

MSDSA Graduation: Spring 2025

INTRODUCTION

Overview

- Online extremism refers to the promotion of extremist ideologies and hateful content on the internet
- With the ever-increasing popularity of social media sites, political and social views often make their way into a variety of online conversations
- Previous research by Costello, et al. (2016) suggests that demographics and distrust of the government affect exposure to hateful material online
- Research by the United States Government Accountability Office (2024) identified associations between online extremism and hate crimes. These reports also found that up to 1/3 of internet users have been exposed to this type of degrading content

Objective

- Identify patterns of personality traits related to impulsivity, trust, and closeness to others
- Predict the production of online hate material and identify the factors that influence online extremism

METHODS

Data Information

- Data was collected from the 2017 wave on the “Online Extremism Survey” (Costello, et al. 2023)
- Data collected from a sample of youth/young adults (15-36)
- Focus on demographics and exposure to online extremism
- 1,076 observations, 158 variables

Data Handling

- 53 variables selected related to the topic of interest (agreement and participation in online extremism)
- Missingness assessed and values imputed using MICE imputation – imputed dataset used for clustering and XG Boost
- Variable correlation assessed to remove select variables with high correlation (3 variables removed)

Descriptive Analysis of Agreement with Extremism

- Demographic Categories
 - Political Views and Gender
 - To what extent do you agree with the degrading material?
 - 1 (Strongly Disagree) to 5 (Strongly Agree)

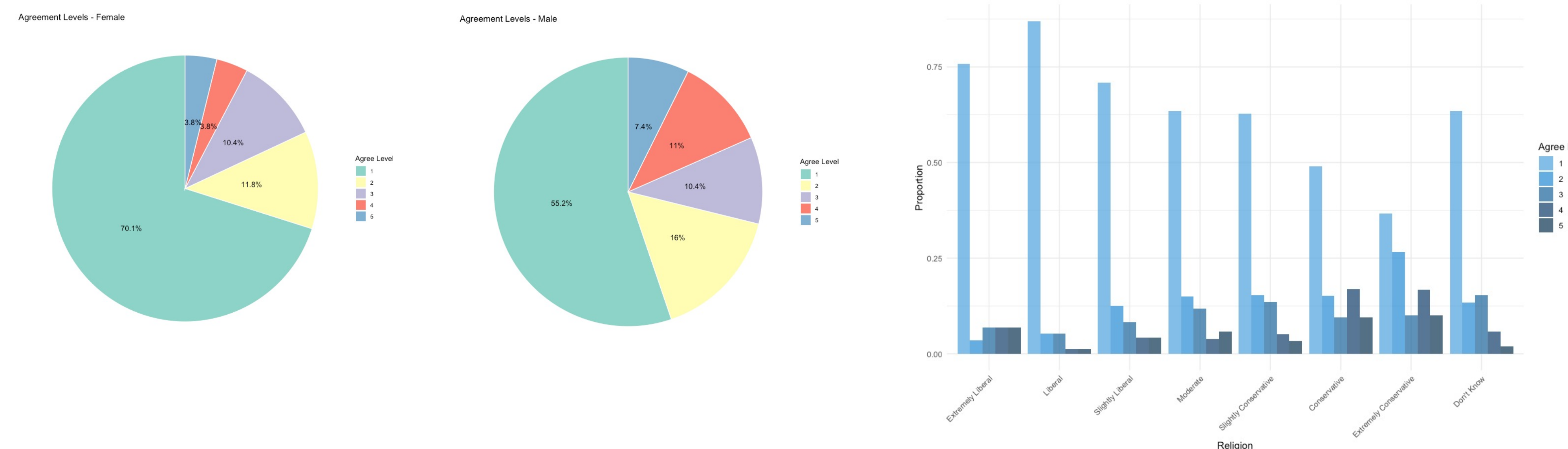
K-Means Cluster Analysis

- K-Means Clustering using responses from 4 different personality questionnaires related to personality, moral views, trust of others, and closeness to different groups
- Responses standardized for consistency across different scales
- Number of clusters determined using the “Elbow Method” to identify the point of diminishing return on variance explained
- 3 clusters selected and grouped based on personality measures
- Centers assessed to identify common patterns within clusters

XG Boost Model to Predict Posting of Hateful Material

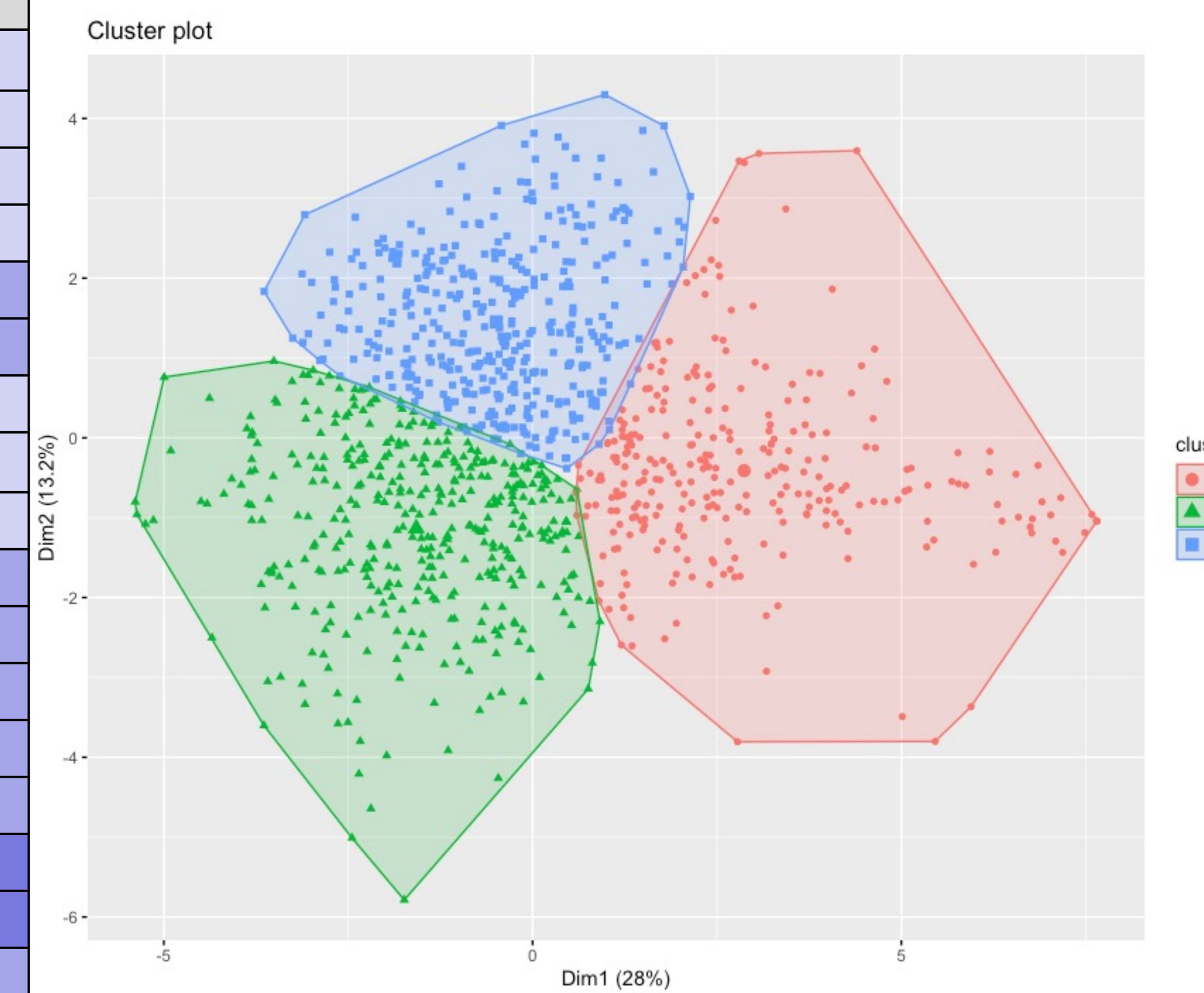
- Outcome: Have you ever produced online hateful/degrading material? (Binary outcome “Yes” or “No”)
- Data split into “Train” (80%) and “Test” (20%)
- Model Training**
 - Tuning Grid to Identify Best Hyperparameters: Rounds = 100, Depth = 6, ETA = 0.3, Gamma = 5, Features = 0.75, Min Child Weight = 2, Subsample = 1
 - Metric for Best Model: Kappa
 - Class weights applied in XG Boost training to improve prediction of positive class (11% of sample)
- Model Interpretation**
 - Variable importance to determine most influential features
 - Confusion matrix to assess correct/incorrect predictions

Demographics and Agreement with Degrading Material

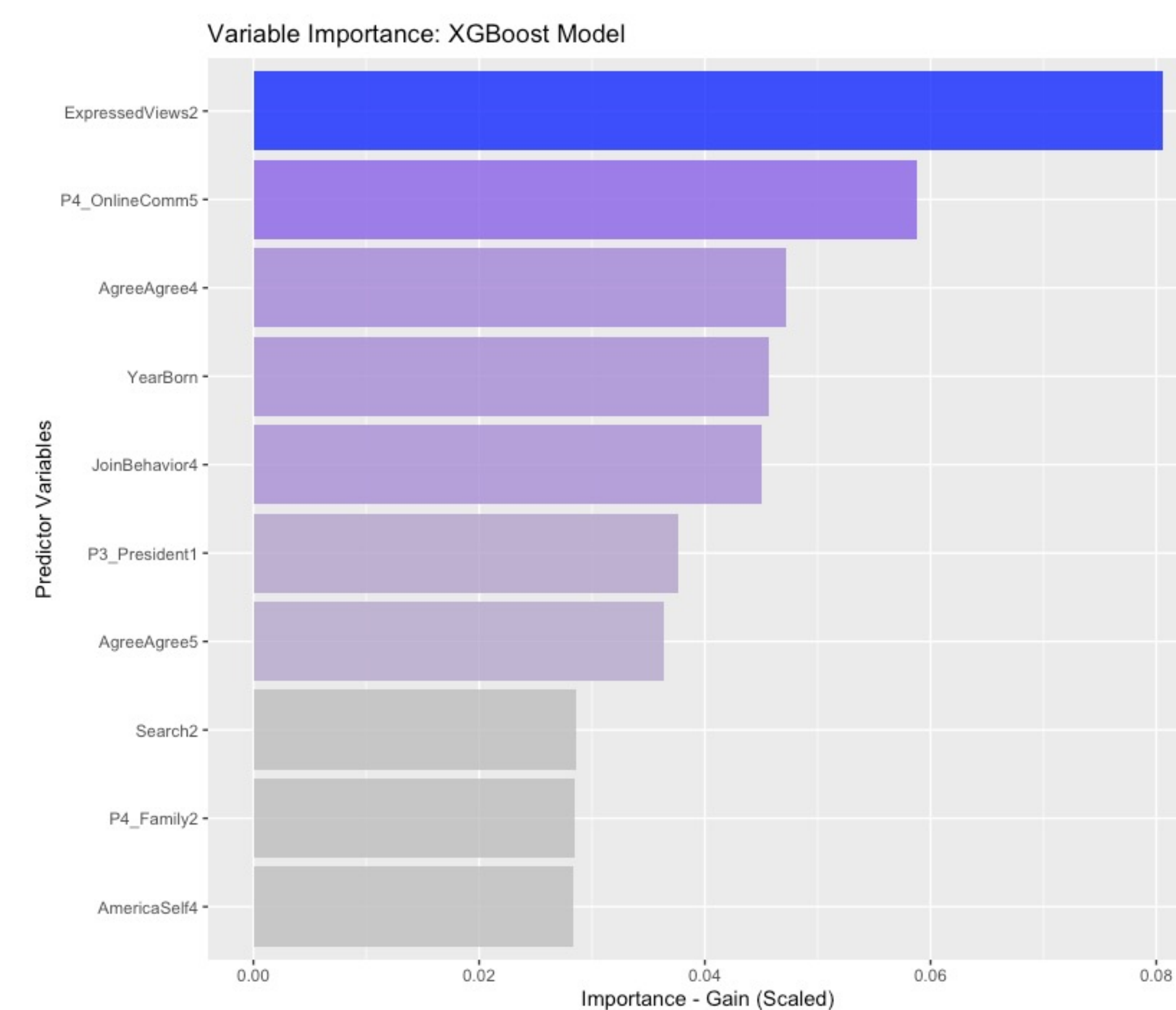


K-Means Cluster Analysis

Item	1	2	3
Friends Downloaded Music/Movies Illegally	0.47	0.18	-0.53
Impulsive Actions	0.72	0.12	-0.65
Joining in with Friends	0.94	-0.07	-0.61
Easier to Talk Anonymously Online	0.56	-0.04	-0.38
High Self Esteem	0.32	-0.48	0.25
Enjoy Taking Risks	0.47	-0.31	-0.03
Do What You Want if No One Gets Hurt	0.65	0.03	-0.51
Family/Friends Fighting is No One's Business	0.09	0.10	-0.17
Laws Made to be Broken	0.89	-0.13	-0.52
Trust in Business Leaders	0.90	-0.81	0.16
Trust in People	0.85	-0.66	0.05
Trust in Police	0.57	-0.76	0.36
Trust in Politicians	1.15	-0.60	-0.23
Trust in President	0.93	-0.64	-0.03
Feel Close to Family	0.04	-0.43	0.41
Feel Close to Friends	0.16	-0.46	0.35
Feel Close to Online Community	0.69	-0.38	-0.12
Feel Close to Religious Community	0.66	-0.65	0.18



XG Boost Model to Predict Producing of Hateful Material



Confusion Matrix

	Yes	No	
Yes	18	22	40
No	6	168	174
	24	190	

RESULTS

Descriptive Analysis of Agreement with Extremism

- Gender**
 - Female: Majority responded “strongly disagree” (70.1%) with only 7.6% reporting some level of agreement
 - Male: Majority responded “strongly disagree” (55.2%) with 18.4% reporting some level of agreement
 - Male respondents had higher rates of agreement with online degrading content compared to female respondents
- Political Views**
 - Those who described their views as “Liberal” had the highest rate of strong disagreement
 - Those who identified as “Conservative” or “Extremely Conservative” had the highest rate of agreement

K-Means Cluster Analysis

- Cluster 1:**
 - Higher levels of risk taking, impulsivity, social influence, rejection of law, trust, and levels of closeness to online/religious communities
- Cluster 2:**
 - Moderate levels of risk taking and impulsivity
 - Lower levels of self esteem, trust, and closeness to others
- Cluster 3:**
 - Lower levels of risk taking and impulsivity
 - Moderate levels of trust in various people/groups and closeness to online and religious communities
 - Higher levels of closeness to friends and family

XG Boost Model

- Model Performance**
 - Accuracy: 0.8692
 - Kappa: 0.4912
 - Sensitivity: 0.7500
 - Specificity: 0.8842
- Variable Importance**
 - Not expressing political views online was the most influential factor in predicting whether someone had produced hateful material
 - Other variables such as high trust in online communities, high agreement with degrading views, and age (year born) were also influential in predicting the outcome of interest

DISCUSSION

- Overall, men and conservatives are more likely to agree with online extremism compared to their counterpart groups
- Among young people from this sample, 3 distinct groups were identified based on key personality and moral indicators
- Understanding the different architectures of personality and how they relate to actions provides important insight into understanding of online behaviors
- Using demographics, personality assessments, and experiences with online extremism, the XG Boost model was able to predict production of online hate material with fairly good accuracy
- This model identified influential factors such as expressing political views and being close with online communities as the main factors influencing those who participate in online extremism
- While online extremism will continue to exist in some capacity, it is important to understand the patterns associated with this type of behavior in order to reduce the prevalence and real-world impact of exposure to this kind of hate material
- This study predominantly focused on demographic and personality characteristics related to online extremism. However, there are many other factors that could provide more useful information on this topic.