

COMPARATIVE STUDY OF MACHINE LEARNING TECHNIQUES TO CLASSIFY EDIBLE AND POISONOUS MUSHROOMS

ABSTRACT

Mushrooms are actually fungi kingdom, but comes under the category of vegetables for cooking purposes. Mushrooms are of different types, both eatable and poisonous. Each mushrooms have unique look and taste. But nutritional benefits of eating mushrooms may change depending upon the type of mushroom we used. They contain essential nutrients such as proteins, vitamins, minerals, amino acids, antibiotics and antioxidants. Mushrooms are good for our health. But all species of mushrooms are not edible, some of them are poisonous and may cause health problems and leads us to death. So before eating, it should be checked whether it is edible or not. Actual determination and proper identification of mushrooms are the only safety way to consume mushrooms. This paper is a comparative study of different machine learning techniques on mushroom dataset in terms of its performance to identify which algorithm is best for proper identification of edible and poisonous mushrooms.

INTRODUCTION

In the world there are millions of different species of mushrooms are present; both are edible and non-edible being poisonous. For a non-expertise person, it is difficult to identify poisonous and edible mushroom of all the species manually. So a computer based system is required to identify poisonous and non-poisonous mushrooms. If a person unknowingly consume poisonous mushrooms may lead to death. Therefore, proper care must be taken to classify edible mushrooms with that of poisonous ones. In this paper, mushroom classification is done with the help of *weka tool* (Waikato Environment for Knowledge Analysis).

TOOL USED

WEKA (Waikato Environment for Knowledge Analysis), a tool for data pre-processing, classification, regression, clustering, association rules, visualization and is made use of for the prediction. It is a collection of machine learning algorithms for data mining tasks. The algorithms are applied directly to a dataset. This paper aims in identifying edible and poisonous mushrooms based on certain attributes like shape, size, color, etc. Here we consider the mushroom dataset which contains many attributes. Apart from these attributes, we can also include other attributes like cap margins, cap size, stem color, ecology, protein content, toxins, taste etc. Here we apply different machine learning techniques on mushroom dataset and is evaluated using its accuracy, mean absolute error, kappa statistics etc. Based on that evaluation, we can easily identify which algorithm is best for the identification of edible and poisonous mushrooms.