Introduction

- Antimicrobial resistance, or AMR, is what we call the development of antibiotic immunity in microbes and pathogens, and has quickly become one of the greatest threats to global health
  - The GRAM Study, published in the Lancet in 2022, predicts nearly 40 million AMR deaths by 2050
- ML is a form of AI that allows computers to learn otherwise obscure patterns and trends in data
  - We could potentially leverage this to create a clinical tool for optimizing treatments and unveiling key insights about AMR onset

Investigative Question(s)

- Can we utilize machine learning to program a tool that can be used to rapidly predict AMR levels prior to treatment faster than previous models based on clinical data inputs?
- Could such a model be used to identify and rank the strongest signal producers for AMR for different antibiotics?

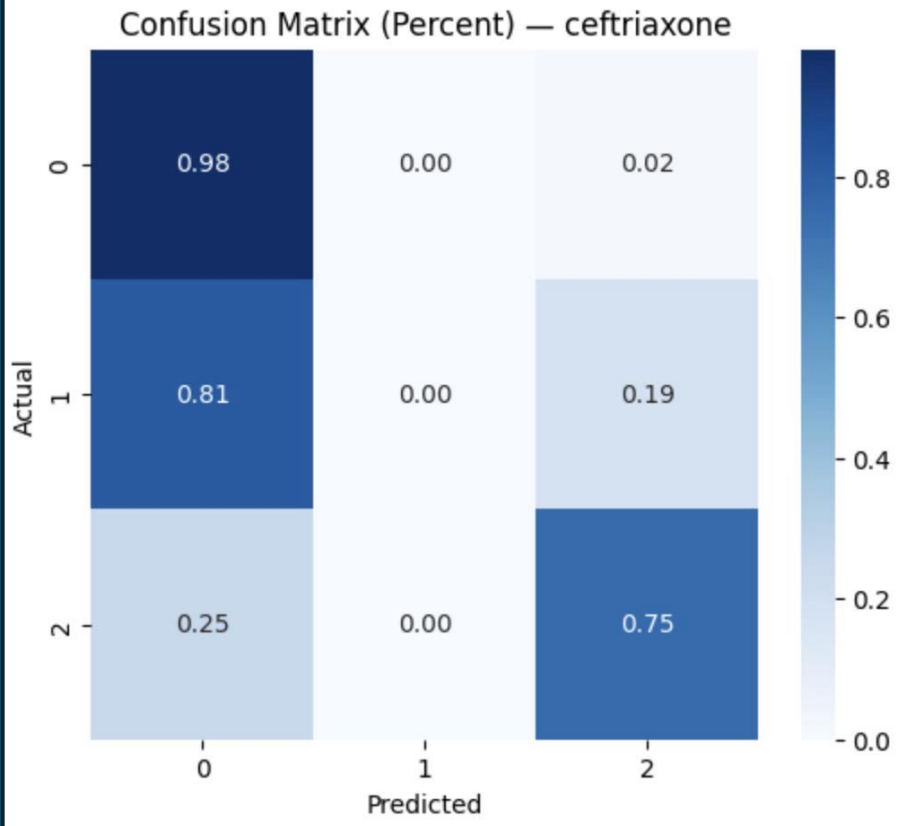
- Hypotheses: “yes” for both questions

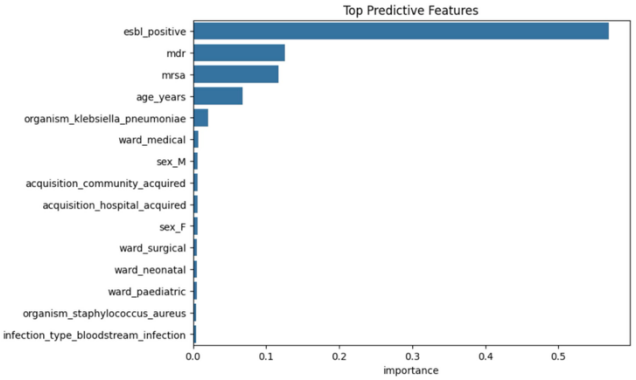
Methodology/Steps

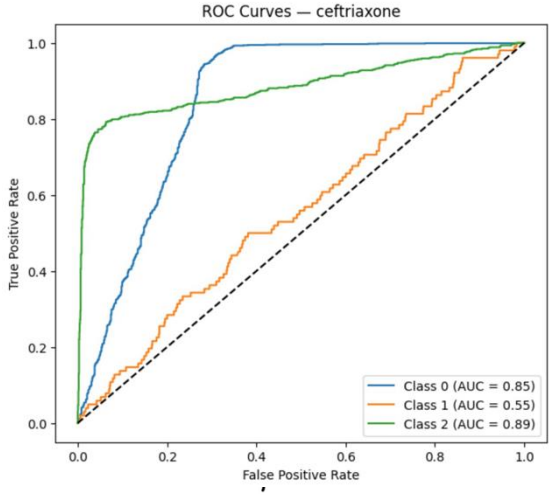
- 1. Data collection
- 2. Cleaning/Curation
- 3. Feature Selection
- 4. Data Encoding
- 5. Build Random Forest
- 6. Model Comparison
- 7. Hyper-Parameter/Threshold Tuning
- 8. Model Evaluation

Ceftriaxone

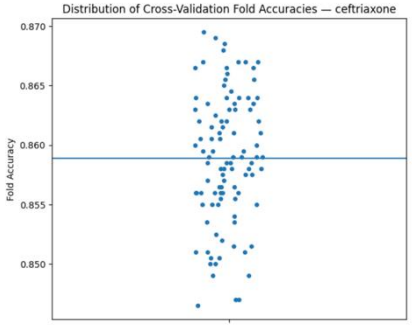
```
Model Performance:
Accuracy: 0.8605
Precision: 0.8214884440874686
Recall: 0.8605
F1 Score: 0.8351112925359583
```

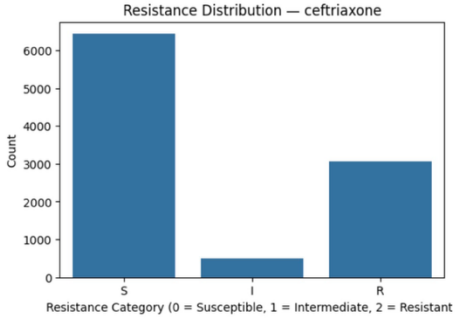






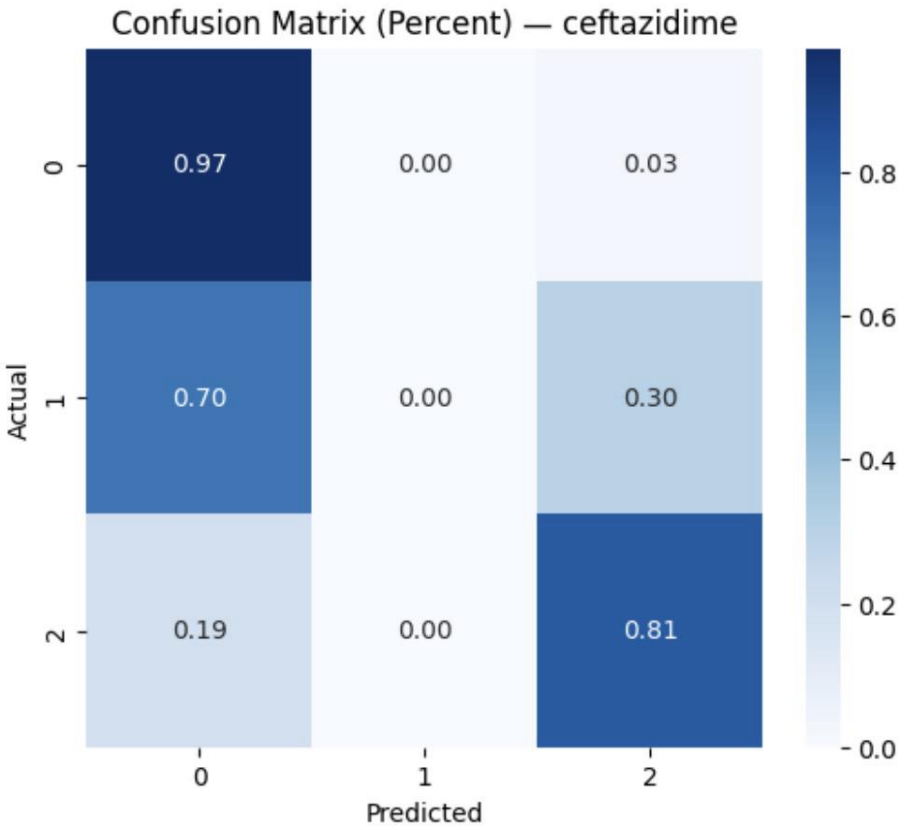

```
Overall ROC-AUC (ceftriaxone): 0.762
```

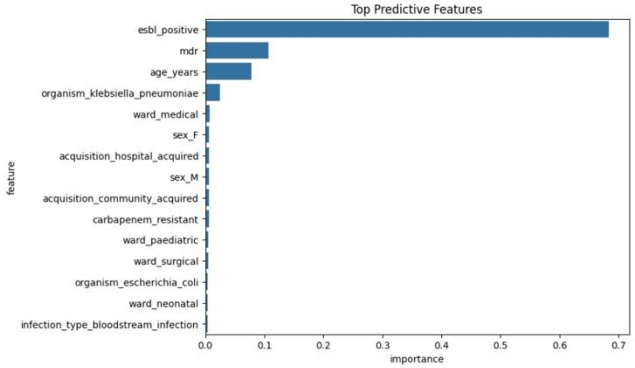


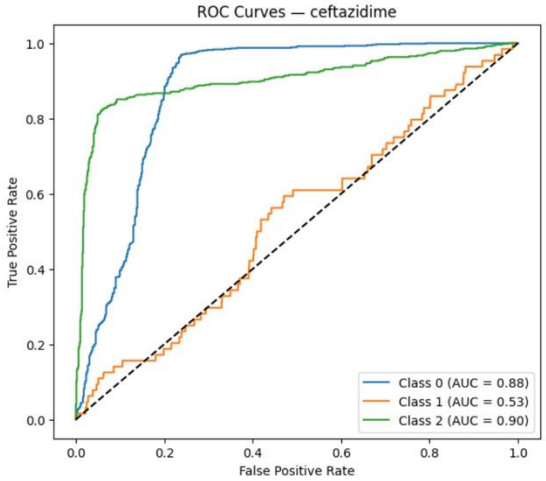


Ceftazidime

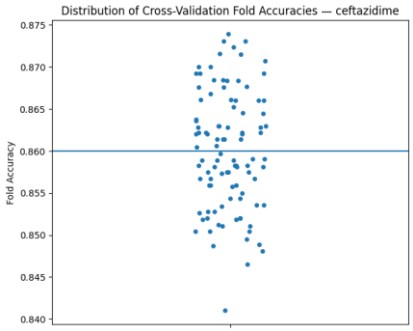
```
Model Performance:
Accuracy: 0.8598277212216131
Precision: 0.8213479509495073
Recall: 0.8598277212216131
F1 Score: 0.8363766912003828
```

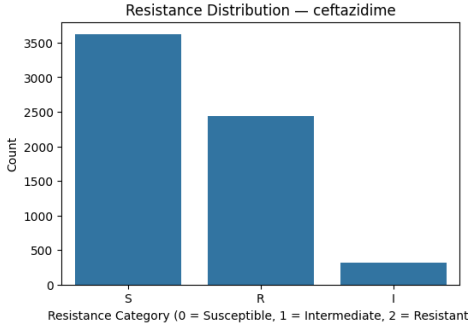






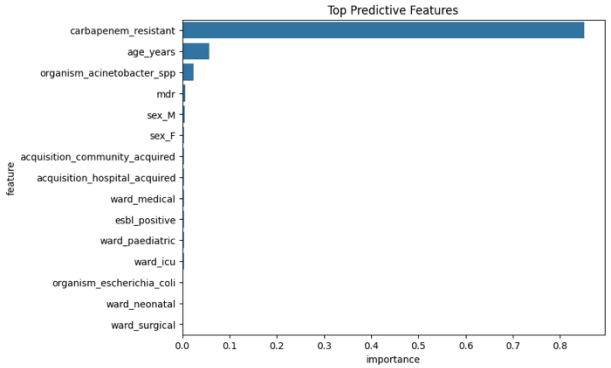

```
Overall ROC-AUC (ceftazidime): 0.769
```

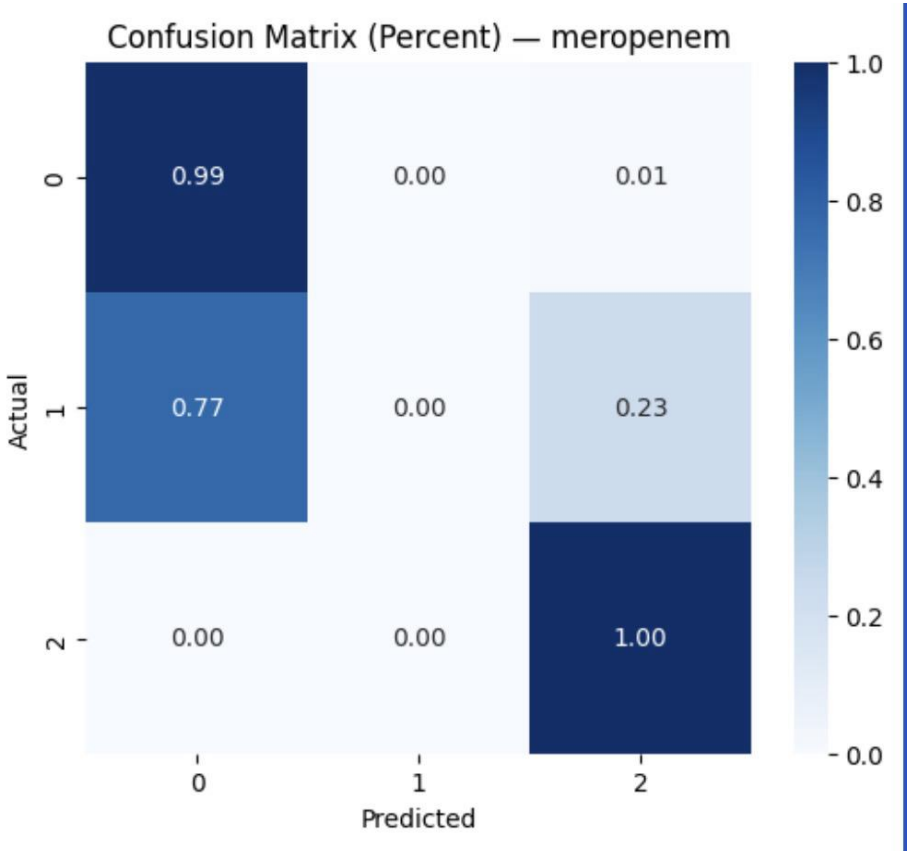


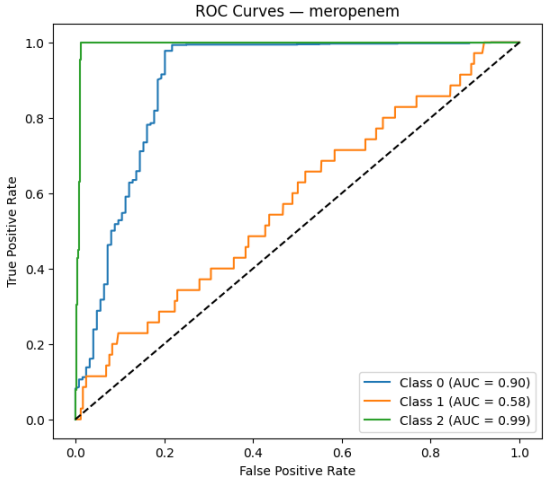


Meropenem

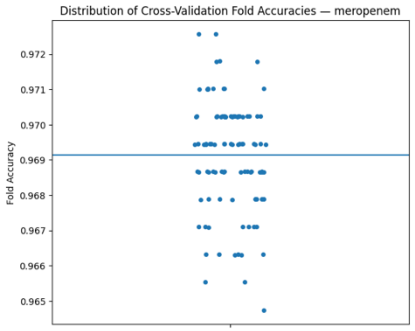
```
Model Performance:
Accuracy: 0.966327329678935
Precision: 0.9411713457703018
Recall: 0.966327329678935
F1 Score: 0.9532520004441043
```

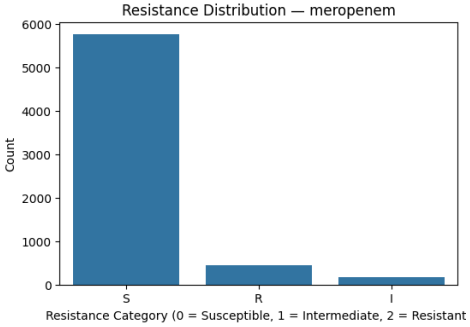







```
Overall ROC-AUC (ceftazidime): 0.769
```

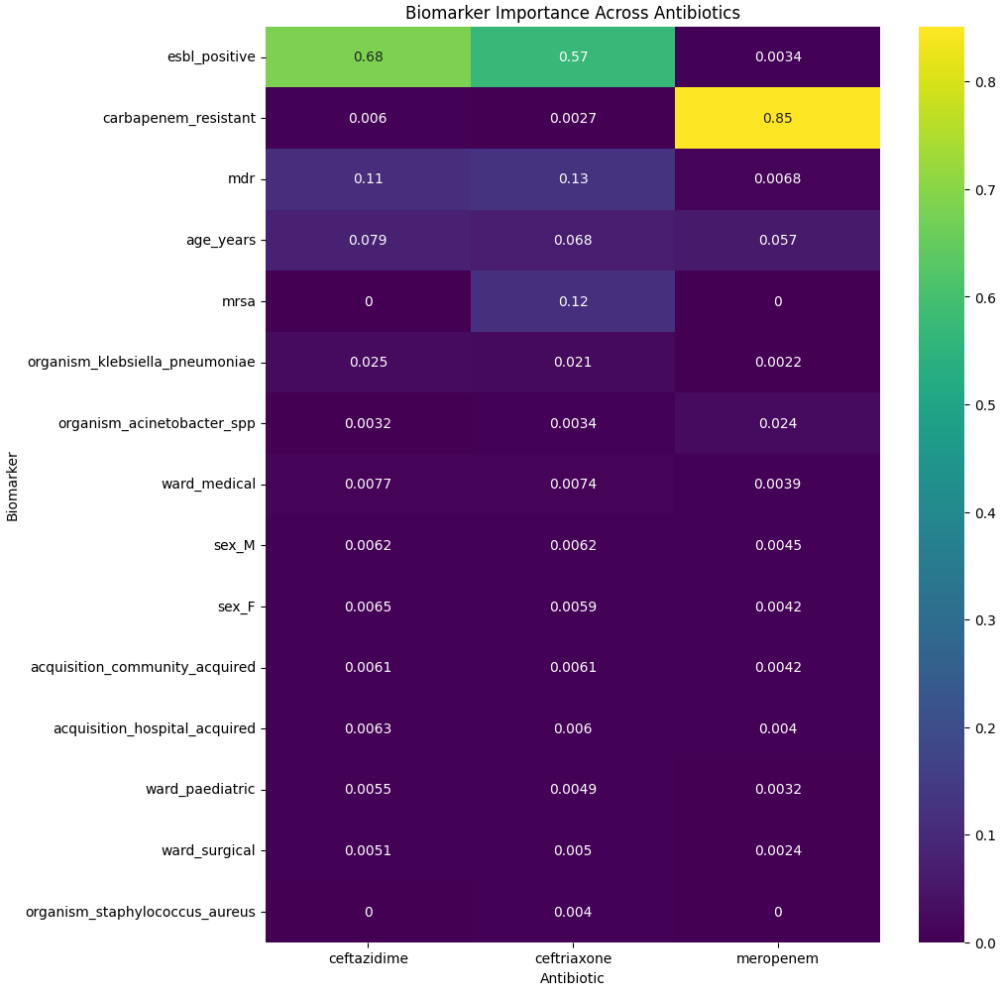




Results and Discussion

- Summary:

- Mean accuracy over several thousand test cases was approximately 90%
- Precision score and jitter chart affirm consistency and reliability
  - Precision above 0.8
  - Jitter chart SE = ~0.005
- Confusion Matrices, ROC Curves, and case distribution are all in alignment
  - Reveals bias against intermediate cases
    - Significantly lower accuracy than other cases
- Graph of predictive features reveals trends of markers per antibiotic



The model revealed key pairs of antibiotic and biomarker, indicated by the heat map; this includes ESBL value for ceftazidime and carbapenem resistance for meropenem. Other factors are shown to be not strong, which is also valuable.

Conclusion (Application, limitations, expansion)

- The developed model:
  - Successfully predicts AMR to a high degree of accuracy
    - This can be used in a clinical setting
      - Patients will receive more effective treatment
  - Revealed trends and biomarkers for AMR
    - Mapped certain antibiotics to specific characteristics
      - Provides key insight
      - Baseline for future research
  - Will help to slow overall mitigation of AMR onset

- This could contribute to saving millions of lives

- Errors and limitations:
  - Data skewness and scarcity
    - Large bias towards susceptibility

- Expansion:
  - This project can be used to produce new biomarkers for investigation
    - Potentially discover new biological mechanisms behind AMR
  - An optimization algorithm can be run using produced data
    - By figuring out what is resistant and what is not, we can run further algorithms to optimize sequences of treatments to objectively mitigate the onset of AMR