

Task:

The MNIST handwritten digit dataset is one of the most commonly used datasets for image classification.

You can access the dataset here: <http://yann.lecun.com/exdb/mnist/>.

It is also available in many ML packages such as TensorFlow, PyTorch, FastAI. Feel free to directly load the dataset via either source and train a model on it (high accuracy is NOT a criteria for evaluation).

Artifacts needed for evaluation:

1. Entire codebase with train and inference script to GitHub with links to access
2. Saved exported file (use any model repo framework or google drive)
3. Dockerfile to build docker image to launch the inference as a REST endpoint (in the same git)
4. Screenshot of the post request done locally using postman or curl

Criteria:

1. Low latency for inference
2. Low inference docker image size
3. Code structure

Bonus Points:

1. YAML file related to deployment on Kubernetes
2. Use of frameworks like Kserve/Triton/MLFlow/Prometheus
3. Architecture diagram or writeup explaining flow of requests