





Business Problem

To build a recommendation system based on drink and cocktail recipes from aroun d the world.

Machine Learning Problem

ONE OF THE CENTRAL AREAS OF RESEARCH IN ML



Recommendation systems are a <u>very well studied area of Machine Learning</u>, generally relying on unsupervised learning techniques or reinforcement learning

CLUSTERING RECOMMENDATION IS WELL STUDIED



For our specific problem, we implement two different clustering models to generate the predictions: K-Means and Mean-Shift





Machine Learning Components

01 Ingredients ingredients Líst Líst [0,1,1,...,0,0,1] [0,1,1,...,0,0,1] 02 05 05 Kernel PCA Kernel PCA Final Final 2 Principal output: output: Components Cluster 2 Cluster 5 6-dimensional 2-dimensional vector vector Mean Shift K-means [-1.45,2.3] [-1.45,...,2.3] clustering clustering Output PCA Output PCA 3 clusters 6 clusters 04 04 03 03

02

6 Principal

Components

Data Ingestion and Feature Enclineering



DATA SOURCES

We call The CocktailD B API, the output is a JSON object with the information that we requested.

CLOUD STORAGE

We access the extracted data in Big Que ry, where we build a table that stores the info.

DATA PROCESSING

After our initial data load, we manipulate and transform the data in SQL as a follow p step to be able to use it.

The first thing we do i s that we turn the cock tails into an occurrence e matrix for ingredients, o if it does not occur in the cocktail, 1 if it does.

System Design



