CS259 Quantitative Methods for Computer Science

Lab 3: Using Bayes in Practice

Assessment

This lab is assessed only formatively: you don't receive a numeric mark. Feedback is provided only if submitted by the deadline. Please note however, that our quizzes may include questions based on the lab instructions, so the best way to prepare for the quiz is to do this lab (even if only partially or submitted late).

You are not required to attend lab sessions, but it is a good opportunity to work on your assignments and get help from the lecturer and the demonstrators.

You can work on <u>this</u> assignment in the <u>groups</u> of your choice no larger than 3 students. But make sure <u>everyone</u> retains the copy of the code and uploads the report to myPlace, since it will be marked <u>individually</u>. You don't need to state the names of the students in the group.

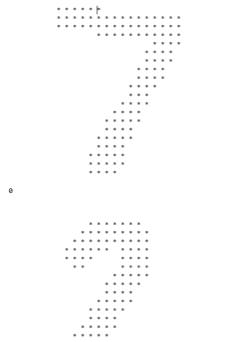
Submission

The report for this lab assignment with your screenshots needs to be uploaded to myPlace <u>before</u> 10am on Tuesday October 22nd as a single PDF file.

INSTRUCTIONS

Task 1. MNIST data

Finish the posted template so it can read the images from the posted dataset test-10.txt and print them using empty space character for zero values and star (*) for non-zero values. The first few lines of your output should look like this:



Note that in Task 2, we will apply our Naïve Bayes model from Lab 2 to this data. Since our model can only handle two possible outcomes ("flu or no flu"), we need to "binarize" the labels by converting them to 0s or 1s. So, we will treat the digits {5,6,7,8,9} as label "1" and the digits {0,1,2,3,4} as label "0". Note that the screenshot above shows binarized labels, so you can see 1 for digit "7" and 0 for digit "2". Our slides this week have additional hints and explanations for this task.

To Submit:

Take the screenshots covering your entire code and the first page of your output. Add them to your report.

Task 2. Applying Naïve Bayes Model for MNIST.

Finish the posted template that applies Naïve Bayes model from our Lab 2 to MNIST data (above). Note that the task has been "binarized": the model predicts if the digit is smaller than 5 or not. You will need to finish some missing parts of the code in the template and to copy your code for the Image class from Task 1 above. Comments within the template will guide you to the appropriate location for insertion. After completing the code, run your program. The model's performance on the testing data should yield results better than a simple random guess.

Hints: Don't forget to download the datasets from the posted archive. You may need to change the path. Our slides this week have additional hints and explanations for this task.

To Submit:

Take the screenshots covering your entire code and all your output and add them to your report. Submit your report to myPlace as stated above under "Submission."