**Machine Learning : Linear Regression**

**A characterization of machine learning problems**

Table

Description automatically generated

**Example of a regression problem**

Chart, line chart, scatter chart

Description automatically generated

**How to estimate how much a house will sell for**

* **Just draw a straight line through the data (linear)**

**Supervise learning because you are given the answer from some sample data and it’s regression because you are predicant a real value output on just a label (classification)**

**Goal**

* **Map data X to real value Y**
* **Linear regression, mapping is linear function of X**
* + **Linear Function**

**In this lecture the linear function has been tuned down to only look at one feature**

Chart, line chart

Description automatically generated

Chart, scatter chart

Description automatically generated

* **We want to find a line such that the y values governed by our line and the y values governed by ground truth are minimized.**
* **Goal :** 
  + **Squared to remove the sign**

Text

Description automatically generated

A picture containing table

Description automatically generated

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

**How do we get ?**

**She didn’t divide by 6 (2n) in the previous 2 examples**

Chart

Description automatically generated

**Ideally our loss function space would look like a bowl shape**

Chart, line chart

Description automatically generated

**Linear regression can only learn linear functions**

Chart, surface chart

Description automatically generated

* **We can see the ideal bowl shape**
* **We always want to change the values of theta that get us to the bottom of the bowl**

**Updates using gradient descent**

* **Used to get to the bottom of the bowl**
* **Start with initial guess for**
* **Repeatedly changed to make smaller**
  + - **A big learning rate can cause you to overshoot the bottom of the bowl and you actually end up going backwards, but the pro is that it is fast**
    - **A small learning rate will definitely get your there, but it will take longer**
* **You can expect gradient descent to work beautifully only when our loss function space is convex, like a smile**

Text

Description automatically generated

Text, letter

Description automatically generated

Shape

Description automatically generated with low confidence

* **Positive number because the derivative is the tangent line at that point and that line has a positive slope**
* **It is good that**

Chart

Description automatically generated with low confidence

* **Negative slope example**

Diagram

Description automatically generated

Chart, line chart

Description automatically generated

* **Once we hit an optima the derivative will be 0, the tangent at that point will be flat so it will not change, it will converge**

Text, letter

Description automatically generated

Text, letter

Description automatically generated