Artificial Intelligence

Introduction

Artificial intelligence is a branch of computer science dedicated to simulating human cognitive behavior. In the realm of artificial intelligence, we have machine learning, which allows a computer to learn without human intervention.

The process of developing an artificial intelligence system has four key steps:

1. Data scientists build an artificial intelligence system using techniques. These techniques can include, but are not limited to, machine learning.
2. The artificial intelligence system then trains machine learning models by studying patterns in the data.
3. Data scientists observe the results of these machine learning models and subsequently optimizes them, most commonly by modifying the artificial intelligence system which generates these models.
4. The process of training and optimization--steps 2 and 3--repeats until the model becomes accurate enough for its intended purpose. At this point, the model is capable of going into production.

New Terms

- **Artificial Intelligence**: the ability for a computer to simulate human cognitive behavior.
- **Weak Artificial Intelligence**: a narrow application of intelligence, focused on solving one problem. An example of this is a system which counts people who enter or leave a bus. All artificial intelligence systems today are considered weak AI.
- **Strong Artificial Intelligence**: a system which exhibits a human's ability to generalize and solve a variety of problems, including learning how to solve

new problems without direct human programming. As of today, there are no strong AI systems in the world, and there remains debate in the academic community around whether strong AI is achievable.

- **Super Artificial Intelligence**: a system which surpasses a human's ability to generalize and solve problems. This is like strong AI, but the expectation is that the system is superior in every domain of problem-solving capability.
- **Machine Learning**: the process of combining algorithms and data to allow a computer to learn without human intervention.

Workloads

A **workload** is a type of problem we want to solve. For example, one workload is **computer vision**, and an example of computer vision to review a folder full of images and tag those images which include pictures of either cats or dogs.

There are five common workloads for artificial intelligence.

Computer Vision

Computer vision is the process of interpreting the world visually. We train models on videos or images, typically to recognize some object, state, or interaction. Six common implementations of computer vision are:

- **Image classification**: Define the subject of an image. For example, an image classification system could pick out the most prominent actor or thing in a photograph, such as a person, a fruit bowl, or London's Big Ben tower clock.
- **Object detection**: Determine whether a particular thing is in an image. An example of this would be a camera placed near a chicken coop which detects foxes or other predators and raises an alert based on this.
- **Semantic segmentation**: Classify individual pixels in a photo as belonging to a particular thing. For example, a camera monitoring a busy road would

generate an image, and then a semantic segmentation model could identify the individual vehicles on the road and highlight each vehicle with a separate color.

- **Image analysis**: Image analysis provides us a textual description of who or what is in a photo. For example, suppose we have a photo of a man playing catch with a dog. An image analysis model might tell us that the man is throwing a tennis ball, that the man is wearing a blue cap, and that the dog is a yellow Labrador Retriever.

- **Facial detection and recognition**: Identify whether there is a face in a photo. For example, you might want to count how many people enter a store. Using a video camera monitoring the entrance to the store, a facial detection model can identify the number of people who pass through into the store based on observing faces.

- **Optical character recognition**: Read text from an image. For example, you might open a phone application which reads a menu written in a language. That application uses optical character recognition to read letters and words on the image. It can further connect to a service which translates those letters and words into another language, showing the menu items in your preferred language.

Prediction, Forecasting, and Anomaly Detection

Prediction, forecasting, and anomaly detection are techniques intended to analyze data--typically numeric values--and generate an estimation. In the case of prediction and forecasting, we attempt to draw conclusions from the data. For example, suppose we want to estimate the value of a house. We might use input data such as the number of bedrooms, number of bathrooms, and square footage, as well as information on the school district, whether it has certain amenities, and the year it was built.

Anomaly detection intends to find errors or unusual activity in a system. It tends to be focused on a single measure, such as the temperature of a machine or the number of products listed on a marketplace. This kind of data is usually--but not always--time series in nature, and we emphasize trends and changes rather than specific values. That the temperature is 80 degrees Celsius is not necessarily a problem; the problem is that it should normally be 30 degrees Celsius.

Natural Language Processing

Natural language processing (NLP) is the process of interpreting written or spoken language. We train models based on written documents or audio clips of speech. Three common implementations of natural language processing are:

- **Text analytics**: Analyze documents. There are several methods of analysis. One method is to extract phrases, such as finding every time a character in a novel uses a particular word. We can also enumerate the characters in a novel or pick out the places these characters visit over the course of a story. We can even evaluate the sentiment of text, understanding the moods of characters over the course of the book.
- **Text translation**: Translate text between languages. This allows a person speaking English to present to an audience of non-English speakers. The translation service interprets the speaker's text, translates it to another language, and displays that text on the screen as live captions.
- **Text to speech**: Interpret written text and speak it aloud. This kind of service has been around for decades, with improvements over time coming primarily in conveying meaning and nuance in the text using inflection and tone.

In addition, Microsoft has a cloud service called the **Language Understanding Intelligent Service (LUIS)**, which accepts written or spoken inputs, processes those inputs, and allows developers to perform some action based on those inputs. For example, a person may speak a command to open the curtains in the living room.

LUIS would interpret this command and could call a function to activate a physical device inside the house, opening the living room curtains.

Knowledge Mining

Knowledge mining is the process of extracting knowledge from vast amounts of information. In Azure, **Azure Cognitive Search** is the primary tool. It tags information in documents, allowing for easy, detailed searches of those documents.

Conversational AI

Conversational AI is another common workload. With this workload, we develop software agents intended to communicate directly with humans, typically in conversational format. One common example of this is a chatbot which interacts with you on a website. Another example of a conversational AI system would be a virtual agent, who might help you search on the market for a particular house based on the criteria most important to you.