scikit-learn or Tensorflow or keras.
- Aim to achieve the highest accuracy model possible without overfitting.
- Please include a requirements.txt file.
- If the tester is unable to run your code, you will not be able to achieve satisfactory marks.

Project evaluation

We are evaluating project according to the features which specified table at the below.

Base requirements

1. Trainable Machine Learning model.
2. README included with code base on how to run Machine Learning model.
3. Python requirements file included.

Bonus points

1. Hyperparameter Tuning
2. Model can be trained and served using a flask/Django service.
3. DockerFile included with codebase to serve machine learning model.

Project evaluation

We are evaluating the task according to the features which are specified below:

1. Accuracy of the Machine Learning model.
2. Ease of model to train and serve.
3. Quality of comments.
4. Quality of code.

Note

1. Please be prepared to walk through the code and demonstrate your thought processes in an interview.
2. If you have any questions, please drop us email via the email u receive this test