

Lecture 4

Machine Learning and the Learning Algorithms

Rahul Bhattacharya

Machine Learning: The New Science

Part I

What is this New Science of Machine Learning?

Putting Machine Learning in the context of Artificial Intelligence (AI)

 Introduction, Lecture 0

Part II

Foundational Concepts of Machine Learning

 Lectures 1 – 12.

Part III

Solving Real Life Problems with Machine Learning

 Lectures 13 – 25.

Part IV

Machine Learning: The Changing Landscape of Knowledge

(Near and the Not So Distant Future)

 Lectures 26 – 27.

Let's Ask the Question Again: What is Machine Learning?

Arthur Samuel, creator of one of the world's first self-learning programs, coined the term “machine learning” in 1959 while working at IBM. He defined machine learning as a “field of study that gives computers the ability to learn without being explicitly programmed.” This means that instead of writing a set of instructions (algorithm) via a formal computer program to tell the computer to do something or recognize something, like an image or a pattern, the machine could simply be shown that “something”, i.e. data, over and over again until it would be able to “learn” on its own from that data.

Machine learning is the science of learning and abstracting knowledge from data. Such learning by a computer happens when a lot of data – big data – is presented to the computer along with a desired result. If I want the computer to recognize a tree or my grandmother then I simply show enough images of a tree or my grandmother telling the machine which one is which and the computer is able to recognize them both. If I want the computer to diagnose lymphoma or any other type of cancer and suggest a treatment then I feed enough data about cancer patients together with what treatment was given to them in the past and the outcome of those treatments and the computer figures out how to diagnose cancer in a new patient and what treatment to offer to a newly diagnosed patient. Anything and everything that we can learn or have learned so far, the computer can also do so on its own provided it has enough examples of it.

In short, machine learning is learning from data. It is the first step towards making the computer intelligent. We just keep training the computer by showing enough examples, enough data, and the computer will do all the rest because it would have gained all the knowledge needed to solve problems.

Machine learning is a way of getting rid of computer programming (and, the computer programmer). The computer does its own programming, writes its own code and does not need any human intervention. All the human being has to do is to train the computer with enough examples (data). Once that training is done, the computer's learning is complete.

Computationally speaking, the goal of machine learning is to generalize the training experience that comes with data (or examples) and generate a hypothesis (output) that estimates the target function.

In the very early days of machine learning, in the late 1940s and early 1950s, when the field was still dominated by the quest for artificial intelligence there was great hope that computers could do all that a human being can do. Frank Rosenblatt, the inventor of perceptron, one of the first machine learning algorithms, believed that multilayer feedforward networks (“perceptrons”) could enable computers to “walk, talk, see, write, reproduce itself and be conscious of its existence.” He and other early researchers in the field wanted to give the computers a brain – a kind of neural thinking system – that would resemble a human brain and the rest would automatically follow.

Rosenblatt's dream hasn't come to fruition yet and perhaps we are still many decades away from a intelligent, human-like computer but his work and the work done by his contemporaries more than six decades ago laid the foundation for machine learning.

The Learning Algorithm: Starting Point for Machine Learning

We have talked about algorithms in the previous lectures. In machine learning also we have algorithms. Within the context of machine learning, i.e. the way a computer learns from data, the algorithm signifies the way a computer learns; algorithm is the process in which a computer extracts or derives knowledge from data. That is why these are called **learning algorithms**.

In machine learning, we talk about a different kind of algorithm, the learning algorithm. It's different from the ones that we talked about in the previous lectures where an algorithm is a precursor to a computer program and is designed as a set of specific instructions given in a particular order for the computer to execute a task. But as we have seen machine learning turns the process around. In machine learning a computer writes its own program. In conventional computer programming, we provide input data and an algorithm – the kind of algorithm that we have talked about thus far – via a computer program and the computer executes the tasks and generates an output. It's a pure computational task. In machine learning, on the other hand, we feed the data and the output (the desired result) into the computer and the computer given out the algorithm that governs the given data.

The way that happens, i.e. the way the computer figures out – learns – from the data and the given output what kind of knowledge is embedded in the data and what algorithm would have turned the input into the output is via a special kind of algorithm, known as the learning algorithm.

A learning algorithm – sometimes called “learners” – is an algorithm that make other algorithms.

Pedro Domingos has talked about a master learning algorithm that would be the mother of all algorithms on earth. In his seminal book *The Master Algorithm*, he states the central theme of the book as “All knowledge – past, present and future – can be derived from data by a single universal learning algorithm.”

We have to keep this in mind: **a machine learning algorithm or a “learning algorithm” is a very special kind of algorithm that can generate all other algorithms inside a computer**. This learning algorithm is precisely how a computer learns from data. A learning algorithm forms the very foundation of machine learning.

References:

- Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev.* 1958;65(6):386–408.
- Domingos, Pedro, *The Master Algorithm*, Basic Books, 2015
- Khanna, Rahul, Mariette Awad, *Efficient Learning Machines, Theories, Concepts, and Applications for Engineers and System Designers*, Apress Open, 2015

Machine Learning: The New Science is an **online course** developed by Risk Latte AI to increase awareness and a very basic understanding of the field of Machine Learning, the new science that is taking our world by storm. In the next ten years, the world that we see around us would have changed completely, thanks to machine learning. In the next 20 to 30 years at most we, the human beings, would be in a totally uncharted territory, once again thanks to machine learning and the much bigger and related discipline of artificial intelligence (AI). In more ways than one, machine learning is the stepping stone to understanding artificial intelligence. Machine learning is a very important sub-field of artificial intelligence.

This online course prepares a person to start off with a more formal **Machine Learning 101** course. This course contains very easy to understand lectures and simple tutorials implemented in Microsoft Excel™ to illustrate real life problem solving using machine learning algorithms.

- This course is primarily targeted towards high school, college and university students and middle to senior level working professionals and executives around the world who have a very basic knowledge of mathematics and a very limited or even **no knowledge** of computer programming.
- For certain segments of our target audience, such as certain specific groups of high school, college and university students in India and other developing countries this online course costs only US\$8.00 (USD Eight only).
- This course comes with on-demand and customized onsite, classroom lectures and workshops at a minimal cost.
- A lot of the lectures and many Excel spreadsheets containing worked examples are **FREE** and can be easily downloaded from our website. These convey the general flavor of the course.

Risk Latte AI is a unit of **Risk Latte Americas Inc.**, a Montreal, Canada incorporated company, that is in the business of developing machine learning algorithms and software for banks, financial institutions, healthcare and the education industry as well as developing gamification and social learning applications for the general public.