**Introduction:**

ML has been around for decades; it has been used in several specialized applications. First ML application became a mainstream, it was a spamfilter. ML is a science of programming computers so they can learn from data. Spamfilter is a ML program that gives examples of spam emails flaged by users and based on that it can learn to flag spam. The examples that system used to learn is called training set. Each training example is called training instance. In this case the task is to flag spam emails. The experience is training data and the performance measure needs to be defined. For example, we can use the ratio of correct classified emails. This particular performance measure is called **Accuracy**, and it is often used in classification tasks.

## **ML** **CATEGORIES**

There are so many types of ML systems.ML system can be classified according to the amount and type of supervision they get during a training. There are 4 major categories:

Supervised learning, Unsupervised learning, Semi-Supervised learning, Reinforcement learning

## **Supervised learning**

In supervised learning, the training set you feed to the algorithms includes desired solutions called labels. Typical supervised learning task is classification. The spam filter is a good example of this: it is trained with many examples of emails with their class spam or ham and it must classify new email. Another typical task is to predict a target numeric value, such as the price of a car, giving a set of features called predictors. This sort of takes is called regression. To train the system, you need to give it many examples of cars, including both their predictors and their labels. Regression algorithms can be used for classification as well and vice versa. Here are some supervised learning algorithms:

1. KNN
2. Linear Regression,
3. Logistic Regression
4. Support Vector machines
5. Decision tree
6. Random Forest
7. Neural Network.

Labeled data

Square

Triangle

Labels

Prediction

Training model

Test Data

## **Unsupervised Learning**

In unsupervised learning the training data is unlabeled. The system tries to learn without teacher, for ex. We have lots of data about blocks visitors, we make a running clustering algorithm to try to detect group of similar visitors. Algorithm finds groups without our help, because we haven’t provided any data, which group a certain visitor belongs to. For ex. It may notice that 40% of visitors are males, who comic books and generally read book in the evening, while 20% young drama lovers who visit during the weekend.

Another important unsupervised task is anomaly detection, for ex. catching manufacturing defects or automatically removing outliers from data set, before feeding another learning algorithm.

## **Semi-Supervised Learning**

Since labeling data is usually time consuming and costly, we will often have plenty of enabled and some labeled instances. This called semi-supervised learning. Most semi-supervised learning combine supervised and unsupervised algorithms.

## **Reinforcement learning**

Reinforcement learning can observe the environment, select and perform actions and get reword in return or penalties in the form of negative reword, it must learn by itself what is the best strategy called policy to get the most reword over time. A policy defines what outlines the system should choose for a given situation. For ex. many robots implement reinforcement learning to learn how to work.

More formally, we describe supervised learning as learning a function that maps an input space X to some output space Y

Mathematical part of learning regression:

Y = β0 + β1x1+ β2x2+ β3x3 …+ βnxn+ ε

* Y is output (dependent variables)
* Xs are inputs (independent variables)
* β0 are coefficients to be estimated
* ε is the error term

**Traditiona approach VS Machine Learning approch**

Study the problem

Write rules

Launch!

Analyze errors

Evaluate

Study the problem

Training ML algorithm

Launch!

Analyze errors

Evaluate sol

solution

Data