1. **Solution to problem 2**
   The weight vector values were largely different as shown below:

old	new
$\theta = -54.0$	$\theta = 513.0$
$w_1 \rightarrow -752.0$	$w_1 \rightarrow -1719.0$
$w_2 \rightarrow 4771.0$	$w_2 \rightarrow 2216.0$
$w_3 \rightarrow 17714.0$	$w_3 \rightarrow 1357.0$
$w_4 \rightarrow 762.0$	$w_4 \rightarrow 1506.0$
$w_5 \rightarrow 6.0$	$w_5 \rightarrow 435.0$
$w_6 \rightarrow 676.0$	$w_6 \rightarrow 1152.0$
$w_7 \rightarrow 3060.0$	$w_7 \rightarrow -300.0$
$w_8 \rightarrow -2004.0$	$w_8 \rightarrow -1928.0$
$w_9 \rightarrow 5459.0$	$w_9 \rightarrow 1462.0$
$w_{10} \rightarrow 9591.2999999989$	$w_{10} \rightarrow 991.9000000000113$
$w_{11} \rightarrow 3832.0$	$w_{11} \rightarrow -775.0$
$w_{12} \rightarrow 14963.0$	$w_{12} \rightarrow 6726.0$
$w_{13} \rightarrow 20912.0$	$w_{13} \rightarrow 11721.0$

The confusion matrix also was also more error prone when testing data set 1. It resulted in the following confusion matrix:

53	1
9	54

The application to data set 3 produced largely different conclusions.

However, the most important feature's weight value stayed relatively the same which shows that it has more influence on the Δw on each iteration of the inputs. This is as expected since because that feature is the most important.