Career Education REVIEW

The Career College Information Source

What is Machine Learning?

Harry Brelsford, Director of Business Development, LeadScorz

What is Machine Learning?

Harry Brelsford, Director of Business Development, LeadScorz

The short answer to the question of "what is machine learning?" is this: Netflix.

When I describe the concept of machine learning, I simply ask if the individual uses Netflix for entertainment. The answer is always yes. So how does Netflix "just know" what movies you might like or enjoy? It's called Machine Learning.

Technically speaking, machine learning is all about testing, testing, testing. It starts with training data (your historic purchase behavior at Netflix) and then it applies predictions to what you might like. When you make a predicted selection, your answer is so noted. And so are your declines. Machine learning is a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. And it all happens automatically. In fact, the automation minimizes the commitment of resources such as developers and administrators after the initial stages.

As an aside the Netflix example is so popular in the big data community that the CalTech University "Machine Learning Video Library" speaks towards a Netflix competition!

Don't fret or be embarrassed if you don't know what machine learning is. By the end of this article you will. This professional and scientific field is less than 10 years old. Consider this a fast-

> **Machine learning explores the study and construction of algorithms that can learn from and make predictions on data.**

track primer to walk into your next meeting conversant in machine learning.

Digging deeper

In order to consider machine learning, three elements must be present.

- A pattern exists. Rating movies on Netflix has a pattern.
- We cannot pin it down mathematically. That's why we are using machine learning to "learn" from data.
- We have to have data. Machine learning always starts with data.

Your goal is to extract knowledge or insight from that data. There is some amount of automation in this process. Instead of trying to divine wisdom manually, you are applying some algorithm to help answer the questions being presented. And that

automation would be a computer! Machine learning is not just machines. This journey requires you to make many smart human decisions as part of the process; it's not only machines.

> **This journey requires you to make many smart human decisions as part of the process; it's not only machines.**

There are two main categories of machine learning.

- **Supervised learning.** Making predictions using data. This is also known as predictive modeling. I recently took an online course where this was presented as "ham" or "spam." The scenario was training a spam filter to review incoming email and determine if it was non-spam ("ham") and should be placed in your Inbox. You know the rest of the story. The other type of email was spam and placed in the Trash folder. This goal is to predict a specific outcome. Another view is that the primary goal of supervised learning is to build a model that "generalizes." It accurately predicts the future rather than the past.

- **Unsupervised learning**. This is the process of extracting structure from data and learning how to best present data. For example, you could segment students (and potential students) into clusters that exhibit similar behaviors. You might find three groups of students displaying similar behaviors within each group (for example the two-year degree seeking students cluster might have renter status as a predictive variable; Masters students might have a homeownership tendency). But this variable is dissimilar between the two groups. Most importantly, there is no right answer!

How does machine learning "work?"

First you train a machine learning model with training data (also known as labeled data). This is data that has been labeled with the outcome (is the email ham or spam?). This is called model training because the model is learning the relationship between the attributes of data and its outcome. Such a relationship, using our email example, might include length of text strings, specific text in an email, length of the message, etc.[1]

Second, you make a prediction on new data for which the label is unknown. It's actually easier than it sounds. A new piece of email arrives.

HARRY BRELSFORD is a senior professional in the technology field directing the business development efforts at Seattle-based LeadScorz, a predictive analytics firm. His raison d'être dating back to the Apple II+ accessing a VAX remotely during college was and is to marry business and technology. He has worked in a variety of industry capacities including technology and management consulting, starting/owning/operating a small business and serving enterprise organizations (Microsoft). His academic credential including a Masters of Business Administration (MBA) from the University of Denver in Project Management. He has 12+ years of college-level adjunct teaching experience at Seattle Pacific University, Central Washington University and for-profit colleges including Alaska Computer Institute and Griffin College. He has authored 19+ books on specific technologies including client server, VoIP, eCommerce and business topics including consulting guides. He has published hundreds of blogs and articles. Learn more about Harry Brelsford here: https://www.linkedin.com/in/harryb.

Contact Information:
Harry Brelsford
Director of Business Development
LeadScorz
Phone: 206-915-3072
Email: harry.brelsford@leadscorz.com

The "machine" hasn't seen this piece of email before. It predicts that this piece of email is "ham" not spam based on everything it knows.

Bottom line. The model is learning from past examples using inputs and outputs and then applying what it has learned to future inputs in order to predict future outputs. Does that make sense?

This isn't your father's data analysis!

For a hint at the complexity of data modeling, these questions provide insights.

- How do I choose which attributes of my data to include in the model?
- How do I choose which model to use?
- How do I optimize this model for best performance?
- How do I ensure that I'm building a model that will generalize to unseen data?
- Can I estimate how well my model is likely to perform on unseen data?

Of course your mileage will vary in answering those questions as every situation is unique. And the mantra of test, test and test applies here.

Writing about learning for an audience of professional educators would, at first blush, appear to be almost elementary if not offensive. But a quick review of the basics always makes sense. The components of learning in the classroom and inside the machine are the same: input, output and target function. Basically the function statement creates the output based on the input. In the case of education, the input is the student application. The output is a score that determines whether the student would enroll and be successful. The function is to score the student lead. Now multiply that by thousands of inputs (even billions

according to some data providers). This activity results in a hypothesis that is the learning algorithm for machine learning. Multiple algorithms would be a hypothesis set. Think of this as having lead scoring algorithms for students who are in three different degree categories: associate, bachelor's, master's. It's that simple.[2]

But going one layer deeper, the bottom line is a threshold value. If the student lead score exceeds a minimum, the lead is approved and followed-up with an admissions representative. This can be charted as a scatter diagram (a common way to present machine learning).

Machine learning for dummies

Spoiler alert. Such a book doesn't exist, likely because you'd have to be a smart dummy to engage in a machine learning motion. But I researched how to simplify the machine learning conversation for mere mortals and found a few resources. One resource was a LinkedIn SlideShare called **Machine Learning for Dummies** by Venkata Reddy Konasani. As you can see below in Figure 1, the notion of grouping objects is a machine learning concept. The machine has to learn what animal groups belong together.

Figure 1: A simplified view of the machine learning premise.
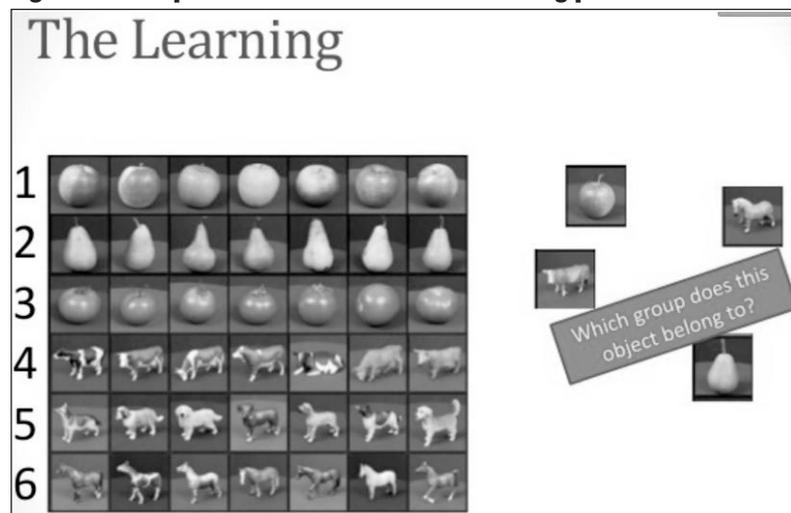
Konasani goes on to articulate the following thoughts. How does our brain process the images and make groupings happen (via learning)? Afterwards, the brain can look at a new image and compare with the groups classified to get the closest match (this is Classification). Because we're talking about machines, performing the same operation occurs via machine learning. "We write programs for learning and then classification, this is nothing but machine learning," Konasani shared.[3]

The idea is this. Machine learning inside an actual university results in higher Customer Satisfaction Scores (CAST).

Compare to rules-based learning

This article would be remiss if it didn't do a quick compare and contrast with rules-based predictive modeling. Just as the name implies, rules-based is based on a static set of rules to predict an outcome. But it's not really dynamic machine learning. The "black box" only knows what it knows. One argument is that rules-based predictive algorithms are a few nano seconds faster than machine learning-based predictive algorithms. True that but you give up a lot of power. The answer to this challenge is the need for speed.

A killer combination

Cray Computers is just a few blocks away from our Seattle office. Cray is best known for its super computers. With machine learning, it's horsepower is now back in fashion. That's because it brings the strength needed to create the killer combination of machine learning + big data. And given that data providers are now proclaiming incorporating over 1 trillion attributes into machine learning-based predictive lead scoring[4], you'll need a lot of horsepower to quickly crunch through that amount of data.

An EDU example

It's all too easy to view big data and machine learning as dehumanizing personal relationships. But I assert machine learning is an enabling technology and my evidence is a mid-sized online university. The digital marketing executive told me "…though we use high technology to gain prospective students, we are truly student centric." The gentleman's point is this. The online university is benefiting students rather than blasting them with messaging. "An example is we use machine learning to better understand our audiences then we isolate this like-group to deliver ads that are relevant to that person or group of people. So we do not target those seeking to get a trade certificate or only want an AA degree."

"On the other side you can say something about using machine learning to better improve internal processes and prospect experience. An example may be that machine learning can help schools isolate those who are in need to talk to someone now about starting school (a high lead score) versus those who just want information (a low lead score) and do not want to be bothered."

The idea is this. Machine learning inside an actual university results in higher Customer Satisfaction Scores (CAST). Furthermore, there is an interesting conversation hosted by IBM Research titled *"The classroom will learn you."* (Here: http://www.research.ibm.com/cognitive-computing/machine-learning-applications/decision-support-education.shtml#fbid=YRjywNwjQki). That discusses how cognitive systems will provide decision support for

teachers. This is a slightly different narrative then we've had in the article on machine learning. Steeped in the academic-side, *"The classroom will learn you."* Talks about classroom that will have technology tools that understand each individual student and create custom curriculums in the cloud. Machine learning meets Massive Open Online Courses (MOOCs).

Or in Chalapathy Neti's (IBM Research) own words "IBM envisions educational institutions adopting cloud-based cognitive systems to collect and analyze all of this data over a long period of time — creating longitudinal student records that would give teachers the information they need to provide personalized learning experiences for their students. These systems would also help teachers identify students who are most at risk, why they are struggling, as well as insight into the interventions needed to overcome those challenges."

"The system could also couple a student's goals and interests with data on their learning styles so that teachers can determine what type of content to give the student, and the best way to present it. Imagine an eighth grader who dreams of working in finance but struggles with quadratic and linear equations. The teacher would use this cognitive system to find out the students learning style and develop a plan that addresses their knowledge gaps." (Neti, 2013)

Resources
1. Android Authority "What is Machine Learning" Gary Sims https://youtu.be/WXHM_i-fgGo http://www.androidauthority.com/what-is-machine-learning-621659/
2. Lecture 01 – The Learning Problem, Professor Yaser Abu-Mostfa, Caltech. https://youtu.be/mbyG85GZ0PI
3. Machine Learning for Dummies, Venkata Reddy Konasani. http://www.slideshare.net/21_venkat/s-32981502)
4. Versium : Launches First Automated Predictive Lead Score Service That Incorporates Over 1 Trillion LifeData(R) Consumer and Business Attributes (4-Traders; Janaury 27, 2016) http://www.4-traders.com/news/Versium-Launches-First-Automated-Predictive-Lead-Score-Service-That-Incorporates-Over-1-Trillion-L--21752239/

Career Education REVIEW

Career Education REVIEW

GET INFORMATION THAT MATTERS

Are you spending valuable time sifting through emails and media sources searching for information that will help you and your school be more successful? Career Education Review (CER) is an easy way to staying informed about the career education sector.

CER keeps career education leaders and managers up-to-date on best practices, emerging technologies, enrollment and new program ideas, news and events. Today, top leaders, like you, subscribe to CER as a way to stay informed and competitive.

WHY SUBSCRIBE

"Career Education Review is the real thing – information on career schools and programs assembled and written by individuals who have lived and know the sector. It is a vital source of important information anyone involved in career education needs to have."

-Henry Herzing, Ph.D.(hc),MBA, MSEE, President, Herzing College, Co-Founder and Chancellor Emeritus, Herzing University

SUBSCRIBE TODAY AND TAKE ADVANTAGE OF:

- ✔ Original articles written by career education leaders

- ✔ Exclusive interviews with key higher education players

- ✔ Articles and best practices in admissions, academics, faculty development, retention, marketing, alumni, career services, placement etc.

- ✔ Articles and webinars by education attorneys on compliance, legal issues, government regulations

- ✔ Monthly Washington news brief

**Subscribe online at www.careereducationreview.net/subscribe/
OR contact Jenny Faubert at 920-264-0199 or jfaubert@careereducationreview.net**

Career Education Review
2221 South Webster Avenue, Suite A #255
Green Bay, WI 54301
Phone: 920-264-0199 | Fax: 920-659-7797
Email: jfaubert@careereducationreview.net | Website: www.careereducationreview.net

Career Education
REVIEW

✓ **LEADER INSIGHTS**

✓ **FEEDBACK FROM CAREER EDUCATION EXPERTS**

✓ **HIGH QUALITY CONTENT**

✓ **FREE WEBINARS**

WHY SUBSCRIBE

"Career Education Review serves as the leading publication to get useful news as well as powerful resources that our campus leaders can use to optimize operational effectiveness and more importantly, student success. This is the reason we feel it's critical all of our campus directors receive CER!" - Jeanne Herrmann, Chief Operating Officer, Globe University/Minnesota School of Business

YES!

Sign us up for an annual subscription to the Career Education Review at a cost of $329.00 Scan and email completed form to jfaubert@careereducationreview.net. The form can also be faxed or mailed to the information below. Or subscribe online at www.careereducationreview.net/subscribe

Name: _____ Title: _____

School/Organization: _____

Shipping Address: _____

City: _____ State: _____ Zip: _____

Phone: _____ Email: _____

Method of Payment: _____Check _____Credit Card Total: _____

Charge to: _____Visa _____Mastercard _____American Express

Card Holder Name: _____

Card Number: _____

Experation Date: _____ Security Code: _____

Billing Address (if different than shipping): _____

Signature: _____

Career Education Review
2221 South Webster Avenue, Suite A #255
Green Bay, WI 54301
Phone: 920-264-0199 | Fax: 920-659-7797
Email: jfaubert@careereducationreview.net | Website: www.careereducationreview.net