# Experiment 11

**Dhruv Singhal || 500075346 || R177219074 || AIML || Sem5**

**Different Anomaly Detection Algorithms on 2D Dataset**

# Importing required libraries

In

[3]:

**import**

time

**import**

numpy

**as**

np

**import**

matplotlib

**import**

matplotlib

.

pyplot

**as**

plt

**from**

sklearn

**import**

svm

**from**

sklearn

.

datasets

**import**

make\_moons

,

make\_blobs

**from**

sklearn

.

covariance

**import**

EllipticEnvelope

**from**

sklearn

.

ensemble

**import**

IsolationForest

**from**

sklearn

.

neighbors

**import**

LocalOutlierFactor

# Data

for this experiment we have use make\_moons, make\_blobs datsets which are already in sklearn.datasets

sklearn.datasets.make\_blobs(n\_samples=100, n\_features=2, \*, centers=None, cluster\_std=1.0, center\_box=(- 10.0, 10.0), shuffle=True, random\_state=None, return\_centers=False) Generate isotropic Gaussian blobs for clustering.

sklearn.datasets.make\_moons(n\_samples=100, \*, shuffle=True, noise=None, random\_state=None) Make two interleaving half circles. A simple toy dataset to visualize clustering and classification algorithms.

**Preprocessing & Visualization**

for this experiment we do not need preprocessing as we need to visualize different anomaly detection algorithms

# Model Building

## Covariance Matrix

This metric assesses how many standard deviations σ away xᵢ is from μ, thereby being dimensionless. An extreme observation has a large distance from the center of a distribution. It is useful in predictive maintenance as it enables finding unusual behavior of a system; upcoming down-times and malfunction can be targeted and identified, which is particularly important since finding faults in equipment, machines, and devices early can reduce the extent of the damage.

## One Class SVM

One Class SVM mode is trained in only one class, referred to as the normal class. The model learns all the features and patterns of the normal class . When a new observation is introduced to the model, then based on its learning, the One Class SVM detects if the new observation deviates from the normal behaviour then it classifies it as an oulier else the new observation will identified as an inlier. One Class SVM is based on Support Vector Machines where the binary classes are separated by a non-linear hyper plane

## Local Outlier Factor

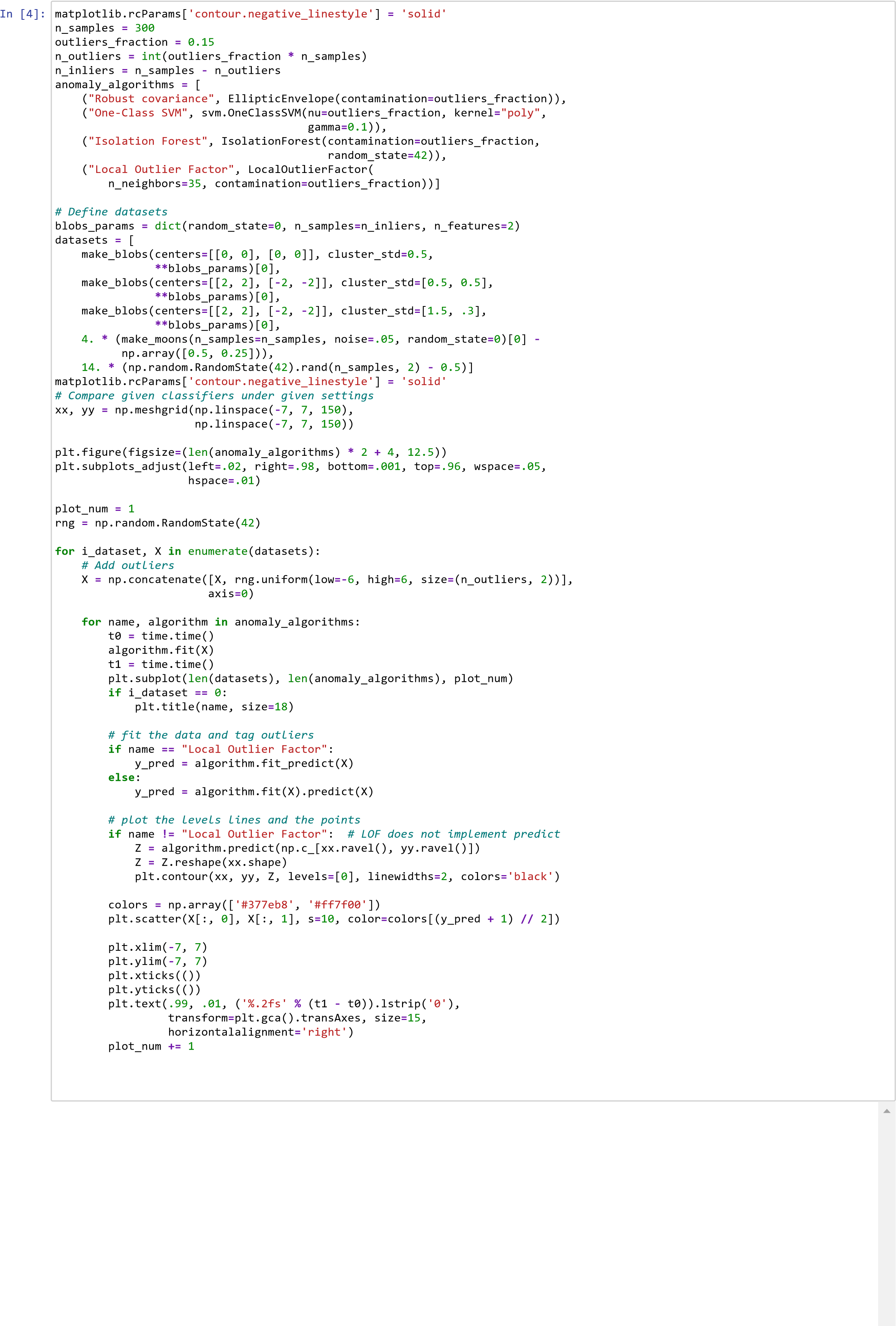
Local outlier factor (LOF) is an algorithm used for Unsupervised outlier detection. It produces an anomaly score that represents data points which are outliers in the data set. It does this by measuring the local density deviation of a given data point with respect to the data points near it.

## Isolation Forest

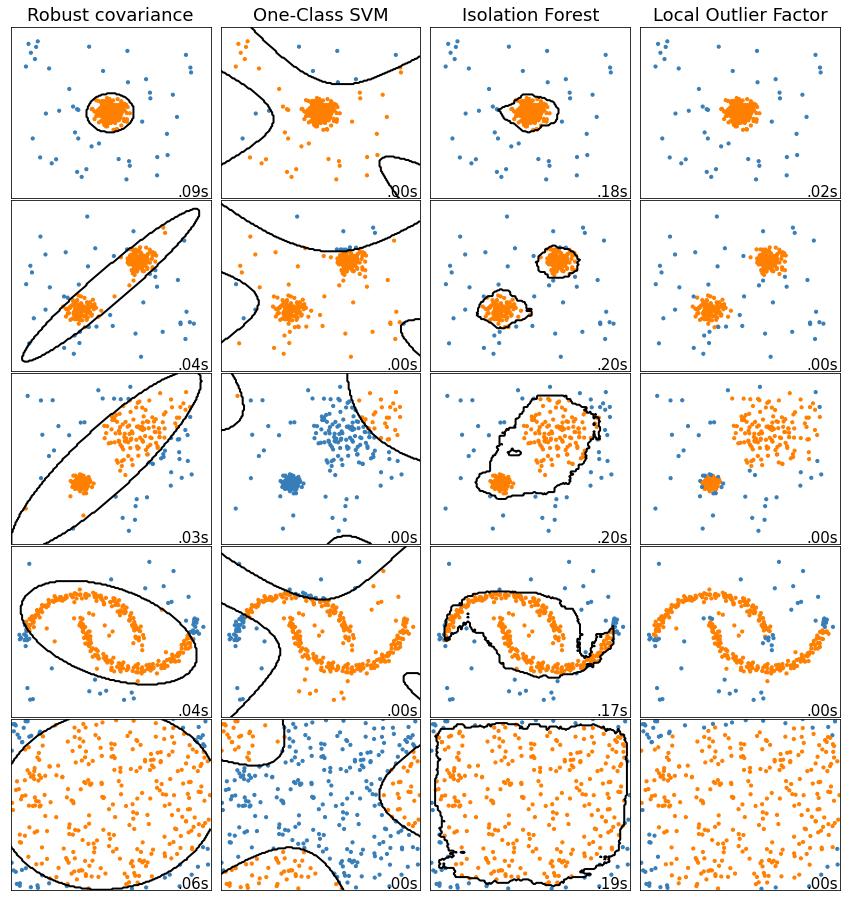
Isolation Forest is based on the Decision Tree algorithm. It isolates the outliers by randomly selecting a feature from the given set of features and then randomly selecting a split value between the max and min values of that feature. This random partitioning of features will produce shorter paths in trees for the anomalous data points, thus distinguishing them from the rest of the data.

## Compile & Train

Altering the hyperparameters to see the effect of the hyperparameter on the working of the algorithm



# Result

By this experiment i came to know the working of different anomaly detection algorithms and how they differ from each other in working and in time perspective how they find anomaly

In [ ]: