A Nonlinear Hierarchical Dimension Reduction Approach Based Generalized Distance Measure for Big Data Analytics

R. Krishnan, S. Jagannathan and V.A. Samaranayake

ABSTRACT
Nonlinearities inherently present in big data can introduce complications in pattern classification problems unless explicitly tackled. This paper introduces a distance measure based on a novel hierarchical nonlinear dimension reduction approach to manipulate nonlinear relationships in large scale scenarios. Specifically, a linearizing transformation and a dimension reduction process is performed here to reduce dimension. This process is shown to be computationally favorable in high dimensional situations. Once the dimensions are reduced efficiently, Mahalanobis distance (MD) is used to calculate the final distance value. It is demonstrated that the proposed distance measure can handle various data types effectively. Simulation results are included to study the performance of the proposed distance measure.