Foundations of machine learning: Assignment 1

The task set was to develop a model that takes a data set of a 5-dimensional input space and outputs a 1-dimensional continuous value. For training, a dataset of 1000 examples was given. To build the model for this regression problem a Radial Basis Function Network (RBF) was chosen over an Multi-Layer Perceptron (MLP). This was because they are quite robust to noise and they are fast to train due to how few parameters that need to be optimized in contrast to an MLP.

RBF’s have a fixed three-layer architecture, that works on Covers theorem on the separability of patterns (Cover, 1965).

*When casting a complex pattern-classification problem, into a high dimensional space nonlinearly, it is more likely to be linearly separable than in a low-dimensional space.*

The idea behind RBFs is being able classify points by characterising data points by the distances from basis or centre vectors chosen when using training data.

Pre-processing Techniques Used

Before designing the network to approximate the function for the given data, a few pre-processing techniques were first applied.

Normalization was used first, to prevent the network from being ill-conditioned. Normalizing the inputs sets the range of each dimensional input to have roughly the same range. Ultimately allowing for a more stable convergence of weights. The equation used below:

In addition to normalization, whitening was

Second was PCA, to apply PCA it was vital that the number of dimensions to reduce to was such that upwards of 95% of the variance could be retained (Ng, 2017). This can be calculated after finding the eigenvectors and eigenvalues of each dimension, one can do a summation of the retained eigenvalues over all the eigenvalues to establish this variance retention.

Thus, by determining the maximum eigenvalues one can deciding on the appropriate dimensions to keep.

PCA

After

Creative optimisation of the process?

What were the free parameters and how they were chosen

Critical Evaluation

Bibliography

Cover, T. (1965). *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*. 1st ed. [S.l.]: [s.n.], pp.326-334.

Ng, A. (2017). *PCA - Ufldl*. [online] Deeplearning.stanford.edu. Available at: http://deeplearning.stanford.edu/wiki/index.php/PCA [Accessed 23 Mar. 2017].

EXTRA

[Figure 1, (Saedsayad.com, 2017)]

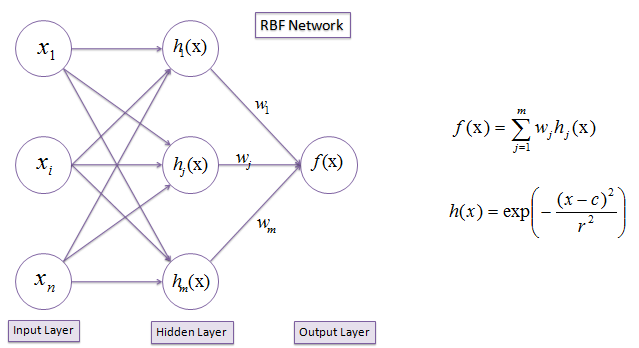


Figure 1

Saedsayad.com. (2017). *Radial Basis Network*. [online] Available at: http://www.saedsayad.com/artificial\_neural\_network\_rbf.htm [Accessed 23 Mar. 2017].