Predicting Online Courses

Abstract

Udemy is an American educational platform for open online courses that professional adults and students can take. It was founded in May 2010. it Contains many categorical fields, such as development, design, business, technology, and marketing.

June of 2021, the platform has more than 44 million students, 183,000 courses, and 65,000 instructors teaching courses in 75 languages. There have been over 594 million course enrollments.

Design

To predict the best course Rating based on the features that we have using Kaggle .com dataset through Linear Regression.

Data

The dataset that we will be using for our prediction is taken form Kaggle , Below is the some of features we will be using ,Enrollment and Num Subscribers, Stars, Course Price and Discount Price.

data set has 8503 rows and 13 columns.

○ Algorithm

- Linear regression

○ Tools

- Pandas and NumPy packages to manipulate data.
- Matplotlib and seaborn library for visualizing data.
- LinearRegression from the sklearn.linear_model class of the sklearn module, to perform the linear regression then predicted the Rating.
- Lasso from the sklearn.linear_model class of the sklearn module, to perform the Lasso regression then predicted the Rating.
- Ridge from the sklearn.linear_model class of the sklearn module, to perform the Ridge regression then predicted the Rating.
- train_test_split to split the training and test data sets.
- Jupyter notebook to execute the code.

○ Communication

In addition to the slides and the jupyter notebook Code submitted, we will deliver a presentation in the final day.