

Research on some data classification methods and application in mushroom classification using Weka tool

Souksakhone Inthavong

Abstract: This study investigates several data classification methods and their application to mushroom classification using the Weka tool. Focusing on the challenge of accurately distinguishing between edible and poisonous mushroom species, the research evaluates the performance of key classification algorithms, including Decision Trees (J48), Naïve Bayes, and Support Vector Machines (SVM). Experiments were conducted on the publicly available UCI Mushroom dataset, which contains categorical features describing mushroom physical characteristics. The methodology involved data preprocessing, feature selection, and model training within the Weka environment. Performance metrics such as classification accuracy, precision, recall, and F-measure were used to compare the algorithms. The experimental results demonstrate that the J48 decision tree algorithm achieved the highest classification accuracy, outperforming other methods in terms of both speed and interpretability. The findings confirm that machine learning techniques, when applied through accessible tools like Weka, provide an effective and reliable approach for automated mushroom identification, with significant practical implications for food safety and mycology.

Keywords: data classification, Weka, mushroom classification, data mining