Daniel Campos September 5th, 2013

RCS:campod2 RIN:660996361

Machine Learning Problem Set 1

1. Exercise 1.3
   1. Show that

Since if

And the is miscategorized then if

Y must be -1 and if Y must be 1 by that nature for it to be miscategorized the signs of Y and will always be opposite, when one is + the other is – and vice versa. Through this any times they are multiplied it will result in a negative number therefore less than zero.

* 1. Show that
     1. ?
     2. Therefore
  2. Argue that the move from w (t) to w(t+1) is a move in the right direction.
     1. Since we are constantly bringing closer to 0 as it is miscategorized we are adding to the w vector. By doing so we are slowly beginning the move to include it within the right section and after enough adjustments to the w vector will be positive and one therefore making each t+1 into a step in the positive direction

1. Exercise 1.5
   1. Learning based approach because we are not sure what full relation a medical test has to age and we can go pinning it down through learning from data.
   2. Design based approach since the problem is well specified
   3. Learning approach because the relation between charges and fraudulent ones is not known but with enough data it can be found.
   4. Design since it is a well specified problem and even without data using gravity and air resistance we can figure out the relation
   5. learning based approach because the relation between traffic and the cycles is not fully known and isn’t necessarily continuous and you need data to predict
2. Exercise 1.6
   1. Supervised since immediately it can be known if the user should have been recommend the book or not. The Training data is the same as online learning and usually the initial responses to book suggestions.
   2. Reinforcement based since there is no correct or incorrect move only some output and a grade for how good the move was. The training data may be a short game to demonstrate a the effect of placing a x or o and how the opponent will respond.
   3. Supervised since each movie will have a set result of genres it is categorized in as long as there is a consistent system. There is a set input and output pair that is supposed to be. For Training data the system may use an online database with movies that have their genres present.
   4. Unsupervised since much like a language one may try to play the guitar at many different points and ways and initially there will be no semblance of what music or the sound should be but you will gradually develop an ability to make music. It can also be argued that this can be Reinforcement because every time your play a cord (weather alone or in context) you have an output (sound) and the relative goodness of the output (did it sound good?)
   5. Supervised since for every occasion where someone is allowed to have a amount of debt there is a set response, either is was too much debt or not. The training set for this will be the bank history for the customer default rate and how much debt they were allowed to acrue
3. Exercise 1.7
   1. One 3

Two 3

Three 1

Zero 1

* 1. One 3

Two 3

Three 1

Zero 1

* 1. One 3

Two 3

Three 1

Zero 1

* 1. One 3

Two 3

Three 1

Zero 1

1. Problem 1.1

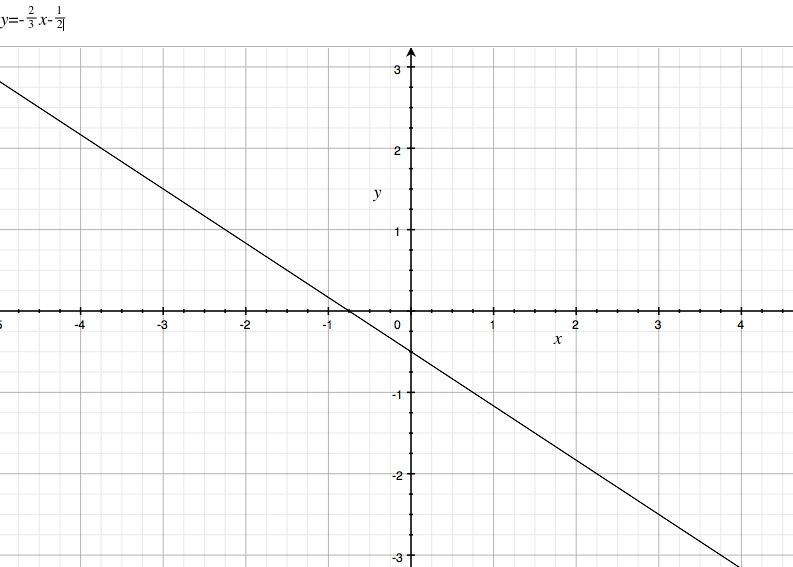
The probability of drawing out a second black ball given that the first ball drawn was black is 2/3.

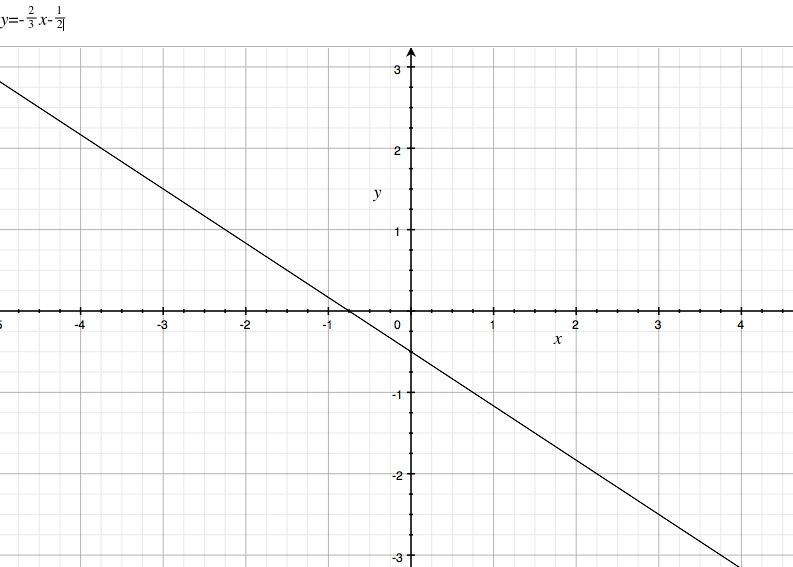
P(Black and Black)= 1- P(White and black)

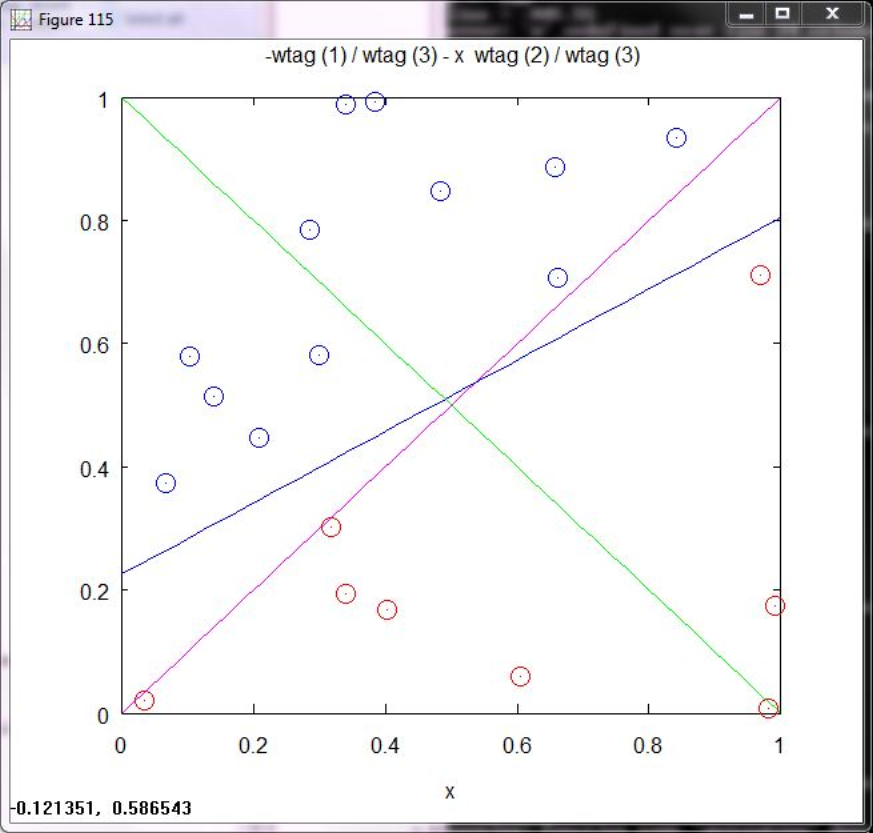
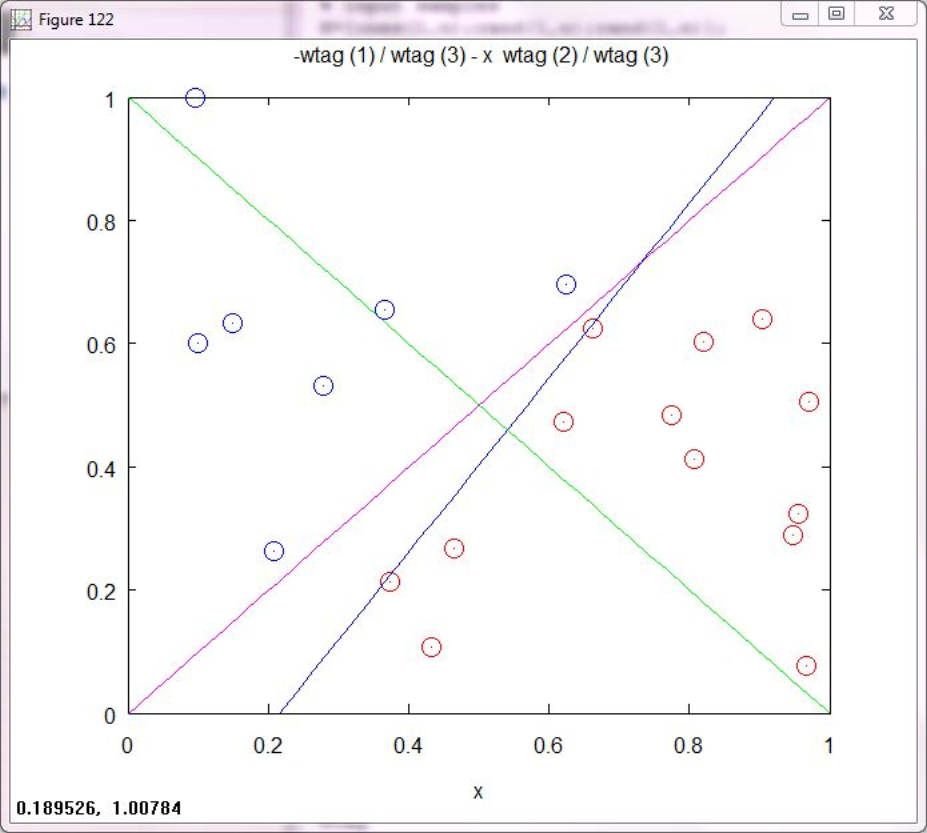
P(White and Black)= P(White Given Black)\*P(Black)= 1/3 \*1=1/3

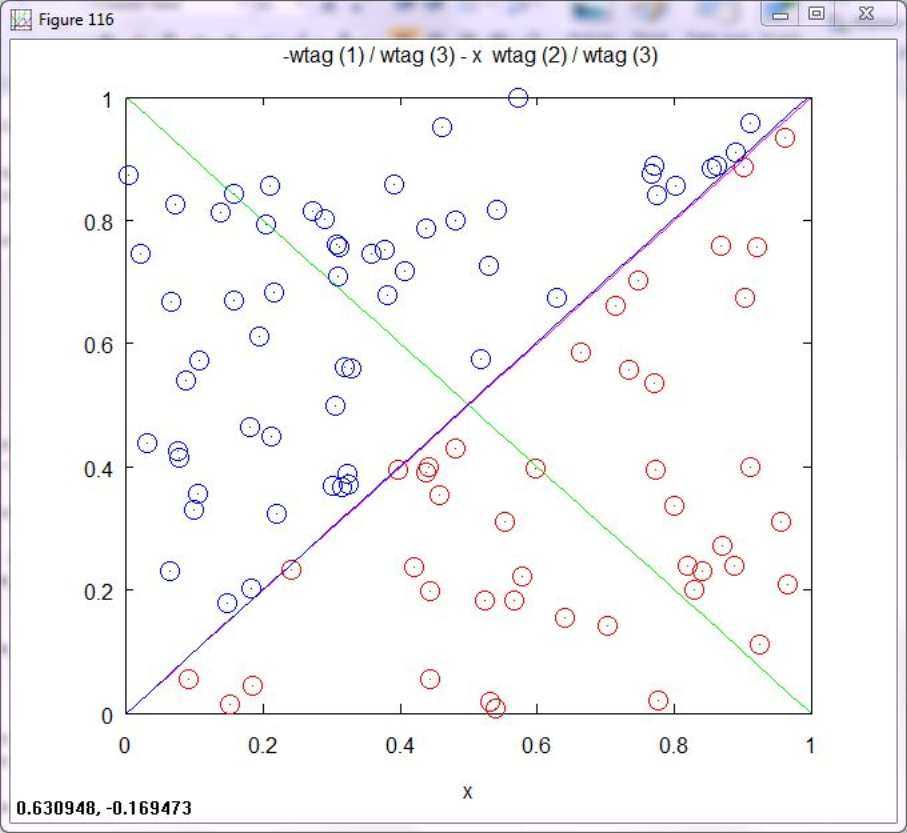
Though we have 4 marbles we have 3 potential outcomes given that the first draw was a black marble. Based on this there are 2 outcomes that generate a second black and one that generates a white so odds 2/3

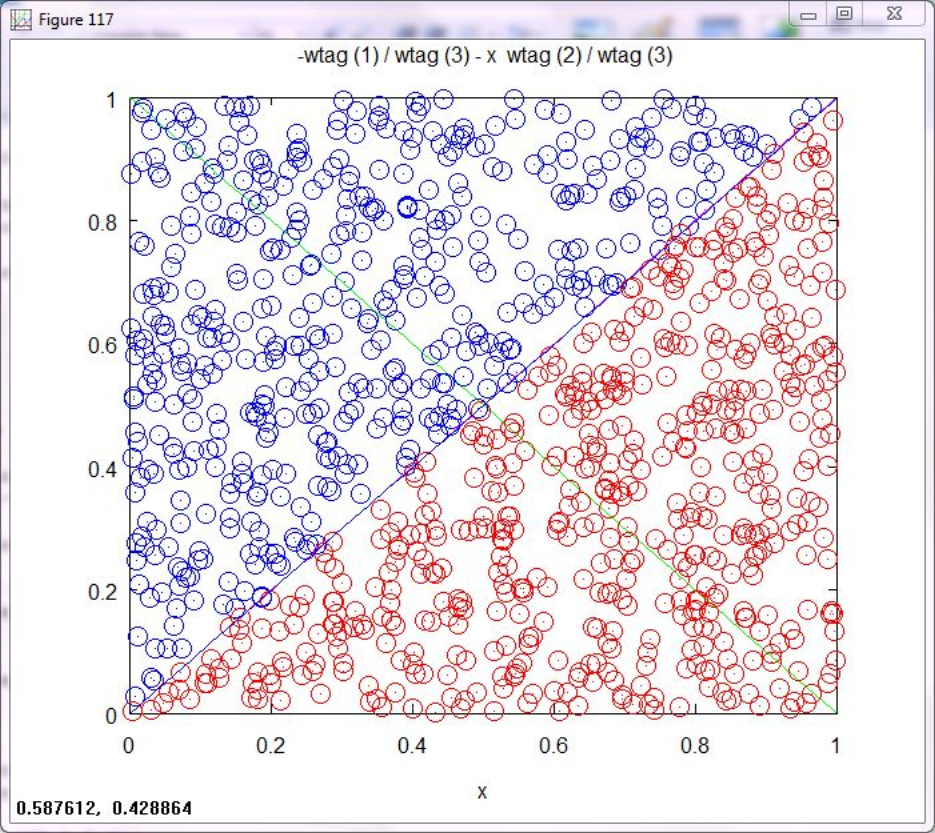
1. Problem 1.2
   1. so then a= and b=
   2. 🡪 



case w=-[1,2,3]🡪  is the same graph because the negatives cancel out.

1. Problem 1.4 (a-e)
   1. I plotted all the points in the graph-displayed bellow. I programed it in octave. Green line is the initial hypothesis the Blue line is the final hypothesis and the purple line is function f.
   2.  Iterations before complete 34. Our final hypothesis G is quite a ways off from the function f but still effectively divides all the points correctly
   3. 
   4. Iterations before complete 36. Our final hypothesis G is quite a ways off from the function f but still effectively divides all the points correctly It is also much closer on most occasions to many of the red points than our function f

 Iterations before complete 75 at this point the initial function f is close to our final hypothesis g but slightly of in the higher range

* 1. 

Iterations before complete 571 at this point the initial function f is almost identical to our final hypothesis g