Lecture 1: Summary

In this lecture, we explored the question **"what is intelligence?"** and how it has been studied over time. In lecture, we said that defining intelligence is not straightforward, but one way to think about it is as the ability to act on the world, adapt, and achieve goals. For example, humans and animals show intelligence in different ways, while plants mostly demonstrate automatic or reactive behaviors, like sunflowers turning toward sunlight. A key point that we emphasized in class was that simply having sensors or reflexes doesn't make something intelligent unless it can also plan or take purposeful actions.

From this perspective, we decomposed intelligence into three main components: perception (gathering information from the environment), action (moving or responding based on that information), and cognition (the "glue" that interprets sensory input and decides what action to take). Our course content will focus on learning, which is a component of cognition. In fact, learning entails looking at past/historical data and making inferences about future data/events.

The lecture also reviewed how ideas about artificial intelligence (AI) developed historically. Early work in the 1940s by McCulloch and Pitts modeled neurons as simple computational units, laying the foundation for neural networks. Around the same time, Alan Turing proposed key ideas about computation and the famous "imitation game," now known as the Turing Test. In the mid-20th century, the field of cybernetics also tried to formalize the perception–cognition–action loop.

Later, in the 1960s through 2000s, research went through ups and downs. Rosenblatt's perceptron introduced machine learning models, but skepticism from Minsky and Papert led to the first "AI winter." Neural networks made a comeback with backpropagation in the 1980s, leading to applications like handwritten digit recognition. At the same time, other machine learning methods like SVMs, random forests, and kernel methods rose in popularity, partly because they were easier to train than neural networks.

Finally, the "deep learning revolution" took off in the 2000s with more data, more compute power (especially with hardware like GPUs), and breakthroughs like convolutional neural networks that transformed performance on benchmarks such as ImageNet. This ushered in today's widespread applications of deep learning in vision, language, robotics, science, and more.

At the end of the lecture, the course goals were laid out: not only to learn how to use modern deep learning models/ideas effectively, but also to understand the mathematics behind them so we can push the field further!