






Quiz Week 1 - Content

Learning Types

Your supervisor at *fakesellingstuffcompany* gives you data on your companys customers, containing age, sex, income, residence, and money used at the company. Note that this is a 5-dimensional data set.

- Your supervisor would like to know if we can predict how much a client is worth for the company from basic information to help marketing target new customers. e.g. he would like to know if we can predict/estimate money used at the company from the attributes age,sex,income residence. Is this a Supervised or Unsupervised learning problem? 
- The supervisor asks you to learn whether you client base can be split into separate profile groups depending on their attributes. Is this a Supervised or Unsupervised Learning problem? 

Regression or Classification

- In the previous question you were asked to make a model that could predict money used at company from (age,sex,income,residence). Is that regression or classification? 
- Recognizing the color wine as white, rose, red. Classification or Regression? 
- Predicting a students grade in machine learning as af function of previous grades (on the 12 scale). Regression or classification. 

Learning and/or Design

The police would like your help with criminal facial detection. They believe you can classify whether a personal is a criminal by looking at his face alone. Unfortunately, all the mugshots taken was at a 45 degree angle. Your job is to build a criminal classifier system. You do as follows.

Step 1: You write an algorithm that rotates the face -45 degrees so all suspects (almost criminals) faces outwards.

Step 2: After looking at the rotated images you discover that the position of the eyes is always 3cm down in the picture. You decide to build a feature extractor that looks 3 cm down in the image and locates the eyes. From this the feature selector computes the length between the eyes, the vertical distance to the nose, the mouth and the horizontal distance to the ears.

Step 3: You enter the features extracted from images of known criminals and innocent people into the perceptron algorithm and construct a classifier (represented by a set of weights).

The final system is then the 3-stage algorithm: Rotate, Feature Extract, Classify with weight from step 3.

Is step 1 learning or design?

Is step 2 learning or design?



Is step 3 learning or design?

Do you think our approach will work?



Can you extend/alter the problem/algorithm to estimate the length of the prison time given?

Ranking Posts Reddit, a well known social collaboration online bulletin board, has an interesting way of ranking posts and comments. The posts are scored as follows

```
s = ups-downs
order = log(max(abs(s), 1), 10)
if s > 0:
    sign = 1
elif s < 0:
    sign = -1
else:
    sign = 0
seconds = epoch_seconds(date) - 1134028003
return round(sign * order + seconds / 45000, 7)
```



see <http://amix.dk/blog/post/19588> for an longer explanation.

Notably, Reddit ranks comments in a completely different way. <https://possiblywrong.wordpress.com/2011/06/05/reddits-comment-ranking-algorithm/>

Hacker News is a little different. The scoring is

$$(P - 1)/(T + 2)^G$$

where, P = points of an item (and -1 is to negate submitter's vote), T = time since submission (in hours), G = Gravity, defaults to 1.8. See <http://amix.dk/blog/post/19574> for more details

Why do these scores look like they do? How do you think they came up with them? Is it a machine learning approach or pure design?


Cost Functions

Least Squares You work at a company, and you have just released a new system based on machine learning using linear regression with the least squares error measure. You did not have time to look much at the final product but in testing everything looked fine. One of your colleagues, Clumsy, was responsible for running the learning algorithm on the final data you collected.

The following morning Clumsy comes to you and says: *I made an error on the few datapoints we received yesterday. I forgot to scale them down as we did with all the other data. But it is less than 1 percent of the data, so it is probably not a problem – the other data will dominate, right, since the error measure is an average.* **Is Clumsy right?**



Generalization

- Does our current generalization bound give us any bounds on the out of sample error for the perceptron learning model? 
- We want to learn $f : X \rightarrow \{-1, +1\}$. Assume f is probabilistic and independent of the input domain X . $P(f(x) = 1 \mid x) = P(f(x) = 1) = 0.2$. What this means is that when f is evaluated it ignores the input and flips its private biased coin independently of all other invocations and returns the result. What is the best out of sample error we can achieve? What if $P(f(x) = 1) = 0.5$? What are the optimal classifiers for these two cases?