



# THE PREDICTION GAME: THE EVER-EVOLVING SCIENCE OF MACHINE LEARNING

**DR GERALD FRIEDLAND, BASED AT THE UNIVERSITY OF CALIFORNIA, BERKELEY, INVESTIGATES THE SCIENCE THAT UNDERPINS THE FAST-CHANGING TECHNIQUE OF MACHINE LEARNING. IT'S A FIELD THAT HAS SEEN RAPID GROWTH IN RECENT YEARS AS ONE OF THE KEY TECHNOLOGIES IN ARTIFICIAL INTELLIGENCE. ITS FUTURE POTENTIAL IS ENORMOUS – IT COULD EVEN TRANSFORM SCIENCE ITSELF**

When you speak into a device to ask for a weather forecast, a football score or a music track, Siri or Google or Alexa usually understand your words and try to provide the best answer. You may not even remember the first time you ever tried this. And at the time, you probably didn't think how amazing it was that a device could recognise your voice, even though it had never heard it before.

The computing method behind voice recognition, and many other applications, is called machine learning. It's one of the key technologies of artificial intelligence, which aims to give computers and robots the ability to perform tasks that previously only humans could do.

While machine learning has been improving in leaps and bounds in recent years, there's still vast potential for improvement. This is the focus of research activity by Dr Gerald Friedland, Adjunct Assistant Professor at University of California Berkeley and co-founder of a technology company called Brainome, Inc. According to Gerald, the technology could even change the way scientists work.

## WHAT IS THE SCIENCE BEHIND MACHINE LEARNING?

Machine learning is a computing method that is used to infer a function, often called a model, from observations. As an example, you could make observations of various aspects of used cars, such as the manufacturer, model, age, size and condition. You could also list the price that each car was sold for. Do this for a lot of cars and you'll have collected a big set of data.

You could then 'train' a computer to create a model of used car sales based on the previous sales figures. This is machine learning. Given various bits of information about a car, the model could be used to predict that, for example, a 2018 Ford Fiesta that was in good condition would sell for £8,350. It would only be a prediction, of course, and many factors determine how accurate the prediction would be. To date, machine learning has been dominated by statistics, a branch of mathematics. Maths is a very powerful tool used by scientists to describe things precisely.

## HOW DOES MACHINE LEARNING WORK?

When experiments are performed, a table records all the factors that go into it in one row. There's another set of rows for the observations. As we vary the experiment and do it again, a new row is created each time.

Remember Isaac Newton? He is said to have discovered gravity when an apple fell on his head. And not only that, he also created a model of how energy is related to gravity. It's fun to imagine him dropping the apple on his head from different heights and noting how much more painful it was when dropped from higher up! (He didn't actually do this, so don't try it at home!)

The result of this imaginary experiment would be a table where the first column records height and the second column records the 'pain level'. Newton trained his brain on this table to figure out that energy (pain) is proportional to height.

In machine learning, the first step in this process, creating a table, is memorisation. The





## DR GERALD FRIEDLAND

Adjunct Assistant Professor at University of California Berkeley and CTO of Brainome, Inc.

.....

## FIELD OF RESEARCH

Computer Science

.....

## RESEARCH PROJECT

Using science and engineering methods to improve the design of neural networks

second step - creating a formula (a model) to describe the table and all experimental outcomes of the same setup in the future - is called generalisation.

In computer science, we want a machine learning model to create the formula and be able to predict future experiments. In other words, we want machine learning to generalise.

### WHAT ARE THE CHALLENGES THAT NEED TO BE OVERCOME?

Often, all a machine learner does is memorise the table of observations as it is not (yet) able to generalise. This is called 'overfitting' and it's one of the most recognised problems in machine learning. However, it is often misunderstood. When the experimental set-up presented to the machine learner in training is too complex for the amount of training data, then all the machine learner can do is overfit - this is memorising. Memorisation is the first step. But it's a worst-case scenario when it comes to generalisation. Such a model would not be able to make future predictions and would therefore be useless.

### WHAT HAVE BEEN THE OUTCOMES OF GERALD'S RESEARCH?

Gerald has been working on measurements for the design and training of neural networks. These mimic the way animals' (including humans') brains work. He treats each neuron (brain cell) like an electrical element in a circuit. By doing this, he can come up with rules similar to Ohm's law that describe how neurons behave when they are combined, just like resistors in a circuit. While the units of an electrical circuit are Volt, Ampere, and Ohm, the units in an information network are bits. All of these units can be translated back into physical measurements. One result is the ability to solve problems using neural networks that are much smaller than it's been possible to use so far. Gerald is currently working on building a speech recogniser with a dozen neurons.

Ultimately, Gerald's research involves applying physics to the world of information. He wants to make machine learning as easy as possible. His ultimate goal is to make it 'plug and play' so that it's as simple to use as selecting rows and columns in a Microsoft Excel spreadsheet and pressing a button.

### HOW WILL MACHINE LEARNING BE APPLIED IN THE FUTURE?

Machine learning will become easier to use, and used in even more AI applications in the future. But Gerald has an even more futuristic vision. He thinks that machine learning will be able to replace part of the job previously done by scientists. Great thinkers of the past, like Einstein, came up with equations as a result of a process inside their brains. Einstein generalised a table of observations to come up with his famous formula  $E=mc^2$ . He did it inside his head, and on paper. But in future, the deduction of scientific formulae will be done by computer. As a result, scientists won't need to come up with formulae. Instead, they will have the job of posing the right questions and collecting the data in a clean way. They will also be needed to explain the results produced by the machine to a broader audience.

# ABOUT MACHINE LEARNING

The applications of machine learning have mushroomed in recent years. Anywhere that data is collected, machine learning can be employed to predict future behaviour.

When you buy from an online store, data is stored of how you found the website, and which products you clicked on. The patterns of behaviour can then be used to identify similar people and send them targeted adverts.

Similarly, machine learning predicts the TV shows you'll want to see on Netflix, and whose posts you'll be interested in seeing on Facebook and other social media platforms. Machine learning is also at work in identifying fraudulent activity on your bank account and in satellite navigation.

## WHAT DOES THE FUTURE HOLD FOR MACHINE LEARNING?

The future of machine learning will be even more exciting than the present. It

will be the technology behind driverless cars, augmented reality, and improved voice recognition and computer interfaces. It will make computer vision an everyday reality, allowing robots to move freely, and safely, around because they will be able to understand what they are seeing through their cameras. And augmented reality, which will overlay what we see through smart glasses or on our phone screens, will understand which information to give us – and when. If you see a friend on the street, augmented reality will remind you if it's their birthday.

## WHAT SHOULD I STUDY TO GET INTO MACHINE LEARNING?

Machine learning is a branch of computer science, but future advances will take it out of the world of mathematics and statistics and into the real world. That's why it's a good idea to study physics at school, while you also learn how to programme computers and participate in open-source projects.

For example, you can experiment with simple hardware such as Arduino and learn the basics of the Unix command-line user interface, using Linux, for example. When you get to college, relevant subjects include linear algebra, statistics, and information theory. You should also take classes in statistical mechanics and thermodynamics, and computer architecture.

## WHAT OTHER QUALITIES ARE USEFUL?

According to Gerald, to be a success in the field of machine learning, it's important not to skip the fundamentals. It's not worth getting too hooked on any particular technology to the exclusion of others. That's because in computing, programming languages, tools and methodologies go in and out of fashion. If you know the fundamentals, you can adapt to the tools that are currently trendy, see the world for what it is, and be an innovator in a fast-changing world.

## OPPORTUNITIES IN MACHINE LEARNING ENGINEERING

According to the careers website Prospects.ac.uk, machine learning is such a new field that there are few academic courses dedicated to it.

Instead, most employers accept a Masters degree or a PhD in a relevant subject. If you want to follow this as a career path, the best advice is to study a subject that has an element of machine learning within it.

Computer programming experience is essential, and employers look for knowledge of programming languages, including C++, Java, and Python.

If you don't have a degree, you could still get into the field if you have experience in data or statistical analysis and take a dedicated Udacity course.

The average UK salary for a machine learning engineer is £52,000, but you could be paid much more if you join a large multinational company.

## GERALD'S TOP TIPS

1. Read Feynman Lectures On Computation by the Nobel prize-winning physicist Richard Feynman.
2. Take the time to actually understand things by asking and looking for answers to the 'why?' question. An answer is valid if, and only if, it fits many viewpoints.
3. It's OK to be a nerd! I got an F in art on my 12th-grade report sheet!



# HOW DID DR GERALD FRIEDLAND BECOME A MACHINE LEARNING ENGINEER?

## DID YOU ALWAYS WANT TO WORK WITH COMPUTERS?

Yes, absolutely – I started programming at the age of 7 years old and machine learning when I went to university as an undergraduate. I did consider physics for a while too. I've written a book for elementary school children called, 'Beginning Programming Using Retro Computing', because I love the idea of young people working with computers as early as I did!

## WHO OR WHAT INSPIRES YOU?

I love to make things that work. I am also inspired by simplifying complicated things - I'm driven by explaining difficult things in a way that makes people say, "OMG, that makes sense! I never thought about it that way."

## WHAT FASCINATES YOU MOST ABOUT COMPUTER SCIENCE?

Computer science is an engineering discipline that builds machines to automate maths. Less than a century ago, maths was the king discipline that governed all sciences. Now for the first time, computer science is able to show that maths itself is physical. We can build a mechanistic machine that solves mathematical problems. So, for the

first time in history, we have a means to question maths and ask the 'why' questions. For example, I have a 10-minute lecture on YouTube that explains why prime numbers exist. Maths defines and observes prime numbers as a phenomenon. But computer science can explain *why* they have to exist - it's a result of multiplication being a compressed representation of addition. This then allows us to reason about their importance for encryption, for example.

## WHAT ARE THE MOST PRESSING ISSUES FOR COMPUTER SCIENTISTS TODAY?

Programming and the fundamentals of information need to be taught in school so that there is time to actually teach the science of computer science in college. Because programming is taught so late, most computer scientists see themselves as programmers when they graduate from college. So, we miss the opportunity to see computer science as a rigorous science and engineering discipline with measurements and reproducible outcomes. I personally think calculus in school should make way for statistics and programming. Newton and Leibniz created an amazing framework for analysing data but it's not compatible

with modern tools, which are based on statistics and information.

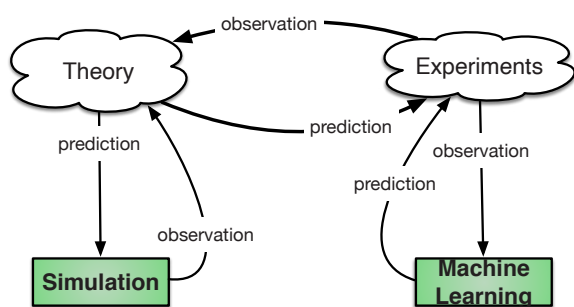
## HOW DO YOU RELAX AFTER LOOKING AT COMPLEX THEORIES?

I am a martial artist (Taekwondo), a scuba diver, and a runner. But whenever I have anxiety, I do some maths because that's a good way of channelling an overactive brain.

## HOW DO YOU APPROACH SCIENTIFIC PROBLEMS AND EVERYDAY PROBLEMS?

Science is the quickest way we know to understand the environment we live in. There are other ways but they are not as efficient. Science is not perfect, but nothing is. The biggest problem with science is that it takes patience and practice to learn, but we all have the genetic predisposition to be able to understand it. Don't let anybody tell you otherwise! If you don't get on with maths and science, don't give up. You may one day land on a teacher who will truly inspire you. My biggest everyday frustration is with ignorance and people making a political point to ignore science, because I personally don't have the patience for anything else but science.

## Machine Learning in the Scientific Method



*The scientific process, enhanced by machine learning.*



*Machine learning engineers using the Brainome Daimensions™ tool to measure and discuss the learnability of a dataset.*