Preventing Opioid Abuse

Team Artificial Intel-Agents

Corey Duncan | Ayodeji Ejigbo | Sarah Roe

Presentation Outline

01 Problem Statement

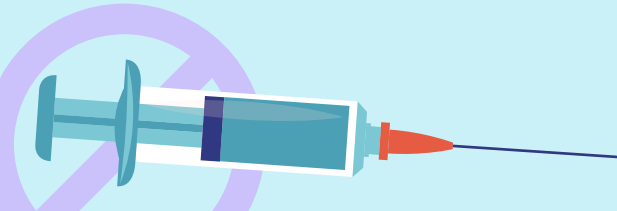Is it possible to build a machine learning model that helps us identify which patients are likely to abuse prescription pain medications using a patient's drug usage history?

Can this information be used to help reduce the chances of patients becoming addicted while simultaneously helping patients at lower risk not see a lapse in their medication?

O2   Why is this Important?

- Many addictions start with prescription opioids.

- Due to the high cost and difficulty of obtaining prescription pain medication, many people turn to cheaper, stronger opioids.

- Fentanyl is one such opioid, that has killed more young Americans (18-45) than Covid-19 in 2021-2022.

- In America, fentanyl fatally poisons someone every 8.57 minutes, for 175 deaths per day.

03 Objective

Our project aims to develop a machine learning model to predict the likelihood of a patient abusing pain medications. This information would allow doctors to:

- Make decisions on what types of pain meds to give, and how much at a time

- Prevent cases of addiction and/or relapse

- Intervene early, create treatment plans and overall improve quality of care for all patients

O4 Ethical Considerations

Outright denying narcotic pain medications based on potential abuse is wrong and not our model's goal. Instead, we aim to aid doctors in making informed decisions on pain medication types and dosages, providing these benefits:

1 Low risk patients could receive more medication at once so they don't see a lapse

2 Moderate risk patients could be given less medication upfront and be offered additional support/monitoring

3 Severe risk patients could be offered non-narcotic alternatives

05

EDA & Models

EDA – Corey

Missing/Mislabeled Data
Dropped columns with incorrect values

Checked Correlations
Explored which features were worth using in our model
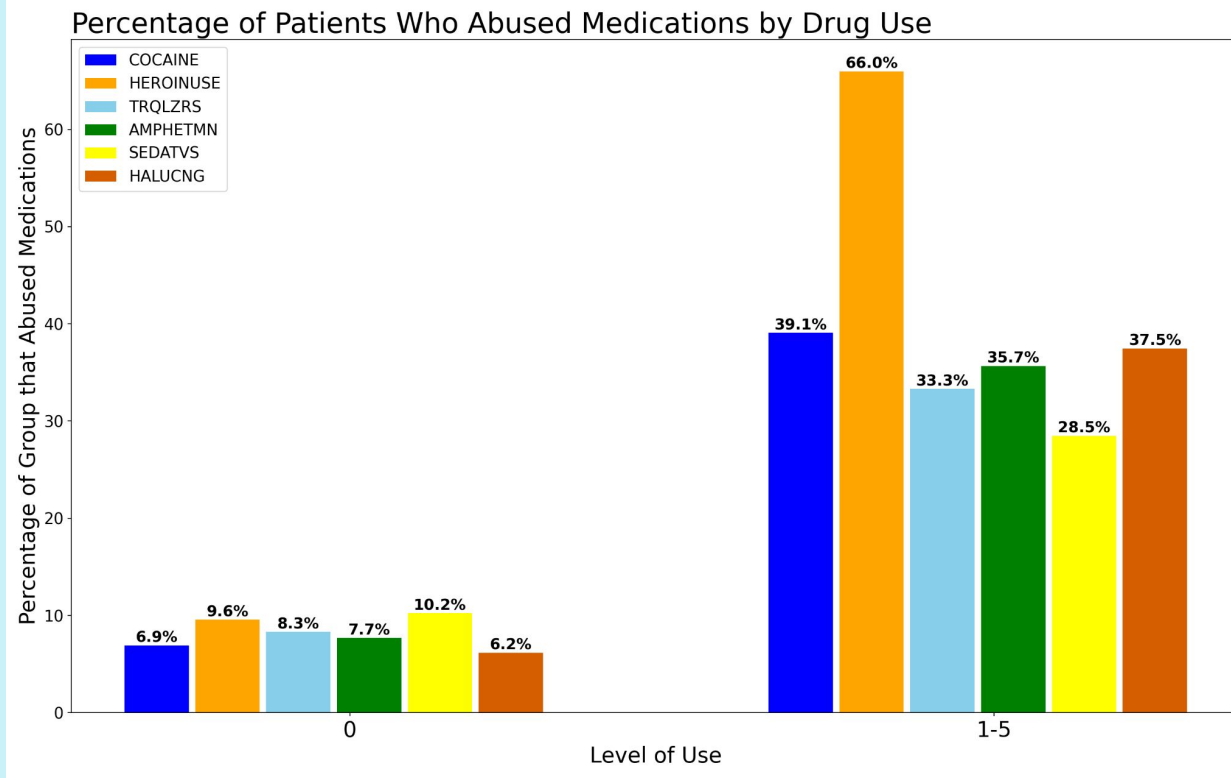
01 — 02 — 03 — 04

Imbalanced Classes
Used oversampling of minority class to fix imbalance

Interaction Terms
Created new columns such as Uppers, Downers, Overall Health

What role does recreational drug use play?

Percentage of Patients Who Abused Medications by Drug Use

Legend:
- COCAINE
- HEROINUSE
- TRQLZRS
- AMPHETMN
- SEDATVS
- HALUCNG

Y-axis: Percentage of Group that Abused Medications
X-axis: Level of Use

Level 0:
- 6.9%
- 9.6%
- 8.3%
- 7.7%
- 10.2%
- 6.2%

Level 1-5:
- 39.1%
- 66.0%
- 33.3%
- 35.7%
- 28.5%
- 37.5%

This graph shows the percentage of patients who abused pain medications and their drug use in the past year.

66% of patients with reported heroin use in the last year also abused pain medications.

Does level of use make a difference?

Ratio of Patients that Abuse Medications by Number of Drug Types Used
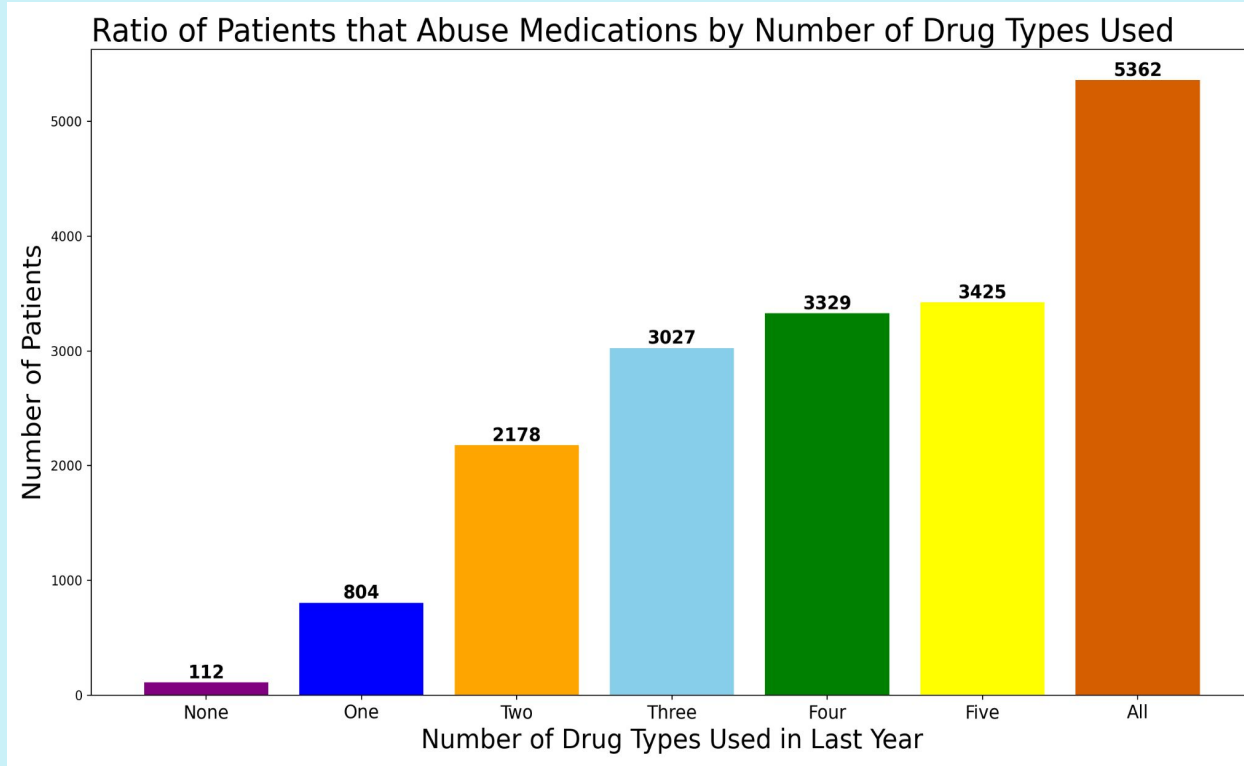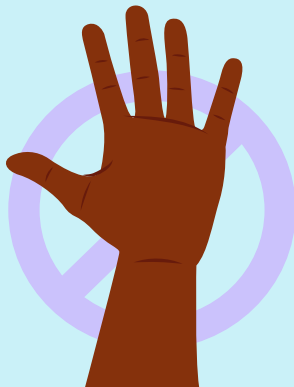

This graph represents all our patients who abused pain medications, grouped by how many drug types they've abused in the last year.

Models – Corey

Tried a variety of models, including

- Random Forest Variations (0.888, Best)
- Logistic Regression (0.749)
- Neural Nets (0.758)
- SVM (N/A)
- Voting Classifier Variations w/ Gradient Boost, Extreme Gradient Boost, Random Forest, etc. (0.847)

Best Model
Random Forest w/ PCA & GridSearch

- Include interaction terms
- Included features with missing values
- Standard Scaled
- GridSearchCV for best params
- PCA

Score: 0.888	Predicted No Abuse	Predicted Abuse
Actual No Abuse	35093	3006
Actual Abuse	5486	32455

EDA – Ayo

Missing/Mislabeled Data
Dropped columns that had incorrect values

Checked Correlations
Explored potential variables to focus on

01　**02**　**03**　**04**　**05**

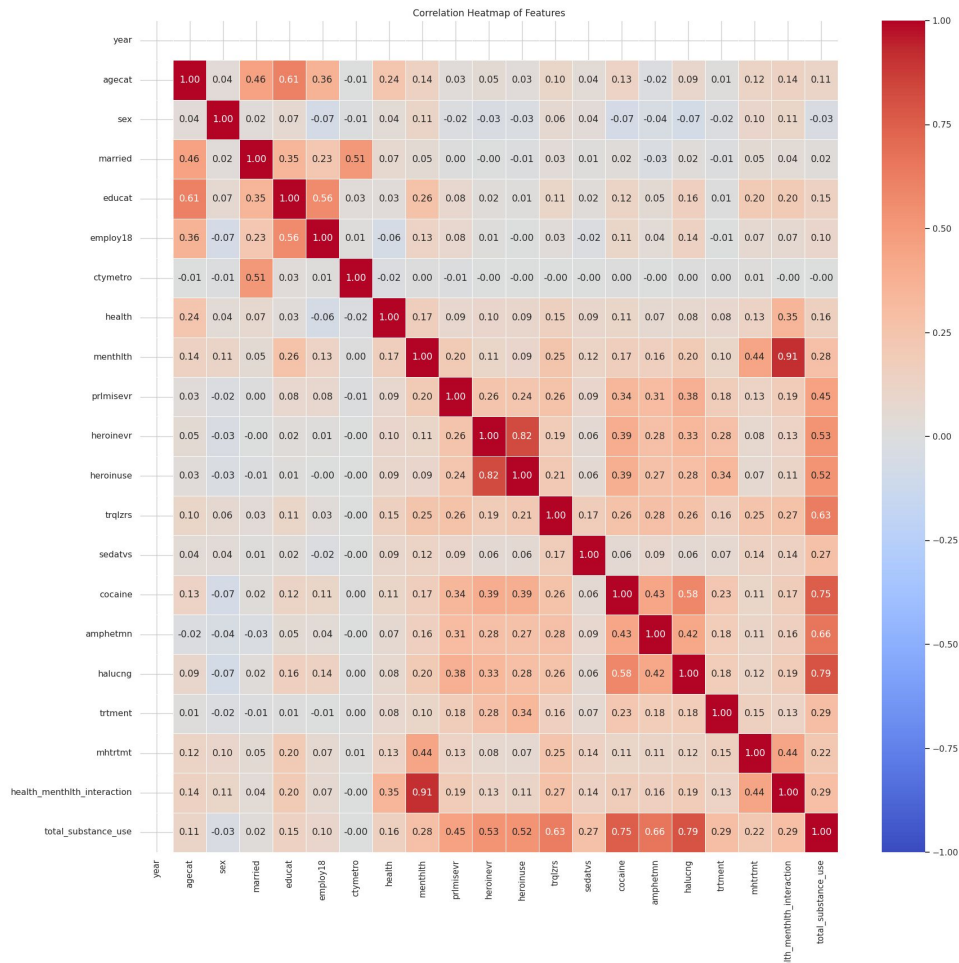**Imbalanced Classes**
Used SMOTE to fix imbalance

Feature Engineering
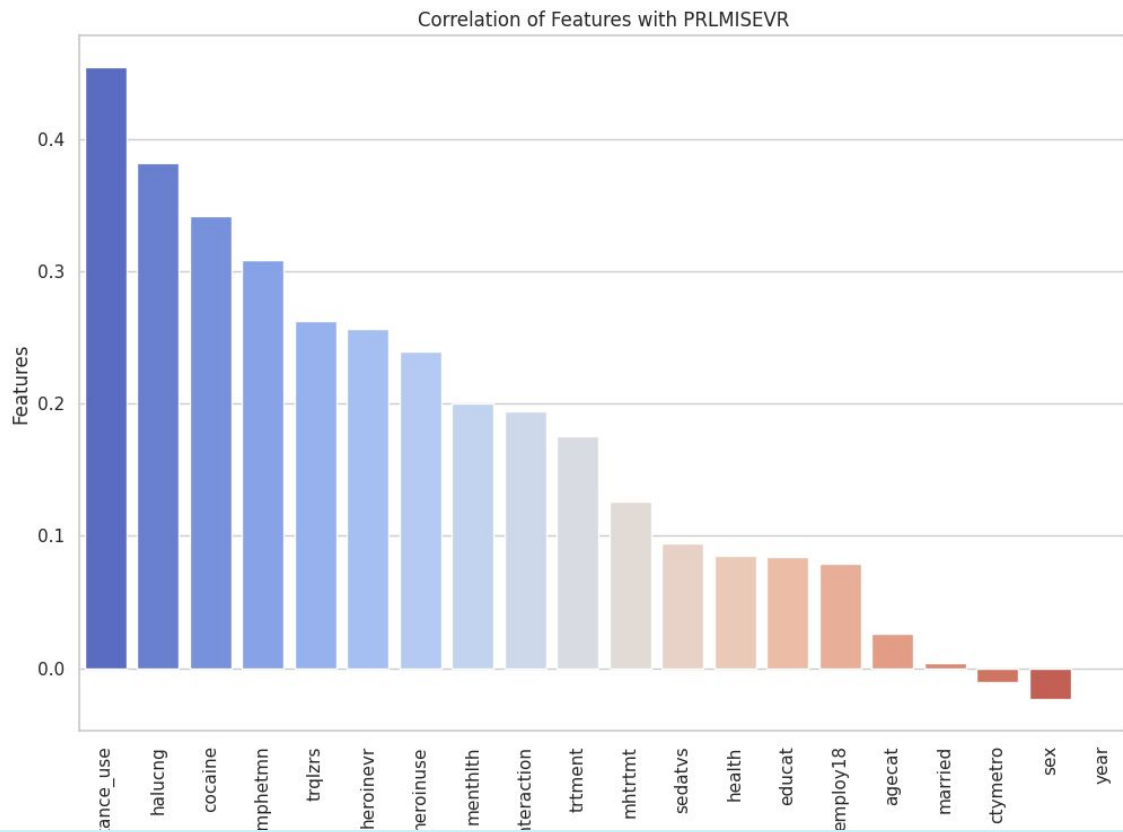Creating new features to improve performance

Principal Component Analysis
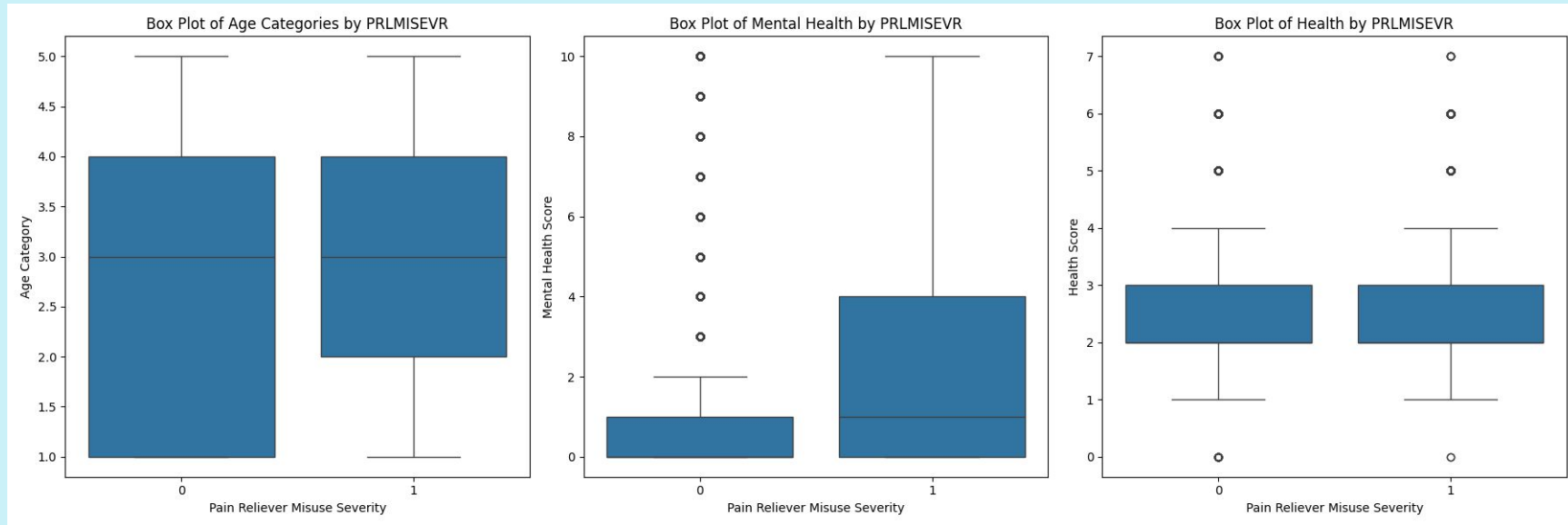Fit a PCA model to understand what variables were significant

Heatmap showing the relationship between all variables

HOW DOES EACH FEATURE CORRELATE WITH DRUG MISUSE SEVERITY?

Correlation of Features with PRLMISEVR

Distribution of age categories, health, and mental health scores for different levels of drug misuse severity (`prlmisevr`).

Models – Ayo

- LogisticRegression(max_iter=1000) - (0.908)
- RandomForestClassifier() - (0.902)
- KNeighborsClassifier() - (0.894)
- SVM - (0.902)
- Neural Nets - (0.903)

Logistic Regression

- max_iter=1000
- PCA (n_components=5)
- Precision: 92%
- Recall: 89%
- F-1 Score = 90.5%

EDA – Sarah

Missing/Mislabeled Data

Dropped columns that had incorrect values

PCA Feature Importance

Fit a PCA model to understand what variables were significant
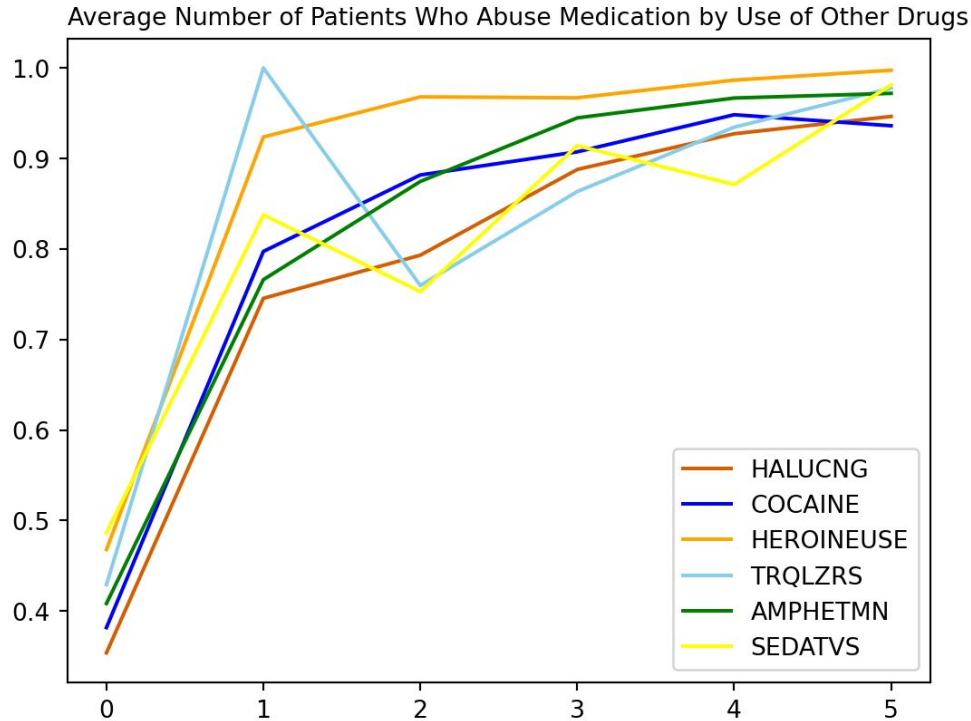
01 — **02** — **03** — **04**

Imbalanced Classes

Used bootstrapping to fix imbalance

Checked Correlations

Explored potential variables to focus on

How does Drug Use impact Drug Addiction

Average Number of Patients Who Abuse Medication by Use of Other Drugs

Legend:
- HALUCNG
- COCAINE
- HEROINEUSE
- TRQLZRS
- AMPHETMN
- SEDATVS

This graph shows a general upward trend indicating as amount of drug use increases, the likelihood of becoming addicted to drugs also increases

PCA Feature Importance

Variable Importance
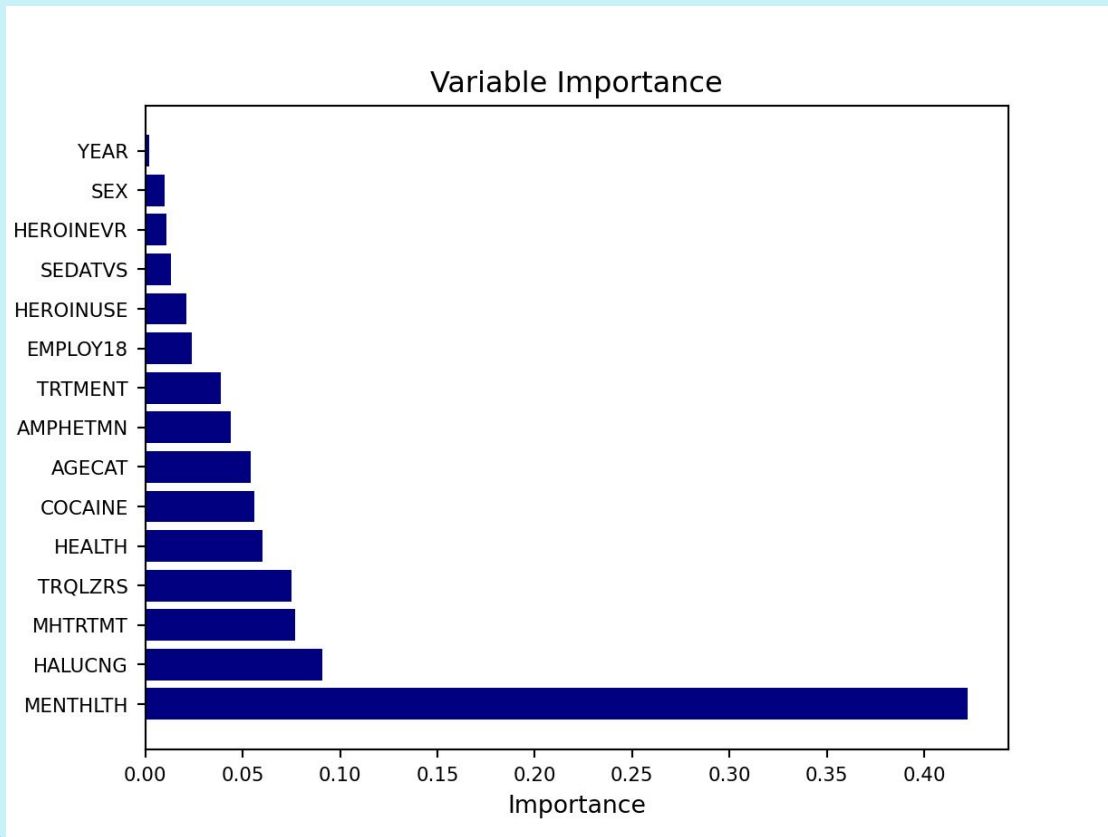
Mental Health is shown to have the highest feature importance

Models – Sarah

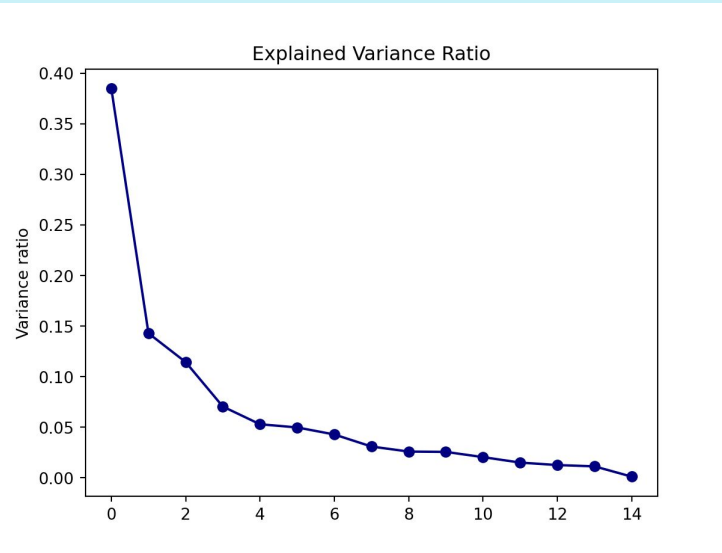Tried a variety of models, including

- **Random Forest Variations (0.85, Best)**
  - With/Without PCA
- **Neural Nets (0.753)**
- **Logistic Regression (0.742)**
  - With/Without PCA
- **Naive Bayes (0.738)**

Best Model
Random Forest w/ PCA

- N_estimators=600
- Max_features=9
- Max_depth=30
- PCA (n_components=5)

Explained Variance Ratio

Limitations/Next Steps

- Hard to fully understand patterns based solely on self reported data
- Understanding the risk of False Negatives
- Data set, while large, had some incorrect values

- Collecting self-reported pain level to better understand 'need'

- Continually updating and re-fitting to new data

Conclusion

Enhancing Patient Safety

Personalized Pain Management

Reducing Addiction Risk

Improving Healthcare Efficiency

Evidence-Based Decision Making

Opioid Crisis

Work Cited

- National Institute on Drug Abuse: Drug Overdose Death Rates
  https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates#:~:text=Nearly%20108%2C000%20persons%20in%20the,drugs%20from%201999%20to%202022.
- Kaggle Dataset on Opioid Abuse
  https://www.kaggle.com/datasets/thedevastator/predicting-pain-reliever-misuse-abuse/data
- Dataset Documentation:
  https://zenodo.org/records/2301844#.Y8OqptJBwUE
- Opioid Crisis in Young Americans:
  https://murphy.house.gov/media/press-releases/murphy-fentanyl-killing-more-young-americans-covid-19

Thanks