MLDM Mini Projekt

Kim Lan Vu, Asha Schwegler

6. Dezember 2022

Inhaltsverzeichnis

1 Introduction

1.1 Goal of the assignment

The goal of this assignment is to predict the type of crime from specific information like time and place. The authors of this paper analysed the data from crime reports from the city of San Francisco provided by the site "Kaggle". Additionaly the data from (Insert data here) was used.

2 Loading the Data

2.1 Getting the data

The San Francisco crime report was loaded into the notebook. At first glance the authors noticed that there are nine columns and 878050 data samples.

3 Processing and cleaning the Data

3.1 Processing and Cleaning the Data

In order to make predictions the authors decided that the PdDistrict would suffice as location information, so the address could be dropped. The description and resolution of the crime and arrest were also unnecessary to make a description of the type of crime. The authors kept the columns dates, category, DayOfWeek, PdDistrict, and the coordinates X and Y. To be able to work with the date the datatype was changed from String to datetime. The authors also assigned for each of the 39 categories a number in order to........

4 Plotting and analysing Data

5 Building the Models