Machine Learning 2014: Project 1 - Regression Report

lukasbi@student.ethz.ch
ajenal@student.ethz.ch
harhans@student.ethz.ch

October 31, 2014

Experimental Protocol

TODO: Luki

Suppose that someone wants to reproduce your results. Briefly describe the steps used to obtain the predictions starting from the raw data set downloaded from the project website. Use the following sections to explain your methodology. Feel free to add graphs or screenshots if you think it's necessary. The report should contain a maximum of 2 pages.

1 Tools

We used Matlab for our project.

TODO: Andrin

2 Algorithm

We tried different algorithms, but we stuck with ridge regression in its closed form. In our opinion it performed better than the ordinary least squares as well as the lasso algorithm. So basically we solved the equation $\beta = (X' * X + \lambda * I)^{-1} * X' * y$

3 Features

Did you construct any new features? What feature transforms did you use? TODO: Andrin

4 Parameters

Vital for feature selection and parametritation is the normalization of the feature vectors. The domain of every dimension of the features have to be dimensionless and normalized to ensure a balanced weighting.

The parameters used here were the mean distribution of the training data t.m. mean feature f_{avg} and standard deviation σ_{avg} to f_{avg}. We normalize the training set by subtracting f_{avg} from all points and diveding by σ_{avg}. This is done with the following code:

```
MEAN = mean(training);
STD = std(training);
averagedata = training-repmat(MEAN, size(training,1),1);
normdata = bsxfun(@rdivide, averagedata, STD);
```

TODO: Hans

5 Lessons Learned

What other algorithms, tools or methods did you try out that didn't work well? Why do you think they performed worse than what you used for your final submission?

TODO: HAns