Problem Set 4
Due on Friday, November 7, 2014 at 11:55 pm

How to Submit

Create one zip file (.zip) of your code and submit it via `ilearn.ucr.edu`. This zip file should have a single script named `ps4.m` that runs all of the examples. It should also contain the code necessary to run these samples. Finally, it should contain a single pdf file (ps4.pdf) that has your plots.

Do not supply any directories in your zip file. Each file should (in comments) list

- Your name
- Your UCR student ID number
- The date
- The course (CS 229)
- The assignment number (PS 4)

Neural Networks
Summary:

- Implement a procedure to train a neural network.
- Train it on the "class2d" dataset from the previous problem set, supplied in a new file with $y \in \{0, 1\}$, because that is more natural for neural networks.
- Plot the results.

All non-linearities should be signoid functions (including the output). It should be trained with cross-entropy (the same loss function used in class for neural-network binary classification).

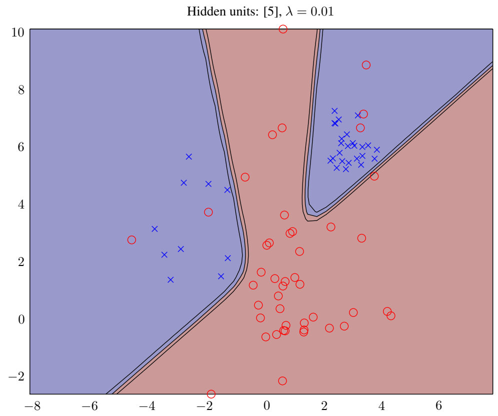
Select a step-size (η) as a function of the iteration number, i, of the form $\frac{\beta_1}{\beta_2+i}$. You'll need to find a suitable value for β_1 and β_2. Pick the weights initially at random from the interval $[-1, 1]$. Be sure to have a "bias" term for each layer. Run until convergence. Make sure you do not stop too soon. If η is very small, then the change might be small, but you still aren't near a minimum (because the gradient is still large).

I would suggest plotting the weights as the procedure learns (and perhaps even the decision surface) to check to see how things are going. You should use gradient descent (batch, not online stochastic gradient descent).

Vary three aspects of your training:

- Number of layers: 2 or 3 (that is, 1 or 2 layers of hidden units)
- Number of hidden units per hidden layer: 1, 5, 20
- Amount of weight-decay (λ in the slides from class): 0.001, 0.01, or 0.1

That totals 18 different experiments. For each experiment the resulting decision surfaces and the training data together. Please give the plot a title that makes is clear which experiment it is. Below is an example plot from a 2-layer, 5 hidden units per layer, result.



Note that the axes are "equal" (same aspect ratio — see the help for `axis`). This will make the results look better. Please do this.

You should look at the help pages for `contourf` (and perhaps `meshgrid`). The following bit of code will color the backgrounds according to their label nicely (it assumes that x and y come from `meshgrid` and that z is a value to be thresholded at 0.5). Before calling this code, plot the data points as blue xs (for y=1) and red circles (for y=0).

```
[c,h] = contourf(x,y,z,[-Inf 0.4 0.5 0.6 Inf]);
ch = get(h,'Children');
for i=1:length(ch)
        if (get(ch(i),'CData')<0.5)
        set(ch(i),'FaceColor',[0.8 0.6 0.6]);
    else
        set(ch(i),'FaceColor',[0.6 0.6 0.8]);
    end;
end;
```

You code will take a while to run to convergence. Collect all of your plots (see `help print` on how to output them as a pdf or similar), with titles (see `help title`) in a single pdf document along. Submit this document and your code according to the instructions above.