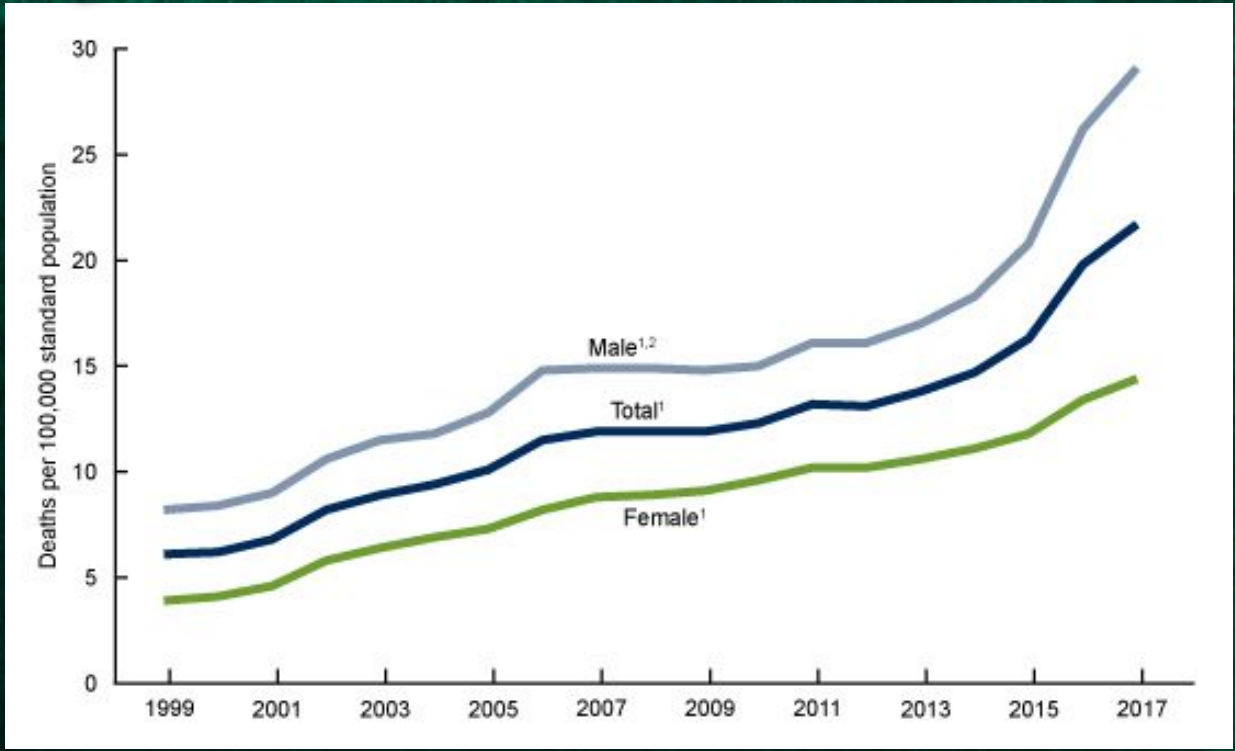
Predicting **Drug Use** with *Machine Learning*

❖ Drug Related Deaths: 1999-2017



Source: cdc.gov

"Research has shown that brain development continues into a person's 20s, and that age of drug initiation is a very important risk factor for developing addiction. This underscores the importance of ***drug use prevention and screening*** *for substance use or misuse among adolescents and young adults.*

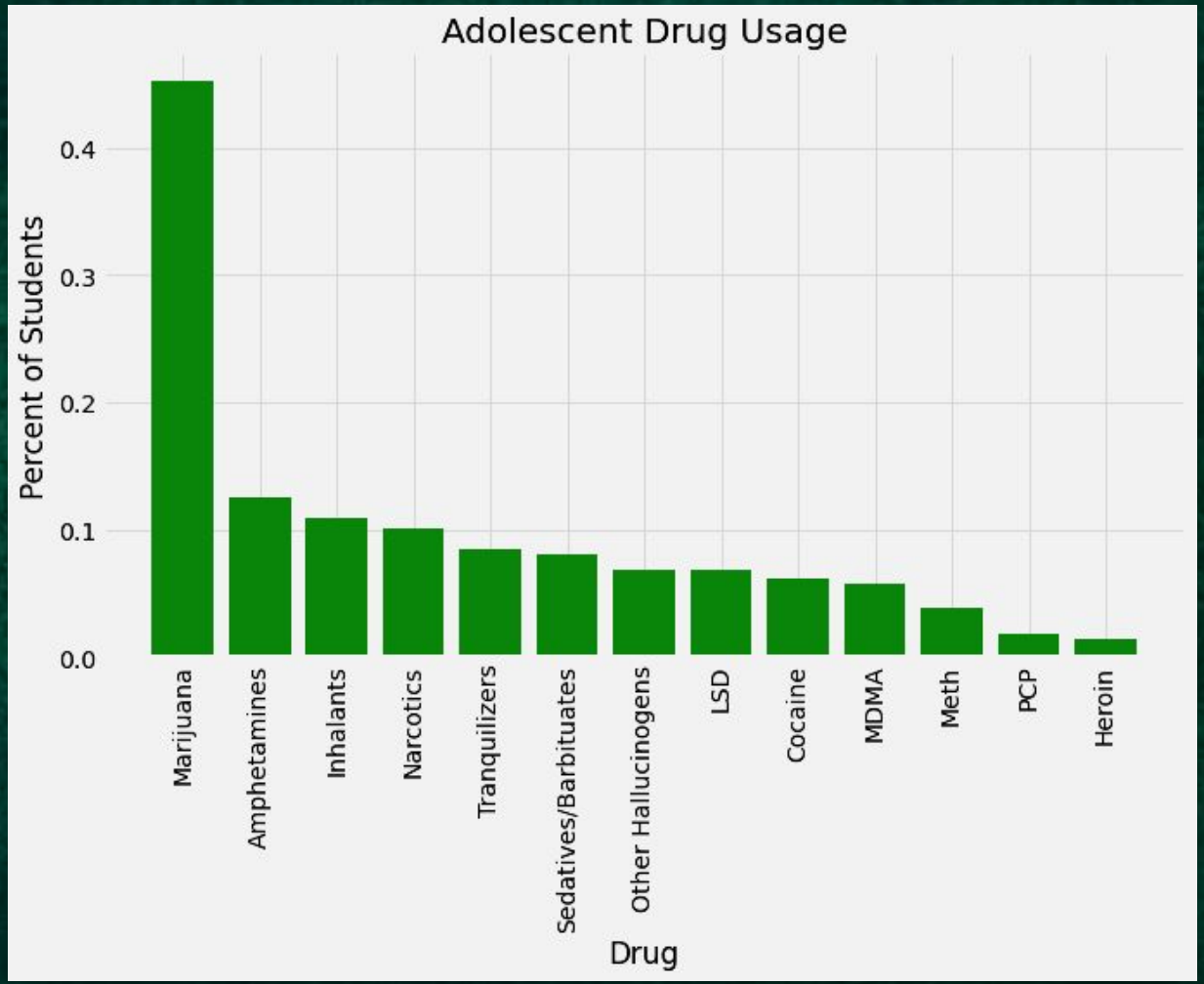
--Emily B. Einstein, Ph.D., chief of NIDA's Science Policy Branch

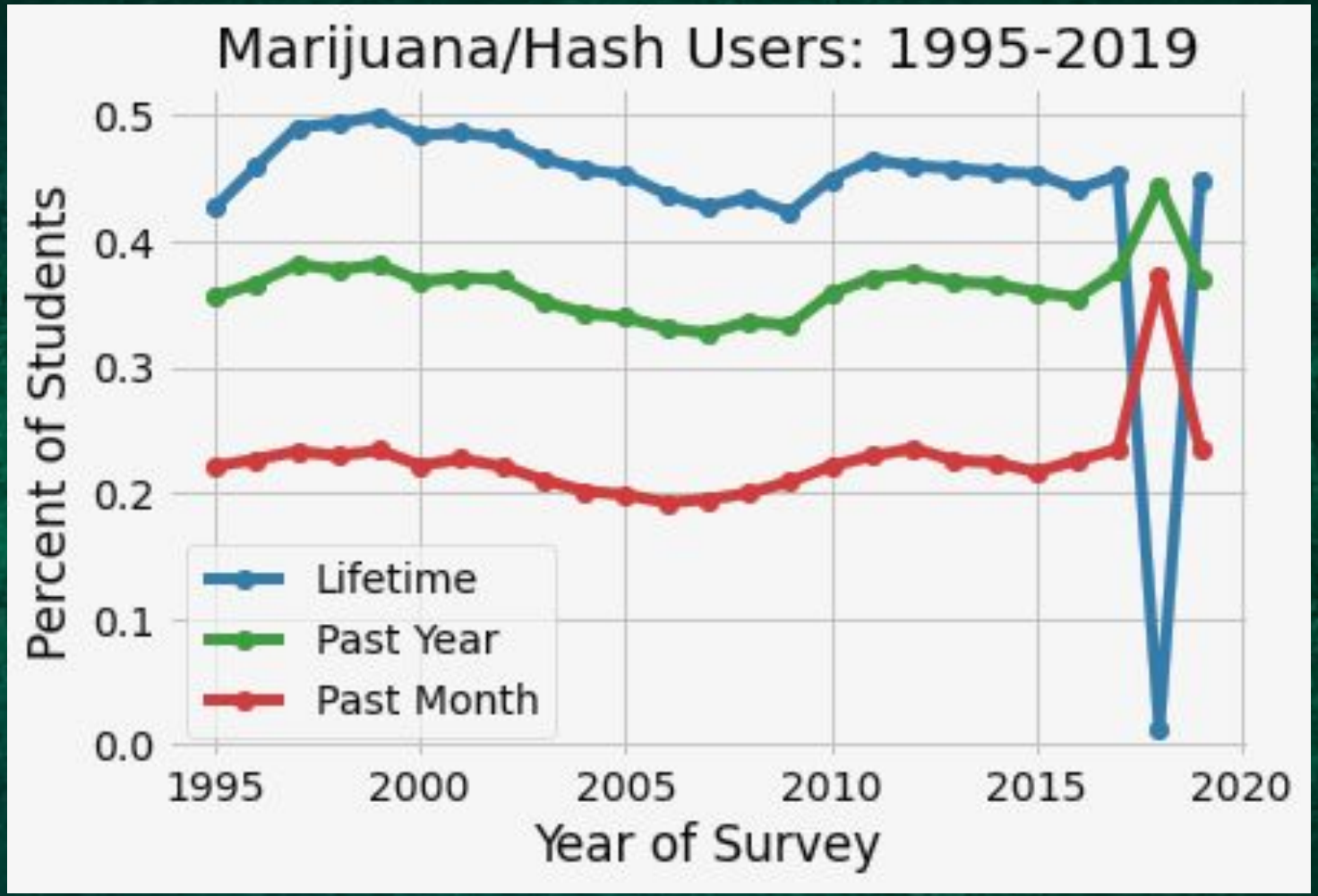
Data

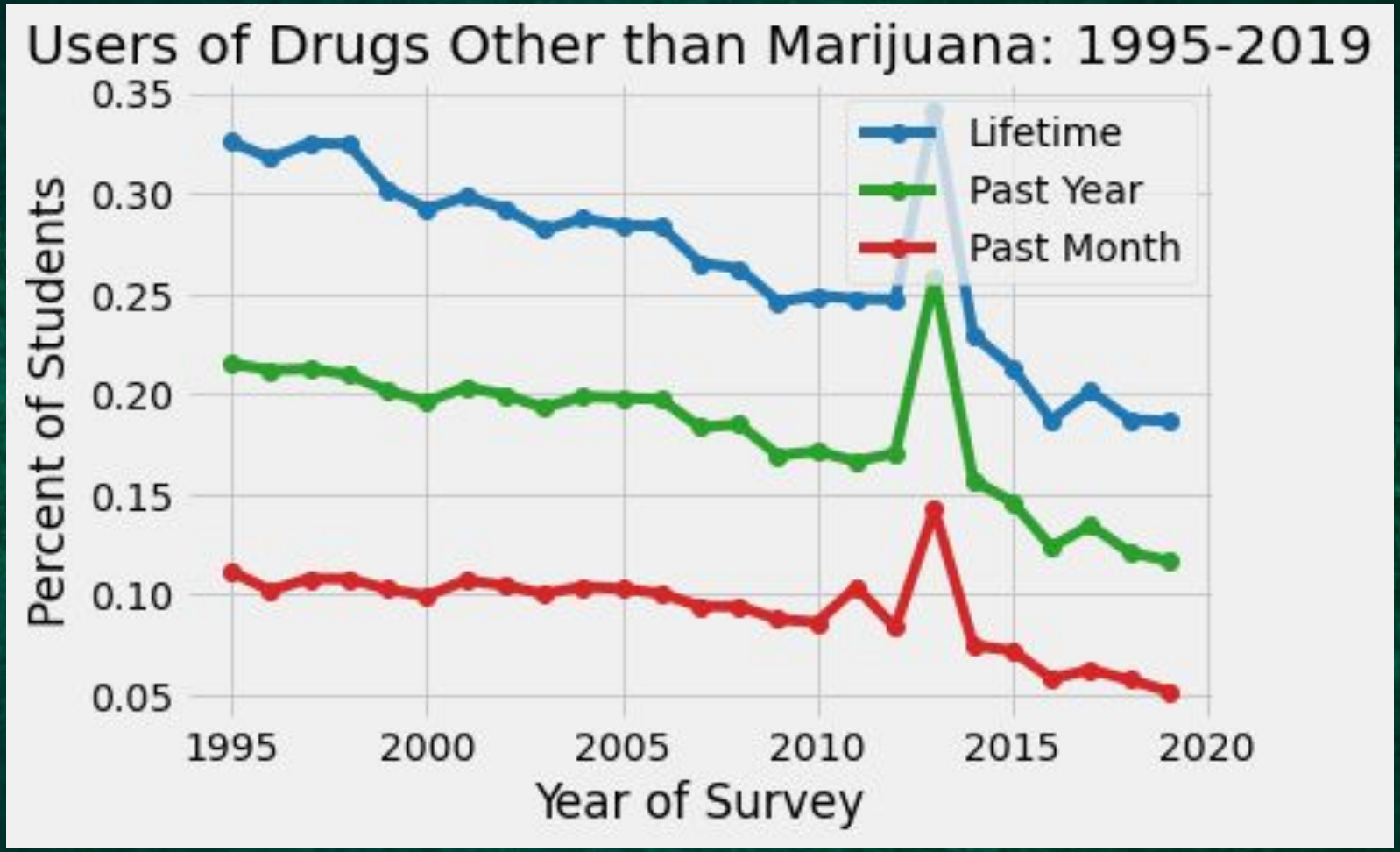
101,628 Surveys

133 Schools

48 Questions

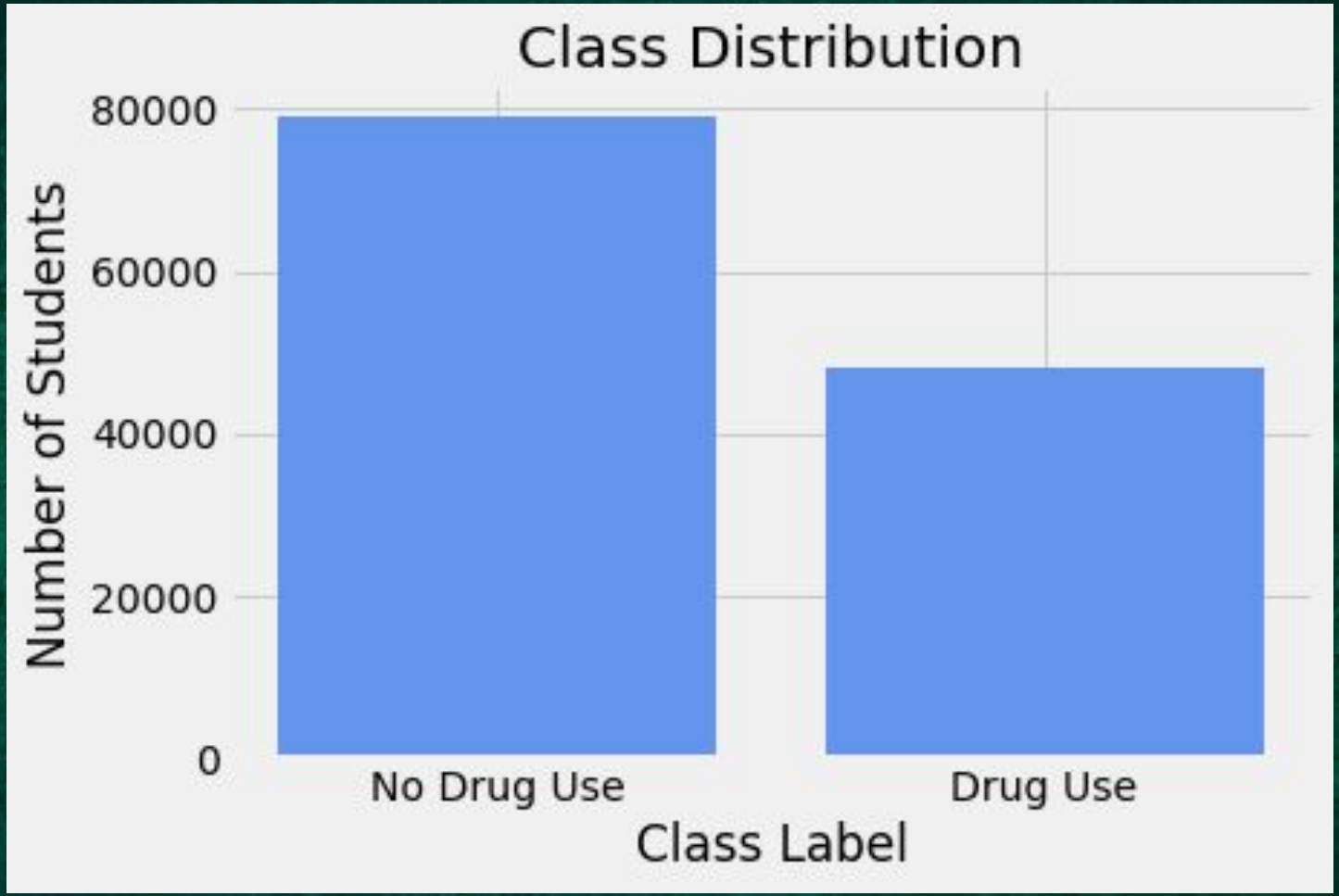

Adolescent Drug Usage
Percent of Students
0.0
0.1
0.2
0.3
0.4
Marijuana
Amphetamines
Inhalants
Narcotics
Tranquilizers
Sedatives/Barbituates
Other Hallucinogens
LSD
Cocaine
MDMA
Meth
PCP
Heroin
Drug


Marijuana/Hash Users: 1995-2019
Percent of Students
0.0
0.1
0.2
0.3
0.4
0.5
1995
2000
2005
2010
2015
2020
Year of Survey
Lifetime
Past Year
Past Month


Users of Drugs Other than Marijuana: 1995-2019
Percent of Students
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1995
2000
2005
2010
2015
2020
Year of Survey
Lifetime
Past Year
Past Month

❖ Target:

"Have you used any illicit drugs in the past *year*?"

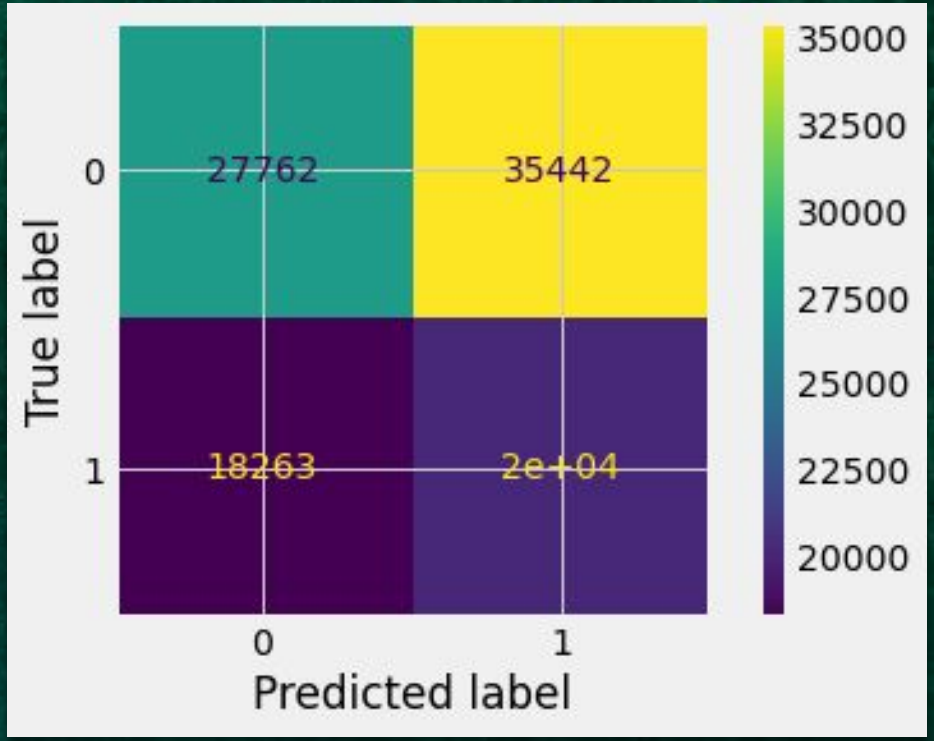

Class Distribution
80000
60000
40000
20000
0
Number of Students
No Drug Use
Drug Use
Class Label

❖ First Simple Model

Decision Tree

61% Recall

70% Accuracy

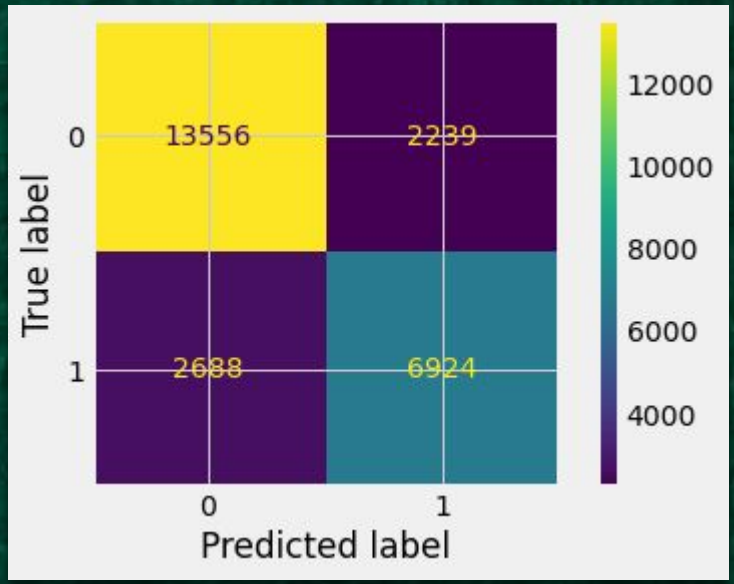


❖ Final Model

Logistic Regression

72% Recall

80.6% Accuracy



Positive Features

1. Year = 2019 **(1.12)**
2. Year = 2015 **(0.71)**
3. Consumed Alcohol *40+ times* in past year **(0.66)**

Negative Features

1. Consumed NO Alcohol in past year (-1.04)
2. Never Smoked a Cigarette (-1)
3. Grew up on a Farm (-0.58)

❖ Next Steps

- Further Subsetting
- Studying Age
- Neural Network

Thank you!

Jamie Dowat

https://github.com/Luv2bnanook44

https://www.linkedin.com/in/jamie-dowat-5a3b5a19b/