

Hospital Infection Rate Classifier

Jackie Gutman, Alex Pine, Maya Rotmensch



NEW YORK UNIVERSITY

Why do HAIs matter?

Hospital-acquired infections are the leading cause of preventable patient mortality.

1 in 25 hospital patients acquire an infection.

Mortality rate is **12%** and higher.

Each infection costs upwards of **\$17,000**.

The Data

Using “Central Line” infections as our measure.

Represented two ways:

- Raw score

- Binary indicator (“Above Average”, or otherwise)

Used Binary indicator as target.

Used 5 different features, including the infection rate of previous years.

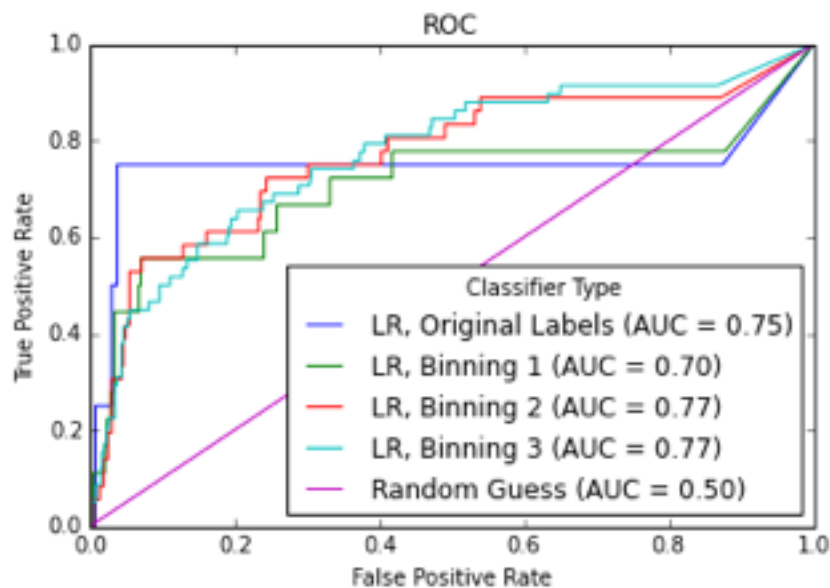
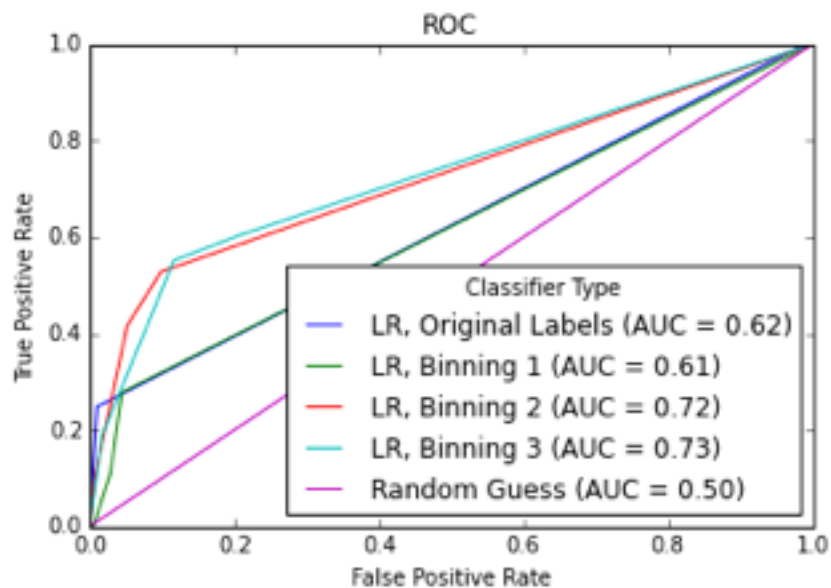
Challenges

- Lots of potential training data to sift through.
- Very few positive labels (only 23 of 2005).
 - Calculated our own, more inclusive, labels from the raw scores to reduce variance.

The Model

Autoregressive Model

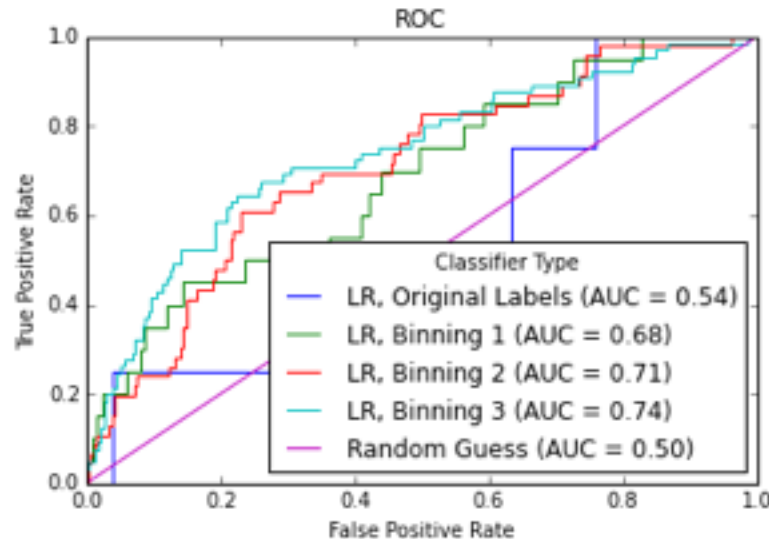
- Logistic Regression
- Trained only on infection rates from previous years.



The Model

Logistic Regression - Full Data Model

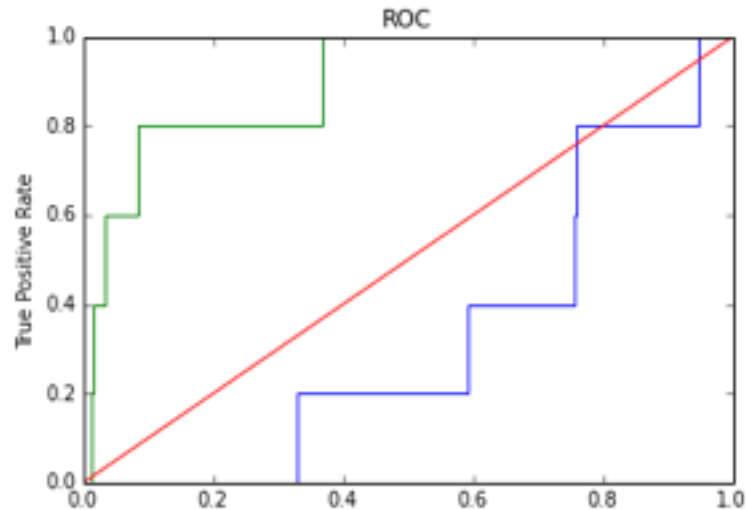
- Dataset with all available features . AUC \sim 0.70
- no significant difference between L2 Vs. L1 regularization.
- CDC supplied labels perform very poorly.



The Model

Proxy Logistic Regression Model

- Train on variable that:
 - Similar in distribution to target variable
 - More frequent
- Test on original target variable

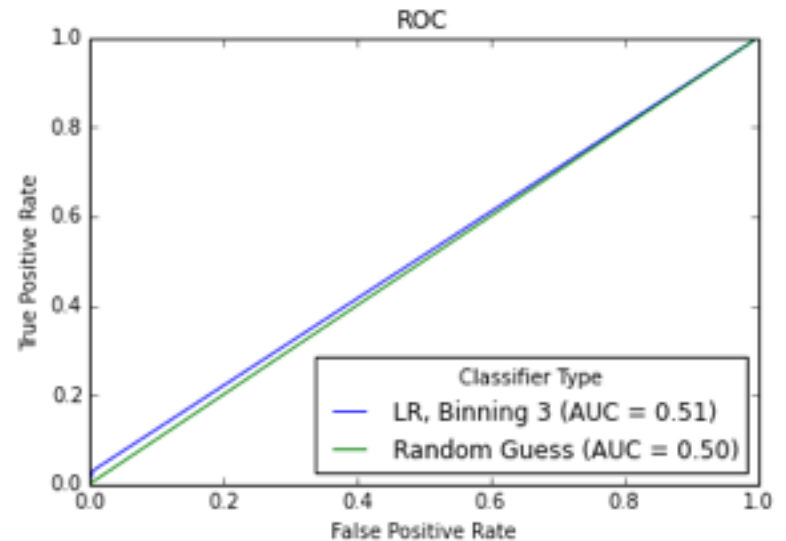
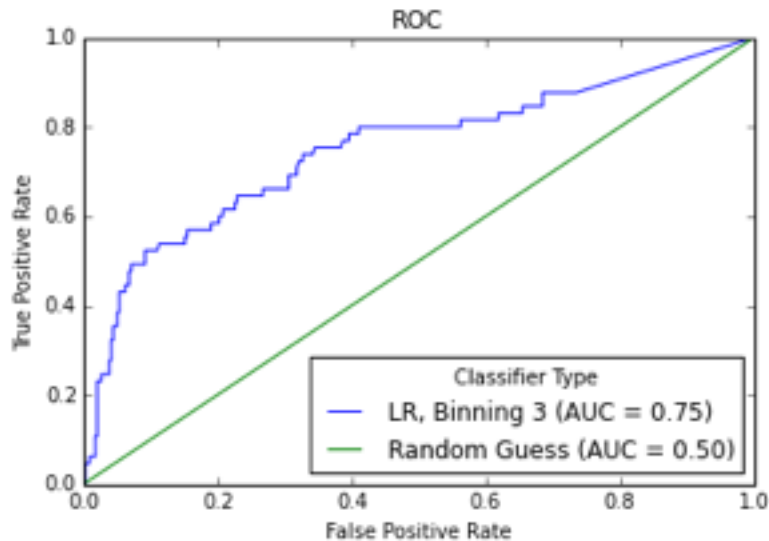


— LR, Regular Labels (AUC = 0.32) — Random Guess (AUC = 0.50)
— LR, Binning Labels Proxy 2 (AUC = 0.90)

The Model

Feature Selection

- Recursive Feature Elimination shows our non-infection data does not yield meaningful features.



Deployment and Future Work

Goal: *where to send our teams of infection control specialists*

Autoregressive Model with our more inclusive binning functions.

If important to predict CDC provided labels

- Proxy Logistic Regression Model

Future Work

- Investigate more features in the hopes of finding more informative features.
- Create autoregressive models for other Infection rate measures.
- Experiment with other proxy measures.

Thank you for listening!