Teacher's section

hw = learned function

for linear regression
hw(x) = w1*x + w0

Quadratic
hw(x)=w2*x^2+w1*x+w0

d(f1(w)+f2(w)+f3(w)+...)/dw

Finding values of w1, w2 that minimize loss can be done using gradient descent as well

y=foo(x,w1,w2,w3,...)

or

y=foo(x1,x2,x3.,x4.., w1,w2,w3,w4...)

for complicated functions like this use gradient descent to find the values of w1,w2,w3,... that minimize loss

loss = sum over all of your data points

sum(y-hw(x1,x2,x3...))^2

Gradient descent

Loss function loss(w1,w2,w3,w4...)

loss(w1,w2,w3,w4...) = sum_data points(h(w1,w2,w3,..., x1,x2,x3,...)-y)^2

Gradient of loss function

gradient(loss(w1,w2,w3,w4...)) =>

{delta_w1, delta_w2, delta_w3, ...} = v

step w values in direction of vector v

w1'=w1+alpha*delta_w1
w2'=w2+alpha*delta_w2
w3'=w3+alpha*delta_w3

For linear example

y=x*w1+w0

loss=sum_j(y_j - h_w(x))^2

{
d(loss)/dw0=sum_j(y_j-h_w(x_j))
d(loss)/dw1=sum_j(j-h_w(x_j))*x_j

gradient = {d(loss)/dw0, d(loss)/dw1}
={sum_j(y_j -hw(x_j)), sum_j(y_j-hw(x_j)*x_j)}

w0'=w0+alpha*sum_j(y_j -hw(x_j))

w1'=w1+alpha*sum_j(y_j -hw(x_j)*x_j)

w0=w0'
w1=w1'

}repeat until it converges

Linear classification
Data points consists of parameter values along with a classification type

Parameters {x1,x2,...}
Type - what you are trying to classify by

In example data set would be
(x1,x2, type)

Example classify football players based on their height and weight
x1=height
x2=weight
type={wider receivers, blocker}

Data
{5'6", 350 lbs, blocker}
{6'4, 280 lbs, wide receiver}
{5'7", 380, blocker}
{6'2", 270 lbs, wide receiver}

Find a function that classifies a player as a wider receiver or blocker based on their weight and height

f(weight,height)={blocker/wide_receiver}

Idea: find a function for which all elements above are of one type and all elements below are the other type

If w1*weight + w2*height > w0=> wide receiver

If w1*weight +w2*height <= w0=> blocker

w1*weight + w2*height - w0 > 0 => wide receiver

w1*weight + w2*height - w0 <=0 => blocker

What are values for w0, w1 and w2
Learn these values from data

In slide, find lines that separate points into black points and white point
find values w0, w1, w2

Define loss function that quantifies how well function does at dividing points into groups
Simple function: just count the number of points on wrong side of line

More complicated function: Take into consideration distance between points on wrong side of line and line.