

Chapter 3 - Datasets

1 Introduction

The bike rental dataset is a regression problem where we use day of week, time, user numbers, and weather to predict how many bikes will be rented.

The YouTube spam dataset is a binary classification problem where we classify a comment as spam or not.

The cervical history dataset is a binary classification dataset that uses medical history features to predict whether a patient will get cervical cancer.