Supervised and Unsupervised learning :

Supervised learning:

- Supervised learning as the name indicates the presence of a supervisor as a teacher. Basically supervised learning is a learning in which we teach or train the machine using data which is well labeled that means some data is already tagged with the correct answer.
- After that, the machine is provided with a new set of examples(data) so that supervised learning algorithm analyses the training data(set of training examples) and produces a correct outcome from labeled data.
- Supervised learning is the learning of the model where with input variable (say, x) and an output variable (say, Y) and an algorithm to map the input to the output.

Types:

It is Two types:

```
    1. Classification - Here the target variable is Discrete In nature.
(E.g. 0,1,2,3,4,5)
    2. Regression - Here the target variable is Continous In nature. (E.
g. 0.5,1.23,2.35,3.22)
```
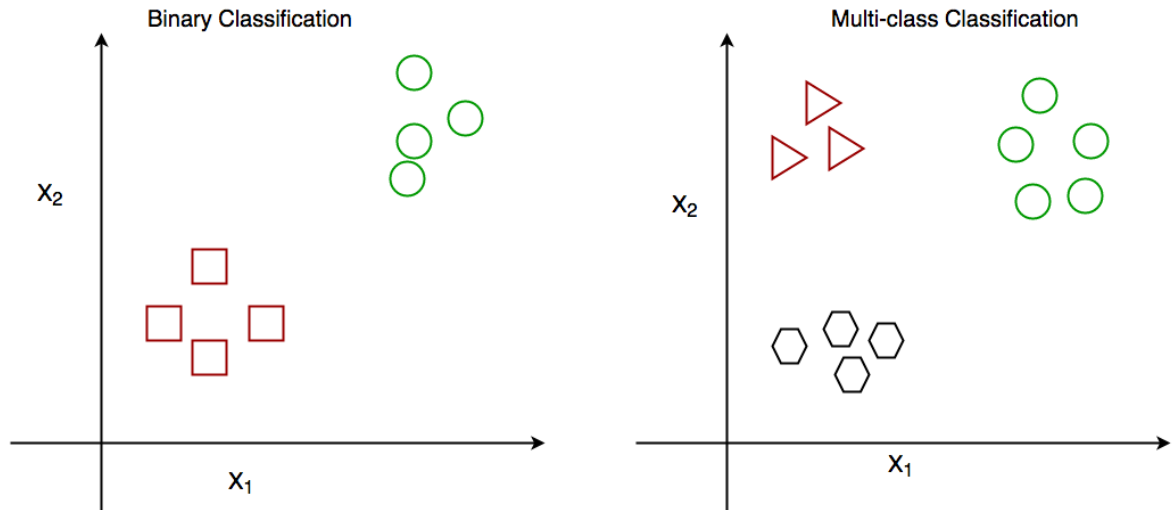
Classificaion:

- A classification problem is when the output variable is a category. it attempts to draw some conclusion from observed values. Given one or more inputs a classification model will try to predict the value of one or more outcomes.
- For example, when filtering emails "spam" or "not spam", when looking at transaction data, "fraudulent", or "authorized". Employee promoted or not.

Classification is of two types:

- Binary Classification : When we have to categorize given data into 2 distinct classes. Example – On the basis of given health conditions of a person, we have to determine whether the person has a certain disease or not. Here target contain only two value.

- Multiclass Classification : The number of classes is more than 2. For Example – On the basis of data about different species of flowers, we have to determine which specie does our observation belong to.

Binary Classification

Multi-class Classification

X_2

X_1

X_2

X_1

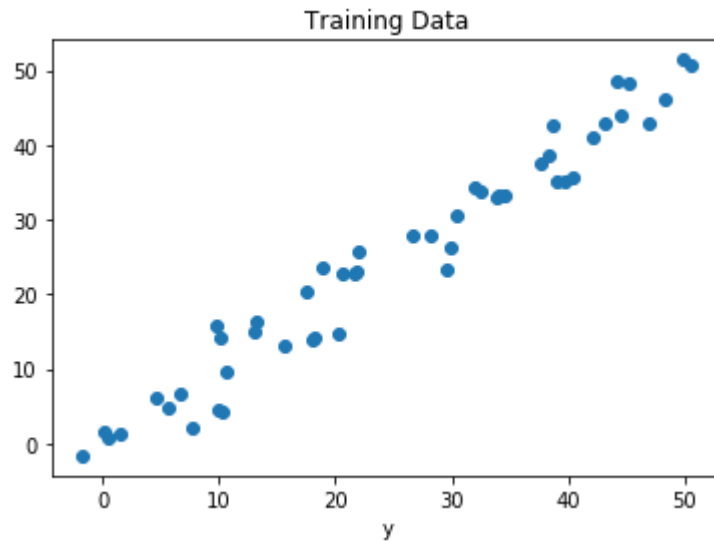## Some Ml Algo Which is used in Classificatio

1. LogisticRegression
2. KNeighborsClassifier
3. SVC
4. DecisionTreeClassifier etc

classification metrics :

- Accuracy, Confusion_matrix, Precision, recall, F1-score, ROC, AUC

Regression :

- A regression problem is when the output variable is a real or continuous value, such as "salary" or "weight". Many different models can be used, the simplest is the linear regression. It tries to fit data with the best hyperplane which goes through the points.
- It can also identify the distribution movement depending on the historical data. Because a regression predictive model predicts a quantity, therefore, the skill of the model must be reported as an error in those predictions

Training Data

## Some Ml Algo Which is used in Classificatio:

1. Linear Regression
2. Polynomial regression
3. Ridge regression
4. Lasso regression
5. SVR
6. DecisionTreeRegressor etc

Regression Metrics :

- Mean Squared error (MSE), Root Mean Squared error (RMSE), Root Mean Squared Log error (RMSLE),Mean Absoulute Error (MAE), R-Squared, Adjusted R-Squared,

Unsupervised learning

- Unsupervised learning is where only the input data (say, X) is present and no corresponding output variable is there.
- Unsupervised learning is the training of machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance.
- Here labled data not present means here no target variable.
- Here the task of machine is to group unsorted information according to similarities, patterns and differences without any prior training of data.

Source: https://www.geeksforgeeks.org/supervised-unsupervised-learning/ (https://www.geeksforgeeks.org/supervised-unsupervised-learning/)
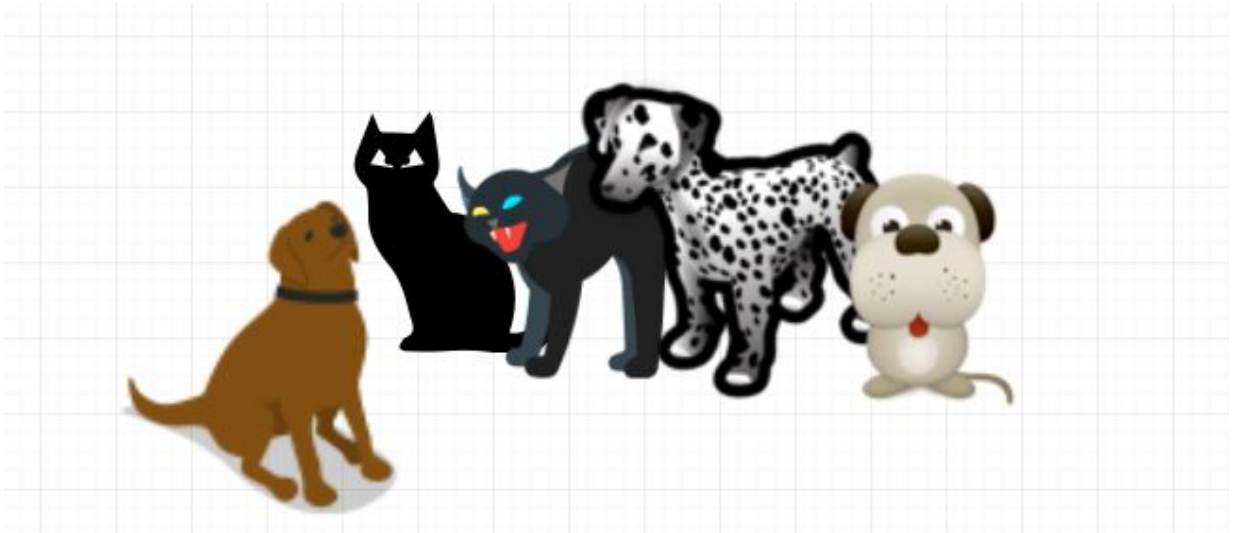
- Thus the machine has no idea about the features of dogs and cat so we can't categorize it in dogs and cats.
- But it can categorize them according to their similarities, patterns, and differences i.e., we can easily categorize the above picture into two parts.
- First may contain all pics having dogs in it and second part may contain all pics having cats in it. Here you didn't learn anything before, means no training data or examples.
- It allows the model to work on its own to discover patterns and information that was previously undetected. It mainly deals with unlabelled data.

Types:

```
    1. Clustering - discover the inherent groupings in the data
    2. Association - discover rules that describe large portions of your
data
```

Applications of Clustering

- Clustering has a large no. of applications spread across various domains. Some of the most popular applications of clustering are:

```
    * Recommendation engines
    * Market segmentation
    * Social network analysis
    * Search result grouping
    * Medical imaging
    * Image segmentation
    * Anomaly detection
```

In upcoming Notebook i will going to cover all the Clustering Algorithm So stay tunned with me.