Week 1-1: Introduction

1. A computer program is said to learn from experience E with respect to some task T and some performance measure P if its performance on T, as measured by P, improves with experience E. Suppose we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather. In this setting, what is T?

   The weather prediction task.

2. Suppose you are working on weather prediction, and your weather station makes one of three predictions for each day's weather: Sunny, Cloudy or Rainy. You'd like to use a learning algorithm to predict tomorrow's weather. Would you treat this as a classification or a regression problem?

   Classification.

3. Suppose you are working on stock market prediction. You would like to predict whether or not a certain company will win a patent infringement lawsuit (by training on data of companies that had to defend against similar lawsuits). Would you treat this as a classification or a regression problem?

   Classification.

4. Some of the problems below are best addressed using a supervised learning algorithm, and the others with an unsupervised learning algorithm. Which of the following would you apply supervised learning to? (Select all that apply.) In each case, assume some appropriate dataset is available for your algorithm to learn from.

   (a) Have a computer examine an audio clip of a piece of music, and classify whether or not there are vocals (i.e. a human voice singing) in that audio clip, or if it is a clip of only musical instruments (and no vocals).

   (b) Given genetic (DNA) data from a person, predict the odds of him/her developing diabetes over the next 10 years.

5. Which of these is a reasonable definition of machine learning?

   Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.