Applied Machine Learning

Introduction

Siamak Ravanbakhsh, Reihaneh Rabbany

COMP 551 (winter 2020)
Objectives

• understanding the scope of machine learning
  ▪ relation to other areas
• understanding types of machine learning
What is Machine Learning?

ML is the set of "algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions"

an inadequate history of ML

- 1950: Turing test
- 1956: checker player that learned as it played (Arthur Samuel)
  - coined the term Machine Learning
- 1958: first artificial neural networks called Perceptron (Frank Rosenblatt),
  - 1959: ADELINE (Widrow and Hoff)
- 1963: support vector machines (Vapnik & Ya)
- 1969: Minskey and Pappert show the limitations of single-layer neural networks
- 1970-80s rule-based and symbolic AI dominates (two AI winters)
- 1980's Bayesian networks (Judea Pearl)
- 1986 Backpropagation rediscovered (Rumelhart, Hinton & Williams)
- 1991 Kernel trick for SVM
- 2012 AlexNet wins Imagenet by a large margin
- 2012 - now deep learning explosion...
- next? AI winter? AGI?
Placing ML: overlapping fields

- Artificial Intelligence: it's a broader domain (includes search, planning, multiagent systems, robotics, etc.)
Placing ML: overlapping fields

- **Artificial Intelligence**: its a broader domain (includes search, planning, multiagent systems, robotics, etc.)

- **Statistics**: historically precedes ML. ML is more focused on algorithmic, practical and powerful models (e.g., neural networks) and is built around AI
Placing ML: overlapping fields

- **Artificial Intelligence**: it's a broader domain (includes search, planning, multiagent systems, robotics, etc.)

- **Statistics**: historically precedes ML. ML is more focused on algorithmic, practical and powerful models (e.g., neural networks) and is built around AI

- **Vision & Natural Language Processing**: use many ML algorithms and ideas
Placing ML: overlapping fields

- **Artificial Intelligence**: its a broader domain (includes search, planning, multiagent systems, robotics, etc.)
- **Statistics**: historically precedes ML. ML is more focused on algorithmic, practical and powerful models (e.g., neural networks) and is built around AI
- **Vision & Natural Language Processing**: use many ML algorithms and ideas
- **Optimization**: extensively used in ML
Placing ML: overlapping fields

- **Artificial Intelligence**: it's a broader domain (includes search, planning, multiagent systems, robotics, etc.)

- **Statistics**: historically precedes ML. ML is more focused on algorithmic, practical and powerful models (e.g., neural networks) and is built around AI

- **Vision & Natural Language Processing**: use many ML algorithms and ideas

- **Optimization**: extensively used in ML

- **Data mining**: scalability, and performance comes before having strong theoretical foundations, more space for using heuristics, exploratory analysis, and unsupervised algorithms
Placing ML: overlapping fields

- **Artificial Intelligence**: it's a broader domain (includes search, planning, multiagent systems, robotics, etc.)
- **Statistics**: historically precedes ML. ML is more focused on algorithmic, practical and powerful models (e.g., neural networks) and is built around AI.
- **Vision & Natural Language Processing**: use many ML algorithms and ideas.
- **Optimization**: extensively used in ML.
- **Data mining**: scalability, and performance comes before having strong theoretical foundations, more space for using heuristics, exploratory analysis, and unsupervised algorithms.
- **Data science**: an umbrella term for the above mostly used in industry when the output is knowledge/information to be used for decision making.
Placing ML: main venues

top computer science conferences

Figure from Dong et al.
Some terminology

\( x \) | input
---|---
features
predictors
independent variable
covariate
Some terminology

\[ \text{input features} \rightarrow \text{ML algorithm (hypothesis)} \rightarrow \text{output targets} \]

\[ \text{predictors} \]

\[ \text{independent variable} \]

\[ \text{covariate} \]

\[ \text{output targets} \]

\[ \text{labels} \]

\[ \text{predictions} \]

\[ \text{dependent variable} \]

\[ \text{response variable} \]
Some terminology

$x$ | input
--- | ---
fractional value
predictors
independent variable
covariate

$\rightarrow$ ML algorithm (hypothesis)

$y$ | output
--- | ---
targets
labels
predictions
dependent variable
response variable

elementary example

$<\text{tumorsize, texture, perimeter}> = <18.2, 27.6, 117.5> \rightarrow \text{cancer = No}$
Some terminology

(labelled) datasets: consist of many training examples or instances

<tumorsize, texture, perimeter>, <cancer, size change>

<18.2, 27.6, 117.5>, <No, +2>
<17.9, 10.3, 122.8>, <No, -4>
<20.2, 14.3, 111.2>, <Yes, +3>

⋮

<15.5, 15.2, 135.5>, <No, 0>
### Some terminology

**datasets:** consist of many training examples or **instances**

<table>
<thead>
<tr>
<th>tumorsize, texture, perimeter</th>
<th>cancer, size change</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.2, 27.6, 117.5</td>
<td>No, +2</td>
</tr>
<tr>
<td>17.9, 10.3, 122.8</td>
<td>No, -4</td>
</tr>
<tr>
<td>20.2, 14.3, 111.2</td>
<td>Yes, +3</td>
</tr>
<tr>
<td>15.5, 15.2, 135.5</td>
<td>No, 0</td>
</tr>
</tbody>
</table>

$x^{(1)}$, $x^{(2)}$, $x^{(3)}$, ..., $x^{(N)}$:

One instance
"Types" of Machine Learning

Supervised Learning: we have labeled data

- classification
- regression
- structured prediction

most of this course!
Supervised Learning

Regression: continuous output

<table>
<thead>
<tr>
<th>&lt;tumorsize, texture, perimeter&gt;</th>
<th>&lt;size change&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;18.2, 27.6, 117.5&gt;</td>
<td>&lt; +2 &gt;</td>
</tr>
<tr>
<td>&lt;17.9, 10.3, 122.8&gt;</td>
<td>&lt; -4 &gt;</td>
</tr>
<tr>
<td>&lt;20.2, 14.3, 111.2&gt;</td>
<td>&lt; +3 &gt;</td>
</tr>
<tr>
<td>&lt;15.5, 15.2, 135.5&gt;</td>
<td>&lt; 0 &gt;</td>
</tr>
</tbody>
</table>
## Supervised Learning

**Classification:** categorical/discrete output

<table>
<thead>
<tr>
<th>&lt;tumorsize, texture, perimeter&gt;</th>
<th>&lt;cancer&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.2, 27.6, 117.5</td>
<td>No</td>
</tr>
<tr>
<td>17.9, 10.3, 122.8</td>
<td>No</td>
</tr>
<tr>
<td>20.2, 14.3, 111.2</td>
<td>Yes</td>
</tr>
<tr>
<td>15.5, 15.2, 135.5</td>
<td>No</td>
</tr>
</tbody>
</table>

**Regression:** continuous output

<table>
<thead>
<tr>
<th>&lt;tumorsize, texture, perimeter&gt;</th>
<th>&lt;size change&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.2, 27.6, 117.5</td>
<td>+2</td>
</tr>
<tr>
<td>17.9, 10.3, 122.8</td>
<td>-4</td>
</tr>
<tr>
<td>20.2, 14.3, 111.2</td>
<td>+3</td>
</tr>
<tr>
<td>15.5, 15.2, 135.5</td>
<td>0</td>
</tr>
</tbody>
</table>
Supervised Learning: Example

Google Unveils Neural Network with “Superhuman” Ability to Determine the Location of Almost Any Image

Guessing the location of a randomly chosen Street View image is hard, even for well-traveled humans. But Google's latest artificial-intelligence machine manages it with relative ease.

by Emerging Technology from the arXiv   February 24, 2016
Supervised Learning: Example

Machine Translation:
data consists of input-output sentence pairs (x,y)
Supervised Learning: Example

Image captioning

**input:** image

**output:** text

![Image 1](image1.png)

The man at bat readies to swing at the pitch while the umpire looks on.

![Image 2](image2.png)

A large bus sitting next to a very tall building.

image: COCO dataset
Supervised Learning: Example

Object detection

**input:** image

**output:** a set of bounding box coordinates

Image: https://bitmovin.com/object-detection/
"Types" of Machine Learning

Unsupervised Learning: only unlabeled data

- clustering
- dimensionality reduction
- density estimation / generative modeling
- anomaly detection
- discovering latent factors and structures

Helps explore and understand the data
Closers to data mining
We have much more unlabeled data
More open challenges
Unsupervised Learning: Example

**clustering**

similar to classification but labels/classes should be inferred and are not given to the algorithm

\[
\begin{align*}
&<\text{tumorsize, texture, perimeter}> \\
&<18.2, \\ &27.6, \\ &117.5> \\
&<17.9, \\ &10.3, \\ &122.8> \\
&<20.2, \\ &14.3, \\ &111.2> \\
&<15.5, \\ &15.2, \\ &135.5>
\end{align*}
\]
Unsupervised Learning: Example

**Generative modeling** (density estimation):
learn the data distribution $p(x)$

---

*Facebook Removes Accounts With AI-Generated Profile Photos*

Researchers said it appears to be the first use of artificial intelligence to support an inauthentic social media campaign.
"Types" of Machine Learning

**Semisupervised learning:** a few labeled examples

- we can include structured problems such as
  - matrix completion (a few entries are observed)
  - link prediction
Netflix Awards $1 Million Prize and Starts a New Contest

BY STEVE LOHR  SEPTEMBER 21, 2009 10:15 AM

Update | 1:45 p.m. Adding details announced close finish to the contest.

Jason Kempin/Getty Images Netflix prize winners, from left, Jim Chabbert, Martin Piotte, Michael Jahrer, Andreas To
Netflix Awards $1 Million Prize and Starts a New Contest

BY STEVE LOHR  SEPTEMBER 21, 2009 10:15 AM

Update | 1:45 p.m. Adding details announced close finish to the contest.
"Types" of Machine Learning

Reinforcement Learning:
- weak supervision through the reward signal
- sequential decision making
- biologically motivated

also related:
imitation learning: learning from demonstrations
- behavior cloning (is supervised learning!)
- inverse reinforcement learning (learning the reward function)
Reinforcement Learning: Example

Human Level Control Through Deep Reinforcement Learning

Abstract

The theory of reinforcement learning provides a normative account deeply rooted in psychological and neuroscientific perspectives on animal behaviour, of how agents may optimize their control of an environment. To use reinforcement learning successfully in situations approaching real-world complexity, however, agents are confronted with a difficult task: they must derive efficient representations of the environment from high-dimensional sensory inputs, and use these to generalize.
Reinforcement Learning: Example

Google’s AlphaGo Defeats Chinese Go Master in Win for A.I.

By PAUL MOZUR  MAY 23, 2017

Ke Jie, the world’s top Go player, reacting during his match on Tuesday against AlphaGo, artificial intelligence software developed by a Google affiliate. China Stringer Network, via Reuters

Related Coverage

A.I. Is Doing Legal Work. But It Won’t Replace Lawyers, Yet.  MARCH 20, 2017

China’s Intelligent Weaponry Gets Smarter  FEB. 3, 2017

The Future of Work
The Future of Not Working  FEB. 23, 2017

Master of Go Board Game Is Outplayed by Google Computer Program  MARCH 13, 2016
Summary

• **Supervised Learning**: we have labeled data
  - classification
  - regression
  - structured prediction

• **Unsupervised Learning**: only unlabeled data
  - clustering
  - dimensionality reduction
  - density estimation / generative modeling
  - anomaly detection
  - discovering latent factors and structures

• **Semisupervised learning**: a few labeled examples
• **Reinforcement Learning**: reward signal