Machine learning in dosimetry:
Why we shouldn’t fear AI

Mark Gooding, DPhil
What is the fear?

Terminator 2, 1991

What might we fear?

©Copyright Mirada Medical Ltd. 2017
Not yet, not for about 40 years

Terminator, 1984

Sarah Connor: “Look, I'm not stupid you know? They cannot make things like that yet... “

Kyle Reese: “Not yet, not for about 40 years”
Plausible fears

• Ethics
• Responsibility
• Transparency

But WHY?

http://moralmachine.mit.edu/

“people should stop training radiologists now”
– Geoff Hinton, Toronto, 2016
Legitimate fears?

Paddington, 2014
Teaching points

- The basic principles of machine learning and AI
- The strength and limitations of AI
- How AI is being deployed in radiotherapy
- Approaches to validation of AI

What is AI?

Methods to enable a computer to mimic human behaviour/intelligence through learning and reasoning

- Reasoning / planning
- Vision
- Speech / language
- Movement / robotics

©Copyright Mirada Medical Ltd. 2017
What is AI?

Methods to enable a computer to mimic human behaviour/intelligence through learning and reasoning

- Machine learning
- Expert systems
- Logic
- Statistical approaches

Methods to enable a computer to mimic human behaviour/intelligence based on examples or experience

Machine learning 101

Input Data → Predictions
Getting some data

<table>
<thead>
<tr>
<th>State</th>
<th>No. of Cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana</td>
<td>890,000</td>
</tr>
<tr>
<td>Colorado</td>
<td>2,700,000</td>
</tr>
<tr>
<td>Wyoming</td>
<td>1,310,000</td>
</tr>
<tr>
<td>California</td>
<td>5,200,000</td>
</tr>
<tr>
<td>New York</td>
<td>1,460,000</td>
</tr>
</tbody>
</table>

Nearest neighbour

<table>
<thead>
<tr>
<th>State</th>
<th>No. of Cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>2,700,000</td>
</tr>
<tr>
<td>Texas</td>
<td>2,700,000</td>
</tr>
</tbody>
</table>
Nearest Neighbour (2)

<table>
<thead>
<tr>
<th>State</th>
<th>Population</th>
<th>Land Area (sq. mi)</th>
<th>No. of Cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana</td>
<td>6,666,818</td>
<td>35,826</td>
<td>890,000</td>
</tr>
<tr>
<td>Colorado</td>
<td>5,607,154</td>
<td>103,642</td>
<td>2,700,000</td>
</tr>
<tr>
<td>Wyoming</td>
<td>579,315</td>
<td>97,093</td>
<td>1,310,000</td>
</tr>
<tr>
<td>California</td>
<td>39,536,653</td>
<td>155,779</td>
<td>5,200,000</td>
</tr>
<tr>
<td>New York</td>
<td>19,849,399</td>
<td>47,126</td>
<td>1,460,000</td>
</tr>
<tr>
<td>Texas</td>
<td>28,304,596</td>
<td>261,232</td>
<td></td>
</tr>
</tbody>
</table>

⚠️ Nearest neighbour is very sensitive to the training data

Linear regression

\[ y = 0.0903x + 1007.1 \]
Different input data can lead to different results

Getting some confidence

<table>
<thead>
<tr>
<th>State</th>
<th>Population</th>
<th>Land Area (sq. mi)</th>
<th>No. of Cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana</td>
<td>6,666,818</td>
<td>35,826</td>
<td>890,000</td>
</tr>
<tr>
<td>Colorado</td>
<td>5,607,154</td>
<td>103,642</td>
<td>2,700,000</td>
</tr>
<tr>
<td>Wyoming</td>
<td>579,315</td>
<td>97,093</td>
<td>1,310,000</td>
</tr>
<tr>
<td>California</td>
<td>39,536,653</td>
<td>155,779</td>
<td>5,200,000</td>
</tr>
<tr>
<td>New York</td>
<td>19,849,399</td>
<td>47,126</td>
<td>1,460,000</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>6,859,819</td>
<td>10,554</td>
<td>38,000</td>
</tr>
<tr>
<td>North Carolina</td>
<td>10,273,419</td>
<td>53,819</td>
<td>800,000</td>
</tr>
<tr>
<td>Florida</td>
<td>20,984,400</td>
<td>53,625</td>
<td>1,690,000</td>
</tr>
<tr>
<td>Michigan</td>
<td>9,962,311</td>
<td>56,539</td>
<td>1,150,000</td>
</tr>
<tr>
<td>Texas</td>
<td>28,304,596</td>
<td>261,232</td>
<td>?</td>
</tr>
</tbody>
</table>
Getting some confidence

Average error: 604,000
Average error: 1,155,000
Average error: 1,009,000

Understanding the expected performance

Training
Validation
Testing
Why should we test?

<table>
<thead>
<tr>
<th>Model type</th>
<th>Average error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y=bx+a</td>
<td>570,000</td>
</tr>
<tr>
<td>Y=cx^2+bx+a</td>
<td>489,000</td>
</tr>
<tr>
<td>Y=dx^3+cx^2+bx+a</td>
<td>497,000</td>
</tr>
<tr>
<td>Y=ex^4+dx^3+cx^2+bx+a</td>
<td>486,000</td>
</tr>
</tbody>
</table>

The golden rule!
Machine learning 101

- Training
- Validation
- Testing
- Use
Cows in Texas?

<table>
<thead>
<tr>
<th>State</th>
<th>Population</th>
<th>Land Area (sq. mi)</th>
<th>No. of Cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas</td>
<td>28,304,596</td>
<td>261,232</td>
<td>7,928,168</td>
</tr>
</tbody>
</table>

Machine learning 101

- Machine learning is a representation of data
- We need training data to learn from
  - Split into training and validation
- We need more data for testing
- We must be careful when extrapolating
- Alaska doesn’t have many cows
Machine learning 102

Input Data → Classification → Prediction

<table>
<thead>
<tr>
<th>State</th>
<th>Land Area (sq. mi)</th>
<th>Cowboy state</th>
<th>attle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana</td>
<td>35,826</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>103,642</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Wyoming</td>
<td>97,093</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>155,779</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>47,126</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kansas</td>
<td>10,554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Carolina</td>
<td>53,819</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>53,625</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>56,539</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alabama</td>
<td>50,645</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alaska</td>
<td>570,641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arizona</td>
<td>11,539</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arkansas</td>
<td>52,035</td>
<td>✔️</td>
<td></td>
</tr>
</tbody>
</table>
Machine learning 102

- Nearest neighbour
- Random forests
- Polynomial regression
- Linear regression
- Knowledge-based
- Decision trees
- Neural networks
- Expert systems

Iterative training
The process stays the same

Training → Validation → Testing → Use

Complexity

• More complex models allow greater capacity to learn
• More complex models need more data to train
• More complex models need more data to test
• More complex models are harder to explain
Simplicity

Data ➔ Apply model ➔ Estimate dependant variable(s)

AI is just a model

Teaching points

• The basic principles of machine learning and AI
• The strength and limitations of AI
Strengths and Limitations

- Well-defined tasks
- Repetitive tasks
- Complex input / finding patterns
- Creativity / extrapolation
- Need for data
- Explainability

Teaching points

- The basic principles of machine learning and AI
- The strength and limitations of AI
- How AI is being deployed in radiotherapy applications

Music from Jukedeck – create your own at http://jukedeck.com

©Copyright Mirada Medical Ltd. 2017
AI in RT

- Auto-contouring
- Auto-planning
- Decision support

⚠ Not all products are cleared for sale in USA

Auto-Contouring

- Well-defined task
- Repetitive task
- Complex input / finding patterns
- Creativity / extrapolation
- Need for data
- Explainability
Auto-Contouring

Deep Learning Contouring (DLC)

Model-based

Atlas-based

Auto-Planning

- Well-defined task
- Repetitive task
- Complex input / finding patterns
- Creativity / extrapolation
- Need for data
- Explainability

©Copyright Mirada Medical Ltd. 2017
Auto-Planning

- Dose assessment
- Knowledge-based
- Multi-criteria optimisation
- Inverse planning

Decision support

- Well-defined task
- Repetitive task
- Complex input / finding patterns
- Creativity / extrapolation
- Need for data
- Explainability
Decision support

Big questions

Dose prediction

Data collation

Cohort analysis

Teaching points

• The basic principles of machine learning and AI
• The strength and limitations of AI
• How AI is being deployed in radiotherapy applications
• Approaches to validation of AI
Can it be used in practice?
In the clinic?

Quantitative assessment

Figure 1

H. Bakker et al.
Quantitative evaluation of deep learning contouring of head and neck organs at risk. ESTRO 2018

©Copyright Mirada Medical Ltd. 2017
Quantitative assessment

- Easy to report
- Easy to assess change

- May be hard to define “better”
- Assumes certainty in “truth”
- May be hard to translate to clinical impact

Task-based assessment

- Prospective timing measurements (min)

Figure courtesy of M. Hussien
Task-based assessment

- Easy to interpret for clinical context
- Difficult to generalise to other clinical contexts
- Can be time consuming to carry out
- Hard to perform repeated assessment

Qualitative assessment

M. Gooding et al. Multi-centre evaluation of atlas-based and deep learning contouring using a modified Turing Test. ESTRO 2018

www.autocontouring.com
Qualitative assessment

- Some degree of clinical interpretation possible
- Some degree of longitudinal assessment possible

- Doesn’t feel very scientific
- Inherently subjective

Commissioning

Assessment of AI-based medical devices is no different to other medical devices*

All validation approaches have both benefits and limitations

*Assuming continuous learning is not taking place
Teaching points

- The basic principles of machine learning and AI
- The strength and limitations of AI
- How AI is being deployed in radiotherapy applications
- Approaches to validation of AI

Plausible fears

- Ethics
- AI errors/responsibility
- Transparency
- Job replacement

AI is just a tool
Job replacement?

“Accounting technology has allowed the accountant to move from a desk, covered with papers making calculations that took hours to be completed, to more dynamic ways of performing and, it has allowed the accountant to find new challenges and much more to offer than in the past”

Agnes Ann Pepe, The evolution of Technology for the Accounting Profession, 2011

Jobs evolve!

- Creativity / extrapolation
- Need for data
- Explainability

“Humans are awesome!”

- Rob Buckingham, Oxford, 2018
A concluding quote

“people should start training dosimetrists now”

– Mark Gooding, Austin, Tx, 2018

Mark.gooding@mirada-medical.com

@SciChief

©Copyright Mirada Medical Ltd. 2017