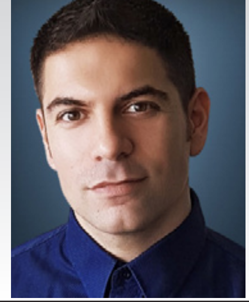
Lior Gazit, Memorial Sloan-Kettering Cancer Center



Trading Cancer for Data: Machine Learning for Cancer Diagnosis

Introduction

Artificial intelligence (AI) is penetrating everything we consume, create, and own. This is an industrial revolution that has been expanding for decades and will continue until it reins in every bit of redundant human labor. The first industrial revolution, starting in the 1700s, took manufacturing out of people's homes and into coordinated factories, setting the ground for efficiency and automation. Machines were the key. A worker with the aid of a machine introduced an unimaginable level of production. The second revolution, which introduced electrification, was based on the premise that the worker was just slowing the machine down. A machine with the aid of electrification introduced a significantly increased level of production. First, we had machines empower the work, then we designed them to precisely do the work, and now we are teaching them how to learn to do the work; they learn it, and then do it. These machines do the work better than we do. A stock trading algorithm replaces stock brokers, plural. It is created by a programmer who typically has a great understanding of logic and decision-making, but probably did not even take "Introduction to Microeconomics." It is then taking years' worth of financial data to derive optimal trading rules. A typical algorithmic trading (algo-trading) program may perform millions of trades a day while it also reads the news, adjusts to changes in the market, and spouts out updates to management. Power consumption is optimal, space consumption is minimal, the program doesn't take vacation days, and is oblivious to ego. If a human can be trained to spot a preferable trade, so can a machine. What else can machines be trained to do? Solve a mathematical equation, hit an enemy target, tell when to flush, signal before hitting a parked car's bumper? Simple enough! Recognize John Doe's face out of 10 hours of surveillance footage, drive a car, and spot a cancerous cell in an MRI?

Artificial Intelligence

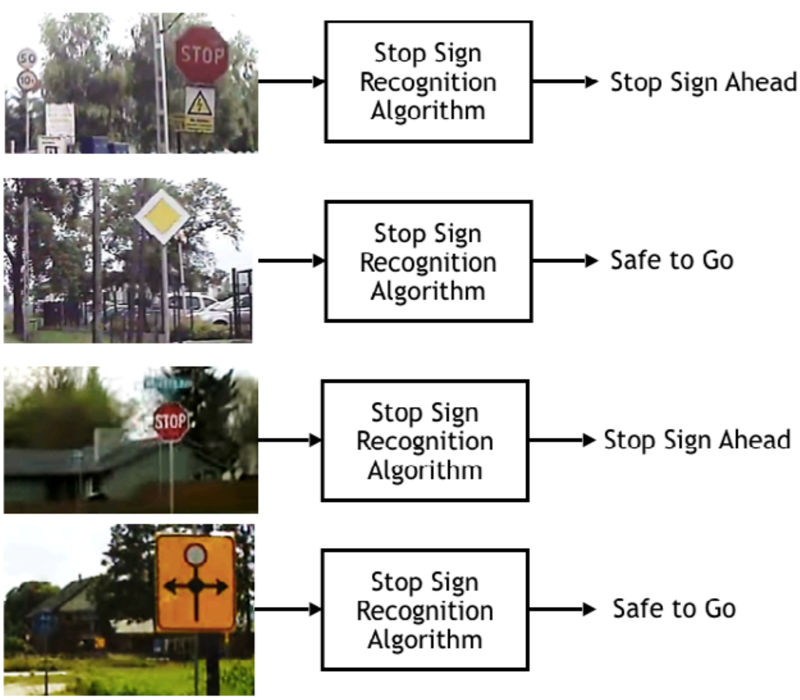
Artificial intelligence is defined as programmed decision-making. As this is a very broad definition, some argue that the label AI only applies to highly sophisticated applications, thus adding a contemporary aspect to it. Spellcheck, for instance, is a solution to a simple problem. When a word is entered, the program searches for it in the dictionary, and if it isn't found, it is flagged as a potential error for the user to correct. Autocomplete will take a partially entered sentence, search past entered sequences that started with identical text, and complete the entered text in the most probable (frequent) manner. Sentiment Analysis observes the text and derives what attitude the user is expressing, utilizing advanced algorithms in Natural Language Processing (NLP). Are they upset, disappointed, happy, confused? The answer will surely come in handy when trying to automatically prioritize the 10,000 customer service queries submitted in the last 24 hours, before the morning shift starts hitting the phones and keyboards.

Spellcheck is a remarkably simple algorithm. Given a preset dictionary, the decision for each word is deterministic and simple. With Autocomplete, the decision is statistical, there may be two completions that are plausible, but one appeared 1,000 times in the past, while the other appeared 50,000 times, thus deeming the latter more likely. The algorithm uses past entries to learn the probability of sequences and then applies the principle of Maximum Likelihood by selecting the most frequent. It learned on its own, no prior insight was programmed into it by the programmer. This entire notion is called machine learning (ML). The program uses past instances to learn about recurring patterns, an empirical study of cause and effect. So far so good! With Sentiment Analysis, we are entering fresh territory. The machine attempts to take a text sequence that it has probably never seen before, and classify it as expressing a specific attitude or emotion. It must break the text down to individual words, as we humans do, and key on those words that express opinions, emotions, moods, etc. Is the mere presence of the sequence "...thrilled when I opened the package..." enough to tell the machine whether this is a positive or negative query? Not quite. It must analyze the rest of the statement to validate. It is enough for that sequence to be preceded by "not" to turn it on its head. This calls for a thick sauce of analyzing keywords and their interactions while utilizing statistical learning of past examples so as to understand how this specific mixture comes together, and hopefully derive what attitude it is suggesting.

Machine Learning

Let's focus on machine learning. As conveyed in the previous section, machine learning drives the current AI wave we are experiencing. It is the force multiplier behind this new revolution. No longer is technology limited to what humans can perceive. Instead of specifying a set of rules for what action to take, the rules are set for how to learn. Inherently, the rules for action immediately follow.

Figure 1. Four examples of typical images from a dashboard camera. When driving, each would be analyzed by the recognition algorithm to identify stop signs.



There is a price, though. You see, in the first industrial revolution one would have to provide space for the workers to come and coordinate their work – factories. With the second, one would have to go through the trouble of building an entire assembly line, only after which it would be put into work, which is a big investment. With machine learning, something has to make up for the fact that you don't specify the solution, but are forcing the machine to learn. It needs something to learn from: data.

I'd like to welcome you to any ordinary graduate class in computer vision (CV). We need to design an algorithm for a digital camera that identifies stop signs in images. This would be a part of an AI of an autonomous car. The code's input is a single image at a time, and the output for each image is either "Stop Sign Ahead," or "Safe to Go."

Assuming we are taking a course in computer vision before taking our first class in machine learning, we would probably choose to teach the algorithm "what a stop sign looks like." You will notice, as depicted in Figure 2, for the machine, an image is a set of ordered numbers (pixel values), thus the analyses are limited to looking for sequences of values in the ordered structure that we portray as an image. So, to mine for a stop sign, we will define rules. It will be a long process filled with analyses of the image and questions that follow. We may scan the image for straight lines. We would check if any eight of them form an octagon. If so, we would try to discriminate it farther so to avoid calling just any octagon a stop sign: Are most of its inner pixels red? Are the inner pixels that aren't red, white? This would be a long piece of code. Eventually, assuming we missed nothing, we would have a solid algorithm. In today's era, we are not thrilled about such approaches. Two main caveats come to mind: R&D time and intellectual limitation.

How much time would it take us to design the stop sign algorithm? Don't be fooled into accounting only for the time it took us to set the rules for recognizing stop signs, we had actually spent quite a substantial amount of time in our lives learning what stop signs look like to begin with. To make this point clearer, what if the task was to recognize something less trivial? For instance, design an algorithm that takes a magnetic resonance image (MRI) of the spine and tells whether the subject suffers from a ruptured disk. If we went with the above approach of setting specific rules, our first step would be to start studying MRIs of the spine so to understand what our algorithm must look for. What do healthy spines look like? What do ruptured disks look like? Only after we grasped the details, would we be able to articulate them and suggest a set of concrete rules. We would probably spend a few weeks before we had something reasonable to start with.

Now let's discuss our own intellectual limitation. While we wouldn't have a hard time suggesting a set of rules for stop signs, how about for dogs? Recognizing a dog in a clear image is an unremarkably easy task for a human, even if their brain hasn't grown completely, such as in young children. But, could we specify a set of rules that would apply to an image and its pixels to identify a dog? Oh man! We would probably not get past the "Does it have four legs?" question. While we do have the intellectual ability to recognize a dog, we can't articulate a set of (pixel-based) rules that would apply to an image's pixels.

These aspects and challenges are not limited to computer vision problems. They exist in every machine-learning endeavor. If, for instance, LinkedIn or Facebook wants to suggest content or potential connections to their members, they would face similar problems. They would want to make a match based on a member's posts, interests, etc. But, articulating rules for how a member's posts relate to other content, what they may find interesting is as complicated as the examples of the dog and the spine.

Machines to the rescue! A machine-learning algorithm will do the job. Quickly too! How would it work? Since I started with the notion of computer vision, I'll stick with it. We would provide the machine with a big enough variety of images. How big? It needs to be big enough for it to get an idea of what a positive input looks like, and a negative one. Much like we would research for ourselves to learn what a ruptured disk looks like, we would have two sets of images to train on, those labeled a "Yes" and those labeled a "No." In machine learning, this is called a labeled data set. Next, a specific machine learning algorithm of the developer's choice will iterate over all the "Yes" images and will find pattern-like consistencies between the pixels. It will then do the same for the "No" images. It will see what consistencies exist in the "Yes" images, but don't also exist in the "No" images. If it finds such consistencies, they become the rules for classification. Voila! Remember the ruptured disk and the R&D time? Well, what would take a human weeks, a machine can do in mere seconds. Remember the dog and our intellectual limitation to articulate pixel-based rules? The machine does not learn first what a dog is and then attempt to explain it in pixel form (like a human would); it directly looks for pixel patterns that are typical of dogs. *C'est tout!*

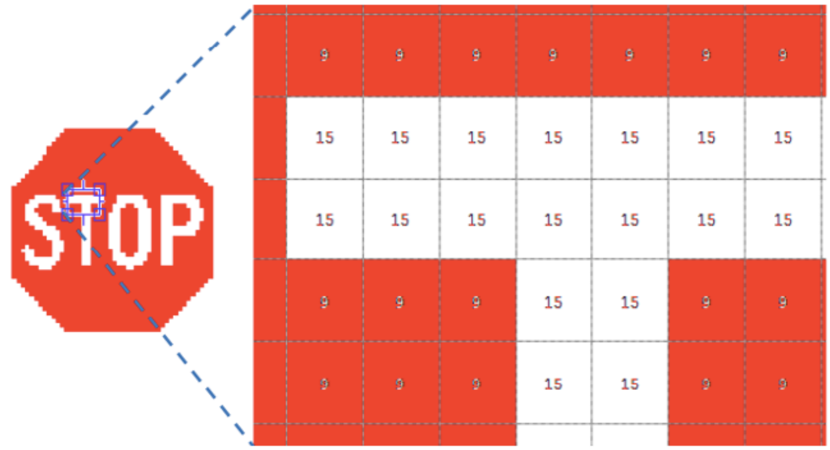


Figure 2. A look into Computer Vision. On the left is what humans see; on the right is what the algorithm "sees." It perceives the image as an array of values. Each pixel has a position and a value. Note that the values in the right image are of the intensity of the brightness, presented for simplicity. In practice, each pixel has not one but three values, typically one for each of the three colors, red, green and blue.

My Pancreas Work

As a data scientist on the Strategy Analytics team in Memorial Sloan-Kettering Cancer Center, I'm like a kid at Six Flags. We utilize some of the most cutting-edge algorithms and we deal with a lot of data, usually two kinds: clinical and operational. Clinical data would be on the patient level, featuring disease condition, physical characteristics, medical history, etc. Operational data would present hospital processes like occupancy, flow of chemotherapy administration, and so on. A typical project would result in a prediction or classification algorithm. The purpose of the algorithm's result may be operational, and the data used is a mix of operational and clinical attributes. In Strategy Analytics, we don't study cancer (what I typically call the "micro" of cancer), we learn about cancer and use the knowledge for a bigger scope such as treatment operations (I call it the "macro" of cancer). We look for connections between hormones, cancer stages, radiation therapy, tumor size, even finance, admission time, nursing scheduling, and more. These connections, often hidden and non-trivial, are exploited to learn and improve.

As a break from routine, we are encouraged to think about ideas of our own. Given my interest in applying machine learning to computer vision applications, I reached out to a researcher in our institution, Dr. Amber Simpson, and allocated a certain amount of time to work with her team on identifying early risk of pancreatic cancer using medical imaging. Pancreatic cancer is notorious for its high mortality rate, mostly due to late discovery. Computer vision to the rescue!

We followed a research paper that spoke about how one could identify discolorations in cysts in the pancreas. We figured, if a human can understand how to find these discolorations, so can a machine…but faster. We first started by collecting images of the pancreas from various patients, then creating a labeled set by letting experts diagnose whether each person is a "Yes" or a "No" for pancreatic cancer risk. Next we made a mixture of the two AI approaches that I described above – self-defined rules, and machine learning-derived rules. The first imitated what the mentioned paper taught us. We told the machine to look at a cluster of pixels that are brighter than the surrounding ones. Then, given these bright clusters, we used machine learning to find what clusters are typical of "Yes" labels, and which are typical of "No" labels. This algorithm produced results that supported the initial paper, thus functioning as a non-invasive contribution to early pancreatic cancer detection.[1]

The future is happening and machines are leading the way for us. They tell us what to buy, how to get to a destination, what to read, and who to watch.

Additional Examples

I have touched on NLP and computer vision. The typical third musketeer is voice recognition. One example application is the authentication of the speaker for security purposes. Some banks create a layer of security for their automated phone services by asking the caller to recite a predetermined line. The algorithm crosses the fresh recording with previous acquaintance with the account's owner and determines whether they are indeed the current caller.

Seizure detection in epileptic patients is a non-invasive way to alert others of an occurring seizure. Such seizures can occur without warning and their consequences may be physically harmful to the patient, and, in extreme cases, cause death. While the prevention of such seizures is still unsolved, there is great value in alerting nearby caregivers, for instance, when the patient is at home, in bed sleeping, and naturally isn't being watched. The non-invasive method is driven by machine learning. A set of sensors are constantly collecting electroencephalogram (EEG) signals through the scalp, and when a pre-learned discriminating pattern occurs, it alerts a caregiver.

Recommendation systems and targeted advertising are two applications that are very similar in their AI back ends. An on-demand TV content provider will utilize a recommendation system for learning the interest of every individual viewer and then suggest the content that this viewer would appreciate. The learning is based on this viewer's past choices. Every watched-content is a "Yes" label in the labeled data set, which is continually growing. The next step is to derive characteristics that this content falls under, and recommend to the viewer similar new content. A characteristic, or rule, could be "Thriller, released in the last five years, starring a female Oscar nominee." It is highly valuable for the provider to be able to match the relevant part of its content to each individual. Targeted advertising is similar. It would track your purchases online and will recommend the next product you are likely to want. One way to do that is to follow the things you bought recently, look for other people who bought the same things in the past, see what those people bought next, and recommend to you that next product. This is very similar to Autocomplete, right? You would order paint and it would recommend a brush. Order a bathing suit and it would recommend sunscreen. Order a wing suit for base-jumping and it would recommend a probate lawyer.

Computer Vision in Medical Imaging for Cancer Diagnosis

When I'm discussing the various aspects and applications of AI and machine learning, I get excited. It is the forefront of an ongoing industrial and social revolution. It is evolving so rapidly that within each decade new social and technological norms are formed and we must either keep up or risk being left behind.

Since I'm in the field of healthcare, and cancer, in particular, I mark these realms as the "next to benefit" from this revolution. Healthcare is no sand box; it is very conservative, and for the right reasons. We had to let many other sectors experiment with AI and prove it ripe before applying its force to healthcare. As a preliminary step, it is safest to apply AI as an aid, a force multiplayer, serving the physician, nurse, technician, or other healthcare professional. An even safer approach is to start with diagnosis, rather than treatment.

Cancer imaging diagnosis is being transformed by the AI revolution. The need is very clear: better diagnoses will dramatically reduce cancer deaths. The most effective way to cure cancer is to catch it in its early stage, much like a cavity in your tooth. Machines will learn massive image sets and extract pixel-based rules to provide radiologists with an uncanny ability to generate many more diagnoses with better precision. The diagnostic ability of AI does not fully overlap with a radiologist's, thus yielding wonderful synergies. It has advantages that a radiologist wouldn't have, but it lacks critical insights the radiologist possesses; much like a pilot and his autopilot. For example, AI can learn texture patterns that are typical of cancerous cells by focusing specifically on the differences between textures in healthy cells and in cancerous ones. A radiologist may miss the fine, pixel-level details. Moreover, radiologists may only take in so much information when observing the patterns, as they are only human. An advantage that the radiologist has is the qualitative understanding and clinical meanings of what is in the image. It also translates to the ability to extrapolate and make sense of something they are seeing for the first time. This is not the case with AI. In most machine-learning cases, if some notion was not presented clearly and frequently in the labeled data set, AI will not understand where it stemmed from and whether it is indicative of cancer. It learns from examples.

The outcome of this specific revolution will be spectacular. Cancer screening will be quicker and more common, which means we will all have easier and cheaper access to cancer-targeted imaging services. Monitoring cancer health will be as common as maintaining dental health. Nowadays, we take for granted our ability to avoid a root canal by having twice-a-year checkups at the dentist. Monitoring cancer health will be similar to that process. Early stages will be the new cavity and later stages will be the new root canal.

Conclusion

The future is happening and machines are leading the way for us. They tell us what to buy, how to get to a destination, what to read, and who to watch. This is a beautiful manifestation of our ongoing yearning to improve and excel. There is no dark side to this. Machines will not turn on us, replace us here on earth, or rule us. They are our force multiplier. While it's easy to get lost in this AI tsunami, remember that machines are only doing what we design them to do. In principle, this revolution will not do something that the past two revolutions didn't do. Yes, it will require adapting and learning. Yes, it will make some services and professions obsolete, thus creating changes in the job markets, forcing some to adjust while others will become irrelevant. Take librarians for instance. As the library's documentation of its inventory went digital, the librarian had to adapt and get familiarized with the new computer-based interface. Things are quite similar for the travel agent, too. Today's youth will have a lot to think about when they choose a profession. Choosing to be a mail carrier 60 years ago is one thing, choosing it 20 years ago is another. What would you think about an 18-year-old banking on being a mail carrier today?

These are opportunistic times and the means necessary to seize this opportunity do not discriminate. You don't need to belong to a specific gender or race, and your physical strength is irrelevant. This revolution is for everyone. Everyone can read, learn, and adapt. It will serve us better, keep us safer, and make us healthier. *Viva la revolución!*

Endnotes

[1] Lior Gazit, Jayasree Chakraborty, Liana Langdon-Embry, Richard K. G. Do, Amber L. Simpson, Quantification of CT images for the classification of high- and low-risk pancreatic cysts (http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=2608967).

About the Author

Lior Gazit is a data scientist in Strategy-Analytics at Memorial Sloan-Kettering Cancer Center, a world leader in cancer care, research and education. The Strategy-Analytics team is a driving force utilizing state-of-the-art analytical tools and algorithms, and was honored the 2012 INFORMS prize. He designs and carries innovational operations driven by predictive modeling and machine-learning programing. He was previously the senior algorithms developer for the Active Noise Control groundbreaker Silentium. He holds a BSc and an MSc in computer engineering and has authored several scientific publications. He is available at ***www.linkedin.com/in/liorgazit***.