Towards personalised empirical antibiotic guidance - a pilot study

Dr B Collyer, University of Warwick
Dr E Moran, Heart of England NHS Trust
Prof M Keeling, University of Warwick

Predicting presence of resistant infections using patient data

1. **THE DATA**: from the Heart of England NHS trust
2. **THE METHOD**: machine learning classification algorithm
3. **PERFORMANCE**: how good are the predictions?
4. **IS THIS USEFUL?**

THE DATA:

Heart of England NHS Foundation trust has uniquely broad electronic data collected over last 8 years.

Including records from **antibiotic susceptibility tests**.

Are we able to predict the results of the tests for a patient on arrival?

THE DATA: Susceptibility testing

EUCAST antimicrobial disc diffusion susceptibility tests

Sample type	Date Collected	ORG	AMP	AUG	GT	TAZ	MEM	ETP	CAZ	CTX	TEM
Urine	2014-05-02	ECOL	S	S	S	S	S	S	S	S	S
Urine	2014-05-03	ECOL	R	R	S	S	S	S	S	S	S
Sputum	2014-04-30	ECOL	S	S	S	S	S	S	S	S	S
Urine	2014-05-02	ECOL	S	S	S	S	S	S	S	S	S
Blood	2014-05-02	MECOL	R	S	S	S	S	S	R	R	S

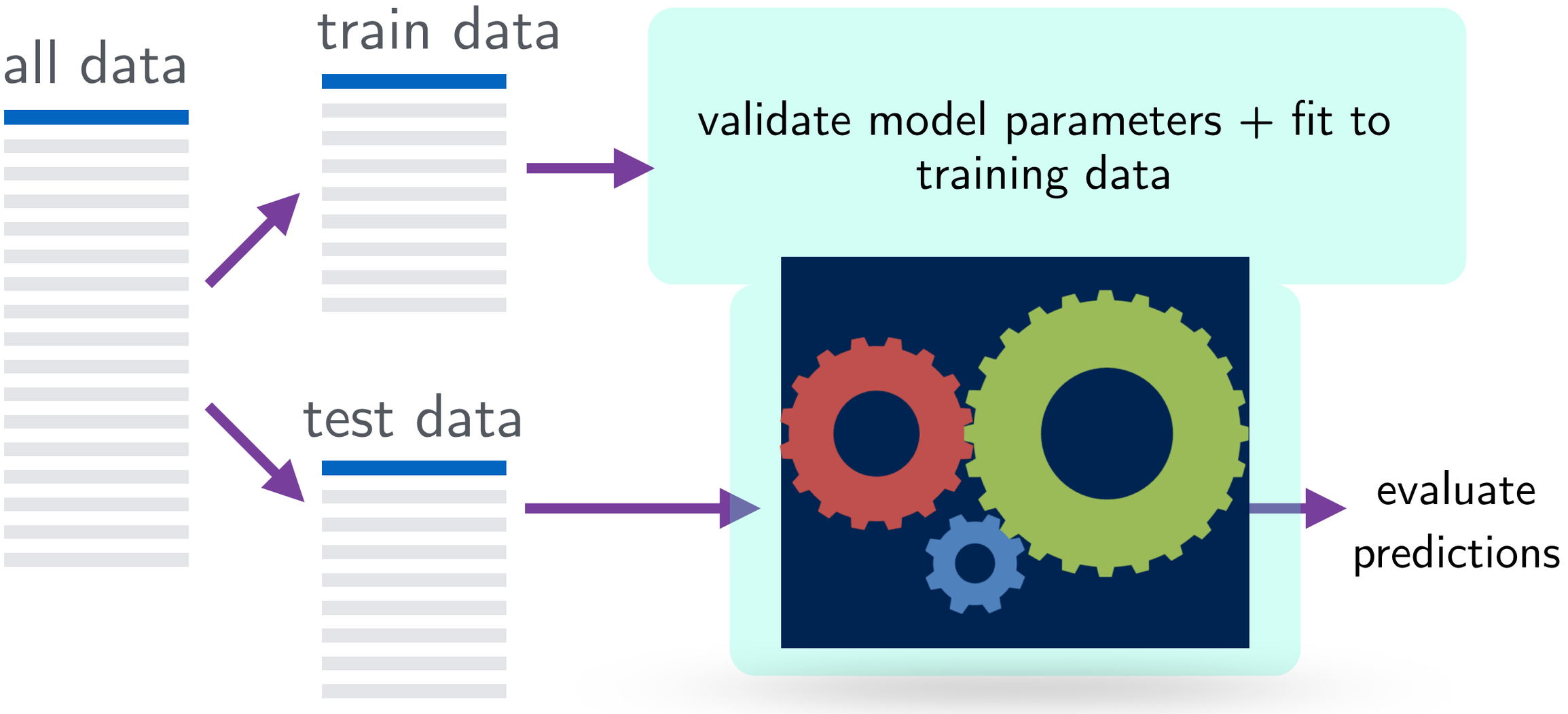
antibiotic test results: susceptible or resistant

What patient data can we use?

- **PERSONAL DATA**:
  - Age, ethnicity, co-morbidities (post-code, GP surgery)
- **ADMISSION DATA**:
  - Initial diagnosis, ward admitted to, previous wards visited, number of previous hospital visits, previous resistant infections
- **PRESCRIPTION DATA**
  - Antibiotics previously given in hospital

Machine learning

- input: 20,000 test results and hospital record matches

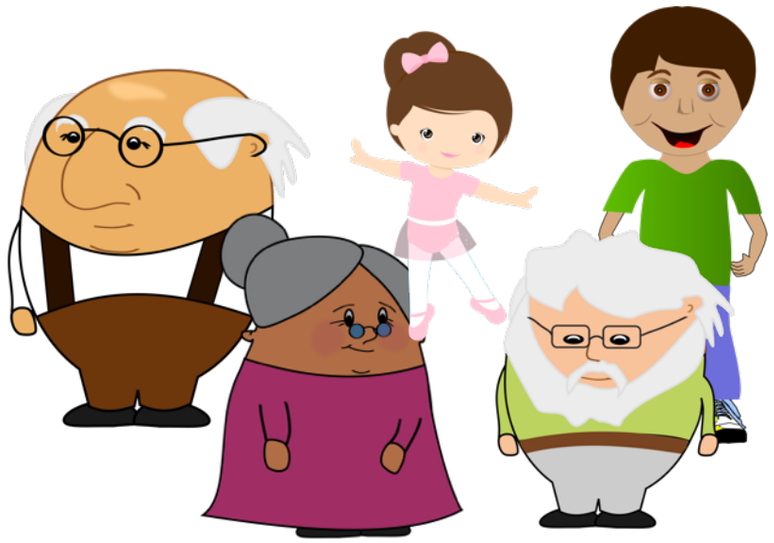


Which algorithm to use?

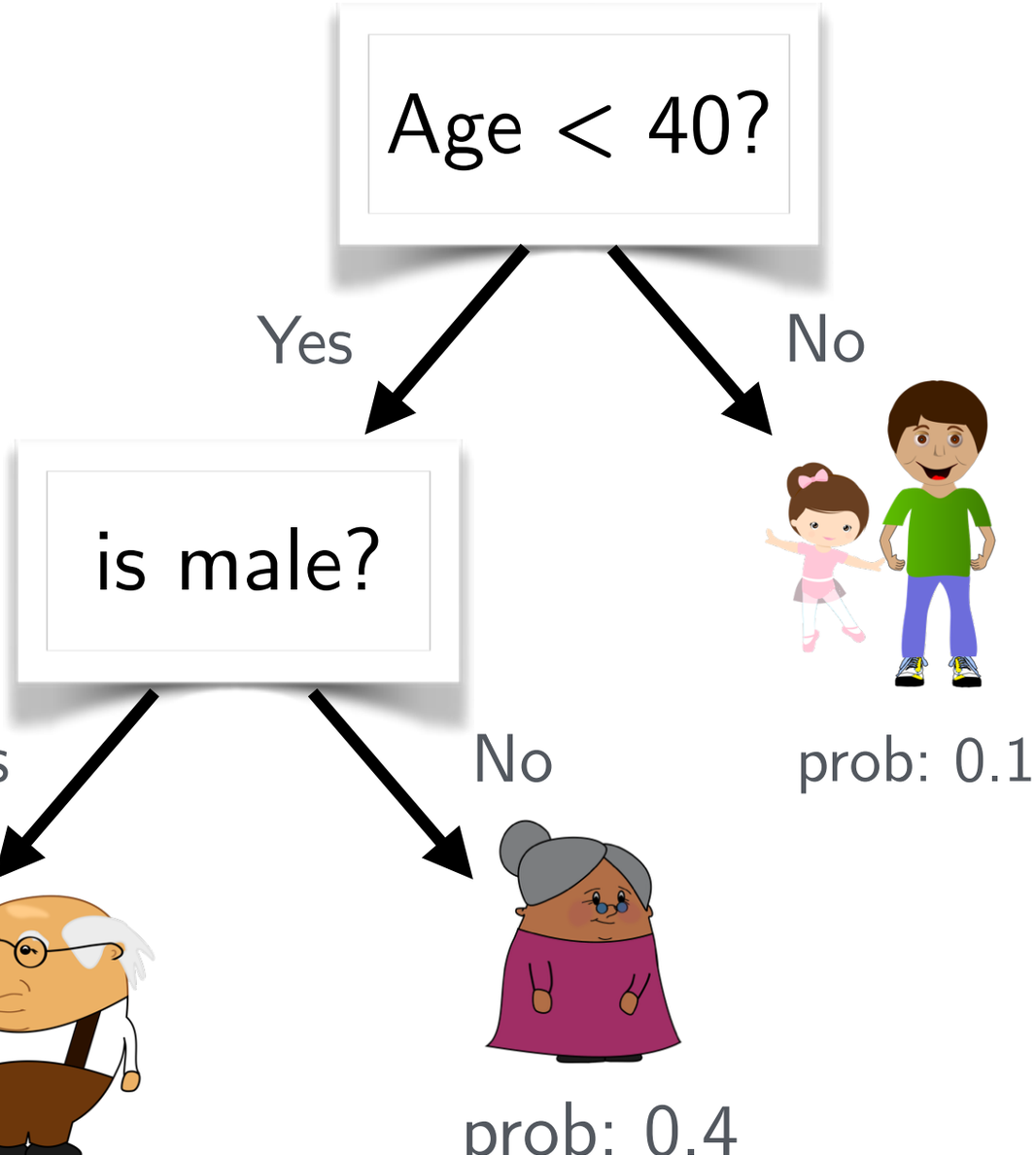
- Gradient boosted tree classifier implemented in XGBoost
- Equally good at handling qualitative & quantitative predictors
- Ignores predictors that are not important, good for sparse data

Simple decision-tree classification e.g.

Inputs: age, gender, occupation



Does the person enjoy golf?



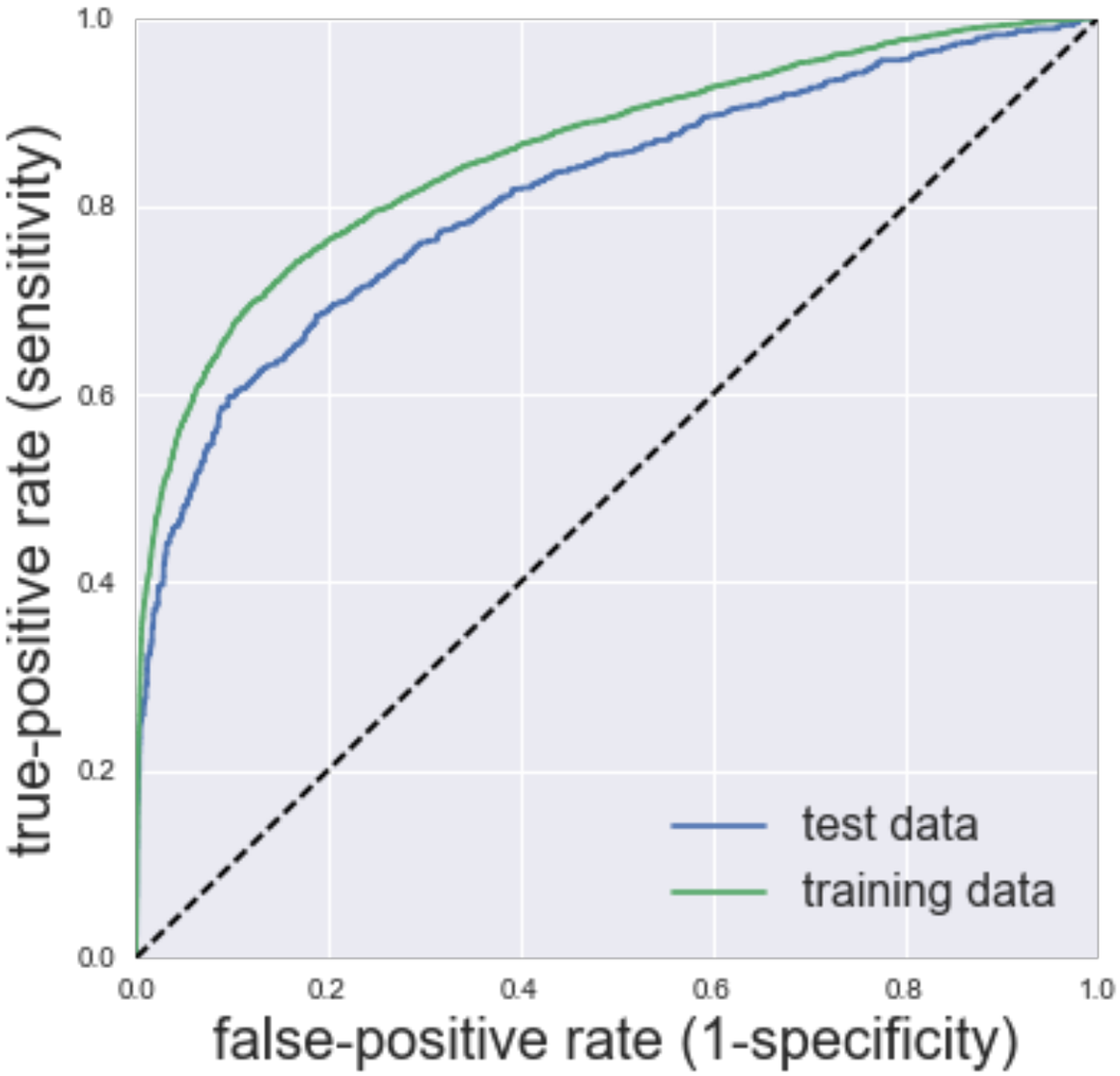
"gradient boosting"

boosting - new trees produced focus on data wrongly classified by combination of previous trees

gradient - refers to method used to find optimal tree (fast)

Evaluating predictions

Receiver operating characteristic (ROC) curve



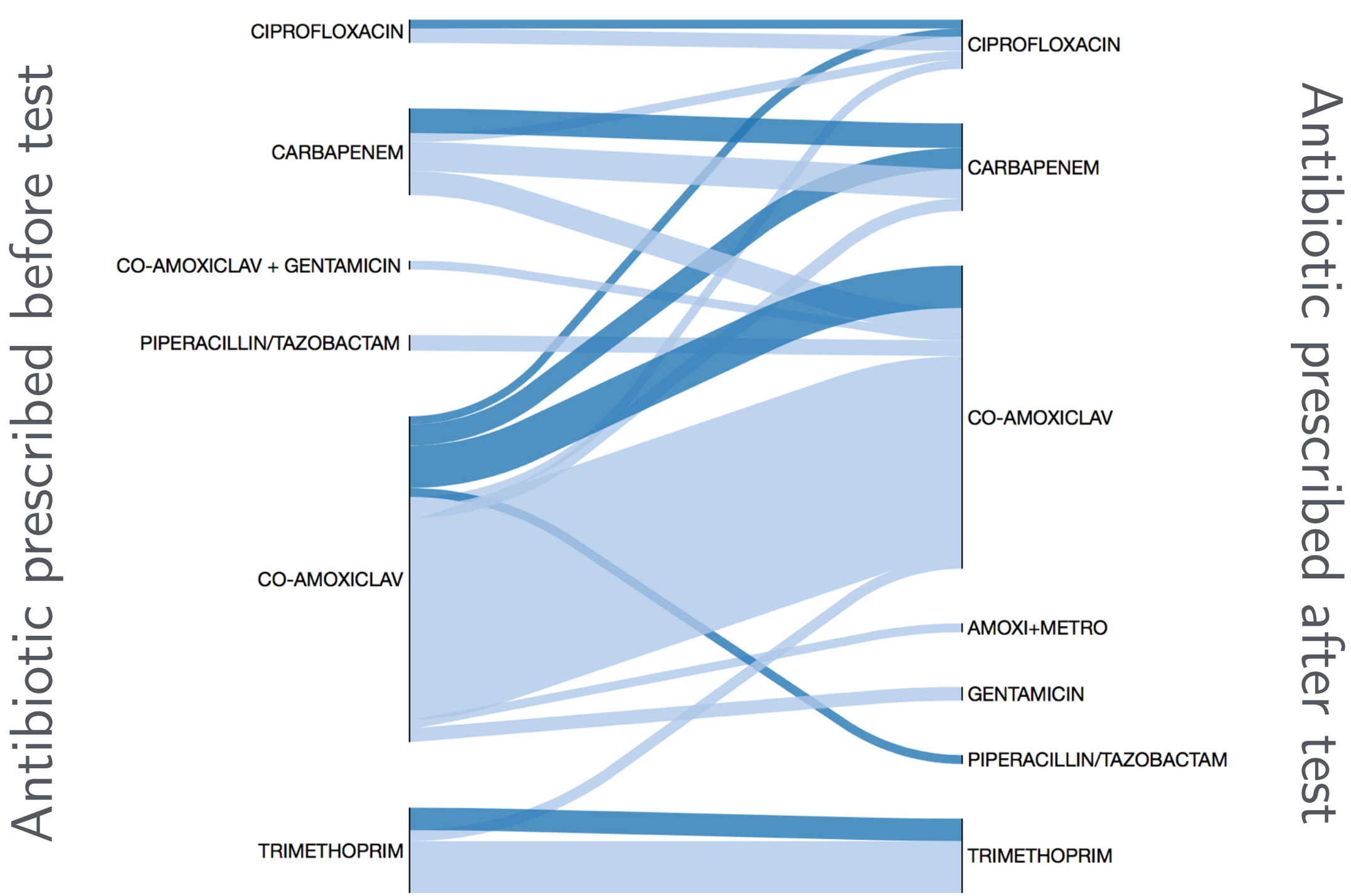

```
Model Report (Train)
Accuracy : 0.8571
AUC Score : 0.857859

Model Report (Test)
Accuracy : 0.8345
AUC Score : 0.813391
```

- But how useful is this really?

How can we judge how useful our predictions are?

From data: Antibiotic therapies for UTIs with positive test for E.coli



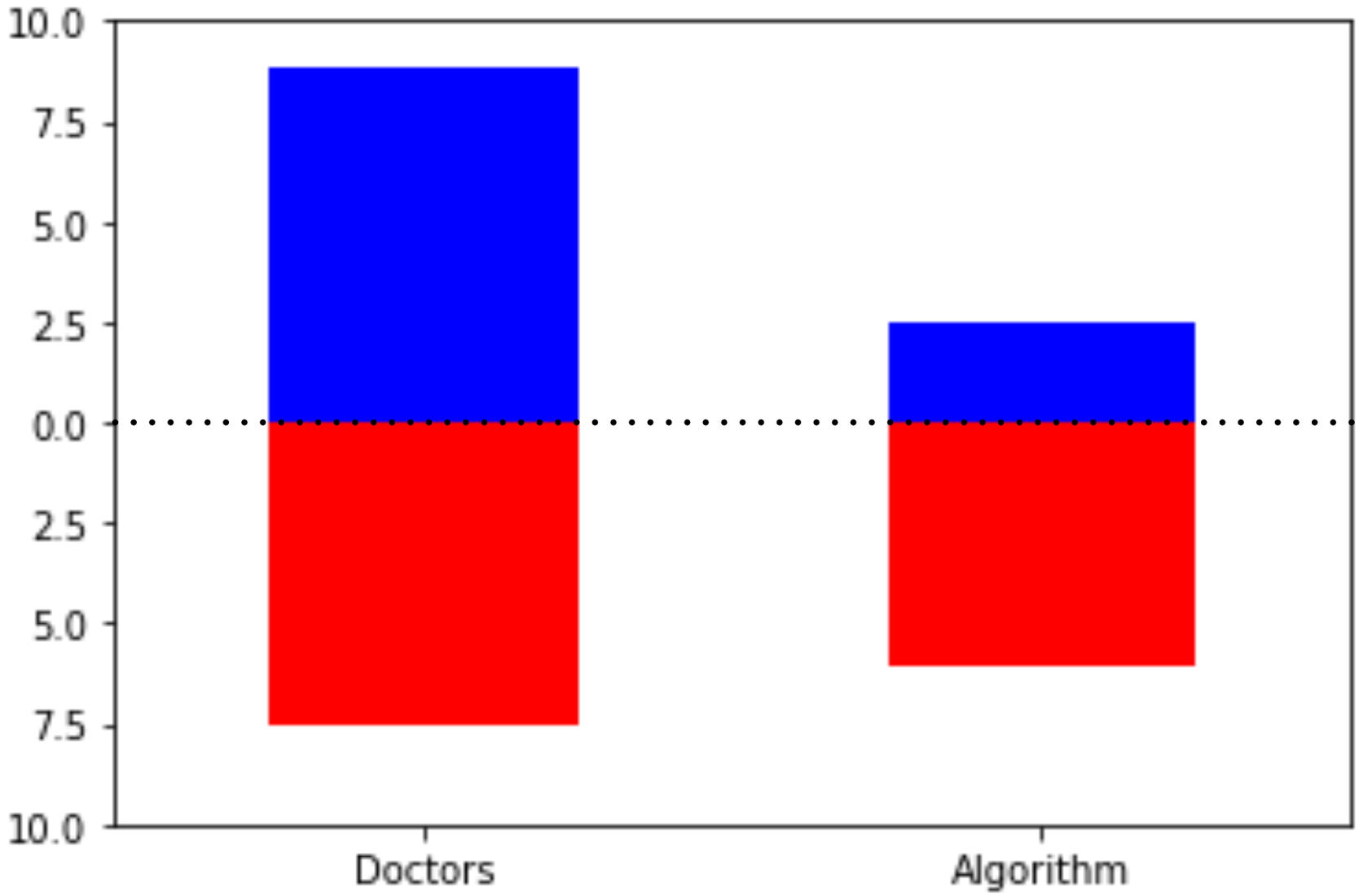
Simplified example

- Co-amoxiclav & carbapenems most common antibiotic therapy for UTIs, so for simplicity let's just consider these
- Co-amoxiclav - medium/broad spectrum, antibiotic "work-horse"
- Carbapenems - very broad spectrum, resistance very rare

Simplified example

% unnecessary carbapenem

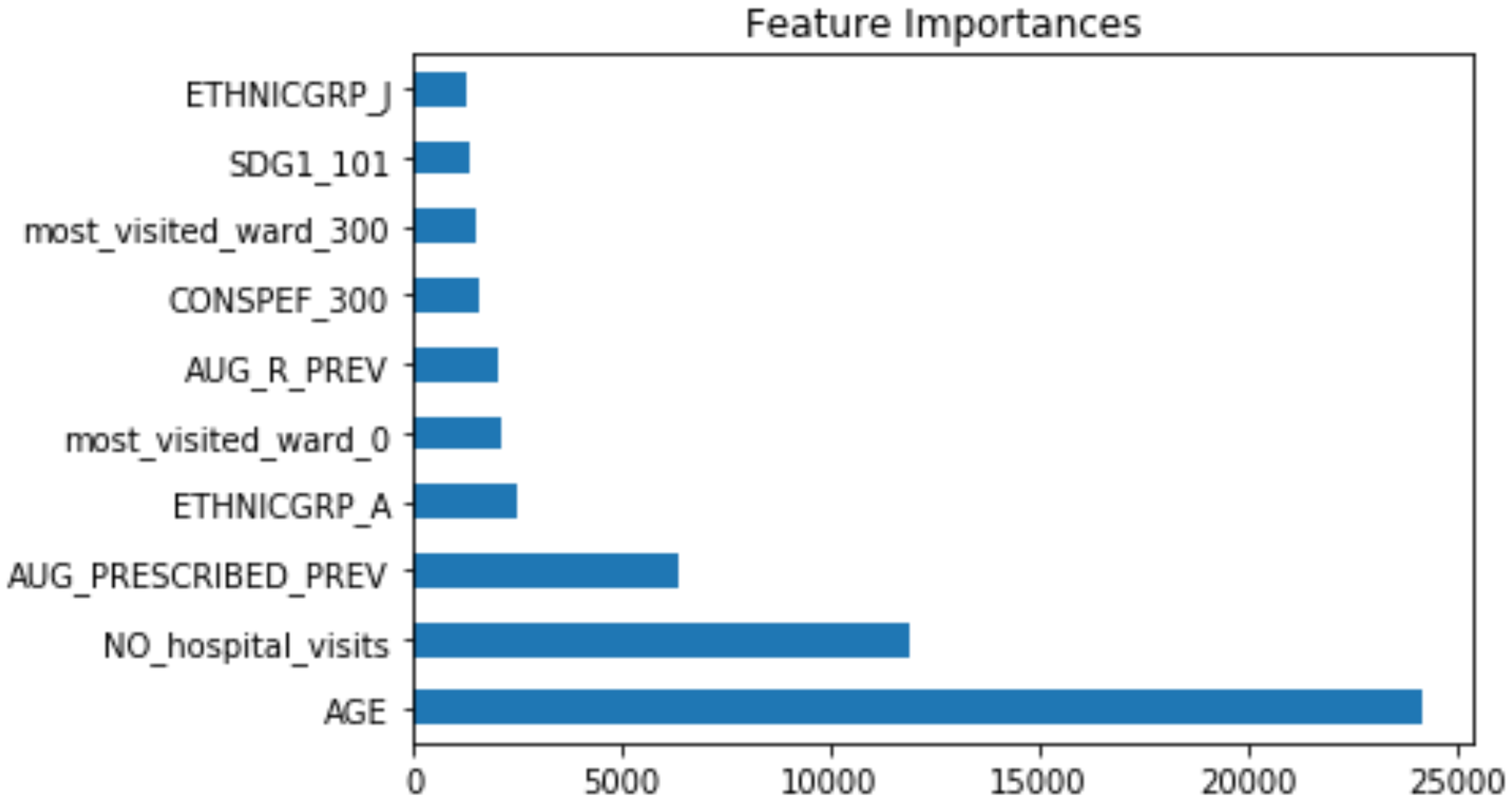
% ineffective initial choice



Finally...

- Looking back on the data is not the best way to test performance
- We want to make an AI that gives doctors a likelihood of resistance for range of antibiotics - doctors should make decisions.
- Heart of England has lots of good quality data, lots of scope for using evidence to aid public health

Thanks for you attention


Feature Importances
ETHNICGRP_J
SDG1_101
most_visited_ward_300
CONSPEF_300
AUG_R_PREV
most_visited_ward_0
ETHNICGRP_A
AUG_PRESCRIBED_PREV
NO_hospital_visits
AGE
0
5000
10000
15000
20000
25000

Simplified example

Before test	After test		Average stay (s.e.)	Number found in dataset	Algorithmic choice
Co-amoxiclav	Co-amoxiclav	**S**	6.0 days (±0.4)	**295**	**321**
Carbapenem	Co-amoxiclav	**S**	8.8 days (±1.4)	**35**	**10**
Co-amoxiclav	Carbapenem	**R**	11.3 days (±1.4)	**30**	**24**
Carbapenem	Carbapenem	**R**	11.3 days (± 1.7)	**36**	**42**

- Potentially fewer total days, and fewer patients on broad spectrums