**data mining** Data mining is carried out by a person, in a specific situation, on a particular data set, with a goal in mind. Typically, this person wants to leverage the power of the various pattern recognition techniques that have been developed in machine learning.

Quite often, the data set is massive, complicated, and/or may have special problems (such as there are more variables than observations). Usually, the goal is either to discover / generate some preliminary insights in an area where there really was little knowledge beforehand, or to be able to predict future observations accurately. Moreover, data mining procedures could be either **'unsupervised'** (we don't know the answer--discovery) or **'supervised'** (we know the answer--prediction). Note that the goal is generally not to develop a more sophisticated understanding of the underlying data generating process. Common data mining techniques would include **cluster analyses**, **classification** and **regression trees**, and **neural networks**.

**Statistics**

Classical statistics (here I mean both frequentist and Bayesian) is a sub-topic within mathematics. I think of it as largely the intersection of what we know about probability and what we know about optimization. Although mathematical statistics can be studied as simply a Platonic object of inquiry, it is mostly understood as more practical and applied in character than other, more rarefied areas of mathematics. As such (and notably in contrast to data mining above), it is mostly employed towards better understanding some particular data generating process. Thus, it usually starts with a formally specified model, and from this are derived procedures to accurately extract that model from noisy instances (i.e., estimation--by optimizing some loss function) and to be able to distinguish it from other possibilities (i.e., inferences based on known properties of sampling distributions). The prototypical statistical technique is regression.

**machine learning**  Machine learning involves the study of algorithms that can extract information automatically

**Artificial Intelligence** program a computer to behave and perform a task as an intelligent agent (say, a person) would.  induce new knowledge from experiences.

**Statistics** is just about the numbers, and quantifying the data. There are many tools for finding relevant properties of the data but this is pretty close to pure mathematics.

**Data Mining** is about using **Statistics** as well as other programming methods to find patterns hidden in the data so that you can explain some phenomenon. Data Mining builds intuition about what is really happening in some data and is still little more towards math than programming, but uses both.

**Machine Learning** uses **Data Mining** techniques and other learning algorithms to build models of what is happening behind some data so that it can predict future outcomes. Math is the basis for many of the algorithms, but this is more towards programming.

**Artificial Intelligence** uses models built by **Machine Learning** and other ways to reason about the world and give rise to intelligent behavior whether this is playing a game or driving a robot/car. Artificial Intelligence has some goal to achieve by predicting how actions will affect the model of the world and chooses the actions that will best achieve that goal. Very programming based.

Neural network with deep learning :

supervised neural network to run your input through several classifications, and an unsupervised neural network optimize feature selection as a pre-training step.

Natural langage analysis with google api

<https://cloud.google.com/natural-language/docs/>

Natural language toolkit python

http://www.nltk.org/