**What is Machine Learning?**

Two definitions of Machine Learning are offered. Arthur Samuel described it as: "the field of study that gives computers the ability to learn without being explicitly programmed." This is an older, informal definition.

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.

## Supervised Learning

In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

Supervised learning problems are categorized into "regression" and "classification" problems. In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function. In a classification problem, we are instead trying to predict results in a discrete output. In other words, we are trying to map input variables into discrete categories.

**Example 1:**

Given data about the size of houses on the real estate market, try to predict their price. Price as a function of size is a continuous output, so this is a regression problem.

We could turn this example into a classification problem by instead making our output about whether the house "sells for more or less than the asking price." Here we are classifying the houses based on price into two discrete categories.

**Example 2**:

(a) Regression - Given a picture of a person, we have to predict their age on the basis of the given picture

(b) Classification - Given a patient with a tumor, we have to predict whether the tumor is malignant or benign.

## Unsupervised Learning

Unsupervised learning allows us to approach problems with little or no idea what our results should look like. We can derive structure from data where we don't necessarily know the effect of the variables.

We can derive this structure by clustering the data based on relationships among the variables in the data.

With unsupervised learning there is no feedback based on the prediction results.

**Example:**

Clustering: Take a collection of 1,000,000 different genes, and find a way to automatically group these genes into groups that are somehow similar or related by different variables, such as lifespan, location, roles, and so on.

Non-clustering: The "Cocktail Party Algorithm", allows you to find structure in a chaotic environment. (i.e. identifying individual voices and music from a mesh of sounds at a [cocktail party](https://en.wikipedia.org/wiki/Cocktail_party_effect)).

## ****General Questions****

***Q: Is the grader server down?*** A: First step is to check [here](http://status.coursera.org/).

***Q: The audio in the videos is quite bad sometimes, muffled or low volume. Please fix it.*** A: You can mitigate the audio issues by turning down the bass and up the treble if you have those controls, or using a headset, which naturally emphasizes the higher frequencies. Also you may want to switch on the English closed captioning. It is unlikely to be fixed in the near term because most students do not have serious problems and therefore it is low on the priority list.

**Q: What does it mean when I see “Math Processing Error?”** A: The page is attempting to use MathJax to render math symbols. Sometimes the content delivery network can be sluggish or you have caught the web page Ajax javascript code in an incomplete state. Normally just refreshing the page to make it load fully fixes the problem

**Q: How can I download lectures?** A: On Demand videos cannot be downloaded.

**Q: Is there a prerequisite for this course?**A: Students are expected to have the following background:

* Knowledge of basic computer science principles and skills, at a level sufficient to write a reasonably non-trivial computer program.
* Familiarity with the basic probability theory.
* Familiarity with the basic linear algebra.

**Q: Why do we have to use Matlab or Octave? Why not Clojure, Julia, Python, R or [Insert favourite language here]?**A: As Prof. Ng explained in the 1st video of the Octave tutorial, he has tried teaching Machine Learning in a variety of languages, and found that students come up to speed faster with Matlab/Octave. Therefore the course was designed using Octave/Matlab, and the automatic submission grader uses those program interfaces. Octave and Matlab are optimized for rapid vectorized calculations, which is very useful in Machine Learning. R is a nice tool, but:

1. It is a bit too high level. This course shows how to actually implement the algorithms of machine learning, while R already has them implemented. Since the focus of this course is to show you what happens in ML algorithms under the hood, you need to use Octave 2. This course offers some starter code in Octave/Matlab, which will really save you tons of time solving the tasks.

**Q: Has anyone figured out the how to solve this problem? Here is my code [Insert code].** A: This is a violation of the Coursera Honor Code. Find the Honor Code [here](https://learner.coursera.help/hc/en-us/articles/209818863-Coursera-Honor-Code).

**Q: I've submitted correct answers for [insert problem]. However I would like to compare my implementation with other who did correctly.** A: This is a violation of the Coursera Honor Code. Find the Honor Code [here.](https://learner.coursera.help/hc/en-us/articles/209818863-Coursera-Honor-Code)

**Q: This is my email: [insert email]. Can we get the answer for the quiz?** A: This is a violation of the Coursera Honor Code. Find the Honor Code [here](https://learner.coursera.help/hc/en-us/articles/209818863-Coursera-Honor-Code).

***Q: Do I receive a certificate once I complete this course?*** A: Course Certificate is available in this course. Click [here](https://www.coursera.org/certificate/machine-learning) to learn about how Course Certificate works and how to purchase.

**Q: Why do all the answers in a multiple correct question say correct response when you submit the answer to an in-video question?** A: Coursera's software is designed to suggest the correctness of each state of the check box. Therefore, an answer having a correct answer tag below it means that the state of that check box is correct.

**Q: What is the correct technique of entering a numeric answer to a text box question ?** A: Coursera's software for numeric answers only supports '.' as the decimal delimiter (not ',') and require that fractions be simplified to decimals. For answers with many decimal digits, please use a 2 digits after decimal point rounding method when entering solutions if not mentioned in the question.

**Q: What is the correct technique of entering a 1 element matrix ?** A: They should be entered as just the element without brackets.

**Q: What does a A being a 3 element vector or a 3 dimensional vector mean?** A: If not described a vector as mentioned in the questions is A =

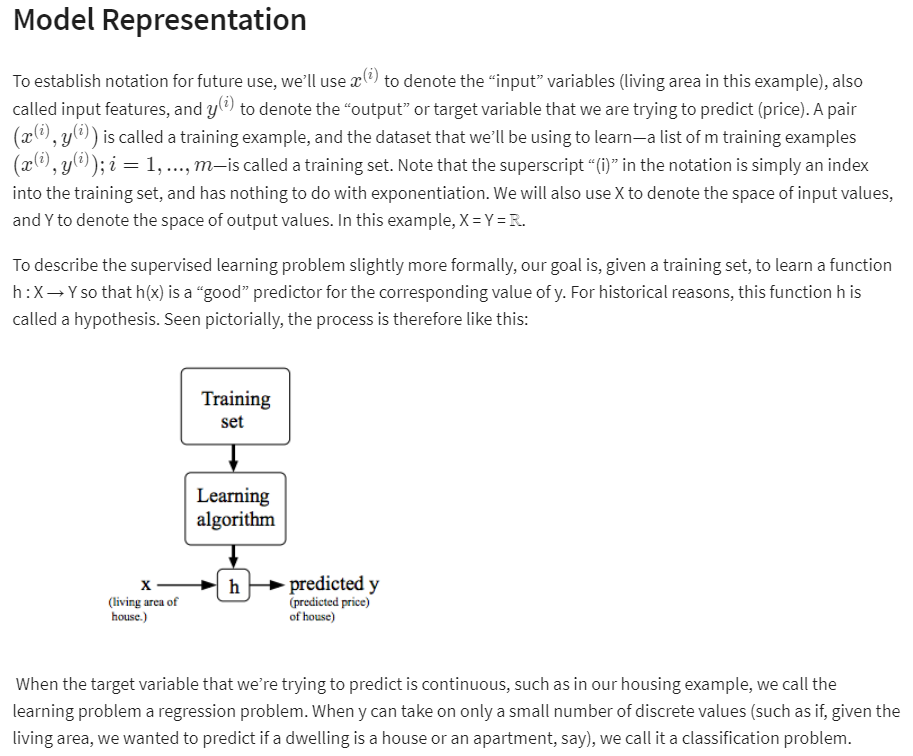
⎡⎣*element*1*element*2*element*3⎤⎦

*A*=⎣⎢⎡​*element*1*element*2*element*3​⎦⎥⎤​.

**Q: I think I found an error in a video. What should I do?** A: First, check the errata section under resources menu. If you are unsure if it is an error, create a new thread in the discussion forum describing the error.

**Q: My quiz grade displayed is wrong or I have a verification issue or I cannot retake a quiz. What should I do?** A: Contact Help Center. These queries can only be resolved by learner support and it is best if they are contacted directly.

.



# Cost Function

We can measure the accuracy of our hypothesis function by using a **cost function**. This takes an average difference (actually a fancier version of an average) of all the results of the hypothesis with inputs from x's and the actual output y's.

