How Machine Learning will transform businesses - Application in Precision Health

The Edmonton Data Management Chapter Association
Annual meeting
November 22, 2018
- There are measure choices to make any organization look like a leader.

- Since 2015, the Chinese national AI strategy has helped Tsinghua move from 10th to 2nd.

- Alberta is also growing, but more slowly, and has been in the top 5 since 2000.

- In 2017 Alberta rose to 2nd, and was displaced by Tsinghua in 2018.
Pan-Canadian

- Founded in 2002
- 14 principal investigators (2018)
- About 80 graduate students
- Chinook: Checker solved in 2007
- Heads-up limit Texas hold’em solved in 2015
- AlphaGo started with David Silver PhD

Edmonton
- Amii: Alberta Machine Intelligence Institute
  - University of Alberta

Toronto
- Vector Institute:
  - University of Toronto
  - University of Waterloo

Montreal
- MILA:
  - Montreal Institute for Learning Algorithms
  - Université de Montréal
  - McGill University

Rich Sutton
Reinforcement Learning

Geoff Hinton
Deep Learning

Yoshua Bengio
Deep Learning
One of the reasons Google acquired DeepMind is their Atari Game learner from pixels. It started at Amii - UofA with Michael Bowling and Larry Page - 2014.
West Africa 2014 Outbreak

A "mysterious" deadly disease silently spreading
Cholera?? (found bacteria)

Started in Meliandou, Guinea in late 2013 and was only identified as virus in March 2014.
How did Emile - or patient zero - contract the disease in the first place?
Using biological characteristics (57 variables) of 21 bat species known to carry filoviruses, with a sample of 1116 bats, a model is built to discriminate known and unknown filovirus carriers.

Fruit eating bats
Ebola reservoir

Insect eating bats
Filovirus-positive bat species estimated with 87% accuracy.

Machine learning and computer modeling can help predict which species have the potential to cause future epidemics.
Machine learning and computer modeling can help predict which species have the potential to cause future epidemics. Using biological characteristics (57 variables) of 21 bat species known to carry filoviruses, with a sample of 1116 bats, a model is built to discriminate known and unknown filovirus carriers.

Fruit eating bats
Ebola reservoir

Insect eating bats

Filovirus-positive bat species estimated with 87% accuracy.

Undiscovered Bat Hosts of Filoviruses, BA. Han, HP. Schmidt, LW. Alexander, SE. Bowden, DTS. Hayman, JM. Drake, PLOS Neglected Tropical Diseases, 2016, 10:1371
Filovirus distribution and diversity is greater than previously thought.

From Hypothesis testing to hypothesis generation.
Search for molecules inhibitors of Ebola virus

• Some pharmaceutical ingredients have been found to have anti-EBOV activity in vitro.
• There are millions of additional commercially available molecules that could be screened.
• How to prioritize the compounds for testing?
• **Machine Learning** was used to score drugs as potential EBOV inhibitor.
  • Three of the highest scoring molecules where not in the training set, Quinacrine, Puronaridine and Tilorone, were tested in vitro and had relevant EC 50 values.
### Hypothetical telecom data

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Phone Number</th>
<th>City</th>
<th>Plan</th>
<th>Avg. 3m Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>John Smith</td>
<td>647 225 8085</td>
<td>Toronto</td>
<td>2y</td>
<td>$12</td>
</tr>
<tr>
<td>3</td>
<td>John Simon</td>
<td>780 886 5053</td>
<td>Edmonton</td>
<td>3y</td>
<td>$369.00</td>
</tr>
<tr>
<td>4</td>
<td>Randy Regal</td>
<td>705 234 6767</td>
<td>Toronto</td>
<td>3y</td>
<td>$77.10</td>
</tr>
<tr>
<td>5</td>
<td>Mary Tasear Smith</td>
<td>780 334 3434</td>
<td>Edmonton</td>
<td>3y</td>
<td>$369.00</td>
</tr>
<tr>
<td>6</td>
<td>Susan Wilcox</td>
<td>780 291 6063</td>
<td>Edmonton</td>
<td>3y</td>
<td>$131.00</td>
</tr>
<tr>
<td>7</td>
<td>Martha Witherby</td>
<td>780 322 9768</td>
<td>Edmonton</td>
<td>3y</td>
<td>$459.37</td>
</tr>
<tr>
<td>11</td>
<td>Kurt Locke</td>
<td>780 654 1121</td>
<td>Edmonton</td>
<td>3y</td>
<td>$830.00</td>
</tr>
<tr>
<td>12</td>
<td>Kent Wafegert</td>
<td>647 631 0348</td>
<td>Toronto</td>
<td>2y</td>
<td>$38.78</td>
</tr>
<tr>
<td>13</td>
<td>Brent Mavka</td>
<td>403 566 7372</td>
<td>Calgary</td>
<td>2y</td>
<td>$299.39</td>
</tr>
<tr>
<td>17</td>
<td>Wayne Jones</td>
<td>780 236 3006</td>
<td>Edmonton</td>
<td>3y</td>
<td>$336.00</td>
</tr>
<tr>
<td>18</td>
<td>Patty Klien</td>
<td>780 550 1819</td>
<td>Edmonton</td>
<td>3y</td>
<td>$50.18</td>
</tr>
<tr>
<td>20</td>
<td>Morris Slevchuk</td>
<td>780 434 6280</td>
<td>Edmonton</td>
<td>3y</td>
<td>$628.01</td>
</tr>
<tr>
<td>21</td>
<td>Patrick Klum</td>
<td>403 337 9291</td>
<td>Calgary</td>
<td>3y</td>
<td>$33.79</td>
</tr>
<tr>
<td>22</td>
<td>Wilma Renton</td>
<td>780 118 2388</td>
<td>Edmonton</td>
<td>3y</td>
<td>$89.11</td>
</tr>
<tr>
<td>24</td>
<td>Ben Rikon</td>
<td>403 261 3134</td>
<td>Calgary</td>
<td>2y</td>
<td>($26.23)</td>
</tr>
<tr>
<td>33</td>
<td>Natalie May</td>
<td>403 409 6223</td>
<td>Calgary</td>
<td>3y</td>
<td>$0.96</td>
</tr>
<tr>
<td>25</td>
<td>Maggie Wong</td>
<td>226 882 0911</td>
<td>Edmonton</td>
<td>2y</td>
<td>$89.11</td>
</tr>
<tr>
<td>26</td>
<td>Karen Pollonta</td>
<td>403 566 7372</td>
<td>Calgary</td>
<td>2y</td>
<td>$299.29</td>
</tr>
<tr>
<td>50</td>
<td>Mary Tasear Smith</td>
<td>780 334 3434</td>
<td>Edmonton</td>
<td>3y</td>
<td>$369.00</td>
</tr>
<tr>
<td>27</td>
<td>Susan Wilcox</td>
<td>780 291 6063</td>
<td>Edmonton</td>
<td>2y</td>
<td>$131.00</td>
</tr>
<tr>
<td>28</td>
<td>Wilma Renton</td>
<td>780 118 2388</td>
<td>Edmonton</td>
<td>3y</td>
<td>$89.11</td>
</tr>
<tr>
<td>29</td>
<td>Karen Pollonta</td>
<td>403 566 7372</td>
<td>Calgary</td>
<td>2y</td>
<td>$299.29</td>
</tr>
<tr>
<td>31</td>
<td>Monica Kwalshuck</td>
<td>403 210 4448</td>
<td>Calgary</td>
<td>3y</td>
<td>$1044.48</td>
</tr>
<tr>
<td>32</td>
<td>Kent Wafegert</td>
<td>647 631 0348</td>
<td>Toronto</td>
<td>3y</td>
<td>$38.78</td>
</tr>
<tr>
<td>33</td>
<td>Natalie May</td>
<td>403 409 6223</td>
<td>Calgary</td>
<td>3y</td>
<td>$0.96</td>
</tr>
</tbody>
</table>

### Assumption:
Customers are independent
Values are identically distributed

- **19 customers up for plan renewal**
- Which one to renew?
- Which one to give incentive to stay?

Sort by profit in the last 3 months
Do not renew or give incentive if profit < $50 (?)

- **6 least profitable customers**
Could be the wrong decision

**Motivating Example about Integration**

- **Not enough profit**
- **Plan**
  - Avg. 3m Profit
    - 3y: ($26.23)
    - 2y: ($12)
    - 3y: $0.96
    - 3y: $33.79
    - 1y: $50.18

**Hypothetical telecom data**

19 customers up for plan renewal
Which one to renew?
Which one to give incentive to stay?
<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Phone Number</th>
<th>City</th>
<th>Plan</th>
<th>Avg. 3m Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>Ben Rikon</td>
<td>403 262 3134</td>
<td>Calgary</td>
<td>3y</td>
<td>($26.23)</td>
</tr>
<tr>
<td>1</td>
<td>John Smith</td>
<td>647 225 0088</td>
<td>Toronto</td>
<td>2y</td>
<td>($12)</td>
</tr>
<tr>
<td>33</td>
<td>Natalie May</td>
<td>403 409 6223</td>
<td>Toronto</td>
<td>2y</td>
<td>$0.96</td>
</tr>
<tr>
<td>22</td>
<td>Wilma Renton</td>
<td>780 118 3388</td>
<td>Edmonton</td>
<td>3y</td>
<td>$8.00</td>
</tr>
<tr>
<td>21</td>
<td>Patrick Klum</td>
<td>403 337 9291</td>
<td>Calgary</td>
<td>3y</td>
<td>$33.79</td>
</tr>
<tr>
<td>12</td>
<td>Kent Wafegert</td>
<td>647 631 0348</td>
<td>Toronto</td>
<td>3y</td>
<td>$38.78</td>
</tr>
<tr>
<td>18</td>
<td>Patty Klein</td>
<td>780 510 1819</td>
<td>Edmonton</td>
<td>3y</td>
<td>$50.18</td>
</tr>
<tr>
<td>34</td>
<td>Aly Huffington</td>
<td>403 255 0304</td>
<td>Calgary</td>
<td>3y</td>
<td>$55.04</td>
</tr>
<tr>
<td>20</td>
<td>Chris Cristie</td>
<td>403 844 4423</td>
<td>Calgary</td>
<td>2y</td>
<td>$64.11</td>
</tr>
<tr>
<td>32</td>
<td>Fred Courou</td>
<td>416 773 2234</td>
<td>Toronto</td>
<td>3y</td>
<td>$73.22</td>
</tr>
<tr>
<td>23</td>
<td>Ryan Waters</td>
<td>403 715 5550</td>
<td>Calgary</td>
<td>3y</td>
<td>$75.50</td>
</tr>
<tr>
<td>1</td>
<td>John Smith</td>
<td>780 510 1819</td>
<td>Edmonton</td>
<td>3y</td>
<td>$77.10</td>
</tr>
<tr>
<td>30</td>
<td>Gunther Twallaby</td>
<td>403 778 3040</td>
<td>Calgary</td>
<td>3y</td>
<td>$78.32</td>
</tr>
<tr>
<td>26</td>
<td>Margaret Wong</td>
<td>226 882 0911</td>
<td>Toronto</td>
<td>3y</td>
<td>$89.11</td>
</tr>
<tr>
<td>25</td>
<td>Sun Liu</td>
<td>226 690 4241</td>
<td>Toronto</td>
<td>3y</td>
<td>$90.42</td>
</tr>
<tr>
<td>9</td>
<td>Wanda Rhymes</td>
<td>403 441 2534</td>
<td>Calgary</td>
<td>3y</td>
<td>$92.00</td>
</tr>
<tr>
<td>28</td>
<td>Karen Pottons</td>
<td>403 750 9201</td>
<td>Calgary</td>
<td>3y</td>
<td>$92.75</td>
</tr>
<tr>
<td>7</td>
<td>Susan Wilcox</td>
<td>780 291 6063</td>
<td>Edmonton</td>
<td>3y</td>
<td>$131.00</td>
</tr>
<tr>
<td>3</td>
<td>John Simon</td>
<td>780 886 5053</td>
<td>Edmonton</td>
<td>3y</td>
<td>$189.45</td>
</tr>
<tr>
<td>17</td>
<td>Wayne Jones</td>
<td>780 236 3026</td>
<td>Edmonton</td>
<td>3y</td>
<td>$236.05</td>
</tr>
<tr>
<td>19</td>
<td>Brent Mavka</td>
<td>403 566 7372</td>
<td>Calgary</td>
<td>3y</td>
<td>$299.20</td>
</tr>
<tr>
<td>10</td>
<td>Mary Tasear Smith</td>
<td>780 334 3434</td>
<td>Edmonton</td>
<td>3y</td>
<td>$369.00</td>
</tr>
<tr>
<td>16</td>
<td>Brian Olson</td>
<td>403 939 7574</td>
<td>Calgary</td>
<td>3y</td>
<td>$436.78</td>
</tr>
<tr>
<td>8</td>
<td>Martha Witherby</td>
<td>780 322 9768</td>
<td>Edmonton</td>
<td>3y</td>
<td>$459.37</td>
</tr>
<tr>
<td>14</td>
<td>Kim Cho</td>
<td>780 434 2399</td>
<td>Edmonton</td>
<td>3y</td>
<td>$542.00</td>
</tr>
<tr>
<td>20</td>
<td>Morris Slevich</td>
<td>780 434 6280</td>
<td>Edmonton</td>
<td>3y</td>
<td>$628.00</td>
</tr>
<tr>
<td>31</td>
<td>Jane Smith</td>
<td>780 233 5045</td>
<td>Edmonton</td>
<td>3y</td>
<td>$673.33</td>
</tr>
<tr>
<td>2</td>
<td>Sue Burns</td>
<td>416 345 6060</td>
<td>Toronto</td>
<td>3y</td>
<td>$724.00</td>
</tr>
<tr>
<td>19</td>
<td>Greg Aderan</td>
<td>403 332 7468</td>
<td>Calgary</td>
<td>3y</td>
<td>$746.82</td>
</tr>
<tr>
<td>13</td>
<td>Megan Potink</td>
<td>780 432 5623</td>
<td>Edmonton</td>
<td>3y</td>
<td>$802.00</td>
</tr>
<tr>
<td>11</td>
<td>Kurt Locke</td>
<td>780 654 1121</td>
<td>Edmonton</td>
<td>3y</td>
<td>$830.00</td>
</tr>
<tr>
<td>10</td>
<td>Julie Austini</td>
<td>403 223 7654</td>
<td>Calgary</td>
<td>3y</td>
<td>$983.12</td>
</tr>
<tr>
<td>31</td>
<td>Monica Kwaishuck</td>
<td>403 210 4448</td>
<td>Calgary</td>
<td>3y</td>
<td>$1,044.48</td>
</tr>
<tr>
<td>27</td>
<td>Joe Garther</td>
<td>416 224 1109</td>
<td>Toronto</td>
<td>3y</td>
<td>$1,100.10</td>
</tr>
</tbody>
</table>

34 customers interconnected with the 19 to renew. Which one to renew? Which one to give incentive to stay?

Inter-call network with call frequency
Additional data was required: Data Linking and Integration
Inter-call network with call frequency

Global centrality based PageRank
Inter-call network with call frequency

Community Mining
Community Mining

Centrality per community

Dropping Natalie: Risk = $3145.32
Community Mining

Centrality per community

Dropping John: Risk = $6324.14
Exploiting additional data and sophisticated analysis could give a different perspective and provide unexpected insights leading to competitive advantage.
Predicting residual value before auction

“Personalized” Residual Value Based on specific features

Fig 4. Auction sales Model XYZ: test dataset scatter plot
Outline

What is AI?

What is ML?

Why Now?

Precision Health

Examples

Outline
What is Artificial Intelligence?

What is Machine Learning?

ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.

Computer mimics human intelligence

Hyper optimism
High expectation for significant disruptions

Disillusionment

1950’s 1960’s 1970’s
What is Artificial Intelligence?

Artificial Intelligence

Early artificial intelligence stirs excitement.

What is Machine Learning?

Machine Learning

Machine learning begins to flourish.

Deep Learning

Deep learning breakthroughs drive AI boom.
What is Artificial Intelligence?
What is Machine Learning?

Machine Learning
- Supervised Learning
- Deep Learning
- Unsupervised Learning
- Reinforcement Learning

Artificial Intelligence
- Knowledge Representation
- Planning
- Natural Language Processing
- Reasoning
- Computer vision
- DL ⊆ SL ⊆ ML ⊆ AI ⊆ CS

High performance & parallel computing
- Networking
- Software Engineering
- Graphics
- Database Management Systems
- Security and Encryption
- Algorithm Analysis
- HCI

Computer Science
What is Artificial Intelligence?

Artificial Intelligence is the computational part of the ability to achieve goals in the World.

– John McCarthy

Artificial Intelligence is an enhancement of human creativity.

“Artificial Intelligence is the science of making machines do things that would require intelligence if done by men.”

-- Marvin Minsky

Artificial Intelligence is a discipline striving to get machines behave intelligently by performing tasks that would normally require human-level intelligence.
Why the resurgence of AI?

- Advancement in High Performance Computing
- Development of Powerful Algorithms
- Availability of Massive Data
What Happens in an Internet Minute?

1,572,877 GB of global IP data transferred

### Where is the Data Coming From?

- **$400 Million** during Alibaba peak day sales
- **438,801** Wiki page views
- **34.7 Million** instant messages (MIM) sent
- **194,064** app downloads
- **38,194** photos uploaded
- **57,870** page views
- **10 Million** in sales
- **31,773** hours of music played
- **347,222** Tweets
- **3.3 Million** pieces of content shared
- **6.9 Million** messages sent
- **10 Million ads displayed**
- **4.1 Million** searches
- **138,889** hours of video watched
- **23,148** hours of video uploaded
- Netflix + Youtube = more than ½ of all traffic

### And Future Growth is Staggering

- By 2017, mobile traffic will have grown **13X** in just 5 years
- In 2017, there will be **3X** more connected devices than people on Earth
- All digital data created reached **4 zettabytes** in 2013

THE WORLD’S EXPANDING PRODUCTION OF DATA

The world’s data supply (e.g., from digital books, photos, X-rays) grew exponentially in recent decades. In 2007, there were enough data produced by everyone in the world to fill a stack of CD-ROMs (1.2 mm thick per CD) stretching from the earth to beyond the moon—and that was six years ago. The stack of CD-ROMs would stretch much farther today.

2005 2014 2016

128MB ➔ 512 GB
4000 X improvement (13 years)

The 64 GB would have cost $1,536,000,000 in 1956

An improvement of (in terms of price) 76 million times (62 years)
Why the resurgence of AI?

- Advancement in High Performance Computing
- Availability of Massive Data
- Development of Powerful Algorithms
- Availability of Massive Data
IBM Summit
Oak Ridge National Laboratory
June 2018 (122 PetaFlops)

Sunway TaihuLight
National Supercomputing Center in Wuxi, China
2016 (94 PetaFlops) $275 MM
Samsung S9
Exynos: octa-core
2.7 GHz, 727 GFlops

iPhone X
A11 2.39 GHz, 600 GFlops
1.1 TeraFlops
10,000 Intel Pentium Pro
size: Tennis-Court
Power: 850 kW
Energy consumption: 800 houses
Cost: $55 million
206,000 X
In 25 years

06/1997: #1 = 1.1 TFlop/s

IBM Summit
Oak Ridge National Laboratory
June 2018 (122 PetaFlops)

CM-5/1024 Thinking Machines
Los Alamos National Laboratory
Nov. 1993 (59 GigaFlops)
128MB → 512 GB
4000 X improvement (13 years)

An improvement of (in terms of price)
76 million times (62 years)
Why the resurgence of AI?

- Advancement in High Performance Computing
- Development of Powerful Algorithms
- Availability of Massive Data

Deep Learning
What is Deep Learning?

• Just a rebranding of Neural Networks
• Neural Network = Layered collections of interconnected simple computing nodes. Each node applies a non-linearity to the weighted sum of its predecessors

DNN (Deep NN): fully connected
CNN (Convolutional NN): break components into sub-components
RNN (Recurrent NN): maintains state across a sequence of inputs
LSTM (Long Short-Term Memory): remembers values over time
Image Classification Accuracy Over Time

Accuracy:
- 100%
- 95%
- 90%
- 85%
- 80%
- 75%
- 70%

Year:
- 2009
- 2010
- 2011
- 2012
- 2013
- 2014
- 2015

Accuracy:
- 70.5%
- 71.8%
- 74.2%

Typical Human: 90%
Human Expert: 95%

14M Images. Crowdsourced labels. 22k Categories (plant, animal, device, food, structure, person)
Image Classification Accuracy Over Time

Deep Learning used

- Human Expert
- Typical Human

Accuracy

- 100%
- 95%
- 90%
- 85%
- 80%
- 75%
- 70%

Year

- 2009
- 2010
- 2011
- 2012
- 2013
- 2014
- 2015

70.5% 71.8% 74.2%

14M Images. Crowdsourced labels. 22k Categories (plant, animal, device, food, structure, person)
Why the success?

- Image Recognition
- Speech Recognition
- Object Detection
- Machine Translation
- Question and Visual Answering

Loud and clear
Speech-recognition

Switchboard corpus is a collection of telephone conversations widely used in speech-recognition systems and research papers.
What is Machine Learning?

• Machine Learning provides means to machine programs to learn from large data, interpret the trends in the data and adapt to the data.

• In the field of Machine Learning, Deep Learning is one of many paradigms of supervised learning (requiring training data sets).

• Reinforcement Learning is also Machine Learning.

• AI is much more than Deep Learning.
Concrete Machine Learning Applications

- Credit Approval
- Spam detection
- Suspicious credit card transaction (Falcon fraud assessment system)
- Diagnosis
- Prognosis
- Treatment recommendation
- Precision Medicine
- Noninvasive brain-machine interface
Reinforcement Learning

Intelligent Agent

ACTION

World (Environment)

REWARD

STATE

EXPERIENCE

Learns a policy:
sequence of actions to maximize $\sum$ rewards

Rich Sutton
The Pioneer in Reinforcement Learning
Adaptive Prosthetics

Bento Arm & HANDi Hand

Modular Prosthetic Limb

Targeted Reinnervation Surgery
Hebert et al., 2014

Dawson et al., MEC, 2014
Industries are ripe for disruption?

- Any with lots of data in computer readable format
- Banking & Finance
- Capital Markets
- Insurance
- Accounting & Audit
- …

Increase efficiency and effectiveness

- Reactive
- Uniform
- Proactive
- Preventive
- Personalized
Public Sector examples

Pittsburgh:

Cincinnati:
- Optimize medical emergency response. Recommend dispatcher appropriate response for medical emergency (treated on-site or taken to hospital)

Other:
- ML to decide which restaurant to inspect
- Crime Prediction (dashboard of key metrics)
Public Sector other possibilities

- Automatically routing/forwarding incoming e-mail to the corresponding officials.
- Initial screening of emergency calls. Would allow emergency call center employees to engage more with callers requiring human attention. (Chatbot otherwise)
- Tenant Assistance / Eviction Fighter Chatbot
- What’s is happening in my city? Social media monitoring
Foundations for AI

• What data do you have?
• What is ground truth?
• What result do I want?
• What actions can I take?

1. Identify and understand the nature of the problem we want to solve;
2. Identify the type of data available to address the problem;
3. Annotate and label data for training
4. Evaluate the predictions and understand the risks of the system before deployment.

• What processes are automated?
• Where do we use data to make decisions?
• Where and or how do you currently use prediction?
What can we do with all this?

Precision Health
What do we do in medicine today?

Medicine is inherently “personal” but different people with the same symptoms are treated the same way.
Precision Medicine...

Towards precision medicine …

Treatment tailored to individual patients
Personalized Medicine

- Use of genetic information
  - Micro-array data analysis + SNP
  - Genomics, Metabolomics, Microbiome, etc.

- Definition now encompasses personalization beyond just genetics or omics
How can this be done?

Machine Learning

<table>
<thead>
<tr>
<th>α-catenin</th>
<th>β-catenin</th>
<th>E-cadherin</th>
<th>...</th>
<th>p120</th>
<th>age</th>
<th>size</th>
<th>...</th>
<th>pten</th>
<th>recur?</th>
</tr>
</thead>
<tbody>
<tr>
<td>m c n</td>
<td>m c n</td>
<td>m c n</td>
<td></td>
<td>m c n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 3 0</td>
<td>2 0 4</td>
<td>0 0 2</td>
<td></td>
<td>2 0 0</td>
<td>60 4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 1 4</td>
<td>3 0 2</td>
<td>0 4 0</td>
<td></td>
<td>0 2 1</td>
<td>2 70</td>
<td>2</td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>1 4 3</td>
<td>0 2 0</td>
<td>4 0 0</td>
<td></td>
<td>2 1 2</td>
<td>62 6</td>
<td>1</td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>1 2 0</td>
<td>1 4 3</td>
<td>0 2 0</td>
<td></td>
<td>4 0 0</td>
<td>71 2</td>
<td>2</td>
<td></td>
<td></td>
<td>N</td>
</tr>
</tbody>
</table>

Historical Patients

Novel patient

Classifier

Learner

recur? N
Microarray Analysis

Biopsy

DNA Microarray

N = 33,000 genes per Patient

<table>
<thead>
<tr>
<th>g1</th>
<th>g2</th>
<th>g3</th>
<th>...</th>
<th>gN</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3</td>
<td>2.1</td>
<td>55.0</td>
<td>...</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Classifier

Hormonal Therapy
As adjuvant treatment

ER positive

ER negative

TAMOXIFEN

Trastuzumab

HER-2 positive

HER-2 negative

Taximophen treatment: Yes/No

A Machine Learned Classifier that uses Gene Expression Data to Accurately Predict Estrogen Receptor Status
M. Bastani, L. Vos, N. Asgarian, J. Deschenes, K. Graham, J. Mackey, R. Greiner,
PLoS One, 2013, 8(12)
Other examples in BioMedical applications

> Breast Cancer
  – Predict who would likely develop breast cancer?
  – Predict if breast cancer would reoccur.

> Kidney Transplant prediction
  – How will transplant behave over time?
  – Will a patient reject a transplant?

> Prostate Cancer
  - What toxicity level will the patient reach with radiotherapy?

> Cancer Cachexia
  – Which cancer patients will “waste away”??
# Nuclear Magnetic Resonance Spectroscopy to Classify Patients

**Collect patient urine**

**Obtain NMR spectrum**

**Compute Metabolic Profile**

<table>
<thead>
<tr>
<th>Gluconate</th>
<th>Hippurate</th>
<th>Histidine</th>
<th>Isoleucine</th>
<th>Isopropanol</th>
<th>Lactate</th>
<th>Lactose</th>
<th>Leucine</th>
</tr>
</thead>
<tbody>
<tr>
<td>414.2</td>
<td>599.3</td>
<td>2.73</td>
<td>10.44</td>
<td>16.01</td>
<td>40.83</td>
<td>90.3</td>
<td>5.6</td>
</tr>
</tbody>
</table>

- **Given:**
  - Metabolic profile of patient (NMR / MassSpec of patient’s urine, blood, …)

- **Predict:**
  - Patient’s disease state (Cachexia, Cancer, Reaction to Rx, …)

- **Machine Learning profiles tools for …**
  - Learn $Profile \rightarrow Dx$ classifier

---

Intelligent Diabetes Management

- Each patient with Type I diabetes must regulate his/her own insulin:
  - Give self a dose of insulin (4x / day), based on
    - Current blood glucose level
    - Anticipated carbohydrate consumption
    - Stress
  - Dose is based on formula
    - specific to patient
    - varies over time
- Diabetes MD can adjust formula based on patient’s “Diabetes Diary”
  ... but only on visits – every 3-6 months? ... none in 3rd world countries...

- Automate this “policy adoption” process
  - Reinforcement Learning!
- w/ Alberta Diabetes Institute (E Ryan, ...)

Improved A1C with smart phone app use in Type 1 diabetes
E. Ryan, J. Holland, E. Strouila, B. Bazelli, S. Babwik, H. Li, P. Senior, R. Greiner.
DSS: Treatment Recommendation

1. Injury
2. Treatment 1
3. Evaluation
4. Treatment 2
5. Evaluation
6. Treatment n
7. Evaluation
8. Return to work

Human experts make errors

Workers’ Compensation Board

1. Injury
2. First Treatment
3. Evaluation
4. Secondary Treatment
5. Evaluation
6. Return to work

Semi-Automatic Tumor Segmentation

An Automatic Brain Tumor Segmentation Tool
I. Diaz, P. Boulanger, R. Greiner, B. Hoehn, L. Rowe, A. Murtha,
IEEE Engineering in Medicine and Biology Society, 2013
AI and Machine Learning

- Machine Learning is already transforming our world faster than most realize.
- AI will impact our lives and jobs (automation, personalization, etc.)
- Industries are ripe for disruption

- Customization can also be done in healthcare
- Treatment tailored to individual patients or cohorts of patients

- Use of Machine Learning is key in many markets

- AI is a tool to enhance our capabilities and creativity

- Amii is a global authority in Machine Learning with unique capabilities
How Machine Learning will transform businesses - Application in Precision Health

The Edmonton Data Management Chapter Association - Annual meeting
November 22, 2018

Dr. Osmar R. Zaïane
University of Alberta Professor
Amii Scientific Director and Fellow

zaiane@cs.ualberta.ca
@ozaiane

Thank you For your attention