1999 Paper 8 Question 9

Neural Computing

Explain why probability theory plays a central role in neural computation. Discuss how the problem of classification can be expressed in terms of the estimation of a probability distribution. [6 marks]

Explain what is meant by a *likelihood function* and by the concept of *maximum likelihood*. [4 marks]

Consider a neural network regression model which takes a vector $\mathbf{x}$ of input values and produces a single output $y = f(\mathbf{x}, \mathbf{w})$ where $\mathbf{w}$ denotes the vector of all adjustable parameters ("weights") in the network. Suppose that the conditional probability distribution of the target variable t, given an input vector $\mathbf{x}$, is a Gaussian distribution of the form

$$p(t|\mathbf{x}, \mathbf{w}) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{\{t - f(\mathbf{x}, \mathbf{w})\}^2}{2\sigma^2}\right)$$

where σ^2 is the variance parameter.

Given a data set of input vectors $\{\mathbf{x_n}\}$, and corresponding target values $\{t_n\}$, where $n = 1, \ldots, N$, write down an expression for the likelihood function, assuming the data points are independent. Hence show that maximisation of the likelihood (with respect to $\mathbf{w}$) is equivalent to minimisation of a sum-of-squares error function. [10 marks]