aws

Tackling our world's hardest problems with machine learning

Table of Contents

A letter from Swami Sivasubramanian

Machine learning has graduated from the realm of science fiction to become a core, transformative technology for organizations across industries and categories. The unique potential and power of machine learning are sparking significant innovation, powering the ideas that are improving lives and protecting our planet right now. With machine learning, organizations are making inroads toward protecting and supporting our veterans, finding homes for homeless people, understanding climate change, improving health outcomes, and more. But this is just the beginning.

The technology is ripe, and it now has the ability to provide new and significant solutions for some of the world's biggest challenges.

More than a hundred thousand companies and organizations worldwide have turned to Amazon Web Services (AWS) for machine learning—to help track disease outbreaks worldwide, find new ways to treat cancer and more.

However, access to machine learning, a new technology to so many of these organizations, can often come with a skills and technology deficit. That's where AWS steps in, partnering with innovators to bridge the gap and bring pioneering solutions that can help tackle our most urgent and important challenges.

– Swami
Vice President, Amazon Machine Learning

Tackling our world's hardest problems with machine learning

Amazon Web Services (AWS) puts machine learning in the hands of every developer, data scientist, and expert practitioner. And every day, these individuals leverage and apply machine learning in new ways for the benefit of society. From decreasing carbon emissions to improving safety in the workplace, the methods by which today's leaders apply machine learning reveal a tremendous opportunity to improve lives and the world in which we live.

What are artificial intelligence and machine learning?

You've probably heard artificial intelligence (AI) and machine learning (ML) described in a number of ways, so let's take a step back and review their exact definitions:

Artificial intelligence is a way to describe any system that can replicate tasks that previously required human intelligence. Almost always, this is related to some kind of complex decision-making task where human judgment would normally be required. Most use cases for AI are looking for a probabilistic outcome—making predictions, classifications, or decisions with a high degree of certainty and in ways that are similar to human judgment.

Almost all AI systems today are created using machine learning. Machine learning uses large amounts of data to create and validate decision logic. This is known as a model. The AI system feeds input data into that model, and then the model outputs humanlike predictions or classifications. Essentially, machine learning is the underlying technology that powers intelligent systems.

Now, we'll take a look at a selection of real-world use cases of companies and organizations using machine learning to change the world for the better.

Building a sustainable future

Amazon is committed to building a sustainable future for our customers and the planet, from creating sustainability in the cloud to the ambitious goals of The Climate Pledge. As companies and organizations across the globe make their commitments to sustainability, machine learning will play its part in ensuring success.

aws

Analyzing cloud patterns to better understand climate change

As a recipient of an Amazon Research Award, the **University of Oxford** and its climate scientists are working to unearth new ways to combat climate change. The award program offers unrestricted funds and AWS Promotional Credits to support research at academic institutions and nonprofit organizations in areas that align with our mission to advance the science of customer obsession.

Machine learning is an essential tool for climate change research since climate science is such a data-intensive field. Climate models are enormous, requiring supercomputers to run them, and analysis requires a huge amount of earth observation data. As data continues to grow, along with its complexity, it becomes impossible to explore all avenues of research manually. This is one of many ways Oxford and AWS continue to work together, including a collaboration to fund a test bed of new research in AI and data science across the university.

The Climate Processes Research Group in the Department of Physics at Oxford hopes that by studying the effects of aerosol pollution on clouds, they'll be able to break new ground in global warming research, leveraging tools like AWS Deep Learning AMIs running on Amazon Elastic Compute Cloud (Amazon EC2). Clouds reflect sunlight back to space, acting like an umbrella that cools the earth. So, even small changes in clouds in response to global warming or air pollution could have a big impact on environmental health and serve to accelerate or dampen the greenhouse effect. Machine learning models can track these changes to understand why clouds change, which could be the key to addressing global warming.

Now Oxford scientists are able to analyze satellite data covering the entire earth multiple times a day, providing countless images of aerosol-impacted clouds, which they're able to process in the AWS Cloud, thanks to the AWS Imagine Grant program. Such scalable machine learning techniques allow experts to make rapid progress in an area where researchers previously spent months of their time manually identifying features in fairly limited datasets.

Using machine learning to eliminate 915,000 tons of packaging to date

Amazon sells hundreds of millions of different products and ships billions of items a year. To ship with minimal packaging at maximum speed and ensure the customer's order arrives at their doorstep undamaged, the team must innovate at a large scale. The goal is to scale decision-making across the hundreds of millions of products that are shipped—not to automatically default to boxes but instead to identify items that can be shipped in flexible packaging options such as mailers (padded paper envelopes) or bags, which are more sustainable.

In practice, this meant creating machine learning algorithms built on terabytes of product data, from product descriptions to customer feedback. Working closely with AWS Professional Services, these terabytes of data are cleaned, cataloged, and ready for mining. The machine learning algorithms then ingest that data to identify the best packaging with the least waste. Using Amazon SageMaker, the packaging team can analyze hundreds of millions of products, billions of customer shipments, and multiple channels of customer feedback, providing actionable insights in real time.

Some of the most impactful machine learning models identify products that don't need any packaging at all—like diapers. Others are designed to determine product categories like toys and differentiate collectibles, where the type of the original packaging is important to the customer. These products are shipped with the protection of an Amazon shipping box versus a stuffed animal, which can be shipped safely in a bag.

Machine learning has accelerated changes in packaging mix significantly, reducing the use of boxes from 69 percent to 42 percent, and it's still Day 1!

Machine learning is an essential tool to drive innovation to support sustainability. It empowers the potential for real business change and innovation, including efficiency improvements, renewable energy, materials reductions, and other carbon emission elimination strategies.

Flexible packaging is **75%** lighter than a similarly sized box and will conform around a product, taking up **40%** less space than a box during shipping—which means a lot fewer trucks on the road.

Tackling one of the biggest single sources of CO_2 emissions

Improving the energy systems of buildings is a complex, dynamic pursuit. Buildings are an inherently chaotic space, with a mixture of control systems, interoperability, data standards, and equipment configurations, worsened by the fact that most buildings standing today predate the end of the Cold War. With machine learning, **Carbon Lighthouse** can analyze, optimize, and monitor the lighting and HVAC (heating, ventilation, and air conditioning) systems to maximize efficiency.

To understand how buildings react and respond to dynamic changes throughout the day, Carbon Lighthouse uses its patented CLUES® platform to analyze real building data from across multiple sources, including utility submeter and building management system (BMS) data, as well as data on occupancy and weather variables.

Factors like weekdays vs. weekends, a typical workday vs. a holiday, or a winter morning at 6 a.m. vs. a warm summer afternoon at 4 p.m. can greatly impact how efficiently a building is managed. The data and machine learning algorithms built on AWS play integral roles in the value that Carbon Lighthouse delivers to its clients. The right building data analyzed with machine learning enables Carbon Lighthouse to translate millions of data points into actionable insights on a building's energy use over time and as the building dynamics change and shift. Carbon Lighthouse has now processed data from over 100 million square feet of commercial real estate (CRE), building a deep dataset of real building insights that allows Carbon Lighthouse to uncover and apply operational, financial, and climate impact at scale.

Creating financial opportunities for more

Machine learning is a valuable tool that can unlock new business potential for organizations around the world, but it also has the power to serve and reach individuals and provide them with new financial opportunities. From small-business owners to underserved consumers looking for access to credit-based economies, machine learning creates new opportunities.

Unlocking the financial potential of millions with machine learning

About 2.5 billion people around the world are underserved by traditional financial institutions. According to one World Bank estimate, approximately 68 percent of adults have no credit data and, therefore, no credit score. And in most of these areas, national IDs do not exist or are inconsistently enforced throughout the country. As a result, about 85 percent of the world's transactions are cash-based, limiting or preventing their access to the global economy and creating a frustrating, vicious cycle.

Without financial data, it's hard for financial services companies to identify this population, to reach them, and to transact with them.

But financial services company **Tala** recognized that with machine learning powered by AWS, they could find this population of unbanked consumers and serve them with a modern credit infrastructure built from scratch. They'd be the first to offer loans to this population of previously unbanked consumers and unlock their financial potential in the marketplace: their ability to participate as both buyers and as business owners and contribute to their community.

Machine learning solutions often hinge on human understanding. In the case of Tala, sitting down with real customers provided the foundational insights, and machine learning provided the ability to scale. With Amazon SageMaker, Tala has adopted automated processes end to end, leveraging machine learning at each step of the decision process to approve or decline a loan, to determine loan amounts, or to make decisions concerning collections and acquisitions. The entire flow uses machine learning to automate as much as possible and reduce the cost, enabling them to serve more customers at a better price.

To date, Tala's innovative machine learning solution has offered more than **5 million customers over $2 billion in loans**.

Keeping employees safe

Australian-based **Bigmate**, a computer vision company focused on enhancing workplace safety, is using machine learning to reduce workplace accidents and help companies detect potentially ill employees as they arrive on-site.

Reducing accidents with a high degree of accuracy

Bigmate developed Warny™ to enhance safety in the workplace and reduce workplace accidents. With Amazon SageMaker, they created sophisticated computer vision algorithms to protect people working around dangerous machines—such as forklifts, trucks, or manufacturing machinery. It can detect instances of spontaneous combustion of materials, overheating of equipment, and fires in the workplace, as well as analyze, report on, and alert machine operators in real time about unexpected events, such as a person being in an unsafe area, even when not in the line of sight of the operator.

Across the board, on average, Bigmate is seeing an **80%** drop in incidents in their clients' work environment.

Making it safe to go back to work

In a pandemic, the choice between returning to work and staying safe has been particularly challenging. Bigmate's prescreening solution, Thermy™, tackles the issues of reopening and operating safely, using thermal imaging that can immediately detect elevated temperatures of people in real time, at scale, scanning 30 people a second, 500–600 people a minute, plus run 8.3 scans per second to validate its readings. The solution, which is based on the Warny platform and technologies, uses thermal cameras and advanced analytics with machine learning, providing real-time information through dashboards hosted on AWS for remote viewing and trend analysis.

Other thermal solutions only capture skin temperature, which doesn't accurately reflect core body temperature. Bigmate's platform calculates a true representation of a person's core temperature. It first uses computer vision technology and the data from a thermal camera and an optical camera to isolate the subject's head to capture skin temperature even when the subject has a beard, glasses, a hard hat, or other features. A machine learning algorithm can then calculate a representation of the core body temperature to determine whether the subject has a fever.

Improving health outcomes

The healthcare industry has long been an early adopter of advances in technology. Today, machine learning is a key tool in revolutionizing the expansive field of health and medicine. From developing new procedures to collecting and analyzing data, machine learning has the unique ability to power new insights and predictions to improve patient care and population health.

Identifying and tracking pandemics

Pandemics pose a complex challenge—and the urgency to solve the problem is growing. **BlueDot** pioneered outbreak detection solutions with goals of anticipating outbreaks and mitigating risks. To track, contextualize, and anticipate infectious disease risks, the company's unique and powerful solution combines public health information and medical expertise with advanced data analytics and machine learning on AWS.

BlueDot's software consists of a machine learning platform that leverages billions of data points from a vast array of sources in over 65 languages. To identify new outbreaks, it constantly scans foreign-language news reports, animal and plant disease networks, official government announcements, and more than 100 datasets with proprietary algorithms.

AWS is key to processing all of this data, using custom machine learning algorithms that rely on natural language processing to make sense of and structure the data. Using Amazon EC2, they can process massive amounts of unstructured text data into organized, structured, spatiotemporal pathogen data—identifying the space, time, and name of the pathogen.

Their machine learning platform, powered by AWS, anticipates the spread and impact of over 150 different pathogens, toxins, and syndromes in near real time. With this critical information, they're able to advise governments, public health organizations, and other clients on how to disrupt the threat of pandemics—and help get ongoing disease spread under control.
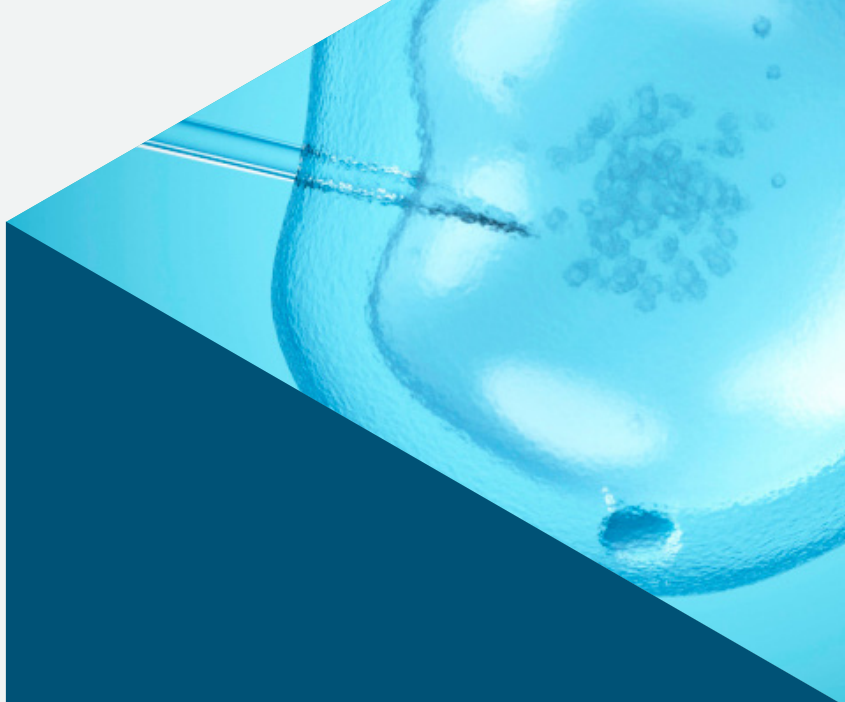
Unlocking the puzzle of infertility

More than eight million babies have been conceived via IVF since the technology was first introduced in 1978. But IVF is not a precise science.

AI healthcare company **Presagen** is one of the companies stepping up in this essential health space, with scalable machine learning that can be used by clinics and patients anywhere in the world.

Currently, embryo selection relies on the embryologist conducting a visual assessment of embryos using a microscope. However, there are only a few macro features the clinician can use to grade those embryos, often resulting in imprecise and subjective decisions. With machine learning, Life Whisperer helps embryologists choose the most viable embryo or embryos the first time—significantly increasing the chance of a successful IVF cycle.

After the embryologist sends captured images of the candidate's embryos to the cloud-based Life Whisperer application, they are assessed by computer vision algorithms. Trained on thousands of historical IVF cases, the algorithm, built using PyTorch on AWS, identifies aspects of the embryo that are most critical in determining viability but invisible to the human eye.

In real time, within 10–15 seconds, the platform returns a report that helps the doctor select the most viable embryos for implantation. The stronger the score, the more likely the chance of pregnancy for an embryo. With their machine learning algorithms, Presagen has improved accuracy in choosing an embryo by 25–30 percent over the current standard of care.

With computer vision technology that helps physicians identify the most viable embryos for implantation, the company's Life Whisperer product can reduce the time to pregnancy for patients on average by 15%.

aws

Treating cancer with the power of the cloud

Cancer Commons is taking the fight against cancer into the cloud. The nonprofit, which helps patients with advanced cancer access personalized treatment options, seeks to overcome a core problem in cancer research: The best cancer treatments involve combinations of therapies, with far too many plausible regimens to test in clinical trials.

Facing heavy computing and data requirements, Cancer Common initially used multiple cloud service providers. But the organization quickly realized the benefits of having all its applications on one infrastructure.

Today, Cancer Commons uses AWS compute infrastructure and services like Amazon Comprehend Medical to enable physicians and patients to leverage the collective knowledge of the world's top institutions.

"Now, almost every part of our process is powered by AWS in some way," said Erika Vial Monteverdi, executive director at Cancer Commons.

When necessary, Cancer Commons convenes virtual tumor boards (VTBs), where experts help patients refine their options based on their individual medical history and preferences. The discussion that happens in the VTBs is mined using natural language parsing and other AI techniques, and a recommended treatment plan is created from that information.

Cancer Commons developed its VTB system on AWS using a HIPAA-compliant architecture. It also relies on Amazon Relational Database Service (Amazon RDS) for the underlying PostgreSQL database services.

"Whether it's raw computing, specialty services such as NLP, specialty computing such as GPU, or optimization services such as Amazon SageMaker, AWS has the tools and services we need," Erika said. "Thanks to AWS, we are helping accelerate cancer research and save lives."
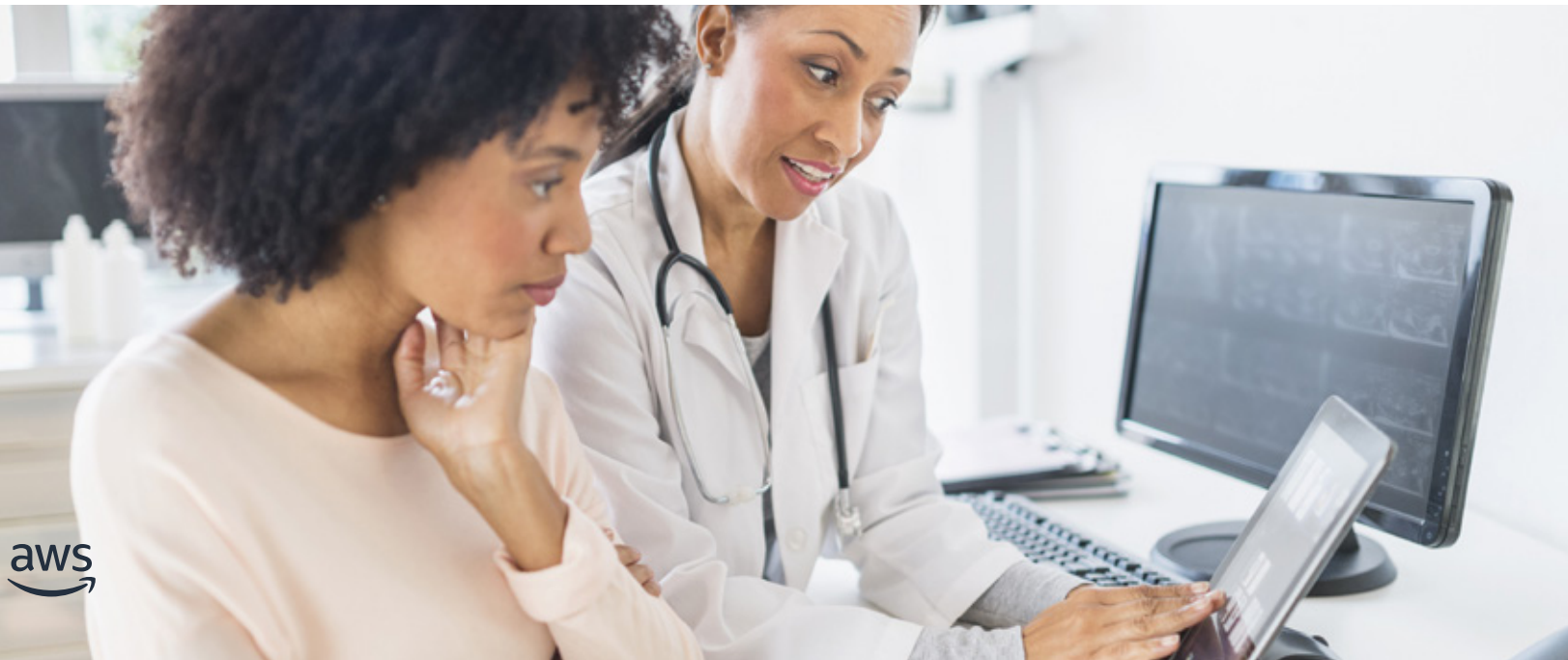
Changing lives for the better

Some of the toughest challenges societies face are managed by nonprofits, government programs, and community-focused organizations. Technology has enabled these organizations to build networks and collaborate on modern solutions to today's biggest challenges. From helping veterans in need to addressing homelessness, machine learning has immense power to improve the lives of individuals and our communities in need.

Addressing homelessness by providing help faster

PATH, a Los Angeles-based organization founded to address the ever-increasing issue of homelessness, applied for an AWS Imagine Grant to develop a way to shorten the time it takes to match homeless individuals and families with homes of their own. With the grant and support from the AWS team, the organization developed LeaseUp! to connect clients with the best possible housing for their needs.

Amazon Personalize captures relevant information about available units of housing so case managers can recommend the best housing option to their clients in real time. By integrating this technology, the organization has been able to match over 600 individuals experiencing homelessness with housing—and reduce the time it takes to do so. Timing in these situations is often critical; a person who is ready to come in and get help one day may not return the next.

LeaseUp! aims to add 2,000 new units to its database over the next year to help even more people make it home. Bringing more existing apartments onto the platform and working more seamlessly with the landlords to list rental units are important steps in not just addressing homelessness but also ending it.

Enabling faster suicide intervention among veterans

Since 2012, **RallyPoint**, a social media platform designed for the broader US military community, has provided an online user experience focused on military service members, veterans, families, caregivers, and survivors to help them lead more successful and fulfilling lives. Among the millions of public discussions on the platform, a small percentage comes from members who share thoughts and behaviors about self-harm. The US Department of Veterans Affairs estimates that approximately 17 military veterans die by suicide each day—and RallyPoint has made it a priority to offer critical mental health resources and support to these men and women when they need it.

Developing a way to quickly and accurately sift through these high-risk public posts created by a small minority of RallyPoint users is a challenge.

In order to speed the discovery of these at-risk public posts, RallyPoint turned to the Amazon Machine Learning Solutions Lab and researchers at Harvard's Nock Lab. The Amazon Machine Learning Solutions Lab worked closely with RallyPoint to develop a machine learning model using Amazon SageMaker that can quickly analyze public posts on the RallyPoint platform and help determine whether there is an indication of self-harm. With the help of this machine learning model, RallyPoint has been able to successfully flag concerning posts quickly and accurately while reducing the amount of manual review needed to enable a potentially lifesaving intervention.

Ongoing, RallyPoint and Harvard will continue to further refine the model while evaluating the best content (e.g., mental health programs, hotlines, support groups) and preferred method to surface information to users. In the long term, the goal of the solution will be to augment the community engagement by RallyPoint member administrators that takes place on the platform today when the risk of self-harm in this audience is identified.

aws

Getting started with machine learning

Organizations are constantly working on innovative techniques to solve the most important issues the world faces today, making a profound and significant impact. We've seen what's possible, and AWS is committed to helping our customers bring their machine learning solutions to life.

Learn more ›