Machine Learning Foundations

# CAPSTONE-PROJECT

## TEACHING-ASSISTANT-PERFORMANCE-EVALUATION

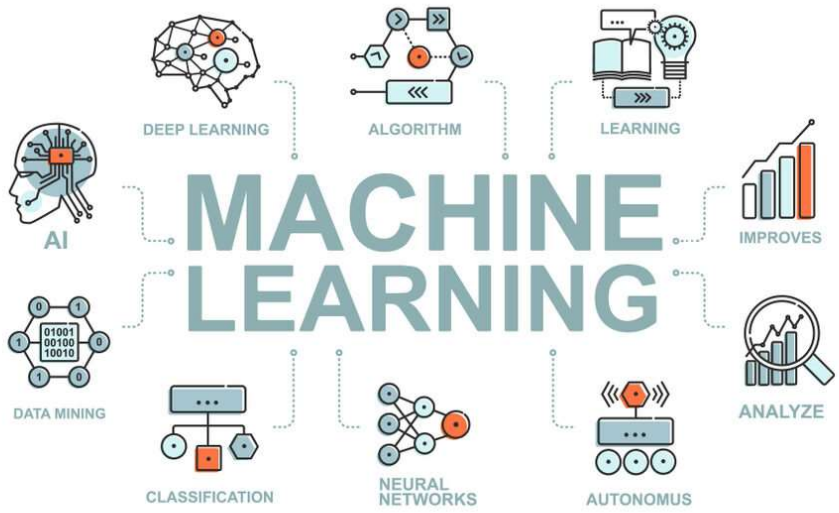
K M L Karunanayake

DSA_0308

2022-07-31

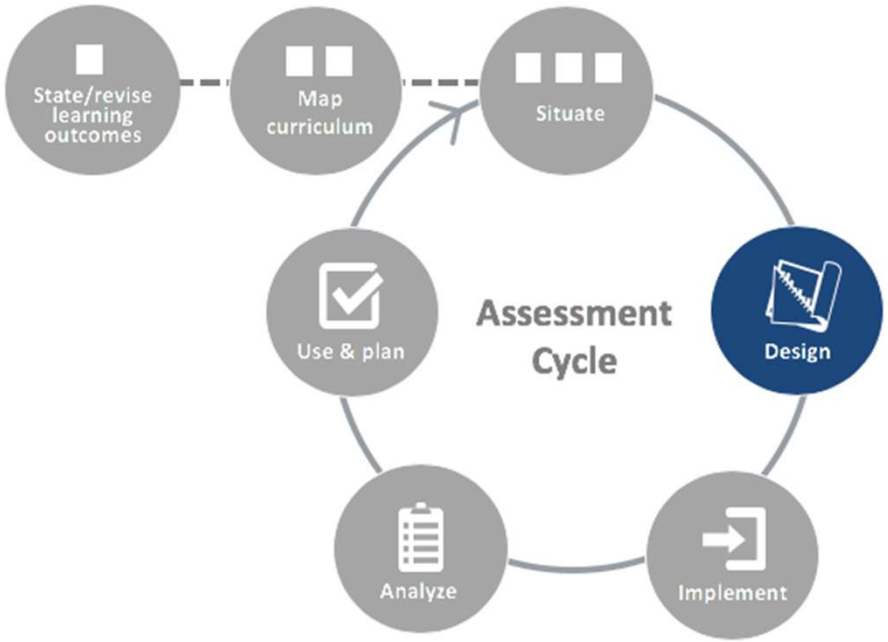
# CONTENT

- Use Case
- Data
- Solution
- Model
- Results
- Implementation



# USE CASE

## Teaching Assistant Evaluation



# DATA

## Teaching Assistant Evaluation Data Set

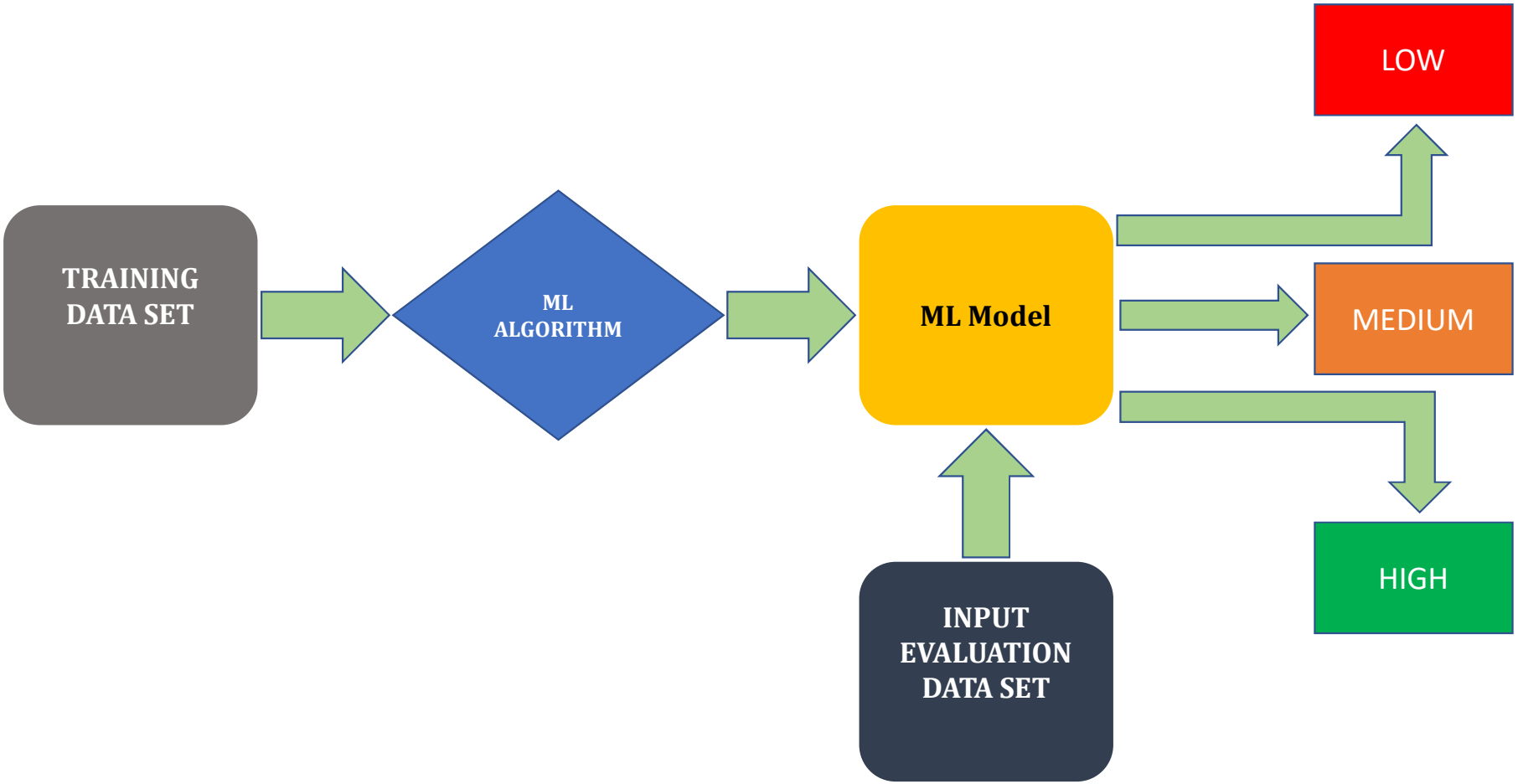
**Data Source – UCI Machine Learning Repository**

**Data Format– CSV file format**

**Number of Attributes – 5**

**Number of Instances – 151**

# SOLUTION


TRAINING DATA SET
ML ALGORITHM
ML Model
LOW
MEDIUM
HIGH
INPUT EVALUATION DATA SET

# TOOLS


scikits
learn


NumPy


seaborn

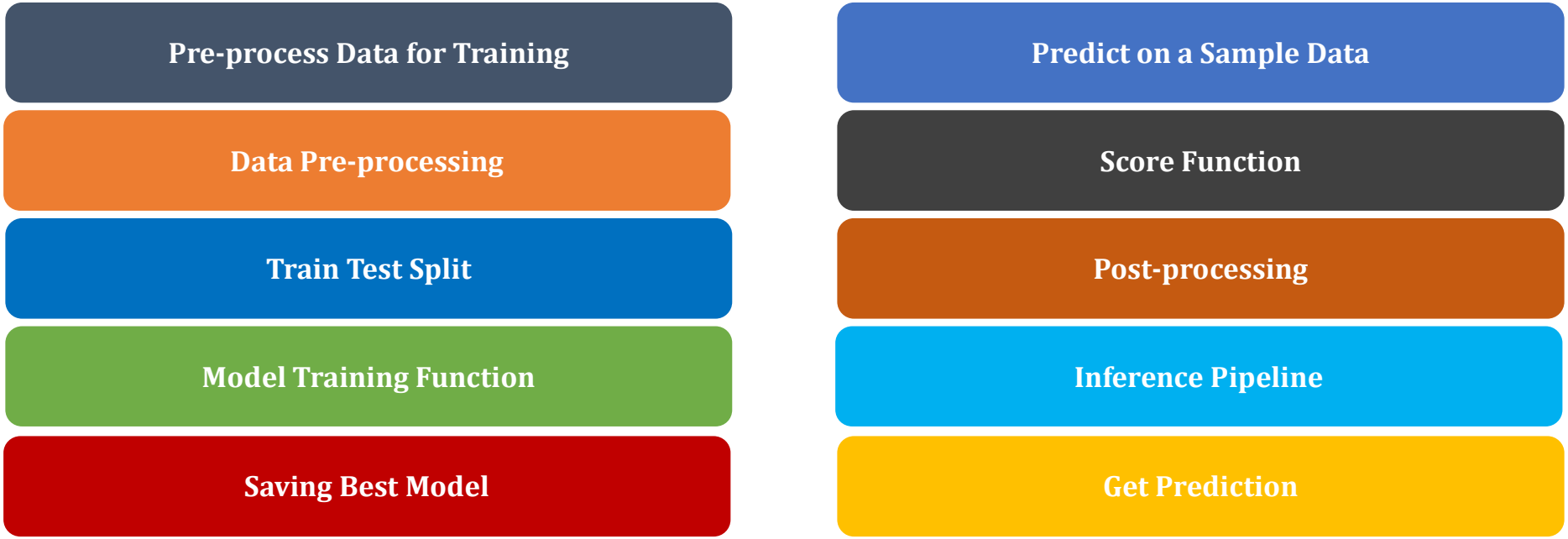

pandas


Joblib


jupyter


matplotlib

# MODEL BUILDING



# SELECTING BEST MODEL

## Models parameters

```
[ ] models = []
    models.append(model_train(LogisticRegression(n_jobs=3, verbose=1), 'lgr1', X_train, y_train, X_test, y_test))
    models.append(model_train(RandomForestClassifier(n_estimators=100, max_depth=None, n_jobs=3, verbose=1), 'rf1', X_train, y_train, X_test, y_test))
    models.append(model_train(RandomForestClassifier(n_estimators=500, max_depth=None, n_jobs=3, verbose=1), 'rf2', X_train, y_train, X_test, y_test))
    models.append(model_train(RandomForestClassifier(n_estimators=500, max_depth=10, n_jobs=3, verbose=1), 'rf3', X_train, y_train, X_test, y_test))
    models.append(model_train(RandomForestClassifier(n_estimators=500, max_depth=20, n_jobs=3, verbose=1), 'rf4', X_train, y_train, X_test, y_test))
    models = pd.DataFrame(models)
    models
```

## Models

| | model_name | model | accuracy | precision | f1_score |
|---|---|---|---|---|---|
| 0 | lgr1 | LogisticRegression(n_jobs=3, verbose=1) | 0.543478 | 0.625000 | 0.529956 |
| 1 | rf1 | (DecisionTreeClassifier(max_features='auto', r... | 0.717391 | 0.746283 | 0.713043 |
| 2 | rf2 | (DecisionTreeClassifier(max_features='auto', r... | 0.717391 | 0.746283 | 0.713043 |
| 3 | rf3 | (DecisionTreeClassifier(max_depth=10, max_feat... | 0.717391 | 0.746283 | 0.713043 |
| 4 | rf4 | (DecisionTreeClassifier(max_depth=20, max_feat... | 0.717391 | 0.746283 | 0.713043 |

## Best model

```
print(gs_model.best_params_)
```

```
{'max_depth': None, 'n_estimators': 500}
```

# RESULTS

Sample Data

| | English Speaking | Course instructor | Course | Semester | Class size |
|---|---|---|---|---|---|
| 91 | 2 | 9 | 5 | 2 | 24 |
| 69 | 2 | 11 | 1 | 2 | 51 |
| 10 | 2 | 10 | 22 | 2 | 9 |
| 137 | 2 | 22 | 1 | 2 | 42 |
| 35 | 2 | 16 | 8 | 2 | 36 |
| 5 | 2 | 23 | 3 | 1 | 20 |
| 92 | 2 | 18 | 25 | 2 | 25 |
| 135 | 2 | 12 | 8 | 2 | 24 |
| 58 | 2 | 7 | 25 | 2 | 42 |
| 100 | 1 | 6 | 17 | 2 | 35 |

Predictions

```
array([[0.062     , 0.146     , 0.792     ],
       [0.862     , 0.02      , 0.118     ],
       [0.014     , 0.138     , 0.848     ],
       [0.64433333, 0.32      , 0.03566667],
       [0.882     , 0.112     , 0.006     ],
       [0.004     , 0.23594848, 0.76005152],
       [0.094     , 0.156     , 0.75      ],
       [0.186     , 0.72      , 0.094     ],
       [0.032     , 0.902     , 0.066     ],
       [0.044     , 0.788     , 0.168     ]])
```

| | English Speaking | Course instructor | Course | Semester | Class size | id |
|---|---|---|---|---|---|---|
| 0 | 2 | 18 | 21 | 2 | 29 | 13 |

```
0.084
```

# MODEL SERVER API AND TEST

## Model Server API

```
(base) PS C:\Users\lahir\Desktop\MLFoundations> C:\Users\lahir\anaconda3\python C:\Users\lahir\Desktop\MLFoundations\model_server.py
C:\Users\lahir\anaconda3\lib\site-packages\sklearn\base.py:310: UserWarning: Trying to unpickle estimator DecisionTreeClassifier from version 1.0.2 w
hen using version 0.24.2. This might lead to breaking code or invalid results. Use at your own risk.
  warnings.warn(
C:\Users\lahir\anaconda3\lib\site-packages\sklearn\base.py:310: UserWarning: Trying to unpickle estimator RandomForestClassifier from version 1.0.2 w
hen using version 0.24.2. This might lead to breaking code or invalid results. Use at your own risk.
  warnings.warn(
 * Serving Flask app "model_server" (lazy loading)
 * Environment: production
   WARNING: This is a development server. Do not use it in a production deployment.
   Use a production WSGI server instead.
 * Debug mode: on
 * Restarting with watchdog (windowsapi)
C:\Users\lahir\anaconda3\lib\site-packages\sklearn\base.py:310: UserWarning: Trying to unpickle estimator DecisionTreeClassifier from version 1.0.2 w
hen using version 0.24.2. This might lead to breaking code or invalid results. Use at your own risk.
  warnings.warn(
C:\Users\lahir\anaconda3\lib\site-packages\sklearn\base.py:310: UserWarning: Trying to unpickle estimator RandomForestClassifier from version 1.0.2 w
hen using version 0.24.2. This might lead to breaking code or invalid results. Use at your own risk.
  warnings.warn(
 * Debugger is active!
 * Debugger PIN: 699-707-837
 * Running on http://127.0.0.1:5001/ (Press CTRL+C to quit)
```

## Test 1 and Result

```
In [11]: import requests

         url = 'http://127.0.0.1:5001/classifier'

         payload = {'English Speaking': 1,
                 'Course instructor':10,
                 'Course': 21,
                 'Semester':  2,
                    'Class size': 29,}

         headers = {}

         files=[

         ]

         response = requests.request("POST", url, headers=headers, data=payload, files=files)

         print(response.text)

         {
           "prediction": 0.442
         }
```

## Test 2 and Result

```
In [16]: import requests

         url = 'http://127.0.0.1:5001/classifier'

         payload = {'English Speaking': 0,
                 'Course instructor':3,
                 'Course': 3,
                 'Semester':  2,
                    'Class size': 29,}

         headers = {}

         files=[

         ]

         response = requests.request("POST", url, headers=headers, data=payload, files=files)

         print(response.text)

         {
           "prediction": 0.078
         }
```