OBJECTIVES

- Define artificial intelligence, machine learning, and deep learning
- Overview of machine learning techniques and tools
- Explore how machine learning is being used in medicine and imaging
ARTIFICIAL INTELLIGENCE (AI)

• Artificial intelligence (AI) is intelligence exhibited by machines

• Ideal "intelligent" machine is a flexible rational agent that perceives its environment and takes actions that maximize its chance of success at some goal

• Term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving"
AI
DEEP BLUE

• Deep Blue won its first game against a world champion on February 10, 1996, when it defeated Garry Kasparov in game one of a six-game match

• Kasparov won three and drew two of the following five games, defeating Deep Blue by a score of 4–2

• Deep Blue was then heavily upgraded, and played Kasparov again in May 1997
DEEP BLUE

• Deep Blue won game six, therefore winning the six-game rematch 3½–2½ and becoming the first computer system to defeat a reigning world champion in a match under standard chess tournament time controls

• Kasparov accused IBM of cheating and demanded a rematch

• IBM refused and retired Deep Blue
MACHINE LEARNING (ML)

• Subfield of computer science that gives computers the ability to learn without being explicitly programmed

• Simply, ML is the science of teaching computers how to learn, in an effort to glean information from data that more conventional statistical approaches may not be able to achieve
ML

• Arises at the intersection of statistics, which seeks to learn relationships from data, and computer science, with its emphasis on efficient computing algorithms

• Marriage between mathematics and computer science is driven by the unique computational challenges of building statistical models from massive data sets, which can include billions or trillions of data points
ML

• Evolved from the study of pattern recognition and computational learning theory in artificial intelligence

• Explores the study and construction of algorithms that can learn from and make predictions on data

• Goal of ML algorithm is to develop a mathematical model that fits the data
Based on your internet history, you might be dumb enough to enjoy extreme sports.

Click here to buy a ticket to base jump from the international space station.

I think the internet is trying to kill me.

We call it "machine learning."
Examples of Machine Learning

Self-driving cars

Recommendation systems

and many, many more ...

Photo search
DEEP LEARNING (DL)

Deep learning (also known as deep structured learning, hierarchical learning or deep machine learning) is a branch of machine learning based on a set of algorithms that attempt to model high level abstractions in data.
Deep Learning learns layers of features
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.
ML IN RADIOLOGY

• Growing availability of imaging data is shifting radiology toward quantitative diagnosis

• Advances in machine learning combined with lower cost and more powerful GPU and data availability are creating new doors for research and development
NVIDIA's insane DGX-1 is a computer tailor-made for deep learning

Do you have an extra $129,000 lying around?

Get faster training, larger models, and more accurate results on deep learning with the NVIDIA DGX-1™. This is the world’s first purpose-built system for deep learning and AI accelerated analytics, delivering performance equal to 250 conventional servers. It comes fully integrated with hardware, deep learning software, development tools, and runs popular accelerated analytics applications. This means you can immediately shorten data processing time, visualize more data, accelerate deep learning frameworks, and design more sophisticated neural networks.

NVIDIA DGX-1 is available through select NPN Accelerated Computing partners.

Learn how AI researchers are advancing deep learning and analytics with DGX-1.
Implementing Machine Learning in Radiology Practice and Research

OBJECTIVE. The purposes of this article are to describe concepts that radiologists should understand to evaluate machine learning projects, including common algorithms, supervised as opposed to unsupervised techniques, statistical pitfalls, and data considerations for training and evaluation, and to briefly describe ethical dilemmas and legal risk.

CONCLUSION. Machine learning includes a broad class of computer programs that improve with experience. The complexity of creating, training, and monitoring machine learning indicates that the success of the algorithms will require radiologist involvement for years to come, leading to engagement rather than replacement.

It is difficult to ignore the growing interest in machine learning (ML). ML algorithms generate public excitement because they include playing games against humans [1], self-driving cars, and identifying the characteristics of a great selfie [2]. In radiology, they may help identify schizophrenia with brain MRI [3] and identify genetic markers in glioblastoma [4]. This article introduces radiologists to ML and describes considerations for initiating and evaluating ML projects.

Machine Learning Overview
ML algorithms are divided into statistical models that are straightforward to build, many topics outside the typical knowledge of either academic or practicing radiologists affect whether their models produce accurate, useful results.

ML algorithms start with a set of available inputs and desired outputs. Common inputs in radiology are image data and report text. Output takes the form of a set of conditions and associated probabilities. As an example, we are training a network called MageNet to identify animals in images. If we feed the network a picture of a domestic dog (input), it returns the following list (output): domestic dog, 92%; wolf, 7%; fox, 0.2%; horse, 0.01% (Fig. 1).
ML LEARNING ALGORITHMS
MEDICAL IMAGE DATA FOR ML

• 3 sets of image data sets (medical knowledge and pixel/image data)
  • Training
  • Test
  • Validation

• Training data sent through model algorithm to establish values for each hyperparameter

• Test data sent through model and accuracy of predictions or classifications is evaluated

• Validation data for final model
STANDARDIZING DATA SET

• Medial image data lack standards for ground truth labeling

• Curation of large data sets can be expensive and time consuming due to labeling and preprocessing

• New techniques can help by incorporating the image preprocessing step into the algorithm itself

• Require tightly focused and well defined data or extremely large data sets
PREPROCESSING

- Features selection and extraction
  (to select parameters for the machine learning process. When the number of parameters is significantly higher than the number of cases in the dataset, a data dimensionality reduction is needed)

- Missing data
  (to impute missing data or remove cases with missing information)

- Data normalization
  (to put data with the same order of magnitude, e.g., age in years and height in cm; after being normalized both will have the same magnitude)

- Noise reduction
  (to remove noise that may be present in some modalities, e.g., noise related to the acquisition process)

- Multimodal data
  (different sources of data, e.g., demographic, medical, imaging, genetics)
DATA AUGMENTATION

• Large, well-curated data sets are hard to come by in radiology

• Morphing techniques that slightly modify the images in a dataset can significantly increased the size of the dataset
  • Skewing
  • Contrast
  • Resolution
  • Flipping/rotating
  • Zoom
  • Changing location of finding
CAD VS ML

• CAD algorithms designed to identify presence or absence of image features known to be associated with disease state (e.g. breast microcalcifications on mammogram)

• ML techniques focus on particular labeled outcome (e.g. ductal adenocarcinoma) and in process of training, node clusters evolve into algorithms for identifying features

• ML has potential to identify useful features not currently known to be associated with disease
TRANSFER LEARNING: IMAGENET-NET.ORG

- ImageNet collaboration maintains a large dataset (currently 14 million images) that are labeled with nouns related to the content of each image.
- ImageNet sponsors annual events at which computer science groups from around the world submit trained algorithms in an attempt to classify images from a subset of the ImageNet data for higher and higher levels of accuracy.
IMAGENET LSVRC2014 Object Detection Dataset

Selection Criteria:
- Validation images fully annotated with all 200 object categories, used in ILSVRC2013 and in ILSVRC2014
- Images with traffic light, train

- train
- traffic light

[Images with object classes and object instances]

Classes:
- truck
- taxi
- tiger
- toaster
- traffic light
- train
- trombone
- trumpet
- turtle
- tv or monitor
- tricycle
- vacuum
- violin
- volleyball
- waffle iron
- washer
- water bottle
- watercraft

Show Images
Since 2010, the annual ImageNet classification challenge has been used to determine the state-of-the-art in computerized image classification. The ImageNet training data set consists of more than 1,000,000 photographs in 1000 object categories.

In 2012, Krizhevsky et al. from the Univ. of Toronto achieved a performance breakthrough (markedly decreased error) using a deep convolutional neural network.

Since 2012, all winning entries (and most entries overall) have used convolutional neural networks.
Create an algorithm to distinguish dogs from cats

In this competition, you’ll write an algorithm to classify whether images contain either a dog or a cat. This is easy for humans, dogs, and cats. Your computer will find it a bit more difficult.


The Asirra data set
The Lung Cancer Detection Challenge
Start Your Submission: January 12 – April 12, 2017

Lung cancer is one of the most common types of cancer, with nearly 225,000 new cases of the disease expected in the U.S. in 2016.

Using a data set of high-resolution scans of lungs provided by the National Cancer Institute, participants will develop artificial intelligence algorithms to accurately determine when lesions in the lungs are cancerous. This will dramatically reduce the false positive rate that prevents low-dose CT scans from being widely used for lung cancer detection.

Competition results have the potential to advance our understanding of how all types of cancer develop and spread in the body. They’ll also free radiologists to spend more time with patients.

The Prize
This year, the Data Science Bowl will award a total prize purse of $1 million—provided by the Laura and John Arnold Foundation—to those who observe the right patterns, ask the right questions, and in turn, create unprecedented impact around this high-priority issue.

![Prizes](image)

In addition, $5,000 will be awarded to each of the top three most highly voted Kernels. (Total of $15,000) and $1,000 in prizes to be awarded for sharing your Data Science Bowl journey on social media – more details to be announced on February 1, 2017.

Compete on Kaggle More Competition Details
Using a data set of high-resolution image scans of lungs from hundreds of patients provided by the National Cancer Institute, Data Science Bowl participants will develop artificial intelligence algorithms to accurately determine when lesions in the lungs are cancerous.

- Dramatically reduce the false positive rate that plagues the current technology.
- Potential to advance our understanding of how all types of cancer develop and spread in the body.
- They’ll also free radiologists to spend more time with patients.
CURRENT ML APPLICATIONS FOR IMAGING

• Automated interpretation of abdominal CT scans (segmentation and analysis)
• Normal vs fibrotic lung disease
• Phenotypes of tumors (e.g. GBM)
• Differentiating benign and malignant lung nodules
• Automatic determination of skeletal bone age
• Determine image quality and positioning on mammography
**Question:** Can a deep learning model perform as well as a human in determining bone age from a hand radiograph?

**Answer:** Yes

**Collaborators:** David B. Larson, MD, MBA, Matthew C. Chen, Matthew P. Lungren, MD, MPH, Safwan S. Halabi, MD, Nicholas V. Stence, MD, Curtis P. Langlotz, MD, PhD

---

**Table 3.** Summary statistics of the cross-validation paired interobserver difference between each reviewer’s bone age estimate and mean of the other three human reviewers’ estimates, compared to that of the model. MAD = Mean absolute difference.

<table>
<thead>
<tr>
<th></th>
<th>Reviewer 1</th>
<th>Reviewer 2</th>
<th>Reviewer 3</th>
<th>Reviewer 4</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>0.65</td>
<td>0.55</td>
<td>0.53</td>
<td>0.69</td>
<td>0.61</td>
</tr>
<tr>
<td>Model</td>
<td>0.51</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.52</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.14</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.16</td>
<td>-0.08</td>
</tr>
<tr>
<td>p value (paired t-Test)</td>
<td>0.00</td>
<td>0.50</td>
<td>0.99</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>
USING ML TO AUTOMATE THE EVALUATION OF BREAST POSITIONING IN MAMMOGRAPHY

• Training and testing of classifiers using deep learning technology: Using Google TensorFlow, 1,000 mammography images currently available will be used to train deep learning classifiers based on the 13 positioning criteria

• Reviewer training: In order to increase the size of the training set, three human reviewers will be trained in evaluating image quality according to 13 image quality assessment criteria

• Iterative refinement of the classifiers with greater numbers of images: The larger training sets will be iteratively applied to refine the model

• Performance of the model will be evaluated at different dataset sizes to determine the incremental benefit of additional images
USING ML TO AUTOMATE THE EVALUATION OF BREAST POSITIONING IN MAMMOGRAPHY

Table 1. Breast position quality criteria. Adapted from (4).

<table>
<thead>
<tr>
<th>No.</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Good visualization of posterior breast tissue on both CC and MLO views</td>
</tr>
<tr>
<td>2</td>
<td>Not too much sagging of the breast on the MLO view</td>
</tr>
<tr>
<td>3</td>
<td>Adequate amount of pectoralis muscle on the MLO view</td>
</tr>
<tr>
<td>4</td>
<td>Inframammary fold visible on the MLO view</td>
</tr>
<tr>
<td>5</td>
<td>Breast not positioned too high on image receptor on the MLO view</td>
</tr>
<tr>
<td>6</td>
<td>Skin folds do not significantly obscure tissue visualization</td>
</tr>
<tr>
<td>7</td>
<td>Posterior nipple line on CC view measures within 1 cm of that of MLO view</td>
</tr>
<tr>
<td>8</td>
<td>No excessive exaggeration of the breast on the CC view</td>
</tr>
<tr>
<td>9</td>
<td>No portion of the breast cut off on the image</td>
</tr>
<tr>
<td>10</td>
<td>No other body parts projected over the breast</td>
</tr>
<tr>
<td>11</td>
<td>Adequate visualization of contralateral breast cleavage on the CC view</td>
</tr>
<tr>
<td>12</td>
<td>Nipple in profile on at least one of the screening views</td>
</tr>
<tr>
<td>13</td>
<td>No motion artifacts</td>
</tr>
</tbody>
</table>
ML NOT JUST FOR IMAGES

Information Extraction from Narrative Radiology Reports

- Stanford Part of Speech Tagger
- Porter stemmer
- Word shape from Stanford CoreNLP toolkit
- ReLex to detect negation
- Radiology ontology class

Report information extraction system

Machine learning/NLP

Annotated radiology reports

Unannotated radiology reports

“A 1cm calcified mass probably is present in the anterior right upper lobe”

Anatomy: “right upper lobe”
Anatomy modifier: “anterior”
Observation: “mass”
Observation modifiers: “calcified”, “1cm”
Uncertainty: “probably is present”

Information linkage, summarization, real-time decision support for radiologists

Machine learning approaches to analysing textual injury surveillance data: A systematic review

Kirsten Vallmuur

Centre for Accident Research and Road Safety – Queensland, School of Psychology and Counselling, Faculty of Health, Queensland University of Technology, Kelvin Grove 4059, Brisbane, Queensland, Australia
HOW TO GET STARTED

• Partner with informatics team and data scientists
• Short cut (aka sell your soul): partner with industry
• Define the problem you are trying to solve
• Data, data, data (with context)
• Supervised and Unsupervised Learning
RADIOLOGIST OBSOLETE?
“...machine learning will displace much of the work of radiologists and anatomical pathologists. These physicians focus largely on interpreting digitized images, which can easily be fed directly to algorithms instead. Massive imaging data sets, combined with recent advances in computer vision, will drive rapid improvements in performance, and machine accuracy will soon exceed that of humans. Indeed, radiology is already part-way there: algorithms can replace a second radiologist reading mammograms and will soon exceed human accuracy”
“The patient-safety movement will increasingly advocate use of algorithms over humans — after all, algorithms need no sleep, and their vigilance is the same at 2 a.m. as at 9 a.m. Algorithms will also monitor and interpret streaming physiological data, replacing aspects of anesthesiology and critical care. The timescale for these disruptions is years, not decades.”
PITFALLS

• Benefits are reduced if one of the predictor variables (fever) or outcome variables (clinical response) requires manual data collection or expert assessment.

• Difficult to convince a reviewer that a hypothesis about an unspecified pattern is legitimate.

• No proven methods to estimate sample-size requirements for models using machine learning.

• Difficulty in interpreting the meaning of a machine learning model.

• Machine learning models are more opaque, with many variables and many potential interactions.
ETHICAL AND LEGAL DILEMMAS

• Who owns the data (patient, health system, department, vendor)?
• Black box / opaque algorithms
• FDA regulation
• Validating proprietary ML algorithms
• Is responsibility for successes and failures on human (radiologist), machine (PC) or algorithm developer (industry)?
EDUCATIONAL RESOURCES

• Andrej Karpathy’s blog
• NVIDIA Deep Learning Institute
• Coursera machine learning course (Stanford University)
• Convolutional Neural Networks for Visual Recognition (CS231n)
• Convolutional Networks (deeplearning4j)
YOUTUBE: DEEP LEARNING SIMPLIFIED SERIES
THANK YOU FOR YOUR ATTENTION!

safwan.halabi@stanford.edu

@radhelper

www.linkedin.com/in/safwanhalabi
REFERENCES


