Practical Machine Learning Exercise

Introduction

This is a short exercise to test your ability to construct and evaluate machine learning models. You should produce a machine learning model and an evaluation metric. The goal of this exercise is not to produce a state-of-the-art machine learning model. Which model you use, and how you evaluate, is up to you. The choice of model is not important (although we will assume that when you choose a model, you understand what it is and how it works). Your solution should be simple, but sensible: you should be able to explain why it tests something of impact to the problem.

Dataset description

In the "**train_set.csv** " you will find information about visits to a mall by customers. Each line represents one user. The first column is unique user identifier. The second column if indices of a day when customer have visited a mall. Enumeration is started from some fixed day; the day with index 1 is Monday (e. g. 7th is Sunday, 8th is again Monday). Indices are in the range from 1 to 1001 which equals to 143 full weeks. You are to predict the first day of the next (144th) week when user will visit a mall. For example, if a user will visit a mall on Wednesday of the next week, then your prediction should be equal to 3. So for each user, you need to predict a number:
- 0: no visit on the next week
- 1: Monday
- 2: Tuesday
- 3: Wednesday
- 4: Thursday
- 5: Friday
- 6: Saturday
- 7: Sunday

In the "**test_set.csv**" you will find correct answers for the next week.

Link to the data:
- https://drive.google.com/file/d/0B4bl7YMqDnVid0V3bW83Mngta00/view?usp=sharing

Tips and Clarifications

- We are not looking for a model that performs perfectly: we are looking to see that you can build a sensible model with a sensible evaluation.
- You should spend up to 3 hours on the problem, and submit a solution with no more than a several hundred lines of code.
- You can use any programming language you like to solve the problem: pick a language suited to the task, and one you are comfortable with.
- You can use of third-party libraries for model building and evaluation if needed.
- It would be great if you will use *jupyter notebook*, *rmarkdown* or similar tool.
- If you do not understand something or have questions, please contact us!

Submitting Your Solution

Please email the code of your solution, and a single-paragraph summary of your model and evaluation result. If you do any pre-processing to the data, please also include the script you use to do this (or a list of the commands run).