Exploring Anomaly Fraud Detection in Medicare Part D

Edward Gu, Ethan Mah, Carnell Zhou, Ellen Li

Agenda

- Background
- Methodology
- Visual Representations
- Discussion
- Conclusion

Background

43 million enrolled in Part D

Health care fraud costs US $100 billion annually

Number of fraud providers caught each year continue to rise

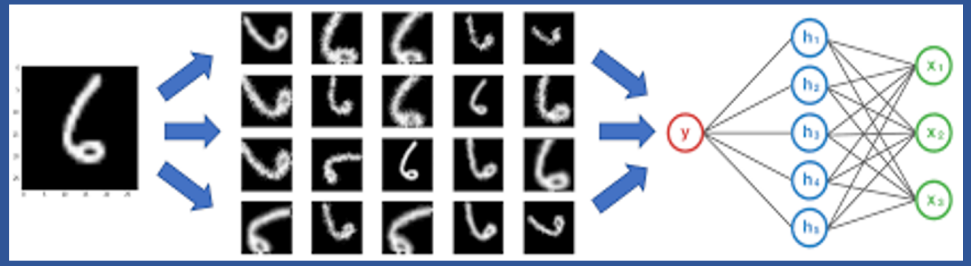
Databases:

Medicare Part D PUF 17 Medicare Part D Summary 17
Office of Inspector General Exclusion 18 RxNav
Health Facts Datasets

101 unique excluded NPI's compared to 904,999 unique non-excluded NPI's

Extremely imbalanced datasets

Background



- Supervised Learning
- ROS + RUS
- Training Data Augmentation
- Train + Test + Split
- Pandas, Numpy, Sklearn, Tensorflow, etc.
- AUC

Medicare Part D Costs in 2019

Coverage Phase	Total Drug Costs	Cost-Sharing Percentage
1. Annual Deductible	$415	*Beneficiary Pays 100%*
2. Initial Coverage	$3,820	*Beneficiary Pays 25%* / *Plan Pays 75%*
3. Secondary Coverage (Donut Hole)	$5,100	Brand Name Medications: (70% Manufacturer Discount) *Beneficiary Pays 25%* Plan Pays 5% Generic Medications: *Beneficiary Pays 37%* Plan Pays 63%
4. Catastrophic Coverage		** Plan Pays 15% Medicare Pays 80%

Gaps in Knowledge

- Multi-year trends of Medicare Part D & exclusion data not examined previously
- Many new features such as scheduled drugs
- Training Data Augmentation
- LSTM RNN

Table 1. Definitions for Schedule I – V Drugs
Schedule I Drugs - High potential for abuse - No currently accepted medical use in the U.S. - Lack of accepted safety for use of the drug under medical supervision
Schedule II Drugs - High potential for abuse - Currently accepted medical use in the U.S. - Abuse may lead to severe psychological or physical dependence
Schedule III Drugs - Potential for abuse less than schedule I and II drugs - Currently accepted medical use in the U.S. - Abuse may lead to moderate or low physical dependence or high psychological dependence
Schedule IV Drugs - Lower potential for abuse less than schedule III drugs - Currently accepted medical use in the U.S. - Abuse may lead to limited physical or psychological dependence relative to schedule III substances
Schedule V Drugs - Low potential for abuse relative to schedule IV substances - Currently accepted medical use in the U.S. - Abuse may lead to limited physical or psychological dependence relative to schedule IV substances

Data

Analytics	Preprocessing	Enrichment	RNN
Combined datasets via NPI's Manually documented 100 Schedule 2 drugs	Created 26 non-specialty features One hot-encoded 188 specialty features	Training data augmentation: 10,000 fraud cases added Ratio of 0.012 % kept ROS, RUS, and ROS	Purpose: identify similarities and differences between frauders/non-frauders throughout time

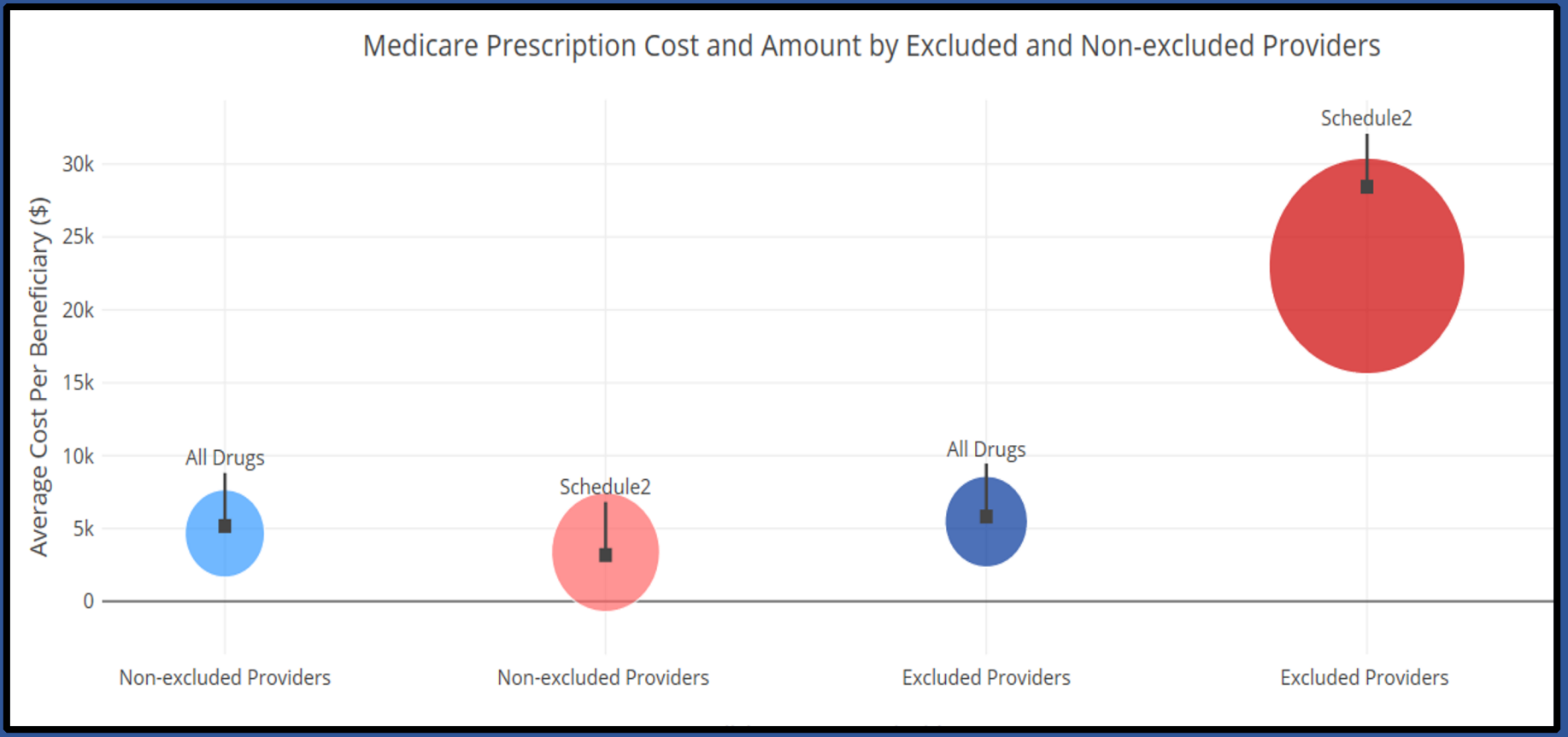
About Our Models

Supervised Classification Models

Logistic Regression
Decision Trees
Gaussian Naive-Bayes
KNN
RBF SVM
Ada Boost
Gradient Boosting
Extra Trees
Random Forest

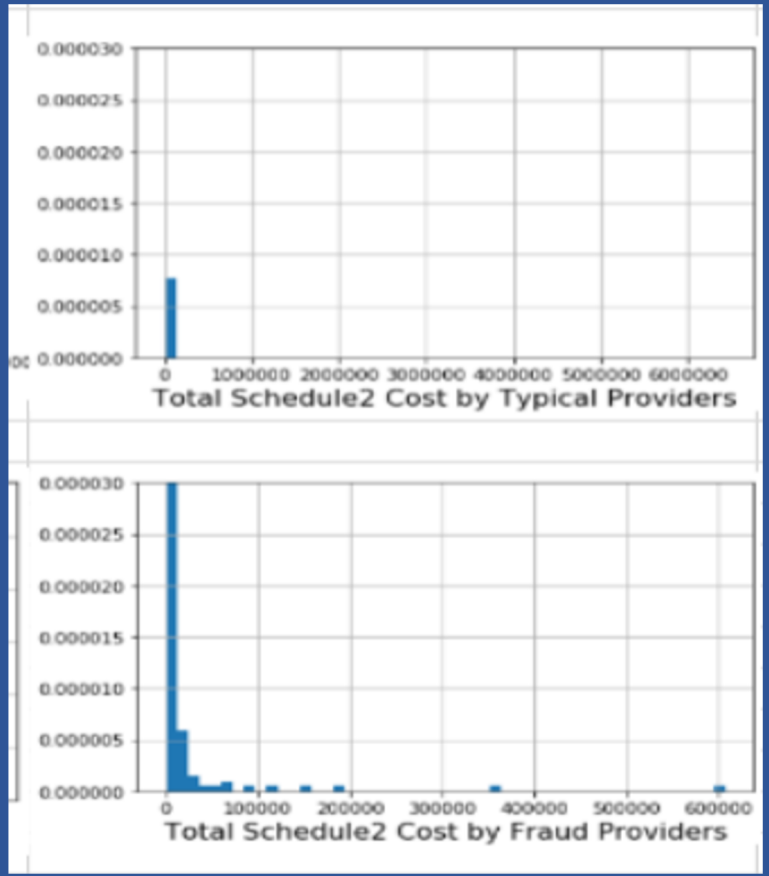
Deep Neural Network

Hyperparameters
(Dense: 16, Dropout: 0.2, Activation: ReLU & Sigmoid
Optimizer: Adam
Loss: Binary Crossentropy,
Batch Size: 400, Epochs: 20 (Early Stopping))

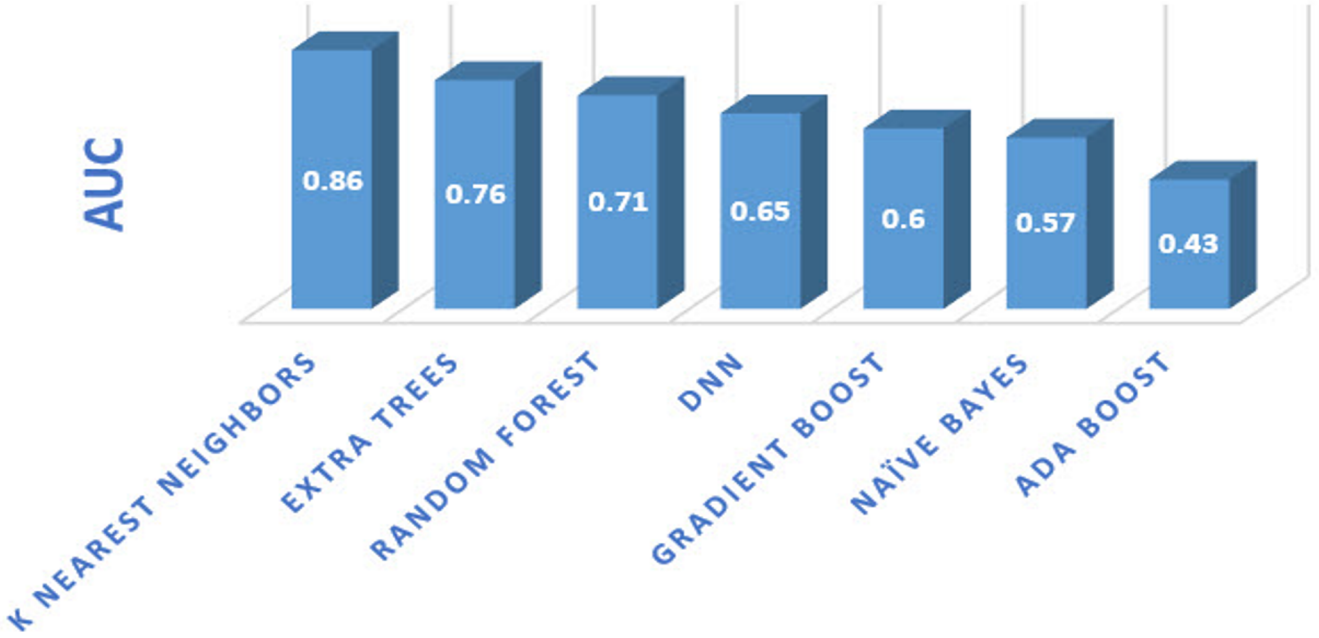

Medicare Prescription Cost and Amount by Excluded and Non-excluded Providers
Average Cost Per Beneficiary ($)
30k
25k
20k
15k
10k
5k
0
All Drugs
Schedule2
All Drugs
Schedule2
Non-excluded Providers
Non-excluded Providers
Excluded Providers
Excluded Providers

Four States (NJ, NY, CA, IL)

```
Dimensions of training set:
(233733, 5, 846)
Dimension of full target:
(233733, 1)
The top 25 variables in the training set
[('fraud', 133264, 0.5701548347901237)]

Predicting: 'fraud'
Dimensions of test set:
(58971, 5, 846)
Dimensions of target:
(58971, 1)

Total positive cases in test set:
41
Total predicted positive cases in test set:
302.0
Total sum of probabilities
314.67322
Computed AUC of the ROC:
0.816226982819633
F1 score:
0.03498542274052478
Average precision score:
0.00792686261628352

Classification report
              precision    recall  f1-score   support

           0       0.99      1.00      1.00     58669
           1       0.15      0.02      0.03       302

   micro avg       0.99      0.99      0.99     58971
   macro avg       0.57      0.51      0.52     58971
weighted avg       0.99      0.99      0.99     58971
```



Machine Learning Models



Discussion Pt. I

1. Top fraud drugs differ greatly to top non-fraud drugs by prices (less so by amount)
2. Fraud providers sell more brand drugs (a top feature)
3. Fraud providers prescribe more Schedule 2 drugs
4. States with highest fraud rates = highest opioid addiction rates
5. Top 3 fraud specialties: Internal medicine, Family practice, Cardiology

Discussion Pt. II

1. Saves taxpayer and government money
2. ROS and RUS both improved RF AUC scores slightly
3. Reduces unnecessary Schedule 2 drugs prescribed
4. K Nearest Neighbor achieved an AUC of 0.86
5. Numerous novel ways of anomaly detection
6. RNN yielded AUC of 0.82 for top 4 states

Thank You!

Future Studies

- Utilize Medicare Opioid Prescriber Summary datasets
- Compare SMOTE, ADASYN, and other enrichment methods to training data augmentation
- Optimize data augmentation
- Improve neural networks
- Determine best distance calculation to determine fraud
- R ratios only based off 10-20 fraud cases
- Unsupervised market basket can be used in machine learning