

Project Proposal

Fred Syata

Alex Westlund

Topic & Motivation

The main goal of our project will be categorization of cannabis products by chemical composition. Traditional methods include grouping by strain and “genus”, but these methods convey very little information in the current marketplace. Different companies have the same names for different products, or different names for the same product. Within a single company’s product, the same product/strain can have variance in its chemical composition. We are proposing to group individual products by their chemical composition, as opposed to their genus or strain. This will be used in a larger model with the final goal of building a recommender system for a national company, as part of our capstone project.

Broad Plan

We are pursuing Option B, building a new dataset with the added value of categorizations for individual products. The intended audience will initially be ourselves, to use for the upcoming capstone project as well as The Cannabist Company (capstone project sponsor). We are planning on producing a dataset that contains categorization data for all products offered by The Cannabist Company.

Challenge

The specific challenge we are taking on is creating a multi-layer machine learning model (deep learning model) for product categorization. This means we will be implementing more than one model and attempting to take advantage of the phenomena of consensus, where the average of many models provides a more accurate and robust prediction than any individual model.

Data

The Cannabist Company will provide lab analysis (chemical composition data) by Wed 16, 2024. There is a team member on the capstone project that is a current employee of the company and will advocate to ensure that data is provided by the deadline.