Laboratory 6: Machine Learning/Deep Learning

Instructor: Young H. Cho T.A.: Arthur Win

Due Oct 12 at 11:59 PM (Report), Oct 14 at 11:59 PM (Video)

Banknote Authentication Dataset

Data was extracted from images that were taken from genuine and forged banknote like specimens. Download the data from UCI ML repository (https://archive.ics.uci.edu/ml/machine-learning-databases/00267/data_banknote_authentication.txt). The dataset has 1372 instances with each instance having 4 features (Skewness, Kurtosis, entropy and variance) and 1 label (Whether forged/ genuine).

Operations to be performed

Task 1. => Exploratory data analysis – Make Scatter plots of features (of all possible permutations) and use different colors to represent class 0 (genuine) and 1 (forged)

Task 2. => Selection of a suitable machine learning algorithm – Explore several devised algorithms and use your analysis to choose one. Explain your choice through proper reasoning. There are several ML algorithms like Logistic regression, Support vector machines, Bayesian classifiers etc.

Task 3. => Select the first 200 rows of Class 0 and the first 200 rows of Class 1 as the test set and the rest of the data as the training set. Build a model based on your chosen algorithm, train and test the model. Report the training accuracy, test accuracy and test F-score.

Q 1) **AWS SageMaker**

Implement the above machine learning task on Amazon SageMaker. Report your results and provide detailed explanation and analysis wherever required.

Step 1. Enter the SageMaker console and create Notebook instance. Make sure to create a new IAM role and select ‘Any S3 bucket’ for easy access to S3 buckets.

Step 2. Once the Notebook instance is created, launch it and open a new Jupyter notebook.

Step 3. Create a new S3 bucket and store the data in it.

Step 4. Download the Banknote authentication data stored in the S3 bucket into the SageMaker instance.

Step 5. Perform the given tasks and report the required results.

Step 6. Make sure the resources are terminated to avoid extra charges.

P.S.: Make a wise decision while configuring the instance. This Machine Learning task doesn’t require heavy resources. Make appropriate decisions to save the AWS promotional education credit.

The following link will provide you with useful information while creating S3 bucket and to transfer data from S3 to SageMaker instance.

https://aws.amazon.com/getting-started/tutorials/build-train-deploy-machine-learning-model-sagemaker/?trk=gs_card

Q 2) **Deep Learning AMI**

Deep Learning Machine AMIs are loaded with deep learning frameworks like TensorFlow, Keras, Caffe etc. which can be used for building models, training, and work with large-scale datasets. Refer the following description and perform the tasks with Banknote authentication dataset. Also present the tutorial implementation with screenshots and required explanation.

https://aws.amazon.com/getting-started/tutorials/get-started-dlami/?trk=gs_card