



Generation Vape: Factors associated with teen e-cigarette use

Maya Remington
Summer/Fall 2021

The Problem: Electronic cigarettes in teens

- Nicotine-containing electronic cigarettes (e-cigs): Popular among middle and high school students in the US
- Problem because nicotine is...
 - Highly addictive
 - Harmful to the adolescent brain



This project



Aim: use a classification model to determine factors associated with e-cig use in teens:

- Use of other tobacco products
- Belief systems
- Exposure to pro- and anti-tobacco influences

Data & Methods



2020 National Youth Tobacco Survey

- ❑ 14,000+ students
- ❑ Grades 6-12
- ❑ 68+ questions (depending on answers)

CSV download:

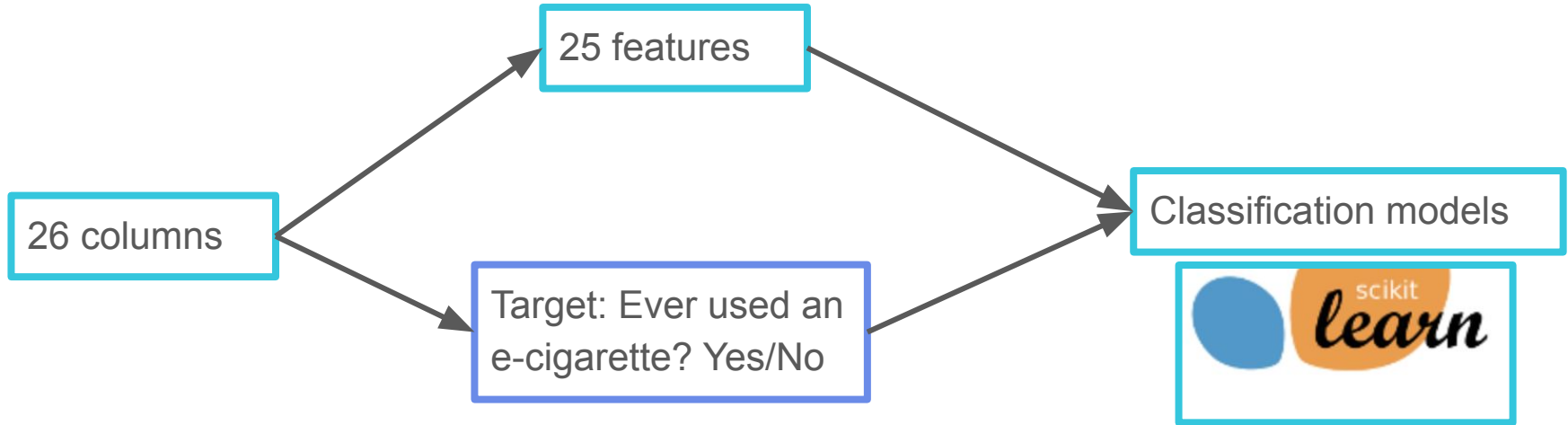
- 994 columns
- 14,531 rows (students)

Feature
reduction &
cleaning

 pandas 

- 26 columns
- 13,114 rows (students)

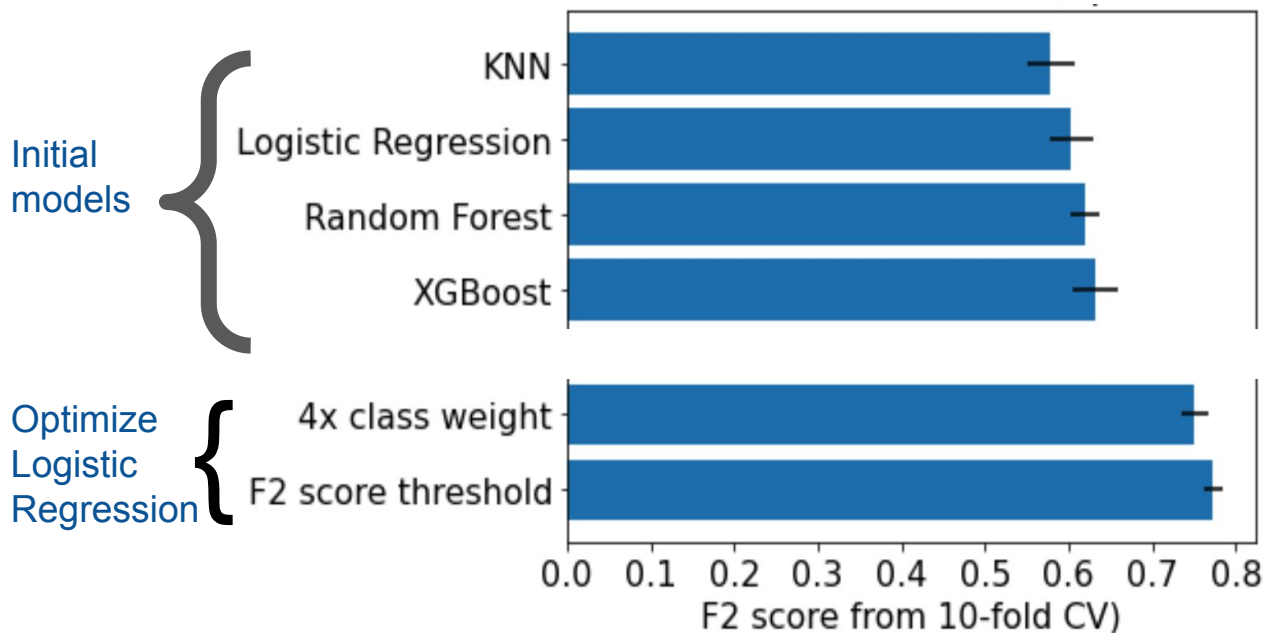
Data & Methods



24.4%

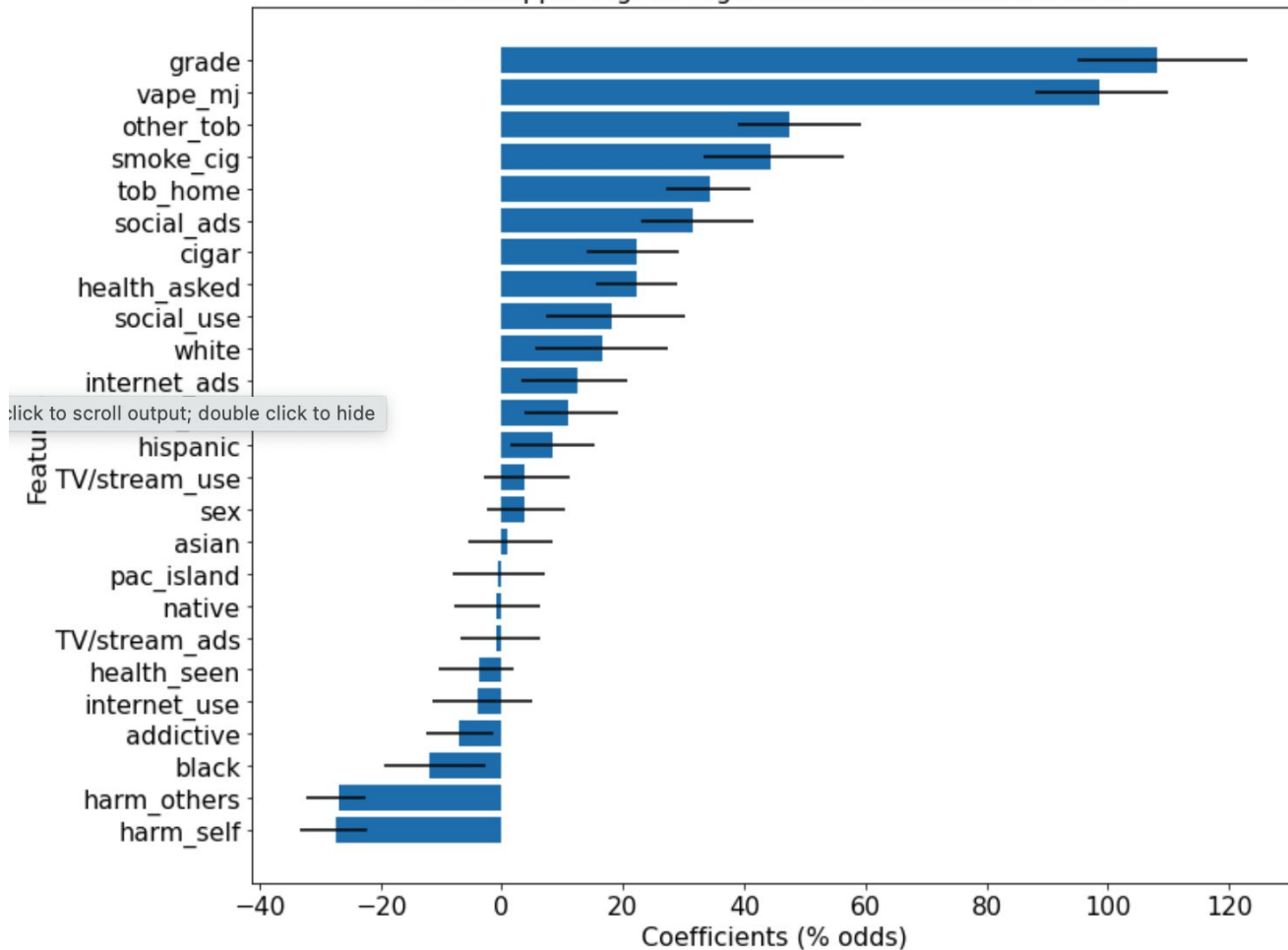
of surveyed students
answered YES

Model selection & optimization



- Primary metric: F2 score (to prioritize recall over precision)
- Results: Initial models were comparable
- Chose to optimize logistic regression because of its interpretability
- Final model F2 score: 0.736

Bootstrapped logistic regression coefficients with 95% CI



- Most highly associated with e-cig use:
 - Grade in school
 - Cannabis vape use
- Negative association: sense of potential harm to self or others

Conclusion

- Optimized a classification model using 2020 National Youth Tobacco Survey data
- Answered the question: what factors are associated with e-cig use in teens?

May be useful for:

- Parents
- Healthcare professionals



Thank you!