Predicting Cannabis Consumption Using Personality Traits, Demographics, and Legal Drug Usage

Economics 573 Research Paper

Hayden Trayah, Monica Rivell, Anthony Hu

UNC Chapel Hill
16 November 2022

Contents

1 Introduction **2**

2 Literature Review **3**

3 Model Assumptions **4**

4 Data **5**

5 Methods **8**

6 Results **9**

6.1 Lasso Regression **9**

6.2 Logistic Regression 12

6.3 Boosting 14

6.4 Decision Tree and Random Forest 16

6.5 k-Nearest Neighbors 18

6.6 Linear Discriminant Analysis 19

6.7 Support Vector Machines 21

7 Discussion **22**

8 References **25**

1 Introduction

Personality, and its predictive power, has been a topic of interest studied dating back thousands of years to the Ancient Greek and Roman Times and further researched since the official emergence of the science of Psychology, with groundbreaking findings from famous names such as Sigmund Freud and Ivan Pavlov. Every individual has their own unique personality fingerprint and it can be used to help predict behavior in particular scenarios.

One scenario in which this can be applied is evaluating an individual's risk of consuming and even abusing drugs. The significance of studying drug usage cannot be overestimated. The use of drugs can introduce a myriad of risks for one's self and their environment. For an individual, this may mean an increase in health risks, including higher likelihood of premature mortality and morbidity. Additionally, drug usage can constitute a serious problem among society and pose stark economic costs. The last available estimate performed in 2007 revealed an economic cost estimated at an annual $193 billion due to drug abuse in the United States. This value includes costs due to lost productivity and its related labor participation costs, turnover/absenteeism, participation in drug abuse treatment, and premature deaths. This figure also includes costs associated with healthcare and criminal justice costs, primarily criminal investigations, prosecution, incarceration, and victim costs (National Archives and Records Administration).

As we can see, drug usage is a topic worth taking a look into. In the following study, we utilize a dataset containing information on 1,885 respondents and their usage of 18 psychoactive substances, one of which is cannabis. Using machine learning techniques, including logistic regression, lasso regression, decision trees and random forests, K-nearest neighbors, boosting, and linear discriminant analysis, we examine the impact that personality characteristics, demographics, and one's usage of legal substances have to predict cannabis use. Based on a range of previous findings, we theorize that these factors hold predictive power in determining one's usage of cannabis. The implications of proving this hypothesis could mean greater hope in developing risk assessing technology to be used in real-life situations for the benefit and safety of individuals at risk for drug usage and thus the betterment of society as a whole, including reduced economic cost implications.

2 Literature review

Personality traits and their predictive power is a useful tool that has been utilized in a myriad of research. (Ferwerda & Tkalčič, 2020) mentions a few areas where a clear correlation can be found using personality traits and can be used to improve or benefit individuals' performance or experience. Just a few examples include using personality traits to personalize music recommendations, tailor online educational tools to better suit an individual's learning style, and creating a treatment plan specific to a person and their unique personality. One study specifically looked at the correlation between personality traits, learning approaches, and academic performance of students (Al-Omari, 2015).

There are many approaches to studying personality. For example, a study used the Zuckerman-Kuhlman Personality Questionnaire (Feldman et al, 2011). Another study simply used an abbreviated personality research form that looked at attributes such as achievement, cognitive structure, affiliation, autonomy, exhibition, and impulsivity (Labouvie & McGee, 1986). While there are many options, the most commonly used method among our research is the Neuroticism Extraversion Openness Five Factor Inventory, commonly known as NEO-FFI (Dale et al, 2020).

A similar study to ours was conducted to identify predictors of response to intravenous ketamine to help identify the risks and benefits of ketamine therapy.

Using the NEO-FFI on 125 participants, they tested the degree of neuroticism in predicting responses to ketamine and which personality factors had an impact on the response. They examined descriptive statistics and t-scores using R and utilized Kruskal Wallis for continuous variables and Fisher for categorical. Using separate logistic regression models for each of the 5 personalities, they found that openness to experience was the only factor that significantly predicted outcomes (Dale et al, 2020).

Using the same dataset as our research, another study took an opposite approach and looked at the predictability of people's personality traits based on their drug consumption profile to help create personalized treatment programs. Using Weka, they used ZeroR as their baseline classifier and for each classifier reported RMSE which indicated prediction performance of the personality traits. They utilized random forest, M5Rules, radial basis function, using 10 fold cross validation with 10 iterations, and which all outperformed the baseline model. They found that prediction of openness to experience gained the highest improvement over the baseline model and agreeableness also gained a little as well (Ferwerda & Tkalčič, 2020).

3 Model Assumptions

We theorize that personality characteristics, demographics, and one's usage of legal psychoactive substances have predictive power in determining one's usage of cannabis. For example, we predict that individuals who are more open to experiences, more impulsive, and have fewer years of education are more likely to use cannabis. Cannabis is a federally banned substance in the United States that carries stigma in certain societies and social groups around the world, meaning that users who are open to experiences, more impulsive, and more connected in social groups where drug use is common and accepted, are theoretically more likely to use a restricted substance despite social, institutional, and health risks.

However, cannabis is very widely used around the world despite its illegality in a majority of countries. Due to the drug's restricted status, in addition to cultural perceptions surrounding cannabis in Western countries many of our observations originate, we believe that personality traits, demographics, and legal drug usage are effective predictors for cannabis use. We use binary predictive models in this analysis using the factors outlined in the data section to predict an individual's use of cannabis in the last 30 days based on the assumptions that we made.

4 Data

The data was collected in 2015 by Elaine Fehrman, Vincent Egan, and Evgeny M. Mirkes. The data consists of 1885 respondents, with 12 known attributes regarding the NEO-FFI-R, BIS-11, and ImpSS personality measurements (neuroticism, extraversion, openness to experiences, agreeableness, conscientiousness, impulsiveness, and sensation seeking), level of education, age, gender, country of residence, and ethnicity. In addition, the data includes the use of 19 drugs – 18 real drugs and Semeron, a fake drug used to filter exaggerated observations – with seven possible responses corresponding to their last use of each respective drug: never used, used over a decade ago, used in the past decade, used in last year, used in past month, used in last week, and used in last day. In total, each respondent has 32 corresponding values, including one value for ID.

We made certain statistical assumptions regarding the nature of the data. We assumed that the data collected was truthful and accurate. Since the data was collected through a survey, there could be a variety of reasons for inaccuracy in the responses, such as survey and response bias, mistrust of the surveyors, and exaggeration of the responses. Most of these could not be controlled, except for the exaggeration of the responses; we deleted any observation that stated usage of the false drug Semeron, from the dataset. Otherwise, we assumed that the observations were truthful in reflecting the drug use of the individual in question.

To answer our main question, we decided to transform the data to orient our models towards binary predictions. First, we transformed the seven categories for each drug's usage into a binary value based on whether the last use was within the previous month. We chose one month, because that is the approximate timeframe in which marijuana stays in a person's body after usage. Next, we dropped all illicit drugs, and dropped legal drugs that we deemed inappropriate for this application – chocolate and legal highs. In total, we dropped 14 drug variables, leaving only 4 drugs: alcohol, caffeine, nicotine, and cannabis. Finally, we transformed the demographic variables into dummy matrices. For each categorical variable with *n* distinct values, we created *n-1* binary variables corresponding to each category. Finally, we split the data into a training and testing set with 1314 observations in the training set and 562 observations in the testing set.

Below, we have included tables to illustrate the distribution of the responses for drug variables. From left to right, we have the binary frequencies for the alcohol, caffeine, cannabis, and nicotine variables respectively.

0	1	0	1	0	1	0	1
332	1544	120	1756	1095	781	1007	869

We also included tables regarding the frequencies of dummy matrices among education, gender, and country of origin. From top to bottom, we have the table showing frequencies of the age, gender, education, ethnicity, and origin of country variables, respectively.

18-24	25-34	35-44	45-54	55-64	65
643	481	355	294	93	18

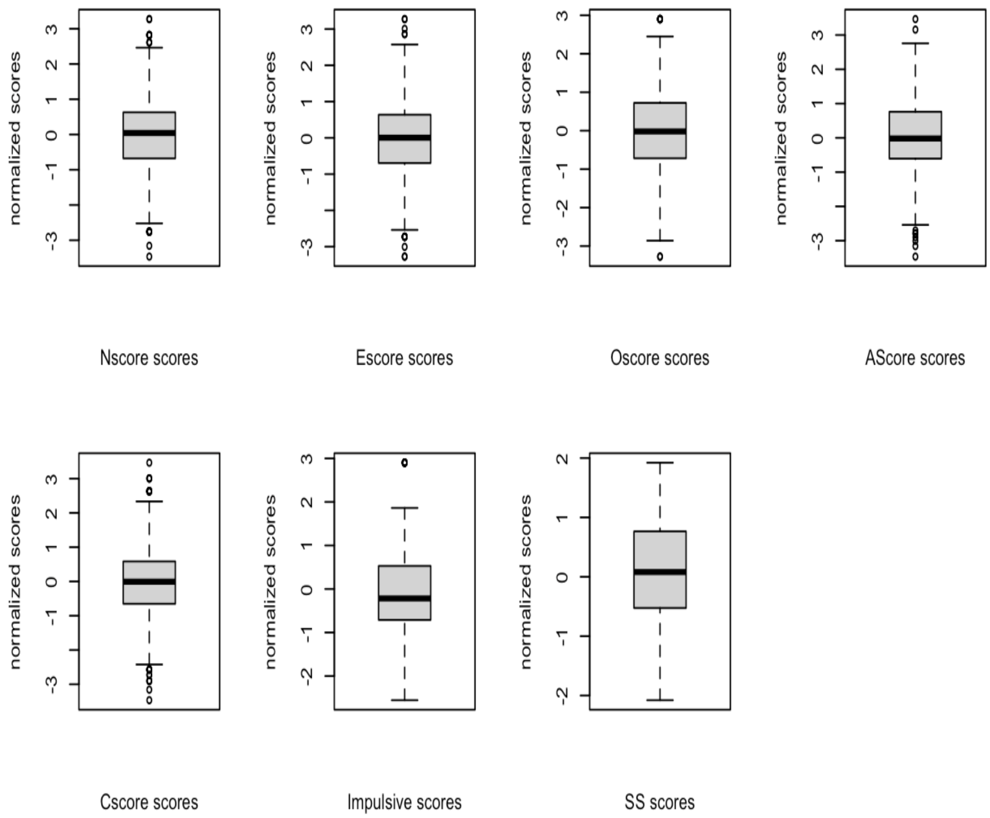
F	M
941	943

Doctorate degree	89
Left school at 16 years	99
Left school at 17 years	30
Left school at 18 years	100
Left school before 16 years	28
Masters degree	283
Professional certificate/ diploma	269
Some college or university, no certificate or degree	506
University degree	480

Australia	Canada	New Zealand	Other	Republic of Ireland
54	87	5	118	20
UK	USA			
1043	557			

Asian	Black	Mixed-Black/Asian	Mixed-White/Asian	Mixed-White/Black
26	33	3	19	20
Other	White			
63	1720			

Below, we illustrate the box and whisker plots for the personality variables, based on the normalized scores. Each personality variable has been denoted in the title label underneath the plot.



5 Methods

Our team randomly split the 1876 usable observations into a rounded 70/30 split of training and test data, resulting in a distribution of 1314 training observations and 562 testing observations. We started our analysis with lasso regression to gain an understanding of the most significant variables in the data set. We then compared our lasso results to a logistic regression using the full set of predictors. We then developed predictive models using boosting, k-nearest neighbors (kNN), support vector machines (SVM), decision trees, random forests, and linear discriminant analysis (LDA). We analyzed the misclassification error of each model by comparing the predicted class to the true class of each test observation. We used k-fold cross-validation to calculate the error for our lasso regression.

6 Results

6.1 Lasso Regression

We used lasso regression first to investigate which variables were significant in the data set. Lasso regression uses a similar approach to standard linear regression by minimizing deviance to calculate coefficients. However, unlike linear regression, lasso regression incorporates a “punishment” term, in addition to the residual sum of squares, in the minimization constraint that calculates the coefficients. The shrinkage parameter (lambda) scales the punishment term which tunes the extent to which the model shrinks coefficients in order to eliminate non-predictive variables. Lasso is a dimension reduction technique and we used it to identify the least powerful predictors in our data set to gain an understanding of which factors explain cannabis use the most. First, we used cross-validation to identify the ideal error-reducing shrinkage parameter for use in the lasso regression. Figure 1 illustrates the values of misclassification error along a range of shrinkage parameters. We identified 0.006588544 as the optimal error-minimizing lambda.

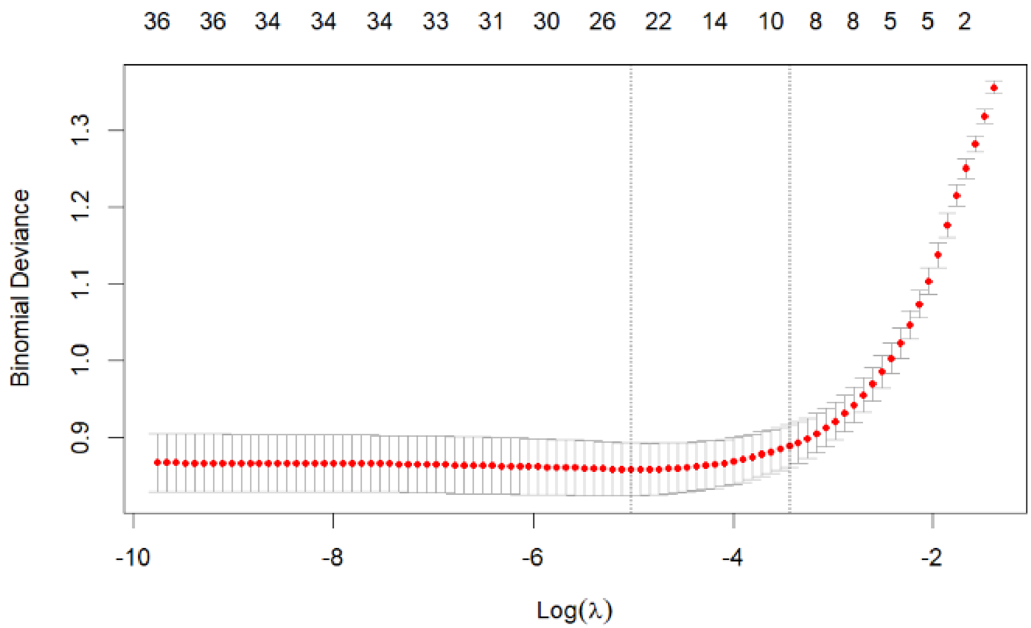


Figure 1: Shrinkage parameters and their corresponding binomial deviance

Our team ran the lasso regression using the ideal shrinkage parameter. We then engaged in further hyperparameter tuning by calculating the misclassification errors among classification boundaries ranging from 5% to 95%, meaning test observations were classified as 0 or 1 by determining if their prediction (a percentage) was above or below the classification boundary (classifying as a 1 or 0, respectively). This process was used to identify the error-minimizing classification boundary to compensate for an unequal distribution of classes in the response variable, which led to imbalanced classification, and to reduce our error rate. Figure 2 shows the misclassification rates along a range of classification boundaries from 5% to 95%. The error-reducing boundary was identified as 0.5, or 50%, leading to a prediction accuracy of 82.74%. Figure 3 illustrates the coefficient shrinkage of our lasso model as lambda increases, with coefficients that shrink to zero as the most significant. At the error-minimizing shrinkage value, the lasso model eliminated the variables: Escore, AScore, Impulsive, Caff, 35-44, Left school at 17 years, Professional certificate/diploma, Canada, Republic of Ireland, Black, Mixed-White/Black, and White. Figure 4 shows the ROC curve of our lasso model which can be interpreted as the true positive rate on the y-axis and the true negative rate on the x-axis.

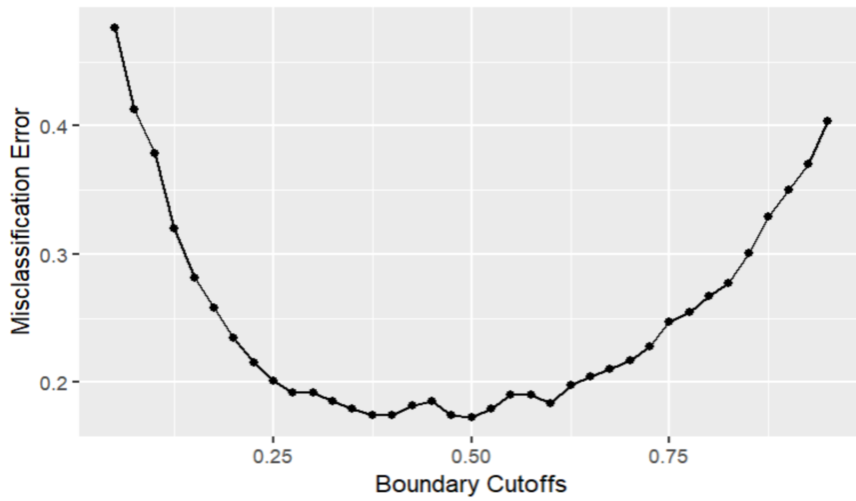


Figure 2: Lasso misclassification error along a range of classification boundaries

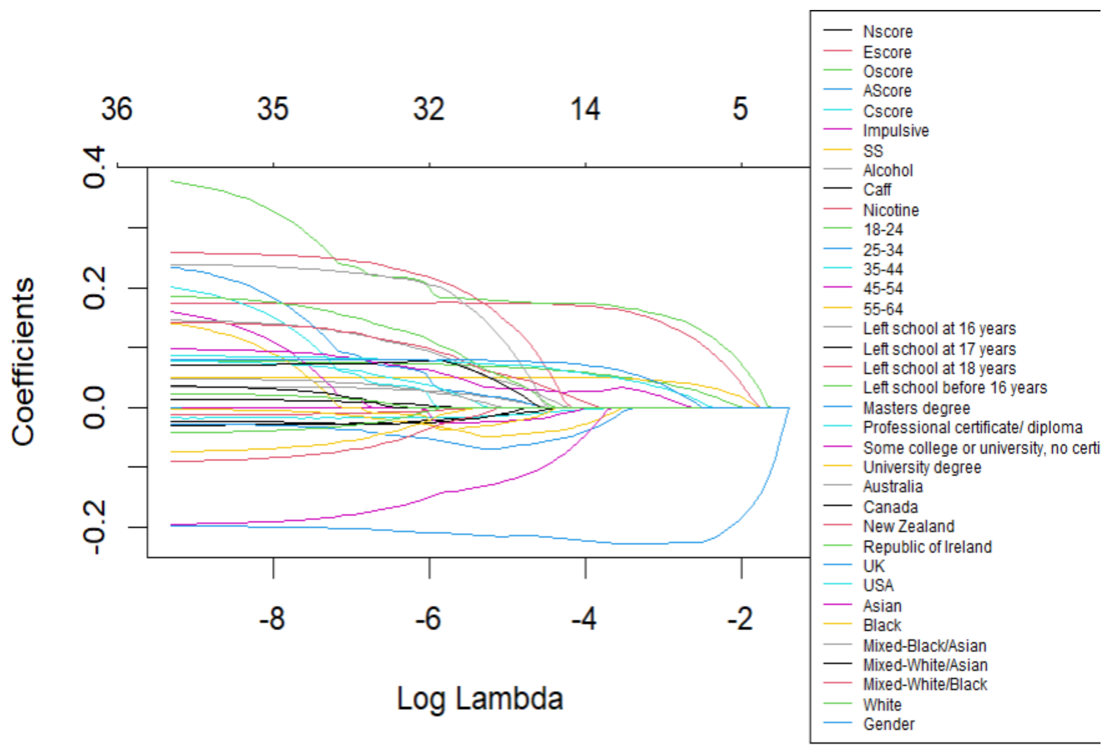


Figure 3: Lasso coefficient shrinkage along increasing shrinkage parameters

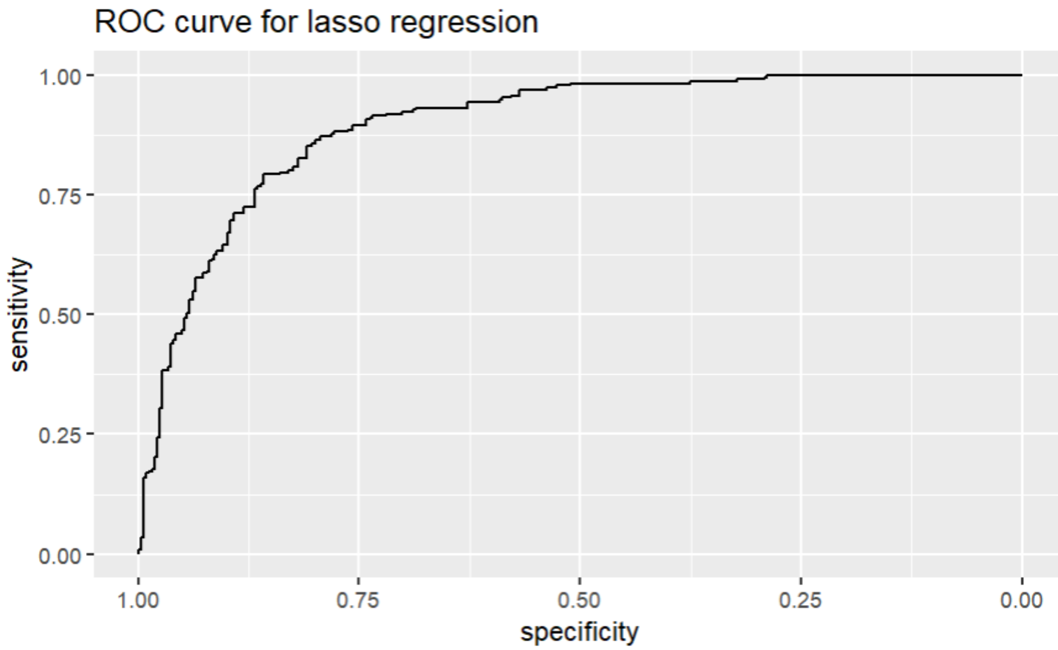


Figure 4: ROC curve for lasso regression: y-axis = TPR, x-axis = TNR. Greater area under the curve (AUC) indicates a more accurate model.

6.2 Logistic Regression

After completing lasso regression, we used logistic regression incorporating the full set of predictors to determine if variable selection produced a more accurate model. Logistic regression applies the logit link to the standard linear regression formula to produce a binary prediction model. We performed the same classification boundary optimization process that was used for lasso to determine the error-minimizing classification boundary between 5% and 95%. However, logistic regression does not incorporate a shrinkage parameter in the coefficient calculation like lasso, so boundary optimization was our only hyperparameter tuning for this model. The optimal boundary cutoff for our logistic regression was 35% which produced a misclassification error of 0.1761566, or a success rate of 82.38%, very slightly below lasso. Figure 5 shows misclassification error along boundary cutoffs from 5% to 95%. Figure 6 shows the ROC curve for logistic regression. Figure 7 shows the confusion matrix for our logistic regression, which produced a false positive rate of 5.69395% and a false negative rate of 11.92171%.

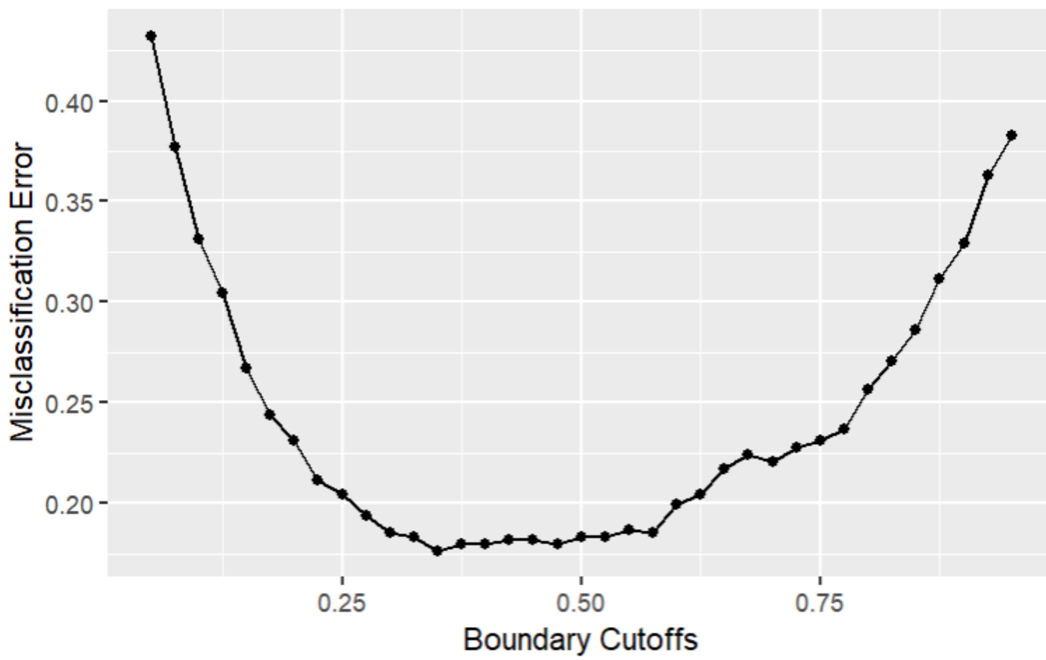


Figure 5: Logistic misclassification error along a range of classification boundaries

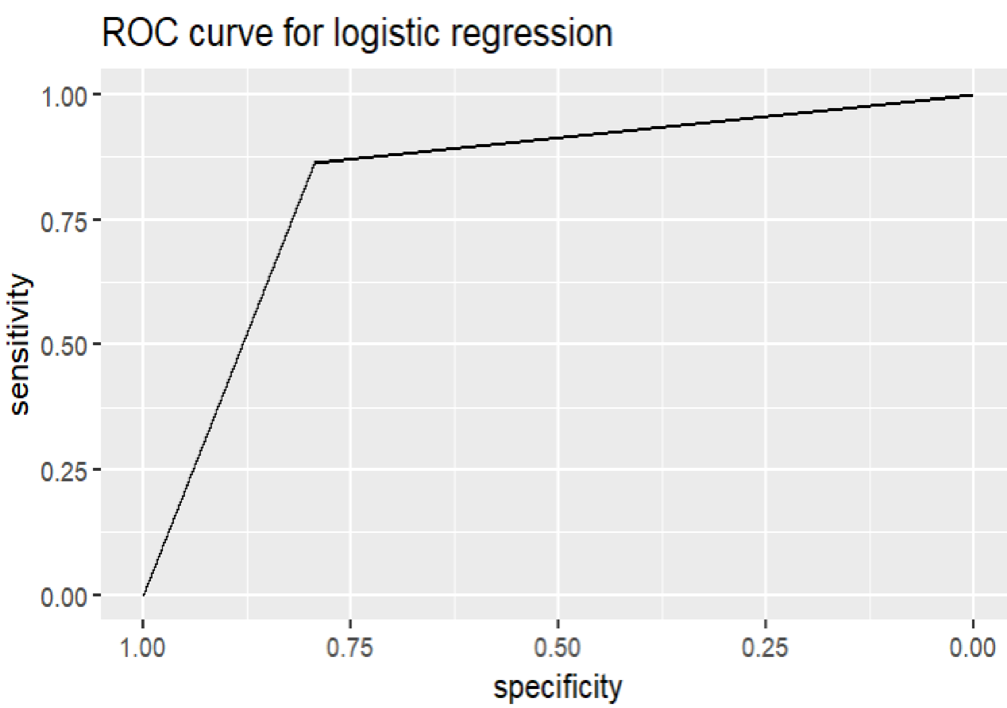


Figure 6: ROC curve for lasso regression: y-axis = TPR, x-axis = TNR

```
Confusion Matrix and Statistics

          Reference
Prediction   0   1
         0 258  67
         1  32 205

               Accuracy : 0.8238
                 95% CI : (0.7898, 0.8545)
    No Information Rate : 0.516
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.6459

 Mcnemar's Test P-Value : 0.0006329

            Sensitivity : 0.8897
            Specificity : 0.7537
         Pos Pred Value : 0.7938
         Neg Pred Value : 0.8650
             Prevalence : 0.5160
         Detection Rate : 0.4591
   Detection Prevalence : 0.5783
      Balanced Accuracy : 0.8217
```

Figure 7: Confusion matrix for logistic regression

6.3 Boosting

Boosting is a prediction technique that involves sampling the data set with replacement to create a number of new data sets, a process known as bagging, and producing a random forest with each tree being generated from a unique bootstrapped data set. In boosting, each tree is generated sequentially to continuously explain parts of the model that were not explained by the previous trees. In the pursuit of reducing misclassification error, we used hyperparameter tuning to optimize the shrinkage parameter used to generate trees. This shrinking parameter is best understood as the learning rate of the model, or the step-size reduction of each tree. Figure 8 shows the misclassification error of the boosting models along a range of shrinkage parameters from 0.01 to 0.5. The error-minimizing shrinkage parameter was .4804, but visual analysis of the graph shows that this value was an outlier among an upward trend in the higher range of the shrinkage parameters. As a result, we used 0.23 as our shrinkage parameter as this value is the approximate bottom of the quadratic misclassification error regression line. Using the shrinkage parameter of 0.23, we performed the classification boundary optimization method that we used in lasso and logistic regression. We found the optimal boosting classification boundary to be 0.425, producing an overall accuracy of 82.56%, which is just in-between lasso and logistic regression. Figure 8 shows the misclassification error along different shrinkage parameters. Figure 9 shows misclassification error along different classification boundaries using the optimal shrinkage parameter of 0.23.

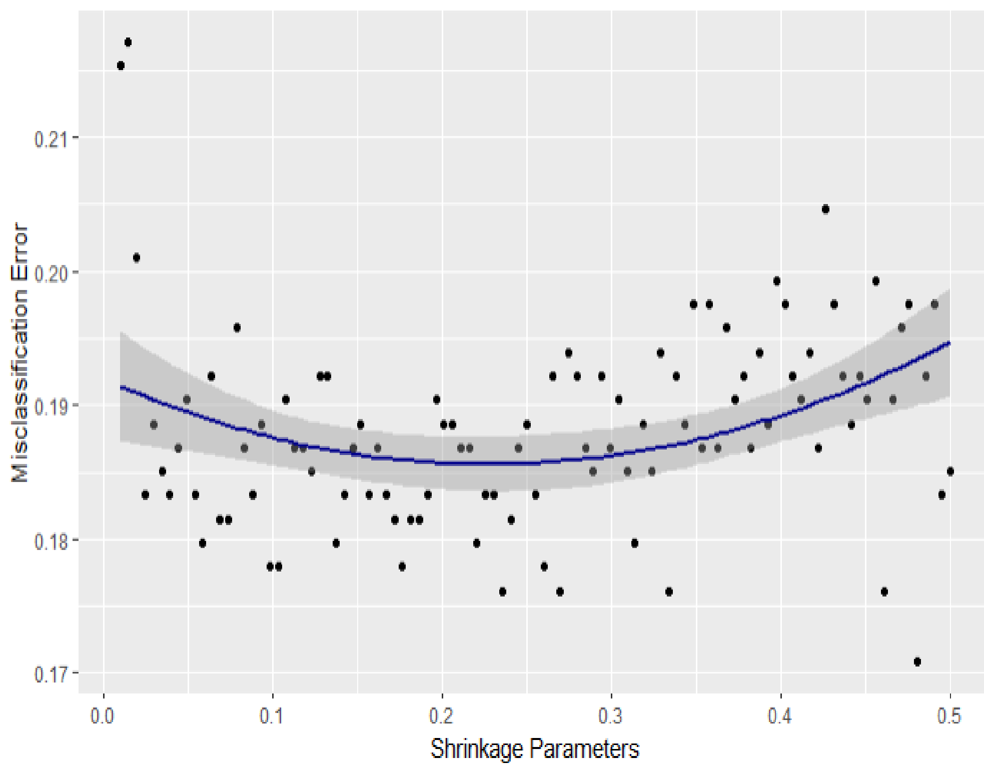


Figure 8: Misclassification error using shrinkage parameters ranging between 0.01 and 0.5

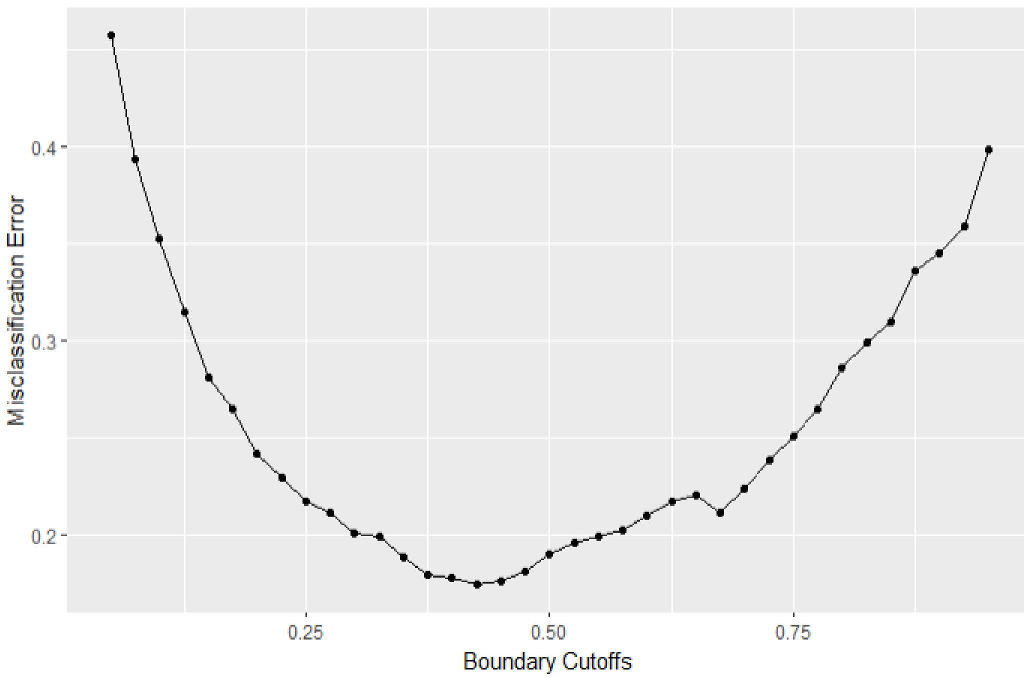


Figure 9: Boosting misclassification error along classification boundaries from 5% to 95%

6.4 Decision Tree and Random Forest

Decision tree is a classification method, where observations are classified using conditional logic based on the variables in the dataset. For each node in a tree, if-else logic is applied to test observational data to be further branched out to other nodes, or to a terminal node, resulting in a classification. For the decision tree method, we found the tree that minimizes the training classification error, and reached a testing classification error of 0.1992883, or a success rate of 80.07%. The four variables used were the UK, nicotine, gender, and 18-24 age binary variables, as well as the personality score for openness to new experience. Figure 10 shows the error-minimizing decision tree we discovered, and Figure 11 denotes the confusion matrix and testing error for the decision tree when applied to the testing set.

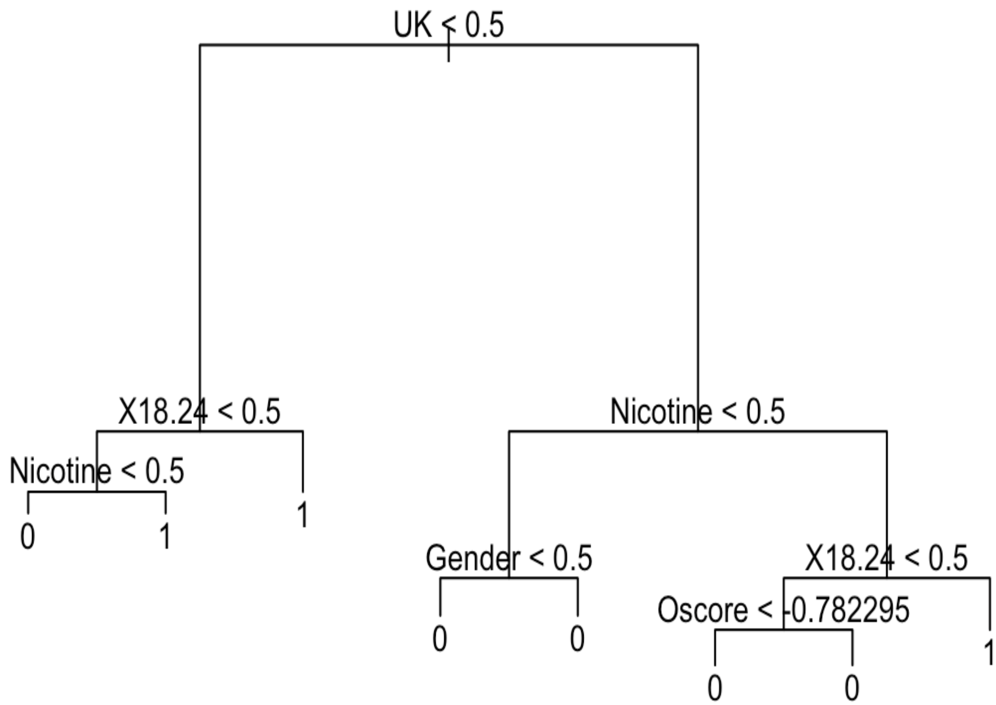


Figure 10: Optimal decision tree

```
  tree_predict
      0   1
  0 271  54
  1  58 179
[1] 0.1992883
```

Figure 11: Confusion matrix for the testing set. The columns of the table represent the predicted classification, and the rows of the table represent the actual classification.

Random forest is an extension of the decision trees, where multiple decision trees are generated, and for each observation the classification depends on the class selected by the most trees. Usually, random forests are an improvement over decision trees, because of the tendency for decision trees to overfit the training data. After running the random forest, we obtain a testing error of .1779359, which is slightly better than the decision tree, as shown in the confusion matrix in figure 12.

```
rf.predict   0   1
         0 272  47
         1  53 190
```

Figure 12: Confusion matrix for the testing set. The columns of the table represent the predicted classifications, and the rows of the table represent the actual classification.

6.5 k-Nearest Neighbors (kNN)

k-Nearest Neighbors is a prediction technique involving the comparison of the test observations to a set number of the “k” nearest observations in the training data. kNN looks at the training observations most similar to each test observation and calculates the proportion of classes among those k similar training observations. The majority classification of the k observations is then chosen as the predicted class for the test observation. We used hyperparameter optimization to find the optimal number of “k” neighbors to minimize misclassification error. We calculated the misclassification error using k values between 1 and 50 and found the error-minimizing number of neighbors to be 22, which resulted in a stand-out accuracy of 90.74%, significantly higher than any other model that we tested. Figure 13 illustrates the misclassification error associated with neighbor values from 1 to 50.

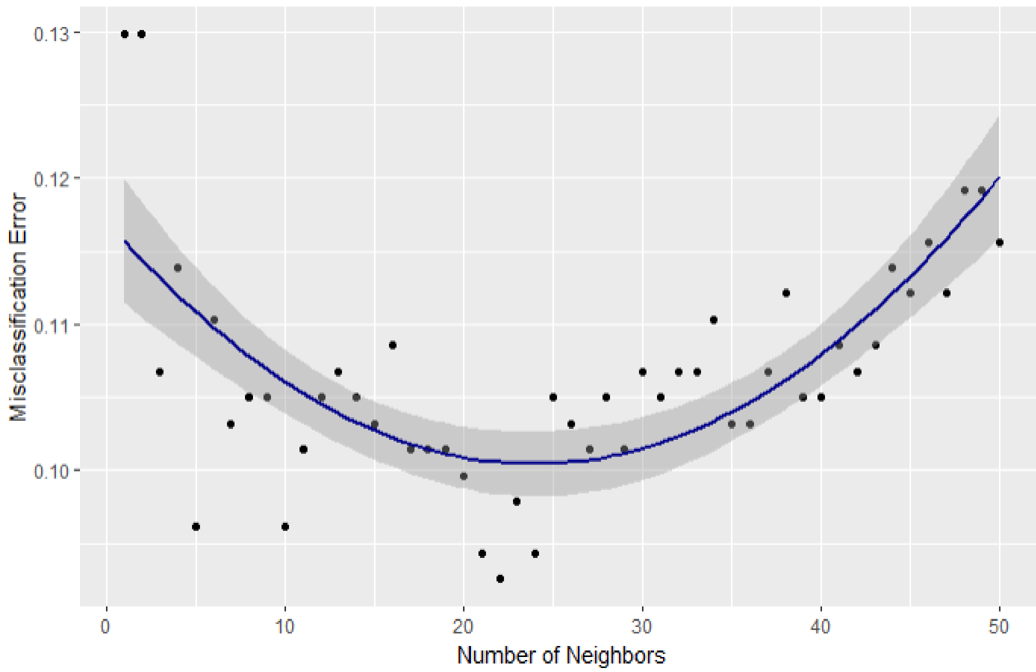


Figure 13: Misclassification error associated with different neighbor quantities

6.6 Linear Discriminant Analysis (LDA)

Linear discriminant analysis (LDA) is a dimensional reduction classification method that depends on the normality of the classes in the response variables, and seeks to maximize the separation between the classes (subsequently minimizing the misclassification error).

We generated an LDA model, and the output is listed below in Figure 14 and 15. The testing error of the LDA model was 0.1886121. In Figure 16, we have the confusion matrix for the LDA model, when applied onto the testing data.

```
Group means:
       Nscore        Escore     Oscore      AScore     Cscore  Impulsive         SS   Alcohol
0 -0.05089368 -0.0002935974 -0.2980838  0.07746401  0.1988430 -0.2055659 -0.3303218 0.8155844
1  0.02314386  0.0304835294  0.4436558 -0.12047708 -0.2295484  0.3013005  0.4434787 0.8400735
       Caff  Nicotine   `18-24`   `25-34`   `35-44`    `45-54`    `55-64`        `65`
0 0.9259740 0.2974026 0.1701299 0.2844156 0.2428571 0.22597403 0.05844156 0.01818182
1 0.9522059 0.6875000 0.5588235 0.2150735 0.1213235 0.07904412 0.02573529 0.00000000
  `Left school at 16 years` `Left school at 17 years` `Left school at 18 years`
0                0.06363636                0.01038961                0.03246753
1                0.03492647                0.02389706                0.07536765
  `Left school before 16 years` `Masters degree` `Professional certificate/ diploma`
0                    0.01168831       0.20259740                           0.1649351
1                    0.01654412       0.08639706                           0.1286765
  `Some college or university, no certificate or degree` `University degree`  Australia     Canada
0                                              0.1441558           0.3090909 0.01558442 0.03896104
1                                              0.4283088           0.1764706 0.04411765 0.05147059
  `New Zealand` `Republic of Ireland`        UK       USA       Asian      Black `Mixed-Black/Asian`
0   0.001298701           0.009090909 0.7597403 0.1298701 0.023376623 0.02597403         0.000000000
1   0.003676471           0.016544118 0.2591912 0.5275735 0.003676471 0.01102941         0.005514706
  `Mixed-White/Asian` `Mixed-White/Black`     White    Gender
0         0.009090909         0.007792208 0.9142857 0.3805195
1         0.016544118         0.011029412 0.8897059 0.6801471
```

Figure 14: The group means of the LDA models for variables based on the two classes of marijuana use.

```
Coefficients of linear discriminants:
                                                               LD1
Nscore                                                -0.128719708
Escore                                                -0.063894941
Oscore                                                 0.309557993
AScore                                                -0.009474028
Cscore                                                -0.044445079
Impulsive                                             -0.006703214
SS                                                     0.205778013
Alcohol                                                0.273200620
Caff                                                   0.099991319
Nicotine                                               0.747972175
`18-24`                                                0.528292204
`25-34`                                               -0.030754587
`35-44`                                               -0.171371758
`45-54`                                               -0.343321680
`55-64`                                               -0.197907859
`65`                                                  -1.129097364
`Left school at 16 years`                              0.388558815
`Left school at 17 years`                              0.814517525
`Left school at 18 years`                              0.522752868
`Left school before 16 years`                          0.944261957
`Masters degree`                                      -0.092648764
`Professional certificate/ diploma`                    0.347386663
`Some college or university, no certificate or degree`  0.344774309
`University degree`                                   -0.038493892
Australia                                              0.476439450
Canada                                                -0.080914690
`New Zealand`                                          1.228977238
`Republic of Ireland`                                 -0.174970800
UK                                                    -0.756964941
USA                                                    0.402876112
Asian                                                 -0.649234781
Black                                                 -0.407009859
`Mixed-Black/Asian`                                    0.989634489
`Mixed-White/Asian`                                    0.109803221
`Mixed-White/Black`                                   -0.393579887
White                                                 -0.282101194
Gender                                                 0.380486085
```

Figure 15: The coefficients of the linear discriminants for the LDA model derived in figure 11.

```
lda.pred   1   2
       0 276  57
       1  49 180
```

Figure 16: The confusion matrix for the classification of testing observations from the LDA model. The column values represent the predicted class, and the row values represent the actual class. In the columns, 1 corresponds with the binary value 0, and 2 corresponds with the binary value 1.

6.7 Support Vector Machines (SVM)

Support vector machine (SVM) is a classification method, where we devise a plane – in this case, a hyperplane – that separates the two classification groups, while seeking to maximize the largest margin possible between data points of different groups. The line spans through *n* dimensions, and the hyperplane derived from the model depends on the distance of data points close to the margins – support vectors – and the cost – the penalty for misclassification of observations.

We attempted to iterate through the cost parameter for the SVM model. However, we could not generate SVM models for all of the costs we desired, due to the limitations in R regarding the number of iterations it can run. Hence, we fixed the cost to 10, and calculated the test misclassification error as expected. In figure 17, we have the summary output for the SVM. There are 588 support vectors – in this model, there were 588 values that the SVM deemed influential towards determining the hyperplane and margins of the model. The SVM model below obtained a testing misclassification error of 0.1779359, and figure 18 shows the confusion matrix for the SVM model on the testing set.

```
Call:
svm(formula = y ~ ., data = SVM_data, kernel = "linear", cost = 10, scale = FALSE)


Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  linear
       cost:  10

Number of Support Vectors:  588

 ( 294 294 )


Number of Classes:  2

Levels:
 0 1
```

Figure 17: The output for the SVM model, with cost fixed at 10. We have 588 support vectors, with 294 in each class.

```
SVM_predict   1   2
          0 276  51
          1  49 186
```

Figure 18: The confusion matrix for the SVM model, applied onto test data for classification. The columns indicate the predicted SVM class of the observations, and the rows indicate the actual class of the observations. In the columns, 1 corresponds with the binary value 0, and 2 corresponds with the binary value 1.

7 Discussion

In Figure 19, we have compiled the models we have performed on the data, as well as their corresponding test error rates. The model with the lowest test error rate is the k-nearest neighbor model, with atest error of 0.0922 when k = 22 neighbors. Even though k-nearest neighbors did yield the smallest test error rate by a substantial amount, the model did not give any insight to the variables that were significant in classifying cannabis use. Models like decision trees and logistic regression have higher test misclassification rates than KNN, but allow us to examine the statistical importance of each variable. Referring to Figure 10 and Figure 20, the most significant variables used in the decision tree were the UK, nicotine, gender, and age 18-24 dummy variables, in addition to the standardized openness personality trait value. In the logistic regression, the most statistically significant variables based on the three asterisks denotation at a less than .01 p-value, are the openness personality variable, nicotine binary variable, the UK binary variable, and the gender binary variable.

Method	Accuracy
Lasso Regression	82.74%
Logistic Regression	82.38%
Boosting	82.56%
Decision Tree	80.07%
Random Forest	81.85%
k-Nearest Neighbors	**90.74%**
Linear Discriminant Analysis	81.14%
Support Vectors Machines	82.02%

Figure 19: The test accuracy rate for each model used in this paper. As seen here, the k-Nearest Neighbors had the greatest predictive accuracy

```
##                                                              z value Pr(>|z|)
## (Intercept)                                                  -0.031 0.975024
## Nscore                                                       -2.126 0.033532 *
## Escore                                                       -0.908 0.363892
## Oscore                                                        5.541 3.01e-08 ***
## AScore                                                        0.010 0.992309
## Cscore                                                       -0.971 0.331679
## Impulsive                                                    -0.153 0.878591
## SS                                                            3.250 0.001152 **
## Alcohol                                                       2.413 0.015842 *
## Caff                                                          0.718 0.472605
## Nicotine                                                      7.403 1.33e-13 ***
## '18-24'                                                       0.030 0.976370
## '25-34'                                                       0.028 0.977540
## '35-44'                                                       0.028 0.977757
## '45-54'                                                       0.027 0.978217
## '55-64'                                                       0.028 0.977656
## 'Left school at 16 years'                                     1.219 0.222805
## 'Left school at 17 years'                                     1.817 0.069291 .
## 'Left school at 18 years'                                     1.753 0.079573 .
## 'Left school before 16 years'                                 2.171 0.029914 *
## 'Masters degree'                                             -0.363 0.716977
## 'Professional certificate/ diploma'                           1.345 0.178714
## 'Some college or university, no certificate or degree'        1.193 0.232836
## 'University degree'                                          -0.337 0.736287
## Australia                                                     1.765 0.077616 .
## Canada                                                        0.084 0.933016
## 'New Zealand'                                                 1.326 0.184769
## 'Republic of Ireland'                                        -0.440 0.659892
## UK                                                           -3.321 0.000897 ***
## USA                                                           2.133 0.032942 *
## Asian                                                        -1.544 0.122688
## Black                                                        -1.198 0.230778
## 'Mixed-Black/Asian'                                           0.012 0.990533
## 'Mixed-White/Asian'                                           0.619 0.535638
## 'Mixed-White/Black'                                          -0.635 0.525312
## White                                                        -0.986 0.324309
## Gender                                                        3.886 0.000102 ***
```

*Figure 20: The p-values of predictors using logistic regression. The most significant variables are denoted with "***" meaning its p-value is below .01*

8 References

1 Al-Omari, W. M. M. (2015, September). Correlation between Personality Traits, Learning Approaches and Academic Performance of Dental Students. CORE. Retrieved November 14, 2022, from https://core.ac.uk/display/41241837

2 Dale RM, Bryant KA, Finnegan N, Cromer K, Thompson NR, Altinay M, Anand A. (2020, January). The NEO-FFI Domain of Openness to Experience Moderates Ketamine Response in Treatment Resistant Depression. J Affect Disord. 1;260:323-328. Doi: 10.1016/j.jad.2019.09.010. Epub 2019 Sep 3. PMID: 31521869.

3 Fehrman, E., Muhammad, A. K., Mirkes, E. M., Egan, V., & Gorban, A. N. (2017). The Five Factor Model of Personality and Evaluation of Drug Consumption Risk. *Data Science*. Doi: 10.1007/978-3-319-55723-6_18.

4 Feldman, M., Boyer, B., Kumar, V. K., & Prout, M. (2011). Personality, Drug Preference, Drug use, and Drug Availability. *Journal of Drug Education*, *41*(1). Doi: 10.2190/de.41.1.c

5 Ferwerda, B., & Tkalčič, M. (2020). Exploring the Prediction of Personality Traits from Drug Consumption Profiles. *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. Doi: 10.1145/3386392.3397589.

6 Labouvie, E. W., & McGee, C. R. (1986). Relation of Personality to Alcohol and Drug Use in Adolescence. *Journal of Consulting and Clinical Psychology, 54*(3), 289–293. Doi: 10.1037/0022-006X.54.3.289.

7 National Archives and Records Administration. (n.d.). *How Illicit Drug Use Affects Business and the Economy*. National Archives and Records Administration. Retrieved November 14, 2022, from https://obamawhitehouse.archives.gov/ondcp/ondcp-fact-sheets/how-illicit-drug-use-affects-business-and-the-economy#:~:text=Economic%20Costs&text=This%20value%20includes%3A,drug%E2%80%90related%20medical%20consequences%3B%20and