Abstract

In this paper, we propose hybrid Random under Sampled Imbalance Big Data (USIBD) framework to extract knowledge from class imbalance big data. A novel under-sampling method for the base learner is also proposed to handle the dynamic class-imbalance problem caused by the gradual evolution of classes in big data. The proposed USIBD knowledge discovery framework is robust and less sensitive to outliers where non-uniform distribution of data is applied. Empirical studies demonstrate the effectiveness of USIBD in various class imbalance big datasets scenarios in comparison to existing methods.

References

5. DOI: 10.4172/2090-4924.1000125.

Index Terms

Computer Science  Information Sciences

Keywords

Classification, Big data, Imbalanced data, Under Sampling, USIBD