Abstract

Supervised machine learning is one of the machine learning task that generates required function from the training data which is labelled. The aim of supervised machine learning is to build or construct a model that makes predictions by using the function inferred from the labelled training data. This paper put a light on how the supervised machine-learning techniques are used to build a predictive model from the dataset of titanic disaster and also a comparative analysis of supervised machine learning methods like Random Forests and Decision Trees are implemented. In this work, with a training dataset containing features or labels like sex, age and class, survivors are predicted from the four test datasets. And from the observations of results a comparative analysis of both supervised machine learning methods namely Decision Trees and Random Forests is implemented.

References

1. A comparative analysis of machine learning methods for classification type decision
A Comparative Analysis of Supervised Machine Learning Methods using Disaster Datasets

problems in healthcare, Emanet et al. Decision Analytics 2014, 1:6, http://www.decisionanalyticsjournal.com/1/1/6

2. Maimon, O., & Rokach, L. (2010). "Data Mining and Knowledge Discovery Handbook". (2nd, Ed.) Springer

3. Kotsiantis, S. B. (2007). "Supervised Machine Learning: A review of classification techniques", Vol 160, No. 3, Frontiers in Artificial Intelligence and Applications

4. Quinlan, J. R. (1987). "Generating production rules from decision trees". Proceedings of the 10th international joint conference on Artificial intelligence, pp. 304-307

5. Tsang, S., Kao, B., Yip, K. Y., Ho, W. -S., & Lee, S. D. (2011, Jan). "Decision Trees for Uncertain Data". IEEE Transactions on Knowledge and Data Engineering, Vol 23, No. 1

6. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (2006, March). "From Data Mining to Knowledge Discovery in Databases". The Knowledge Engineering Review, Vol 21, No. 1, pp. 1-24

7. Breiman, L. 2001a. Random forests. Machine Learning 45:5

Index Terms

Computer Science

Information Sciences

Keywords

Supervised machine learning, Decision Trees, Random Forests.