Task 1.1. Supervised Learning: Standard Classifier

Garoe Dorta-Perez
CM50246: Machine Learning and AI

November 11, 2014

1 Introduction

Given pictures from the world and been asked to classify them in several groups, we are faced with a problem of multi-class classification. One of the options would be to create N one-against-all binary classifiers.

$$Pr(w|\mathbf{x}) = Bern_w[\lambda] \tag{1}$$

However a better one involves using a categorical distribution to model our world.

$$Pr(w|\mathbf{x}) = Cat_w[\lambda[\mathbf{x}]] \tag{2}$$

Since we are solving for multi-class classification a logistic sigmoid function as activation will not be valid. Therefore a softmax function is used instead for each activation.

$$a_n = \phi_n^T x \tag{3}$$

$$\lambda_n = softmax_n[a_1, a_2 \cdots a_N] = \frac{exp[a_n]}{\sum_{m=1}^{N} exp[a_m]} = \tag{4}$$

2 Implementation