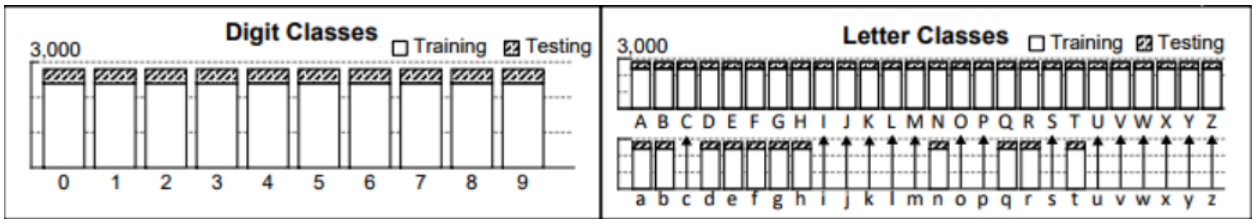
INTERNSHIP PROGRAMMING ROUND 1

You are given a dataset containing 28x28 grayscale images. Each image is either a handwritten letter or digit. The dataset can be downloaded from here:
https://drive.google.com/file/d/12OYCKGQp1VybvLM157ioLU4Bjt7PWpt-/

Format of dataset:
The dataset is well-balanced and contains 47 classes, as described in the image below.
(10 digits, 26 capital letters and 11 small letters)



The dataset is present as a CSV file. You'll find two CSV files: Train-set and Test-set.
You are supposed to train only using the train-set and use test-set only for calculating accuracy.

Number of samples:

Train set	: 112,800	(2400 images per class)
Test set	: 18,800	(400 images per class)

CSV format:
Each line in the csv file corresponds to 1 sample. Each line will contain 785 values.
The first value in all lines indicate the label ID, and the remaining 784 values corresponds to the individual pixel values of the 28*28 image (serialized in column-major order)

The ASCII value of each label ID can be found in the mapping.txt file.
For example, a label ID of 10 has an ASCII value of 65, which means that it corresponds to the character '**A**'.

You are supposed to use all the train samples (lines) to complete the following tasks:

Note: Each task must be submitted as independent runnable codes/Notebooks in a single GitHub repo. (i.e Don't squeeze in all tasks in a single file/Model).

Task 1: Letter/Digit Classifier (Easy)
Given an image, you must be able to classify whether the image is a letter or a digit.

Expected outcome: You are expected to use a ML-based model (like CNNs, etc.) to solve the problem with a reasonably high accuracy.

Task 2: Vowel/Consonant and Even/Odd Classifier (Moderate)
Given an image, you are supposed to design model(s) which does the following:
1. If the image is a letter, you are supposed to predict if it is a vowel or consonant.
2. If the image is a digit, you are supposed to predict if it is an even or odd number.

You are supposed to use only ML models that directly predicts the above, instead of doing manual predictions like using modulus operator on top of digit predictions.

Expected outcome: Given an image, your end-to-end setup must print whether it is a letter or digit, and based on that, it must automatically run the corresponding model to print if it is vowel/consonant or even/odd respectively.

Task 3: Character Classifier
Given an image, you are supposed to predict what digit or letter the image contains.
That is, you will be doing a classification task for 47 classes.

Expected outcome: Given an image, you have to print what character it is (just using a single model). Also, report the class-wise accuracy if possible.

Submission Rules:

1. The **deadline** for this round is **17-Nov-2019** (Sunday 11:59PM IST)
2. It's OK to submit how much ever you have completed within the deadline.
3. You are required to **push all your source code** to your **Github** repository.
4. After you have completed as many tasks as possible, you are required to fill the following Google Form with all the details required.
   Round-1 Submission Form
5. You are expected to **report the accuracies** you obtained for all the tasks and the techniques you used by explaining them all in your Github repository's README file.
6. You are also expected to push your **trained models** to the repo, and add instructions on how to load the model and run the code.
7. Do not use a single end-to-end model to solve all 3 tasks. Train different models specific to each task.

Note:

1. This is just a toy problem to evaluate the candidate's basic knowledge on programming and basic ML.
2. Feel free to use any programming language or framework that's comfortable to you.
3. You can also use Jupyter Notebooks / Google Colab Notebooks (but push it to Github)
4. Feel free to refer any code from the Internet. (**but with credits to the source**)
5. For any further doubts/clarifications, contact us at internship@onefourthlabs.com with subject "[Query] *yourFullName* - Internship Round 1"

ALL THE BEST, FOLKS! 😎