Task

- Show average perceptron equivalence
- Show work (Naive vs improved versions)
- Explain

```
public int train(int[] x, int y)
{
   int argmax = decode(x).getInt();

   if (y != argmax)
   {
      double delta = alpha * y;
      update(x, delta);
   }
    return argmax;
 }

private void update(int[] x, double delta)
{
        for (int j : x)
                weight_vector[j] += delta;
}
```

```
public int train(int[] x, int y, int s)
{
   int argmax = train(x, y);

   if (y != argmax)
   {
      double delta = alpha * y * s;
      updateAverage(x, delta);
   }

   return argmax;
}

void updateAverage(int[] x, double delta)
{
   for (int j : x)
     average_vector[j] += delta;
}
```

Quiz 6: Show the above method is equivalent to the naive approach below.

```
for (int i=0; i<weight_vector.length; i++)
        average_vector[i] += weight_vector[i];
```

Alpha is the learning rate constant; the weight_vector is equal to the number of features we have; y is the ground truth.

Delta is what we want to update our hyperplanes with (the move the hyperplane to a more correct position)

We use delta to update our weight vector which holds the current slope/value of each hyperplane

The average vector contains elements that represent the sum of each corresponding previous hyperplane

In the naïve version we are adding the delta to each weight_vector and then adding the weight vector to average. The weight_vector grows larger with each iteration putting more weight on the most recent iteration.

The other version that uses updateAverage works backward in the sense as it puts more weight on the first iterations and decreases as it's going along. Take for example three iterations.

In both cases delta = alpha * y;

	Naïve	other average perceptron
1)	Weight_vector += delta	s=3 so average+=3*D means average = delta *3
a.	Average = delta	
2)	Weight_vector += delta so it equals 2*delta	s=2 so after the += average now equals delta*5
a.	Average = 3* delta	
3)	Weight_vector += delta again so it equals 3*D	s=1 so after the += average now equals delta*6
a.	After Average+= Weight vector ; Average = 6*delta	

In the end both come to the same average of delta*6