

Statistical Methods
for
Machine Learning
Assignment 3

Kristina Aluzaite (ljg630)
Morten Winther Olsson (hkv518)

18. March 2014

Introduction

III.1 Neural Networks

III.1.1 Neural network implementation

Code is almost done, but haven't finished gradient descent, see "src/Neuron.py".

$$h(a) = \frac{a}{1+|a|} = a \cdot (1+|a|)^{-1}$$

Product and chain rule:

$$h'(a) = (1+|a|)^{-1} - \frac{a}{(1+|a|)^2} \cdot 1 = \frac{1+|a|}{(1+|a|)^2} - \frac{a}{(1+|a|)^2} = \frac{1}{(1+a)^2}$$

III.1.2 Neural network training

See "Neuron_stdout". Currently need to stop gradient descent before overfitting (weights growing to extremes).

III.2 Support Vector Machines

Done, see "src/svm_prob.py".

III.2.1 Data normalization

Done, see "SVM_stdout".

III.2.2 Model selection using grid-search

Done, see "src/svm_prob.py" and "SVM_stdout".

III.2.3 Inspecting the kernel expansion

TODO!

.1 III.2.3.1 Support vectors

.2 III.2.3.2 Scaling behavior