Abstract

This paper introduces a new method which employs the concept of “Orientation Vectors” to train a feed forward neural network. It is shown that this method is suitable for problems where large dimensions are involved and the clusters are characteristically sparse. For such cases, the new method is not NP hard as the problem size increases. We ‘derive’ the present technique by starting from Kolmogrov’s method and then relax some of the stringent conditions. It is shown that for most classification problems three layers are sufficient and the number of processing elements in the first layer depends on the number of clusters in the feature space. This paper explicitly demonstrates that for large dimension space as the number of clusters increase from $N$ to $N+dN$ the number of processing elements in the first layer only increases by $d(\log N)$, and as the number of classes increase, the processing elements increase only proportionately, thus demonstrating that the method is not NP hard with increase in problem size. Many examples have been explicitly solved and it has been demonstrated through them that the method of Orientation Vectors requires much less computational effort than Radial Basis Function methods and other techniques wherein distance computations are required, in fact the present
method increases logarithmically with problem size compared to the Radial Basis Function method and the other methods which depend on distance computations e.g statistical methods where probabilistic distances are calculated. A practical method of applying the concept of Occum’s razor to choose between two architectures which solve the same classification problem has been described. The ramifications of the above findings on the field of Deep Learning have also been briefly investigated and we have found that it directly leads to the existence of certain types of NN architectures which can be used as a “mapping engine”, which has the property of “invertibility”, thus improving the prospect of their deployment for solving problems involving Deep Learning and hierarchical classification. The latter possibility has a lot of future scope in the areas of machine learning and cloud computing.

References

Some Theorems for Feed Forward Neural Networks


29. K Eswaran: A system and method of classification etc. Patents filed IPO No.(a) 1256/CHE July 2006 and (b) 2669/CHE June 2015


Index Terms

Computer Science  Artificial Intelligence

Keywords

Neural Networks, Neural Architecture, Deep Learning, Orientation Vector, Kolmogorov