DUAL CITIZEN

Can AI & Machine Learning Accelerate a 2020s Green Breakthrough?

Author: Karuna Ramakrishnan, Facebook Public Policy
Contributor: Jeremy Tamanini, Dual Citizen LLC
October 2020

Machine learning (ML) and big data are here to stay. In 2020, machine learning and its sub-stream of natural language processing (NLP) are already proving to be disruptive due to innovations in fields as diverse as transport planning, healthcare, disaster prevention and public service delivery. We have also seen some of these cutting-edge applications piloted in innovative projects related to the environment. Machine Learning and the green economy are a seamless fit, given that numerous outputs linked to the environment, climate change, green industrial sectors and investors are in terms of data, indices, and written documents, such as impact studies or economic analyses.

ML can be used effectively to mine huge amounts of data in order to reveal insights and connections that will help optimize the design and outcomes of projects. It can also generate skewed outcomes through faulty design or inherent bias. As is the case with traditional data analysis, the motto "a model is only as good as the underlying assumptions and data" holds true for ML too. However, ML may offer a new approach to accelerating a green breakthrough in the 2020s, something the Dual Citizen practice has been focused on in 2020 through a recent survey of our experts, an update to our climate modeling project linked to the upcoming COP26 and soon, a 10th anniversary Global Green Economy Index™ (GGEI) looking at progress over the past decade.

The purpose of this Insight is to advance the discussion around AI, ML and the green economy by offering an overview of the basic concepts in these fields. It also explores whether AI can reveal new, sharper insights from unstructured data sets related to the commitment of actors and institutions to green growth. Can it accelerate data collection and localization from data sets related to renewable energy, air quality or forests? And how has the sheer quantity of data now available changed the debates around climate change? It also provides background on what the basics of Machine Learning are, how companies are taking advantage of ML, and what led to its proliferation.

Keywords[1]

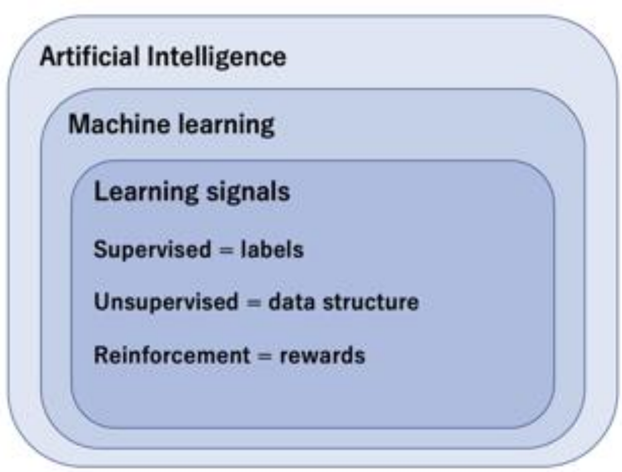
Artificial intelligence = machines that perceive the environment and take actions to achieve goals.

Machine learning = a branch of AI, that gives computers the ability to learn high dimensional patterns from data.

Deep learning = a family of machine learning models, that use multi-layered neural networks to approximate functions.

Machine Learning Basics

Machine Learning as a field deals with computer programs that try to learn from experience. Usually, it is for a purpose, typically prediction, modeling and trying to understand data or trying to control something. Previously, one of the hardest parts about using computers is that you somehow had to tell the machines in painstaking detail how to do what you wanted them to. But machine learning is a promising approach, that instead of writing detailed instructions for each different kind of problem, you just write a very general program. This general program includes instructions for how the machine can learn from experience. This way, you can teach the machine what to do by giving them lots of experience, and by telling them the right and wrong answers along the way. Machine Learning is generally categorized into three types: Supervised Learning, Unsupervised Learning, Reinforcement learning:



Source: MIT School of Engineering and Computer Sciences

[1] Definitions derived from the MIT School of Engineering and Computer Sciences.

Supervised Learning: In supervised learning, the machine experiences the examples along with the labels or targets for each example. The labels in the data help the algorithm to correlate the features. In supervised learning, the machine uses labelled training data to learn how to predict the labels of unseen data. Examples include time series forecasting, computer vision and language translation. Supervised learning is the reason why Facebook can tell which of your friends is in your photo, or why Google can translate text on your smartphone. Two of the most common supervised machine learning tasks are classification and regression. In classification problems, the machine must learn to predict discrete values. That is, the machine must predict the most probable category, class, or label for new examples. Applications of classification include predicting whether a stock's price will rise or fall or deciding if a news article belongs to the politics or leisure section. In regression problems the machine must predict the value of a continuous response variable. Examples of regression problems include predicting the sales for a new product or the salary for a job based on its description.

Unsupervised Learning: The second is unsupervised learning, where the machine is able to generate new data without the supervision of labels. Examples include artistic style transfer and generating realistic faces. Generative Adversarial Networks (GANs) learn to generate realistic images using two competing neural networks. One network generates images (the generator) and a second network has to decide if the image is real or fake. This kind of adversarial learning can be effective. When we have unclassified and unlabeled data, the system attempts to uncover patterns from the data . There is no label or target given for the examples. One common task is to group similar examples together called clustering.

Reinforcement Learning: Adversarial learning can also be used in our final branch of machine learning: reinforcement learning. In reinforcement learning the machine learns to select actions with the supervision of a scalar reward signal. Reinforcement learning is applicable to a wide range of decision-making problems with a reward signal, such as cost or carbon emissions.

Drivers of Modern Machine Learning[2]

The performance of modern deep learning is driven by the interaction of two processes — the increased availability of data and the computing power to train large models with lots of data.[3] The rise of the Internet and devices that generate raw data (sensors, images and text) led to the curation of massive datasets. These massive datasets are the fuel of deep neural networks — without the data, the models can't learn.

The ability to train large models rests upon the ability to access specialized hardware in the cloud. In the 2000s researchers repurposed hardware designed for video games (graphics processing units, or GPUs) to train neural networks.[4] This led to dramatic speed-up in training times, which is important — all our understanding of machine learning is empirical (learned through experiment).

The second hardware trend is cloud computing. The cloud gives access to computation on a fully variable cost basis. Platforms such as Amazon Web Services allow on-demand access to a large amount of GPU-enabled computing power with cheap data storage alongside it. This access to computing power works both vertically within large technology companies and for smaller companies.

A final trend driving modern machine learning is access to algorithms and tools. Almost all the relevant literature for machine learning is available for free on sites like arXiv. It is also possible to access high quality implementations of machine learning tools on GitHub.[5] This tendency for openness stands in stark contrast with the paywalls and licensed software of the energy industry.

---

[2] *BBC.com*, "15 key moments in the story of artificial intelligence," https://www.bbc.co.uk/teach/ai-15-key-moments-in-the-story-of-artificial-intelligence/zh77cqt

[3] Stanford University, "One Hundred Year Study on Artificial Intelligence (AI100)", https://ai100.stanford.edu/ 2016.

[4] Hao, Karen, *MIT Technology Review*, "The start-up making deep learning possible without specialized hardware," https://www.technologyreview.com/2020/06/18/1003989/ai-deep-learning-startup-neural-magic-uses-cpu-not-gpu/, June 2020.

[5] Kinsella, Brett, "Artificial Intelligence Primer from Frank Chen of Andreessen Horowitz," https://voicebot.ai/2017/01/08/artificial-intelligence-primer-frank-chen-andreessen-horowitz/ January 2017.

How are businesses leveraging ML techniques?

There are two main ways of using machine learning in business – sensing and predicting[6]. Sensing means perceiving large amounts of data from sensors in the world and learning how to recognize what's there. This is especially useful with images, but it can also be used for sensing things like sounds, vibrations, and air quality; recognizing faces, retinas or fingerprints; and ultimately analyzing what is present in a scene. For example, machine learning helps self-driving cars recognize other cars, pedestrians and other objects in a visual landscape.

The second way of using machine learning in business is for predicting things. Programs can use large amounts of data, of any kind, to predict what will happen in the future. For instance, programs based on massive amounts of data can help predict financial fraud; disease probabilities based on symptoms and lab tests; chance of mechanical failure based on various kinds of sound and vibration data; and crop yields based on photographs of the fields and information about weather.

Big companies are employing ML because they're seeing it as positive return on investment and not only from products like Siri and Amazon Echo that are already widely adopted by consumers. Corporate deployment of ML is not concentrated in companies that we normally think of as having significant research and development budgets like Google and Microsoft. In 2020, nearly every Fortune 500 company is already running more efficiently[7] — and making more money — because of machine learning.

[6] Derived from coursework at MIT Sloan School, Session 2: AI and Machine Learning.
[7] Biewald, Lukas, *Tech Crunch*, "How real businesses are using machine learning," https://techcrunch.com/2016/03/19/how-real-businesses-are-using-machine-learning/ March 2016.

Applying Machine Learning to the Green Economy Space

We developed a matrix that addresses how we can use machine learning to reinforce the use of AI to accelerate the green economy transition. The examples in the matrix show how AI might transform traditional sectors and systems to address climate change, deliver food and water security, build sustainable cities, and protect biodiversity and human well-being. Specifically, the matrix addresses how we can use three types of data - sovereign, company level and individual data - in tracking, predicting and optimizing[8] our lives and the environment. There is some overlap between the three categories: for example, citizen data for weather patterns is data collected at the individual level but taken collectively provides a country level picture of weather conditions.

Tracking: Tracking data is becoming more effective and inventive as ever. In one case, seals were outfitted with specialized sensors that resemble lumpy metal yarmulkes with antennae[9], to collect data that's helping researchers track how heat moves through ocean currents. In a paper published this week in *Nature Geosciences*, a team of climate scientists led by Caltech oceanographer Lia Siegelman used this clever technique to track changes in temperature as the seal swam the icy waters of the Antarctic. Seals aside, the Internet of Things phenomena has enabled us to collect vast quantities of climate data in cost-effective ways. At the country level, this takes the form of Microsoft's AI for March project that extracts building footprints from satellite imaging to get city level energy consumption data. Companies like BlueSource are tracking countries' compliance to global climate agreements. In finance, with sustainable investing growing in importance, TruValue Labs and Sensefolio track ESG data at the country level to inform green portfolios. At the firm level, we see machine learning tracking data from drones that monitor weather conditions and for precision agriculture[10]; we see companies tracking shipping vessel algorithmic patterns to identify illegal fishing[11], and other agriculture firms such as Blue River detecting and removing weeds automatically using AI algorithms and cameras. Tracking at the individual level ranges from the commonplace, such as household smart

---

[8] Pyle, Dorian and San José, Christina; "An Executive's Guide to Machine Learning," https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/an-executives-guide-to-machine-learning June 2015.

[9] Wu, Catherine J., *Smithsonian Magazine*, "Seals with high tech hats collect climate data in the arctic," December 2019.

[10] Muzar, Michal, *MIT Technology Review*, "Six Ways drones are revolutionizing agriculture," July 2016.

[11] Revkin, Andrew, *Dotearth Blog New York Times*, "How digital tracking of rogue fishing can safeguard vast ocean reserves," September 2016.

water meters that spot inconsistencies and check leaks to citizen data that is tracked to monitor air quality data or even interestingly, whale migration patterns.[12]

Functions/ Data Types	Sovereign Data	Company Data	Individual Data
Tracking	Extracting building footprints from satellite imaging to feed algorithms to extract city level energy consumption (Microsoft AI for Earth).[13]	Detecting and removing weeds automatically using AI algorithms and cameras in agriculture (Blue River)	Mapping global biodiversity through contributions from citizen scientists, with algorithms providing species verification (iNaturalist)[14]
	Analyzing and interpreting unstructured ESG datasets to better inform green investments (TruValue Labs, SenseFolio, Arabesque S-Ray)	Drones used to monitor weather conditions and for precision agriculture that communicate with the sensors and AI-enabled systems (SKycision, American Robotics, SenseFly)[15]	Citizen science in Norway to track and collect data, with examples as far ranging as air quality data and whale migration patterns.
	Tracking compliance under climate agreements (Verra, BlueSource, compliance.ai)	Tracking shipping vessel algorithmic patterns could identify illegal fishing (Global Fishing Watch), biological sensors could monitor the health of coral reefs (Allan Coral Atlas by Vulcan) and ocean current patterns could improve weather forecasting.	Household smart water meters can produce large volumes of data that can be used to predict water flows, spot inconsistencies and check leaks.

Predicting: Machine learning predictions allow businesses to make highly accurate guesses as to the likely outcomes of a question based on historical data, which can be about all kinds of things – customer churn likelihood, possible fraudulent activity, and more. These provide the business with insights that result in tangible business value. With a climate change lens, we see a lot of prediction models using ESG to price country bonds. On climate models, public agencies like NASA and large tech firms like IBM are using machine learning to enhance the predictability of weather data. Other companies are using data intelligence to forecast energy demand with a weather and household behavior lens. At an individual level, there are companies like AutoGrid who are predicting electric vehicle charging behavior so that grid operators can better manage

[12] "Transforming citizen science for biodiversity," *Norwegian University of Science and Technology*, https://citizenscience.no/
[13] "AI for Earth," *Microsoft*, https://www.microsoft.com/en-us/ai/ai-for-earth
[14] "Transforming citizen science for biodiversity," *Norwegian University of Science and Technology*, https://citizenscience.no/
[15] PwC, "Clarity from above: commercial applications of drone technology," https://www.pwc.pl/pl/pdf/clarity-from-above-pwc.pdf , May 2016.

their load; and food apps are getting better at predicting what nutrition plan might work for particular individuals.

Functions/ Data Types	Sovereign Data	Company Data	Individual Data
Predicting	Using ESG data guidance to price country bonds	Forecasting energy demand taking into account finer local weather patterns and household behavior (NEC Turkey	EV adoption: modeling and predicting aggregate charging behavior to help grid operators manage load (AutoGrid)
	Public agencies (UK Met, NASA), and private sector (IBM and Microsoft), are machine learning to enhance the performance and efficiency of weather and climate models.	Acting on data intelligence gathered across the electrical grid to optimize energy efficiency (Agder Energi)	Monitor individual food intake, by applying machine learning to this data could generate personalized nutrition plans optimized for individuals. (MyFitnessPal)

Optimizing: the ideal end-use case for ML is to optimize processes and usage. This can range from IA capable virtual power plans that optimizes the use of solar panels, micro grids and other energy storage installations. A step further, solar roads in countries as diverse as China and France allow roads to learn to heat up in order to melt snow or to adjust traffic lanes based on vehicle flows. On the weather side, Met UK has been optimizing available information so that a chatbot app can extract frictionless data queries. At a firm level, logistics is an area rich with process and route optimization such as bundling shipments. Energy consumption by smart building is being optimized by adjusting heating and cooling according to the weather by companies like clevair.io and IBM. For individuals, ML is helpful with smart meters that can also help forecast and optimize urban energy generation and demand at homes and buildings, which are being adapted by utilities such as UMass Smart, ISO New England, and Ausgrid Residents.

Functions/ Data Types	Sovereign Data	Company Data	Individual Data
Optimizing	AI-capable "virtual power plants" (VPPs) can integrate, aggregate, and optimize the use of solar panels, microgrids, energy storage installations and other facilities (Next Kraftewerk).	Bundling shipments, optimizing freight shipping routes (nuro, TNX, Transmetrics, Shyp)	Lower barriers to EV adoption by improving battery energy management to increase mileage per each charge
	AI-enabled "solar roads" to expand, connect and optimize the grid further (SolaRoad, Colas, Hannah Solar). In solar roads, for example, AI could allow a road to learn to heat up to melt snow, or to adjust traffic lanes based on vehicle flow.	Reduce smart building energy consumption (clevair.io, IBM IoT) by taking weather and occupancy into account and adjusting heating/cooling and ventilation	Smart meters that can also help forecast and optimize urban energy generation and demand at the level of individual homes and buildings. (UMass Smart, ISO New England, Ausgrid Residents)
	Weather data: Chatbot application to demonstrate how "frictionless" data or queries can be extracted from complex big datasets, using sophisticated AI in real time and communicating to the user through a simple interface (UK Met).	Real-time AI-optimized energy efficiency can have an immediate and substantial impact on energy consumption (Google, for example, cut power use in its data centres by 40% by using DeepMind's reinforcement learning algorithms to optimize cooling.	Using AI to collect millions of water related data points each day, making the resulting analysis available to investors looking to build green portfolios. (Aquantix)

Conclusion

As the 2020s begin, we remain stuck in a large gap between the ambition required to hold global temperature rise under 1.5 degrees Celsius and the ability of current efforts to achieve this target. Action will emerge from all directions this decade as grassroots pressure, public policy interventions, private sector initiatives and technological innovation interact to hopefully reorient the global economy to a more sustainable model of living. As this Insight reveals, the new intersection of AI and ML with this green economy sometimes offers more questions that clear direction for action. But at the same time, these fields show promise in helping to accelerate progress, something that is desperately needed as the climate crisis intensifies.

About Dual Citizen LLC

We are a consultancy working globally at the intersection of data measurement and the green economy. To support this work, we have published the Global Green Economy Index™ (GGEI) for the past decade, tracking how 130 countries perform in the green economy based on 20 underlying topics, contained within four key dimensions: leadership & climate change, efficiency sectors, markets & investment and the environment. As part of consulting assignments and to contribute knowledge leadership in this field, we often publish Insights like this one containing new data and information for the sustainability community. All of these recent Insights can be accessed in the GGEI Library by clicking here.