

Pharmacy Claims – Fraud Detection

Giridharan Gurumoorthy – Kavi Global

Rajesh Inbasekaran – Kavi Global

#DS9SAIS

Outline

- Background
 - Fraud Waste and Abuse
 - Current Status and Challenges
- Project
 - Objective
 - Steps
- Implementation
- Initial Results
- Next steps
- Q&A

Background



Healthcare spend

\$3.35 Trillion
\$10K



Pharmacies

67,000
700



Prescriptions

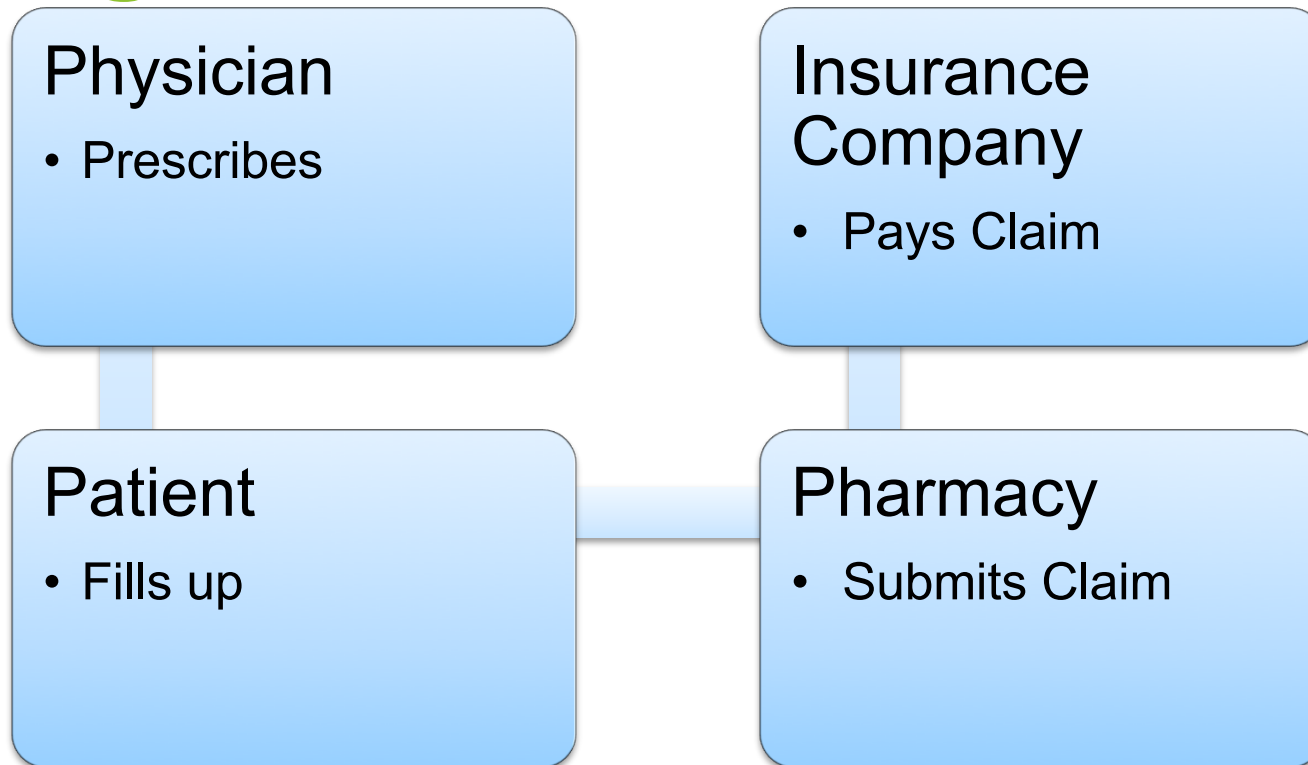
4 billion
11.9%



Physicians

950K
\$6.5B

Background



Fraud Waste and Abuse

- 3% to 10% of healthcare spend
- Possible Actors (Alone or in collusion)
 - Patient
 - Physician
 - Pharmacy

Current Status and Challenges

- Transaction level checks are in place
- Actor level checks are primitive
- Rule based checks are based on historical fraud
 - Fraudsters are innovative
- False positives are expensive

Project

Objective

- Identify anomalous / possible fraudulent actors
- Rank them
 - Prioritize investigations

What did we do?

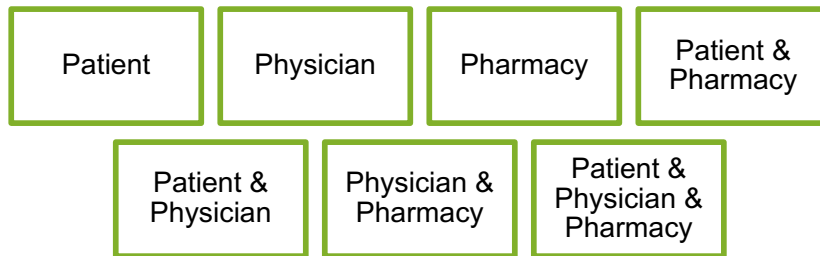
- Identify and rank anomalous behavior
- Identify and rank anomalous relationships between actors
- Generate consolidated scores to rank

Steps

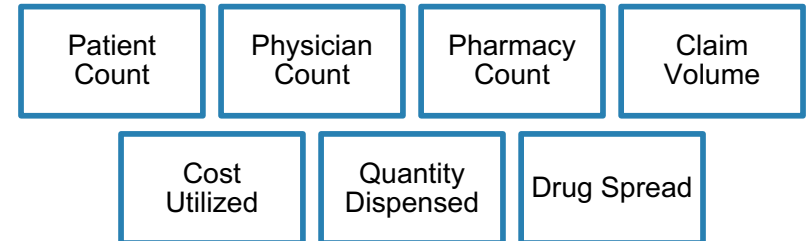


Data Summary

Categories



Metrics



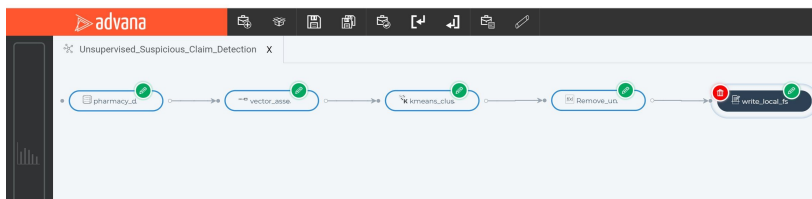
Timeframe



Anomalous behavior

- Grouping similar behavior
- Group Density
- Distance from Group Center

K Means Model



PREVIEW DATA FOR kmeans_clustering1

kmeans_clustering1	id	gtr_dt	total_mbr	avg_clm_per_mbr	avg_days_per_mbr	avg_qty_per_mbr	totalLupi	avg_npi_per_clm	avg_dist_per_clm
12	49	4	1	1	30	30	1	1	0
12	50	4	1	4	120	120	2	0.5	0
12	50	4	1	4	120	120	2	0.5	0
1	1	1	1	5	130	160	3	0.6	0
1	1	1	1	5	130	160	3	0.6	0
1	1	1	1	5	130	160	3	0.6	0
3	9	1	1	1	5	10	1	1	0
6	24	2	1	1	30	4	1	1	1.3
7	26	3	1	3	90	78	1	0.333333333	1.3
7	30	3	1	1	30	18	1	1	1.3

KMEANS_CLUSTERING1

General

Parameter Setting

Temporary Storage



Select Column of Feature Variable* ?

input_features x

Enter a New Column Name for Prediction Output* ?

cluster_output x

Number of Cluster* ?

20

Max Iteration* ?

20

Random Seed ?

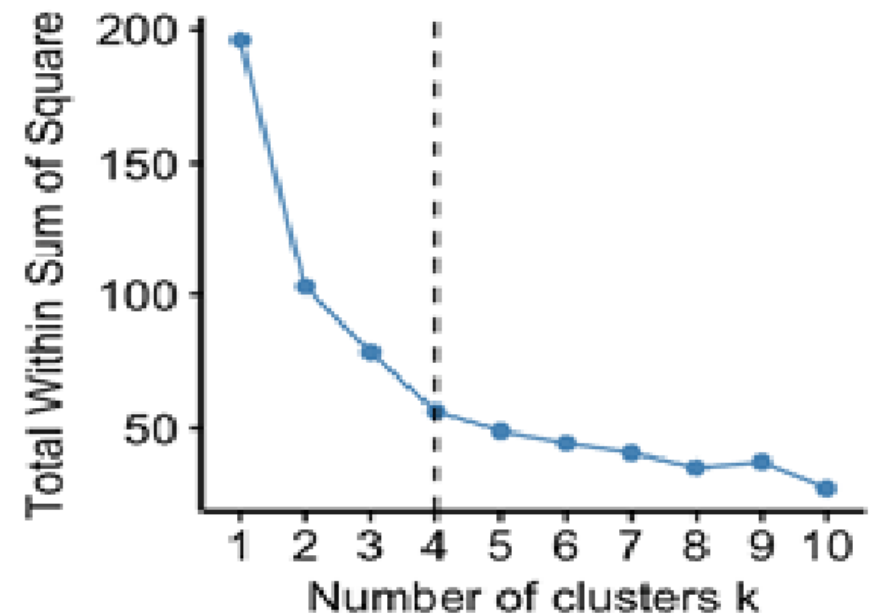
Convergence Tolerance* ?

0.0001

Optimal 'K'

- Elbow Method
 - Compute cluster algorithm for different values of K
 - Calculate WSS for each K and plot the curve
 - Location of the Bend will be the optimal value of 'K'

Optimal Number of Clusters



Anomaly Score

- Density Factor = Size of the cluster / Total Size
- Distance Factor = Distance between data point and center of cluster / Distance of the farthest point in the cluster
- Score = Max of two scores
- Actor Score = Sum of level wise scores with weightages

Example

Patient Level Score card

Patient ID	Cluster_Size_Factor Group 1	Cluster_Distance_Factor Group 1	Cluster_Size_Factor Group 2	Cluster_Distance_Factor Group 2	Cluster_Size_Factor Group 3	Cluster_Distance_Factor Group 3	Cluster_Size_Factor Group 4	Cluster_Distance_Factor Group 4	Anomaly_Score
Patient - 1	1.0000	1.0000	0.9974	1.0000	0.9979	1.0000	0.9979	1.0000	0.9992
Patient - 2	1.0000	0.9985	0.9974	0.9997	0.9979	1.0000	0.9979	1.0000	0.9989
Patient - 3	0.9971	0.9997	0.9974	0.9997	0.9979	0.9998	0.9979	0.9998	0.9987
Patient - 4	0.9971	0.9990	0.9974	0.9997	0.9979	0.9998	0.9979	0.9998	0.9986
Patient - 5	0.9971	0.9972	0.9974	1.0000	0.9979	1.0000	0.9979	1.0000	0.9984
Patient - 6	0.9971	0.9962	0.9974	1.0000	0.9979	1.0000	0.9979	1.0000	0.9983

Physician Level Score card

Physician ID	Cluster_Size_Factor Group 1	Cluster_Distance_Factor Group 1	Cluster_Size_Factor Group 2	Cluster_Distance_Factor Group 2	Cluster_Size_Factor Group 3	Cluster_Distance_Factor Group 3	Cluster_Size_Factor Group 4	Cluster_Distance_Factor Group 4	Anomaly_Score
Physician - 1	0.9965	1.0000	0.9973	1.0000	0.9979	1.0000	0.9979	1.0000	1.0000
Physician - 2	0.9997	1.0000	0.9973	1.0000	0.9979	1.0000	0.9979	1.0000	1.0000
Physician - 3	0.9965	1.0000	0.9973	1.0000	0.9979	1.0000	0.9979	1.0000	1.0000
Physician - 4	0.9997	0.9998	0.9973	1.0000	0.9979	1.0000	0.9979	1.0000	1.0000
Physician - 5	0.9965	0.9998	0.9973	1.0000	0.9979	1.0000	0.9979	1.0000	0.9999
Physician - 6	0.9965	0.9998	0.9973	1.0000	0.9979	1.0000	0.9979	1.0000	0.9999

Pharmacy Level Score card

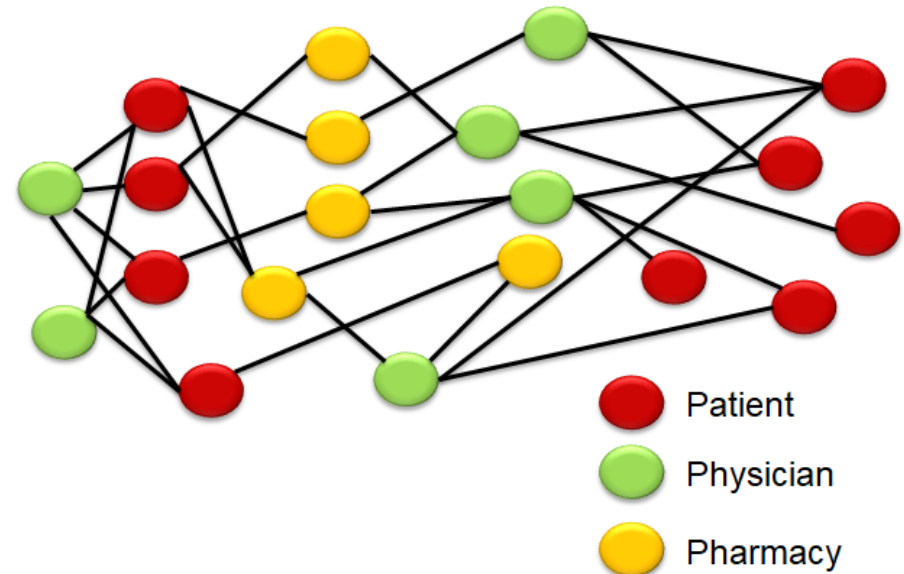
Pharmacy ID	Cluster_Size_Factor Group 1	Cluster_Distance_Factor Group 1	Cluster_Size_Factor Group 2	Cluster_Distance_Factor Group 2	Cluster_Size_Factor Group 3	Cluster_Distance_Factor Group 3	Cluster_Size_Factor Group 4	Cluster_Distance_Factor Group 4	Anomaly_Score
Pharmacy - 1	0.9776	1.0000	0.9974	1.0000	0.9973	1.0000	0.9979	1.0000	1.0000
Pharmacy - 2	0.9776	1.0000	0.9974	1.0000	0.9973	1.0000	0.9979	1.0000	1.0000
Pharmacy - 3	0.9776	1.0000	0.9974	1.0000	0.9973	1.0000	0.9979	1.0000	1.0000
Pharmacy - 4	0.9994	0.9998	0.9998	1.0000	0.9998	1.0000	0.9999	1.0000	1.0000
Pharmacy - 5	0.9776	0.9998	0.9974	1.0000	0.9973	1.0000	0.9979	1.0000	0.9999
Pharmacy - 6	0.9776	0.9998	0.9974	1.0000	0.9973	1.0000	0.9979	1.0000	0.9999

Anomalous Relationship

- Analyze
 - # of Connected Neighbors
 - # of Neighbors' Neighbors

GraphX

- Create a graph with multiple node types
 - Physician
 - Patient
 - Pharmacy



Calculate Neighbor Count

- First level Neighbor
 - Compute Immediate neighbor's degree from the each vertex
- Second Level Neighbor
 - Identify Neighbor's Neighbor degree and bring to the parent vertex
- Third Level Neighbor
 - Identify Second Level Neighbor's Neighbor degree and bring to the parent vertex

Anomaly Score

- Based on position in each count
- Consolidated score = Max of all the counts

Consolidate Score

- Sum of Behavior Score and Relationship Score with configurable weightages

Implementation – How it works?

- Web Application for the use of Analysts
- Features
 - Upload Claim Data
 - Run models
 - View Results – Actor wise ranks
 - Action – Tag as False Positive or Initiate Case
 - Feedback – Input Investigate status

Initial Results

- Many findings would have escaped Rule based checks
- Initial investigation results prove less false positives
- Ranking weightages might need some tuning

Next steps

- Supervised techniques with investigation results
- Use additional data
 - Social media – Twitter, Facebook, Review data
 - Address and Property data - Zillow

Q & A



A close-up photograph of a fountain pen with a dark, textured barrel and a silver-colored nib. The pen is positioned diagonally from the upper right, having just finished writing the word 'Contact' in a cursive script. It is now writing the word 'us!' in the same style. The ink is a dark brown color. The background is a plain, light-colored surface.

Contact us!

Rajesh Inbasekaran : Rajesh@kaviglobal.com
Giridharan Gurumoorthy : Giridharan@kaviglobal.com