CS 418 Final Project

Manasa Kandimalla, Shyam Patel, Carlos Antonio McNulty

December 3, 2019

1. **Problem Selection**

2. **Data Collection**

3. **Data Preparation**

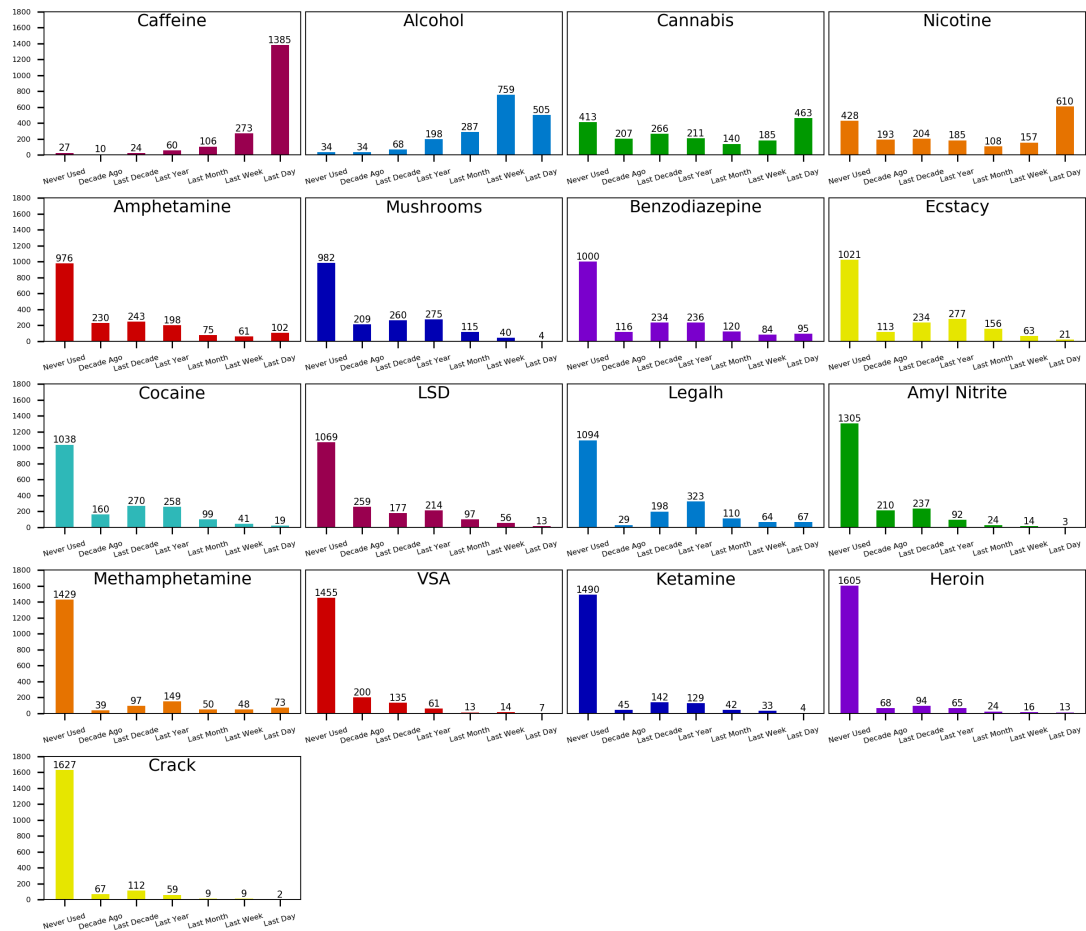4. **Data Exploration**

Figure 1: Drug Usage

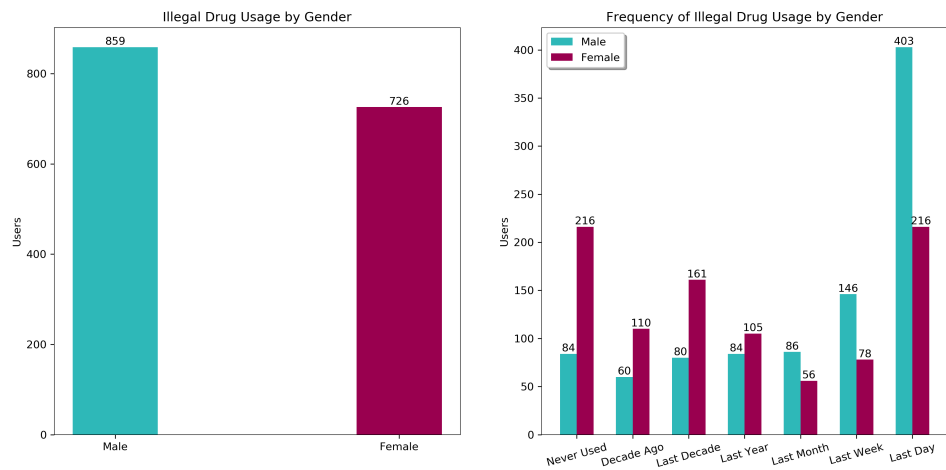Figure 2: Illegal Drug Usage and Frequency by Gender
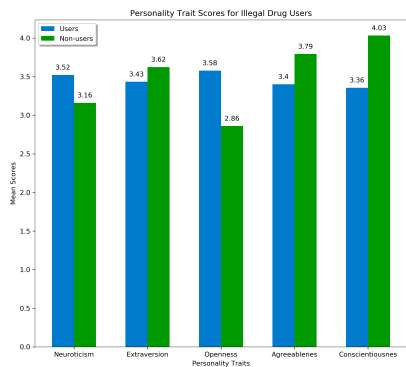

Figure 3: Illegal Drug Usage by Personality Trait
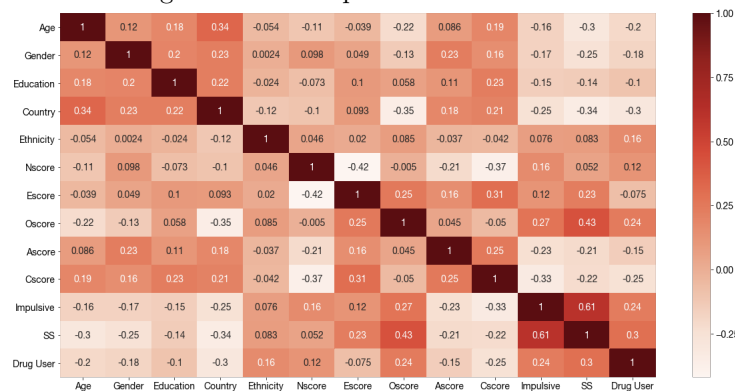

Figure 4: Heatmap of Correlation Matrix
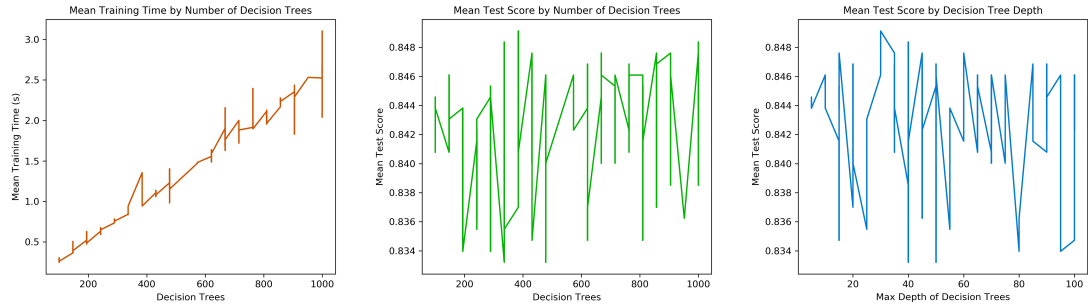
5. Data Modeling

Decision Tree Classifier

Random Forest Classifier

Feature Selection

Grid Search

Figure 5: Hyperparameter Tuning with Grid Search

We performed grid search on 30,240 random forests by varying the number of estimators and the configuration for the base estimators provided to each random forest.

Clustering

6. Results