COMPSCI 389
Introduction to Machine Learning

Days: Tu/Th.  Time: 2:30 – 3:45  Building: Morrill 2  Room: 222

Topic 1.1: Introduction to Machine Learning
Prof. Philip S. Thomas (pthomas@cs.umass.edu)
What is machine learning (ML)?

- Subfield of *artificial intelligence* (AI)
  
  "Al is a *field* concerned with *intelligent behavior in artifacts.*"
  
  – Nilsson 1998

  Like math, physics or theology

- AI is *not* a thing/object.
- The thing/object using AI methods is called an *agent*.
  - Agent: Something that acts, from Latin *agere*, which means “to do.”
  - E.g., a robot or software program
What is machine learning (ML)?

• Subfield of \textit{artificial intelligence} (AI)
  
  \begin{quote}
  \textit{“AI is a field concerned with intelligent behavior in agents.”}\nonumber \\
  \textit{~Nilsson 1998}
  \end{quote}

• What is intelligent behavior?
  • No agreed upon definition.
  • How then do we know when we have created an AI?
  • How then do we know whether a topic belongs in the AI field?
    • Consensus.
    • Not always obvious or intuitive.
    • Not always agreed upon.
Questions

• Is soccer a sport?
• Is chess a sport?
• Is rebooting computers a sport?

• Notice that we determine whether something is a sport or not by consensus.
• How can we use the word “sport” if it’s not well defined?
  • If we think there’s ambiguity, we clarify our statements.
• The term “AI” is like “sport” in this way.
Example: Program A

• **Input:** Easy to read source code
  ```java
  while (value < 100)
  item = 10
  value = value + item
  ```

• **Output:** Fast to run source code
  ```java
  while (value < 100)
  value = value + 10
  ```

• **Question:** Can you produce even more efficient code?
  • **Answer:** `value = 100 + (value \mod 10)`

• **Question:** Does Program A exhibit intelligent behavior?
  • **Answer:** No right/wrong answer. I would say “yes.”

• **Question:** Does the study of programs like this fall within AI?
  • **Answer:** The general consensus is “no.”
    • This is part of compilers ⊂ programming languages ⊂ systems.
Example: Program B

- **Inputs:**
  - A directed graph \((V, E)\)
  - A vertex \(s \in V\) (start)
  - A vertex \(g \in V\) (goal)

- **Output:** Does there exist a path from \(s\) to \(g\)? (Yes/No)

- **Question:** Does Program B exhibit intelligent behavior?
- **Answer:** No right/wrong answer. I would say “no”.

- **Question:** Does the study of programs like this fall within AI?
- **Answer:** Yes! This is a “search” algorithm.
**AI is a field concerned with “intelligent behavior” in agents.**

- Don’t get stuck on the name “artificial intelligence” including “intelligence”.
  - Is “computer science” only about computers?
- Rule of thumb: Be inclusive!
  - Avoid arguments saying, “this doesn’t belong in AI because it’s not about intelligent behavior.”
• ML is a subfield of AI “concerned with the question of how to construct computer programs that automatically improve with experience.” [Tom Mitchell, 1997]
ML is a subfield of AI

- ML is a subfield of AI “concerned with the question of how to construct computer programs that automatically improve with experience.” [Tom Mitchell, 1997]
- Improve = learn
ML is a subfield of AI “concerned with the question of how to construct computer programs that automatically improve with experience.” [Tom Mitchell, 1997]

- Improve = learn
- Experience = data
ML is a subfield of AI

- ML is a subfield of AI “concerned with the question of how to construct computer programs that automatically improve with experience.” [Tom Mitchell, 1997]
- Improve = learn
- Experience = data
- Computer = unnecessary
ML is a subfield of AI concerned with the question of how to construct programs that learn from data.

**Question:** Does Program B fall within ML?

**Answer:** No. It doesn’t learn from data.
Computer Science

- Systems
  - Programming Languages
    - Compilers
  - ...
- Theory
- AI
  - Search
  - ML

Program A (code optimization)
Program B (path finding)
Example: Program C

• **Input:**
  - **Data:** Images of handwritten letters with labels.
  - **Query:** An image of a letter.

• **Output:** Prediction of the label for the query.

![Data](image)

<table>
<thead>
<tr>
<th>Data</th>
<th>Query 1:</th>
<th>Query 2:</th>
<th>Query 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>wolf</td>
<td><img src="image" alt="wolf" /></td>
<td><img src="image" alt="seax" /></td>
<td><img src="image" alt="pacul" /></td>
</tr>
<tr>
<td>seax</td>
<td><img src="image" alt="seax" /></td>
<td><img src="image" alt="seax" /></td>
<td><img src="image" alt="pacul" /></td>
</tr>
<tr>
<td>pacul</td>
<td><img src="image" alt="pacul" /></td>
<td><img src="image" alt="pacul" /></td>
<td><img src="image" alt="pacul" /></td>
</tr>
</tbody>
</table>
Program A (code optimization)

Program B (path finding)

Program C (label prediction)
Summary

• **Artificial Intelligence (AI)**: Field concerned with agents that exhibit intelligent behavior.

• **Machine Learning (ML)**: Subfield of AI concerned with agents that learn from data.

• These distinctions are vague, but they provide some structure for thinking about different types of programs and algorithms we might want to create.
I would like to create a program that takes as input video captured from a camera on a car, and outputs a prediction of whether there are any pedestrians near the car, and if so, where they are.

I have gathered thousands of hours of video recording, and hired people to manually label where pedestrians are at all times.

**Question:** Is the program I create an ML program?

**Question:** What if I have thousands of hours of video, but no labels describing where pedestrians are?
I want to create a program that takes as input the rules of a board game, like chess. When presented with a state of the game, it should then be able to produce as output a prediction of what the best move would be.

**Question**: Is this program an ML program?

I have access to millions of games played by strong players, and I want to use this data to create a program that plays like humans by predicting the move that a player would have made.

**Question**: Is this program an ML program?
Intermission

• Class will resume in 5 minutes.
• Feel free to:
  • Stand up and stretch.
  • Leave the room.
  • Talk to those around you.
• Write a question on a notecard and add it to the stack at the front of the room.
Thank you.