

# What Factors affect Drug Consumption?

---

Prestige Worldwide  
Alex, Ran, Karolina & Enzo

# OVERVIEW

Used data from the University of California Irvine MCL Repository

- Number of Participants: 1885
- Date of Research: October 2016

## INDEPENDENT VARIABLES

$x$

- Age
- Gender
- Education
- Country
- Ethnicity
- Nscore: neuroticism
- Escore: extraversion
- Oscore: openness
- Ascore: agreeableness
- Cscore: conscientiousness

## DEPENDENT VARIABLES

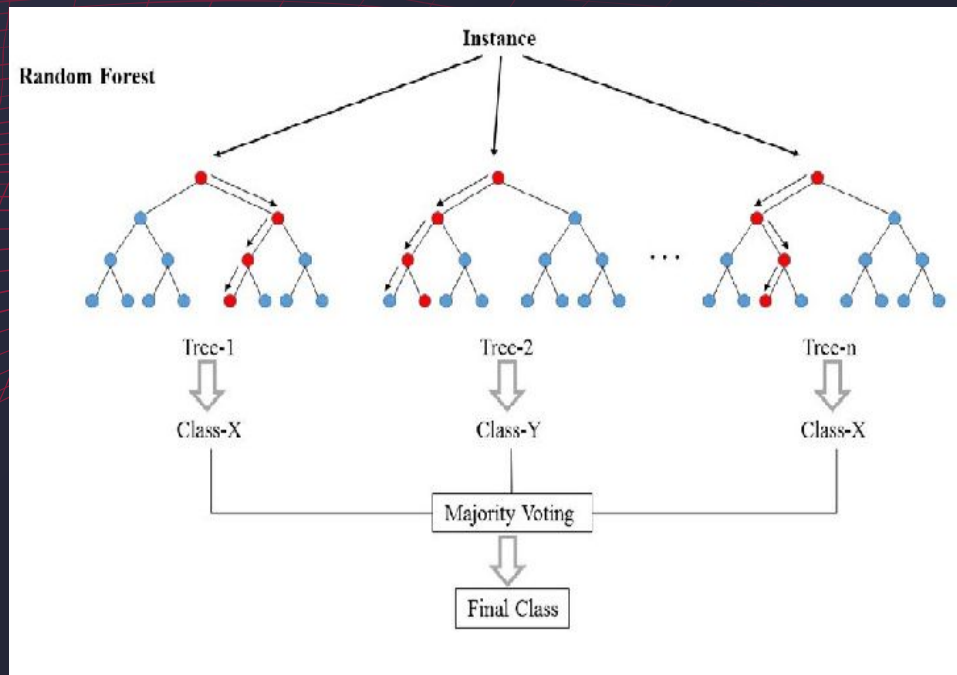
$y$

Level of consumption of:

- Alcohol
- Cannabis
- Cocaine
- Heroin
- LSD

# RANDOM FOREST

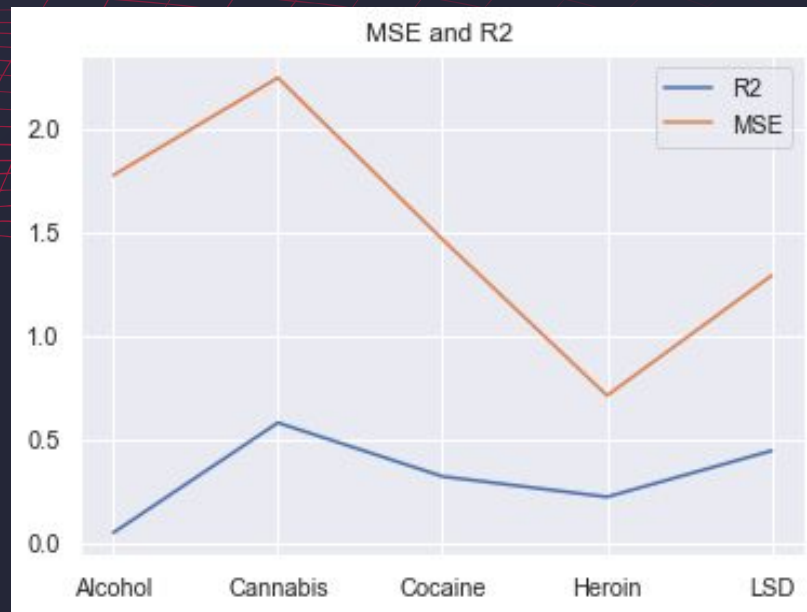
## Overview



How it works: Many decision trees are made and output the class of mean prediction among them

# Random Forest

Results



# K NEAREST NUMBERS

Predicting from Classification

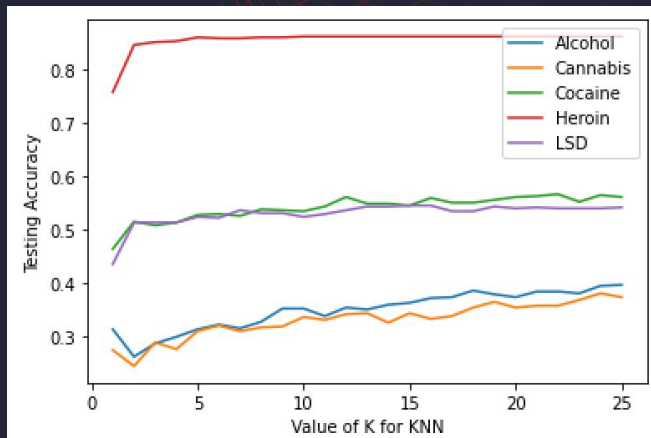
MSE for Alcohol with 5 nearest: 2.027  
MSE for Alcohol with 10 nearest: 1.864

MSE for Cannabis with 5 nearest: 3.507  
MSE for Cannabis with 10 nearest: 3.398

MSE for Cocaine with 5 nearest: 2.057  
MSE for Cocaine with 10 nearest: 1.943

MSE for Heroin with 5 nearest: 0.997  
MSE for Heroin with 10 nearest: 0.952

MSE for LSD with 5 nearest: 1.879  
MSE for LSD with 10 nearest: 1.899



# LOGISTIC REGRESSION

Alcohol

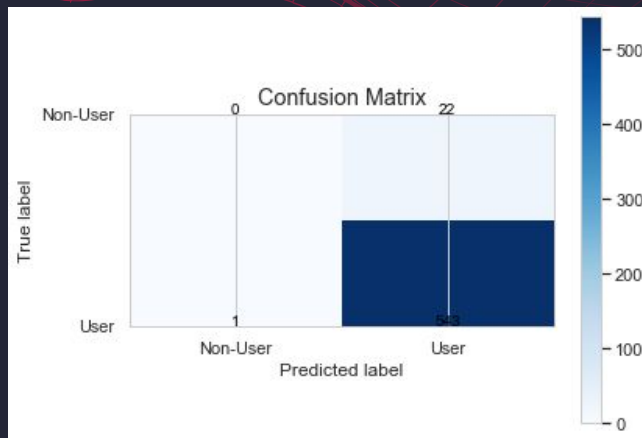
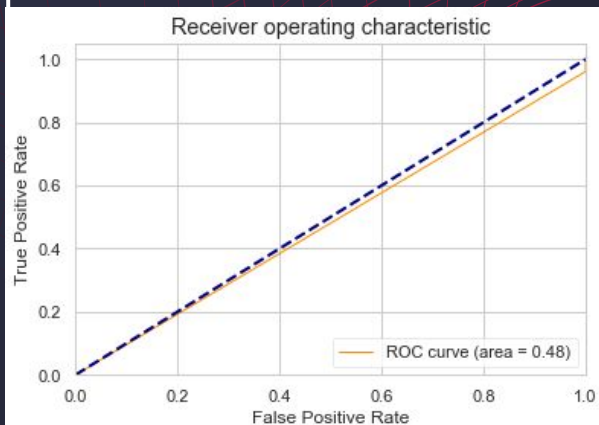
Alcohol

Confusion Matrix:

```
[[ 0 22]
 [ 1 543]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	22
1	0.96	1.00	0.98	544
accuracy			0.96	566
macro avg	0.48	0.50	0.49	566
weighted avg	0.92	0.96	0.94	566





# LOGISTIC REGRESSION

Cannabis

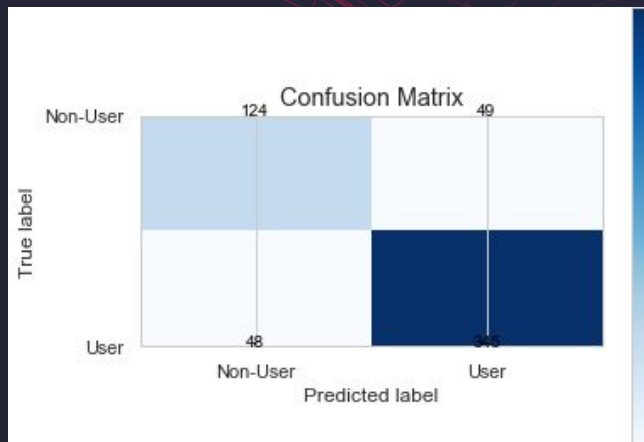
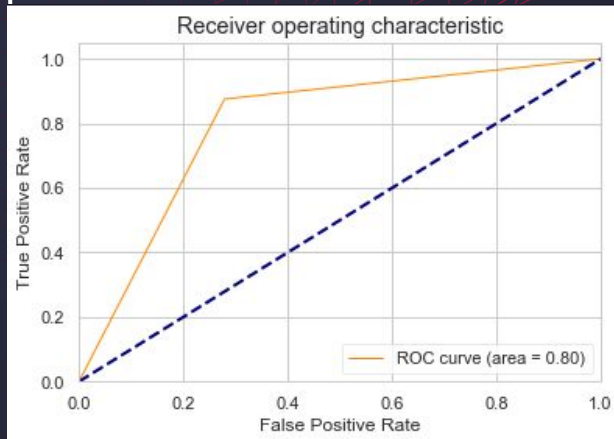
Cannabis

Confusion Matrix:

```
[[124  49]
 [ 48 345]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.72	0.72	0.72	173
1	0.88	0.88	0.88	393
accuracy			0.83	566
macro avg	0.80	0.80	0.80	566
weighted avg	0.83	0.83	0.83	566



# LOGISTIC REGRESSION

Cocaine

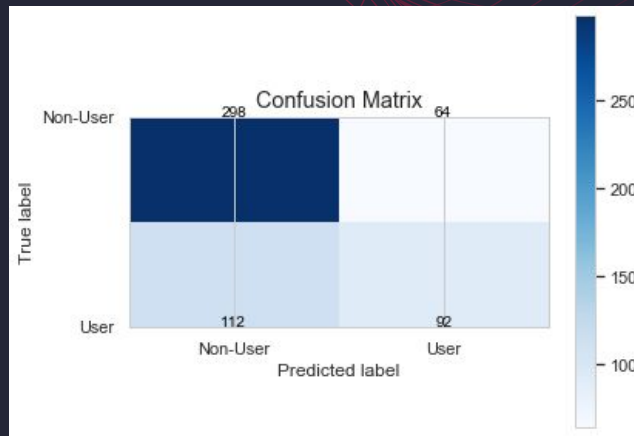
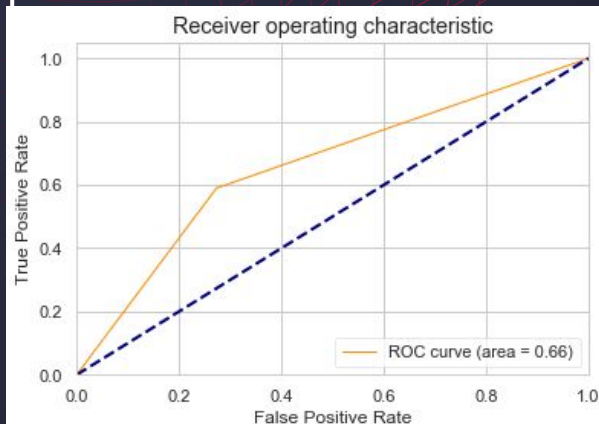
Cocaine

Confusion Matrix:

```
[[298  64]
 [112  92]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.73	0.82	0.77	362
1	0.59	0.45	0.51	204
accuracy			0.69	566
macro avg	0.66	0.64	0.64	566
weighted avg	0.68	0.69	0.68	566





# LOGISTIC REGRESSION

Heroin

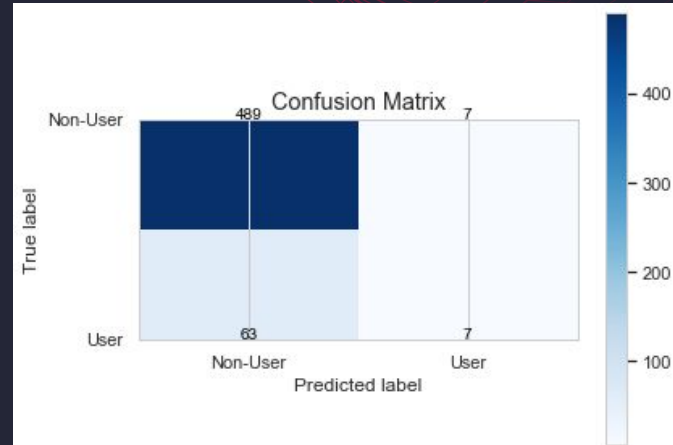
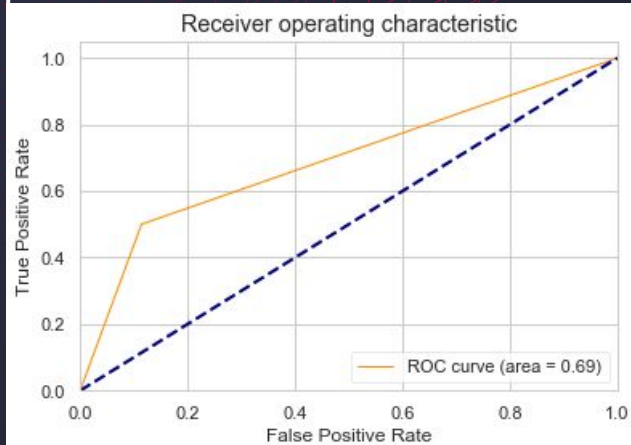
Heroin

Confusion Matrix:

```
[[489  7]
 [ 63  7]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.89	0.99	0.93	496
1	0.50	0.10	0.17	70
accuracy			0.88	566
macro avg	0.69	0.54	0.55	566
weighted avg	0.84	0.88	0.84	566



# LOGISTIC REGRESSION

LSD

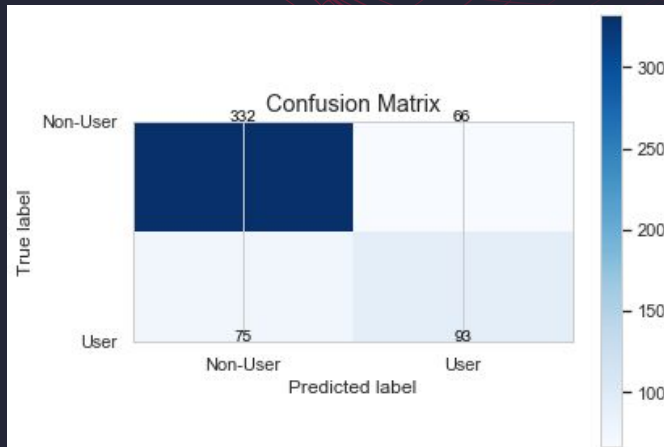
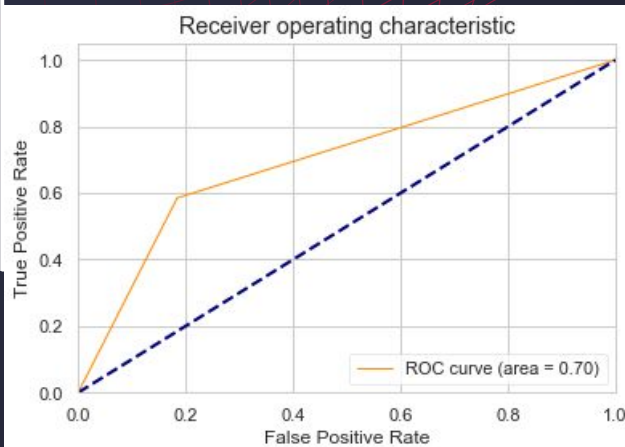
LSD

Confusion Matrix:

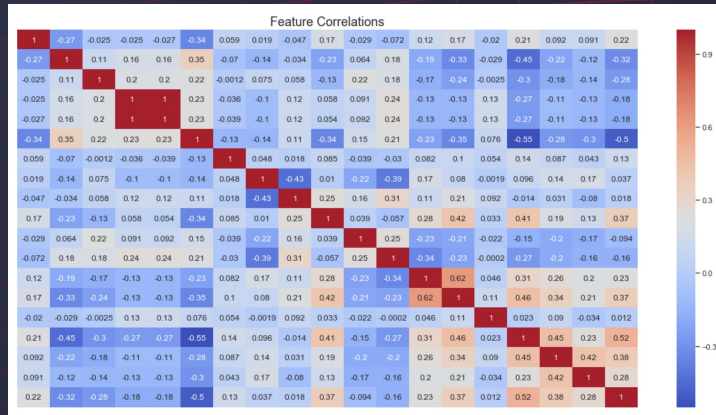
```
[[332  66]
 [ 75  93]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.82	0.83	0.82	398
1	0.58	0.55	0.57	168
accuracy			0.75	566
macro avg	0.70	0.69	0.70	566
weighted avg	0.75	0.75	0.75	566

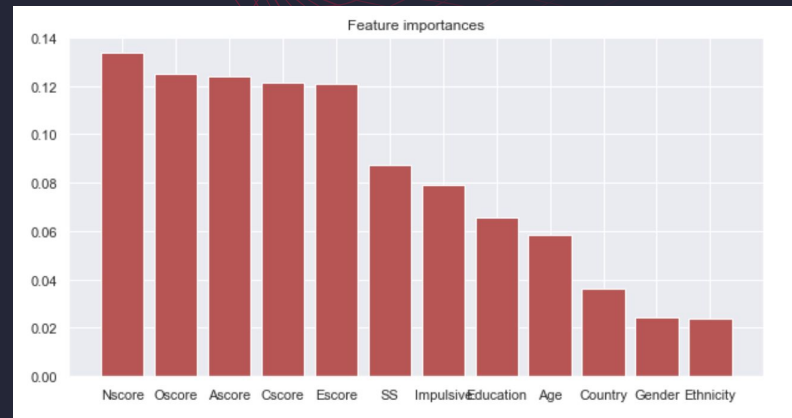


# Evaluation for features



Feature Correlation

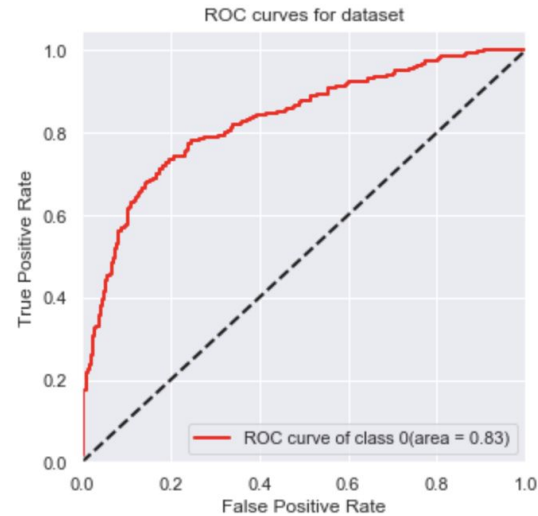
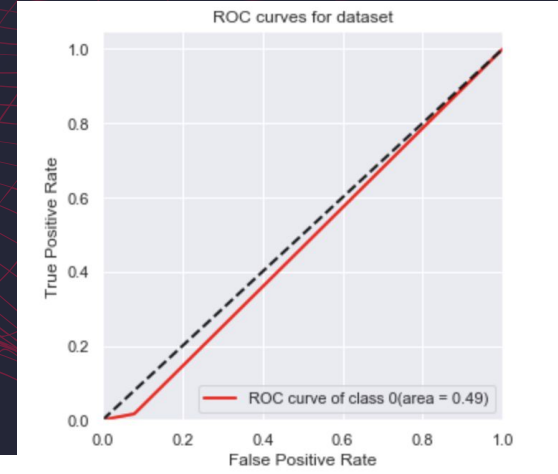
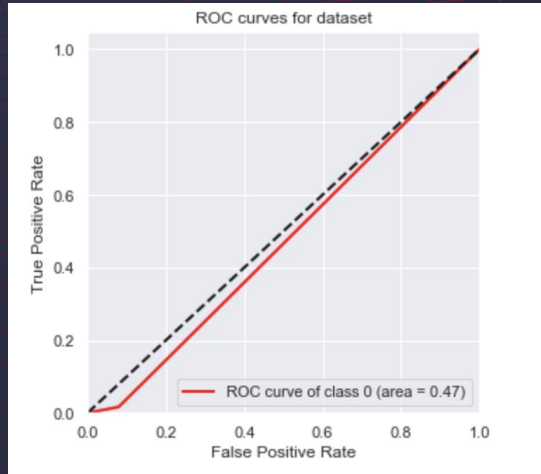
Feature Importance



## Evaluation for Classifiers

	RUC
KNN	0.468527
RandomForest	0.487984
LogisticRegression	0.82805

# Evaluation for Classifiers



**THANKS!**

---