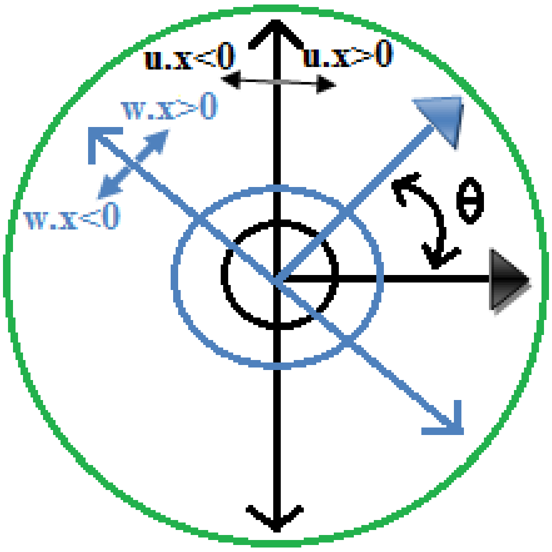
E0 270 Machine Learning *Due:* Apr 17, 2012

Assignment 6 - Report

Sachin Nagargoje *04-04-00-10-41-11-1-08449*

1 Solution: Problem 1

Consider the unit circle in the figure below.



The black line represents unit vector u and the green line represents unit vector w. The angle between them is θ. As we can see from figure, each of these vectors partitions the whole space $\mathbb{R}^2$ into two subspaces. One of these subspaces contains the vectors for which the dot-product is positive and the other subspace contains the vectors for which the dot-product is negative. Thus, given two vectors u and w, partitions S into 4 subspaces containing the vectors x,such that

1. $u.x \geq 0$ and $w.x \geq 0$ subtending angle $\pi - \theta$ at the center of unit circle.
2. $u.x \leq 0$ and $w.x \geq 0$ subtending angle θ at the center of unit circle.
3. $u.x \geq 0$ and $w.x \leq 0$ subtending angle θ at the center of unit circle.
4. $u.x \leq 0$ and $w.x \leq 0$ subtending angle $\pi - \theta$ at the center of unit circle.

where the equality holds only at the corresponding lines and θ is given by,

$$cos(\theta) = \frac{u.w}{\|u\|_2 \|w\|_2} = u.w$$

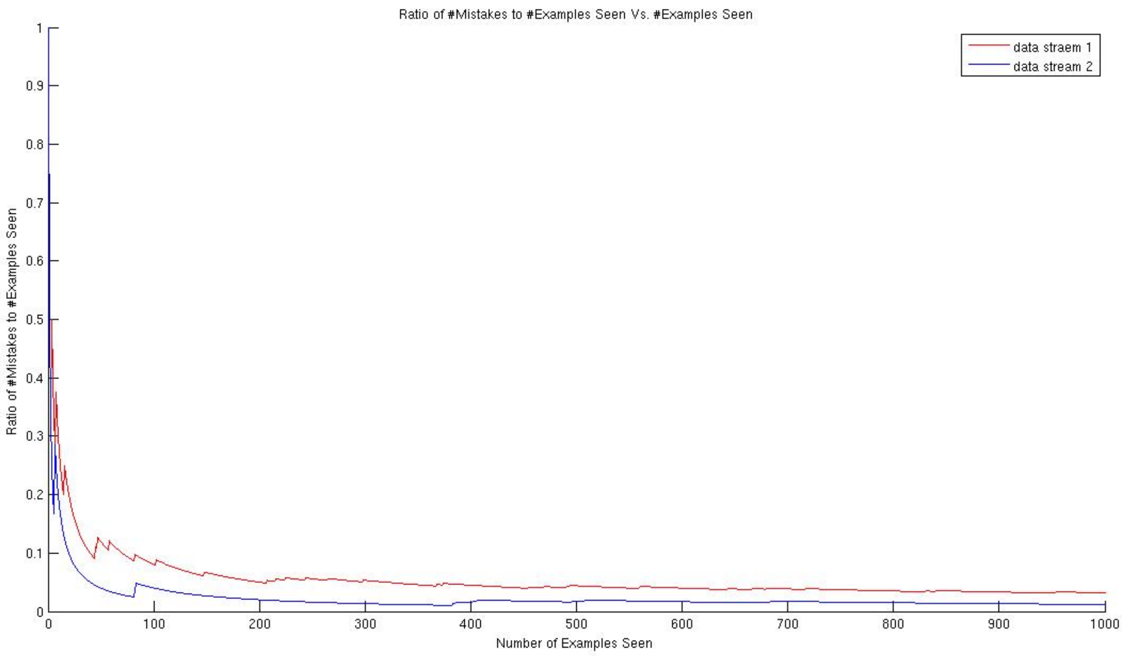
since $\|u\|_2 = \|w\|_2 = 1$.
Thus the probability that $w.x$ and $u.x$ will have different signs , i.e.

$$\boxed{\mathbb{P}_{x \in S}[sign(w.x) \neq sign(u.x)] = \frac{2\theta}{2\pi} = \frac{\theta}{\pi} = \frac{cos^{-1}(u.w)}{\pi}}$$

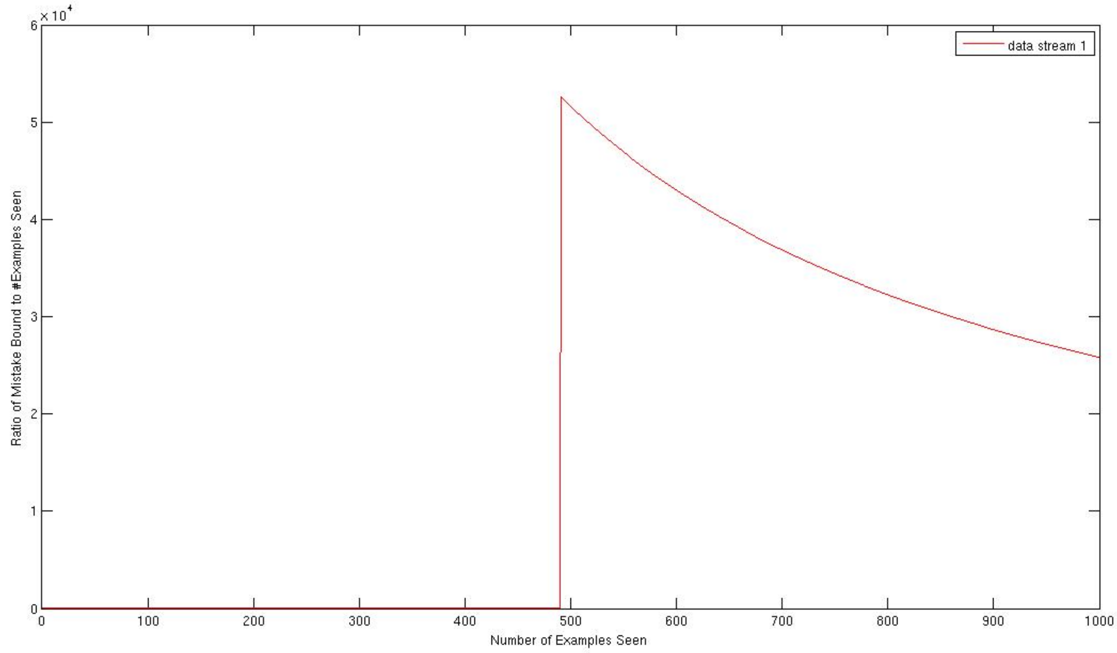
2 Solution: Problem 2

2.1 Part a

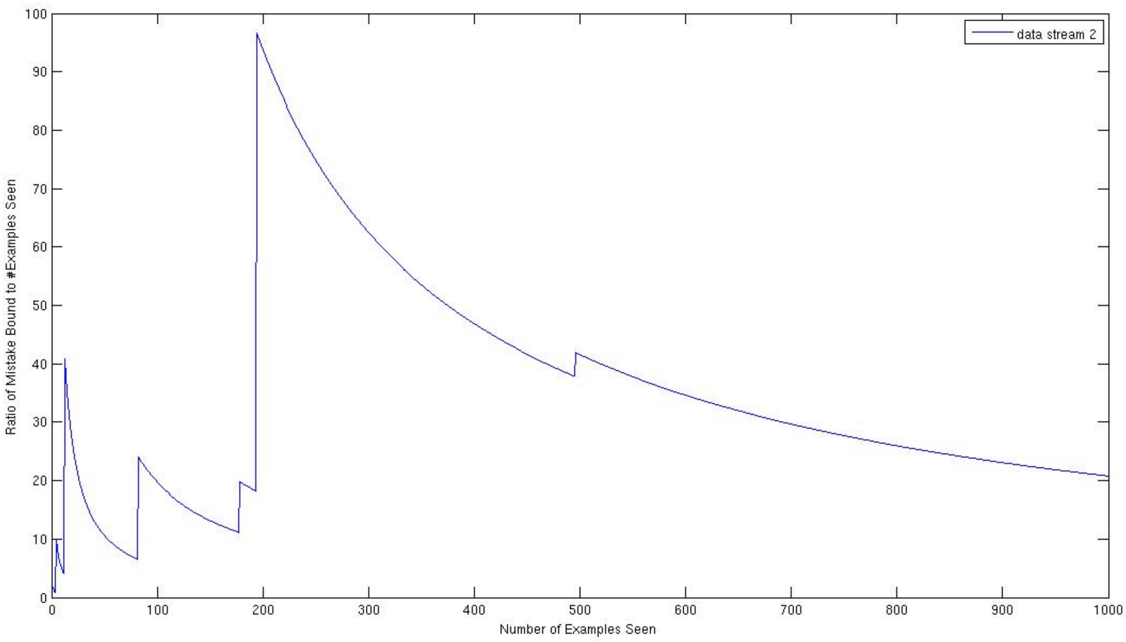
Ratio of the number of mistakes made relative to the total number of examples seen as a function of the number of examples seen



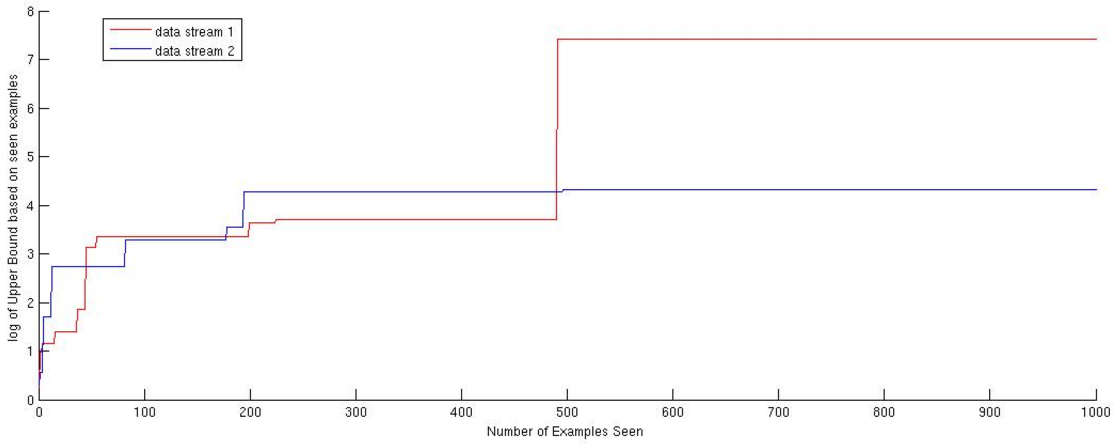
Upper bound on this ratio obtained from the mistake bound for data stream 1



Upper bound on this ratio obtained from the mistake bound for data stream 2

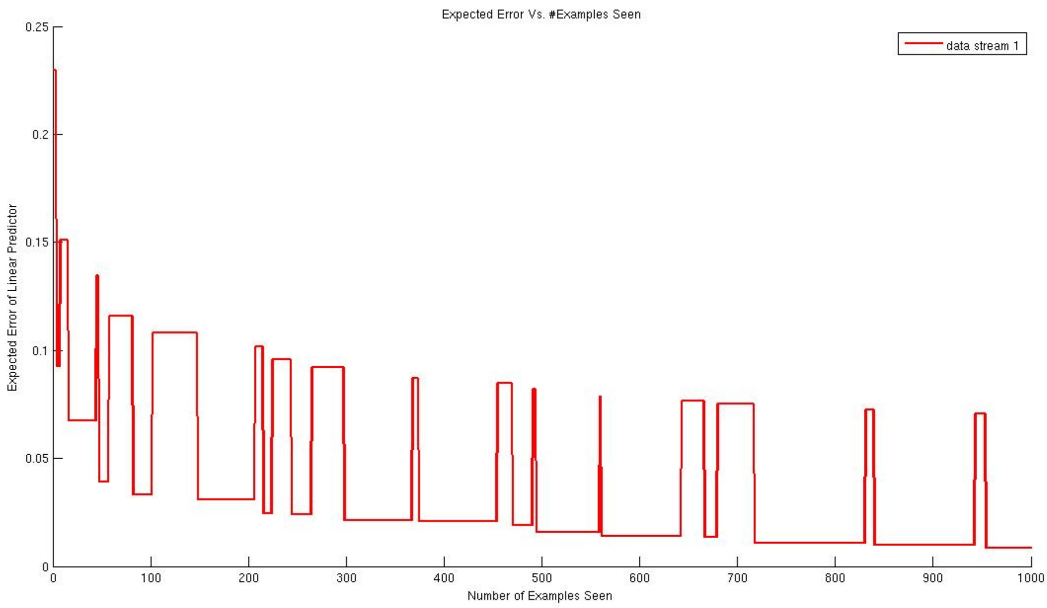


Upper bounds for data stream 1 and 2, comparison on log scale.



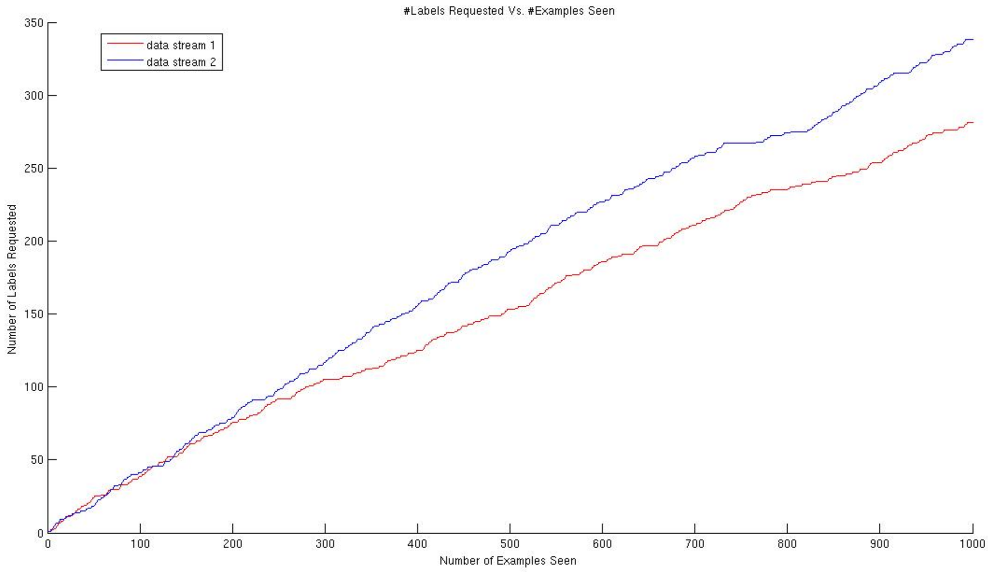
2.2 Part b

Expected error of the linear predictor learned by the perceptron algorithm as a function of the number of examples seen.

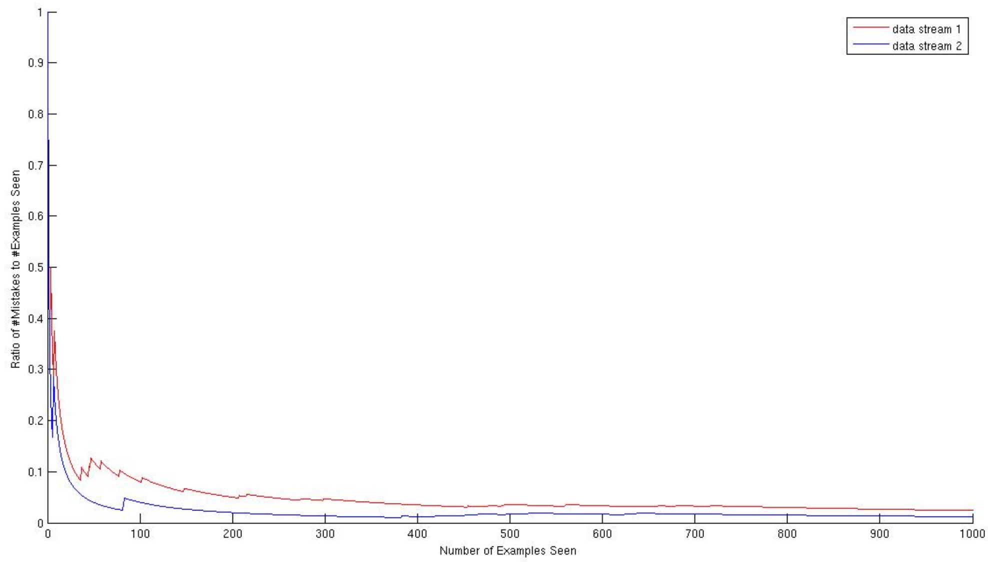


3 Solution: Problem 3

Number of labels requested as a function of the number of examples seen

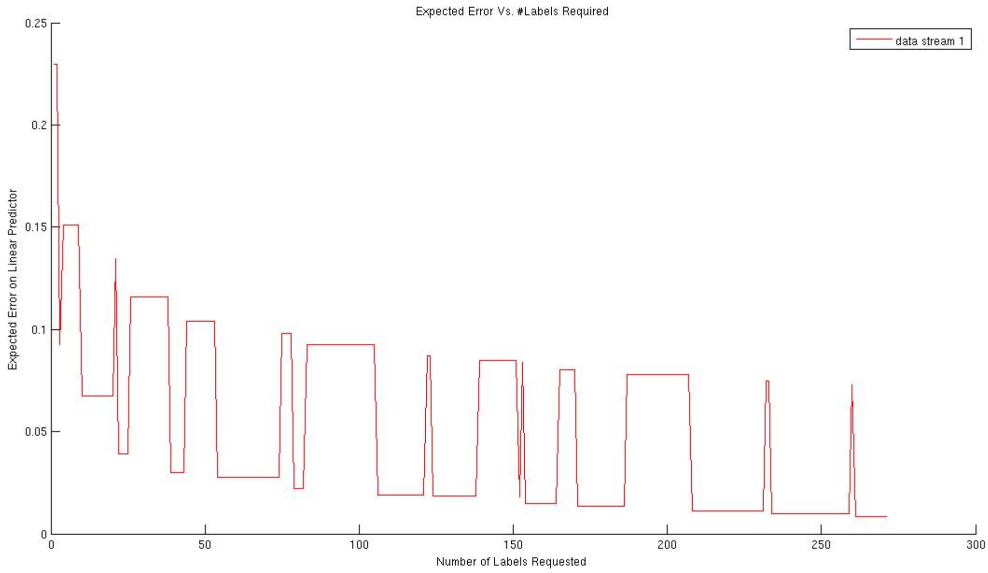


Ratio of the number of mistakes made relative to the total number of examples seen as a function of the number of examples seen



Comparing this plot with plot in 2(a), we get that the final value is lower in this case.

Expected error as a function of the number of labels requested



4 Solution: Problem 4

The solution to problem 1 holds good here for $n-dimensioal$ space as well.
Let U and W be the two hyperplanes, defined by the vectors u and w respectively i.e. planes containing all the vectors which are *orthogonal* to vectors u and w respectively. These two hyperplanes will partition S into four parts containing vectors x (similar to the solution 1) such that,

1. $u.x \geq 0$ and $w.x \geq 0$.
2. $u.x \leq 0$ and $w.x \geq 0$.
3. $u.x \geq 0$ and $w.x \leq 0$.
4. $u.x \leq 0$ and $w.x \leq 0$.

Let θ be the angle between the vectors u and w in the $2-dimensioal$ plane containing vectors u and w. The intersection of the $2-dimensional$ plane containing u and w with the unit sphere will be a circle. Now the projections of all the vectors in S on the plane containing u and w will form concentric circles and exactly $2^{(n-2)}$ vectors in S will have same projections.
Thus, distribution of vectors in S is still uniform and this problem reduces to a problem similar to problem-1. Thus the probability that $w.x$ and $u.x$ will have different signs , i.e.

$$\boxed{\mathbb{P}_{x \in S}[sign(w.x) \neq sign(u.x)] = \frac{2\theta}{2\pi} = \frac{\theta}{\pi} = \frac{cos^{-1}(u.w)}{\pi}}$$