

In [31]:

```
%time
import boto3
import os
import sagemaker
from sagemaker.xgboost.estimator import XGBoost
from sagemaker.session import Session
from sagemaker.inputs import TrainingInput
import pandas as pd
import sklearn as sk
import numpy as np
import sagemaker as sg
import matplotlib as mt
from sklearn.model_selection import train_test_split
import re
import urllib.request
from sklearn.datasets import load_svmlight_file
```

CPU times: user 38 µs, sys: 0 ns, total: 38 µs
Wall time: 42 µs

In [26]:

```
file="data.csv"
urllib.request.urlretrieve('https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians-diabetes.csv',file)
columns = ["number_times_pregnant", "plasma_glucose", "diastolic_blood_pressure", "triceps_skin_fold", "serum_insulin", "body_mass_index", "pedigree_function", "age", "will_get_diabetes"]
data=pd.read_csv(file,header=None,names=columns)
data['will_get_diabetes']=data['will_get_diabetes'].astype('category').cat.codes
data=data[ ['will_get_diabetes']] + [col for col in data.columns if col != 'will_get_diabetes']]
```

In [28]:

```
train,test= np.split(data.sample(frac=1),[int(0.8 * len(data))])
```

In [29]:

```
train.to_csv('train.csv',header=False, index=False)
test.to_csv('test.csv',header=False, index=False)
```

In [32]:

```
## Uploads training and validation csv for XGBoost to train on
boto3.Session().resource('s3').Bucket(bucket).Object(os.path.join(prefix, 'train/train.csv')).upload_file('train.csv')
boto3.Session().resource('s3').Bucket(bucket).Object(os.path.join(prefix, 'test/test.csv')).upload_file('test.csv')
```

In [33]:

```
from sagemaker.inputs import TrainingInput
s3_train_data=TrainingInput(s3_data='s3://{}:{}//train'.format(bucket,prefix),content_type=
'csv')
s3_test_data=TrainingInput(s3_data='s3://{}:{}//test/'.format(bucket,prefix),content_type=
'csv')
```

In [35]:

```
xgboost_container = sagemaker.image_uris.retrieve("xgboost", boto3.Session().region_name,
"latest")
print(xgboost_container)
```

811284229777.dkr.ecr.us-east-1.amazonaws.com/xgboost:latest

In [36]:

```
output_location='s3://{}:{}//output'.format(bucket,prefix)
```

In [40]:

```
sess=sagemaker.Session()
role=sagemaker.get_execution_role()
# construct a SageMaker estimator that calls the xgboost-container
xgboost= sagemaker.estimator.Estimator(xgboost_container,
                                         role,
                                         instance_count=1,
                                         instance_type='ml.m5.large',
                                         output_path=output_location,sagemaker_session=se
ss)
xgboost.set_hyperparameters(gamma='4',
                            min_child_weight ='6 ',
                            subsample = '0.7',
                            objective='binary:logistic',
                            num_round = '50')
```

In [41]:

```
xgboost.fit({'train':s3_train_data})
```

```
2021-10-13 16:23:12 Starting - Starting the training job...
2021-10-13 16:23:41 Starting - Launching requested ML instancesProfilerReport
-1634142192: InProgress
.....
2021-10-13 16:24:42 Starting - Preparing the instances for training.....
2021-10-13 16:26:09 Downloading - Downloading input data...
2021-10-13 16:26:42 Training - Downloading the training image..Arguments: tra
in
[2021-10-13:16:26:55:INFO] Running standalone xgboost training.
[2021-10-13:16:26:55:INFO] Path /opt/ml/input/data/validation does not exist!
[2021-10-13:16:26:55:INFO] File size need to be processed in the node: 0.02m
b. Available memory size in the node: 163.75mb
[2021-10-13:16:26:55:INFO] Determined delimiter of CSV input is ','
[16:26:55] S3DistributionType set as FullyReplicated
[16:26:55] 614x8 matrix with 4912 entries loaded from /opt/ml/input/data/trai
n?format=csv&label_column=0&delimiter=,
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 20 extra
nodes, 0 pruned nodes, max_depth=5
[0]#011train-error:0.228013
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 18 extra
nodes, 0 pruned nodes, max_depth=6
[1]#011train-error:0.193811
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 18 extra
nodes, 0 pruned nodes, max_depth=6
[2]#011train-error:0.203583
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 16 extra
nodes, 0 pruned nodes, max_depth=5
[3]#011train-error:0.198697
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 16 extra
nodes, 0 pruned nodes, max_depth=4
[4]#011train-error:0.203583
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 18 extra
nodes, 0 pruned nodes, max_depth=5
[5]#011train-error:0.190554
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=6
[6]#011train-error:0.192182
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=6
[7]#011train-error:0.184039
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 16 extra
nodes, 0 pruned nodes, max_depth=6
[8]#011train-error:0.180782
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=4
[9]#011train-error:0.180782
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 18 extra
nodes, 0 pruned nodes, max_depth=6
[10]#011train-error:0.174267
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=6
[11]#011train-error:0.167752
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=5
[12]#011train-error:0.164495
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=6
```

```
[13]#011train-error:0.159609
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=6
[14]#011train-error:0.149837
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=6
[15]#011train-error:0.151466
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=4
[16]#011train-error:0.143322
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=5
[17]#011train-error:0.149837
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=5
[18]#011train-error:0.140065
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=3
[19]#011train-error:0.141694
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=6
[20]#011train-error:0.135179
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=5
[21]#011train-error:0.136808
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=5
[22]#011train-error:0.144951
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=3
[23]#011train-error:0.131922
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[24]#011train-error:0.143322
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[25]#011train-error:0.131922
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=5
[26]#011train-error:0.130293
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=3
[27]#011train-error:0.123778
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[28]#011train-error:0.12215
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=6
[29]#011train-error:0.12215
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[30]#011train-error:0.12215
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=4
[31]#011train-error:0.123778
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=4
```

```
[32]#011train-error:0.120521
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[33]#011train-error:0.115635
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=5
[34]#011train-error:0.114007
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=6
[35]#011train-error:0.105863
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=5
[36]#011train-error:0.115635
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=5
[37]#011train-error:0.105863
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[38]#011train-error:0.105863
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=6
[39]#011train-error:0.107492
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[40]#011train-error:0.104235
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=6
[41]#011train-error:0.114007
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=3
[42]#011train-error:0.112378
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[43]#011train-error:0.110749
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=3
[44]#011train-error:0.107492
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[45]#011train-error:0.110749
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[46]#011train-error:0.104235
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=4
[47]#011train-error:0.104235
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 12 extra
nodes, 0 pruned nodes, max_depth=5
[48]#011train-error:0.09772
[16:26:55] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 10 extra
nodes, 0 pruned nodes, max_depth=4
[49]#011train-error:0.09772
```

2021-10-13 16:27:09 Uploading - Uploading generated training model

2021-10-13 16:27:09 Completed - Training job completed

Training seconds: 60

Billable seconds: 60

In [91]:

```
%time
import io
import os
import boto3
import sagemaker
import time
role=sagemaker.get_execution_role()
region= boto3.Session().region_name

bucket= sagemaker.Session().default_bucket()
prefix="sagemaker/xgboost-hw"
```

CPU times: user 80.3 ms, sys: 7.35 ms, total: 87.7 ms
Wall time: 305 ms

In [92]:

```
!pip install -Uq pyarrow

/opt/conda/lib/python3.7/site-packages/secretstorage/dhcrypto.py:16: Cryptogr
aphyDeprecationWarning: int_from_bytes is deprecated, use int.from_bytes inst
ead
    from cryptography.utils import int_from_bytes
/opt/conda/lib/python3.7/site-packages/secretstorage/util.py:25: Cryptography
DeprecationWarning: int_from_bytes is deprecated, use int.from_bytes instead
    from cryptography.utils import int_from_bytes
WARNING: Running pip as the 'root' user can result in broken permissions and
conflicting behaviour with the system package manager. It is recommended to
use a virtual environment instead: https://pip.pypa.io/warnings/venv
WARNING: You are using pip version 21.2.4; however, version 21.3 is availabl
e.
You should consider upgrading via the '/opt/conda/bin/python -m pip install -
-upgrade pip' command.
```

In [93]:

```
pip install -U sagemaker
```

```
/opt/conda/lib/python3.7/site-packages/secretstorage/dhcrypto.py:16: Cryptogr
aphyDeprecationWarning: int_from_bytes is deprecated, use int.from_bytes inst
ead
    from cryptography.utils import int_from_bytes
/opt/conda/lib/python3.7/site-packages/secretstorage/util.py:25: Cryptography
DeprecationWarning: int_from_bytes is deprecated, use int.from_bytes instead
    from cryptography.utils import int_from_bytes
Requirement already satisfied: sagemaker in /opt/conda/lib/python3.7/site-pac
kages (2.59.5)
Collecting sagemaker
  Downloading sagemaker-2.63.0.tar.gz (447 kB)
    ██████████ | 447 kB 27.4 MB/s eta 0:00:01
Requirement already satisfied: attrs in /opt/conda/lib/python3.7/site-package
s (from sagemaker) (19.3.0)
Requirement already satisfied: boto3>=1.16.32 in /opt/conda/lib/python3.7/sit
e-packages (from sagemaker) (1.18.51)
Requirement already satisfied: google-pasta in /opt/conda/lib/python3.7/site-
packages (from sagemaker) (0.2.0)
Requirement already satisfied: numpy>=1.9.0 in /opt/conda/lib/python3.7/site-
packages (from sagemaker) (1.20.3)
Requirement already satisfied: protobuf>=3.1 in /opt/conda/lib/python3.7/site
-packages (from sagemaker) (3.18.0)
Requirement already satisfied: protobuf3-to-dict>=0.1.5 in /opt/conda/lib/pyt
hon3.7/site-packages (from sagemaker) (0.1.5)
Requirement already satisfied: smdebug_rulesconfig==1.0.1 in /opt/conda/lib/p
ython3.7/site-packages (from sagemaker) (1.0.1)
Requirement already satisfied: importlib-metadata>=1.4.0 in /opt/conda/lib/py
thon3.7/site-packages (from sagemaker) (1.5.0)
Requirement already satisfied: packaging>=20.0 in /opt/conda/lib/python3.7/si
te-packages (from sagemaker) (20.1)
Requirement already satisfied: pandas in /opt/conda/lib/python3.7/site-packag
es (from sagemaker) (1.0.1)
Requirement already satisfied: pathos in /opt/conda/lib/python3.7/site-packag
es (from sagemaker) (0.2.8)
Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in /opt/conda/lib/pytho
n3.7/site-packages (from boto3>=1.16.32->sagemaker) (0.10.0)
Requirement already satisfied: s3transfer<0.6.0,>=0.5.0 in /opt/conda/lib/pyt
hon3.7/site-packages (from boto3>=1.16.32->sagemaker) (0.5.0)
Requirement already satisfied: botocore<1.22.0,>=1.21.51 in /opt/conda/lib/py
thon3.7/site-packages (from boto3>=1.16.32->sagemaker) (1.21.51)
Requirement already satisfied: urllib3<1.27,>=1.25.4 in /opt/conda/lib/python
3.7/site-packages (from botocore<1.22.0,>=1.21.51->boto3>=1.16.32->sage
maker) (1.26.7)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in /opt/conda/lib/
python3.7/site-packages (from botocore<1.22.0,>=1.21.51->boto3>=1.16.32->sage
maker) (2.8.1)
Requirement already satisfied: zipp>=0.5 in /opt/conda/lib/python3.7/site-pac
kages (from importlib-metadata>=1.4.0->sagemaker) (2.2.0)
Requirement already satisfied: six in /opt/conda/lib/python3.7/site-packages
 (from packaging>=20.0->sagemaker) (1.14.0)
Requirement already satisfied: pyparsing>=2.0.2 in /opt/conda/lib/python3.7/si
te-packages (from packaging>=20.0->sagemaker) (2.4.6)
Requirement already satisfied: pytz>=2017.2 in /opt/conda/lib/python3.7/site-
packages (from pandas->sagemaker) (2019.3)
Requirement already satisfied: dill>=0.3.4 in /opt/conda/lib/python3.7/site-p
ackages (from pathos->sagemaker) (0.3.4)
Requirement already satisfied: pox>=0.3.0 in /opt/conda/lib/python3.7/site-pa
```

```
ckages (from pathos->sagemaker) (0.3.0)
Requirement already satisfied: multiprocessing>=0.70.12 in /opt/conda/lib/python
3.7/site-packages (from pathos->sagemaker) (0.70.12.2)
Requirement already satisfied: ppft>=1.6.6.4 in /opt/conda/lib/python3.7/site
-packages (from pathos->sagemaker) (1.6.6.4)
Building wheels for collected packages: sagemaker
  Building wheel for sagemaker (setup.py) ... done
    Created wheel for sagemaker: filename=sagemaker-2.63.0-py2.py3-none-any.whl
    size=622844 sha256=ee4c677c3692585a47fb8435f1a227963284d7afc79dd47b2d8b90abad
    b5cf16
      Stored in directory: /root/.cache/pip/wheels/e9/12/f6/25d1d50b7fe9489d0a998
    a7caa1418f6cdaf71ce80f2979ca5
Successfully built sagemaker
Installing collected packages: sagemaker
Attempting uninstall: sagemaker
  Found existing installation: sagemaker 2.59.5
  Uninstalling sagemaker-2.59.5:
    Successfully uninstalled sagemaker-2.59.5
Successfully installed sagemaker-2.63.0
WARNING: Running pip as the 'root' user can result in broken permissions and
conflicting behaviour with the system package manager. It is recommended to
use a virtual environment instead: https://pip.pypa.io/warnings/venv
WARNING: You are using pip version 21.2.4; however, version 21.3 is available.
You should consider upgrading via the '/opt/conda/bin/python -m pip install -
-upgrade pip' command.
Note: you may need to restart the kernel to use updated packages.
```

In [94]:

```
output_location= 's3://{} / {} / output'.format(bucket,prefix)
```

In [95]:

```
sess=sagemaker.Session()
role=sagemaker.get_execution_role()
# construct a SageMaker estimator that calls the xgboost-container
xgboost= sagemaker.estimator.Estimator(xgboost_container,
                                       role,
                                       instance_count=1,
                                       instance_type='ml.m5.large',
                                       output_path=output_location,sagemaker_session=se
ss)
xgboost.set_hyperparameters(max_depth='5',eta='0.2',gamma='4',
                            min_child_weight ='6',
                            subsample = '0.7',silent='0',
                            objective='binary:logistic',
                            num_round = '5')
xgboost.fit({'train':s3_train_data})
```

```
2021-10-13 19:49:45 Starting - Starting the training job...
2021-10-13 19:49:54 Starting - Launching requested ML instancesProfilerReport
-1634154585: InProgress
.....
2021-10-13 19:51:16 Starting - Preparing the instances for training.....
2021-10-13 19:52:36 Downloading - Downloading input data...
2021-10-13 19:53:16 Training - Training image download completed. Training in
progress.
2021-10-13 19:53:16 Uploading - Uploading generated training model. Arguments:
train
[2021-10-13:19:53:10:INFO] Running standalone xgboost training.
[2021-10-13:19:53:10:INFO] Path /opt/ml/input/data/validation does not exist!
[2021-10-13:19:53:10:INFO] File size need to be processed in the node: 0.02m
b. Available memory size in the node: 167.03mb
[2021-10-13:19:53:10:INFO] Determined delimiter of CSV input is ','
[19:53:10] S3DistributionType set as FullyReplicated
[19:53:10] 614x8 matrix with 4912 entries loaded from /opt/ml/input/data/trai
n?format=csv&label_column=0&delimiter=,
[19:53:10] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 20 extra
nodes, 0 pruned nodes, max_depth=5
[0]#011train-error:0.228013
[19:53:10] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 16 extra
nodes, 0 pruned nodes, max_depth=5
[1]#011train-error:0.203583
[19:53:10] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 20 extra
nodes, 0 pruned nodes, max_depth=5
[2]#011train-error:0.205212
[19:53:10] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 18 extra
nodes, 0 pruned nodes, max_depth=5
[3]#011train-error:0.197068
[19:53:10] src/tree/updater_prune.cc:74: tree pruning end, 1 roots, 14 extra
nodes, 0 pruned nodes, max_depth=5
[4]#011train-error:0.197068

2021-10-13 19:53:36 Completed - Training job completed
Training seconds: 57
Billable seconds: 57
```

In [96]:

```
from sagemaker.serializers import CSVSerializer
from sagemaker.deserializers import JSONDeserializer
xgboost_predictor = xgboost.deploy(initial_instance_count=1, instance_type='ml.m5.large', serializer=CSVSerializer(),)
```

-----!

In [102]:

```
testn=np.array(test)
def predict(data, rows=500):
    split_array = np.array_split(data, int(data.shape[0] / float(rows) + 1))
    predictions = ''
    for array in split_array:
        predictions = ','.join([predictions, xgboost_predictor.predict(array).decode('utf-8')])
    return np.fromstring(predictions[1:], sep=',')
pred=predict(testn[:,1:])
```

In [103]:

```
print(pred)
```

```
[0.33513761 0.19165556 0.49315163 0.28474674 0.71769851 0.35202909
 0.28192854 0.49914482 0.31284589 0.43077078 0.34810442 0.29981658
 0.34301195 0.33186477 0.67376214 0.41907459 0.26679343 0.30077562
 0.62112433 0.31058583 0.35202909 0.21205755 0.65629113 0.3358888
 0.46753541 0.30077562 0.19165556 0.19165556 0.33255315 0.19165556
 0.33186477 0.24481148 0.38396621 0.52842742 0.2117146 0.62112433
 0.19165556 0.63169646 0.53711772 0.58896905 0.28614372 0.44628003
 0.75296295 0.61009103 0.662453 0.24434941 0.19165556 0.33453274
 0.68609709 0.63906443 0.59516221 0.2928164 0.24481148 0.4487218
 0.608913 0.48058185 0.70666736 0.2928164 0.27769023 0.60815918
 0.33988249 0.19165556 0.25121298 0.19165556 0.20211647 0.26679343
 0.32261595 0.21205755 0.3292675 0.44650674 0.72566879 0.44628003
 0.28614372 0.36235783 0.31323367 0.70666736 0.48358312 0.5087797
 0.66546053 0.21449159 0.25879988 0.3292675 0.33186477 0.4979153
 0.25121298 0.53731835 0.73699176 0.56565362 0.58896905 0.46230173
 0.42625076 0.58402932 0.71769851 0.19165556 0.30024788 0.22980398
 0.41638255 0.27596655 0.69586313 0.45703706 0.19165556 0.19165556
 0.39677283 0.2117146 0.24481148 0.30581355 0.68234271 0.32261595
 0.37774026 0.22585064 0.20375496 0.26679343 0.29332098 0.35202909
 0.19165556 0.39500839 0.33186477 0.70305252 0.71386546 0.50709975
 0.4610846 0.27251813 0.57642817 0.75537378 0.64095229 0.2651853
 0.30077562 0.27404007 0.33513761 0.7077809 0.19165556 0.662453
 0.19165556 0.32261595 0.21205755 0.56989282 0.62696296 0.27251813
 0.47508821 0.42176282 0.25121298 0.36048114 0.58896905 0.70713216
 0.69111431 0.71879643 0.19165556 0.22585064 0.21449159 0.42800304
 0.53731835 0.19165556 0.19165556 0.19165556]
```

In [104]:

```
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn import metrics
from sklearn.metrics import roc_auc_score
from sklearn.metrics import roc_curve, auc
from sklearn.model_selection import cross_validate
from sklearn.metrics import precision_recall_curve
from sklearn.metrics import f1_score
from sklearn.metrics import auc
from sklearn.metrics import log_loss
```

In [105]:

```
index=test.iloc[:,0]
print(index)
```

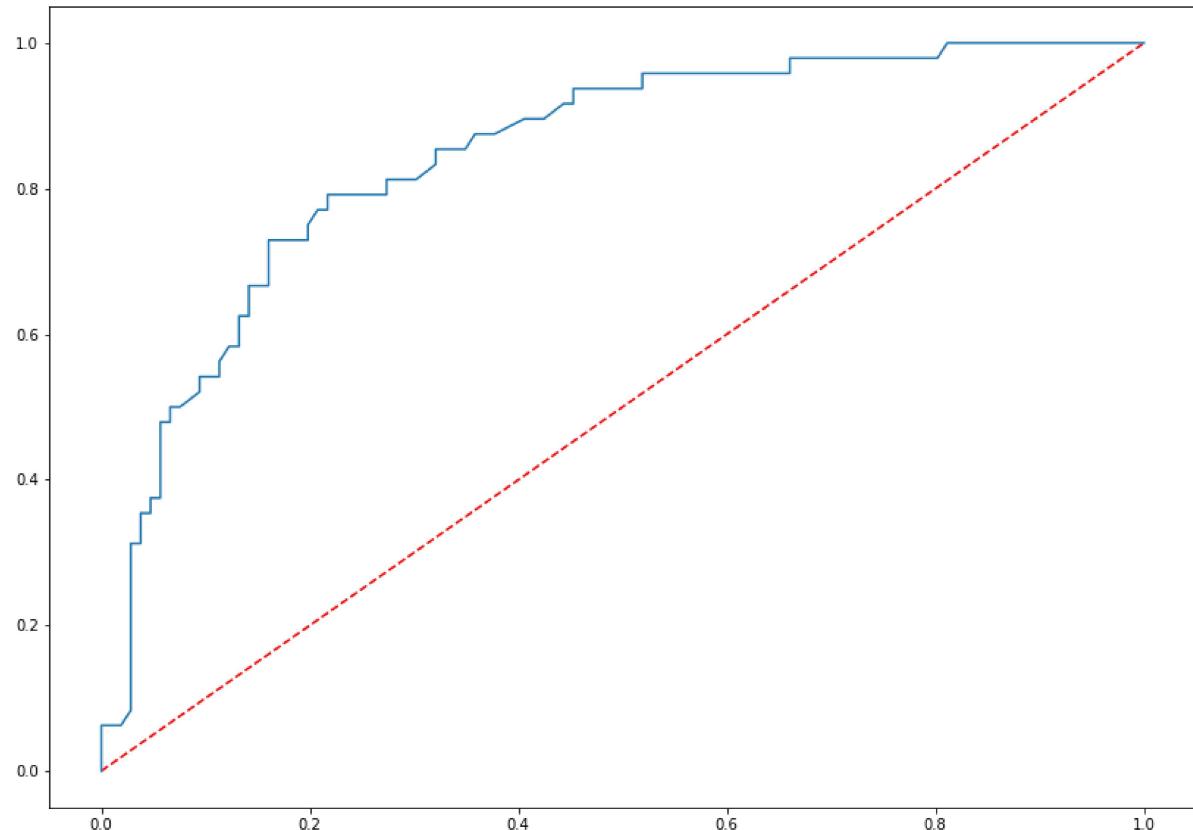
699 0
112 0
30 0
501 0
404 1
..
397 1
211 0
461 0
767 0
334 0
Name: will_get_diabetes, Length: 154, dtype: int8

In [108]:

```
fig = plt.figure(figsize=(14,10))
plt.plot([0, 1], [0, 1], 'r--')
fpr, tpr, thresh = metrics.roc_curve(index,pred)
aucxg= roc_auc_score(index, pred)
plt.plot(fpr, tpr, label=f'Linear learner, AUC = {str(round(aucxg,3))}')
```

Out[108]:

[<matplotlib.lines.Line2D at 0x7f8956799fd0>]



In [110]:

```
from sklearn.metrics import log_loss

loss = log_loss(index,pred)
loss
```

Out[110]:

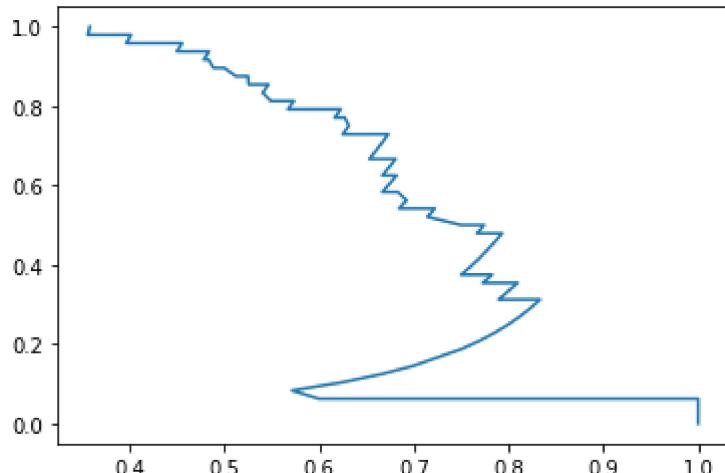
0.5006897226452417

In [112]:

```
from sklearn.metrics import average_precision_score
precision, recall, _ = precision_recall_curve(index, pred)
plt.plot(precision,recall)
```

Out[112]:

[<matplotlib.lines.Line2D at 0x7f8956818790>]



In []:

In []: