Naive bayes-guided bat algorithm for feature selection.

Abstract

When the amount of data and information is said to double in every 20 months or so, feature selection has become highly important and beneficial. Further improvements in feature selection will positively affect a wide array of applications in fields such as pattern recognition, machine learning, or signal processing. Bio-inspired method called Bat Algorithm hybridized with a Naive Bayes classifier has been presented in this work. The performance of the proposed feature selection algorithm was investigated using twelve benchmark datasets from different domains and was compared to three other well-known feature selection algorithms. Discussion focused on four perspectives: number of features, classification accuracy, stability, and feature generalization. The results showed that BANB significantly outperformed other algorithms in selecting lower number of features, hence removing irrelevant, redundant, or noisy features while maintaining the classification accuracy. BANB is also proven to be more stable than other methods and is capable of producing more general feature subsets.

Keyword: Naive bayes; Bat algorithm; Feature selection.