A classic example of a task that requires machine learning: It is very hard to say what makes a 2
What is Machine Learning?
(by examples)
Classification

from data to discrete classes
Spam filtering

data

Welcome to New Media Installation: Art that Learns

Carlos Guestrin to 10615-announce, Osman, Miche

Hi everyone,

Welcome to New Media Installation: Art that Learns

The class will start tomorrow.

***Make sure you attend the first class, even if you are on the Wait List.***

The classes are held in Doherty Hall C316, and will be Tue, Thu 01:30-4:20 PM.

By now, you should be subscribed to our course mailing list: 10615-announce@cs.cmu.edu.
You can contact the instructors by emailing: 10615-instructors@cs.cmu.edu

Natural _LoseWeight SuperFood Endorsed by Oprah Winfrey, Free Trial 1 bottle, pay only $5.95 for shipping mfw rik

Jaquelyn Halley to nherlein, bcc: thehorney, bcc: ang

*** Natural WeightLOSS Solution ***

Vital Acai is a natural WeightLOSS product that enables people to lose weight and cleansing their bodies faster than most other products on the market.

Here are some of the benefits of Vital Acai that You might not be aware of. These benefits have helped people who have been using Vital Acai daily to achieve goals and reach new heights in their dieting that they never thought they could.

* Rapid WeightLOSS
* Increased metabolism - Burn Fat & calories easily!
* Better Mood and Attitude
* More Self Confidence
* Cleanse and Detoxify Your Body
* Much More Energy
* Better Sex Life
* A Natural Colon Cleanse
Object detection

Example training images for each orientation
Weather prediction
Application: Discovering DNA motifs

\[
\text{...TTGGAACAACCATGCACGGTTGATTCGTGCCTGTGACCCACGGAAGACGCCAGCCACCGGTTGTGATG}
\]
\[
\text{TCATA}\textcolor{red}{\text{GGGAATTTCC}}\text{CAGTCGTGAAATAATGCGCTCAGATGAGTAGTAATAGTAAACGCAGGAGGGAGTATTTTGGCCAAGTGCCCTAAGCTATTTCTGTTAT}
\]
\[
\text{CTTACGGACGAAATCCACTATAAGATTATAGAAGAAAGGCGGCGGCGACGAATCTGAAGTACATTAAAGTGGTATGTGCAACAGG}
\]
\[
\text{GTTCGTTATTCCTATTTCTCTTAAGATTTTGATCTGTGTGTATGACTGGAATTAATAAAACCCCTCCTGCACTGACCTTGGACTGAAT}
\]
\[
\text{AACTGTGATACGACCGAAGACTGAACGCTGCGGTTTCTTTATGACCCACGATACGACCAGCTTTAGACCTGAGTTGAGTTG}
\]
\[
\text{GATAACATCCGGCAGGCAGCCAAATCTTTTTGATTTGTATAGGAGACGGATTTGCTAAAGTGTTAAGACTGTTTCCACTACTAGTAAACGGAAGACGCTGAGTTGGAGTT}
\]
\[
\text{GATTTCCGACGTGTTATGCGGGATGAACAGGCGGTCTCTATCATCGGCTCGGCTGTTAAGCGATTCCGTCGCGGAGATAAC}
\]
\[
\text{TGTTGATTTGGCACTGAGTGGCGAGGTCTTTAACAAGGCGGGGTGTACTAAACCCAAAGACCAGGGCCCGAGCGTCAGTGA...}
\]
Regression
predicting a numeric value
Stock market
Weather prediction revisited

Temperature
72° F
Ranking

comparing items
Web search

- **Learning to rank**
- **Learning to rank for information retrieval**
- **Learning to rank using gradient descent**
- **Learning to rank tutorial**

**Yahoo! Learning to Rank Challenge**
[learningtorangechallenge.yahoo.com](http://learningtorangechallenge.yahoo.com)

**Learning to Rank** Challenge is closed! Close competition, innovative ideas, and fierce determination were some of the highlights of the first ever Yahoo!

**Large Scale Learning to Rank**
[www.eecs.berkeley.edu/~dsculley/papers/large-scale-rank.pdf](http://www.eecs.berkeley.edu/~dsculley/papers/large-scale-rank.pdf)

Pairwise learning to rank methods such as RankSVM give good performance, ... In this paper, we are concerned with learning to rank methods that can learn on ...

**Microsoft Learning to Rank Datasets**

We release two large scale datasets for research on learning to rank: L2R-WEB30k with more than 30000 queries and a random sampling of it L2R-WEB10K ...

**LETOR: A Benchmark Collection for Research on Learning to Rank**
[research.microsoft.com/~letor/](http://research.microsoft.com/~letor/)

This website is designed to facilitate research in LEarning TO Rank (LETOR). Much information about learning to rank can be found in the website, including ...
Given image, find similar images

http://www.tiltomo.com/

Slide by D. Sontag (NYU)
Collaborative Filtering
Recommendation systems
Recommendation systems

Machine learning competition with a $1 million prize

Leaderboard

<table>
<thead>
<tr>
<th>Rank</th>
<th>Team Name</th>
<th>Best Score</th>
<th>Improvement</th>
<th>Last Submit Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Ensemble</td>
<td>0.8553</td>
<td>10.10</td>
<td>2009-07-26 18:38:22</td>
</tr>
<tr>
<td>2</td>
<td>BellKor's Pragmatic Chaos</td>
<td>0.8544</td>
<td>10.09</td>
<td>2009-07-26 18:18:28</td>
</tr>
<tr>
<td>3</td>
<td>Grand Prize Team</td>
<td>0.8571</td>
<td>9.91</td>
<td>2009-07-26 13:07:49</td>
</tr>
<tr>
<td>4</td>
<td>Opera Solutions and Vandelay United</td>
<td>0.8573</td>
<td>9.89</td>
<td>2009-07-25 20:05:52</td>
</tr>
<tr>
<td>5</td>
<td>Vandelay Industries I</td>
<td>0.8579</td>
<td>9.83</td>
<td>2009-07-26 02:49:53</td>
</tr>
<tr>
<td>6</td>
<td>PragmaticTheor</td>
<td>0.8582</td>
<td>9.80</td>
<td>2009-12-15:09:53</td>
</tr>
<tr>
<td>7</td>
<td>BellKor in BigChaos</td>
<td>0.8590</td>
<td>9.71</td>
<td>2009-07-26 12:57:25</td>
</tr>
<tr>
<td>8</td>
<td>Dace</td>
<td>0.8603</td>
<td>9.58</td>
<td>2009-07-24 14:18:43</td>
</tr>
<tr>
<td>9</td>
<td>Opera Solutions</td>
<td>0.8611</td>
<td>9.49</td>
<td>2009-07-26 18:02:08</td>
</tr>
<tr>
<td>10</td>
<td>BellKor</td>
<td>0.8612</td>
<td>9.48</td>
<td>2009-07-26 17:19:11</td>
</tr>
<tr>
<td>11</td>
<td>BigChaos</td>
<td>0.8613</td>
<td>9.47</td>
<td>2009-06-23 23:06:52</td>
</tr>
<tr>
<td>12</td>
<td>Feeds2</td>
<td>0.8613</td>
<td>9.47</td>
<td>2009-07-24 20:06:46</td>
</tr>
</tbody>
</table>

Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos

Progress Prize 2007 - RMSE = 0.8712 - Winning Team: KorBell

Cinematch score on quiz subset - RMSE = 0.9514

Slide by D. Sontag (NYU)
Clustering

discovering structure in data
Clustering Data: Group similar things
Clustering images

Set of Images

Goldberger et al.

[Goldberger et al.]

Slide by D. Sontag (NYU)
Clustering web search results
Embedding
visualizing data
Embedding images

- Images have thousands or millions of pixels.
- Can we give each image a coordinate, such that similar images are near each other?
Embedding words
Embedding words (zoom in)
Plugging my own project: NLPlotlib

http://nlplotlib.cs.mcgill.ca/
Traditional Programming

Data → Computer → Output
Program → Computer

Machine Learning

Data → Computer → Program
Output → Computer
Image Categorization

Training

Training Images

Image Features

Classifier Training

Training Labels

Trained Classifier

Slide by Derek Hoiem
Image Categorization

**Training**

- Training Images
- Image Features
- Classifier Training
- Trained Classifier
- Training Labels

**Testing**

- Test Image
- Image Features
- Trained Classifier
- Outdoor Prediction

Slide by Derek Hoiem
ML in a Nutshell

Every machine learning algorithm has three components:

- Representation
- Evaluation
- Optimization
Representation

- Decision trees
- Sets of rules / Logic programs
- Instances
- Graphical models (Bayes/Markov nets)
- Neural networks
- Support vector machines
- Model ensembles
- Etc.

Slide by Domingos (UW)
Evaluation

- Accuracy
- Precision and recall
- Squared error
- Likelihood
- Posterior probability
- Cost / Utility
- Margin
- Entropy
- K-L divergence
- Etc.
Optimization

• Combinatorial optimization
  – E.g.: Greedy search

• Convex optimization
  – E.g.: Gradient descent

• Constrained optimization
  – E.g.: Linear programming
Gradient Descent

https://hackernoon.com/gradient-descent-aynk-7cbe95a778da
Types of Learning

• **Supervised (inductive) learning**
  – Training data includes desired outputs

• **Unsupervised learning**
  – Training data does not include desired outputs

• **Semi-supervised learning**
  – Training data includes a few desired outputs

• **Reinforcement learning**
  – Rewards from sequence of actions
Learning a classifier

Given some set of features with corresponding labels, learn a function to predict the labels from the features.
Example: curve fitting
Example: curve fitting
Example: curve fitting
Example: curve fitting

Overfitting!
Artificial Neural Network

An attempt to imitate the human brain construction, (assuming this is the way it works).

When do we use it?

When we can’t solve the problems ourselves!!!
Artificial Neural Network

The neural network basic structure:

- Big amount of processors – “neurons”.
- Highly connected.
- Working together.

Slide by J. Xu (TTI)
Artificial Neural Network

What does a neuron do?

- Gets “signals” from its neighbors.
- Each signal has different weight.
- When achieving certain threshold - sends signals.
Artificial Neural Network

General structure of ANN:

• One input layer.

• Some hidden layers.

• One output layer.

• Our ANN have one-direction flow!
Artificial Neural Network

• A neuron may be:

- NOT gate
- gate AND
- OR gate

• Because this is a complete system, a neural network can compute anything.

Slide by J. Xu (TTI)
Artificial Neural Network

**Network training and testing:**

- **Training set** - inputs for which we know the wanted output.
- **Back propagation** - algorithm for changing neurons pulses “power”.
- **Test set** - inputs used for final network performance test.
Deep Learning
Deep Art

From: Google DeepDream
Convolutional Neural Networks

https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/