Abstract

In this paper we empirically investigate various sizes of training sets with the aim of determining the optimum training set size for generalization ability of an ANN trained on forecasting TCP/IP network traffic trends. We found from both the simulation experiments and literature that the best training set size can be obtained by selecting training samples randomly, between the interval 5×N_W and 10×N_W in number, depending on the difficulty of the problem under consideration.

References

- R. Aamodt, &quot;Using Artificial Neural Networks To Forecast Financial Time Series,&quot; Norwegian university of science and technology, 2010.
- E. Richards, &quot;Generalization in Neural Networks, Experiments in Speech
Recall the importance of training set size for the generalization ability of Artificial Neural Networks (ANNs) in forecasting TCP/IP traffic trends, as highlighted in the work by H. Leung and W. Zue (1991). Their study emphasizes the role of multi-layered networks in the extraction of speech properties. Another significant contribution is the research by A. Weigend, D. Rumelhart, and B. Huberman (1990), illustrating the prediction of the future with a connectionist approach.

S. Haykin's comprehensive foundation on the generalization capability of ANNs (1999) further underscores the relevance of training set size. T. Mitchell's work on machine learning (1997) also acknowledges the pivotal role of training set size in the generalization capabilities of ANNs.


Index Terms

Computer Science

Networks

Keywords

Generalization ability  Artificial Neural Networks and Training set size.