3003 North 1st Street #311
San Jose, California 95134
info@qeexo.com
Fax +1 408.228.3787

QEEXO INTERVIEW TASK: A BASIC FINGERSENSE CLASSIFIER

Your task is to write a program that takes as input files describing containing information about a touch point and determine whether the touch was from a pad or knuckle touch. Your program will classify approximately 10,000 touches. This task is a simplified version of the real classification task performed by the FingerSense engine.

DATASETS

The task_data zip file contains two additional zip files, train.zip and test.zip. Both train.zip and test.zip have the same directory structure.

DIRECTORY STRUCTURE

The structure of train.zip and test.zip is as follows:

```
root
  user_folder: [hand,table]-timestamp
    instance_folder: timestamp-[pad,knuckle]
      audio.wav
      touch.csv
```

An instance represents data from a single finger tap. An instance contains information about the touch (x,y, touch major and minor axes, pressure, and orientation). Each tap may be from a pad or knuckle. Each instance folder represents a single instance; its label (pad/knuckle) is specified in directory name. The timestamp for each instance is guaranteed to be unique.

Each user folder contains a collection of instance subfolders. A user folder represents data collected from a single user. The hand, table prefix on the user folder specifies whether the data for this user was collected when the user was holding a device in his/her hand, or resting the device on the table. This extra information is provided because it may be useful to you, but don't read too deeply into this (there is no guarantee that it will improve accuracy for your classifier). The set of users in train and test are disjoint—no user in train is also in test.

You will notice that the instance folders in the test dataset do not have labels. This is on purpose. Your task is to generate labels for these test instances.

Your train.zip file should contain 20,659 training instances (10,255 knuckle, 10404 pad).
Your test.zip file should contain 10,528 test instances (5,263 knuckle, 5,265 pad).

OUTPUT

Your program will generate a file `fingersense-test-labels.csv`, which provides the classification of each instance in the train.zip folder of the form.

```
timestamp,label
20140213_134923882,pad
20140213_134922770,pad
20140213_134923882, knuckle
…
```

LANGUAGE/LIBRARIES

You may use any language or libraries you wish. Our recommendation is to use Python and SciPy toolkit (or other machine learning libraries written in Python); however, using other language will not count against you if you are more comfortable in another environment.

DELIVERABLE

Please email us a zip file containing the following:

- `fingersense-test-labels.csv`
- README describing your algorithm along with any other notes.
- Your source code with instructions for how to build and run the code.

ADDITIONAL NOTES

Please let us know if the requirements are unclear or ambiguos.

TIME EXPECTATION

We expect this task to take between 4 and 6 hours for a developer experienced with machine learning. Of course, as with most tasks in machine learning, you can always improve your algorithm. Feel free to spend as much time as you wish on the task, as long as you submit within a week of your starting date.