Tom Mitchell provides a more modern definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

Example: playing checkers.

E = the experience of playing many games of checkers

T = the task of playing checkers.

P = the probability that the program will win the next game.

In general, any machine learning problem can be assigned to one of two broad classifications:

Supervised learning and Unsupervised learning.

# Definition:

Machine learning is is the kind of programming which gives computers the capability to automatically learn from data without being explicitly programmed. This means in other words that these programs change their behaviour by learning from data.

# Type of Machine Learning:

Machine learning can be roughly separated into three categories:

## Supervised learning

The machine learning program is both given the input data and the corresponding labelling. This means that the learn data has to be labelled by a human being beforehand.

## Unsupervised learning

No labels are provided to the learning algorithm. The algorithm has to figure out a clustering of the input data.

## Reinforcement learning

A computer program dynamically interacts with its environment. This means that the program receives positive and/or negative feedback to improve it performance.

# Machine Learning Terminology

## Classifier:

A program or a function which maps from unlabeled instances to classes is called a classifier.

## Confusion Matrix

A confusion matrix, also called a contingency table or error matrix, is used to visualize the performance of a classifier.

The columns of the matrix represent the instances of the predicted classes and the rows represent the instances of the actual class.

In the case of binary classification, the table has 2 rows and 2 columns.

Example:

|  |  |  |  |
| --- | --- | --- | --- |
| **Confusion Matrix** | | **Predicted classes** | |
| male | female |
| **Actual classes** | male | 42 | 8 |
| female | 18 | 32 |

This means that the classifier correctly predicted a male person in 42 cases and it wrongly predicted 8 male instances as female. It correctly predicted 32 instances as female. 18 cases had been wrongly predicted as male instead of female.

## Accuracy (error rate)

Accuracy is a statistical measure which is defined as the quotient of correct predictions made by a classifier divided by the sum of predictions made by the classifier.

The classifier in our previous example predicted correctly predicted 42 male instances and 32 female instance.

Therefore, the accuracy can be calculated by:

accuracy = (42+32)/(42+8+18+32)(42+32)/(42+8+18+32)

which is 0.72

## Precision and Recall

|  |  |  |  |
| --- | --- | --- | --- |
| **Confusion Matrix** | | **Predicted classes** | |
| negative | positive |
| **Actual classes** | negative | TN | FP |
| positive | FN | TP |

Accuracy: (TN+TP)/(TN+TP+FN+FP)(TN+TP)/(TN+TP+FN+FP)

Precision: TP/(TP+FP)TP/(TP+FP)

Recall: TP/(TP+FN)

## Supervised learning

The machine learning program is both given the input data and the corresponding labelling. This means that the learn data has to be labelled by a human being beforehand

## Unsupervised learning

No labels are provided to the learning algorithm. The algorithm has to figure out the a clustering of the input data.

## Reinforcement learning

A computer program dynamically interacts with its environment. This means that the program receives positive and/or negative feedback to improve it performance.

# k-Nearest-Neighbor Classifier

"Show me who your friends are and I’ll tell you who you are?"

The concept of the k-nearest neighbor classifier can hardly be simpler described. This is an old saying, which can be found in many languages and many cultures. It's also metnioned in other words in the Bible: "He who walks with wise men will be wise, but the companion of fools will suffer harm" (Proverbs 13:20 )

This means that the concept of the k-nearest neighbor classifier is part of our everyday life and judging: Imagine you meet a group of people, they are all very young, stylish and sportive. They talk about there friend Ben, who isn't with them. So, what is your imagination of Ben? Right, you imagine him as being yong, stylish and sportive as well.

The principle behind nearest neighbor classification consists in finding a predefined number, i.e. the 'k' - of training samples closest in distance to a new sample, which has to be classified. The label of the new sample will be defined from these neighbors. k-nearest neighbor classifiers have a fixed user defined constant for the number of neighbors which have to be determined. There are also radius-based neighbor learning algorithms, which have a varying number of neighbors based on the local density of points, all the samples inside of a fixed radius. The distance can, in general, be any metric measure: standard Euclidean distance is the most common choice. Neighbors-based methods are known as non-generalizing machine learning methods, since they simply "remember" all of its training data. Classification can be computed by a majority vote of the nearest neighbors of the unknown sample.