Drug Consumption

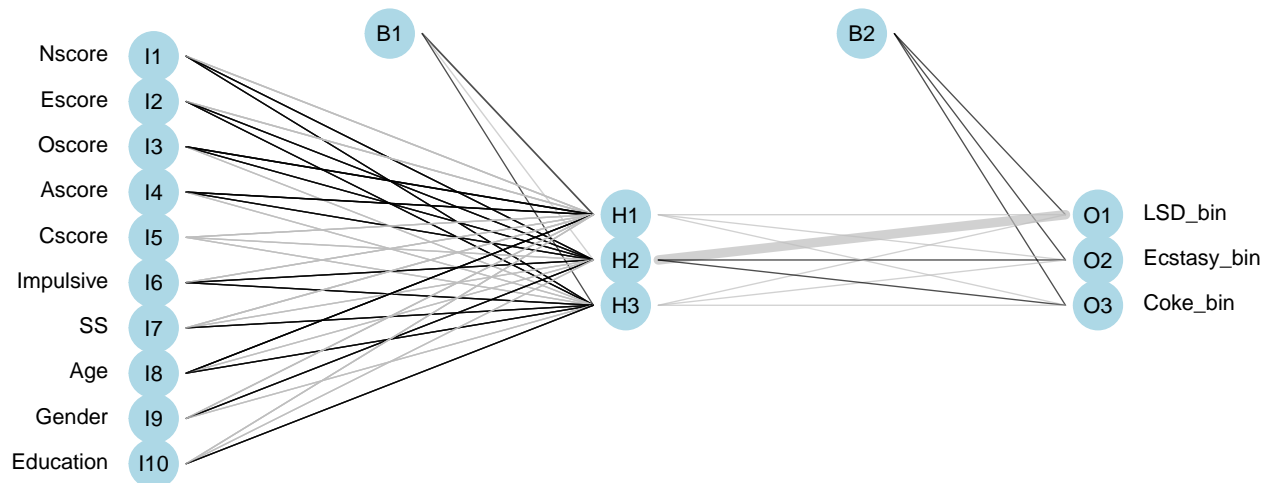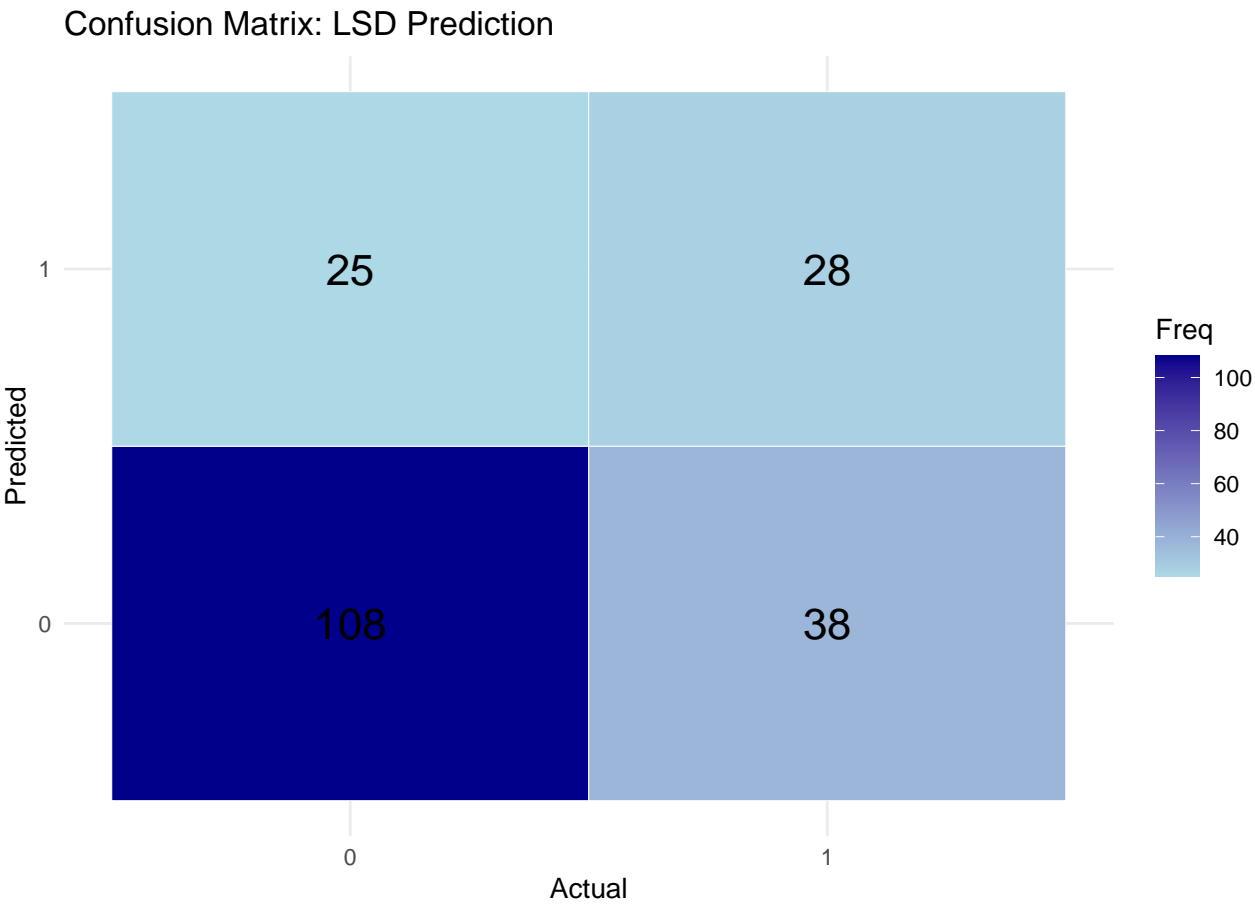Nhat Bui, Johan Ferreira, Thilo Holstein

2025-03-06

Contents

Figure 1: Neural network with three hidden neurons

```
## Error: 236.65
## Steps: 22306
```
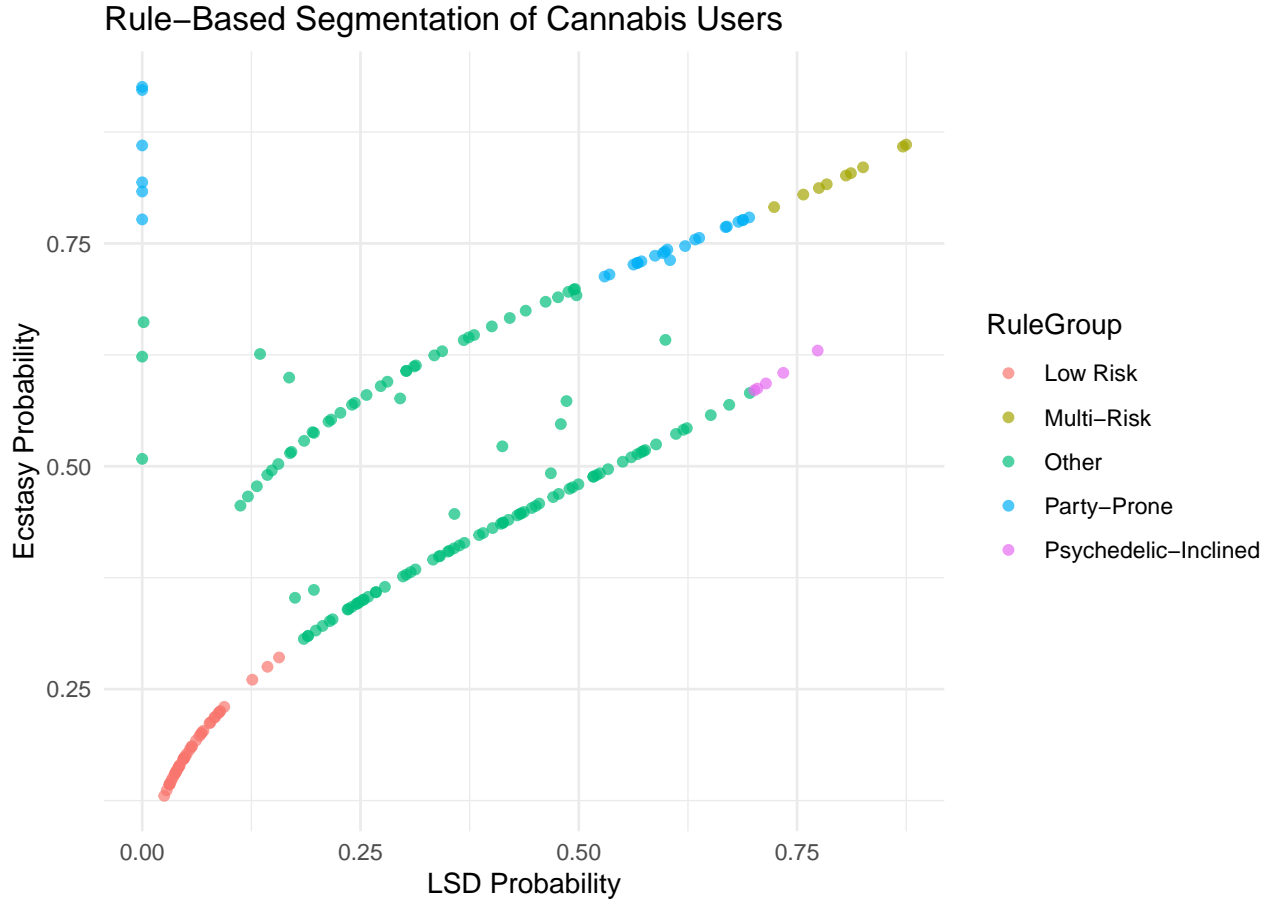
Table 1: Confusion Matrix: LSD Prediction

	0	1
0	108	38
1	25	28

Confusion Matrix: LSD Prediction

Rule−Based Segmentation of Cannabis Users

Table 2: Average predicted probabilities by k-means cluster

Cluster	LSD_avg	Ecstasy_avg	Coke_avg
Cluster 1	0.4861015	0.4792119	0.2575844
Cluster 2	0.1187120	0.2389790	0.2225929
Cluster 3	0.6291449	0.7501268	0.5932870
Cluster 4	0.1881014	0.6116553	0.6053562

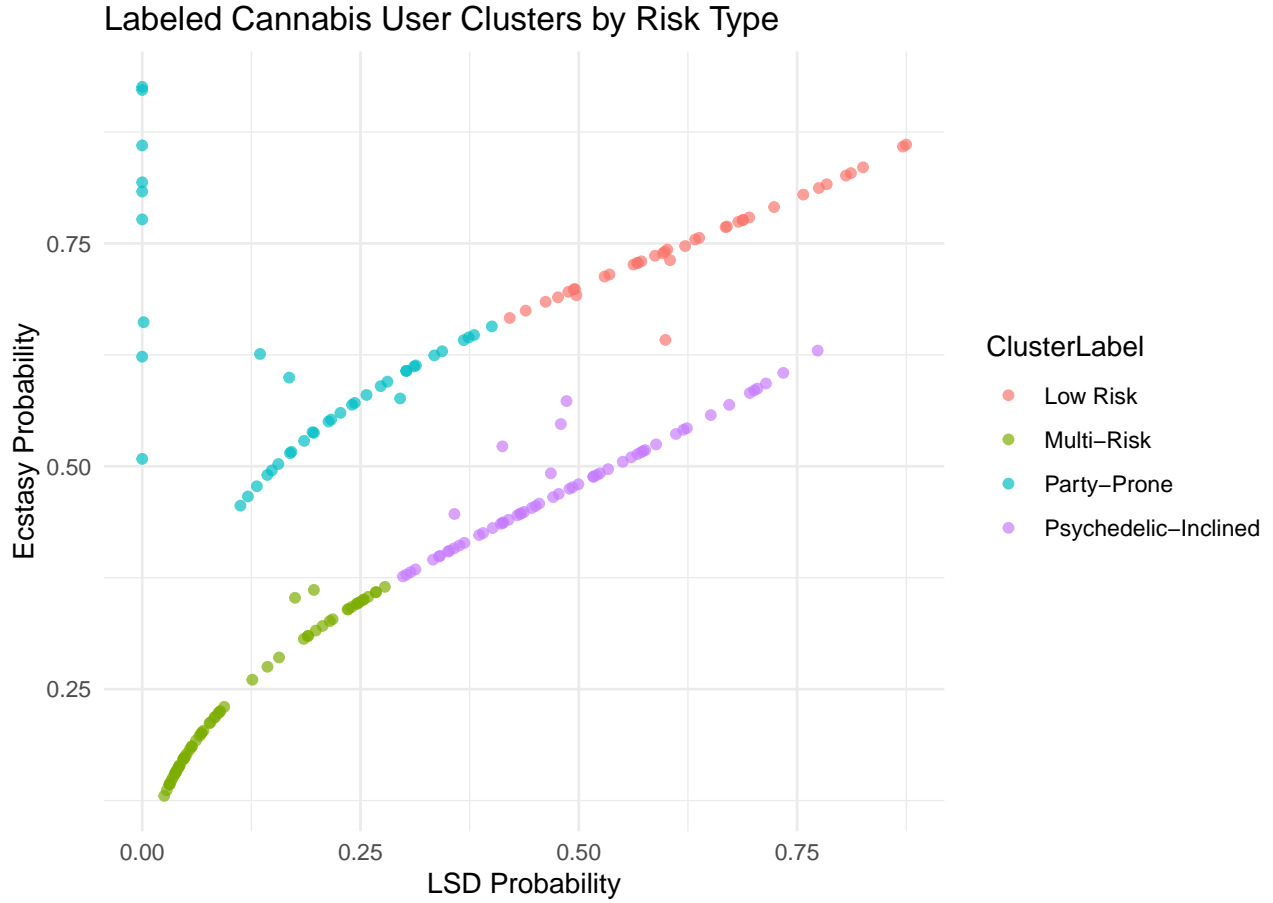## Labeled Cannabis User Clusters by Risk Type

Table 3: Average personality traits by rule-based profile

	Nscore	Cscore	Oscore	Ascore	Escore	Impulsive	SS	Age	Gender	Education
Low Risk	-0.0036255	-0.0448120	-0.2330163	-0.0861743	-0.4316005	-0.0051876	-0.5259474	1.6176809	0.5750000	7.025000
Multi-Risk	-0.0245733	0.1480370	0.7489982	0.0455476	0.5176912	0.6021230	1.0741244	-0.6518947	0.8888889	6.777778
Other	0.2153841	-0.1233864	-0.2243843	-0.1228750	-0.1255741	-0.0052508	-0.0183698	-0.3826926	0.6134454	6.411765
Party-Prone	0.2708529	-0.2392578	0.0540744	-0.2066216	0.6109786	0.5359331	0.8575281	-0.6125300	0.6923077	6.423077
Psychedelic-Inclined	-0.4494501	-0.2116513	1.5638150	0.5071308	-0.0067667	-0.1457635	-0.3429392	-0.7542431	0.8000000	5.000000

Personality Traits by Profile
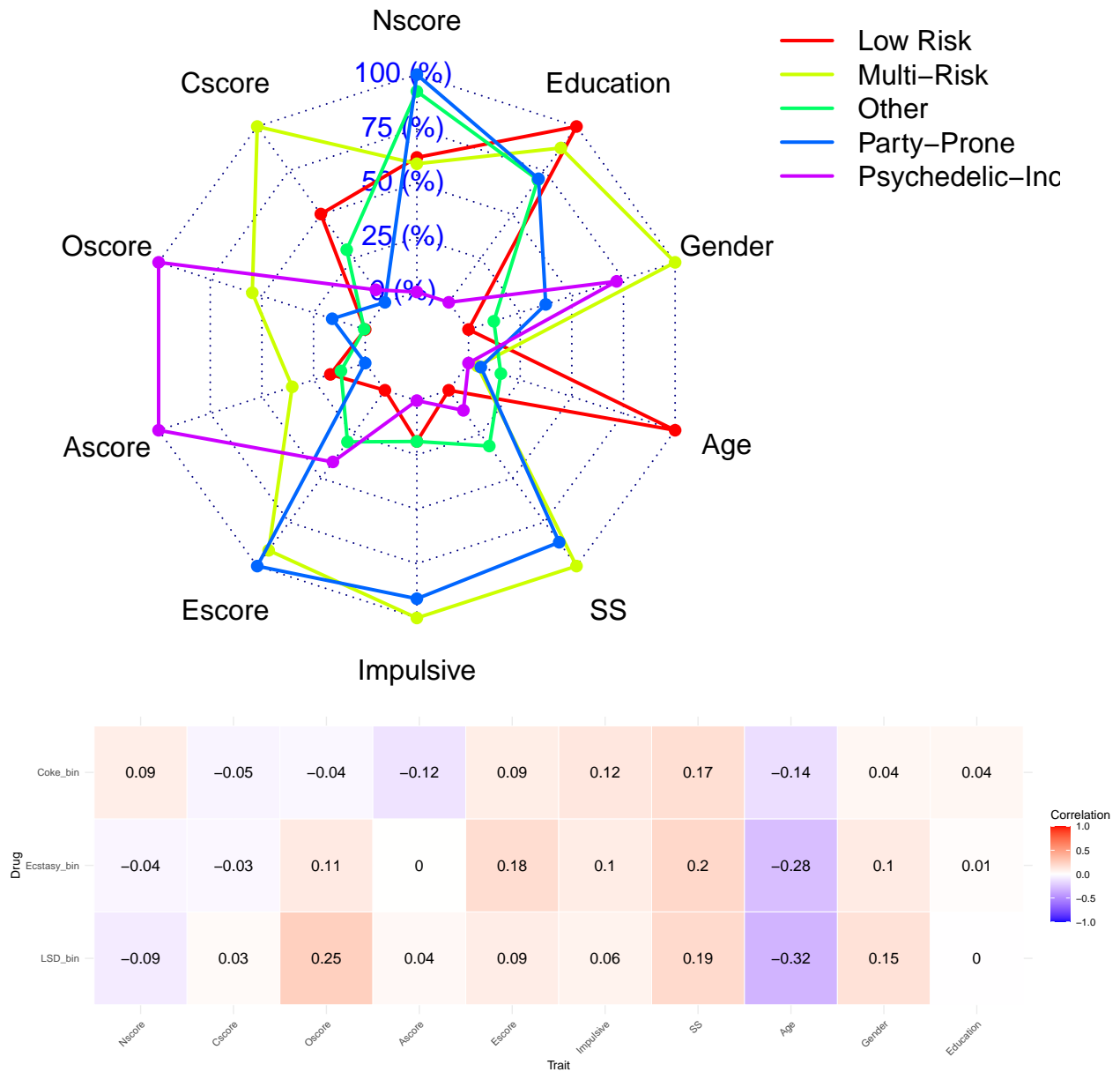
Figure 2: Trait–Drug Correlation Heatmap

Table 4: Binary SVM Confusion Matrix (Accuracy = 0.854)

	No	Yes
No	207	25
Yes	30	115

Table 5: Multi-class SVM Confusion Matrix (Accuracy = 0.830)

	None	Moderate	Poly
None	10	2	0
Moderate	4	188	25
Poly	1	32	115

1 Data Preparation

2 Binary SVM Model
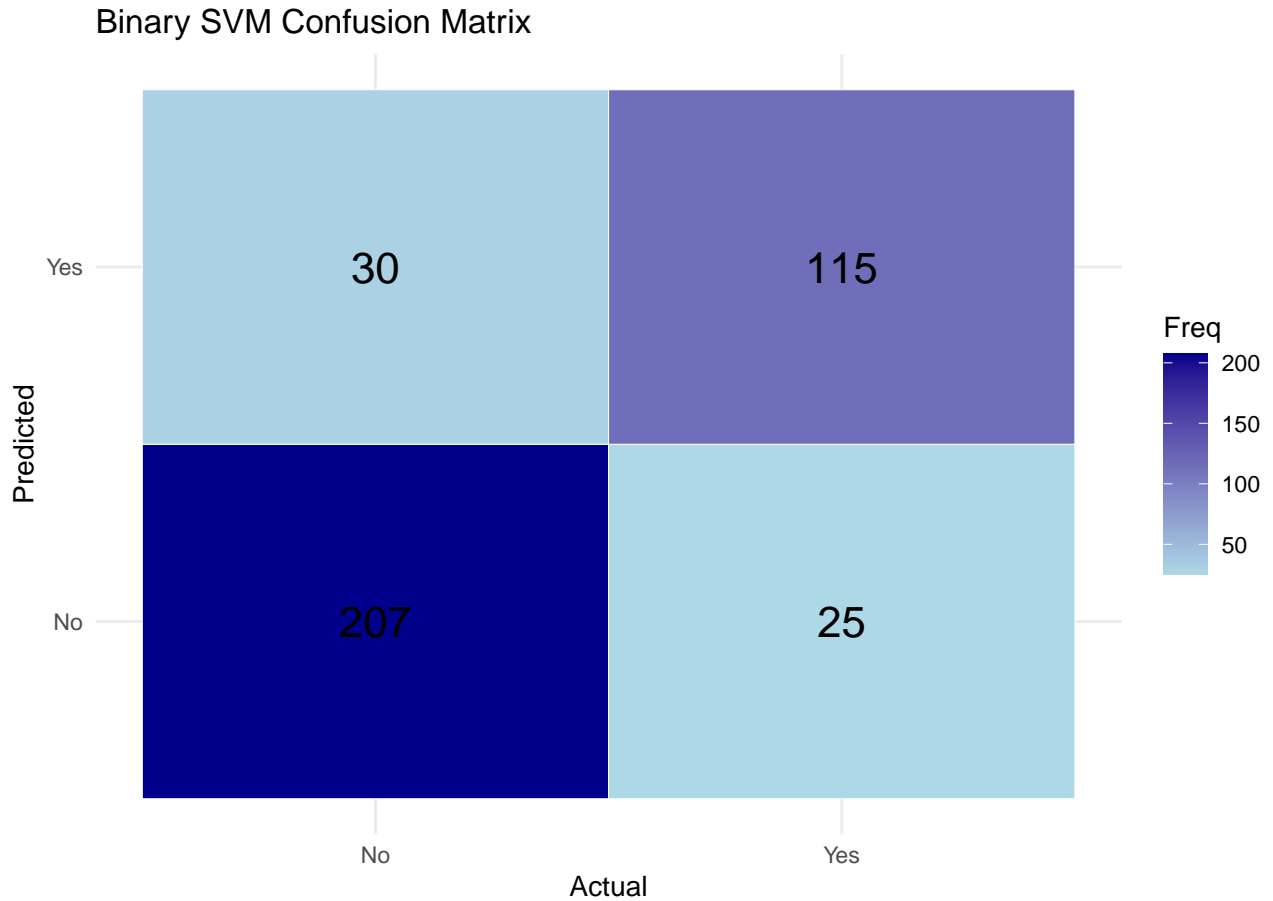
Table 6: Top 3 Highest-Risk Synthetic Profiles

Nscore	Cscore	Oscore	Ascore	Escore	Impulsive	SS	Age	Gender	Education	Alcohol	Cannabis	PredictedRisk	Gender_label	Education_label	Profile
-0.5303591	0.3743403	0.6891178	0.6248764	-0.5093128	-0.2639730	0.7654000	1	1	5	5.200818	4.863939	0.9492896	Male	Some college	Profile 1
0.1117710	-0.2465118	0.7092764	0.0228573	-0.4071493	-0.4827895	0.4269726	1	1	4	5.079785	4.842409	0.9485839	Male	18 years	Profile 2
0.5375388	0.0032667	0.6085776	-0.0433439	0.0838128	-0.1334711	0.5674478	1	1	4	5.118698	5.000000	0.9474303	Male	18 years	Profile 3

3 Multi-class SVM Model

4 Top-Risk Profile Simulation

Actual prediction: 1.00
Average prediction: 0.42
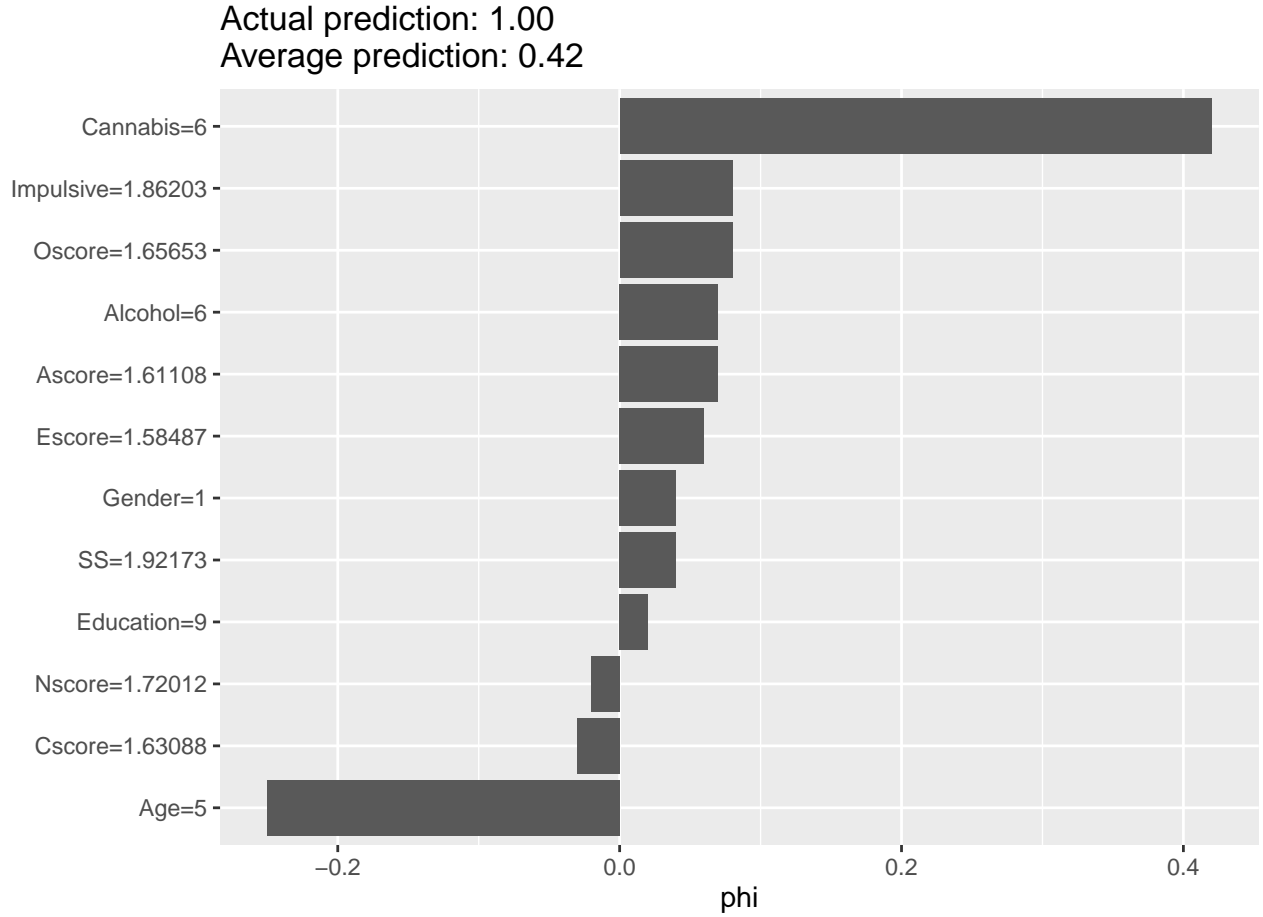
5 Feature Importance (Single-Feature Accuracy)

Prediction Accuracy by Single Feature
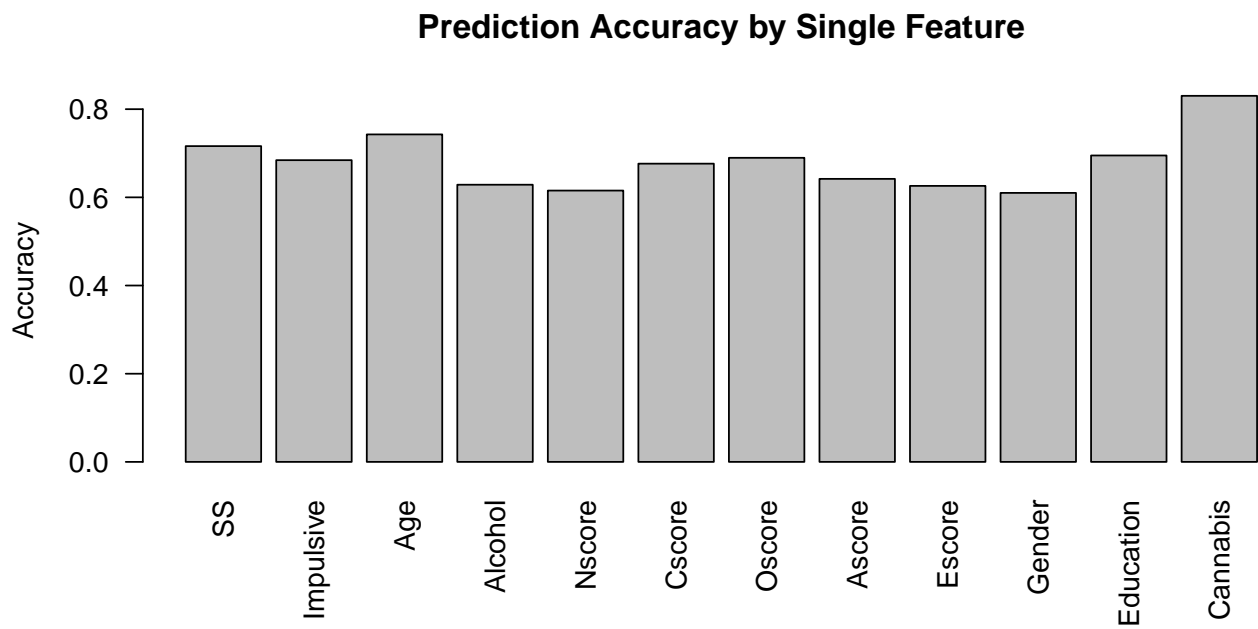
6 Model Confidence Histogram

Model Confidence Distribution
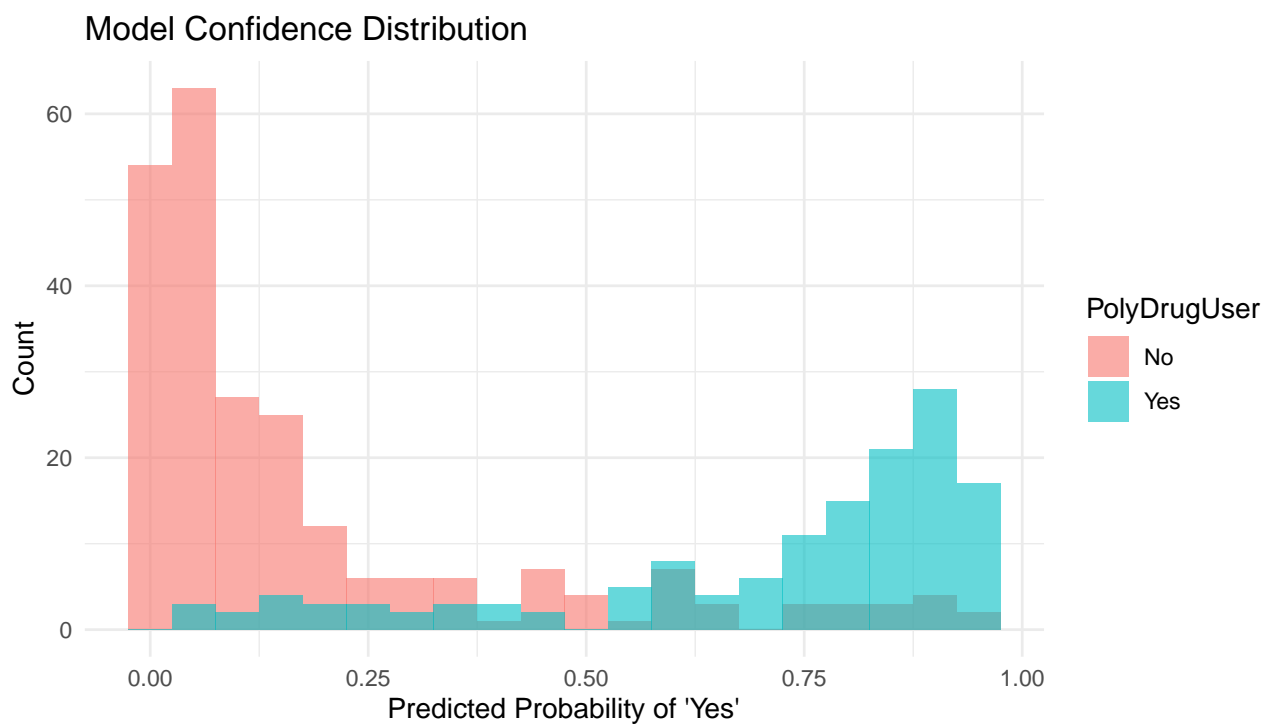
7 PCA Visualization of SVM Output

SVM Output in PCA−Reduced Feature Space

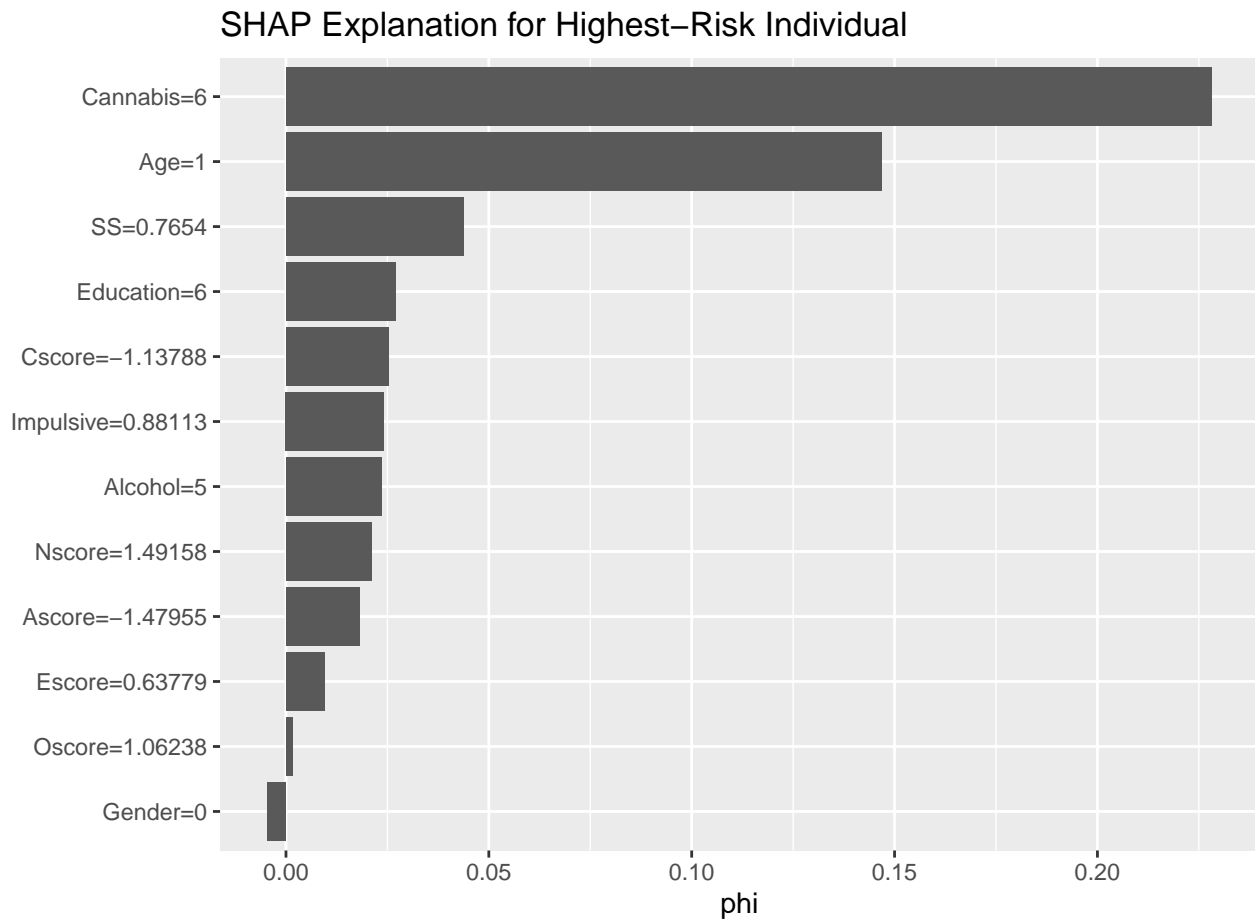Impulsiveness by Predicted PolyDrugUser Class
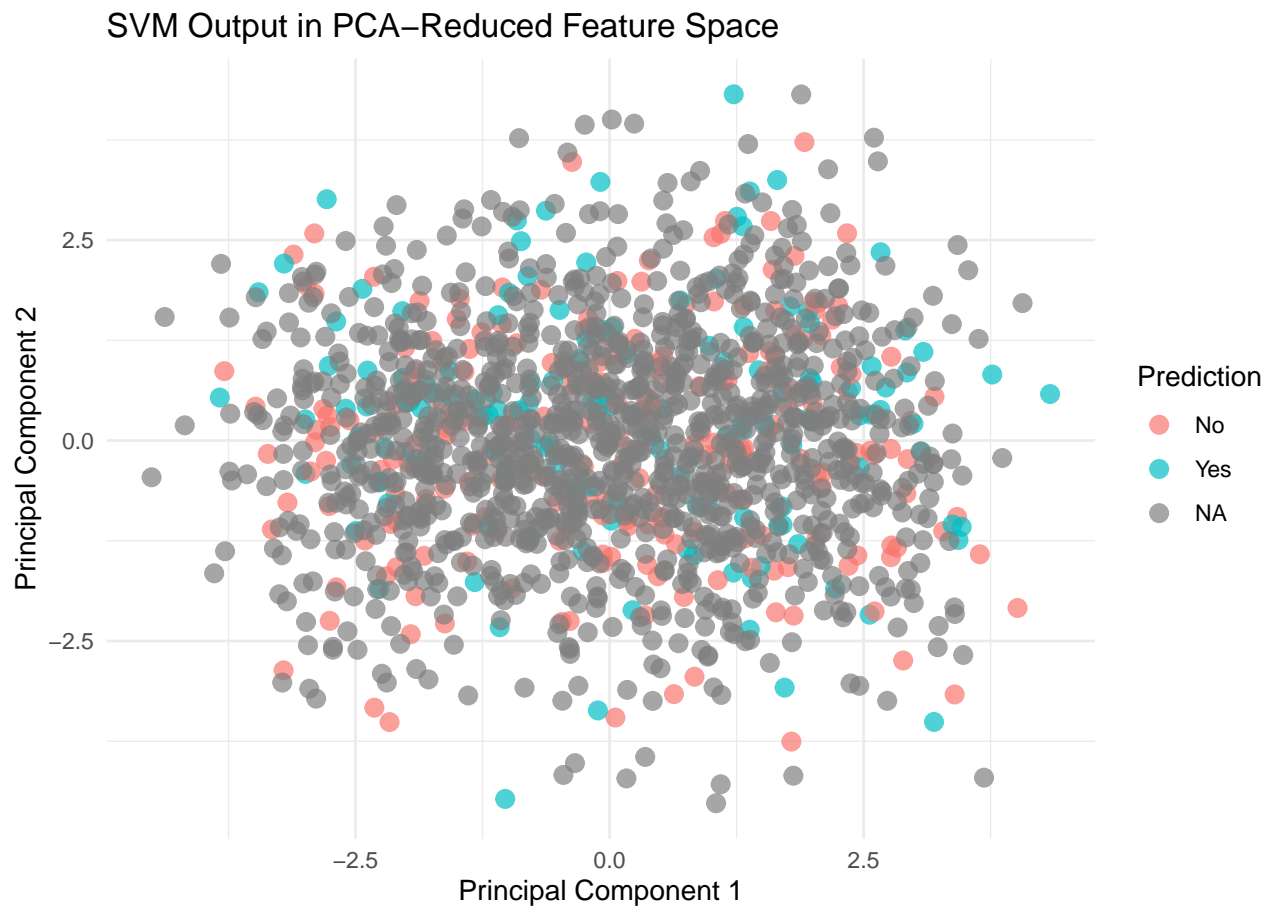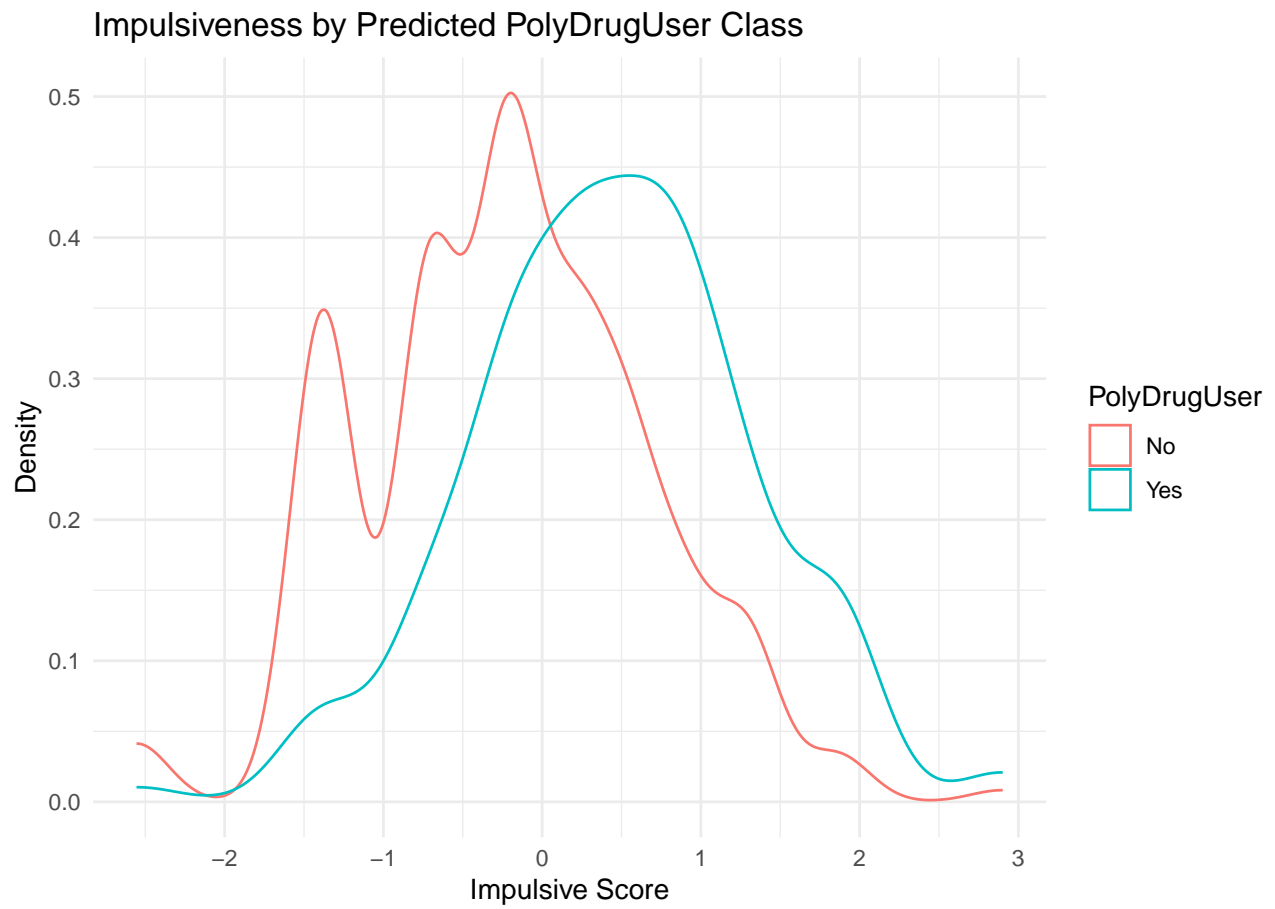
8 Interpretation: Comparing ANN and SVM Predictions

Pearson correlation between SVM and ANN mean probabilities: 0.042

9 Explainability with SHAP Logic

SHAP Explanation for Highest−Risk Individual


```
## 
## **Highest-Risk Individual (test set):**
##        Nscore    Cscore  Oscore   Ascore  Escore Impulsive      SS Age Gender
## 1366  1.49158 -1.13788 1.06238 -1.47955 0.63779   0.88113 0.7654   1      0
##       Education Alcohol Cannabis
## 1366         6       5        6
## 
## Predicted SVM risk probability: 0.956
```