ACTUA'S **ARTIFICIAL INTELLIGENCE (AI)** EDUCATION HANDBOOK

Developed January 2020

With support from Google.org and
Canadian Internet Registration Authority (CIRA)

actua

Youth · STEM · Innovation
Jeunesse · STIM · Innovation

CONTENTS

About Actua **3**
Message from the CEO **4**
How to use this Handbook **5**
Artificial Intelligence: A Primer for Educators **6**
Why AI? 6
AI: A Brief History 6
The Turing Test 7
Narrow vs. General Intelligence 8
Applications of AI 9
1. Recognition 9
2. Conversational Interfaces 10
3. Predictive Analytics 10
4. Personalization 11
5. Autonomous Vehicles / Systems 11
6. Anomaly Detection & Pattern Recognition 12
7. Goal-Driven Systems 12
Combined Applications 13
AI Technologies 13
Machine Learning 13
Supervised Machine Learning 13
Unsupervised Machine Learning 15
Clustering 15
Anomaly Detection 17
Association 17
Reinforcement Learning 17
Deep Learning 17
Other Areas of AI: Natural Language Processing & Computer Vision 18
Precision, Recall and Error Recovery 19
Actua's AI for Education Framework **21**
Actua's AI for Education Framework: An Overview 22
Bringing AI into the K-12 Classroom **25**
AI's Importance in Education 25
Responsible Use of AI 25
Working with Actua's AI activities for youth 26
Additional Resources **26**
Glossary **27**
Acknowledgements **33**

actua

Youth · STEM · Innovation
Jeunesse · STIM · Innovation

ABOUT ACTUA

Actua is Canada's largest science, technology, engineering and mathematics (STEM) outreach organization representing a growing network of 40 university and college based member programs. Each year 300,000 young Canadians in over 500 communities nationwide are inspired through hands-on educational workshops, camps and community outreach initiatives. Actua focuses on the engagement of underrepresented youth through specialized programs for Indigenous youth, girls and young women, at-risk youth and youth living in Northern and remote communities.

For more information or to find a network member program near you, please visit us online at **www.actua.ca** and on social media: **Twitter**, **Facebook**, **Instagram** and **YouTube**.

TERMS OF USE

This work falls under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. For more information, please see **https://creativecommons.org/licenses/by-nc-sa/4.0/**.

That means that you are welcome to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

Under the following terms:

Attribution — You must give appropriate credit, to Actua and provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

Published January 2020

Ottawa, Ontario, Canada

MESSAGE FROM THE CEO

Artificial intelligence (AI) is everywhere, and is only getting more prevalent as technology transforms the ways in which we work and live. From intelligent assistants to recommended products, smart playlists, and facial recognition, AI transcends all industries, and has a multitude of applications and implications.

While youth interact with AI on a daily basis, AI concepts are not yet embedded into Canada's elementary and high school curricula. In order for youth to be prepared for the jobs of today and tomorrow, they need foundational digital skills and literacy, which now includes understanding and applying AI. AI education not only gets students to think critically about their personal interactions with everyday technology, but also how they can leverage AI to solve global challenges.

In early 2019, Actua set out to address this need in engaging youth in AI. With support from Google.org and the Canadian Internet Registration Authority (CIRA), we created Actua's AI Project, designed to contribute to the development of a strong AI training ecosystem in Canada. Our goal is to have AI widely taught in schools, broadly understood by society, acknowledged as an important part of Canada's economy, and something in which Canadians have deep pride.

Actua's AI Education Handbook is intended to support educators with background information on AI, a curriculum-aligned framework, and ideas for classroom implementation. It is designed to accompany Actua's workshop series for educators which provide hands-on opportunities to explore AI concepts in action, both with technology and in unplugged environments. You can learn more about our programs and resources at **actua.ca/ai**.

Actua and our growing network of **40+ network members** are elevating AI education across the country. By the end of 2020, we will have reached over 1,000 K-12 teachers and 30,000 high school students with AI programming. However, the work does not stop there. Continuing to create meaningful learning opportunities for youth to develop skills and competencies for leveraging emerging technology is a critical part of what we do each day.

By using this handbook, you are taking an active step to help grow a vibrant AI ecosystem in Canada and beyond. We are grateful for your support!

Jennifer Flanagan,
President and CEO, Actua

HOW TO USE THIS HANDBOOK

Welcome to AI education! This handbook is designed to support you as an educator in bringing AI concepts and activities into the classroom. We created the content in this handbook with the following in mind:

- **Cross-grade:** Language that resonates with K-12 teachers.
- **Interdisciplinary:** Applications of AI across subjects, not just computer science.
- **Relevant:** Curriculum connections are made at the conceptual level to identify entry points for youth and teachers.
- **Accessible:** Eliminate technical jargon - no coding experience or computer science background is required to understand the concepts and begin teaching them.
- **Canadian content:** Examples of AI innovation that are relevant to Canadians (although we encourage teachers in any country to use this handbook!).

We encourage pre-service and classroom teachers to read this handbook and become oriented to AI fundamentals, as well as AI connections to complement what is currently taught in classrooms. Take time to try some of the activities and recommended interactive materials found on **actua.ca/ai**. There are so many possibilities where you might find the best entry point for you and your students!

This handbook is also intended to accompany Actua's **AI Teacher Training** in-person workshops. These are face-to-face, interactive professional development opportunities offered throughout the year by Actua and our network members. The workshop series has been created based on **Actua's AI for Education Framework** (described later in this handbook), with workshops for each of the six AI themes. For more information, or to connect with your local program, please contact **teachers@actua.ca.**

WHERE TO START:

Depending on your needs, you might want to jump in to this Handbook at different points - it is not necessary to read it in a linear fashion. Here are a few recommendations.

If you are:

- **Looking to learn AI fundamentals.** The AI Primer (starting on page 7) is intended to provide teachers with a brief overview of the fundamental concepts and content needed to understand AI before bringing it into classroom instruction. The primer takes a two-pronged approach to introducing the fundamentals of AI. First, AI will be presented through an applications lens to better understand the range of use cases for this group of technologies. The second half of this primer will introduce some of the underlying technologies that make up the AI landscape and help break down complex jargon and terminology.
- **Interested in making connections between AI and K-12 curriculum.** Following the AI Primer, the handbook introduces Actua's AI for Education Framework, explaining how we structure our approach to AI education for K-12 classrooms. This framework provides relevant, actionable steps for implementing AI activities with students.
- **Ready to discover classroom activities to teach AI.** Start at "Bringing AI into the K-12 Classroom", followed by Actua's recommendations for additional resources, to jump straight into ideas for hands-on way to make AI come alive for students.

Note: When you see **bold, green, underlined words** in this Handbook, you can explore these terms further in the Glossary section.

ARTIFICIAL INTELLIGENCE: A PRIMER FOR EDUCATORS

WHY AI?

There are countless formal definitions of **artificial intelligence.** At its core, **AI is a branch of computer science that deals with a computer's ability to simulate intelligent behaviour**. AI as a "catch-all" term, in fact, represents a range of different technologies, applications, and algorithms.

The AI industry is growing rapidly, and Canada has been recognized as a leader in AI innovation and research. Canada's place in the worldwide AI ecosystem is partially attributable to the **Pan-Canadian Artificial Intelligence Strategy** (PCAIS), a federal government initiative launched in 2017 that has helped attract top AI talent to the country, in both private and public sectors[1]. As a global market, AI is projected to experience a ten-fold increase, or 40% annual growth, by 2026.[2] As the opportunity for AI innovation rapidly grows, Canada will play a significant role in shaping this disruptive set of technologies.

In 2018, two Canadian researchers won the Turing Award, one of the most prestigious prizes in computer science, for their work in advancing the field of AI.[3] Dr. Geoffrey Hinton and Dr. Yoshua Bengio shared the prize with an American researcher, Dr. Yann LeCun. Dr. Hinton and Dr. Bengio were the first Canadians to claim the prize in over three decades.

AI: A Brief History

The term "artificial intelligence" stems back to its first use in the mid 1950s, where it was coined at a conference at Dartmouth College in Hanover, New Jersey.[4] Initial ambitions for AI were quite high, but over the decades, there have been several periods during which optimizing around AI's potential faded and research was focused elsewhere.

Interestingly, many of the core **algorithms** used today were initially described in the 1950s and 1960s. While these algorithms have evolved and improved over time, other changes have had a more profound impact on the field of AI:

- The ability to collect and store vast quantities of data
- Cloud-based storage and retrieval of data
- Exponential increases in computer processing power
- Faster communications networks for moving this data
- An open research community that enables faster research and building within the AI field

---

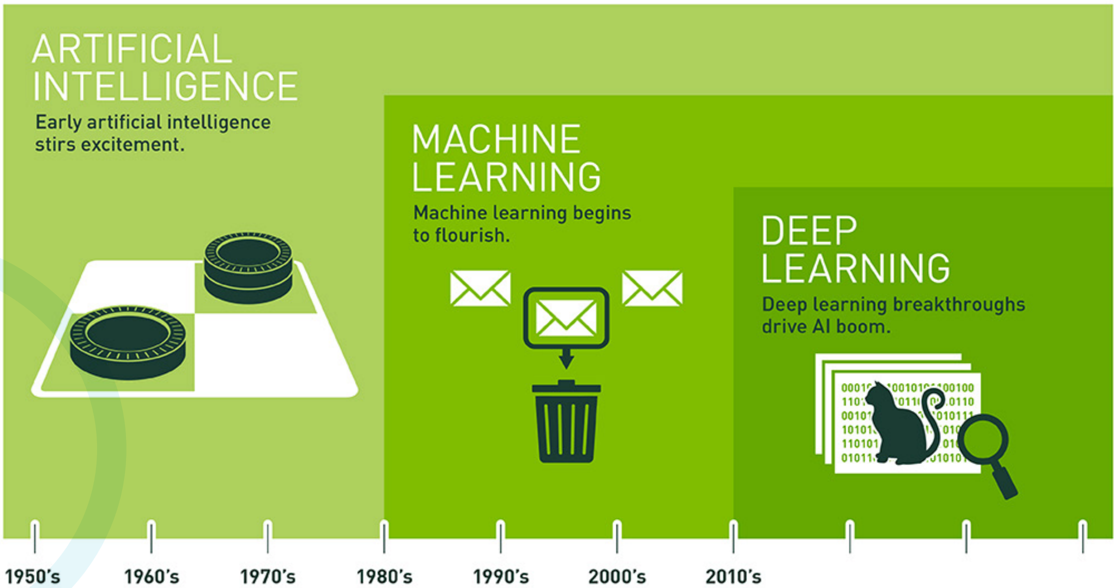
[1] Source: CIFAR (2019, December 9). Canada's top international AI talent grows to 80. Retrieved December 12, 2019 from https://www.newswire.ca/news-releases/canada-s-top-international-ai-talent-grows-to-80-811709991.html.

[2] Source: FinancialNewsMedia.com (2019, November 14). Artificial intelligence (AI) global market projected to exceed $200 billion by 2026. Retrieved December 12, 2019 from https://www.prnewswire.com/news-releases/artificial-intelligence-ai-global-market-projected-to-exceed-200-billion-by-2026-300958067.html.

[3] Source: Semeniuk, I. (2019, March 27). Canadian AI leaders win Turing Award for computer science. Retrieved December 12, 2019 from https://www.theglobeandmail.com/canada/article-canadian-ai-leaders-win-turing-award-for-computer-science/.

[4] Source: Anyoha, R. (2017, August 28), The history of artificial intelligence. Retrieved December 10, 2019 from http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/.

These days, the optimism around the potential of AI lies heavily in the field of **deep learning.** Simply described, deep learning is based around computational algorithms inspired by how human brains work. While deep learning goes back to the earliest days of AI, the power of this approach has really only been uncovered over the past decade or so. AI technologies (including machine learning and deep learning) are further discussed beginning on page 15 of this Handbook.



Source: Copeland, M. (2016, July 29). What's the difference between artificial intelligence, machine learning and deep learning? Retrieved December 10, 2019 from **https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/.**

THE TURING TEST

One of the earliest tests for artificial intelligence is known as the Turing Test, developed in 1950 by Alan Turing[5]. The Turing Test occurs between a human evaluator, a computer, and a human participant. In this test, the human participant interacts (via text interface) with either the human evaluator or the computer. If the participant can't ascertain whether they are interacting with the computer or another machine, the machine is considered to have passed the test.[6]

The Turing Test is an effective demonstration of human-machine interface and interaction, and shows how critical thinking and questioning can be used for deduction. However, since the development of this test, many have argued that this type of test only demonstrates one aspect of simulating human intelligence (i.e., conversation). Further development of other similar tests with increased relevance for modern AI is currently an active area of research.

[5] Source: Stanford Encyclopedia of Philosophy (2016, February 8). The Turing Test. Retrieved December 10, 2019 from https://plato.stanford.edu/entries/turing-test/.

[6] Photo source: Alan Turing statue found in Bletchley Park, UK. "07-turing" by Draig, licensed under CC BY-NC 2.0.

Narrow vs. General Intelligence

One way that the field of AI can be branched is by describing just how intelligent a system is. The following table articulates the differences between ANI (**artificial narrow intelligence**, often referred to as weak AI), and AGI (**artificial general intelligence**, often referred to as strong AI).

	ANI (Artificial Narrow Intelligence)	**AGI** (Artificial General Intelligence)
Also known as	Weak AI	Strong AI
Example	Object recognition, self-driving vehicles, credit card fraud prediction, voice-assistants	No real-world examples currently exist; domain of science fiction
Comparison to human intelligence	Often able to work orders of magnitude faster or more accurately than humans, but only for a well-defined, specific task	On par with human intelligence in all ways

Currently, there is no such thing as AGI. All currently known or developed AI systems would qualify as ANI in the sense that the systems are only able to exert intelligent behaviour within a set number of use cases. For example, a game-playing AI that can beat humans at chess would not be able to predict the future value of the stock market.

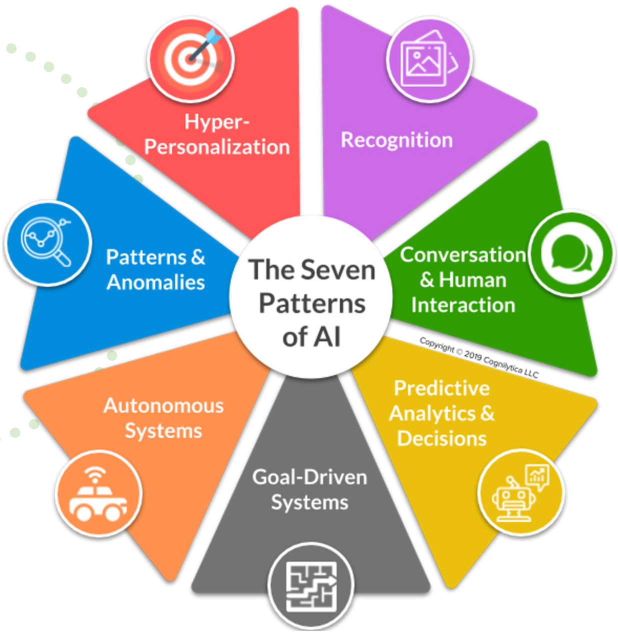
While it may be tempting to consider home-based voice-assistant devices such as Amazon's Alexa or Google Home as examples of AGI, a brief interaction quickly reveals that these systems may seem intelligent in certain contexts, but quickly fail tests in terms of comparison to human cognition. These devices are still in the domain of ANI.

> ***"While machines may exhibit stellar performance on a certain task, performance may degrade dramatically if the task is modified even slightly"***
>
> - ***Yoav Shoham, Stanford Artificial Intelligence Laboratory***

Finally, it's worth noting that one other acronym exists - **ASI - Artificial Super Intelligence**. While currently more science fiction than fact, ASI represents systems that exceed human intelligence.

There are many theories about how ASI might be manifested, but one such hypothesis suggests a Super Intelligent agent would have the ability to recursively create and train more such agents, potentially even controlling humans. But no need to worry; most experts agree that AGI is still decades away, and ASI is still science fiction. It's important to remember that AI algorithms are developed by humans and reflect what they have been trained on. Because human training can include flaws and inherent biases, humans need to be kept accountable, in order to keep AI responsible. This is why it's important to have a variety of inputs from experts in the fields, but also from philosophers, teachers, regulators, artists and civil society with diverse backgrounds, to hold one another accountable to high standards when using AI.

APPLICATIONS OF AI



Ahead of diving into the technical foundations of AI, it's important to understand how AI is being used in the world all around us on a daily basis. One way of exploring this is through applications of AI, examining what dimensions and areas where AI is being leveraged.

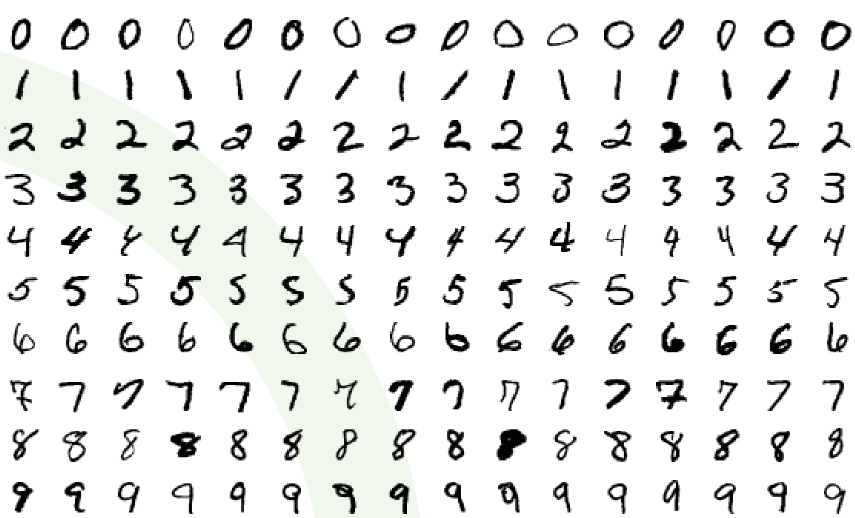
Based around the diagram that follows, seven categories of AI applications or use cases have been identified.

Source: Walch, K. (2019, September 17). The seven patterns of AI. Retrieved December 10, 2019 from **https://www.forbes.com/sites/cognitiveworld/2019/09/17/the-seven-patterns-of-ai/#4005853512d0.**

1. RECOGNITION

This application pertains to a computer or system's ability to recognize one or more *things*. Some of the *things* that a computer may recognize include:

- **Objects**: Images in photos or videos (e.g. is this a cat or a dog?).
- **Voice recognition**: How a computer detects what words are being said.
- **Text or characters**: A.k.a. optical character recognition, how a computer detects what alpha-numeric characters are represented (either printed or hand-written).
- **Facial recognition**: Identifying an individual based on facial or other distinguishing visual characteristics.



As part of the Modified National Institute of Standards and Technology (MNIST) data set, algorithm developers are challenged to correctly identify as many of the hand-drawn digits as possible. Note how many different ways the number 7 can be represented.

Image source: **https://commons.wikimedia.org/wiki/File:MnistExamples.png,** Retrieved December 12, 2019.

2. CONVERSATIONAL INTERFACES

Generally speaking, conversational interfaces take the form of chatbots (text-based), or voice-assistants. In both cases, the underlying technology that makes these experiences possible is called **Natural Language Processing,** or NLP.

NLP is the process of a system understanding the *intent* of a word, phrase, sentence, paragraph or even an entire conversation. Here's just one example that articulates the importance of understanding the intent of a word or phrase, rather than just knowing what the word or phrase is:

Imagine the question: "Do I need an umbrella today?" If you were to ask a human this question, they would quickly understand that you're inquiring about the weather. However, it's worth recognizing that question itself is actually missing critical information.

A system that merely understands what words are represented would not be able to make sense of this question. In fact, it may not even realize that it is in fact, a question.

A well-designed natural language interface would however understand that there is additional contextual information available and (in many cases) be able to give a helpful answer (in part by looking up current or future weather conditions).

Both chatbots and voice assistants have a wide range of business and practical uses ranging from customer service to shopping recommendations to order processing.

3. PREDICTIVE ANALYTICS

The field of predictive analytics deals with taking in past & current data to make predictions about future outcomes.

In the field of inventory management, predictive analytics is being actively used to forecast near and long-term inventory needs. Based on historical data as well as third party data, accurate estimates of future inventory needs can be made that help increase efficiency with respect to holding inventory in warehouses or distribution centres.[7]

In the financial services industry, predictive analytics is being used to estimate a customer's credit worthiness. By analyzing a range of historical factors, accurate estimates can be made about whether a particular customer is at risk of defaulting on a loan or credit card debt.

[7] Image source: http://www.publicdomainfiles.com/show_file.php?id=13511499014922

4. PERSONALIZATION

This field deals with creating personalized experiences, both online and offline, for users.

Online experience personalization involves creating tailored experiences for groups of users, or individuals interacting with a digital platform. This can take many forms ranging from personalized product recommendations in eCommerce, to showing blog readers the type of content they're most likely to consume.

In most cases, online experience personalization is based on understanding the history of what a user has viewed, listened to, or shopped for and then predicting what they might be most interested in in the future.

Offline personalization is newer and still evolving. Common applications for personalization exist in the retail industry where creating unique shopping experiences that are tailored to an individual buyer can help differentiate brands. AI is used in everything from facial recognition (to understand who the customer is) to product recommendations (predicting what they might want to buy next).

5. AUTONOMOUS VEHICLES / SYSTEMS

While the field of autonomous vehicles is well known in the case of self-driving cars, it's important to recognize that there are many other examples of autonomous systems.[8]

Autonomous vehicles and systems are used widely in the manufacturing industry. Autonomous robots that increase the speed and accuracy of assembly lines create significant savings and efficiencies for manufacturers. As these robotic systems have evolved, they've eclipsed or even surpassed human capabilities in terms of speed and accuracy.

Robotic systems also have the ability to work in environments where it may be unsafe for humans, creating additional advantages.

With respect to self-driving cars, several companies have made major strides over the last few years. Companies like Tesla, Waymo, and Google are all actively investing in the technology that will power the next generation of self-driving vehicles. Many of the current challenges around this technology actually surround the regulations for self-driving vehicles on public roads, rather than the capabilities of the technology.

[8] Image source: https://www.flickr.com/photos/romanboed/9572198632 [licensed for reuse]

6. ANOMALY DETECTION AND PATTERN RECOGNITION

Most people who have used a credit card have received a fraud alert or warning, but few have questioned the technology behind these alerts. In many cases, credit card fraud alerting is the result of an anomaly detection algorithm which looks for proverbial needles in haystacks of data.

Anomaly detection and pattern recognition algorithms are often associated with the field of **unsupervised learning**. Generally, unsupervised learning deals with discovering structures or patterns in large unlabelled **data sets**.

Another common application for uncovering structure or patterns in large data sets has to do with the marketing industry. The ability to segment populations of users into groups of individuals who share certain characteristics allows marketers to better target campaigns at groups who are most likely to respond positively to the campaign or promotion.

7. GOAL-DRIVEN SYSTEMS

Goal driven systems have a wide range of uses and are often based on a subfield of **machine learning** called **reinforcement learning.** In goal-driven systems, the algorithm seeks the optimal solution to a given problem through a trial and error process.

In the online advertising industry, where effective campaigns are based on optimal real-time bids for digital advertising space, goal-driven systems or algorithms are being used to increase the performance of these campaigns.

Goal-driven systems are also used in the gaming industry as the underlying technology behind AI-based opponents in video games. Famously, DeepMind's AlphaGo AI beat world Go champion Lee Se-dol in a regulation match of Go, based on a goal-driven system approach;[9] Lee So-dol subsequently retired from professional play after declaring AI unbeatable.[10]

What makes this match of Go between human and machine so interesting is the nature of the game Go itself. Go is a complex game and an algorithm can't use brute force methods (essentially trying every combination of moves) to beat a human (in the game of chess, this is typically how algorithms work). Instead, in Go, AlphaGo had to *learn* strategies by looking at data from previous matches in order to beat the human opponent.

---

[9] Source: Si-soo, P. & Han-soo, L. (2016, March 10). AlphaGo victorious once again. Retrieved December 12, 2019 from http://www.koreatimes.co.kr/www/news/tech/2016/03/325_200068.html.

[10] Source: Vincent, J. (2019, November 27). Former Go champion beaten by DeepMind retires after declaring AI invincible. Retrieved December 12, 2019 from https://www.theverge.com/2019/11/27/20985260/ai-go-alphago-lee-se-dol-retired-deepmind-defeat. Image source from Google/Getty Images.

COMBINED APPLICATIONS

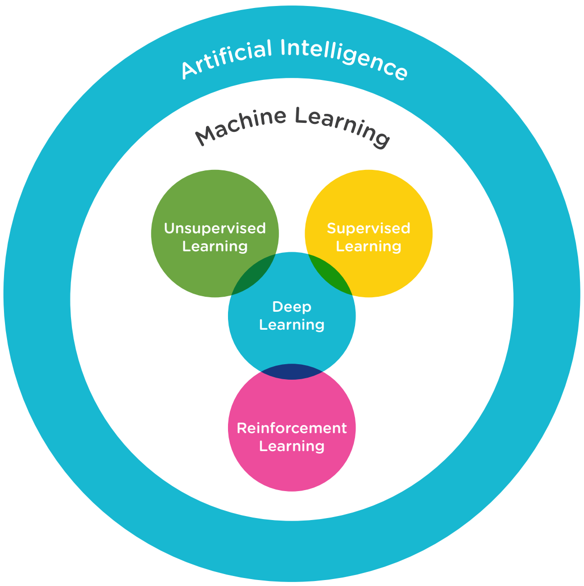
Reviewing the seven key applications or uses cases above, it's worth noting that many real-world systems are actually based on a combination, or two or more of these technologies.

Consider the use of a tool such as Google Lens. This smartphone application allows a user to "look" at just about anything and get contextual, real-time information about what they are looking at (through the camera lens on their smartphone).

If pointed at a menu in another language, for example, Google Lens is able to provide a real-time translation of that menu for the user. This simple example in fact involves recognition (Application 1), natural language processing (Application 2), and often, personalization (Application 4).

AI TECHNOLOGIES

Artificial intelligence includes machine learning (a subset of AI). Within machine learning, there are further sub-disciplines including unsupervised learning, supervised learning, reinforcement learning, and deep learning (deep learning is, in part, represented in all three sub-disciplines of machine learning).



MACHINE LEARNING

Machine learning is probably the most developed subfield of AI, and within the field of machine learning, there are also several subfields. Machine learning can be described as giving computers the ability to uncover patterns or make predictions about data, without being explicitly programmed.

Supervised Machine Learning

What will happen to the stock market tomorrow? What price should we sell this home for? How much should this new employee earn? Is this photo a dog or a cat? These are the types of questions that can be answered (or attempted to be answered) using **supervised machine learning.**

Core to supervised machine learning is having large quantities of *labelled* data. Let's illustrate this through an example.

Home Price Data

Build Year	Square footage	# bedrooms	# bathrooms	Price
1972	1600	3	3	**$335,000**
1985	950	2	2	**$465,000**
1957	2000	5	4	**$650,000**
2001	2200	4	4	**$620,000**
1989	800	2	1	**?**

Before exploring this table further, it's important to understand the key terminology.

- The inputs (build year, square footage, bedrooms, bathrooms), in the language or machine learning, are known as **features**.
- The outputs (price), in the language of machine learning, is known as the **label** or target.

In the example above, there are four rows of labelled data (think of that as historic data) and one row of unlabeled data (think of that as the thing we're trying to predict). The fundamental challenge of machine learning is as follows:

Given labelled data, allow the machine to learn the relationship between features and labels, such that it's able to predict the label for future data.

Within the arena of supervised machine learning, there is yet another division to consider. **Classification** problems deal with labelling something as one type of thing or another - *is this photograph a dog, a cat, or something else?* **Regression** problems deal with predicting a numerical value - *what will the temperature in Ottawa be tomorrow?*

Going back to our house pricing example, we can now see that this is a regression problem. In this example, the challenge is: knowing enough historic information about past home sales, can we predict the value of another home based on that same information.

The importance of feature engineering

If in reading the example above, you're thinking that certain information is missing - you're right! For example, the geographic location or physical condition of a home would be highly correlated to the market value. One of the major challenges in creating an accurate predictive system is that of understanding which features (or inputs) in the data set matter and which ones don't - in the context of what you're trying to predict.

Elaborating on that further, it's entirely possible that the number of bathrooms turns out not to be a predictor of the value of a home. This process of uncovering potential features, and their relationship to the target or label, is a major part of data science or data engineering.

Having reviewed this example, it's now easy to imagine countless other examples of supervised machine learning:

Classification	Regression
Based on a customer's spending patterns, can we predict when they might seek a credit card or line of credit product?	Given historic information about roads where there are a high number of accidents, can we predict risks for new roads being designed?
With enough examples of MRI scans that show both benign and malignant tumours, can we predict whether other scans are indicative of future health risks?	Based on past precipitation data and current conditions, how likely is it to rain in the next 24 hours?

Unsupervised Machine Learning

If you've ever observed that certain online services (video/music streaming, online shopping, etc.) are really good at predicting what you might like, listen to or buy next, you've seen the impacts of **unsupervised learning.**

Unsupervised learning seeks to uncover structure in large data sets. Quite often, this means looking for patterns in that data, or looking for a "needle in a haystack" (something that doesn't look or behave like the rest of the data).

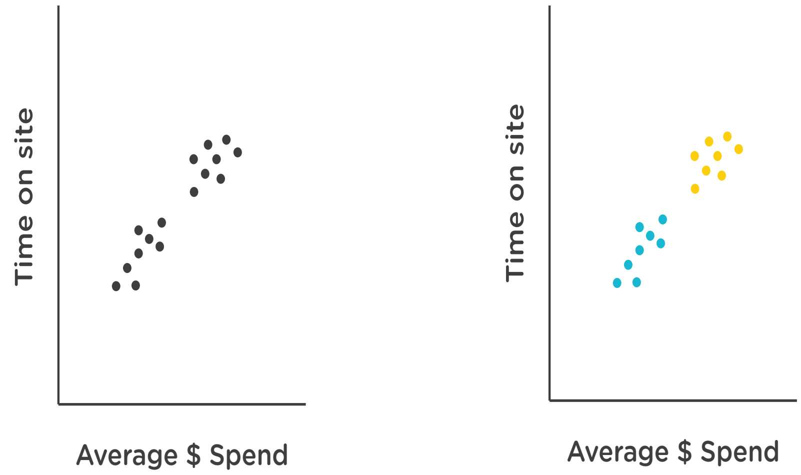
The core difference between the data used for unsupervised learning versus for supervised learning is that the data is unlabelled. If we look at the example of housing data, imagine there was no pricing information in the table. Instead, unsupervised learning (given a lot more data) in this case might help uncover that most houses over 2,000 square feet tend to have four or more bathrooms.

This example grossly oversimplifies the power of unsupervised learning. If you imagine that rather than four features and five rows of data, you had 40 features and 5,000,000 rows of data, suddenly finding patterns in that data less far less obvious. That's the power of unsupervised learning.

Clustering

One of the most common application areas of unsupervised learning is called **clustering.** In clustering, we seek to group data points together.

Looking at the retail industry, one could look at online shopping behaviour and seek ways to group certain types of shoppers together (into clusters) based on how they purchase. This kind of analysis might find that there is a distinct group of shoppers that spend less on average browsing, and tend to buy inexpensive items. This insight could be used to make certain product upgrade recommendations to a subset of buyers, based on how they tend to shop.



The figure above represents online shopping behavioural data, showing how long users spent on the site and how much they spent. Through cluster analysis (a form of unsupervised machine learning) it was uncovered that there is a group of users who spend very little and transact quickly, and a group of users who spend more, but take more time to do so.

While this example may seem obvious, if one were to imagine eight features (instead of the two above), and millions of data points, the ability to cluster the data and uncover behavioural segments could open up new doors for retailers. They would then be better able to target their products and services to the ideal customers - at the right time.

What cluster are you?

A commonly used case for clustering is in quickly categorizing a user into a particular segment, or group of users. In the field of marketing, there's a US database marketing company called Acxiom which has segmented the US population into 70 individual clusters, and subsequently labelled and described those clusters.

Of particular interest is that they have created a ***web-based tool*** *that allows users to fill in just a handful of data points and then learn which cluster or segment they fit into.*

By knowing a bit of information about a user's household income, address, and who resides in the home, an entire description of that individual can be extrapolated. It should go without saying that the description of the cluster is highly subjective, but this example is intended to show some of the practical aspects of unsupervised machine learning.

It should be easy to understand how powerful it would be, for example, for a company doing direct mail campaigns to know which postal codes to target for a certain set of products or services that are being promoted.

Anomaly Detection

Alongside clustering is the ability to find data that doesn't fit the pattern. This is the field of **anomaly detection.** Anomaly detection has a lot of highly practical use cases, for example, in detecting fraudulent transactions in financial services or credit cards.

Association

One extrapolation from the field of clustering is that of **association.** If you imagine an individual who uses an online music streaming service, that service learns from the user's listening habits that they fit within a particular "cluster" of music listeners who tend to like a mix of rock, soul, and rap music.

Well, by looking at songs that others in that "cluster" listen to frequently, it's easy to make highly relevant recommendations for songs that the user may not have yet discovered, but are likely to enjoy.

Reinforcement Learning

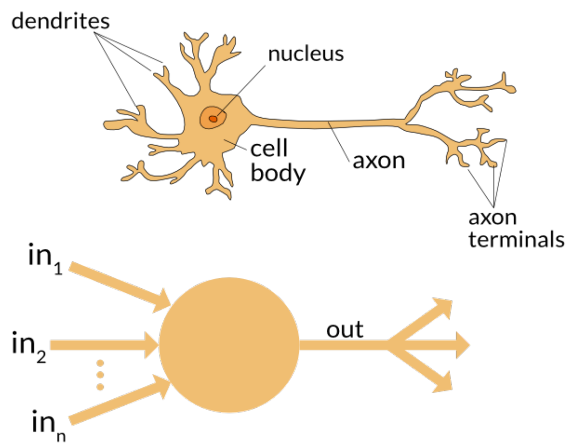
This subset of machine learning is most commonly used in the creation of goal-driven systems. In **reinforcement learning,** the algorithm essentially learns through the process of trial and error.

It's important to understand that reinforcement learning is used primarily when there is a clearly defined end goal or state. If you've ever played a video game against an AI opponent, chances are you were competing with a reinforcement learning algorithm.

DEEP LEARNING

Deep learning is a subset of machine learning that is conceptually based on the way in which our brains work. The precursor to deep learning is called a **neural network**. Neural networks, first developed in the 1950s, were inspired by biological neurons in our brains, though they differ in several ways.

Biological neurons and the structures in our brains can be useful analogies in introducing neural networks to students and connecting the topic back to core curriculum. In the diagram below, the physical structure and function of a biological neuron - with dendrites receiving inputs and the axon terminals responsible for outputs - is analogous to the neural network with layers of inputs transmitting information to lead to an output.



The structure of a neural network is based on biological neurons. Source: Applied Go (2016, June 9). Perceptrons - the most basic form of a neural network. Retrieved December 11, 2019 from **https://appliedgo.net/perceptron/.**

Deep learning builds on the concept of neural networks by stacking consecutive "layers" of these networks to achieve better results.

> *It should be noted that human brains are far more complex than even the most advanced deep learning networks. It's best to consider the relationship between biological intelligence and deep learning as an analogy, rather than an exact representation.*

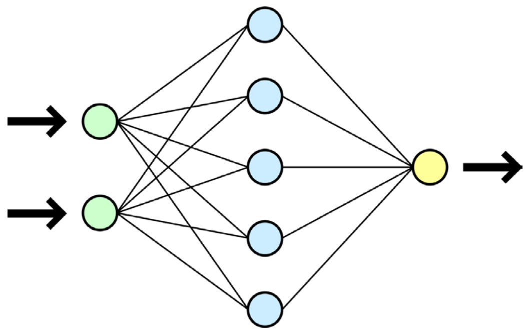


Image source: **https://fr.m.wikipedia.org/wiki/Fichier:Neural_network.svg** [licensed for reuse, modification], Retrieved January 2, 2020.

In this diagram of an oversimplified neural network, one might imagine an input being an image or photograph and the output being a descriptive label, such as a dog or cat. The green circles represent the **input layer,** the blue the **hidden layer** of the network, and the yellow the **output layer.**

While the mechanics of how deep learning works are beyond the scope of this handbook, there are a few important takeaways that should be kept in mind:

- Deep learning is currently considered the field of AI with the most potential for moving us towards strong or general intelligence.
- This type of learning requires large amounts of data and often exponentially more computer processing than statistical machine learning models.
- Deep learning has proven especially powerful where statistical machine learning models have not succeeded, such as object recognition in photos or video.

OTHER AREAS OF AI: NATURAL LANGUAGE PROCESSING AND COMPUTER VISION

There are a few fields of AI that, depending on the context, may or may not be considered as forms of machine learning.

Natural Language Processing

While parts of the field of natural language processing (NLP) use various forms of supervised and unsupervised learning, it's often considered its own discipline. NLP deals with the analysis and understanding of language. NLP is also sometimes referred to as **natural language understanding,** or NLU.

Some of the applications of NLP include:

- Intent extraction: Understanding the meaning of words, phrases, or sentences in context. For example, asking a voice-based assistant "do I need an umbrella?" is translated to an inquiry about rain in the weather forecast.

- Sentiment analysis: Classification about the emotions behind text, words, or sentences. For example, processing an online product review and understanding if the reviewer was ultimately happy or dissatisfied with the product.

It should be noted that classical machine learning or deep learning methods can be, and often are, used in NLP applications.

Computer Vision

Again, while **computer vision** is in some ways part of the overall machine learning landscape, due to the unique problems it solves, it's also often categorized as its own discipline. Computer vision is the field of study that allows computers or other devices to "see" and understand visual information.

> *"The goal of computer vision is to extract useful information from images. This has proved a surprisingly challenging task; it has occupied thousands of intelligent and creative minds over the last four decades, and despite this we are still far from being able to build a general-purpose "seeing machine."*
>
> - Simon Prince, *Computer Vision: Models, Learning, and Inference*

One of the reasons that computer vision is such an active area of research and development is because of its use in self-driving or autonomous vehicles. Self-driving vehicles require the ability to "see", though they will often use information from a range of sensors such as cameras.

PRECISION, RECALL AND ERROR RECOVERY

Working with any kind of predictive algorithm brings a unique set of challenges. Simply put, there are two ways of being wrong. To better understand this, the table below provides an example of an AI spam detection algorithm.

		AI / System Prediction	
		Spam	**Not Spam**
Real World Answer/Result	**Spam**	**True Positive** e.g. AI correctly marks an unwanted email as spam	**False Negative** e.g. AI incorrectly marks an unwanted email as not being spam
	Not Spam	**False Positive** e.g. AI incorrectly marks a real email as spam	**True Negative** e.g. AI correctly acknowledges a real email as not being spam

This table, known as a **confusion matrix,** helps organize information about what the possible scenarios in a predictive algorithm represent. What is important here is to recognize that there are two ways in which a prediction can be wrong. A prediction, in the case of spam detection, can incorrectly mark a real email as spam but also incorrectly mark a spam email as real.

This is an important consideration in terms of tuning an algorithm to give the right performance for a particular situation. In this case, it's worth asking whether it is worse to have a real email get flagged as spam, or a spam email get flagged as real (and thus end up in the inbox).

Precision and **Recall**

Predictive algorithms can be tuned to favour high precision or high recall.

A high precision system, in the case of spam prediction, can ensure that every email that is classified as spam is in fact spam. The trade off here is that this means some spam emails might get missed and end up in the inbox.

A high recall system, on the other hand, will ensure that everything that looks even a little bit like spam ends up in the spam inbox. Of course the tradeoff here is that some real messages will get caught in spam as well.

Depending on the use case, it's easy to imagine situations where either high precision or high recall would give an algorithm a distinct advantage. It's a contextual decision that requires collaboration with all the stakeholders involved in building the AI solution.

actua

Youth · STEM · Innovation
Jeunesse · STIM · Innovation

Actua's Artificial Intelligence (AI) Education Framework

Created with support from Google.org and CIRA, and informed by research from Google's Applied Digital Skills Team, Google Brain, AI4K12.org, CSTA, Microsoft, AI4ALL, and K-12 educators working in AI.

THEME		DATA	PERCEPTION	REPRESENTATION & REASONING	LEARNING	NATURAL INTERACTION	SOCIETAL IMPACTS
Understanding		**Understanding data is foundational to artificial intelligence (AI).**	**Computers sense and perceive the world around them.**	**AI creates models to represent other concepts and uses these models for reasoning.**	**Machine learning happens with data over time.**	**Interaction between AI and humans mimics communication between people.**	**AI can impact society in both positive and negative ways.**
Investigations		What is data, and how do humans use it? What are types of data used in data science? In what ways is data applied in careers and society?	How do machines use sensors to perceive data? How do machine learning tools classify data? What are the limitations of machine perception?	How is data used in AI models? How can models represent other concepts? How do machine models inform decision making?	How do algorithms demonstrate learning? What are neural networks? How does training data influence machine learning?	What does machine-human interaction look like? How do machines understand natural language? What is affective computing; what is consciousness?	What ethical considerations arise when we use and create AI in society? What biases exist in AI algorithms? How can AI be leveraged to face global challenges?
Curriculum Connections		Math: Qualitative and quantitative data, aggregating and analyzing data Science/Interdisciplinary: Data collection, applications of AI (e.g., STEM careers and research) Social: Decision making and reasoning	Math: Data collection and categorizing data inputs Science: Human senses and environmental stimuli, brain processes during sensing and perception Language: Homophones and speech ambiguity	Math: Mapping, graphing, modeling, efficiency Science: Classification systems (e.g., biological) Social: Decision making and reasoning/argument	Math: Sequencing and logic, variables, functions, non-linear graphing Science: Neural processes for learning (e.g., brain structure and function, neural pathways) Social: Understanding bias and critical thinking	Science: Neuroscience (intelligence, consciousness), how living things display intelligence Language: Semantics, communication, language ambiguity Social/Wellness: Non-verbal communication	Math: Detecting bias in data Science/Interdisciplinary: Applications of AI (e.g., STEM careers and research) Social: Decision making, ethics, digital literacy (critical thinking, bias)
PROGRESSION	**Novice/Entry**	Define data; identify data sources or types	Identify sensors; interact with AI agents	Create models; use decision trees	Use a machine learning program; describe learning	Identify verbal and non-verbal communication cues	Identify AI uses and applications in society
	Apprentice	Use data to answer a problem; interpret datasets	Create applications using perception; describe inputs	Design basic decision tree; describe model use	Describe types of machine learning	Compare AI and human performance on tasks	Identify bias potential; describe inclusive AI design
	Practicioner	Describe data analysis; categorical vs numerical data	Describe sensor limitations; use multiple sensors	Design complex decision tree; map efficient paths	Identify bias in data; describe neural network training	Build a chatbot; identify AI	Understand how design impacts function; AI biases
	Expert	Apply data science to solve relevant problems	Use and create complex applications with perception	Describe, use and create search algorithms	Manipulate a neural net/ machine learning algorithm	Identify language ambiguity; debate consciousness	Critically debate social issues and ethics of AI
Applications		Recognition Predictive Analytics Anomaly Detection & Pattern Recognition	Recognition Anomaly Detection & Pattern Recognition	Recognition Predictive Analytics Autonomous vehicles/systems	Conversational Interfaces Predictive Analytics Personalization Anomaly Detection & Pattern Recognition Goal Driven Systems	Conversational Interfaces Personalization Goal Driven Systems	Recognition Conversational Interfaces Predictive Analytics Personalization Autonomous vehicles/systems Anomaly Detection & Pattern Recognition Goal Driven Systems

ACTUA'S AI EDUCATION FRAMEWORK: AN OVERVIEW

Actua's Artificial Intelligence (AI) Education Framework was developed in 2019, using the five big ideas of AI developed by AI4K12[11] as a jumping-off point to build an educational framework that would expand on curriculum connections. Actua developed this framework with additional research and input from Google's Applied Digital Skills Team, Google Brain, CSTA, Microsoft, AI4ALL, as well as K-12 educators working in AI. A key recommendation from this group was to add the data theme (a sixth theme in addition to AI4K12's perception, representation and reasoning, learning, natural interaction, and societal impacts). An understanding of data and its applications is critical to artificial intelligence, and in the context of K-12 education, we felt that educators needed additional resources to introduce data and data science concepts into the classroom.

The framework is intended to be flexible, with curriculum connections speaking to concepts and competencies as opposed to specific programs of study connected to a given grade or course. In this way, we see artificial intelligence as something that can be embedded across subjects and grades in an interdisciplinary manner, not just isolated to computer science instruction.

While Actua is sharing this framework and using it to guide activities and workshops for youth and teachers, it is also part of an iterative design cycle in which we continually revisit the concepts and ideas and update based on emerging needs and developments in AI technology and education. We acknowledge that this framework will likely change in the coming years, just as AI applications, careers, and hopefully as education evolves as well. For questions about the framework or to provide input for future iterations, please email our team at **teachers@actua.ca**.

Themes

There are six main Themes in Actua's AI Education Framework: Data, Perception, Representation and Reasoning, Learning, Natural Interaction, and Societal Impacts. Each Theme is intended to provide a lens through which to view AI education, and is based on AI4K12's work in developing the five big ideas in AI - with the addition of a data Theme. Each of Actua's six Themes has an associated workshop for educators, and all youth activities connect to learning outcomes in one or more of the thematic areas.

Understandings

For each theme, there is an associated Understanding, which is the main takeaway or key learning outcome connected to the theme. Through hands-on exploration and learning, students engaged in thematic activities and/or teachers in Actua's workshops should come away grasping the Understanding associated with that Theme. Understandings are broad, high-level outcomes that change in complexity or level of depth based on student age and competency. For example, the Understanding that corresponds to the *Data Theme* - "data

[11] "Five Big Ideas in AI", retrieved December 5, 2019 from https://bit.ly/ai4k12-five-big-ideas.

comes in many forms and can be used in decision making" - could be approached by an elementary teacher as the difference between qualitative and quantitative data, whereas a high school teacher may choose to delve into the data categorization of variables used in data science.

The *Representation & Reasoning Theme*'s Understanding - "AI creates models to represent other concepts and uses these models for reasoning" - can introduce simple if/then decision trees at the elementary level, which can progress into greater complexity to support students creating algorithms by the high school level. Educators should reference Investigations, Curriculum Connections, and Progressions to inform the appropriate depth at which an Understanding might be addressed for any given student group.

Investigations

Investigations represent guiding questions for inquiry and exploration. These are not exhaustive, but rather represent possible questions for each Theme that provide a path for teachers when creating AI activities. Investigations are intended to provide more specific learning outcomes connected to the Understanding. Depending on student grade and ability, the level of depth or degree of specificity, both in language and complexity explored, will vary. This potential for variation is further reflected in the "Progressions" section.

Curriculum Connections

Canadian curriculum differs by province, and Actua's AI Education Handbook is intended to support teachers across grades and subjects in an interdisciplinary manner. For this reason, the framework is structured in such a way that recommended connections to math, science, language (literacy), social studies, wellness, and interdisciplinary or competency-based outcomes. Depending on the program of studies being followed, the specific grade or course in which concepts are covered may shift; however, these curriculum connections are universally taught and foundational skills and knowledge for students. By providing educators with suggested curriculum connections to core subjects (beyond computer science), teachers are better equipped to make connections to what they are already teaching in the classroom and help students see relevance for AI in multiple disciplines, with multiple applications.

Progressions

Depending on the grade level AI content is introduced to students, it is important to meet them at an appropriate entry point and then work through content that moves them from novice through to expert level (in a K-12 education context). As such, the framework does not provide recommended grades but rather a Progression that begins at novice/entry, then apprentice, practitioner, and expert levels. If AI concepts are introduced in early elementary, the novice/entry level is appropriate, with the higher levels of practitioner and expert being most appropriate at the middle and high school levels. However, for high school students with no previous AI knowledge, it is important to start at the novice/entry level as well before moving into more sophisticated AI concepts. Learning Progressions for each Theme provide scaffolding with general recommendations for the type(s) of activities that would be appropriate for teachers to use with their students in order to engage in the Investigations and work towards the understanding for that particular Theme.

Applications

Applications of AI, as described earlier in this handbook, transcend the AI Themes. To various extents, all seven (recognition, conversational interfaces, predictive analytics, goal-driven systems, autonomous systems, anomaly detection & pattern recognition, and personalization) have relevance in every Theme. For the purpose of the framework, Applications have been placed where they have the greatest relevance so as to guide educators towards seeing where real-world examples and case studies can be woven into content.

BRINGING AI INTO THE K-12 CLASSROOM

AI'S IMPORTANCE IN EDUCATION

While AI is increasingly a top priority in industry and emerging job markets, a corresponding increased presence in K-12 classrooms has not yet been realized. This is in part due to the fact that AI concepts are not yet reflected in most provincial curricula, and there is a skills and knowledge gap for non-technical K-12 teachers looking to bring AI into their instruction.

Both professional development as well as advances in frameworks and/or standards (for AI as well as computer science) would support increased uptake of AI education. Preparing students for relevant post-secondary paths and the future of work necessitates AI literacy, not just immediately before graduation but building a strong foundation that progressively builds as students advance through the education system. It's also important to note that the long term success and adoption of AI will depend on more than just technical roles.

As AI impacts more of our day to day lives, new needs will arise for designers, ethicists, legal expertise and countless other disciplines. This is why AI education has relevance beyond computer science, touching concepts and learnings across grades and subjects.

RESPONSIBLE USE OF AI

Bringing AI into Canadian classrooms isn't without its challenges. There are a few topics that should be considered when thinking about how to introduce this topic to a group of students.

Given the tremendous potential impacts that AI can have on consumers and citizens, it's important to consider how to best harness this group of technologies. The topic of **Ethical AI** or **Responsible AI** is a growing field of research. Private organizations, nonprofits, researchers and governments alike are working to develop frameworks for ensuring that the impacts of AI are good for businesses, individuals and societies as a whole.

Recently, the Canadian government has published a framework and a set of guidelines around the **responsible use of AI in government (https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai.html)**. The guidelines provide a framework through which AI applications can be evaluated and while they were intended for government applications, the core principles apply well to other industries and use cases.

A few key topics that are important to understand with respect to the responsible use of AI:

- **Data bias**: Because AI and machine learning algorithms are trained from large data sets, if there is bias in the data, this bias can be carried forward into the performance of the algorithm. Data bias is currently one of the biggest challenges facing AI.
- **Interpretability**: Because AI algorithms "learn" the relationships between input and output data, there may be a need to understand this relationship and, essentially, what the algorithm is doing. The extent to which an algorithm's action or result can be predicted and understood relates back

to the algorithm's interpretability. In cases where personal information is being used, it may be necessary for algorithm designers to be able to explain how the data is used in decision making.

- **Responsible use of data:** Use of large amounts of data, from a range of sources, carries with it certain ethical and often legal responsibilities. Whenever possible, for example, personal data should be anonymized in order to protect the personal data of individuals. Responsible use of data also relates to providing transparency about when and how data is being used.

USING ACTUA'S AI ACTIVITIES FOR YOUTH

As part of Actua's AI project, a suite of activities for youth is in development (publication will be online at **actua.ca/en/activities**). These activities, created for a high school audience are intended to provide interdisciplinary, hands-on exploration of AI concepts based on Actua's AI Education framework. These activities are appropriate for a variety of environments, including camps, clubs and classroom based workshops delivered by Actua's network members.

Each activity will have a recommended progression for educators who would like to work with students on AI through a multitude of lenses, culminating in a final action project connected to leveraging AI to create social change, connected to the **United Nations' Sustainable Development Goals**. Alternatively, smaller-scale AI explorations can be undertaken by using single activities, each of which is designed to take 1-3 hours, in order to investigate a single AI concept.

To view activities, visit **actua.ca/en/activities** and see those marked with the "Actua AI" label. Note that AI content will continue to be added throughout 2020 and beyond!

ADDITIONAL RESOURCES

If you want to learn more about artificial intelligence, or are curious about other approaches to AI education you can find up to date resources on **actua.ca/ai**. There you will find a curated library of external resources including background information, online interactives, lessons, and other AI content developed by trusted organizations and educators.

GLOSSARY

All glossary terms courtesy of Google's Machine Learning Glossary, licensed under the Creative Commons Attribution 4.0 License, and found at **https://developers.google.com/machine-learning/glossary**, with the exception of those terms marked with an asterisk (*), developed by Actua.

* ALGORITHM

A specific set of steps and or rules to be followed, most commonly in computer science.

* ANOMALY DETECTION

Within the field of unsupervised learning, anomaly detection deals with finding unexpected events or data points within a data set. For example, anomaly detection can be used to predict fraudulent claims in the insurance industry.

ARTIFICIAL GENERAL INTELLIGENCE (AGI)

A non-human mechanism that demonstrates a *broad range* of problem solving, creativity, and adaptability. For example, a program demonstrating artificial general intelligence could translate text, compose symphonies, *and* excel at games that have not yet been invented.

* ARTIFICIAL NARROW INTELLIGENCE (ANI)

A non-human mechanism that demonstrates a *narrow range* of problem solving abilities. For example, a program demonstrating artificial narrow intelligence could predict the next day's weather but be unable to classify an image.

* ARTIFICIAL SUPER INTELLIGENCE (ASI)

A non-human mechanism that demonstrates a broad range of problem solving, creativity, and adaptability skills that exceed human intelligence. ASI is still a hypothetical concept and as such, there are no real-world examples.

ARTIFICIAL INTELLIGENCE

A non-human program or model that can solve sophisticated tasks. For example, a program or model that translates text or a program or model that identifies diseases from radiologic images both exhibit artificial intelligence.

Formally, **machine learning** is a sub-field of artificial intelligence. However, in recent years, some organizations have begun using the terms artificial intelligence and machine learning interchangeably.

* ASSOCIATION

Creating relationships between classes of similar entities, particularly during unsupervised learning. For example, if a group of users with similar musical preferences all like a particular song, it can be deduced that another user who shares these tastes will also like that song.

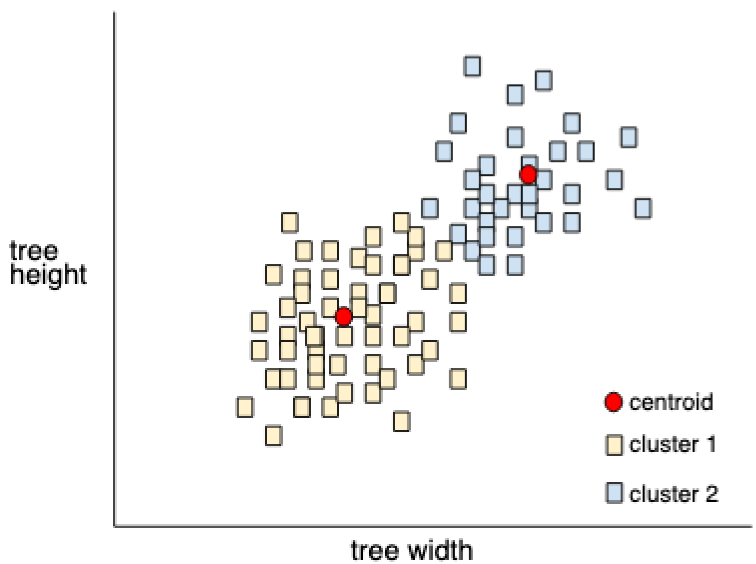
CLASSIFICATION (CLASSIFICATION MODEL)

A type of machine learning model for distinguishing among two or more discrete classes. For example, a natural language processing classification model could determine whether an input sentence was in French, Spanish, or Italian. Compare with **regression model**.

CLUSTERING

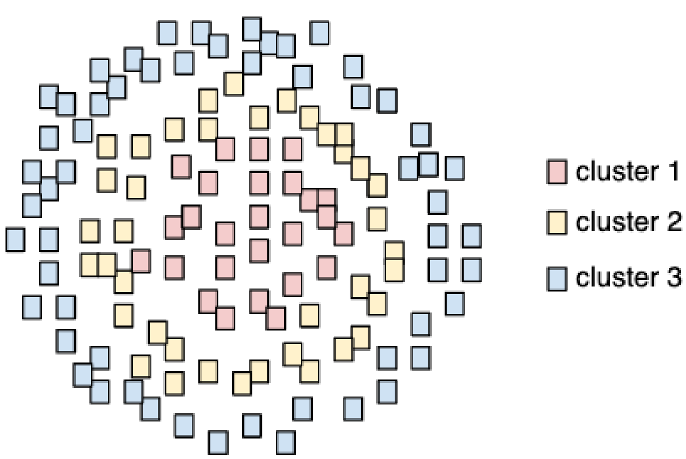
Grouping related **examples**, particularly during **unsupervised learning**. Once all the examples are grouped, a human can optionally supply meaning to each cluster.

Many clustering algorithms exist. For example, the **k-means** algorithm clusters examples based on their proximity to a **centroid**, as in the following diagram:



A human researcher could then review the clusters and, for example, label cluster 1 as “dwarf trees” and cluster 2 as “full-size trees.”

As another example, consider a clustering algorithm based on an example’s distance from a center point, illustrated as follows:



COMPUTER VISION

Computer vision is the field of study that allows computers or other devices to “see” and understand visual information.

CONFUSION MATRIX

An NxN table that summarizes how successful a **classification model’s** predictions were; that is, the correlation between the label and the model’s classification. One axis of a confusion matrix is the **label** that the model predicted, and the other axis is the actual label. N represents the number of **classes**. In a **binary classification** problem, N=2. For example, here is a sample confusion matrix for a binary classification problem:

	Tumor (predicted)	Non-Tumor (predicted)
Tumor (actual)	18	1
Non-Tumor (actual)	6	452

The preceding confusion matrix shows that of the 19 samples that actually had tumors, the model correctly classified 18 as having tumors (18 **true positives**), and incorrectly classified 1 as not having a tumor (1 **false negative**). Similarly, of 458 samples that actually did not have tumors, 452 were correctly classified (452 **true negatives**) and 6 were incorrectly classified (6 **false positives**).

The confusion matrix for a **multi-class classification** problem can help you determine mistake patterns. For example, a confusion matrix could reveal that a model trained to recognize handwritten digits tends to mistakenly predict 9 instead of 4, or 1 instead of 7.

Confusion matrices contain sufficient information to calculate a variety of performance metrics, including **precision** and **recall**.

DATA SET OR DATASET

A collection of **examples**.

DEEP MODEL (DEEP LEARNING)

A type of **neural network** containing multiple **hidden layers**.

Contrast with **wide model**.

FEATURE

An input variable used in making **predictions**.

HIDDEN LAYER

A synthetic layer in a **neural network** between the **input layer** (that is, the features) and the **output layer** (the prediction). Hidden layers typically contain an activation function (such as **ReLU**) for training. A **deep neural network** contains more than one hidden layer.

INPUT LAYER

The first layer (the one that receives the input data) in a **neural network**.

INTERPRETABILITY

The degree to which a model's predictions can be readily explained. Deep models are often non-interpretable; that is, a deep model's different layers can be hard to decipher. By contrast, linear regression models and **wide models** are typically far more interpretable.

LABEL

In supervised learning, the "answer" or "result" portion of an **example**. Each example in a labeled dataset consists of one or more features and a label. For instance, in a housing dataset, the features might include the number of bedrooms, the number of bathrooms, and the age of the house, while the label might be the house's price. In a spam detection dataset, the features might include the subject line, the sender, and the email message itself, while the label would probably be either "spam" or "not spam."

MACHINE LEARNING (ML)

A program or system that builds (trains) a predictive model from input data. The system uses the learned model to make useful predictions from new (never-before-seen) data drawn from the same distribution as the one used to train the model. Machine learning also refers to the field of study concerned with these programs or systems.

NATURAL LANGUAGE UNDERSTANDING (PROCESSING/NLP)

Determining a user's intentions based on what the user typed or said. For example, a search engine uses natural language understanding to determine what the user is searching for based on what the user typed or said.

NEURAL NETWORK

A model that, taking inspiration from the brain, is composed of layers (at least one of which is **hidden**) consisting of simple connected units or **neurons** followed by nonlinearities.

* OUTPUT LAYER

The final layer (the one that presentes the output) in a neural network.

PRECISION

A metric for **classification models**. Precision identifies the frequency with which a model was correct when predicting the **positive class**. That is:

Precision=True Positives / (True Positives+False Positives)

RECALL

A metric for **classification models** that answers the following question: Out of all the possible positive labels, how many did the model correctly identify? That is:

Recall= True Positives / (True Positives+False Negatives)

REGRESSION MODEL

A type of model that outputs continuous (typically, floating-point) values. Compare with **classification models**, which output discrete values, such as “day lily” or “tiger lily.”

REINFORCEMENT LEARNING (RL)

#rl

A family of algorithms that learn an optimal **policy**, whose goal is to maximize **return** when interacting with an **environment**. For example, the ultimate reward of most games is victory. Reinforcement learning systems can become expert at playing complex games by evaluating sequences of previous game moves that ultimately led to wins and sequences that ultimately led to losses.

SENTIMENT ANALYSIS

Using statistical or machine learning algorithms to determine a group’s overall attitude—positive or negative—toward a service, product, organization, or topic. For example, using **natural language understanding**, an algorithm could perform sentiment analysis on the textual feedback from a university course to determine the degree to which students generally liked or disliked the course.

SUPERVISED MACHINE LEARNING

Training a **model** from input data and its corresponding **labels**. Supervised machine learning is analogous to a student learning a subject by studying a set of questions and their corresponding answers. After mastering the mapping between questions and answers, the student can then provide answers to new (never-before-seen) questions on the same topic. Compare with **unsupervised machine learning**.

UNSUPERVISED MACHINE LEARNING

Training a **model** to find patterns in a dataset, typically an unlabeled dataset.

The most common use of unsupervised machine learning is to cluster data into groups of similar examples. For example, an unsupervised machine learning algorithm can cluster songs together based on various properties of the music. The resulting clusters can become an input to other machine learning algorithms (for example, to a music recommendation service). Clustering can be helpful in domains where true labels are hard to obtain. For example, in domains such as anti-abuse and fraud, clusters can help humans better understand the data.

Another example of unsupervised machine learning is **principal component analysis (PCA)**. For example, applying PCA on a dataset containing the contents of millions of shopping carts might reveal that shopping carts containing lemons frequently also contain antacids.

Compare with **supervised machine learning**.

ACKNOWLEDGEMENTS

With funding from

Canada

With gratitude to our advisors, reviewers, and collaborators on this work:

Alexander Chen, Google Teachable Machine
Ali McComb, Shine Creative
Andy Forest, Steamlabs
Barron Webster, Google Teachable Machine
Caitlin Quarrington, Actua
Chris Henry, Google Canada
Doug Braden, InkSmith
Emmanuelle Provost, Actua
Hector Mujica, Google.org
Jordan Griffith, Google Teachable Machine
Katarina Smith, UBC Geering Up Engineering Outreach
Kristina Martin, Actua
Luiza Staniec, Google Canada
Luke Persaud, Toronto District School Board
Matthew Dawson, Google Applied Digital Skills Team
Matthew Grupp, Northwest School Division
Mohamed Abdelazez, Carleton University
Natacha Mainville, Google Brain
Natalie Lao, Google Teachable Machine
Ramy Nassar, 1000 Days Out
Vanessa Ironside, Ontario Tech University Engineering Outreach
Zaitoon Murji, Google Canada

With acknowledgement to the many individuals from groups including AI4K12, AI4ALL, Google.org, Microsoft Canada, Canadian Internet Registration Authority (CIRA), and educators from school districts across North America who provided input that helped to shape this project.

And many thanks to our network members, in particular those who participated in piloting this work in their local communities:

Creative Encounters with Science, University of Guelph
Engineering Outreach, University of Toronto
First Peoples' Centre, Canadore College
Folie Technique, École Polytechnique de Montréal
Geering Up, University of British Columbia
Genitrucs, Université du Québec à Trois-Rivières
Ontario Tech University Engineering Outreach, Ontario Tech University
Queen's Engineering Outreach, Queen's University
Red Hot Science at RDC, Red Deer College
Science Venture, University of Victoria
SuperNOVA, Dalhousie University
WISE Kid-Netic Energy, University of Manitoba

NOTES

NOTES

NOTES

Youth · STEM · Innovation
Jeunesse · STIM · Innovation

© Actua, 2020