California Housing Price Prediction

The housing price prediction problem

The house price prediction problem involves using data related to houses (such as their characteristics, location, size, etc.) to predict the price at which they could be sold. The primary objective of this problem is to create a model that accurately estimates the price of a house based on its features.

Data description

This dataset contains information about houses within a specific area in California, along with summary statistics derived from the 1990 census data.

This information provides insights into the characteristics of houses within the California, including their sizes, age, the surrounding population, income, and the estimated value of these houses based on the census data from 1990.

 However, it's important to perform data cleaning steps such as handling missing values, or outliers before conducting any analysis to ensure accurate and reliable results.

The columns are as follows, their names are pretty self explanitory:

longitude

latitude

housing_median_age

total_rooms

total_bedrooms

population

households

median_income

median_house_value

ocean_proximity