

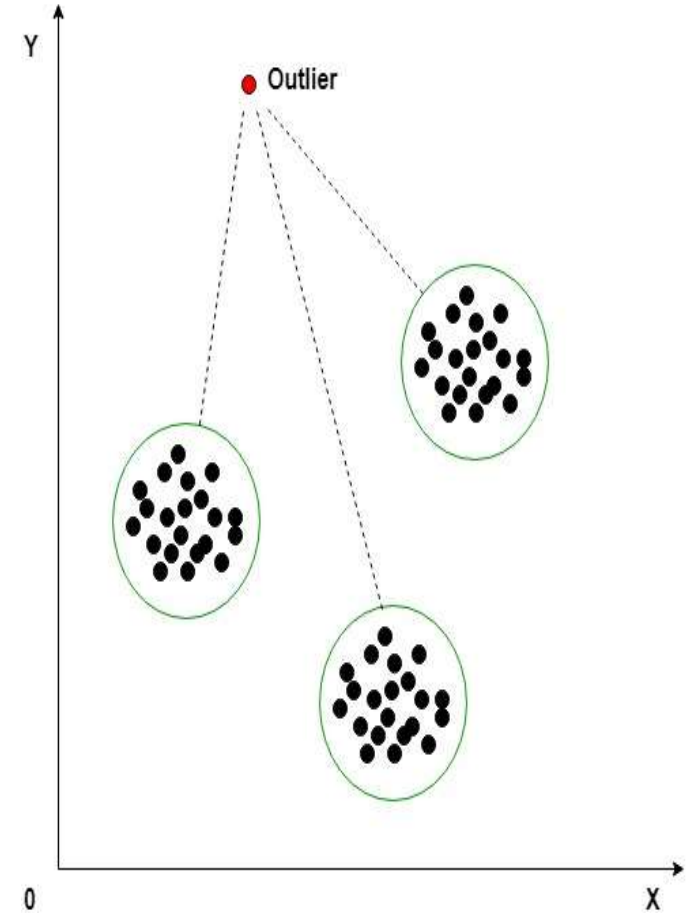
COURSE NAME: ARTIFICIAL INTELLIGENCE
COURSE CODE: CIS 412

SYED TANGIM PASHA
LECTURER,

DEPARTMENT OF COMPUTING AND INFORMATION SYSTEM (CIS)
DAFFODIL INTERNATIONAL UNIVERSITY (DIU)
DHAKA, BANGLADESH

OUTLIERS

- **Outlier:** An outlier is an observation in a given dataset that lies far from the rest of the observations. That means an outlier is vastly larger or smaller than the remaining values in the set.
- **Detecting Outliers:**
 - (1) Boxplots
 - (2) Z-score
 - (3) Inter Quantile Range (IQR)

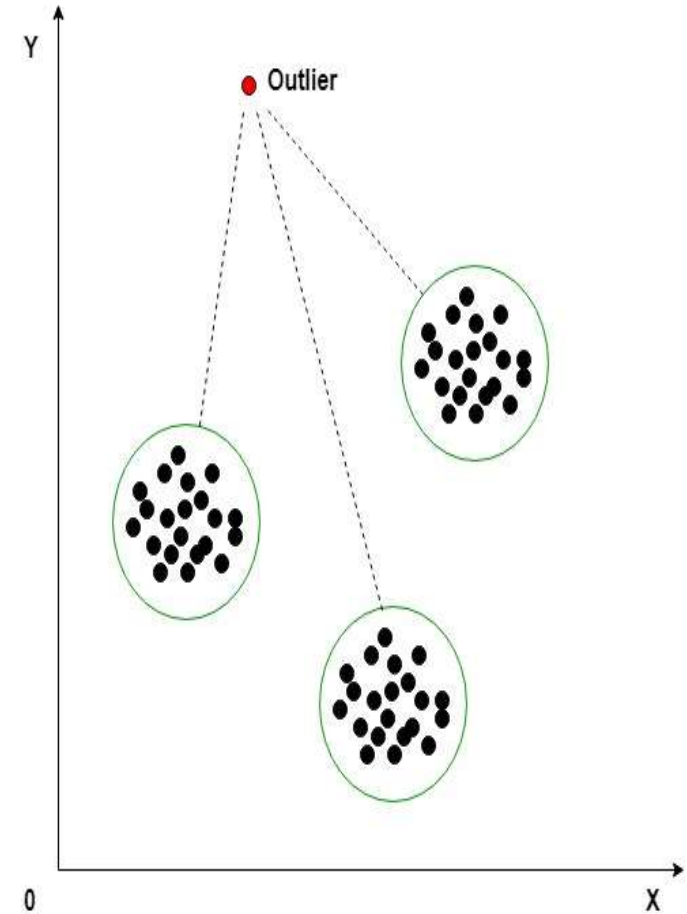


OUTLIERS

- **Handling Outliers:**

(1) Remove Outliers: In this technique, we remove the outliers from the dataset. Although it is not a good practice to follow.

(2) Cap your outliers data: Another way to handle true outliers is to cap them. For example, if you're using income, you might find that people above a certain income level behave in the same way as those with a lower income. In this case, you can cap the income value at a level that keep that intact.

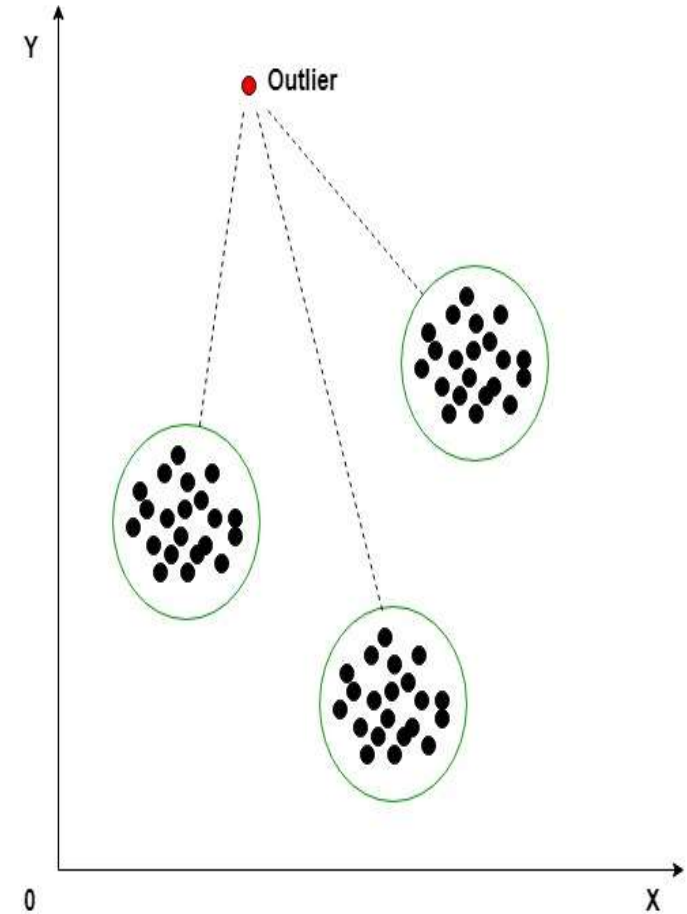


OUTLIERS

- **Handling Outliers:**

(3) **Assign a new value:** If an outlier seems to be due to a mistake in your data, try imputing a new value. Common imputation methods include using the mean of a variable or utilizing a regression model to predict the missing value.

(4) **Mean/Median imputation:** As the mean value is highly influenced by the outliers, it is advised to replace the outliers with the median value.



PARAMETER

- **Parameter:** A model parameter is a configuration variable that is internal to the model and whose value can be estimated from data.
- **Characteristics of Parameter:**
 - (1) They are required by the model when making predictions.
 - (2) They are estimated or learned from data.
 - (3) They are often not set manually by the practitioner.
- **Example:** weights or coefficients of Linear or Logistic Regression models.

HYPERPARAMETER

- **Hyperparameter:** Hyperparameters are parameters whose values control the learning process and determine the values of model parameters that a learning algorithm ends up learning.
- **Characteristics of Parameter:**
 - (1) They are often used in processes to help estimate model parameters.
 - (2) They are often specified by the practitioner.
 - (3) They are often tuned for a given predictive modeling problem

HYPERPARAMETER

- **Example:**

- (1) Learning rate (α) in Gradient Descent optimization algorithm.
- (2) Choice of Activation Function in Neural Network layers (Sigmoid, ReLU, Tanh, etc).
- (3) Number of hidden layers in a Neural Network
- (4) Number of activation units in each layer in Neural Network.
- (5) The drop out rate in Neural Network.
- (6) Number of iterations (epochs) in a training a Neural Network.

HYPERPARAMETER OPTIMIZATION

- **Hyperparameter Optimization:** In machine learning, hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm.
- **GridSearchCV:** It is the process of performing hyperparameter tuning in order to determine the optimal values for a given model. GridSearchCV is a function that comes in Scikit-learn's (Sk-learn) `model_selection` package.