Overview: Machine Learning

Machine Learning:
What is Machine Learning?

- “Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the task or tasks drawn from the same population more efficiently and more effectively the next time.”
  – Herbert Simon
- Closely related to
  - Statistics (fitting models to data and testing them)
  - Data mining / exploratory data analysis (discovering models)
  - Adaptive control theory
  - And of course AI

Machine Learning:
Abstractions from Observation

Learning is the process of automatically constructing abstractions of the real world from a set of observations and past experiences

\[ h: \]

“horse”

\[ d: \]

[Diagram showing images of horses]
Machine Learning: Learning Concepts and Words

Can you pick out the tufas?

Machine Learning: Recognizing Noisy Input

Machine Learning: Classic Recognition Problem

Training examples of a person

Test images
Machine Learning: Information Theory Perspective

- Data compression and transmission over a noisy channel

200 800 2000 4000

h

d

- Which compression captures the essence of the image?
- Which one is best to recognize the same subject in a different photo?

Machine Learning: Why Learn?

- Special Approach to Programming
  - To optimize a performance using example data or past experience.
- Not always needed
  - There is no need to "learn" to calculate payroll
- But used when
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

Types of Machine Learning

- Supervised Learning
  - Classification (pattern recognition)
  - Regression
- Unsupervised Learning
- Reinforcement Learning
Supervised Learning

Supervised Learning:
Classification

- Example: Credit scoring
  Differentiating between low-risk and high-risk customers from their income and savings

- Input data is two dimensional, output is binary

Discriminant:
IF income > θ AND savings > θ THEN low-risk
ELSE high-risk

Supervised Learning:
Classification

<table>
<thead>
<tr>
<th>Color</th>
<th>Shape</th>
<th>Size</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>Square</td>
<td>Small</td>
<td>Yes</td>
</tr>
<tr>
<td>Red</td>
<td>Ellipse</td>
<td>Small</td>
<td>Yes</td>
</tr>
<tr>
<td>Red</td>
<td>Ellipse</td>
<td>Large</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Color</th>
<th>Shape</th>
<th>Size</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>Crescent</td>
<td>Small</td>
<td>?</td>
</tr>
<tr>
<td>Yellow</td>
<td>Ring</td>
<td>Small</td>
<td>?</td>
</tr>
</tbody>
</table>
Supervised Learning:
Classification - Decision Tree

Blue
true false
YES

Hypothesis

Oval
true false
YES

Big
true false
NO

Supervised Learning:
Classification - Decision Tree

Blue
true false
YES

Hypothesis

Oval
true false
YES

Big
true false
NO

Supervised Learning:
Classification - Decision Tree

Blue
true false
YES

Hypothesis

Oval
true false
YES

Big
true false
NO

Supervised Learning:
Classification - Decision Tree

Blue
true false
YES

Hypothesis

Oval
true false
YES

Big
true false
NO
Supervised Learning:
Hypothesis – Quadratic Separation

Supervised Learning:
Hypothesis – Noisy/Mislabeled Data

Supervised Learning:
Hypothesis – Overfitting
Supervised Learning:
Hypothesis – Underfitting?

Supervised Learning:
Hypothesis – More data

Supervised Learning:
Hypothesis – More complex
Supervised Learning:

Linear Regression

- Example: Price of a used car
  \( x \): car attribute
  \( y \): price

- \( y = g(x | \theta) \)
  model: \( g() \)
  parameters: \( \theta = (w, w_0) \)

Supervised Learning:

Polynomial Regression

- Example: Growth of a species
  \( x \): age
  \( y \): length

- \( y = g(x | \theta) \)
  model: \( g() \)
  parameters: \( \theta = (w_3, w_2, w_1, w_0) \)

Supervised Learning:

Piecewise Linear 2D Regression
Supervised Learning:

Some Regression Applications

- Cost estimation
  - Energy consumption
- Control
  - Angle of steering wheel for robot car
  - Kinematics of a robot arm
- Predicted response
  - Surface materials

Supervised Learning:

Range of Methods

- Methods differ in terms of
  - The form of hypothesis space
  - The way to find best hypothesis given data
- There are many successful approaches
  - Decision trees
  - Support vector machines
  - Neural networks
  - Case-based reasoning
  - ...

Supervised Learning:

General Uses

- Prediction of future cases
  Use the rule to predict the output for future inputs
- Knowledge extraction
  The rule is easy to understand
- Compression
  The rule is simpler than the data it explains
- Outlier detection
  Exceptions that are not covered by the rule (e.g. fraud)
Unsupervised Learning

Overview

- General characteristics
  - Learning “what normally happens”
  - No output available
  - Can be formalized in terms of probability density estimation
- Examples
  - Clustering
  - Dimensionality reduction
  - Abnormality detection
  - Latent variable estimation

Unsupervised Learning: K-means clustering

Desired output

K=3 is the number of clusters, here chosen by hand
Unsupervised Learning:
Dendrogram Creation

Unsupervised Learning:
Image Clustering

Unsupervised Learning:
Active Learning – Asking Questions
Reinforcement Learning

Overview

• Characteristics
  – Learning a Policy: A sequence of outputs
  – No supervised output, but a delayed reward
  – Credit assignment problem:
    • Which action led me to winning the game?

• Examples
  – Elevator scheduling
  – Backammon and Chess
  – Robot control