1. Data set generation.

For generating data set, I used a synthetic text generator (refer to GitHub link for more details).
The main purpose of it, is to get labelled dataset for training.
Refer the configs file in example_data/example.py file for all detailed configurations of generated text images.

2. Initially, I trained it on 50000 example, with 128 hidden units in the decoder. The training error obtained was very close to Bayesian error (0.05%).

3. For validation (train-dev set) I used another dataset, having the same distribution but that is not used for training. The above trained model performed poorly. Hence the bias was less but the variance was very high. The validation accuracy was close to just 10%.

4. Then I used a larger data set, 100000 examples and reduced the complexity of model by using 32 hidden units (Hyperparameter optimization). Now the training error was (0.2%). When performed on the above validation data, it gave close to (75%) accuracy.
I could only optimize the above parameters due to limited time constraints.

5. However, when this model (0.2%) error was tested on public test set, it gave poor performance (12%) accuracy.
This is may have happened because the generated dataset was not very realistic with respect to public test set. This is a conventional Data mismatch problem.

6. I exported the above model and this model is merged in Git (master branch).
Later steps such as developing REST API could not be executed due to shortage of time.