cs120_lab3_ctr_df

databricks

Home

Workspace

Recent

Tables

Click-Through Rate Prediction Lab

Clusters

This lab covers the steps for creating a click-through rate (CTR) prediction pipeline. You will work with the Criteo Labs (http://labs.criteo.com/) dataset that was used for a recent Kaggle competition (https://www.kaggle.com/c/criteo-display-ad-challenge).

Jobs

This lab will cover:

Search

- *Part 1:* Featurize categorical data using one-hot-encoding (OHE)

- *Part 2:* Construct an OHE dictionary

- *Part 3:* Parse CTR data and generate OHE features
  - *Visualization 1:* Feature frequency

Send feedback