CS 550 -- Machine Learning
Homework #2
Due: 15:40 (class time), March 29, 2016

In this homework, you will implement a multilayer perceptron (MLP) classifier with a single hidden layer. You will run this classifier on the “Thyroid data set”, which you also used for the first homework. Please see the first homework description for the details of this data set.

Followings are stochastic delta rules derived when cross entropy is used as an error function.

\[
\Delta V_{kj} = \eta \left( r_{ik} - z_{ik} \right) y_{ij}
\]

\[
\Delta W_{ji} = \eta \left( \sum_{k=1}^{C} (r_{ik} - z_{ik}) V_{kj} \right) y_{ij} \left( 1 - y_{ij} \right) x_{ti}
\]

If you want to use these delta rules in your implementation, please be sure that you are consistent with the following notation; otherwise you may obtain weird results.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{ti}$</td>
<td>i-th dimension (input unit) of instance t</td>
</tr>
<tr>
<td>$y_{ij}$</td>
<td>j-th hidden unit for instance t</td>
</tr>
<tr>
<td>$r_{tk}$</td>
<td>k-th actual output for instance t</td>
</tr>
<tr>
<td>$z_{tk}$</td>
<td>k-th computed output for instance t</td>
</tr>
<tr>
<td>$C$</td>
<td>number of classes</td>
</tr>
<tr>
<td>$V_{kj}$</td>
<td>weight from hidden unit j to output unit k</td>
</tr>
<tr>
<td>$W_{ji}$</td>
<td>weight from input unit i to hidden unit j</td>
</tr>
</tbody>
</table>

You are also free to derive your own delta rules (for example, using a different error function). If you do so, give the details of what you modify, and write down the rules that you derive after these modifications.

In this homework,
- Give the details of your implementation. List all the parameters that your implementation uses.
- Explain if you use any normalization on the features. If so, give the details of how you normalize the training and test sets. Discuss how this normalization affects your results. Also provide the results obtained with and without normalization.
- Explain if you need to do anything for handling the situation of having unbalanced class distributions. Discuss how the unbalanced class distributions affect your results. Provide the results obtained when you use the data as it is and when you “handle” this situation.

In your experiments, you will run your MLP classifier multiple times, using different number of hidden units. For each run,
(a) Report the accuracies obtained on the training and test sets. Also report the class-based accuracies as well as the confusion matrices for the training and test sets.
(b) Explain how you select the values of the parameters. Particularly, explain what values you consider, how you decide on these values, and how they affect your results.

Here you need to design your own experiments. That is, you need to decide which values of the hidden unit number you will use, and how you select the parameters (as explained in Part b). This will be part of your homework grade.

In this homework, you should submit a report for the details of your implementation and experimental findings. In your report, you SHOULD NOT give the screen shots of your program. In addition to giving your quantitative results (accuracies), you are expected to summarize and interpret your results and findings. You also need to email the source code of your implementation. The subject line of your email should CS 550: HW2. Do not submit the printout of your source code.