7th Capacity Building Seminar in Health and Care Insurance

Gurgaon

13-Dec-2019

Data Science in Health Insurance - Need of the hour

Sumit Ramani, Consulting Actuary, Actuaria Consultants

Rajasekhar Maddireddy, Founder, PillStore

Institute of Actuaries of India

Agenda

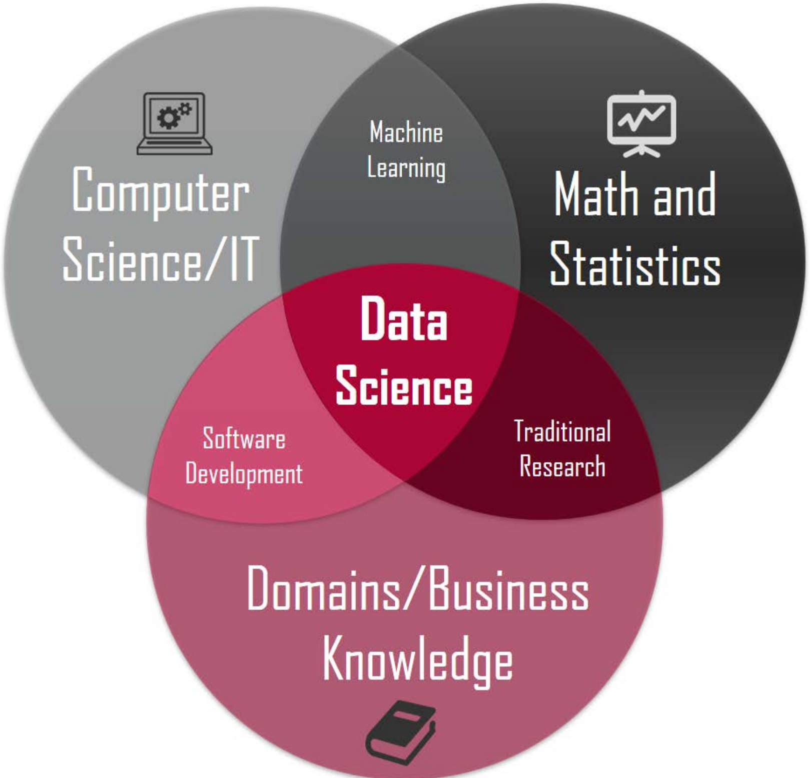
- What is Data Science?
- Why Data Science is need of the hour?
- How Data Science is impacting Health Insurance?
- Case Study: Fraud Detection
  - Structured data
  - Unstructured data to Structured dat
  - Data science comes in many flavours
  - Machine Learning
    - Usecase - An example


What is Data Science?

Definition

Meaningful insights from data using

- Domain Expertise
- Programming Skills
- Mathematics & Statistics



Okay! But what does it mean?

Institute of Actuaries of India

- Application of machine learning algorithms and more to produce artificial intelligent systems that would ordinarily require human intelligence

- In turn, these systems generate insights which analysts and business users can translate into tangible business value.

Data Science is a comprehensive process & Artificial Intelligence is a tool to implement it

So what is Artificial Intelligence?

Artificial Intelligence

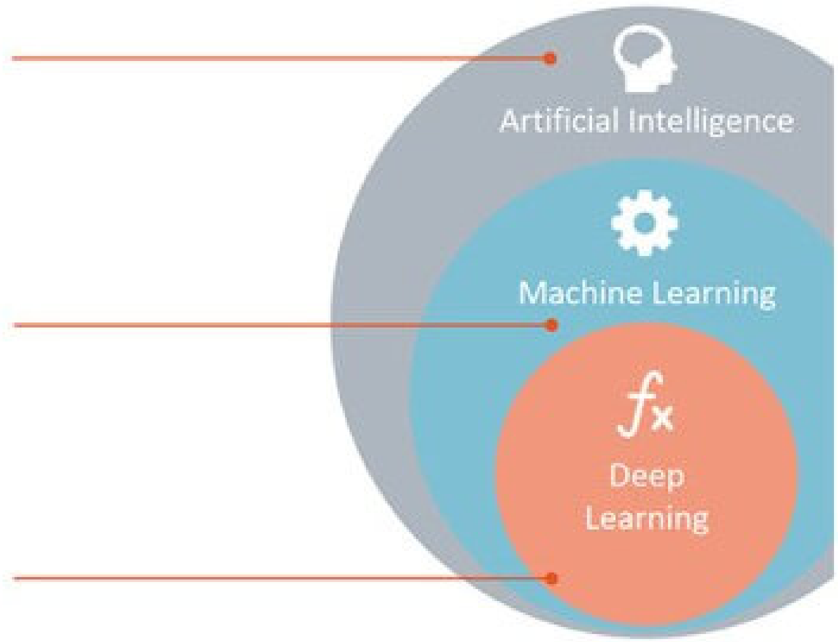
Any technique which enables computers to mimic human behavior.

Machine Learning

Subset of AI techniques which use statistical methods to enable machines to improve with experiences.

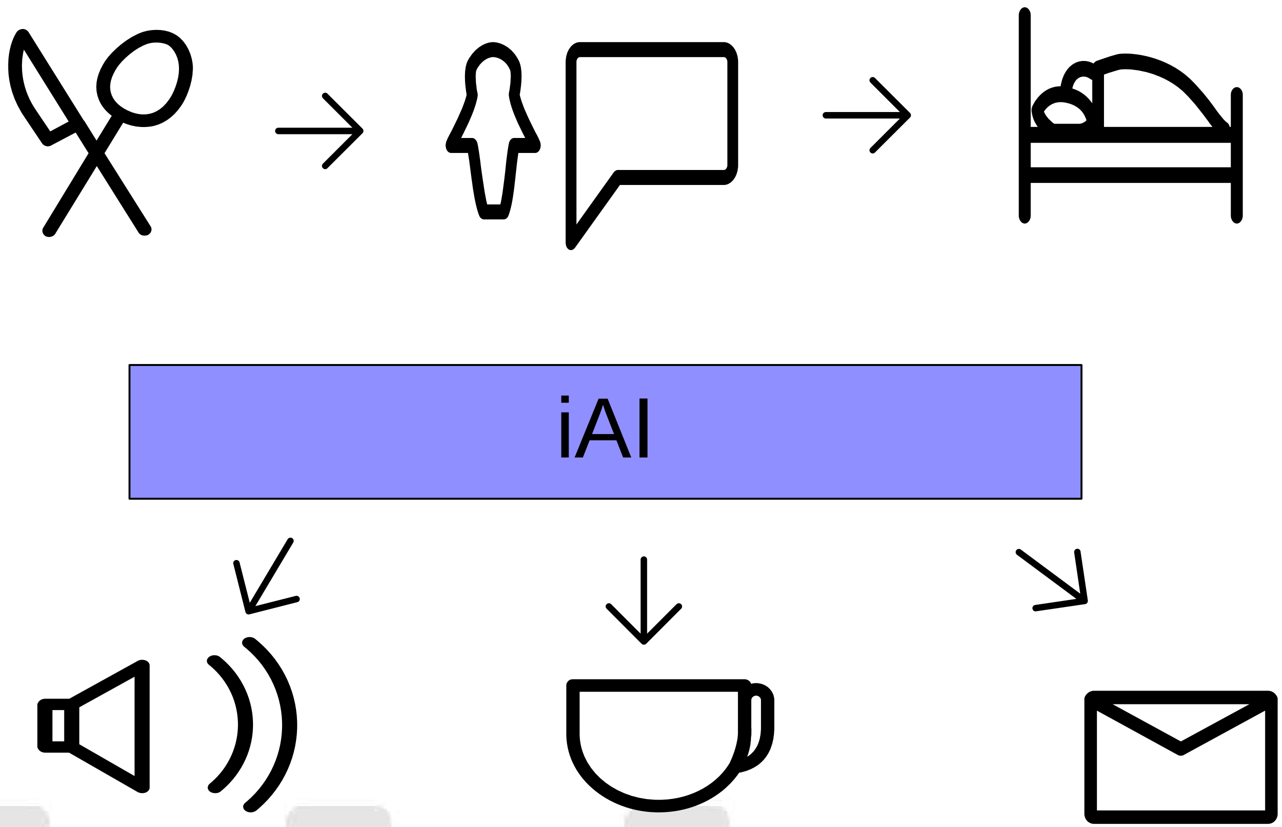
Deep Learning

Subset of ML which make the computation of multi-layer neural networks feasible.



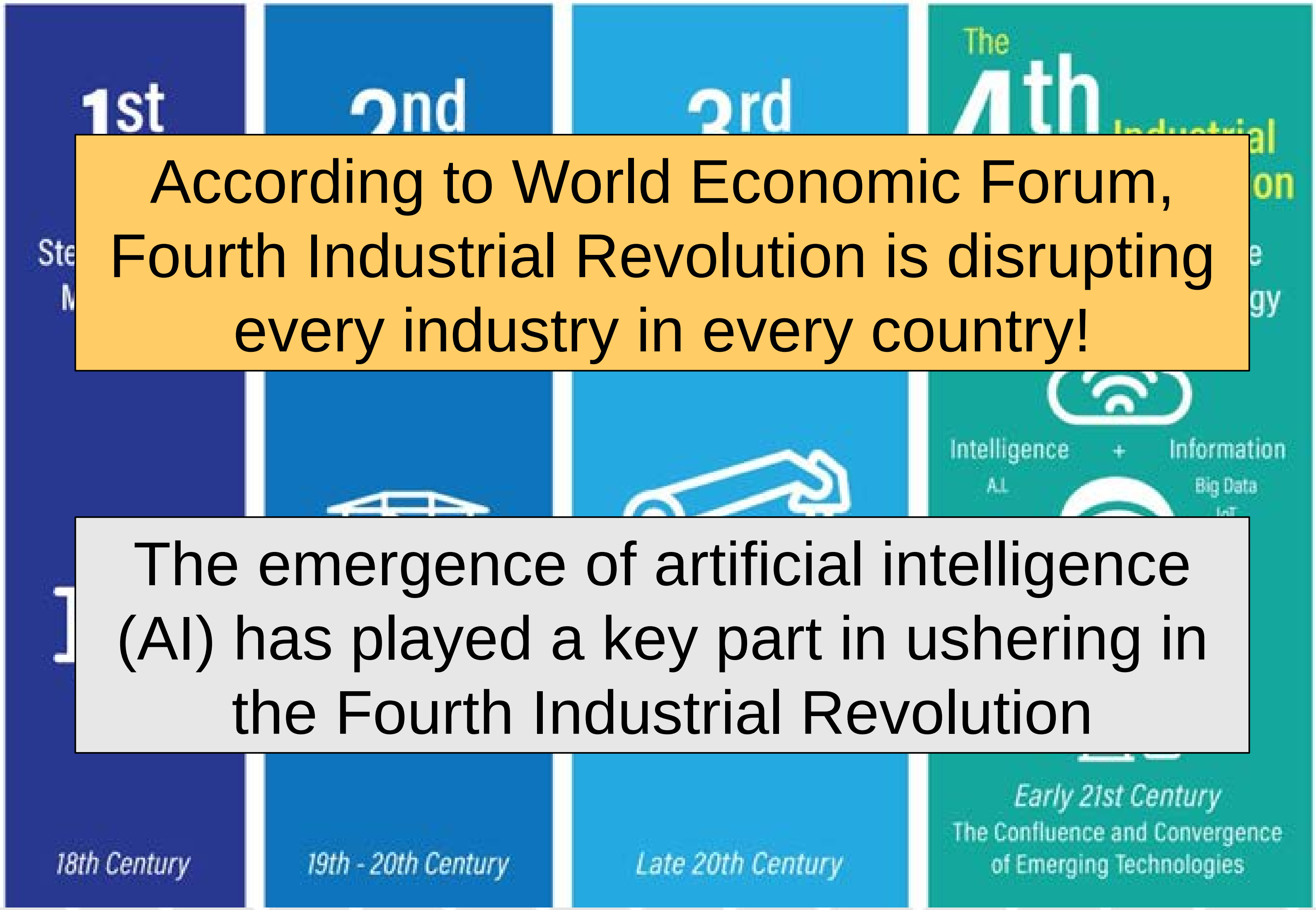
Source: [KDNuggets]

Example please!



Need of the hour!

Fourth Industrial Revolution



According to World Economic Forum, Fourth Industrial Revolution is disrupting every industry in every country!

The emergence of artificial intelligence (AI) has played a key part in ushering in the Fourth Industrial Revolution


AI in Health Insurance

AI in Customer Analysis

- Universal Health Score
- Aggregates health and lifestyle data
- Predicts the risk of preventable lifestyle diseases like Cardiovascular, COPD, Diabetes

AI in Customer Engagement

- AI Bots
- 97% interaction happens with AI Bots
- Type of interaction include
  - Applying for coverage
  - Checking benefits
  - File medical claims

AI in Claims Management

- Efficient Claims Management

Kirontech

- Uses machine learning to detect patterns which relate to
  - Fraud – inaccurate billing
  - Waste – underutilization of services

AI in predicting Emergency Visit

- Predicts when a customer is likely to visit Emergency Room

- Examples
  - Knee or hip replacement 6 months in advance
  - Diagnosis of depression 3 month before any antidepressant is prescribed.

Case Study: Fraud Detection

In God we trust, all others bring data.

–William E. Deming

Sample Medical Record

Institute of Actuaries of India

(#21)

STURDY MEMORIAL HOSPITAL
211 Park Street
Attleboro, MA 02703

RILEY,JOSEPH AUGUSTUS
Medical Record Number: 030459
Account Number: 28568012
Service Date: 10/02/03

EMERGENCY CARE CENTER REPORT

CHIEF COMPLAINT: Left knee pain.

HISTORY OF PRESENT ILLNESS: This is a 49-year-old, white male with a history of COPD, hepatitis C, cirrhosis, and GERD who is complaining of left knee pain. The patient states on Saturday he slipped. He was falling back, he twisted, and had a little bit of pain in the knee. It got worse on Sunday and yesterday it became much more pronounced. He saw his primary care doctor and they were concerned about the swelling he had. He had an ultrasound of the leg to rule out DVT which he says was negative, but he is still having pain. The patient states he is unable to straighten his leg - it hurts too much. He is able to bear weight upon it. He denies any direct trauma to the knee. He denies any recent history of illness. Review of systems are negative.

PAST MEDICAL HISTORY:
1. COPD.
2. Cirrhosis.
3. Hepatitis C.
4. GERD.
5. Depression.

MEDICATIONS: See list.

ALLERGIES: None.

SOCIAL HISTORY: Negative for tobacco. Occasional alcohol.

PHYSICAL EXAMINATION:
VITAL SIGNS: Temperature is 96.7. Pulse is 82. Breathing rate is 20. Blood pressure is 124/70.
GENERAL: The patient is alert and cooperative. He is in mild distress secondary to pain in the left knee.
RIGHT LOWER EXTREMITY: Neurovascularly intact. No swelling. No tenderness.
LEFT LOWER EXTREMITY: There is mild, diffuse swelling of the leg. It is neurovascularly intact. There is tenderness in the infrapatellar region with moderate soft tissue swelling there and bogginess. The patient is unable to extend the left leg at the knee. He states this is not necessarily secondary to pain, but he is just unable to. He is using his arm to lift the leg.

INTERVENTION: An x-ray was performed of the left knee. No fracture or dislocation noted. I discussed the case with the on-call orthopedist, Dr. Fathallah, and my concern for an infrapatellar tendon rupture. He instructed me to put the patient into a knee immobilizer, crutches, and he will see the patient this morning in the office. The patient is being discharged on crutches with a knee immobilizer on the left leg, two Vicodin tablets now for pain and two to go.

Structured Data

	episode_id	episode	line_leve	qua	proce	Bill_Type	plac	reve	episode_	privateNe	episode_	episode_	episode_	billing_pr	billing	icd_diagn	all_icd
2	4548	74220	74220	1	--	Unknown	11	0	339.12	339.12	339.12	339.12	339.12	55905	1404	R6881	R6881;;;;
3	8246	93971	93971	1	--	Unknown	11	0	213.83	213.83	213.83	213.83	213.83	94598	3634	I872	I872;I800
4	13909	73630	73630	1	RT--	Unknown	11	0	138	138	151.394	151.394	151.394	94904	3738	M79671	M79671;
5	25285	73502	73502	1	--	Unknown	11	0	57.75	57.75	57.75	57.75	57.75	98188	5251	M25852	M25852;
6	3316	73221	73221	1	RT--	Unknown	11	0	428.35	428.35	428.35	428.35	428.35	60612	1960	M25521	M25521;
7	24630	73660	73660	1	FY--	Unknown	11	0	59.7	59.7	71.9	71.9	71.9	98101	4666	M25579	M25579;
8	5495	76830	76830	1	--	131	22	402	297.69	297.69	408.87	408.87	408.87	98029	4375	D251	D251;D2
9	17544	72050	72050	1	TC-LT-	Unknown	11	0	110.02	73.91	110.02	110.02	110.02	83616	2472	M25512	M25512;
10	17544	72050	72050	1	26--	Unknown	19	0	110.02	36.11	110.02	110.02	110.02	83616	2472	M50323	M50323;
11	2031	73721	73721	1	LT--	Unknown	11	0	627	627	775.03	775.03	775.03	94015	2886	S83012A	S83012A
12	3599	73521	73521	1	--	131	22	320	400	400	435.11	435.11	435.11	98122	5114	M5442	M5442;M
13	4504	77067	77067	1	--	131	22	403	341.25	341.25	383.77	383.77	383.77	94118	3198	Z1231	Z1231;;;;
14	29765	73522	73522	1	26--	Unknown	22	0	28.09	28.09	28.09	28.09	28.09	98052	4554	M545	M545;;;;
15	31704	72100	72100	1	--	131	22	320	20.47	0	20.47	20.47	20.47	98122	5114	M5442	M5442;M
16	31704	72100	72100	1	26--	Unknown	22	0	20.47	20.47	20.47	20.47	20.47	98122	5114	M5116	M5116;;;;
17	587	70553	70553	1	26--	Unknown	22	0	2139.3	881.3	2139.3	2139.3	2139.3	60540	1825	G253	G253;;;;
18	587	70553	70553	1	--	131	22	611	2139.3	1258	2139.3	2139.3	2139.3	60540	1825	G253	G253;;;;
19	587	70553	A9575	150	--	131	22	636	2139.3	0	2139.3	2139.3	2139.3	60540	1825	G253	G253;;;;
20	1531	72081	72081	1	--	131	22	320	824.71	693.1	824.71	824.71	824.71	94304	3311	Q763	Q763;;;;
21	1531	72081	72081	1	--	Unknown	11	0	824.71	131.61	824.71	824.71	824.71	94304	3311	Q763	Q763;;;;
22	1759	73721	73721	1	RT--	Unknown	11	0	703.58	703.58	823.64	823.64	823.64	98104	4856	S93401A	S93401A
23	1912	73562	73562	1	26-LT-	Unknown	22	0	655.35	29.61	655.35	655.35	655.35	94305	3319	S86912A	S86912A
24	1912	73562	73562	1	LT-	131	22	320	655.35	625.74	655.35	655.35	655.35	94305	3319	S86912A	S86912A

Structured Data

- user_insurance_id
- episode_id
- episode_code
- episode_sub_group
- episode_description
- episodeName
- line_level_procedure_cpt_code
- quantity
- procedure_modifiers
- CPT_description
- Bill_Type
- place_of_service_code
- revenue_code
- episode_cost
- privateNegoRate
- episode_cost_with_imported_cpt_recent
- episode_unit_cost_with_imported_cpt_recent
- episode_cost_with_imported_cpt_recent_unil
- service_date
- billing_provider_zip
- billing_provider_id
- icd_diagnosis_code_1
- all_icd
- age
- gender
- claims_medical_id
- prim_class_ep_code_ind
- unInsuredCharge
- paidByInsurance
- employeeResponsibility
- allowed_amount
- status_of_claim
- not_covered_amount
- cpt_Class
- required_cpt_class
- complete_ep_ind
- pc_tc_ind
- pc_tc_use_ind
- pc_tc_component
- required_pc_tc_comp
- complete_pc_tc_ind
- complete_ep_ind_combined
- still_incomplete_ep
- bilat_ind
- unit_price
- unit_price_outlr_ind
- imported_cpt
- imported_cpt_unit_cost
- imported_cpt_cost_pc_tc_comp
- unit_price_recent
- cost_recent

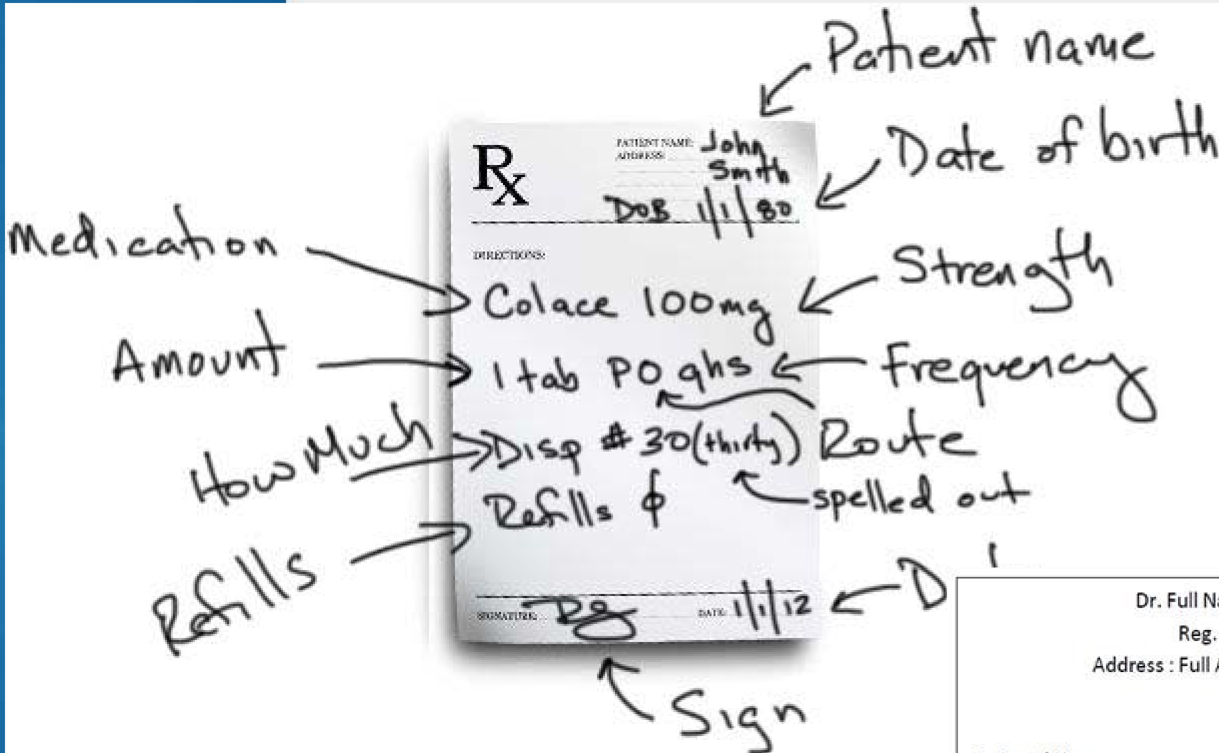
Data Analytics

- Pricing Variation
- Procedure Variation
  - Based on co-morbidities
- Practice variation
  - Outpatient/Inpatient
- Outcomes and quality
  - Readmission
- Gaps in care
- Real time intervention

PillStore

Institute of Actuaries of India

Split of healthcare expenditure by services		
Health Care services consumed	**Amount in Crores**	**Proportion**
Inpatient Curative Care	1,70,407	34.4%
Outpatient Curative Care	85,750	17.3%
Patient Transportation	21,604	4.4%
Laboratory and Imaging Services	21,315	4.3%
Prescribed Medicines	1,36,364	27.5%
Over the Counter Medicines	1,697	0.3%
Therapeutic Appliances & Medical Goods	792	0.2%
Preventive Care	34,033	6.9%
Others	7,742	1.6%
Governance and Health System Administration	15,483	3.1%
Total	**4,95,187**	**100.0%**



Institute of Actuaries of India

Dr. Full Name, M.B.B.S., M.D.
Reg. No.: GMC xxxx
Address : Full Address, Tel No.: xxxxxxx

Date : 29/8/2009

Patients' Name : ____________________ Sex ___
Patient's Address : ____________________________

Age ___

Rx

1. Valium 5 mg
   1 tab at night x 20 --- 20 tab

DISPENSED
Date : _____ Pharmacist :____
Name of Pharmacy
City

Doctor's usual sign,
Dated

Dr.'s Full name,
Qual(MBBS),Reg.No

DO NOT REFILL (DISPENSE ONLY ONCE)

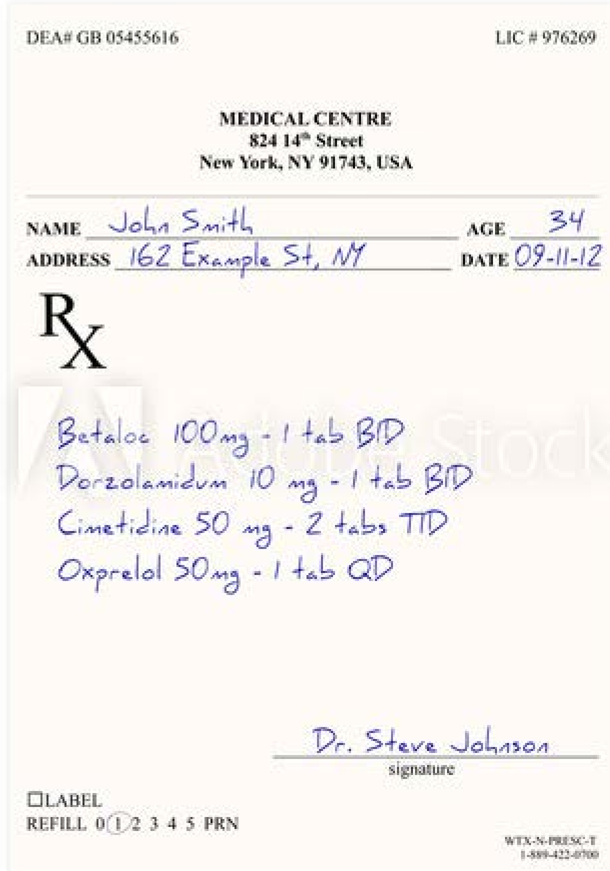
- Prescription on a letterhead, with Doctor's full name, Qualifications, GMC Reg. No., Full address, Tel. No.
- Date
- Patient's Full name & address
- Name of the drug and its potency, total quantity recommended
- Space for Pharmacy to put a "Dispensed Stamp"
- Usual signature of Doctor (not a scribble), & dated by Dr.
- Rubber Stamp of Dr. with name, Qual, Reg.No.
- To avoid refilling of prescription

Unstructured To Structured

Anatomical main group	XX
Therapeutic subgroup	XX-XX
Pharmacological subgroup	XX-XX-XX
Chemical subgroup	XX-XX-XX-XX
Chemical Substance	XX-XX-XX-XX-XX
Chemical Name	XX-XX-XX-XX-XX-XX
Chemical Form	XX-XX-XX-XX-XX-XX-XX
Chemical Route Of Administration	XX-XX-XX-XX-XX-XX-XX-XX
Chemical Strength	XX-XX-XX-XX-XX-XX-XX-XX-XX

Data Science Comes in Many Flavours

- Computer Vision
  - Image Fingerprinting
  - Image Classification
  - OCR
- Machine Learning
  - OCR
  - Document Identification
  - Image Classification



DEA# GB 05455616 LIC # 976269

MEDICAL CENTRE
824 14th Street
New York, NY 91743, USA

NAME John Smith AGE 34
ADDRESS 162 Example St, NY DATE 09-11-12

Rx

Betaloc 100mg - 1 tab BID
Dorzolamidum 10 mg - 1 tab BID
Cimetidine 50 mg - 2 tabs TID
Oxprelol 50mg - 1 tab QD

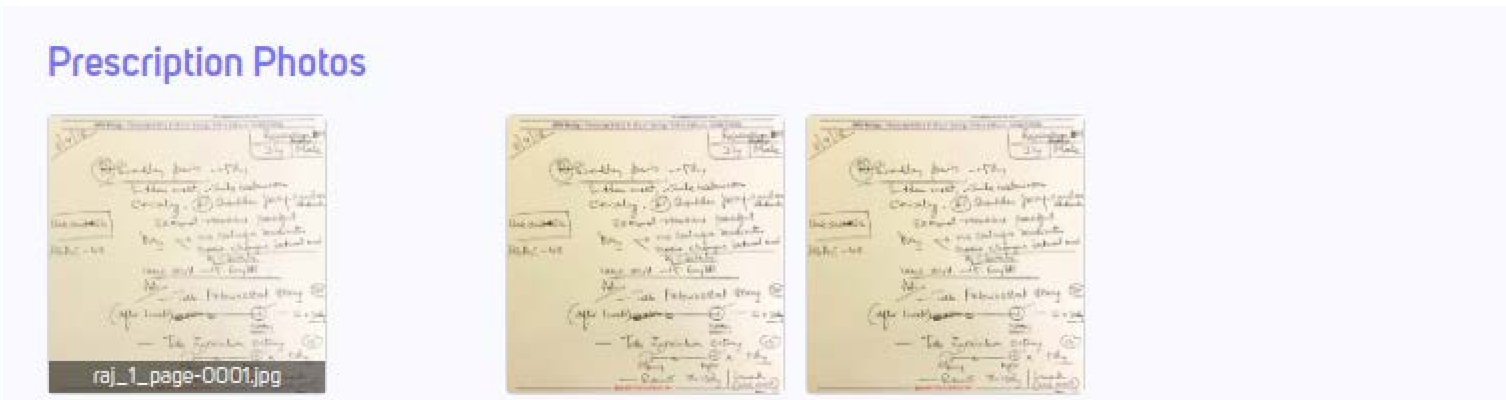
Dr. Steve Johnson
signature

☐LABEL
REFILL 0 1 2 3 4 5 PRN

WTX-N-PRESC-T
1-888-422-0700

#56617167

Image Fingerprinting / Image Classification



Original

Duplicate with 100% match

Duplicate with 73% match

Image Fingerprinting

- Image fingerprinting generated a unique code
- This code is used for identifying duplicates
- Now each prescription image is tied up with the patient insurance claim
- And so each claim is unique with image fingerprint, drug codes, patient profile etc

Image Classification

For this given problem, this can be done in two different approaches:

1. Computer vision
   a. Image fingerprinting
   b. OCR
2. Machine Learning
   a. May not be the right solution here
      i. Because of lack of initial data
      ii. Lots of manual labeling/annotation

Image Classification

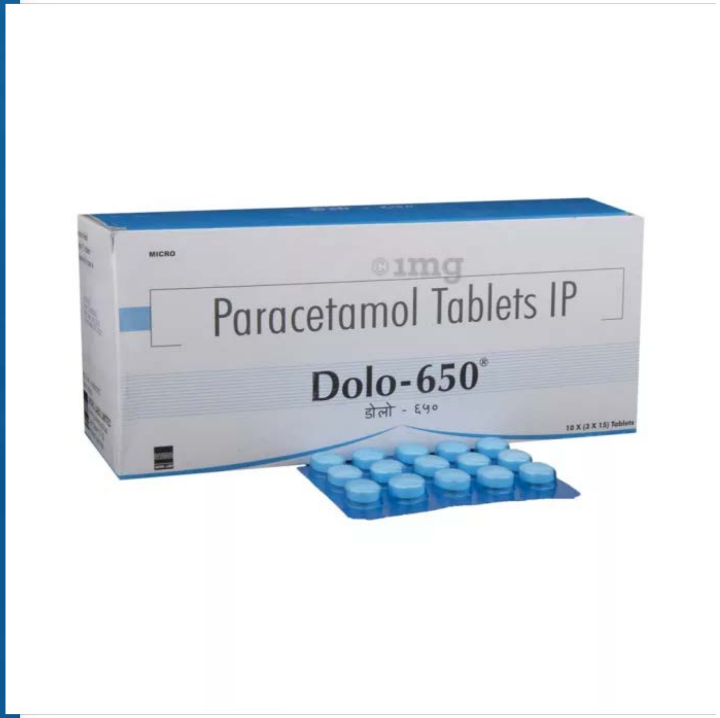
Advantages for Computer vision approach:

1. A strong solution to begin with - not too many false positives
2. Does not need data in one go
3. Can be improved as the data comes in

Disadvantages:

1. Manual intervention is needed at times when the picture mostly empty
2. Document identification may be needed to remove noise

Machine Learning



Summary

Key Takeaways!

- Data Science is analysis and study of data to generate meaningful insights
- Artificial Intelligence is a tool to implement Data Science
- AI has impacted every industry in every country including Health Insurance

Q&A

sumit@actuaria.in

raj@pillstore.in