What Are We Talking about When We Talk about Data as Insurers?

Xiaodong DENG http://seekingQED.com

Advanced Analytics at Manulife Singapore

March 2018

Short-bio

Education

- M.Sc in Mathematics, National University of Singapore, Singapore (2014-2016)
- B.Sc in Applied Mathematics, Beijing Forestry University, China (2010-2014)

Working Experience

- **2017 May - Present:** Data Analytics Function, Manulife Insurance Singapore
- **2016 April - 2017 April:** Data Analytics Specialist, AXA Insurance Singapore
- **2015 January - 2016 March:** Research Assistant, National University Health System

*: All views expressed here are my own and do not represent any entity or my employer.

Xiaodong DENG

Takeaways from the Competition

1. You may find data science immediately useful to you.

Predicting purchase of
a term-deposit product
through telemarketing campaign

"Yes, I'll buy."   "No. Thanks."

*Migrate to
our (your) scenarios*

- If one customer is preferred in terms of risk?
- If one claim involves fraud?
- If one customer should be preferred for cross-sell/up-sell?
- Predict the smoking status of applicants?
- You name it…

Nature: a binary-classification problem
Objective: to improve efficiency of marketing

Takeaways from the Competition

2. *Data/Features* are important (what data you have; what feature engineering can be done)

Example-1: What's the likelihood of a person to be pregnant given specific information?

- What if you only know age, education, occupation, etc?

- How about after I tell you the gender?

- How about after I tell you the waist size?

- How about after I tell you the urine test result?

Image source: https://www.bostonglobe.com/magazine/2015/05/07/women-really-can-have-all-soon-men-get-pregnant/g8qMuRYwiHL6PUwvacZ7PK/story.html

Takeaways from the Competition

2. *Data/Features* are important (what data you have; what feature engineering can be done)
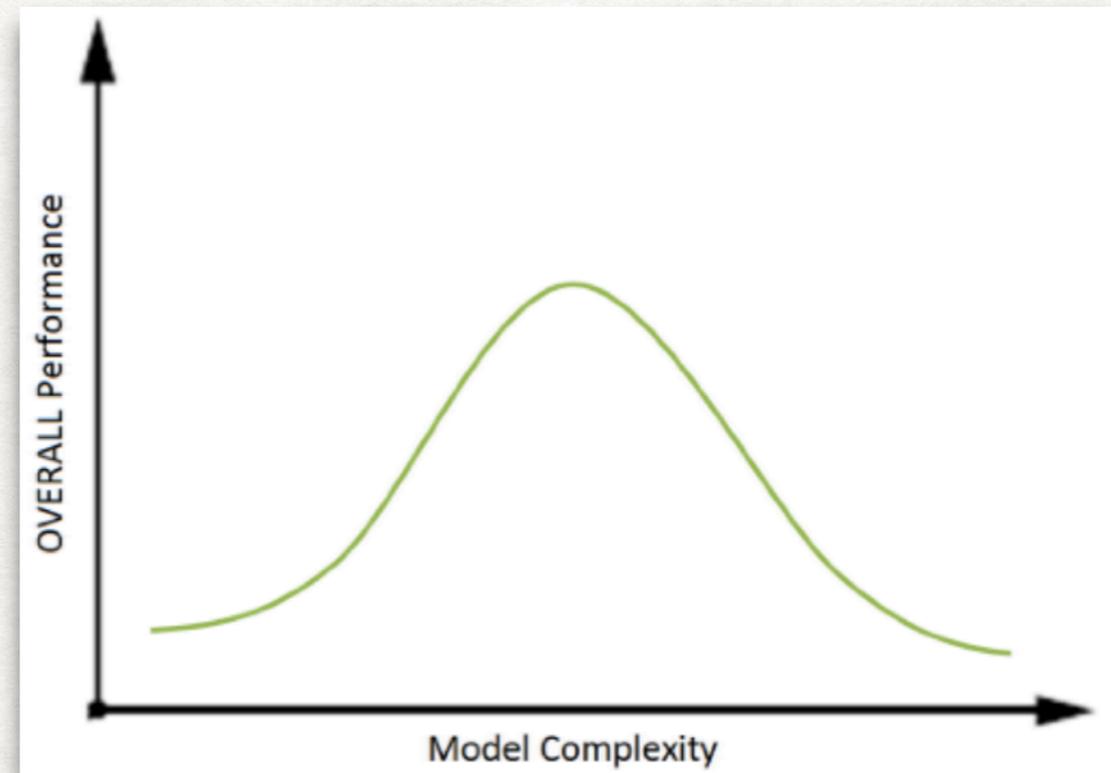
Example-2: New information from old data

ID	Spend on housing	Income	% of Spend on housing
1	1,000	3,000	33.3%
2	500	1,500	33.3%
3	3,000	50,000	6.0%
4	1,500	20,000	7.5%
5	5,000	50,000	10.0%
...	...	...	...

Xiaodong DENG

Takeaways from the Competition

3. It's always about finding the correct *tradeoff*.
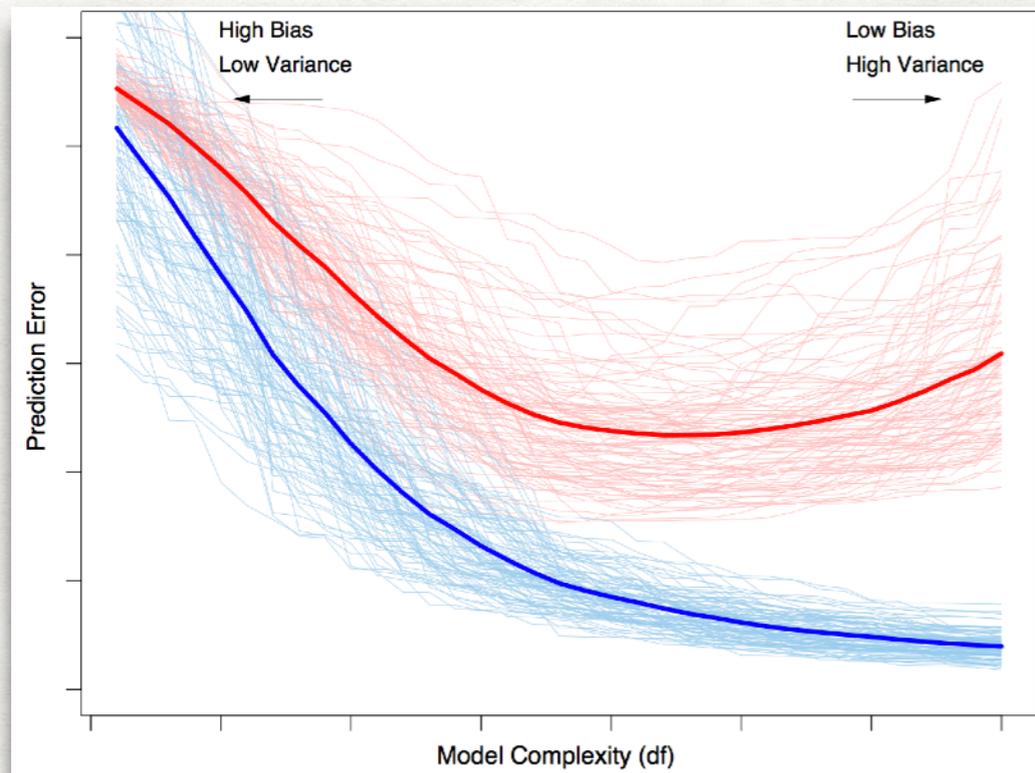
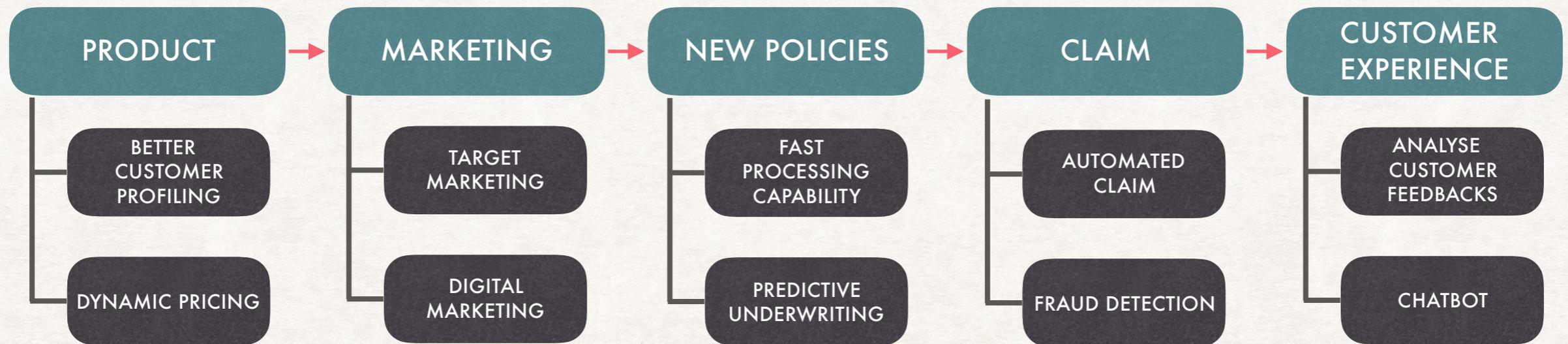
Bias-variance tradeoff*

Model complexity-(project) OVERALL performance tradeoff

*: Trevor Hastie, Robert Tibshirani, Jerome Friedman, *The Elements of Statistical Learning*, Second Edition, Springer

Xiaodong DENG

Applications of Data Science in Insurance Industry

PRODUCT → MARKETING → NEW POLICIES → CLAIM → CUSTOMER EXPERIENCE

PRODUCT
- BETTER CUSTOMER PROFILING
- DYNAMIC PRICING

MARKETING
- TARGET MARKETING
- DIGITAL MARKETING

NEW POLICIES
- FAST PROCESSING CAPABILITY
- PREDICTIVE UNDERWRITING

CLAIM
- AUTOMATED CLAIM
- FRAUD DETECTION

CUSTOMER EXPERIENCE
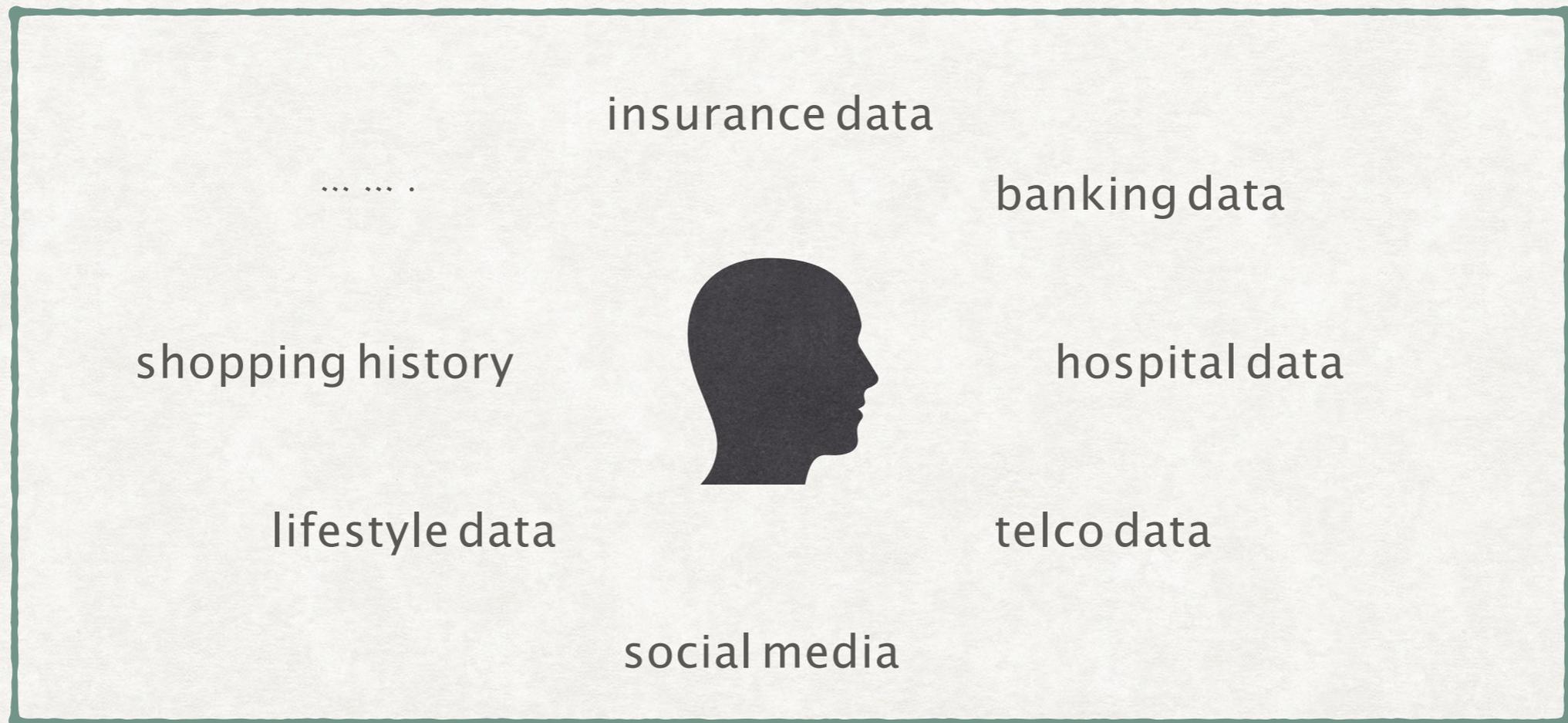- ANALYSE CUSTOMER FEEDBACKS
- CHATBOT

Xiaodong DENG

Applications of Data Science in Insurance Industry

PRODUCT

└ **BETTER CUSTOMER PROFILING**

Use "big data" to better Know Your Customers (KYC).

"Big" in terms of *dimension*, rather than *size*.

insurance data

… … .   banking data

shopping history   hospital data

lifestyle data   telco data

social media

Applications of Data Science in Insurance Industry

PRODUCT

└─ BETTER CUSTOMER PROFILING

Use "big data" to better Know Your Customers (KYC).

"Big" in terms of *dimension*, rather than *size*.

Singapore Insurance Companies Who Already Have This Advantage:

Company	Life Insurance	General Insurance	Others
Manulife	✓		banca collaboration with DBS
AXA	✓	✓	
income	✓	✓	Very wide business under NTUC group
Great Eastern	✓	✓	OCBC
…	…	…	…

Xiaodong DENG

Applications of Data Science in Insurance Industry

PRODUCT

DYNAMIC PRICING

Tailor your offer.

Flight Delay Compensation Insurance
Using data to predict the likelihood of a specific flight being delayed, in order to adjust the premium or stop selling for that flight.
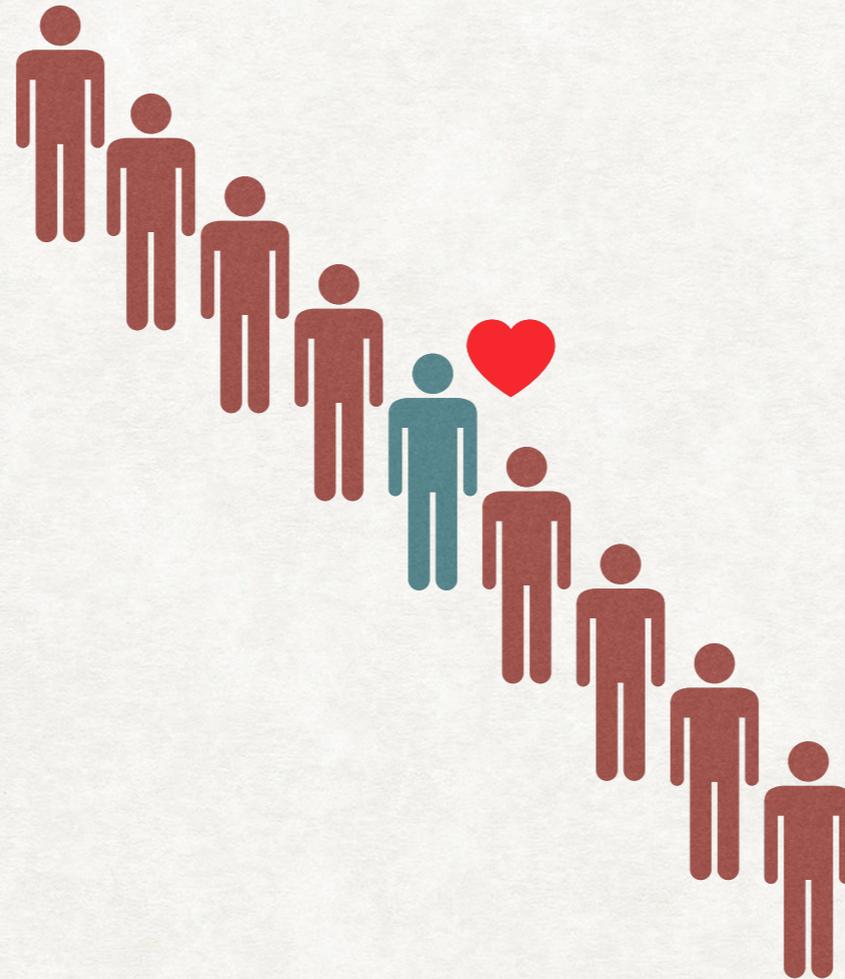
Pay As You Behave
Adjust the premium based on lifestyle or driving behaviour.

1. http://news.chubb.com/2017-09-08-Chubb-and-App-in-the-Air-launch-fully-automated-and-real-time-Flight-Delay-Insurance-in-partnership-with-Swiss-Re-and-FlightStats
2. http://www.straitstimes.com/business/invest/new-car-insurance-policy-tracks-driving-habit

Xiaodong DENG

Applications of Data Science in Insurance Industry

MARKETING

Where is your "Prince Charming"?

Applications of Data Science in Insurance Industry

NEW POLICIES

FAST PROCESSING CAPABILITY

Will you be worried when you have "too many" customers?

众安保险

(ZhongAn Insurance)

First Year of Operation

630 million+ policies underwritten
(about 1.72 million policies per day)

150 million clients

On One Single Day

128 million Yuan
(~26.6 million SGD)
on 2015-11-11

September 2017

Raised $1.5 billion on the HKSE;
HK's second largest IPO
&
Asia's largest fintech IPO
in 2017

1. https://www.the-digital-insurer.com/dia/zhong-an-chinas-first-complete-online-insurance-company/
2. http://hk.jrj.com.cn/2017/09/28104423181572.shtml
3. https://www.cnbc.com/2018/01/29/chinese-tech-unicorns-angle-for-2018-ipos.html

Xiaodong DENG

Applications of Data Science in Insurance Industry

NEW POLICIES

PREDICTIVE UNDERWRITING

How can we "bypass" medical checkup (in life insurance setting)?

Application → Medical Checkup → Policy Issued

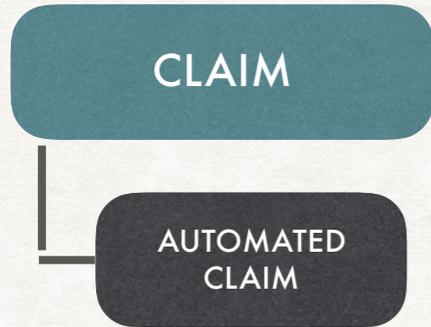Xiaodong DENG
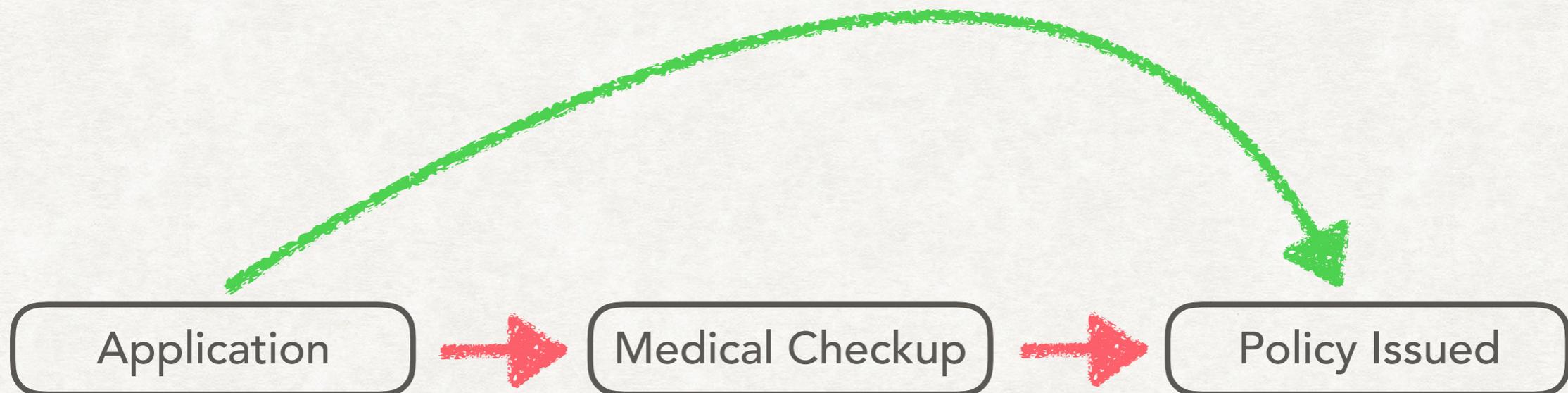
Applications of Data Science in Insurance Industry

NEW POLICIES

PREDICTIVE UNDERWRITING

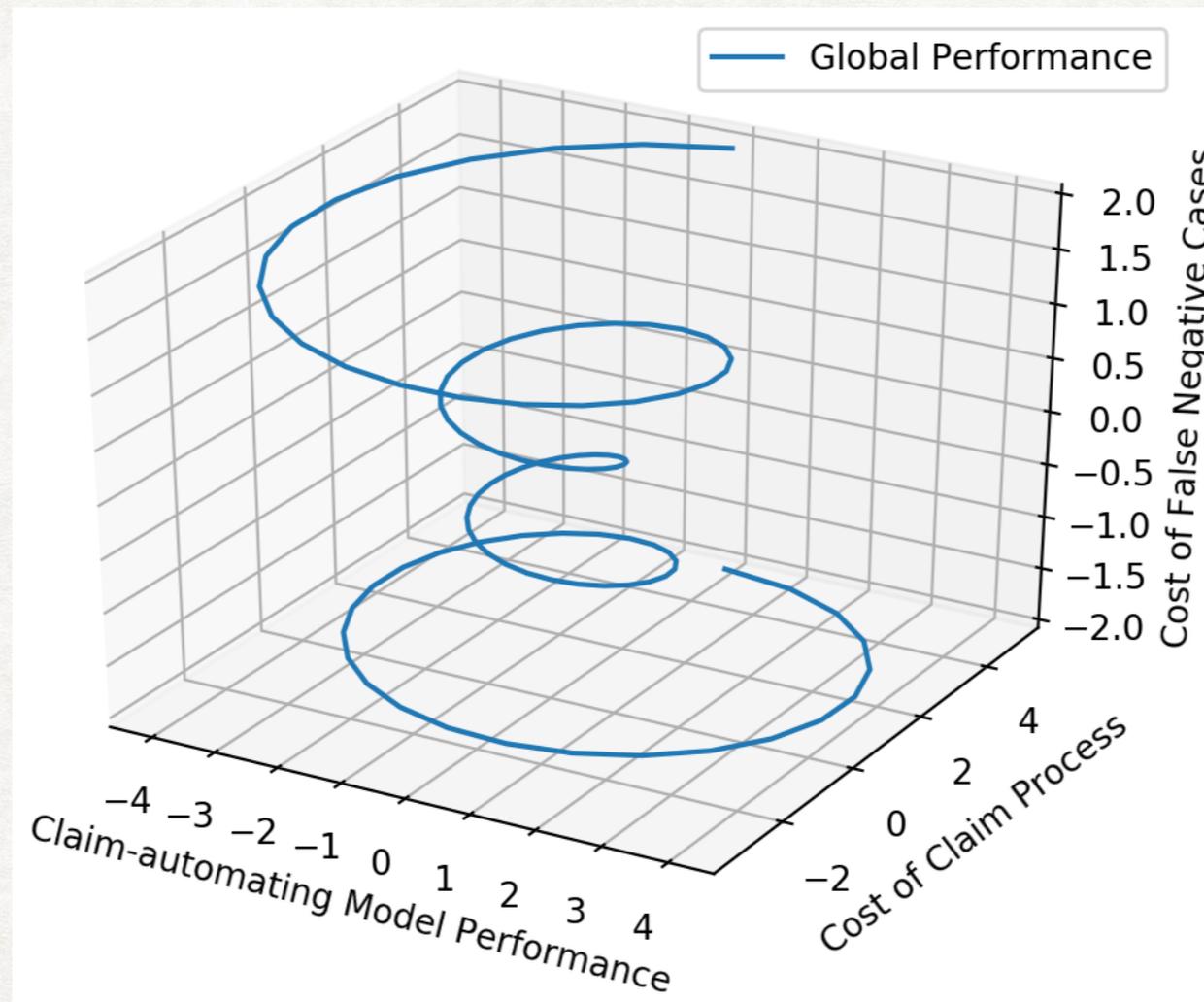How can we "bypass" medical checkup (in life insurance setting)?

	Auto Insurer	Life Insurer
Target Variable	Claims over six-month contract	Mortality experience over life of product (10, 20+ years)
Modeling Data	Underwriting requirements supplemented by third-party data	Underwriting requirements supplemented by third-party data
Frequency of Loss	Approximately 10 percent of drivers make claims annually	Typically, fewer than 1 first year death per 1,000 new policies issued
Business Action	Underwriting Decision	Underwriting Decision

Source: *Predictive Modeling for Life Insurance, Deloitte Consulting LLP*

1. https://www.soa.org/files/pdf/research-pred-mod-life-batty.pdf

Xiaodong DENG

Applications of Data Science in Insurance Industry

CLAIM

AUTOMATED CLAIM

What is the optimum tradeoff?

Bottlenecks/Limitations/Difficulties

Data Availability/Data Quality/Level of Digitisation

Bottlenecks/Limitations/Difficulties

Mindset/Skill Set

I dare you to say […] one more time

- Big Data
- Data Science
- Machine Learning
- Artificial Intelligence
- Hadoop
- Apache Spark
- Distributed Computing
- Data Lake
- Blockchain
- Deep Learning
- Neural Network
- ...

Bottlenecks/Limitations/Difficulties

Mindset/Skill Set

Are you really talking about data science? There are a few questions I was asked

- *"How can I track the traffic of my website?"*

  My answer: This doesn't have to relate to data science/data analytics. This is an IT question.
  You can use Google Analytics by embedding a JavaScript script into your web page source codes.

- *"We do have the scan copies of all the application forms. Why can't we use it?"*

  My answer: Again, this is an IT problem.
  If your process is not digitised enough, what we can do is very limited.
  OCR? It's not accurate enough (even myself can't read some forms)

- *"Why do you use mathematical modelling? We have data science now."*
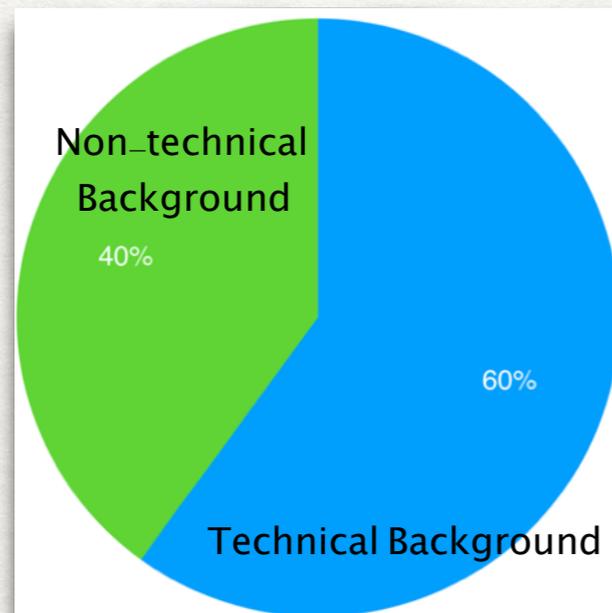
  My answer: Normally mathematical modelling tries to find an exact solution after making some assumptions.
  Data science uses the data we have to help obtain an "empirical" solution.
  What if you don't have data? What if it's possible to reason out the exact solution?

*: http://www.clickandinc.com/blog/wp-content/uploads/2013/01/Screen-shot-2013-01-24-at-12.55.52-PM.png

Bottlenecks/Limitations/Difficulties

Mindset/Skill Set

众安保险

(ZhongAn Insurance)

Non-technical Background
40%

Technical Background
60%

Employee Background

7%

5.25%

3.5%

1.75%

0%

2.8%   2.8%   6.3%
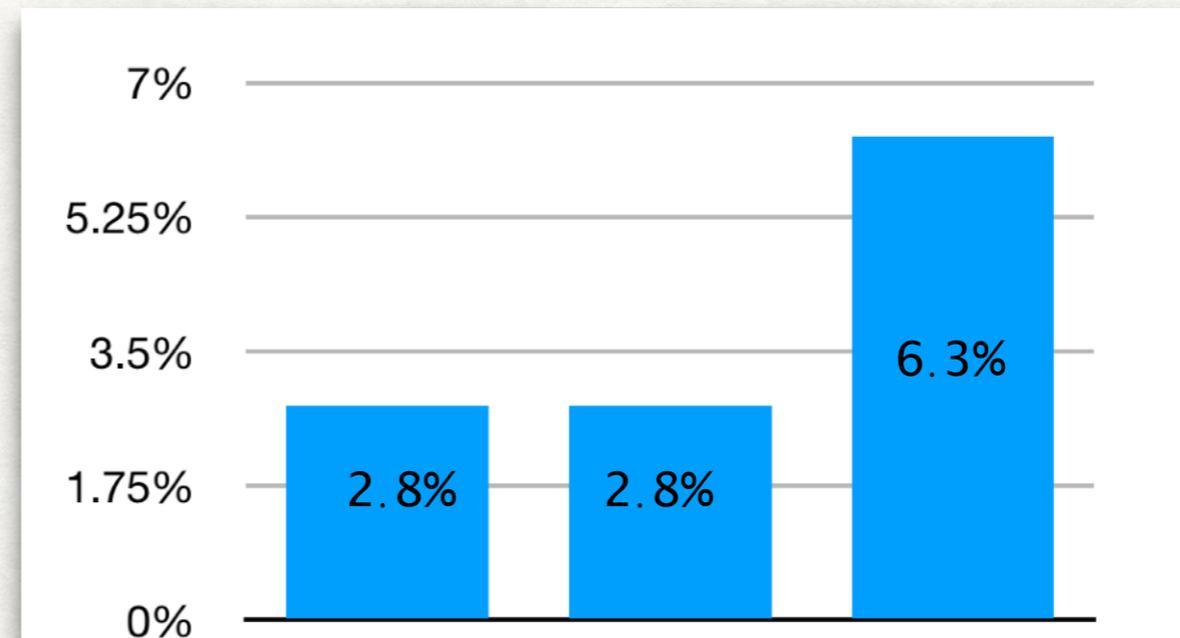
Spend on R&D / Total Premium Received

1. https://www.the-digital-insurer.com/dia/zhong-an-chinas-first-complete-online-insurance-company/
2. http://hk.jrj.com.cn/2017/09/28104423181572.shtml

Xiaodong DENG

Bottlenecks/Limitations/Difficulties

Mindset/Skill Set

R&D Intensity in 2017
(R&D expenditure as a percentage of Revenue)

Company	R&D Intensity
Amazon	11.8%
Alphabet	15.5%
Intel	21.5%
Samsung	7.6%
Microsoft	14.1%
Apple	4.7%
Facebook	21.4%
IBM	7.2%
Netflix	9.6%

1. https://www.strategyand.pwc.com/innovation1000

Bottlenecks/Limitations/Difficulties

Regulation/Privacy Protection

Are you considering regulators as your friends or someone stopping you?

Regulation is a must, and privacy protection should be one of our top priorities.

Bottlenecks/Limitations/Difficulties

Limited Contact Points with Customers (especially for life insurance)

Fast-food restaurant

Fine-dining places

sell more == sell better ?

Maybe we should ask ourselves: *how can we sell/serve **better**, rather than simply sell/serve **more**?*

There is no silver bullet.
Data science is not.
AI is not.
Blockchain is not.

(at least to me)

Thanks!

The PDF version of this slides can be downloaded at http://seekingQED.com/sas2018