cs120_lab2_linear_regression_df

databricks

(http://creativecommons.org/licenses/by-nc-nd/4.0/)

Home

Workspace

Recent

Tables

Linear Regression Lab

Clusters This lab covers a common supervised learning pipeline, using a subset of the Million Song Dataset (http://labrosa.ee.columbia.edu/millionsong/) from the UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets/YearPredictionMSD). Our goal is to train a linear regression model to predict the release year of a song given a set of audio features.

Jobs

Search

This lab will cover:

- *Part 1:* Read and parse the initial dataset
  - *Visualization 1:* Features
  - *Visualization 2:* Shifting labels

- *Part 2:* Create and evaluate a baseline model

Send feedback