Model Card

Model Details

- Developed by Kamran Chitsaz and Alireza Razaghi for the IFT6390 first project.
- Out of distribution classifier
- Model has been developed on 2021-03-10
- Using Variational Autoencoder to identify if a sample is coming from a similar distribution as the training dataset or not

Intended Use

- Intended to be used for detecting if a sample is coming from a similar distribution as the training dataset or not
- Particularly intended for investigating if a picture is of a number or not
- Not suitable for classifying other types of images such as animals or objects
- Should not be used for essential topics such as navigation in smart cars

Factors

- The model is trained on the MNIST dataset and it is mostly accurate for images with similar resolution, colour and content.
- The model is evaluated on images from MNIST and Fashion MNIST datasets to test the accuracy of identifying in distribution and out of distribution samples

Metrics

- We calculate the log likelihood of a new sample and by comparing it to a threshold, classify it as Out of Distribution or In distribution. The choice of a threshold depends on the particular application. We have tune the threshold, as same as other model hyper parameters, by validation set.
- The decision threshold is -500

Training Data

- MNIST, training data split

Evaluation Data

- MNIST, test data split and Fashion MNIST
- We used a combination of two datasets to test our model

Ethical Considerations

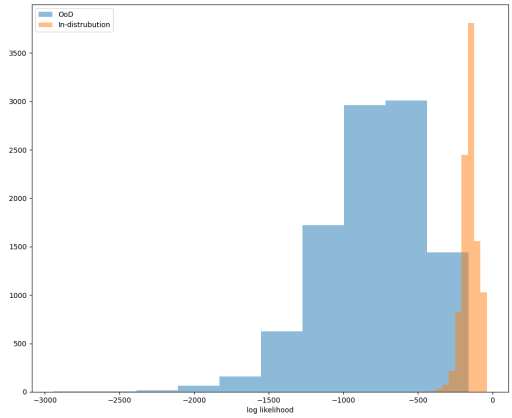
- This project shows that VAE can assign spuriously high likelihood to OoD samples which make them unreliable metrics for OoD detection. This model should not be used for any essential topics

Caveats and Recommendations

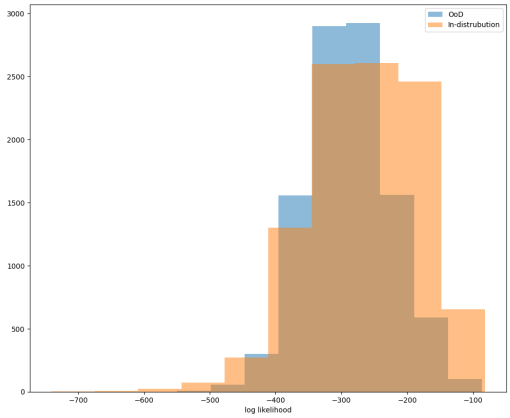
- While the idea of identifying out of distribution sample can be used in vast number of topics, a robust solution is still lacking. By introducing an efficient OoD score for VAEs, we can obtain better performance.
- Another application of this approach is preventing to input out of distribution data to models to maintain the performance of the model and ensure the quality of the collected data.

Quantitative Analyzes

- Histogram of the log likelihood of test samples for (a) VAE traind on Mnist and Fashion-Mnist as OoD, and (b) VAE traind on Fashion-Mnist and Mnist as OoD.



a



b