Kevin Chen

7/24/17

Machine Learning Notes

Course: <https://www.coursera.org/learn/machine-learning>

Welcome to Machine Learning

* Machine learning is the science of getting computers to learn, without being explicitly programmed.
* Neural networks mimic how the human brain works. These type of “learning algorithms” can help make truly intelligent machines.

**Introduction**

Welcome

* Machine learning grew out of work in AI.
* Aimed to provide a new capability for computers
* Examples of uses of Machine Learning:
  + Database mining: processing large datasets
  + Applications that can’t be programmed by hand
  + Self-customizing programs. (e.g. product recommendations)
  + Understand human learning

What is Machine Learning?

* Arthur Samuel’s definition: field of study that gives computers the ability to learn without being explicitly programmed.
* Tom Mitchell’s definition: A computer program learns 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.
* Machine learning algorithms:
  + Supervised learning: teach the computer how to do something.
  + Unsupervised learning: let the computer learn by itself.
  + Other algorithms: reinforcement learning and recommender systems

Supervised Learning

* Example: apply a regression line (linear, quadratic, etc) to a data set to interpolate a point.
* In supervised learning, you are provided the “right answers” for certain inputs.
  + They are categorized into “regression” or “classification” problems.
* Regression: predict a continuous valued output
* Classification problem: identifying which discrete category an example belongs to.
  + Or rather might find the chance that the example might belong to each category
* An example may consist of more than just one feature (e.g. age and tumor size for predicting whether the cancer is malignant)
* Sometimes you want an infinite number of features (i.e. attributes).
  + The Support Vector Machine can deal with truly infinite number of features

Unsupervised Learning

* We’re given data that doesn’t have any labels. We are not told what to do with it, and we’re not told what each data point is.
  + We are simply told that this is a data set: can you find some structure in the data?
* Clustering algorithms break data into different groups that you don’t know in advance.
* Examples of applications for clustering algorithms: grouping news on the same topic, amount of gene expression, organizing large computer clusters, social network analysis, market segmentation, astronomical data analysis
* We are not giving the algorithm the right answer for the examples in the data set.
  + With unsupervised learning there is no feedback based on the prediction results.
* Cocktail party problem: we are given a set of people speaking and some microphones that (depending on where they are placed) records each person’s voice at a certain level.
  + Clustering algorithms will separate each person’s voice.
  + It doesn’t have to be different people talking – it can be background music, etc.
* Depending on the programming environment, many learning algorithms can be really short programs.
* This class uses the Octave programming environment.
  + Really fast environment – great at prototyping.

**Model and Cost Function**

Model Representation

* In supervised learning, we have a data set that is the training set.
* Notation:
  + m = Number of training examples
  + x’s = “input” variable/features. X is the space of input values.
  + y’s = “output” variable/“target” variable. Y is the space of output values.
  + (x, y) is a single training example.
  + (x(i), y(i)) is the ith training example.
* Supervised learning: start with training set, feed into learning algorithm, which produces an output function *h* (stands for hypothesis). This function takes in an input x and predicts an output y.
* How do we represent *h*?
  + . Shorthand: .
  + This is a linear regression with one variable, or a univariate linear regression.

Cost Function