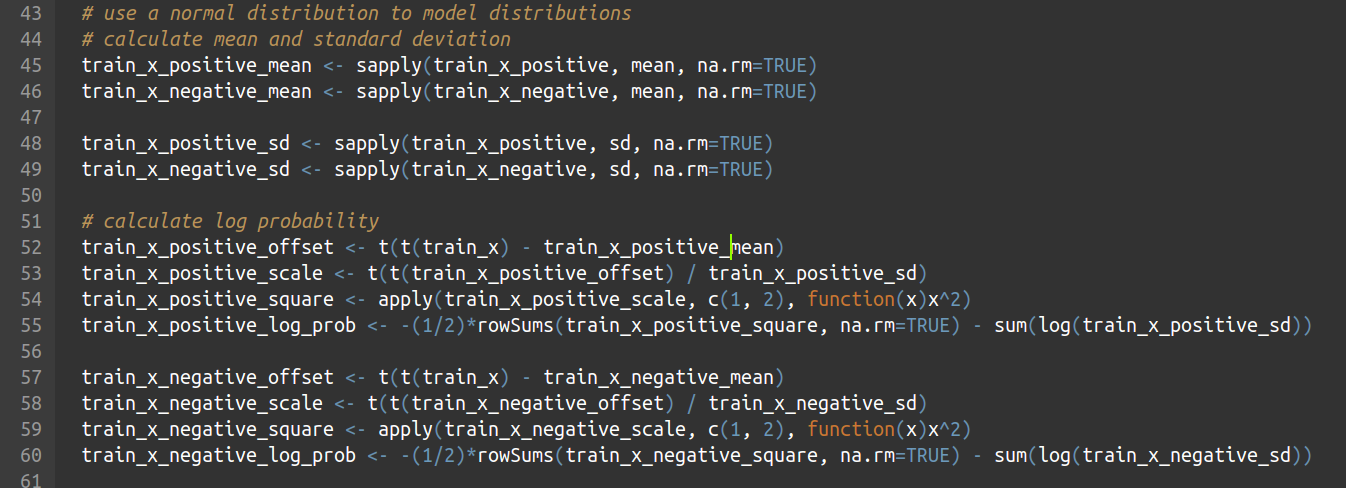
# Homework 1

**Problem 1：**

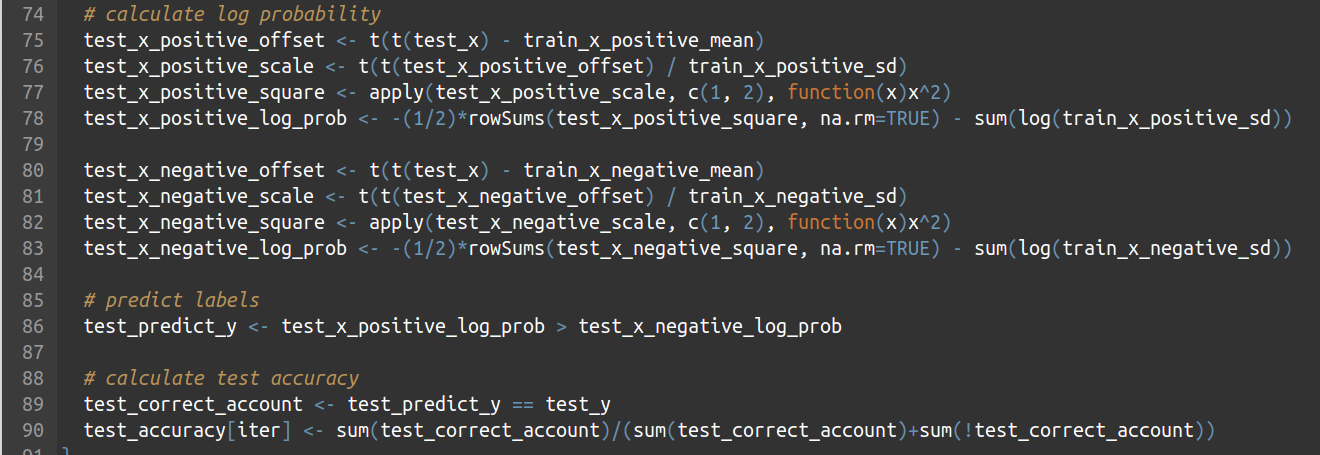
1. **Three accuracies for 1A, 1B, 1D**
   1. **Part 1A:** 74.31373%
   2. **Part 1B:** 75.88235%
   3. **Part 1D:** 77.45098%
2. **Screenshot of code:**
   1. **Test-Train Split:**



* 1. **Probability Calculations:**



* 1. **Evaluations:**

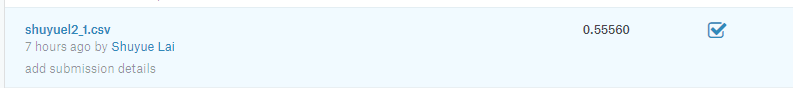


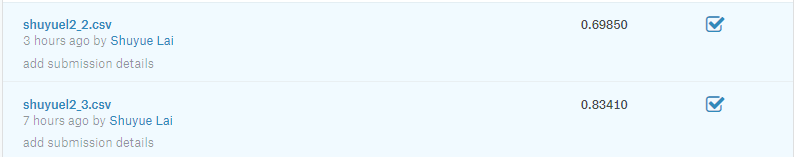
**Problem 2：**

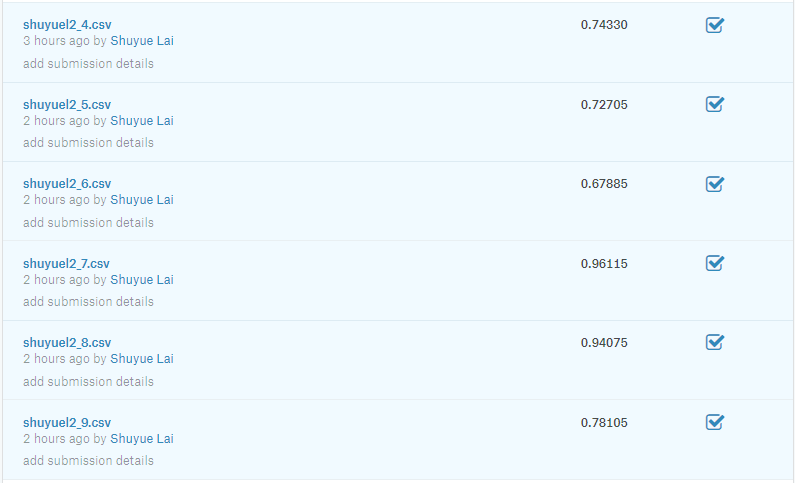
* 1. **Table of accuracies for all 12 cases:**

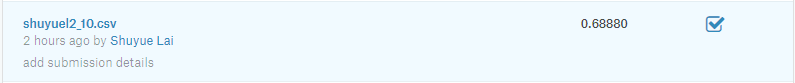
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Case | 1 | 2 | 3 | 4 |
| Accuracy | 55.560% | 69.850% | 83.410% | 74.330% |
| Case | 5 | 6 | 7 | 8 |
| Accuracy | 72.705% | 67.885% | 96.115% | 94.075% |
| Case | 9 | 10 | 11 | 12 |
| Accuracy | 78.105% | 68.880% | 97.190% | 95.285% |

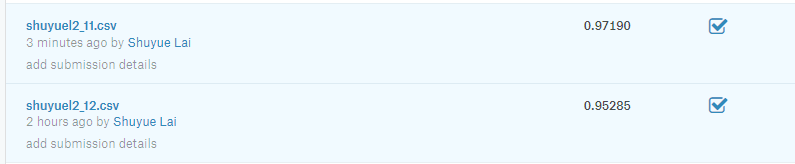
* 1. **Screenshot of Kaggle**











* 1. **A brief explanation of which model is better and why:**
     1. According to the result, Bernoulli Naïve Bayes is better than Gaussian Naïve Bayes, because Bernoulli distribution works better in discrete cases which is MNIST acts.
     2. According to the result, Random Forest Classifier works better with larger number of trees and larger maximum depth.

1. **40 mean images (4 \* 10 of part 2A)**
   1. **Case 1: Gaussian + Untouched**

**         **

* 1. **Case 2: Gaussian + Stretched**

**         **

* 1. **Case 3: Bernoulli + Untouched**

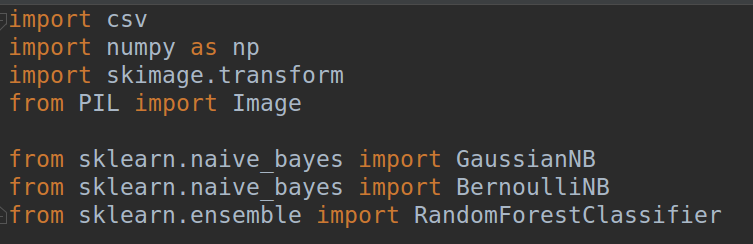
**         **

* 1. **Case 4: Bernoulli + Stretched**

**** **** **** **** **** **** **** **** ****



1. **Screenshot of code:**
   1. **Library:**



* 1. **Evaluations:**

