PONDERA

NAMPI 2019

AUGUST 25-28, 2019 ATLANTA, GEORGIA

1

NAMPI 2019

AUGUST 25-28, 2019 ATLANTA, GEORGIA

HUMANS V. ROBOTS

How Artificial Intelligence and Machine Learning Help Medicaid Investigators fight Fraud

PONDERA

2

NAMPI 2019

Session Overview

- Introductions & Session Background
- What is Artificial Intelligence (AI) and Machine Learning (ML)?
- How are AI and ML used in healthcare data analytics?
- Case Studies
- Best Practices & Take Home Techniques

PONDERA

3

NAMPI 2019

Panel Introductions

- **Mr. Gene D'Angelo, Director of Analytics Innovation**
  - Advanced Analytics for Big Data
  - Fraud Detection Specialist
- **Ms. Keri Geiger, Azure Data and AI Specialist, Microsoft**
  - Data & Artificial Intelligence Architect
  - Cloud Specialist, Microsoft
- **Mr. Jacob Gray, Studies Manager, General Dynamics**
  - Healthcare Analytics Leader
  - Studies Manager, Healthcare Fraud Prevention Partnership, Trusted Third Party
- **Ms. Amanda Huston, Chief Customer Officer, Pondera Solutions**
  - Former FL MFCU Lieutenant
  - Former FL Benefit Program Integrity Director

PONDERA

4

WHAT IS ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING?

5

What is Artificial Intelligence & Machine learning?

- Artificial Intelligence is the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages
- Machine Learning is the scientific study and use of algorithms and statistical models to perform a specific task effectively. It relies on patterns and inference instead of explicit instructions

6

NAMPI 2019

What is Artificial Intelligence & Machine Learning? (Continued)

AI technology has the power to:

- Amplify human ingenuity
- Extend our capabilities
- Transform industries
- Increase productivity
- Help solve society's biggest challenges

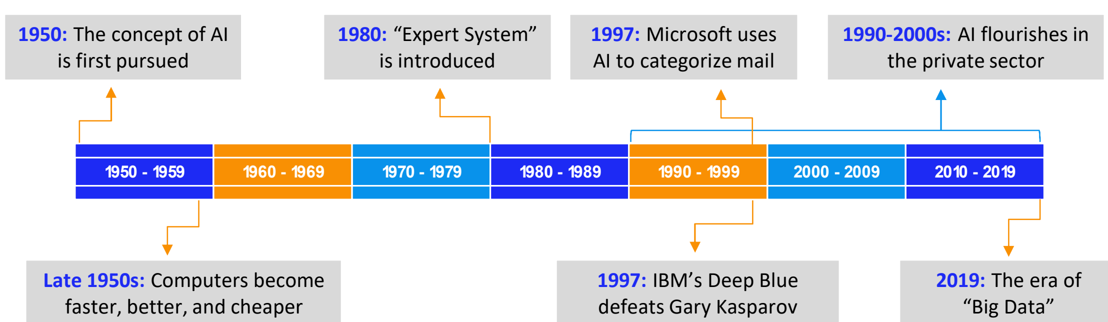
This intelligent technology is already improving our lives today and promises to change the world tomorrow in ways unimaginable to us now.

PONDERA

7

NAMPI 2019

What is Artificial Intelligence & Machine Learning? (Continued)



PONDERA

8

NAMPI 2019

HOW ARE THEY USED IN HEALTH CARE DATA ANALYTICS?

PONDERA

9

NAMPI 2019

How are they used in Health Care Data Analytics?

- Machine Learning – The Beginning
  - 1994 at Blue Cross Blue Shield of Florida
  - Medicare Fraud Branch
  - Ph.D. student in AI at University of Central Florida
  - HOPS – Heuristic Optimized Processing System
  - Custom machine learning model
    - Network Concentration
    - Maximized links to entity ratio

PONDERA

10

How are they used in Health Care Data Analytics? (Continued)

- Large fraud ring (level 47) found in SE Florida
  - Billing for services not rendered
  - Over 150 providers
  - Over $200 million in 18 months
- Very few financial incentives for fraud detection in 1994
  - Claims volume for BCBSFL significantly impacted
  - Pressured to cease fraud detection efforts

PONDERA

11

How are they used in Health Care Data Analytics? (Continued)

- Supervised Machine Learning
  - Identify the characteristics of fraudulent providers using a training set of providers labeled as fraudulent and non-fraudulent
  - Training sets for healthcare fraud are hard to find
    - High degree of class imbalance
      - Random Under-sampling (RUS) is one fix
      - High performance measures can be fool's gold
    - High error rate in non-fraudulent class
  - Separate models might be needed for each type of fraud

PONDERA

12

NAMPI 2019

How are they used in Health Care Data Analytics? (continued)

- Unsupervised Machine Learning
  - Group providers together based on similarities
  - Network Analysis – group providers with similar connections
    - Different ways to define connections
    - Shared beneficiaries is a popular approach
  - Cluster Analysis – group providers with similar practices
    - Hierarchical clustering since goal is to find aberrancy
    - Computationally demanding
    - Good data preparation is vital
  - Good for finding new and unknown types of fraud

PONDERA

13

NAMPI 2019

How are they used in Health Care Data Analytics? (continued)

The Hybrid Approach: Apply an unsupervised technique, then use a training set to identify suspicious groups of providers

1. Form networks and clusters using unsupervised learning
2. Apply to past years to capture fraudulent behavior
3. Use your own case data to create a training set
4. Identify networks and clusters with fraudulent providers
5. Follow providers from past clusters and networks
6. Identify suspicious networks and clusters in recent months

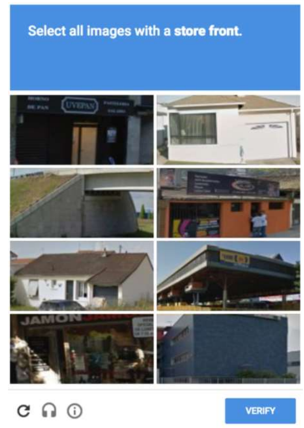
PONDERA

14

NAMPI 2019

How are they used in Health Care Data Analytics? (Continued)

- Expertise = Pattern Recognition
- Scaling and Automating Data Points used by SMEs
- Single SME identified one pattern at a time
- Instead of building machines, "build machines that make machines"



PONDERA

15

NAMPI 2019

CASE STUDIES

PONDERA

16

NAMPI 2019

Case Study 1
Detection Method

- Cluster analysis to identify aberrant providers
  - Focus on larger providers with sufficient activity to form a profile based on the total utilization of procedures and diagnoses
  - Cluster analysis identified an atypical cluster of providers (one suspect NPI was the largest in this cluster)
  - Utilizing the Paired Providers report with suspect NPI, Pondera determined the strongest similarity (based on shared recipients) was 42% with suspect NPI

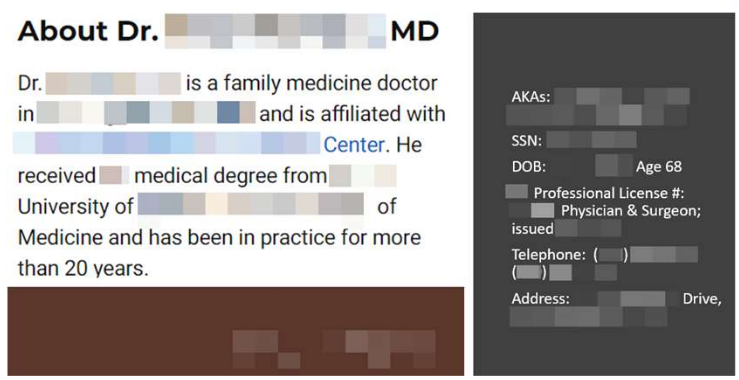
PONDERA

17

NAMPI 2019

Case Study 1
Suspect Provider in the Network

- 3 Billing NPIS
  - Physician
  - Group
  - Clinical Lab
- 2018 paid amounts ranged from $250K to over $2M



PONDERA

18

NAMPI 2019

Case Study 1
Overview of Concerns

- Pondera utilized the Cluster Provider Model to search for other providers in close proximity – one was less than a block from a suspect practice location
  - Suspect NPI's Administrator was the "Authorized Official" (active govt investigation)
  - Heavy similarity of beneficiaries between two other facilities/labs
- Provider's Network had history of complaints and investigation
  - Regulators indicated they have been "a problem for years"
- Mostly same program code for services
- Non-enrolled providers rendering services
- High percent of services related to relative referrals

PONDERA

19

NAMPI 2019

Case Study 1
Overview of Concerns

- Excessive services: up to 87 services per recipient for 10-month period
- Unexplained billing spikes
- Suspect Patient ID Number behavior:
  - Billing for one recipient using multiple IDs on the same date of service
  - Alternating use between different IDs (often issued by the provider of interest)
  - Over 50K distinct combined IDs (3K+ using multiple IDs)
- Many recipients never had single claim from other providers
- Advertisement for free services

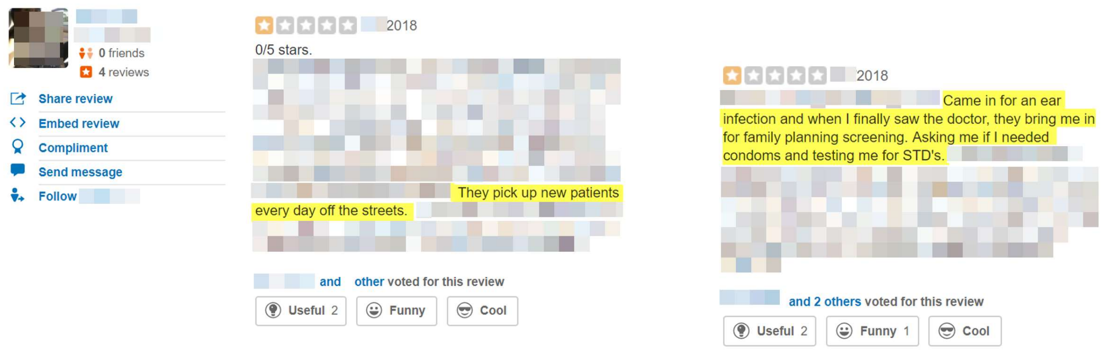
PONDERA

20

NAMPI 2019

Case Study 1
Open Source Reviews

Concerns about recruiting patients & unnecessary services



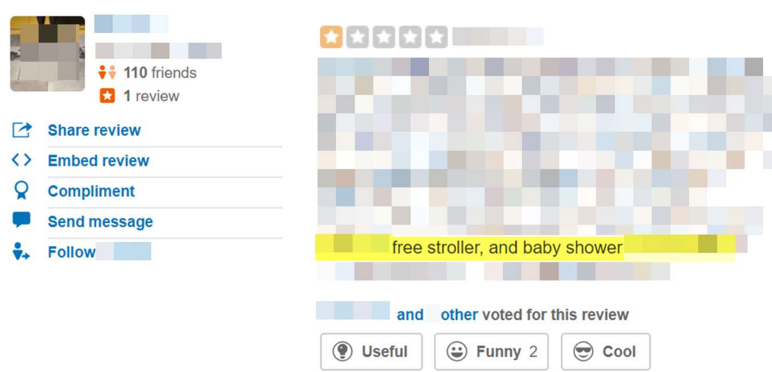
PONDERA

21

NAMPI 2019

Case Study 1
Free Services

Free services offered in conjunction with government benefits programs are generally prohibited



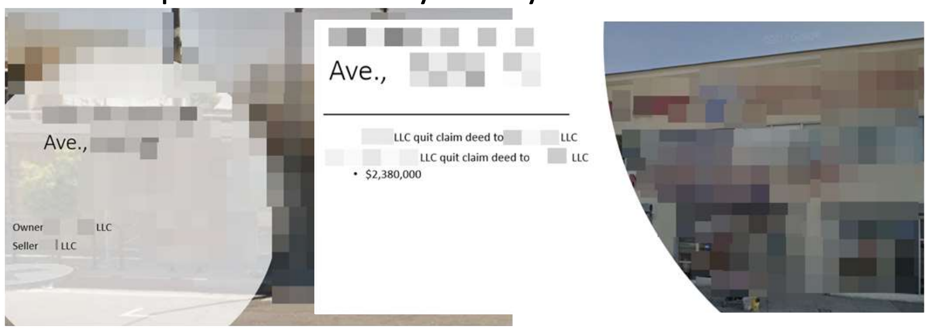
PONDERA

22

NAMPI 2019

Case Study 1
Property Swapping

Public records showed pattern where clinic properties changed ownership names with companies owned by family members about every six months



PONDERA

23

NAMPI 2019

Case Study 2
Policy Drives Change in FWA

- Therapy Provider
- Data Points Utilized by SMEs
  - Procedure Codes, their Levels and Combinations, same Patient/same Day

Provider	Patient	Day	Procedure
PHYS1	MBR1	8/26/19	99213
PHYS1	MBR1	8/26/19	90838
PHYS1	MBR2	8/26/19	99213
PHYS1	MBR2	8/26/19	90838
PHYS1	MBR3	8/27/19	99213
PHYS1	MBR3	8/27/19	90838
PHYS1	MBR4	8/27/19	99213
PHYS1	MBR4	8/27/19	90838

PONDERA

24

NAMPI 2019

Case Study 3
Hiding in Plain Sight

- Provider Ownership is a complicated web
  - Identifiers: Tax IDs, NPIs, Provider IDs
  - Ownership: Names
  - Other: Addresses, Phones, Emails
- Shared Patients
  - Money
  - Time
- Dense concentrations within small clusters
- LA Physician Example
  - Align your investigative practices to best support the data analytics

PONDERA

25

NAMPI 2019

Case Study
How are Facebook Ads like Fraud Detection?

- NLP ability to 'read' strings
- Procedures, Diagnoses, Provider Types
- Interpret patient history
- Predict next phase of health journey
- Flag unexpected cost inflections

PONDERA

26

NAMPI 2019

Supervised Machine Learning Naïve Bayes Classification Case Study

Rank	HCPS Code	Paid Amount	HCPCS Description	Predicted Provider Type	Listed Provider Type
1	88305	90,180.91	Pathology examination of tissue using a microscope	Pathology	Orthopedic Surgery
2	88341	20,193.36	Special stained specimen slides to examine tissue	Pathology	Orthopedic Surgery
3	88342	16,730.66	Tissue or cell analysis by immunologic technique	Pathology	Orthopedic Surgery
4	88307	14,812.32	Pathology examination of tissue using a microscope,	Pathology	Orthopedic Surgery
5	88313	13,562.55	Special stained specimen slides to examine tissue in	Pathology	Orthopedic Surgery
6	G0416	13,352.29	Surgical pathology, gross and microscopic examinati	Pathology	Orthopedic Surgery
7	88189	10,390.74	Flow cytometry technique for DNA or cell analysis	Pathology	Orthopedic Surgery
8	88360	9,941.75	Microscopic genetic analysis of tumor	Pathology	Orthopedic Surgery
9	88333	6,510.41	Pathology examination of tissue specimen during sur	Pathology	Orthopedic Surgery
10	88112	6,488.73	Cell examination of specimen	Pathology	Orthopedic Surgery

PONDERA

27

NAMPI 2019

Supervised Machine Learning Naïve Bayes Classification Case Study

Rank	HCPS Code	Paid Amount	HCPCS Description	Predicted Provider Type	Listed Provider Type
1	92014	45,105.57	Eye and medical examination for diagnosis and tre	Ophthalmology	Pathology
2	92012	33,579.65	Eye and medical examination for diagnosis and tre	Ophthalmology	Pathology
3	65855	9,590.16	Laser repair to improve eye fluid flow, 1 or more se	Ophthalmology	Pathology
4	66984	8,834.24	Removal of cataract with insertion of lens	Ophthalmology	Pathology
5	66982	8,243.19	Removal of cataract with insertion of lens	Ophthalmology	Pathology
6	92083	7,438.84	Measurement of field of vision during daylight cond	Ophthalmology	Pathology
7	92004	4,041.78	Eye and medical examination for diagnosis and tre	Ophthalmology	Pathology
8	92133	3,695.28	Diagnostic imaging of optic nerve of eye	Ophthalmology	Pathology
9	66821	2,964.97	Removal of recurring cataract in lens capsule using	Ophthalmology	Pathology
10	92020	1,810.96	Examination of cornea and iris using lens device a	Ophthalmology	Pathology

PONDERA

28

NAMPI 2019

Unsupervised Machine Learning Cluster Analysis Case Study

Rank	HCPS Code	Total Paid	HCPCS Description
1	66984	517,523,577.46	Removal of cataract with insertion of lens
2	92014	457,015,395.73	Eye and medical examination for diagnosis and treatment, established patient, 1 or more visits
3	92012	204,500,108.68	Eye and medical examination for diagnosis and treatment, established patient
4	66821	101,192,681.94	Removal of recurring cataract in lens capsule using laser
5	92004	85,060,539.05	Eye and medical examination for diagnosis and treatment, new patient, 1 or more visits
6	92083	73,931,787.10	Measurement of field of vision during daylight conditions
7	66982	63,578,475.52	Removal of cataract with insertion of lens
8	92250	52,852,265.93	Photography of the retina
9	92136	50,713,897.82	Measurement of corneal curvature and depth of eye
10	92133	45,382,921.78	Diagnostic imaging of optic nerve of eye

PONDERA

29

NAMPI 2019

Unsupervised Machine Learning Cluster Analysis Case Study

Rank	HCPS Code	Total Paid	HCPCS Description
1	J0178	1,301,308,100.53	Injection, aflibercept, 1 mg
2	J2778	853,404,543.21	Injection, ranibizumab, 0.1 mg
3	67028	197,607,245.73	Injection of drug into eye
4	92014	171,472,510.79	Eye and medical examination for diagnosis and treatment, established patient, 1 or more
5	92134	128,957,258.53	Diagnostic imaging of retina
6	92235	79,643,906.54	Examination of retinal blood vessels by ophthalmoscope
7	92012	73,447,393.55	Eye and medical examination for diagnosis and treatment, established patient
8	92250	40,692,894.11	Photography of the retina
9	67228	35,891,268.65	Laser destruction of leaking retinal blood vessel, 1 or more sessions
10	J9035	31,199,673.08	Injection, bevacizumab, 10 mg

PONDERA

30

NAMPI 2019

Unsupervised Machine Learning
Cluster Analysis Case Study
Small but Aberrant Cluster

Rank	HCPS Code	Total Paid	HCPCS Description
1	66762	151,098.27	Creation of opening in iris for eye fluid drainage using laser, 1 or more sessions
2	92014	79,312.20	Eye and medical examination for diagnosis and treatment, established patient, 1 or more visits
3	65855	70,413.01	Laser repair to improve eye fluid flow, 1 or more sessions
4	68801	50,518.16	Dilation of tear-draining opening
5	92012	44,526.33	Eye and medical examination for diagnosis and treatment, established patient
6	92250	26,760.79	Photography of the retina
7	92235	24,660.42	Examination of retinal blood vessel by ophthalmoscope
8	92226	21,075.51	Examination of eye by ophthalmoscope with retinal drawing
9	68110	20,237.01	Removal of (up to 1 centimeter) growth of sclera
10	66984	15,063.54	Removal of cataract with insertion of lens

PONDERA

31

NAMPI 2019

BEST PRACTICES & TAKE HOME TECHNIQUES

PONDERA

32

NAMPI 2019

Best Practices & Take Home Techniques

Partnership of Humans and Machines

Humans are required to:

1. Develop the training set based on past cases
2. Design the appropriate way to prepare the data
3. Identify true and false positives and give feedback
4. Iteratively modify the algorithm and the data preparation

PONDERA

33

NAMPI 2019

Best Practices & Take Home Techniques

Suggestions

- Beware of fool's gold
- Consider a hybrid approach
- Recast your claims data to express relevant patterns
- Make a cost-effective analytical tool available to your data scientists
- Approachable AI
  - https://sway.office.com/DwdxT9ud839ZSyET?ref=email

PONDERA

34

Audience Q & A

PONDERA

35