

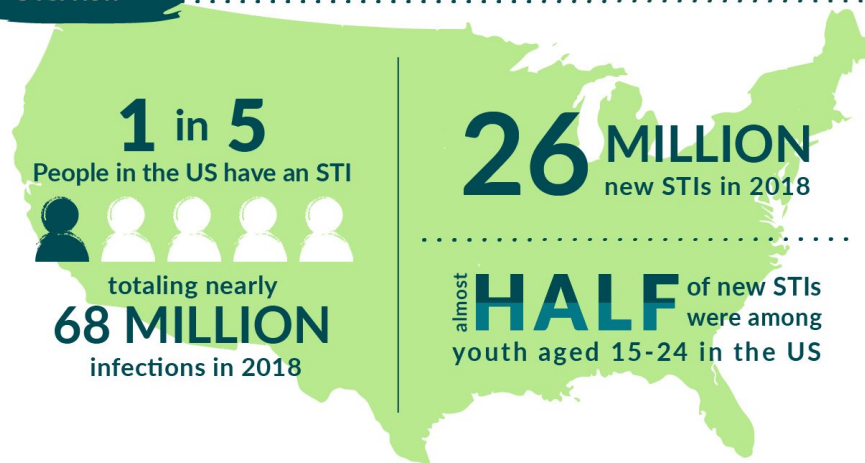


Predicting Antibiotic Resistance in Gonorrhea

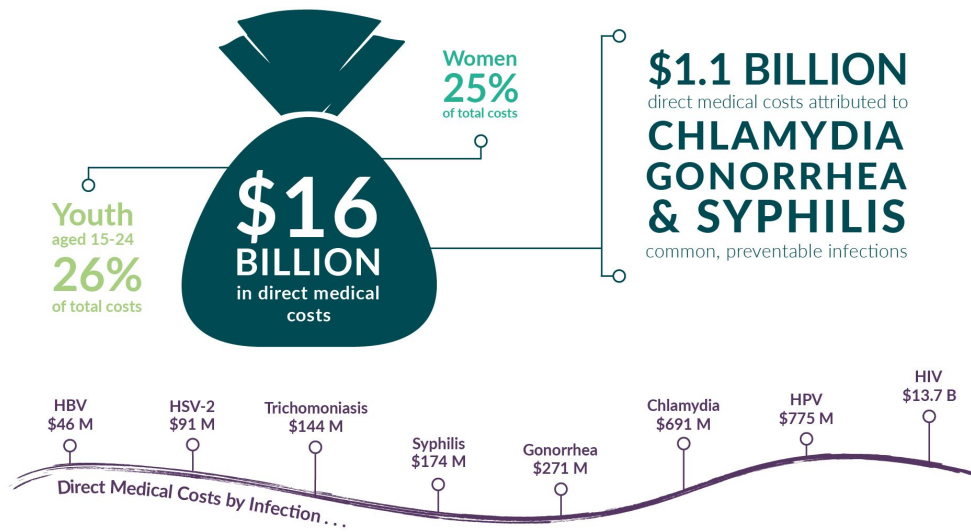
Brendan Nugent



The Problem



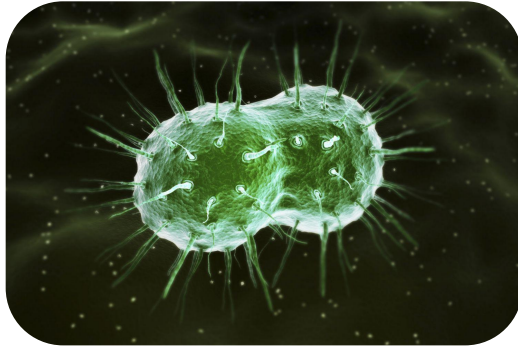
New STIs
total nearly **\$16**
BILLION
in direct medical
costs



**1,568,000 new infections per
year in the US**

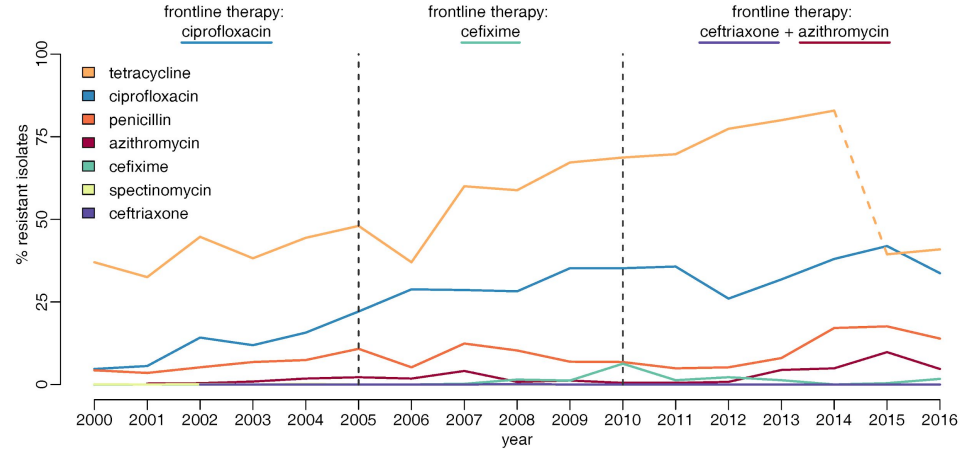
**\$271 million in medical costs
per year**

The Problem



N. gonorrhoeae

<https://www.verywellhealth.com/symptoms-of-gonorrhea-2329059>



Epidemiological Trends of Antibiotic Resistant Gonorrhoea in the United Kingdom by Whittles et al.

HEALTH

New 'Superbug' Strain of Gonorrhea Is Outsmarting Most Antibiotics

Two cases in Massachusetts involve a novel strain more impervious to existing antibiotics than other strains in the U.S.

The Opportunity

Cheap and fast DNA sequencing

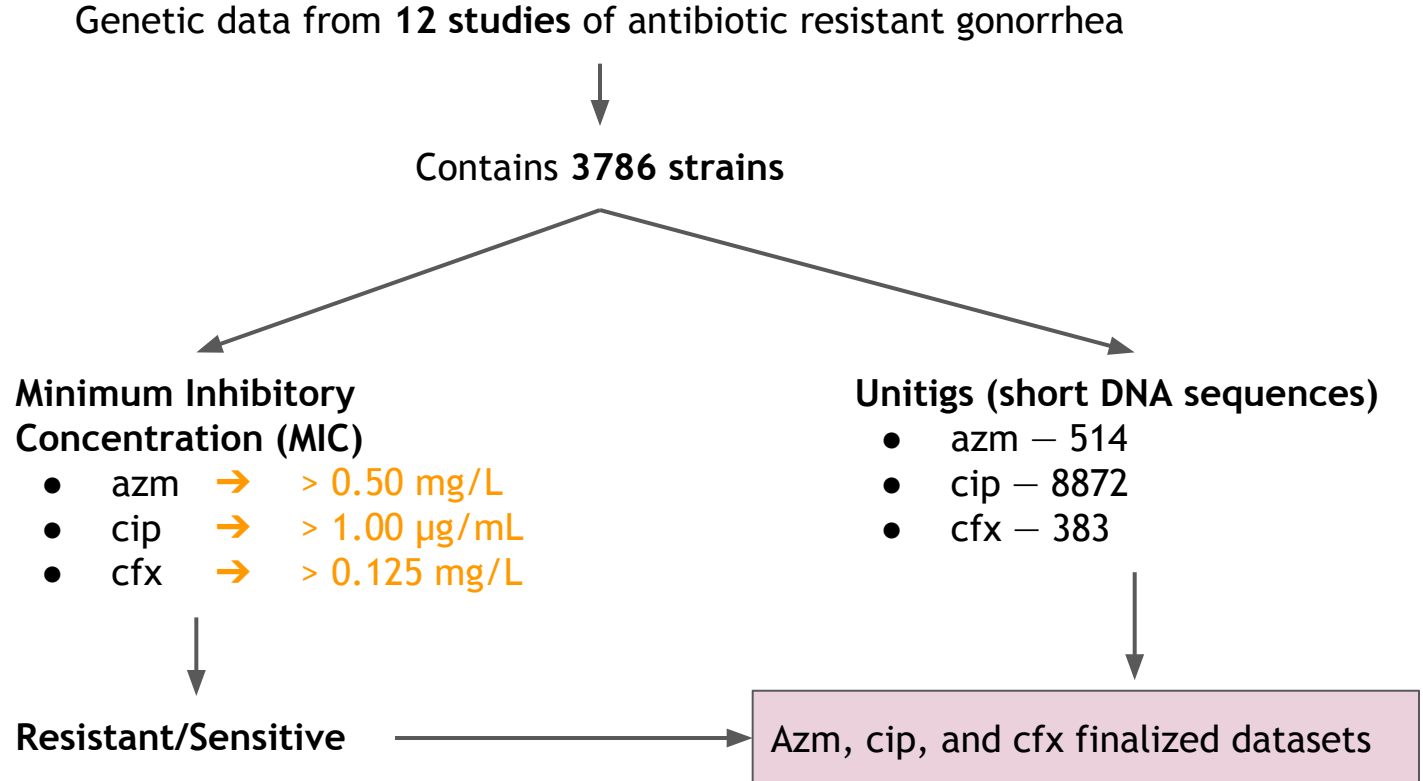
Identification of effective antibiotics

Develop models that predict a strain of gonorrhea's resistance to antibiotics

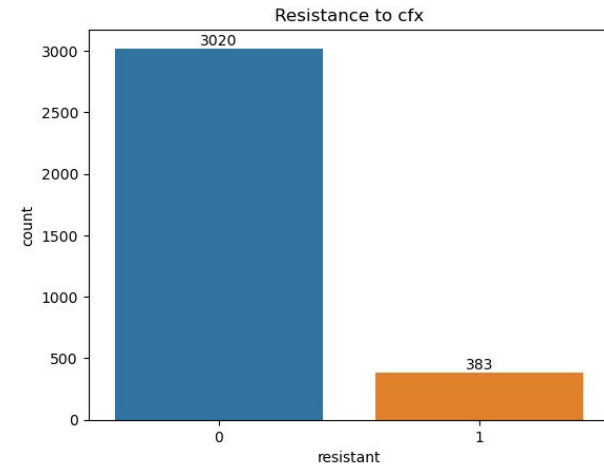
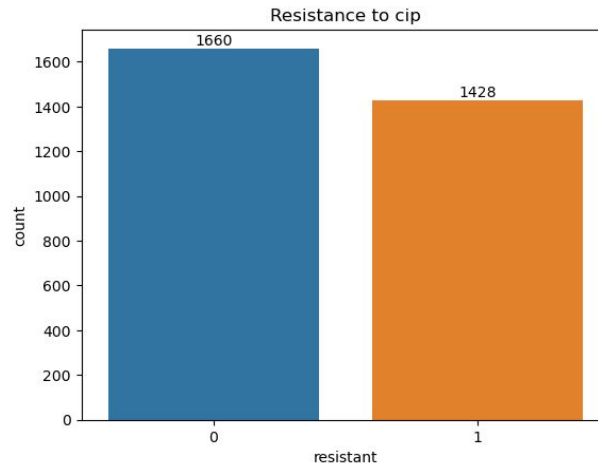
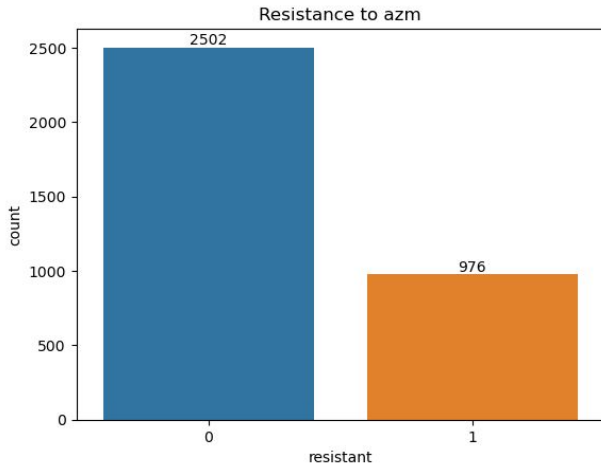
- Azithromycin (azm)
- Ciprofloxacin (cip)
- Cefixime (cfx)



The Data

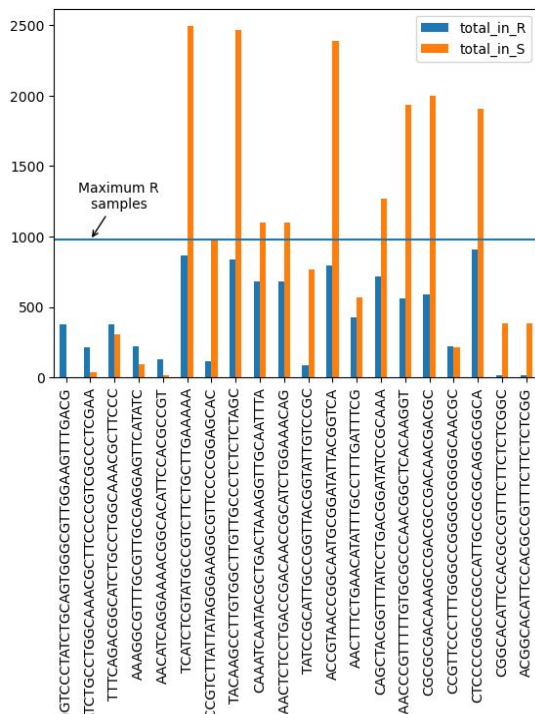


Prevalence of Antibiotic Resistance

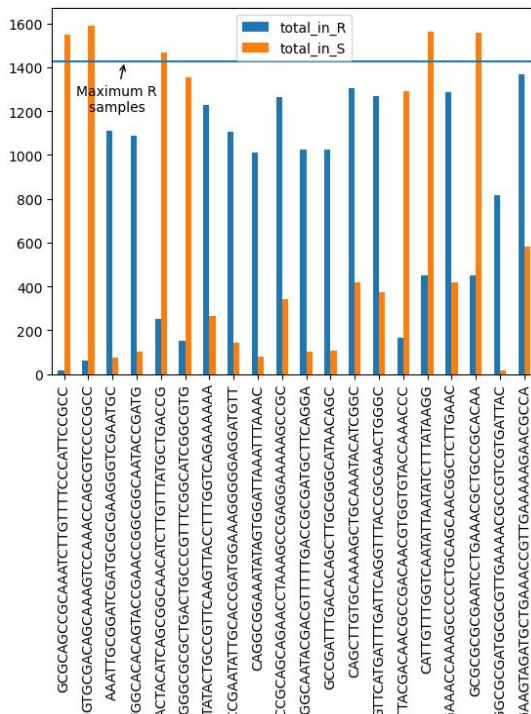


Unitigs Most Correlated with Resistance

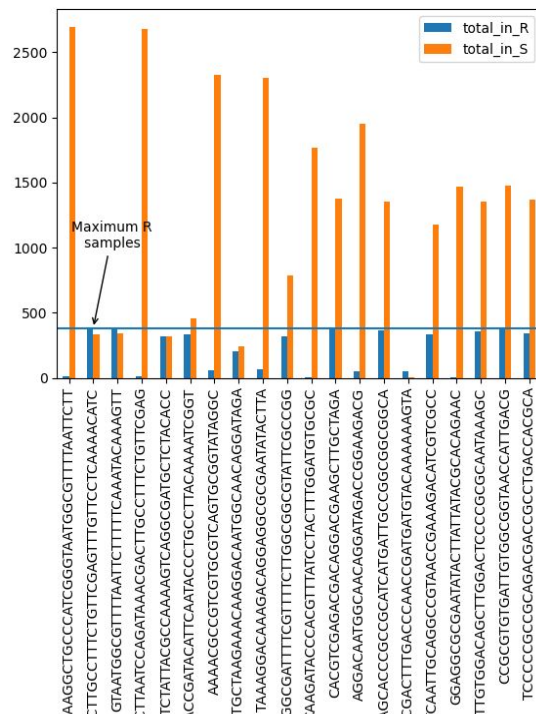
azm: least correlation of all antibiotics



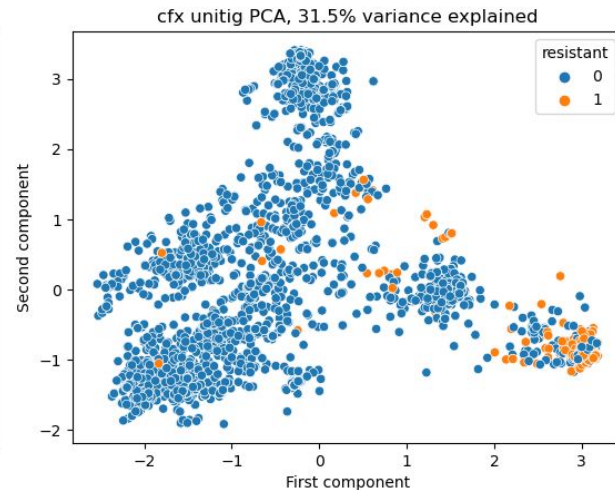
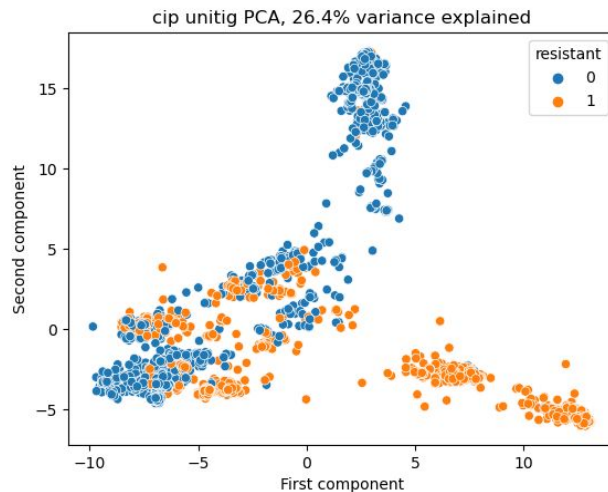
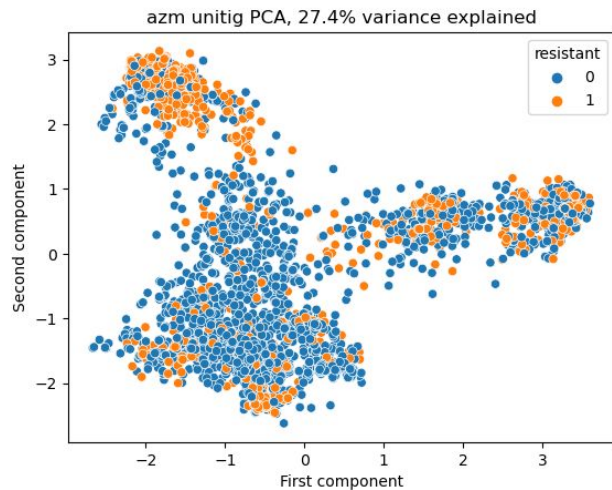
cip: highest correlation of all antibiotics



cfx: high correlation given low prevalence of resistance



PCA



Model Selection

Models:

- Random Forest Classifier
- XGBoosting Classifier

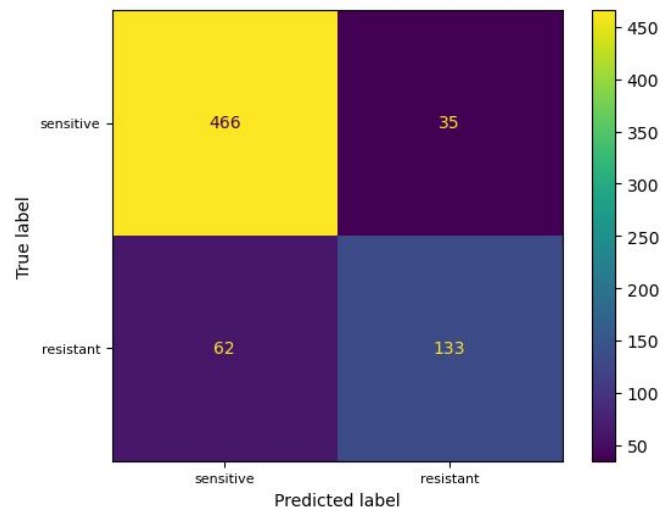
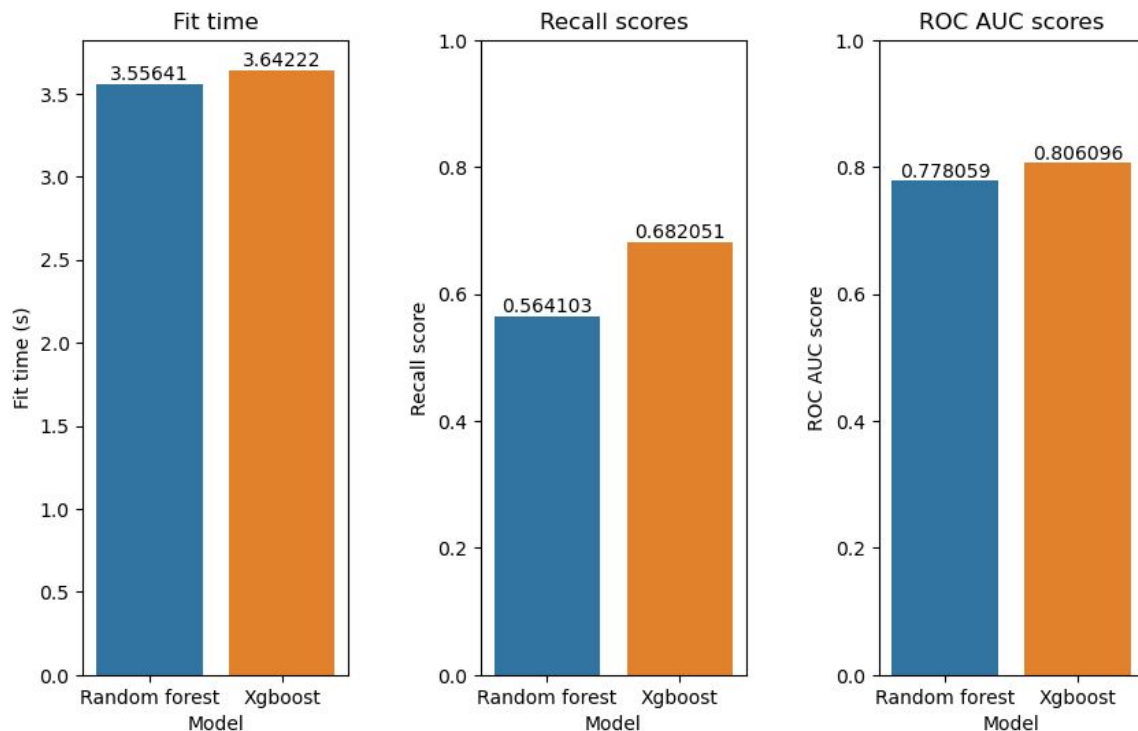
Train/Test Splits:

- 80/20
- 70/30

Conducted a random grid search with 5 fold cross validation for each model on each antibiotic dataset for each split, optimizing for recall.

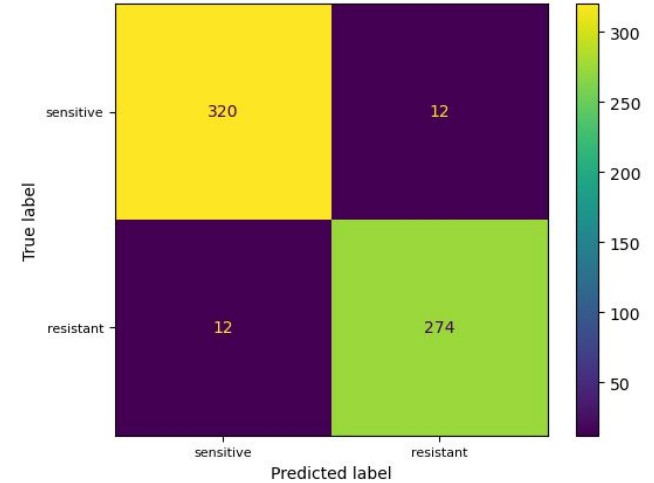
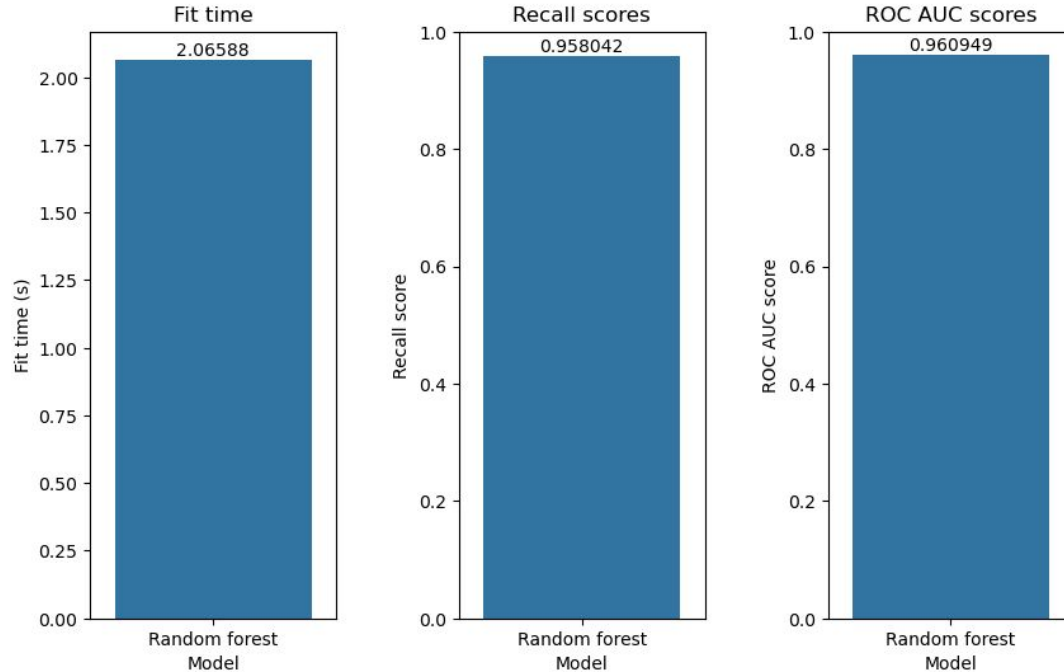
Final azm Model

Model analysis for azm



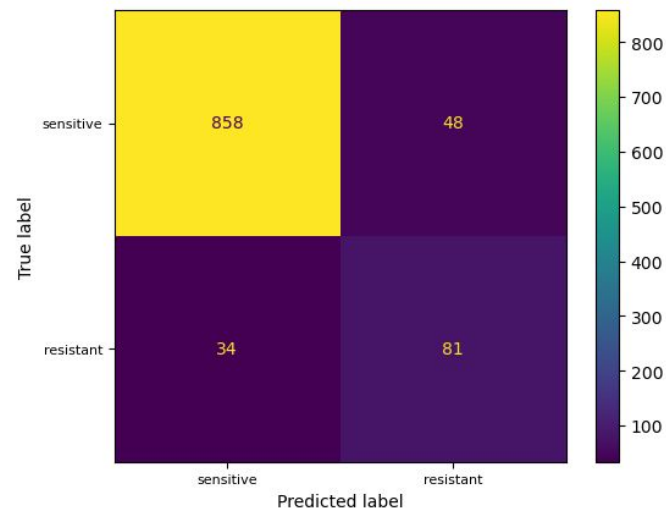
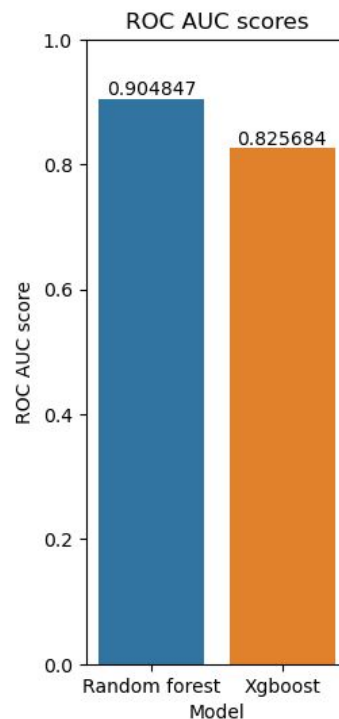
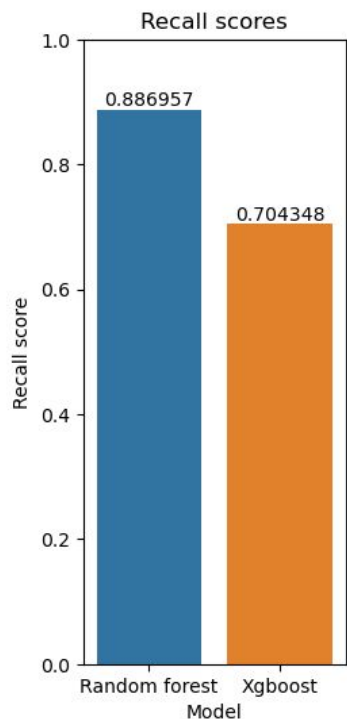
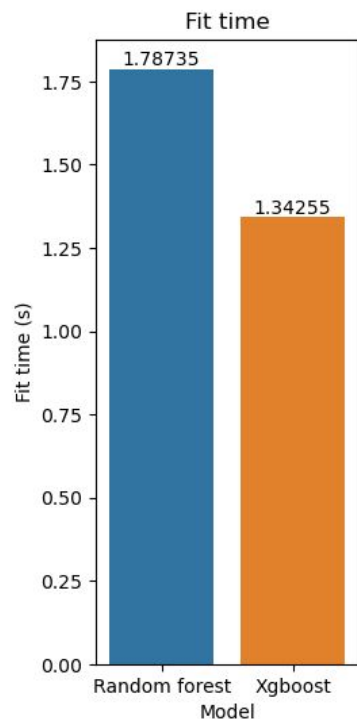
Final cip Model

Model analysis for cip



Final cfx Model

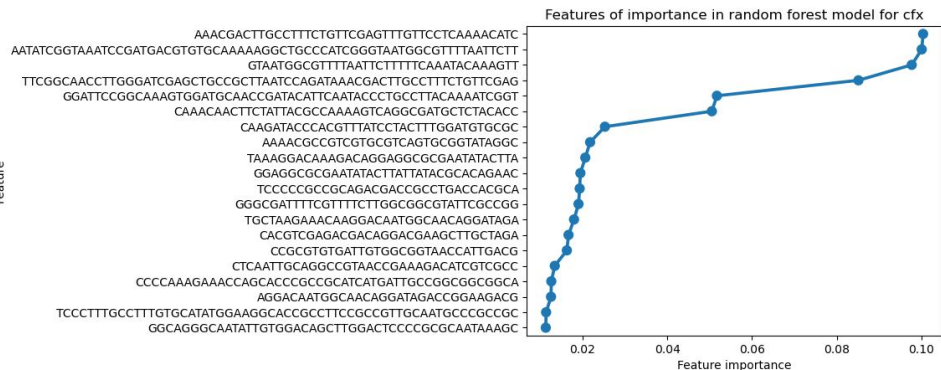
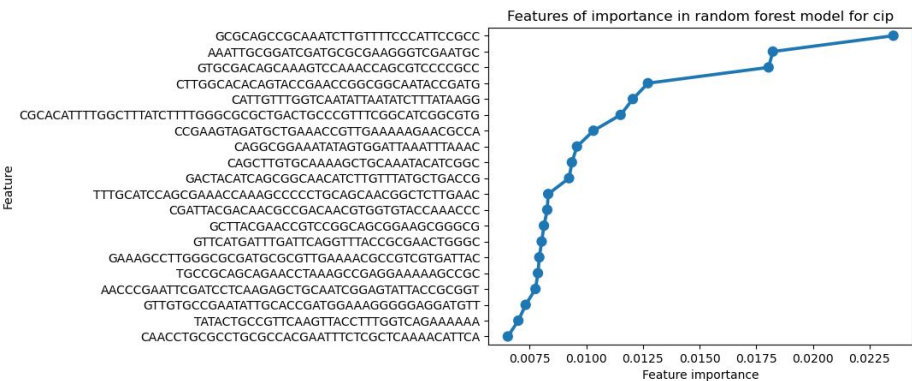
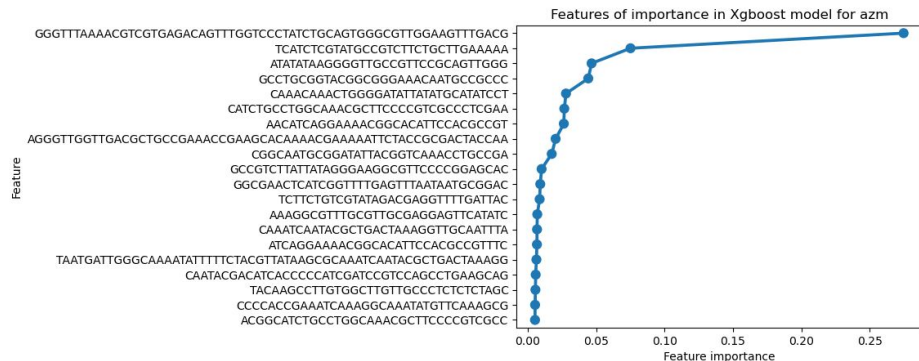
Model analysis for cfx



Feature Importance

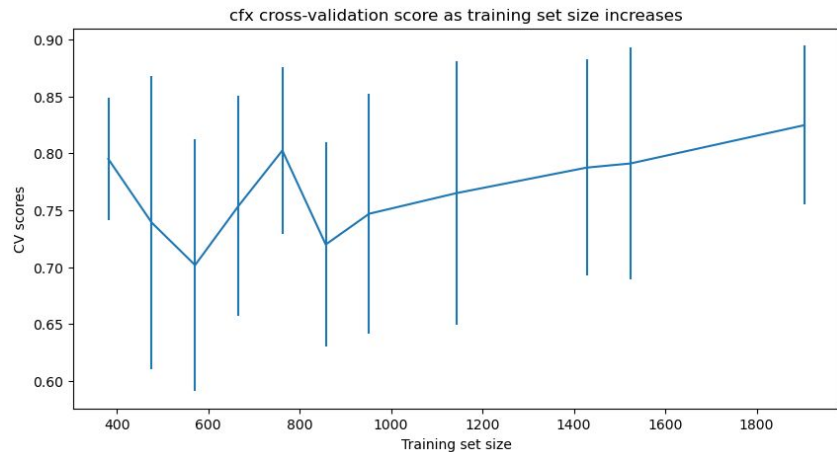
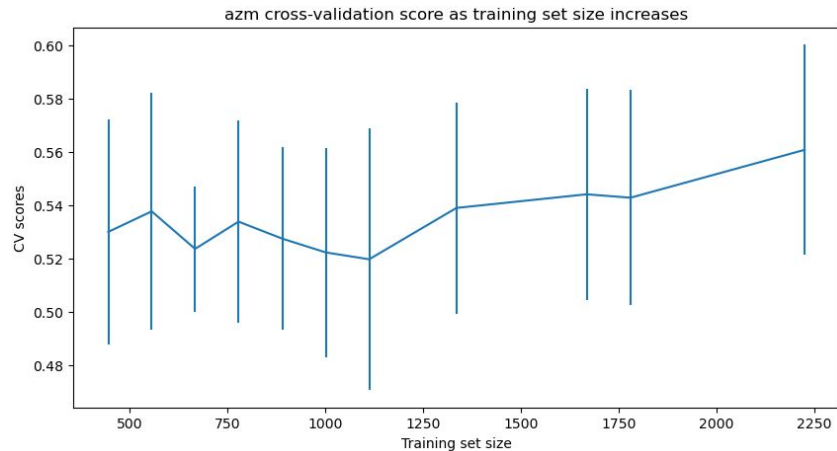
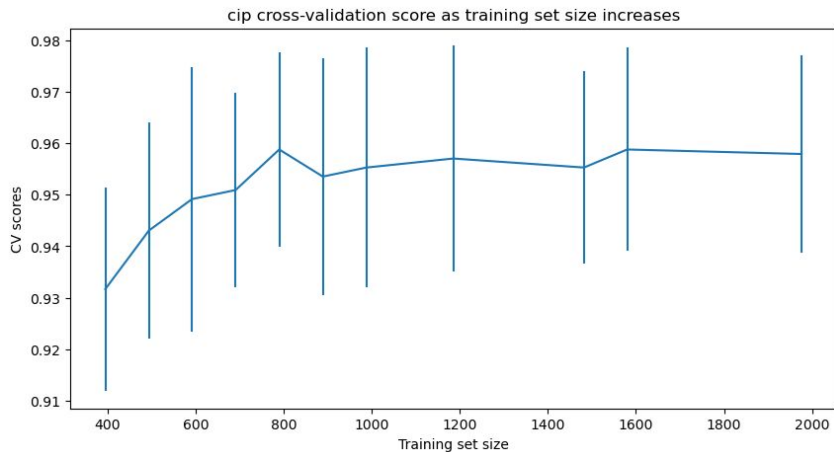
Features matched the unitigs with high Chi-squared values.

azm relied heavily on one unitig whereas the others had multiple strong predictors.

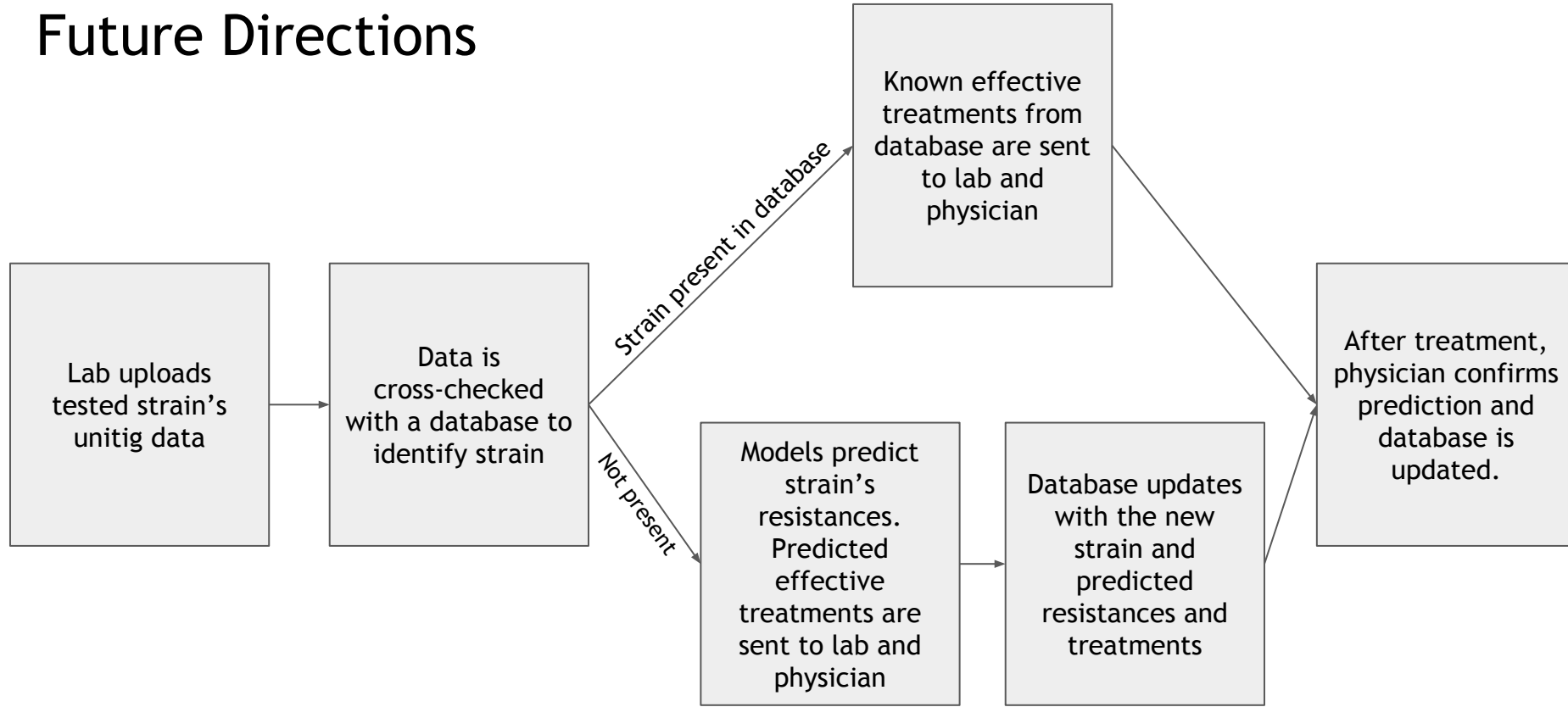


Learning Curves

azm and cfx could benefit from larger training sets.



Future Directions



Future Directions

Improve the models:

- Further feature engineering
- More data
- Exploring other models
(AdaBoost)

Apply this methodology to other
antibiotics and other bacterial infections

Utilize the most important features in
the models to explore genes causing
the resistance to develop drugs to
further combat antibiotic resistance

Thank you!

Brendan Nugent

Email: bjnugent@mac.com

LinkedIn: <https://www.linkedin.com/in/bjnugent/>

Github: <https://github.com/bjnugent>

Project report: https://github.com/bjnugent/antibiotic_resistance_in_gonorrhea/blob/main/reports/final_report_antibiotic_resistance.pdf