**CS 229, Fall 2018**

**Problem Set #3 Solutions:   
Deep Learning & Unsupervised learning**

1. **A Simple Neural Network**
2. Gradient w.r.t .

Forward propagation:

Cost function:

Derivative of (Not needed in this problem).

Derivative of

Add subscription and sum up

1. In the dataset plot, find three points clock wise which can form a triangle to separate the dataset in two:

Subtract with the next to form three vectors, and add intercept in front:

Equations for the above three vectors. For and :

Which can re rewritten as:

This gives us the first component of matrix :

Similarly, for and we have:

And for and we have:

Stack them to from a matrix :

Explanation: given any point x, if its projection on the normal of vector  is positive then it is on the right hand of , otherwise on the left. If it lies on the right of all the vectors, it is inside the triangle. (note the intercept is the last component of W which can be easily adjusted to be the first component)

Output layer matrix can be:

Because the input can only be one of the eight values of h:

This guarantees that only for dataset inside the triangle could the product be negative, which will then make f(x) = 0. All the other cases will have , which makes f(x) = 1.

1. The activation function for is changed to , provide a set of weights that makes it achieve 100% accuracy.