

Using classification to learn about teen electronic cigarette use

Maya Remington September 2, 2021

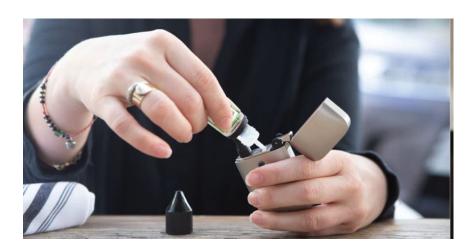
The Problem: E-cigarettes in teens

- Nicotine-containing electronic cigarettes: popular among middle and high school students in the US
- Nicotine is...
 - Highly addictive
 - Harmful to the adolescent brain
- E-cigs may also contain other potentially harmful compounds



2020 National Youth Tobacco Survey

- 14,000+ middle and high school students
- Survey questions cover multiple factors:
 - Demographics
 - Use of tobacco products (in detail!)
 - Tobacco-related beliefs and influences



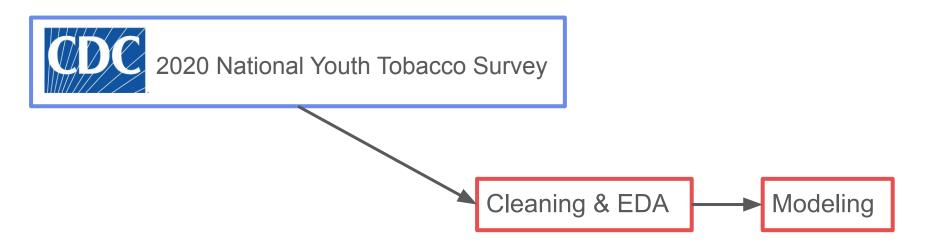
Classification target:

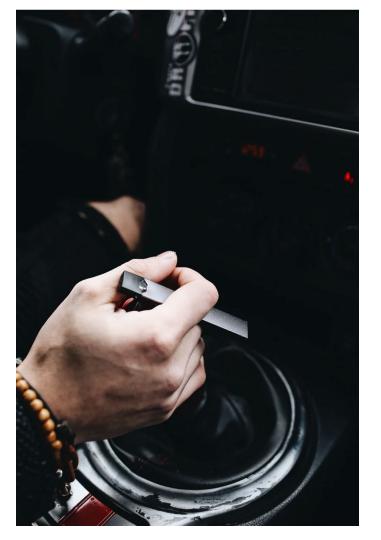
 Has the student ever used an e-cigarette? Yes or No

Aim:

 Determine the factors most associated with e-cigarette use

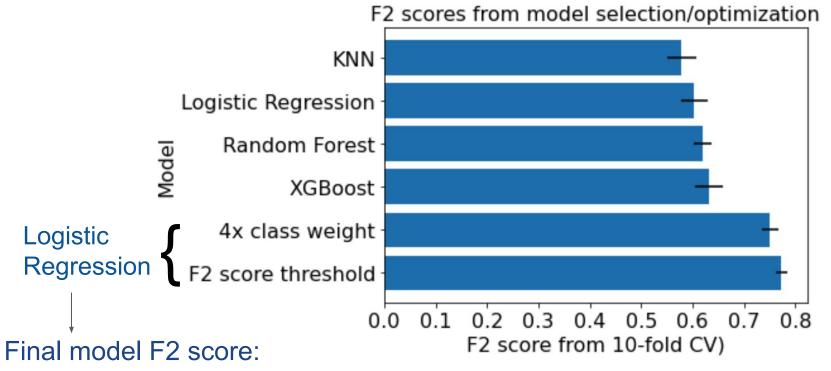
Data & Methods



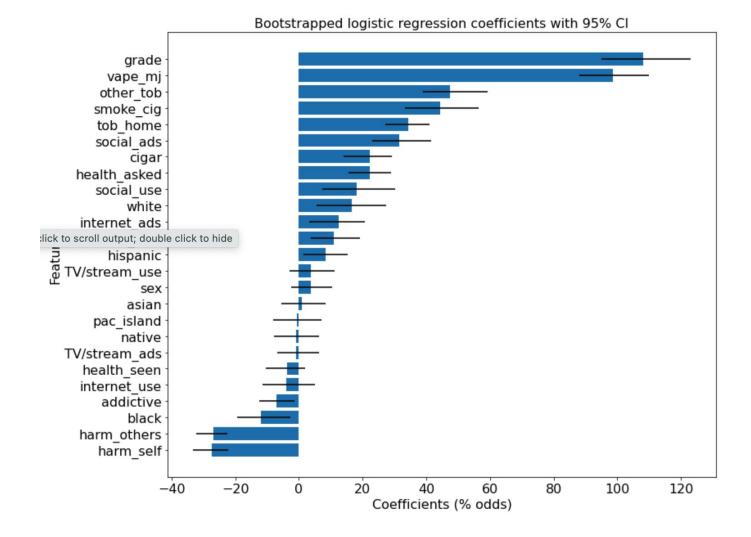


Findings

24.4% of 13,000+ students surveyed have used e-cigarettes



- Train: 0.753
- Test: 0.736



Conclusion

- A students grade, use of cannabis vape, and other tobacco products are most highly associated with e-cigarette use
- Also associated: decrease sense of potential harm to self or others



Future directions

- Try logistic regression on a smaller number of features those that appear to be highly associated with the target
- Try optimizing parameters for other model types (KKN, RF, XGBoost)

Thank you!















