

Introduction to Classification and Regression

Classification and regression are two types of supervised learning methods in machine learning and statistics. In supervised learning, the goal is to learn a mapping function from inputs x to an output y , given a labeled set of input-output pairs.

1.1 Classification Systems

In classification, the output y is a categorical or discrete value. For example, if we are developing a system to predict whether an email is spam or not, y can take two values: "spam" or "not spam". This is an example of a binary classification problem.

Classification problems that have more than two categories are known as multi-class classification problems. For example, predicting the species of an iris flower from a set of measurements of its petals and sepals is a multi-class classification problem, as there are three species of iris flowers.

1.2 Regression Systems

In regression, the output y is a continuous value. For example, if we are developing a system to predict the price of a house given features like its size, location, number of rooms, etc., the output is a continuous number which represents the price.

1.3 Algorithms

There are many algorithms used to solve classification and regression problems, ranging from simple ones like linear regression for regression problems and logistic regression for binary classification problems, to more complex ones like neural networks, which can be used for both classification and regression problems.

1.4 Performance Metrics

Performance of classification and regression models is evaluated using different metrics. For classification, these include accuracy, precision, recall, and F1 score. For regression, common metrics include mean absolute error, mean squared error, and R-squared.

This is a draft chapter from the Kontinua Project. Please see our website (<https://kontinua.org/>) for more details.

Answers to Exercises



INDEX

Classification, 1

Regression, 1