Instacart Market Basket Analysis

Project Idea 1 ::

Instacart, a grocery ordering and delivery app, aims to make it easy to fill your refrigerator and pantry with your personal favorites and staples when you need them. After selecting products through the Instacart app, personal shoppers review your order and do the in-store shopping and delivery for you. Currently they use transactional data to develop models that predict which products a user will buy again, try for the first time, or add to their cart next during a session.The prediction is about which products will be in a user's next order.

Data Resource : kaggle
Data Set   : https://www.kaggle.com/c/instacart-market-basket-analysis/data

CS985/6 Spotify Classification Problem

Project Idea 2 ::

We have a dataset with various attributes from a selection of songs available in Spotify's playlist "All out ..s" starting from the 50s up to the newly ended 10s.This data repo contains 7 datasets (`.csv` files), each representing a Spotify's "All out ..s" type of playlist. Those playlists collect the most popular/iconic songs from the decade. For each song, a set of attributes have been reported in order to perform some data analysis.We need to build a machine learning model that is able to predict the genre of a song.For this problem we actually used/collected data set from here: https://www.kaggle.com/cnic92/spotify-past-decades-songs-50s10s.

Data Resource : kaggle
Data Set   : https://www.kaggle.com/c/cs98xspotifyclassification/data

Expedia Hotel Recommendations

Project Idea 3 ::

Expedia wants to take the proverbial rabbit hole out of hotel search by providing personalized hotel recommendations to their users. Currently, Expedia uses search parameters to adjust their hotel recommendations, but there isn't enough customer specific data to personalize them for each user.Expedia has provided you logs of customer behavior. These include what customers searched for, how they interacted with search results (click/book), whether or not the search result was a travel package. The data in this competition is a random selection from Expedia and is not representative of the overall statistics.For this data we predict the booking outcome (hotel cluster) for a user event, based on their search and other attributes associated with that user event.

Data Resources : kaggle

Data Set : https://www.kaggle.com/c/expedia-hotel-recommendations/data