TRAINING MODEL ON IBM

Date	05 November 2022
Team ID	PNT2022TMID05432
Project Name	Satistical Machine Learning Approaches To Liver Disease Prediction

Create a Training Model

Procedure

1. From the cluster management console, select **Workload** > **Spark** > **Deep Learning**.
2. From the **Models** tab, click **New**.
3. Select a model and click **Next**.
   - To use a previously added model, select one from the list.
   - To import a new model, add the location of the new model before selecting it.
     a. Click **Add Location**.
     b. Specify the framework.
     c. Specify the location of the model. Depending on the framework selected, make sure that the location specified has the correct files.
        - For a Caffe model, you must have at least two files: solver.prototxt and train_test.prototxt. For

inference models, a inference.prototxt file is required.

- For a TensorFlow model, you must have at least a main.py file. If you want to use the Distributed training with IBM Fabric option as a training engine, your model must also have a fabricmodel.py file. For inference models, a inference.py file is required.

  d. Click **Add**.
4. Specify the model name.
5. Specify the model description.
6. Select a training engine.

The following options are available:

- Single node training uses Caffe or TensorFlow.

- **Distributed training with Caffe** uses distributed CaffeOnSpark.

- **Distributed training with TensorFlow** uses native distributed TensorFlow.

- **Distributed training with IBM Fabric** combines Caffe or TensorFlow with a fabric layer for distribution.

- **Distributed training with IBM Fabric and auto-scaling** combines Caffe or TensorFlow with a fabric layer for distribution with auto-scaling enabled utility to both, the doctors and the patients. There are three preliminary steps Keywords Chronic diseases, Classification schemes, that serve as the elementary foundation of any medical treatment Training datasets, Machine learning.

Receiver Operating Characteristic
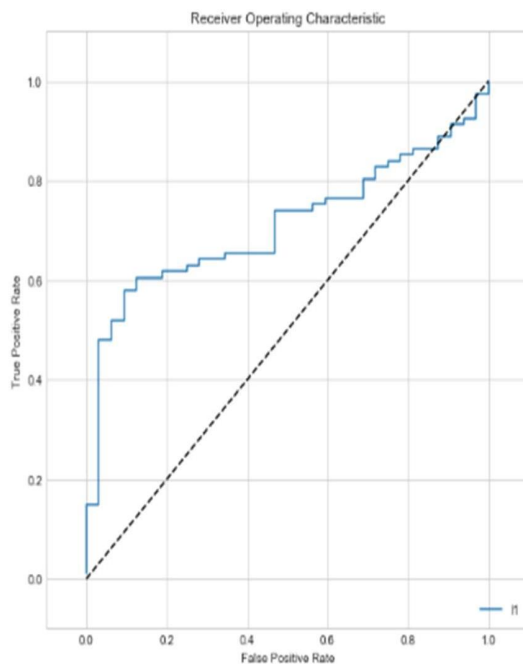
```
In [52]:   1 # Calculate AUC for Train set
           2 print(roc_auc_score(y_train, y_train_pred))
```

0.581751737359214

```
In [53]:   1 # Calculate AUC for Test set
           2 print(auc(fpr, tpr))
```

0.7102623456790124

```
┌─────────────────────────────────────────┐
│   44,  854 individuals in the study      │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│   129 items                              │
└─────────────────────────────────────────┘
                    │
                    │        ┌─────────────────────────────────┐
                    ├───────►│ 58 items with too many missing  │
                    │        │            values               │
                    │        └─────────────────────────────────┘
                    │
                    │        ┌─────────────────────────────────┐
                    ├───────►│ 39,884 individuals underwent    │
                    │        │    ultrasound examination       │
                    │        └─────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│   39,884 individuals with 71 items       │
└─────────────────────────────────────────┘
           ╱                       ╲
          ▼                         ▼
┌──────────────────────┐   ┌──────────────────────┐
│ 70% as training set  │   │  30% as test set     │
└──────────────────────┘   └──────────────────────┘
```