Greetings PPO Software Developer candidate!

Our development stack and environment at PPO is fairly unique and challenging. We've prepared this interview challenge to both gauge how well you would mesh with our team and codebase, and also to give you a firsthand taste of what we do. We also do this so that you get a chance to explain to us something you understand better than us. (Your code!)

There isn't a hard deadline for this exercise; take as long as you need to complete it. However, we ask that you not spend more than a few hours. We value your time and are happy to leave things open to discussion in the next interview.

Project Description:

You will be given a Python 3.7 pickle file containing one Python object. The object will contain training data and labels for you to generate a chemometric classification model exactly like we do for our production system at PPO. In addition to the pickle file, you will be given a labels.txt file which will contain the mapping of our categorical Y variables to their 'real world' names.

Submission:

We would like you to send us a Python 3.7 module file `challenge.py` containing a class that will be called in the following way:

```
> from challenge import SpecPredict
>
> model = SpecPredict('path/to/your/model')
> model.predict(testing_data)
> # testing_data will be a [sample x feature] numpy array and should
> # return a [sample x class] numpy array.
```

Please write your code, comments, and documentation in a 'ready-for-review' style. That is, the point of this is not just that your code works, but that it is understandable and ready for addition to our production codebase. (We will not actually be adding your code to our codebase, of course.)

In addition to the `challenge.py` module, please send us:

- The script you used to train the model
- A model file
- An explanation of how to set up an environment to run your code(s)
- An explanation how you trained the model - what techniques etc.
- An explanation of how you would profile your solution for speed optimizations and why they might be important.
- An explanation of how you would evaluate the quality of classification of your solution.
- Any other utility code you wrote.

Your application should be easy to set up, and should run on Linux. It should not require any non-open-source software.

Please send us an email if you have any questions. We look forward to seeing your solution!