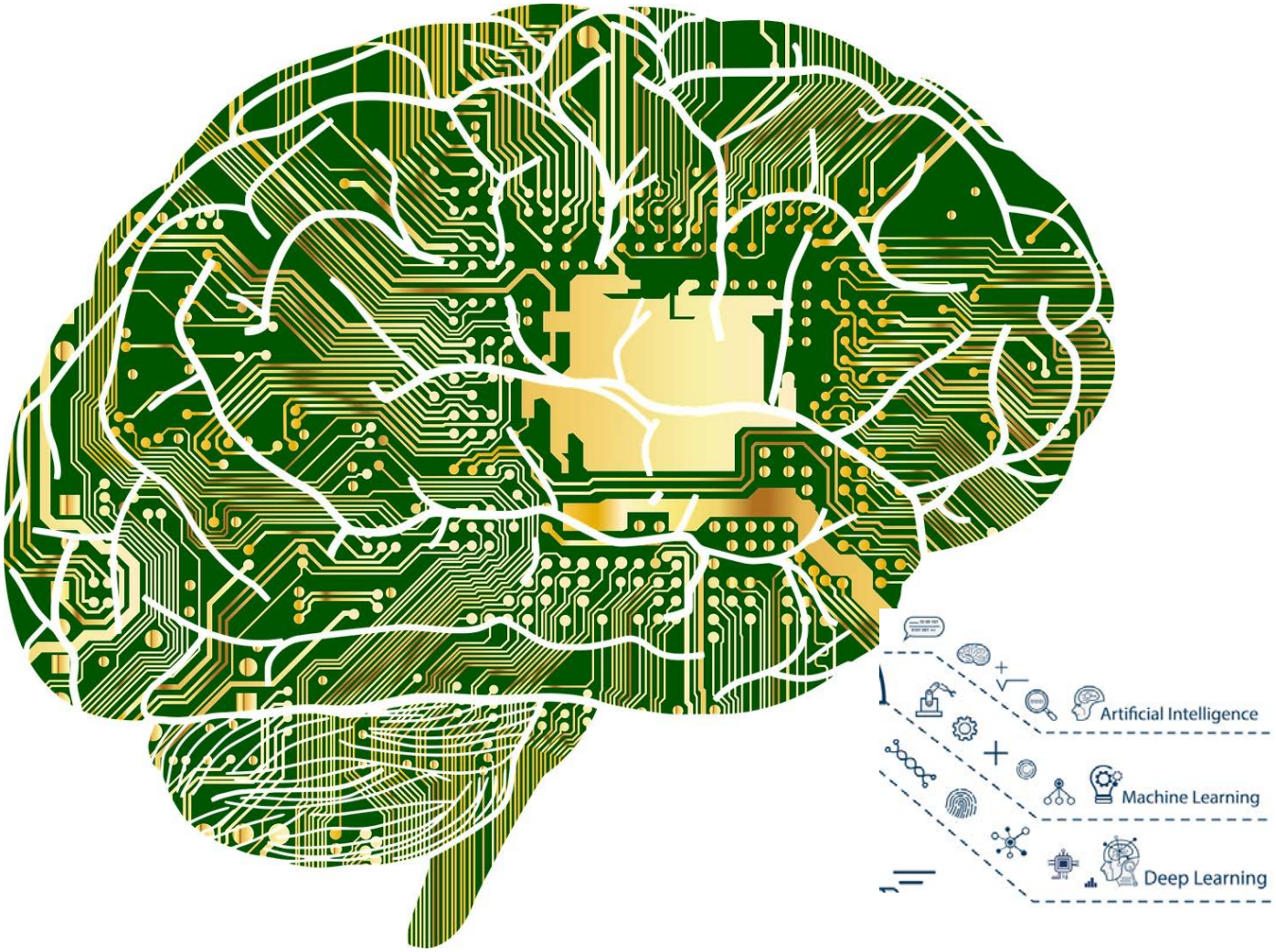


Artificial Intelligence/ Machine Learning Explained

Author: Steve Blank

Stanford | Gordian Knot Center for National Security Innovation

https://gordianknot.stanford.edu

Artificial Intelligence/Machine Learning– Explained

AI is a once-in-a lifetime commercial and defense game changer

Hundreds of billions in public and private capital is being invested in AI and Machine Learning companies. The number of patents filed in 2021 is more than 30 times higher than in 2015 as companies and countries across the world have realized that AI and Machine Learning will be a major disruptor and potentially change the balance of military power.

Until recently, the hype exceeded reality. Today, however, advances in AI in several important areas (here, here, here, here and here) equal and even surpass human capabilities.

If you haven't paid attention, now's the time.

AI and the DoD

The Department of Defense has thought that AI is such a foundational set of technologies that they started a dedicated organization- the JAIC - to enable and implement artificial intelligence across the Department. They provide the infrastructure, tools, and technical expertise for DoD users to successfully build and deploy their AI-accelerated projects.

Some specific defense related AI applications are listed later in this document.

We're in the Middle of a Revolution

Imagine it's 1950, and you're a visitor who traveled back in time from today. Your job is to explain the impact computers will have on business, defense and society to people who are using manual calculators and slide rules. You succeed in convincing one company and a government to adopt computers and learn to code much faster than their competitors /adversaries. And they figure out how they could digitally enable their business – supply chain, customer interactions, etc. Think about the competitive edge they'd have by today in business or as a nation. They'd steamroll everyone.

That's where we are today with Artificial Intelligence and Machine Learning. These technologies will transform businesses and government agencies. Today, 100s of billions of dollars in private capital have been invested in 1,000s of AI startups. The U.S. Department of Defense has created a dedicated organization to ensure its deployment.

But What Is It?

Compared to the classic computing we've had for the last 75 years, AI has led to new types of applications, e.g. facial recognition; new types of algorithms, e.g. machine learning; new types of computer architectures, e.g. neural nets; new hardware, e.g. GPUs; new types of software developers, e.g. data scientists; all under the overarching theme of artificial intelligence. The sum of these feels like buzzword bingo. But they herald a sea change in what computers are capable of doing, how they do it, and what hardware and software is needed to do it.

This brief will attempt to describe all of it.

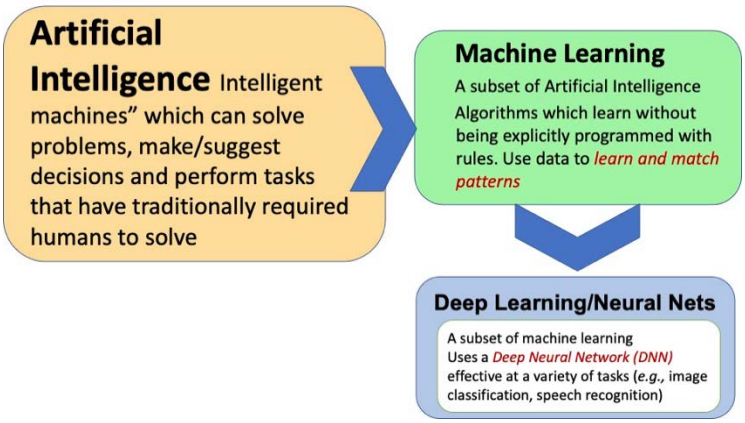
New Words to Define Old Things

One of the reasons the world of AI/ML is confusing is that it's created its own language and vocabulary. It uses new words to define programming steps, job descriptions, development tools, etc. But once you understand how the new world maps onto the classic computing world, it starts to make sense. So first a short list of some key definitions.

AI/ML - a shorthand for Artificial Intelligence/Machine Learning

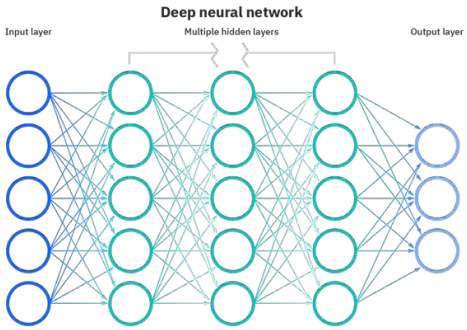
Artificial Intelligence (AI) - a catchall term used to describe "Intelligent machines" which can solve problems, make/suggest decisions and perform tasks that have traditionally required humans to do. AI is not a single thing, but a constellation of different technologies.

Machine Learning (ML) - a subfield of artificial intelligence. Humans combine data with algorithms (see here for a list) to *train a model* using that data. This trained model can then make predications on new data (is this picture a cat, a dog or a person?) or decision-making processes (like understanding text and images) without being explicitly programmed to do so.



> *Machine learning algorithms* - computer programs that adjust themselves to perform better as they are exposed to more data. The "learning" part of machine learning means these programs change how they process data over time. In other words, a machine-learning algorithm can adjust its own settings, given feedback on its previous performance in making predictions about a collection of data (images, text, etc.).

Deep Learning/Neural Nets – a subfield of machine learning. Neural networks make up the backbone of deep learning. (The "deep" in deep learning refers to the depth of layers in a neural network.) Neural nets are effective at a variety of tasks (*e.g.,* image classification, speech recognition). A deep learning neural net algorithm is given massive volumes of data, and a task to perform - such as classification. The resulting model is capable of solving complex tasks such as recognizing objects within an image and translating speech in real time. In reality, the neural net is a logical concept that gets mapped onto a physical set of specialized processors. See here.)



Data Science – a new field of computer science. Broadly it encompasses data systems and processes aimed at maintaining data sets and deriving meaning out of them. In the context of AI, it's the practice of people who are doing machine learning.

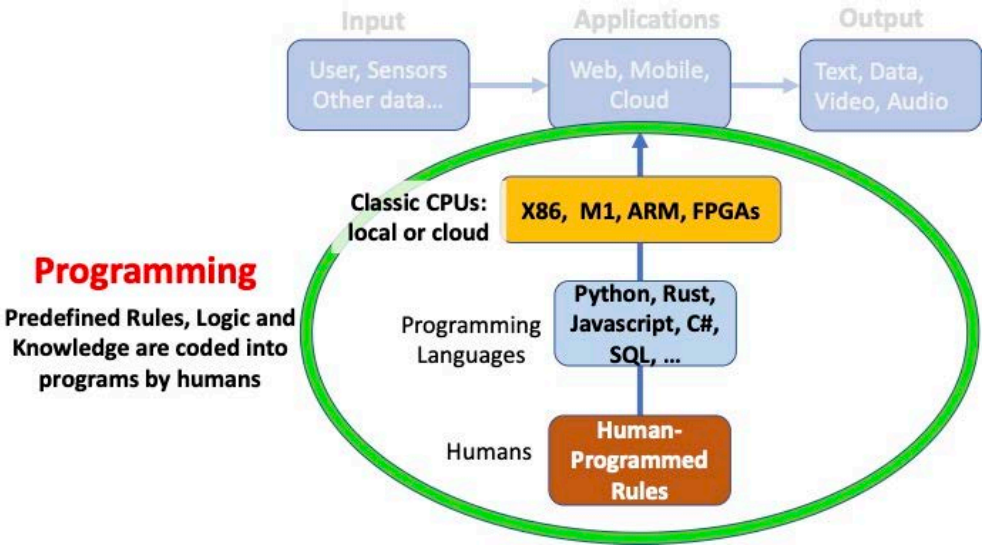
> *Data Scientists* - responsible for extracting insights that help businesses make decisions. They explore and analyze data using machine learning platforms to create models about customers, processes, risks, or whatever they're trying to predict.

What's Different? Why is Machine Learning Possible Now?

To understand why AI/Machine Learning can do these things, let's compare them to computers before AI came on the scene. (Warning – *simplified* examples below.)

Classic Computers

For the last 75 years computers (we'll call these *classic computers)* have both shrunk to pocket size (iPhones) and grown to the size of warehouses (cloud data centers), yet they all continued to operate essentially the same way.



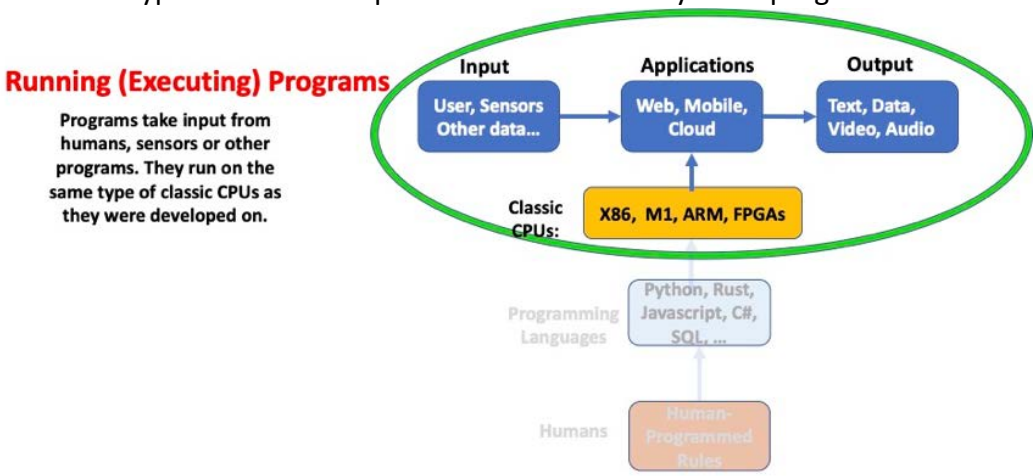
Classic Computers - Programming

Classic computers are designed to do anything a human *explicitly tells them to do*. People (programmers) write software code (programming) to develop applications, thinking a priori about all the rules, logic and knowledge that need to be built in to an application so that it can deliver a specific result. These rules are explicitly coded into a program using a software language (Python, JavaScript, C#, Rust, …).

Classic Computers - Compiling

The code is then *compiled* using software to translate the programmer's source code into a version that can be run on a target computer/browser/phone. For most of today's programs, the computer used to develop and compile the code does not have to be that much faster than the one that will run it.

Classic Computers - Running/Executing Programs
Once a program is coded and compiled, it can be deployed and run (executed) on a desktop computer, phone, in a browser window, a data center cluster, in special hardware, etc. Programs/applications can be games, social media, office applications, missile guidance systems, bitcoin mining, or even operating systems e.g. Linux, Windows, IOS. These programs run on the same type of classic computer architectures they were programmed in.



Classic Computers – Software Updates, New Features
For programs written for classic computers, software developers receive bug reports, monitor for security breaches, and send out regular software updates that fix bugs, increase performance and at times add new features.

Classic Computers- Hardware
The CPUs (Central Processing Units) that write and run these Classic Computer applications all have the same basic design (architecture). The CPUs are designed to handle a wide range of tasks quickly in a *serial* fashion. These CPUs range from Intel X86 chips, and the ARM cores on Apple M1 SoC, to the z15 in IBM mainframes.

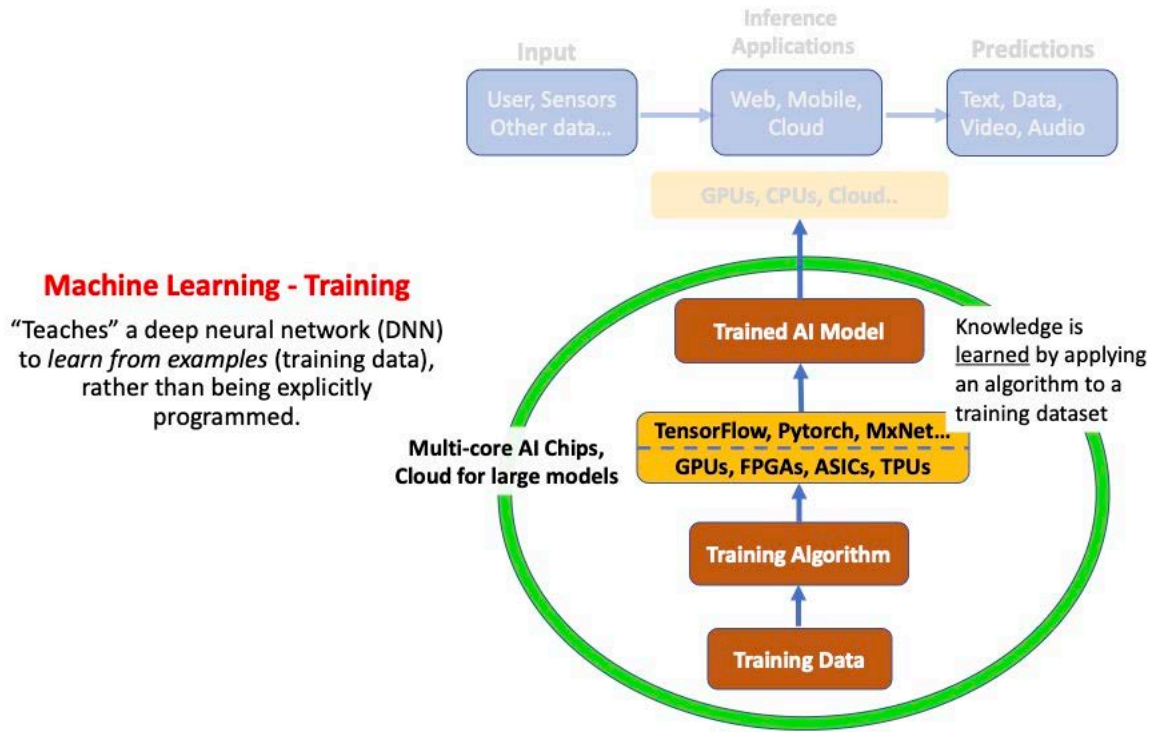
Machine Learning
In contrast to programming on classic computing with fixed rules, machine learning is just like it sounds – we can train/teach a computer to "learn by example" by feeding it lots and lots of examples. (For images a rule of thumb is that a machine learning algorithm needs at least 5,000 labeled examples of each category in order to produce an AI model with decent performance.) Once it is trained, the computer runs on its own and can make predictions and/or complex decisions.

Just as traditional programming has three steps - first *coding* a program, next *compiling* it and then *running* it - machine learning also has three steps: *training* (teaching), *pruning* and *inference* (predicting by itself.)

Machine Learning - Training

Unlike programing classic computers with explicit rules, training is the process of "teaching" a computer to perform a task e.g. recognize faces, signals, understand text, etc. (Now you know why you're asked to click on images of traffic lights, cross walks, stop signs, and buses or type the text of scanned image in ReCaptcha.) Humans provide massive volumes of "training data" (the more data, the better the model's performance) and select the appropriate algorithm to find the best optimized outcome.

(See the detailed "machine learning pipeline" later in this section for the gory details.)



By running an algorithm selected by a data scientist on a set of training data, the Machine Learning system generates the rules embedded in a trained model. *The system learns from examples* (training data), rather than being explicitly programmed. (See the "Types of Machine Learning" section for more detail.) This self-correction is pretty cool. An input to a neural net results in a guess about what that input is. The neural net then takes its guess and compares it to a ground-truth about the data, effectively asking an expert "Did I get this right?" The difference between the network's guess and the ground truth is its *error*. The network measures that error, and walks the error back over its model, adjusting weights to the extent that they contributed to the error.)

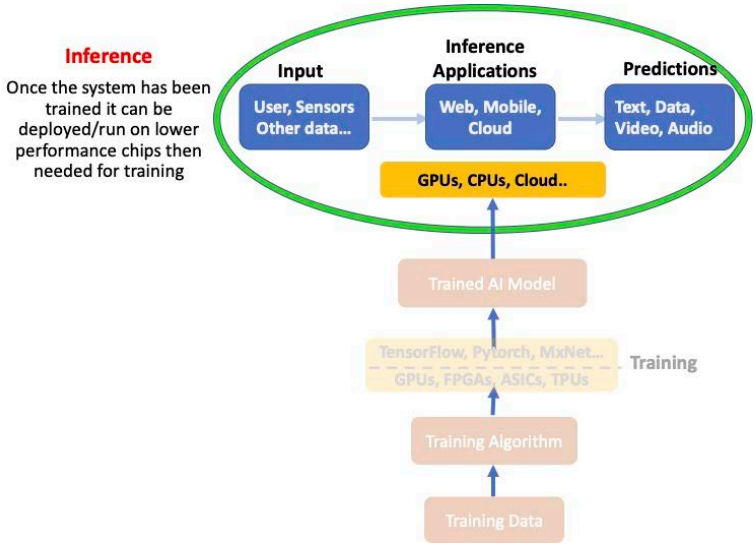
Just to make the point again: *The algorithms combined with the training data - not external human computer programmers - create the rules that the AI uses*. The resulting model is capable of solving complex tasks such as recognizing objects it's never seen before, translating text or speech, or controlling a drone swarm.

(Instead of building a model from scratch you can now buy, for common machine learning tasks, *pretrained models* from others and here, much like chip designers buying IP Cores.)

Machine Learning Training - Hardware
Training a machine learning model is a very computationally intensive task. AI hardware must be able to perform thousands of multiplications and additions in a mathematical process called matrix multiplication. It requires specialized chips to run fast. (See the AI hardware section for details.)

Machine Learning - Simplification via pruning, quantization, distillation
Just like classic computer code needs to be compiled and optimized before it is deployed on its target hardware, the machine learning models are simplified and modified (pruned) to use less computing power, energy, and memory before they're deployed to run on their hardware.

Machine Learning – Inference Phase
Once the system has been trained it can be copied to other devices and run. And the computing hardware can now make inferences (predictions) on new data that the model has never seen before.



Inference can even occur locally on edge devices where physical devices meet the digital world (routers, sensors, IOT devices), close to the source of where the data is generated. This reduces network bandwidth issues and eliminates latency issues.

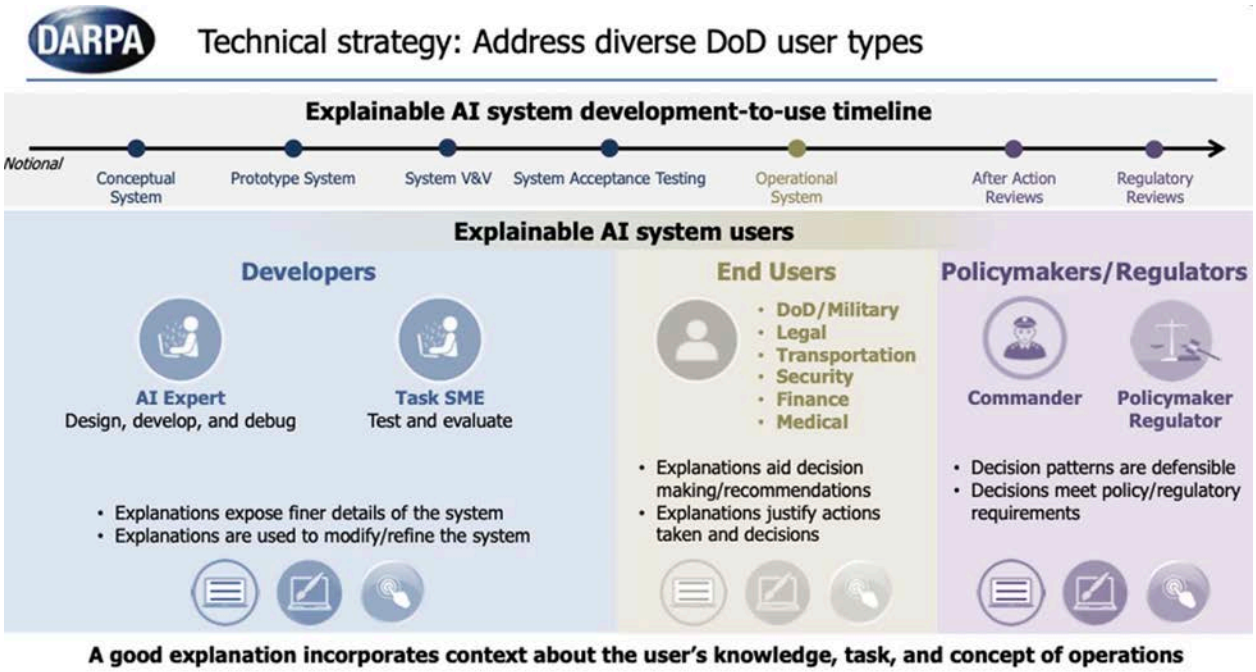
Machine Learning Inference - Hardware
Inference (running the model) requires substantially less compute power than training. But inference also benefits from specialized AI chips.

Machine Learning – Performance Monitoring and Retraining
Just like classic computers where software developers do regular software updates to fix bugs and increase performance and add features, machine learning models also need to be updated regularly by adding new data to the old training pipelines and running them again. Why?

Over time machine learning models get stale. Their real-world performance generally degrades over time if they are not updated regularly with new training data that matches the changing state of the world. The models need to be monitored and retrained regularly for data and/or concept drift, harmful predictions, performance drops, etc. To stay up to date, the models need to re-learn the patterns by looking at the most recent data that better reflects reality.

One Last Thing – "Verifiability/Explainability"
Understanding how an AI works is essential to fostering trust and confidence in AI production models.

Neural Networks and Deep Learning differ from other types of Machine Learning algorithms in that they have low explainability. They can generate a prediction, but it is very difficult to understand or explain how it arrived at its prediction. This "explainability problem" is often described as a problem for all of AI, but it's primarily a problem for Neural Networks and Deep Learning. Other types of Machine Learning algorithms – for example decision trees – have very high explainability. The results of the five-year DARPA Explainable AI Program (XAI) are worth reading here.



A good explanation incorporates context about the user's knowledge, task, and concept of operations

So What Can Machine Learning Do?[1]

It's taken decades but as of today, on its simplest implementations, machine learning applications can do some tasks better and/or faster than humans. Machine Learning is most advanced and widely applied today in processing text (through Natural Language Processing)

[1] https://databricks.com/discover/pages/the-democratization-of-artificial-intelligence-and-deep-learning

followed by understanding images and videos (through Computer Vision) and analytics and anomaly detection. For example:

Recognize and Understand Text/Natural Language Processing
AI is better than humans on basic reading comprehension benchmarks like SuperGLUE and SQuAD and their performance on complex linguistic tasks is almost there**.** Applications: GPT-3, M6, OPT-175B, Google Translate, Gmail Autocomplete, Chatbots, Text summarization.

Write Human-like Answers to Questions and Assist in Writing Computer Code
An AI can write original text that is indistinguishable from that created by humans. Examples GPT-3, Wu Dao 2.0 or generate computer code. Example GitHub Copilot, Wordtune

Recognize and Understand Images and video streams

An AI can see and understand what it sees. It can identify and detect an object or a feature in an image or video. It can even identify faces. It can scan news broadcasts or read and assess text that appears in videos. It has uses in threat detection - airport security, banks, and sporting events. In medicine to interpret MRI's or to design drugs. And in retail to scan and analyze in-store imagery to intuitively determine inventory movement. Examples of ImageNet benchmarks here and here

Detect Changes in Patterns/Recognize Anomalies

An AI can recognize patterns which don't match the behaviors expected for a particular system, out of millions of different inputs or transactions. These applications can discover evidence of an attack on financial networks, fraud detection in insurance filings or credit card purchases; identify fake reviews; even tag sensor data in industrial facilities that mean there's a safety issue. Examples here, here and here.

Power Recommendation Engines

An AI can provide recommendations based on user behaviors used in ecommerce to provide accurate suggestions of products to users for future purchases based on their shopping history. Examples: Alexa and Siri

Recognize and Understand Your Voice
An AI can understand spoken language. Then it can comprehend what is being said and in what context. This can enable chatbots to have a conversation with people. It can record and transcribe meetings. (Some versions can even read lips to increase accuracy.) Examples: Siri/Alexa/Google Assistant. Example here

Create Artificial Images
AI can create artificial images (DeepFakes) that are indistinguishable from real ones using Generative Adversarial

Networks. Useful in entertainment, virtual worlds, gaming, fashion design, etc. Synthetic faces are now indistinguishable and more trustworthy than photos of real people. Paper here.

Create Artist Quality Illustrations from A Written Description
AI can generate images from text descriptions, creating anthropomorphized versions of animals and objects, combining unrelated concepts in plausible ways. An example is Dall-E

Generative Design of Physical Products
Engineers can input design goals into AI-driven generative design software, along with parameters such as performance or spatial requirements, materials, manufacturing methods, and cost constraints. The software explores all the possible permutations of a solution, quickly generating design alternatives Example here.

Sentiment Analysis
An AI leverages deep natural language processing, text analysis, and computational linguistics to gain insight into customer opinion, understanding of consumer sentiment, and measuring the impact of marketing strategies. Examples: Brand24, MonkeyLearn

What Does this Mean for Businesses?

Skip this section if you're interested in national security applications

Hang on to your seat. We're just at the beginning of the revolution. The next phase of AI, powered by ever increasing powerful AI hardware and cloud clusters, will combine some of these basic algorithms into applications that do things no human can. It will transform business and defense in ways that will create new applications and opportunities.

Human-Machine Teaming
Applications with embedded intelligence have already begun to appear thanks to massive language models. For example - Copilot as a pair-programmer in Microsoft Visual Studio VSCode. It's not hard to imagine DALL-E 2 as an illustration assistant in a photo editing application, or GPT-3 as a writing assistant in Google Docs.

AI in Medicine
AI applications are already appearing in radiology, dermatology, and oncology. Examples: IDx-DR, OsteoDetect, Embrace2. AI Medical image identification can automatically detect lesions, and tumors with diagnostics equal to or greater than humans. For Pharma, AI will power drug discovery design for finding new drug candidates. The FDA has a plan for approving AI software here has a list of AI-enabled medical devices here.

Autonomous Vehicles
Harder than it first seemed, but car companies like Tesla will eventually get better than human autonomy for highway driving and eventually city streets.

Decision support
Advanced virtual assistants can listen to and observe behaviors, build and maintain data models, and predict and recommend actions to assist people with and automate tasks that were previously only possible for humans to accomplish.

Supply chain management
AI applications are already appearing in predictive maintenance, risk management, procurement, order fulfillment, supply chain planning and promotion management.

Marketing
AI applications are already appearing in real-time personalization, content and media optimization and campaign orchestration to augment, streamline and automate marketing processes and tasks constrained by human costs and capability, and to uncover new customer insights and accelerate deployment at scale.

Making business smarter: Customer Support
AI applications are already appearing in virtual customer assistants with speech recognition, sentiment analysis, automated/augmented quality assurance and other technologies providing customers with 24/7 self- and assisted-service options across channels.

AI in National Security[2]

Much like the dual-use/dual-nature of classical computers AI developed for commercial applications can also be used for national security.

AI/ML and Ubiquitous Technical Surveillance
AI/ML have made most cities untenable for traditional tradecraft. Machine learning can integrate travel data (customs, airline, train, car rental, hotel, license plate readers…,) integrate feeds from CCTV cameras for facial recognition and gait recognition, breadcrumbs from wireless devices and then combine it with DNA sampling. The result is automated persistent surveillance.

China's employment of AI as a tool of repression and surveillance of the Uyghurs is a dystopian of how a totalitarian regimes will use AI-enable ubiquitous surveillance to repress and monitor its own populace.

[2] https://www.nscai.gov/2021-final-report/

AI/ML on the Battlefield
AI will enable new levels of performance and autonomy for weapon systems. *Autonomously collaborating assets* (e.g., drone swarms, ground vehicles) that can coordinate attacks, ISR missions, & more.

Fusing and making sense of sensor data (detecting threats in optical /SAR imagery, classifying aircraft based on radar returns, searching for anomalies in radio frequency signatures, etc.) Machine learning is better and faster than humans in finding targets hidden in a high-clutter background. Automated target detection and fires from satellite/UAV.

For example, an Unmanned Aerial Vehicle (UAV) or Unmanned Ground Vehicles with on board AI edge computers could use deep learning to detect and locate concealed chemical, biological and explosives threats by fusing imaging sensors and chemical/biological sensors.
Other examples include:

Use AI/ML countermeasures against adversarial, low probability of intercept/low probability of detection (LPI/LPD) radar techniques in radar and communication systems.

Given sequences of observations of unknown radar waveforms from arbitrary emitters without a priori knowledge, use machine learning to develop behavioral models to enable inference of radar intent and threat level, and to enable prediction of future behaviors.

For objects in space, use machine learning to predict and characterize a spacecrafts possible actions, its subsequent trajectory, and what threats it can pose from along that trajectory. Predict the outcomes of finite burn, continuous thrust, and impulsive maneuvers.

AI empowers other applications such as:

- Flight Operations Planning Decision Aid Tool for Strike Operations Aboard Aircraft Carriers
- Automated Battle management – air and missile defense, army/navy tactical…

AI/ML in *Collection*
The front end of intelligence collection platforms has created a firehose of data that have overwhelmed human analysts. "Smart*" sensors* coupled with inference engines that can pre-process raw intelligence and prioritize what data to transmit and store –helpful in degraded or low-bandwidth environments.

Human-Machine Teaming in *Signals Intelligence*
Applications with embedded intelligence have already begun to appear in commercial applications thanks to massive language models. For example - Copilot as a pair-programmer in

Microsoft Visual Studio VSCode. It's not hard to imagine an AI that can detect and isolate anomalies and other patterns of interest in all sorts of signal data faster and more reliably than human operators.

AI-enabled natural language processing, computer vision, and audiovisual analysis can vastly reduce manual data processing. Advances in speech-to-text transcription and language analytics now enable reading comprehension, question answering, and automated summarization of large quantities of text. This not only prioritizes the work of human analysts, it's a major force multiplier

AI can also be used to automate data conversion such as translations and decryptions, accelerating the ability to derive actionable insights.

Human-Machine Teaming in *Tasking and Dissemination*
AI-enabled systems will automate and optimize tasking and collection for platforms, sensors, and assets in near-real time in response to dynamic intelligence requirements or changes in the environment.

AI will be able to automatically generate machine-readable versions of intelligence products and disseminate them at machine speed so that computer systems across the IC and the military can ingest and use them in real time without manual intervention.

Human-Machine Teaming in *Exploitation and Analytics*
AI-enabled tools can augment filtering, flagging, and triage across multiple data sets. They can identify connections and correlations more efficiently and at a greater scale than human analysts, and can flag those findings and the most important content for human analysis.
AI can fuse data from multiple sources, types of intelligence, and classification levels to produce accurate predictive analysis in a way that is not currently possible. This can improve indications and warnings for military operations and active cyber defense.

AI/ML *Information warfare*
Nation states have used AI systems to enhance disinformation campaigns and cyberattacks. This included using "DeepFakes" (fake videos generated by a neural network that are nearly indistinguishable from reality). They are harvesting data on Americans to build profiles of our beliefs, behavior, and biological makeup for tailored attempts to manipulate or coerce individuals.

But because a large percentage of it is open-source AI is not limited to nation states, AI-powered cyber-attacks, deepfakes and AI software paired with commercially available drones can create "poor-man's smart weapons" for use by rogue states, terrorists and criminals.