UNRAVEL

Machine Learning, Explained

Begin

UNRAVEL
Machine Learning, Explained

Machine Learning Primer | Machine Learning Example Walkthroughs | Machine Learning Interpretability Tool | About & Contact

Select a tab to begin!

Welcome to UnravelML

We make machine learning make sense.

We have gathered instructional materials, examples, and tools to help you understand what machine learning is, how it can be used to solve problems, and what goes under the hood. No programming, math, or computer science experience required.

Please choose a learning tool from the top to begin. We recommend you explore them in order for the best learning experience.

The **Machine Learning Primer** is a flipbook of ideas and concepts about machine learning. What is it? How and when was it developed? What is all the hype about? We talk through these concepts using some real-world examples to give you a feel for the context surrounding machine learning. We want you to walk away from the primer with an understanding of the value of machine learning, as well as an introduction to its inner workings.

In **Machine Learning Example Walkthroughs**, we will walk through the Titanic dataset problem from end to end, with a particular focus on the interpretability tools mentioned in the Machine Learning Primer.

Finally, we have created a prototype tool for you to begin exploring machine learning with your own data. The **Machine Learning Interpretability Tool** will guide you through a basic machine learning workflow while providing live tips and hints to help maximize your understanding.

We hope you enjoy our tool! Please feel free to contact us with any comments or questions

- Kevin, Tianhao & Kathryn

UNRAVEL
Machine Learning, Explained

Machine Learning Primer | Machine Learning Example Walkthroughs | Machine Learning Interprebility Tool | About & Contact

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

Select a topic from the menu below!

Welcome to the Machine Learning Primer!

This tab is a flipbook of information and concepts to help introduce you to machine learning. Please choose a topic from the menu to the left to begin.

We want you to walk away from this primer with increased trust and understanding in machine learning. We will do that by providing you with the following:

1. An indication of what machine learning is capable of and how it fits into the data and computer science ecosystem

2. A high-level understanding and appreciation of the mathematical concepts behind machine learning. Data goes in and insight comes out... but what happens in the middle? We want to shed some light on what might seem at first glance like a black box.

3. Some tools to help you interpret the results of a machine learning algorithm. How do we convert numbers and charts to action and impact?

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

Demystifying Machine Learning

In this chapter, we will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

Specifically, we will cover the following topics:

What is ML?
This section gives general context to machine learning. If you;ve never heard of ML before, this is the place to start.

What Problems Can ML Solve?
Thi section gives examples of machine learning, and help you understand how it can be applied your projects.

ML and Artificial Intelligence
This section helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

History of ML
This section captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

Select a topic from the left to begin.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What is ML? gives general context to machine learning. If you've never heard of ML before, this is the place to start.

Demystifying Machine Learning
What is ML?
What Problems Can ML Solve?
ML and Artificial Intelligence
History of ML
Machine Learning Mechanics
How does ML work?
Types of ML
Data Science Workflow
Common ML Algorithms
Machine Learning Interpretability
What is Interpretability?
Local Interpretability
Global Interpretability
Interpretability Tools

What is Machine Learning?

Machine learning uses math to model a relationship between different types of data.

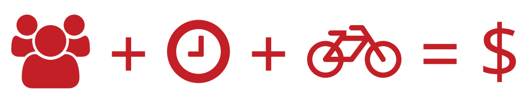
These models can help improve our understanding of how and why something occurred in the past and also help us predict what might occur in the future.

For example, imagine that you own a store that sells bicycles.

You could create a machine learning model that quantifies your daily profits as function of various aspects of your business.

For example:
- how many customers visit the shop
- how many hours your shop is open
- whether certain models are on sale, and the available stock of each bicycle type

You can use your model to better understand the performance of certain bicycle models you have sold in the past, or to help predict the performance of a new bike listed in the future.



While conventional economic and business models exist to solve this same problem, they are not custom-fit to the unique nuances of your business.

Machine learning is a method of creating custom, one-of-a-kind models that can give you more accurate and personalized results than conventional models.

Click the arrow to move through the flipboard!

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What is ML? gives general context to machine learning. If you've never heard of ML before, this is the place to start.

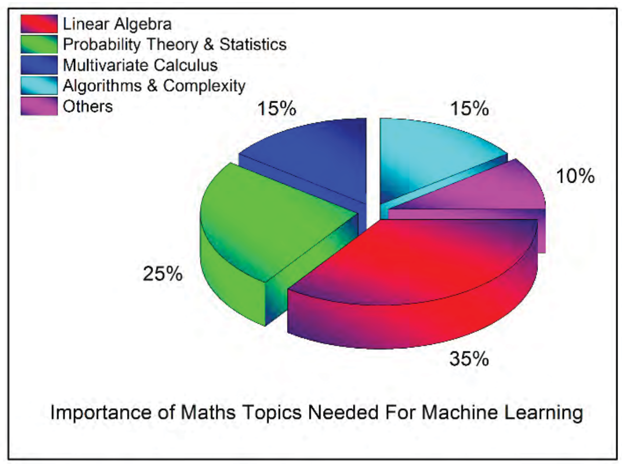
What is Machine Learning?

Machine learning is a blanket term that covers a variety of different algorithms.

For example, you may have heard of algorithms like “naive bayes”, “neural network”, or “random forest”. Each algorithm is based on a different combination of techniques from calculus, statistics, and linear algebra.

Because each algorithm comes from a different mathematical context, they are each suited to solving different types of problems.

One of the main roles of a data scientist is understanding which models are most applicable to the problem at hand, in addition to being able to quantify how well a given model has performed.



This pie chart is an excellent demonstration of the many mathematical topics that feed in to Machine Learning. Click here to read the Towards Data Science article where we got this image.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What is ML? gives general context to machine learning. If you've never heard of ML before, this is the place to start.

What is Machine Learning?

This popular XKCD comic pokes fun at how untrustworthy Machine Learning can seem. We want to get rid of this stigma.

Many machine learning algorithms use very sophisticated math and hence it is easy for those outside of data science to be put off by their "black box" appearance.

This can lead people to be skeptical about an algorithm's results and in some cases even ignore the powerful wisdom that machine learning can provide.

At UnravelML, we want to demystify the machine learning black box.

While the math in some cases may be tricky, the overall concept and application of a machine learning algorithm is generally quite straightforward.

We want everyone from any background to grasp the basic concepts of machine learning.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What is ML? gives general context to machine learning. If you've never heard of ML before, this is the place to start.

Demystifying Machine Learning
What is ML?
What Problems Can ML Solve?
ML and Artificial Intelligence
History of ML
Machine Learning Mechanics
How does ML work?
Types of ML
Data Science Workflow
Common ML Algorithms
Machine Learning Interpretability
What is Interpretability?
Local Interpretability
Global Interpretability
Interpretability Tools

What is Machine Learning?

So we know that machine learning can help us recognize patterns and create models.
For example, we talked about being able to predict sales at a bike store. But what else can it do?

Let's talk about the learning aspect of machine learning.

Did you know Machine Learning is a form of bio-mimicry?
That's right. Machine learning is in some ways built to replicate how humans learn.

When a person is born, they do not know much about anything, however they do have a brain and their five senses to help keep track of what's going on around them. A child might watch their parents walk or talk for a long time, taking in information until they have enough to try on their own. Once a baby takes its first steps or says its first word, they begin the long process of refining what they say and do every day until they speak fluently and walk with perfect balance.

Machine learning is a very similar concept - give a computer the ability to learn, some direction, and a chunck of data for it to make sense of the problem it is trying to solve. It will give you back its best guess, which may or may not be good enough to solve your problem. But as you feed it more data and give it more direction, the model it is creating will become more and more sophisticated until it can capture each aspect of your problem with precision.

Click here to watch a great video by OxfordSparks that explains this concept in depth.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

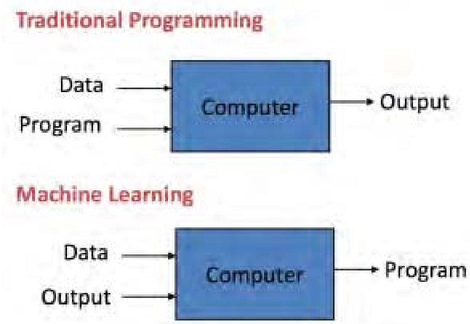
We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What is ML? gives general context to machine learning. If you've never heard of ML before, this is the place to start.

What is Machine Learning?

Okay, so machine learning can make models and it lets computers learn.
But what's the hype? How is this different from regular computer programming?



In traditional programming, a human is at the wheel and the computer is just doing what it is told. In this case, the programmer needs to have an indepth understanding of what exactly is going on in the data and how the problem will be solved. The programmer already has a model built in their head, they just need the computer to execute it. Really, the only thing the programmer is using the computer for is to help crunch big numbers.

In machine learning, on the other hand, it is the computer that is figuring out exactly what's going on and solving the problem for you. It's giving you a platform to solve similar problems every time you see it again in the future. Maybe you don't know what the exact solution is or maybe you do at it would just take forever to write down because the problem is really, really hard (think self driving vehicles and facial recognition).

Do you want to sit around programming a computer to find the precise geometry of a human eye, hand coding each pixel and shape? No. You want to give the computer a few pictures of different eye shapes and colors and send it on its merry way, letting the computer create its own definition. THAT is machine learning.

In the case of predicting sales in the bike shop example, this is the difference between using economic forecasting principles that some economist made eons ago to generalize supply and demand curves (traditional programming), versus building a fully customized sales model for your unique business, unique customers, and unique sales history (machine learning).

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What is ML? gives general context to machine learning. If you've never heard of ML before, this is the place to start.

What is Machine Learning?

Let's summarize:

Machine learning can make **custom-fit models** to help us solve problems or understand a system

Machine learning is about **teaching computers to solve problems on their own**, rather than having a human guide them through the process

Machine learning is an **umbrella term for many different types of models**, all which pull from different areas of mathematics, statistics, and computer science.

Machine learning is **often incremental in accuracy**: the output model or program improves over time with more input data or direction.

Interested in learning more about data science buzzwords? Click here for a great resource.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What Problems Can ML Solve? gives examples of machine learning, and help you understand how it can be applied your projects.

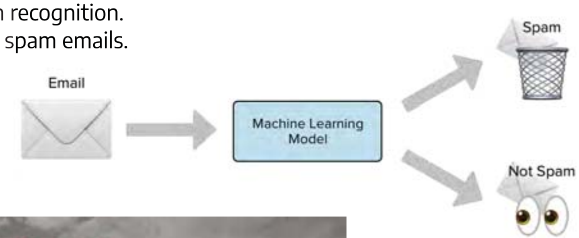
What Problems can Machine Learning Solve?

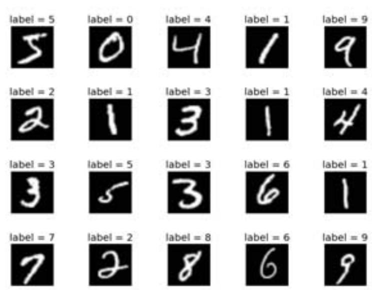
Machine Learning as a field is really only just getting started, so it's hard to have a concrete answer for this question. There is still so much left to be explored and developed.

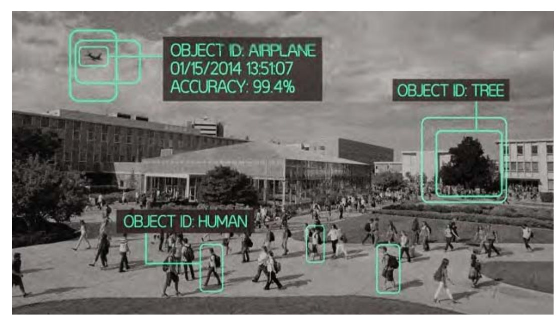
What we can give you, though are some key examples of how machine learning has been used over the past few years to revolutionize many different industries.

Here are the examples we will walk through:

Machine Learning can be used to correct text.
Machine Learning can be used to identify objects in an image.
Machine Learning can be used to profile customers.
Machine Learning can be used for speech recognition.
Machine Learning can be used to identify spam emails.







What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What Problems Can ML Solve? gives examples of machine learning, and help you understand how it can be applied your projects.

What Problems can Machine Learning Solve?

Machine learning can be used to correct text.

iPhone, Android and many other mobile phones feature auto-correcting keyboards, where the mistakes you make in typing are automatically fixed. When you type a search query into Google, it sometimes comes back with suggestions about alternative spellings. This is powered by machine learning.

Text correction is developed through a very repetitive learning process on the machine's end. As users type sentences, the algorithm is recording different charactersitics - whether the word matches a word in the dictionary, which word came before it, how it is used according to sentence structure and grammar. Sometimes it even documents to the character-by-character level. Given enough examples, the algorithm starts to remember patterns and recommend changes when it senses something uncommon.

What's more is that these algorithms can be fine tuned to a particular user. If you own a mobile phone with the autocorrect feature, odds are it is learning about your particular typing habits and altering its algorithm to best match how you talk and type.

Google's "Did you mean" feature can help you find what you need, even if you don't know how to spell it.

Apple's Autocorrect is the source of many internet memes.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What Problems Can ML Solve? gives examples of machine learning, and help you understand how it can be applied your projects.

Demystifying Machine Learning
What is ML?
What Problems Can ML Solve?
ML and Artificial Intelligence
History of ML
Machine Learning Mechanics
How does ML work?
Types of ML
Data Science Workflow
Common ML Algorithms
Machine Learning Interpretability
What is Interpretability?
Local Interpretability
Global Interpretability
Interpretability Tools

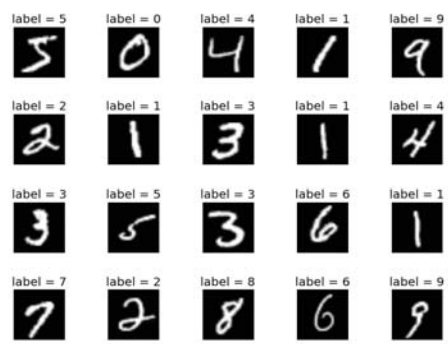
What Problems can Machine Learning Solve?

Machine learning can be used to identify objects in an image.
This is typically done with neural networks, a topic we will touch on later.

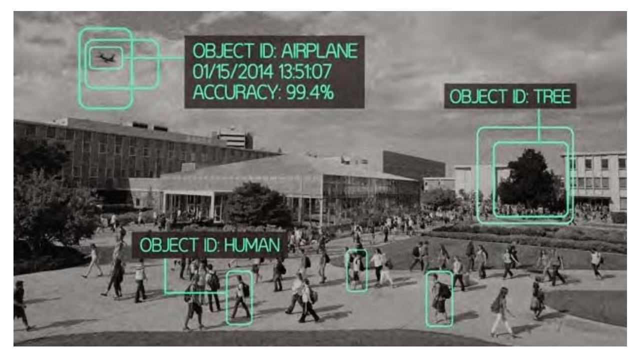
Computer vision works in a very similar way to most other machine learning algorithms. For example, if we are looking to recognize hand written letters, we feed the algorithm a corpus (data set) of pre-labeled examples for each possible letter, preferrably from many different people on many different types of paper using many different types of pens. The computer goes through these images and creates its own definition of what consitutes each letter. When given a new image with an un-labled letter, it can then sort through its existing definitions and choose the most likely match, thereby providing a suggested label.

Computer vision algorithms often work in layers. They start by identifying lines or borders in an image. Then they try to connect the lines into a shape, like a square or circle. Then they try to combine the shapes into a complex object: dog, tree, car, etc.

This type of technology has huge applications. Self-driving cars can use this to distinguish between pedestrians, cars, and bicycles. Manufacturers can use cameras to check for defects in high-speed production lines. Pictures and scans of documents can be converted automatically in to computer-readable text. Miscroscope images can be searched for the presence of certain cells or organisms.



MNIST is a great example of text interpretation.



Computer vision can also be used in surveillance.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What Problems Can ML Solve? gives examples of machine learning, and help you understand how it can be applied your projects.

What Problems can Machine Learning Solve?

Machine learning can be used to profile customers.

You finish watching a series on Netflix and recieve a recommendation to start another title, which Netflix thinks will be of interest to you. This recommendation feature is powered by machine learning, and there is no genre-expertise required!

The algorithm in this case is simply keeping track of what kind of shows you have watched in the past, or are watching currently, what kind of shows you have browsed through or watched the trailer for, and (if applicable) how you have rated the shows you have watched. From this information, aggregated over the entire Netflix viewrship, the algorithm can “cluster” different movies and shows together. For example, it can start to recognize what kinds of shows are sci-fi and which are documentaries, without having any inate understanding of what the difference is between the two. All it knows is what other people like you have watched and enjoyed.

When you finish your show, Netflix will recommend something that “clusters” well with what you just watched or have watched previously.

This type of technology can benefit the customer, but there are also many controvertial applications, most notably advertising.

Netflix Original “House of Cards” was written and cast based on what machine learning algorithms thought would be most successful. Click here to read the New York Times article.

Retail giant Target predicted a teen woman was pregnant based on her purchases, and sent baby-related advertising to her family’s home, which exposed her to her parents. Click here to read the Business Insider article.

What am I looking at?

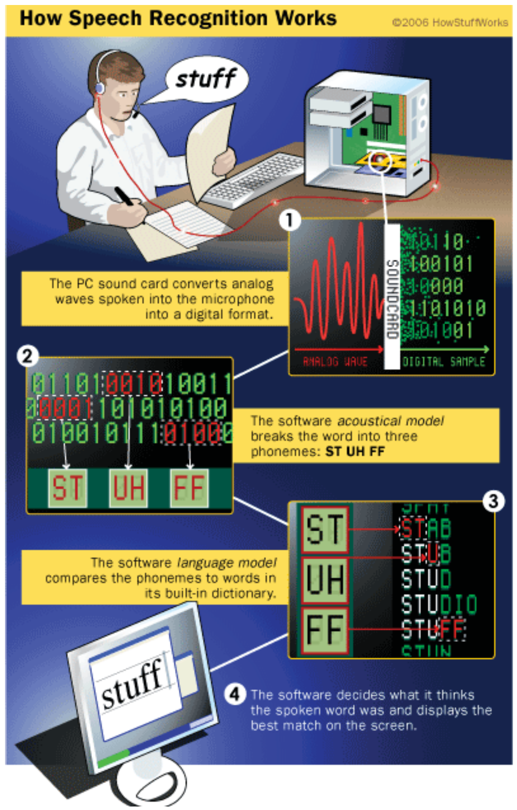
The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What Problems Can ML Solve? gives examples of machine learning, and help you understand how it can be applied your projects.

What Problems can Machine Learning Solve?



Machine Learning can be used for speech recognition.

Technologies like Amazon's Alexa or Google's Home let users speak out loud to computers, asking them to tell the time, check the weather, or add items to a grocery list. These technologies also use machine learning.

Similar to the image recognition case, here the machine learning algorithms are looking for patterns in sound vibrations. Each word you speak is comprised of different notes, pitches, and inflections that the computer has been trained to understand. When the algorithm picks up enough note patterns to match a phrase of words, it can then take an appropriate action according to some preset capabilities - for example, searching the phrase online.

Speech recognition has interesting applications where hands-free operation of a device is necessary. For example, being able to control the radio while driving.

The graphic on the left is a great visualization of the speech recognition process. Click here to learn more on HowStuffWorks.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

What Problems Can ML Solve? gives examples of machine learning, and help you understand how it can be applied your projects.

What Problems can Machine Learning Solve?

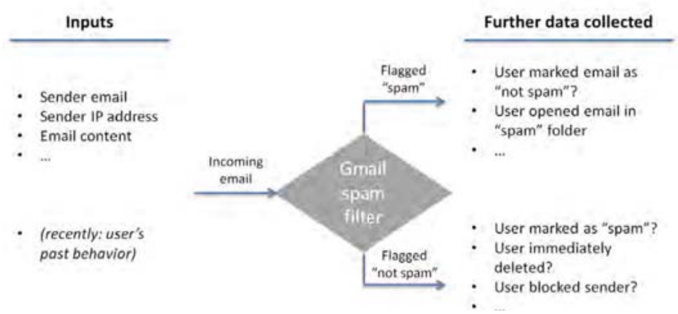
Machine Learning can be used to identify spam emails.

Email platforms like Outlook, Gmail and Yahoo have become so robust to spam emails that we hardly see them any more. But that doesn't mean they're not still being sent.

Email providers use machine learning algorithms to look for patterns in the wording of an email's subject or body that match those of common spam emails. They are also looking to key words or hyperlinks to suspicious sites. Before machine learning, this was a manual process which required preset, rigid filters. However, now it is an algorithm that is constantly changing and growing. Each email that gets flagged as spam get sent back to retrain the algorithm.

Classification like this (spam vs not spam) is a very common application of machine learning.

If you have Python experience and are looking for a first project in machine learning, building a spam detector is a great place to start! Click here to see our recommended tutorial on Medium.



The above flow chart represents Gmail's spam classification algorithm. Click here to learn more on the Harvard Business School website.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

ML and Artificial Intelligence helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

Machine Learning & Artificial Intelligence

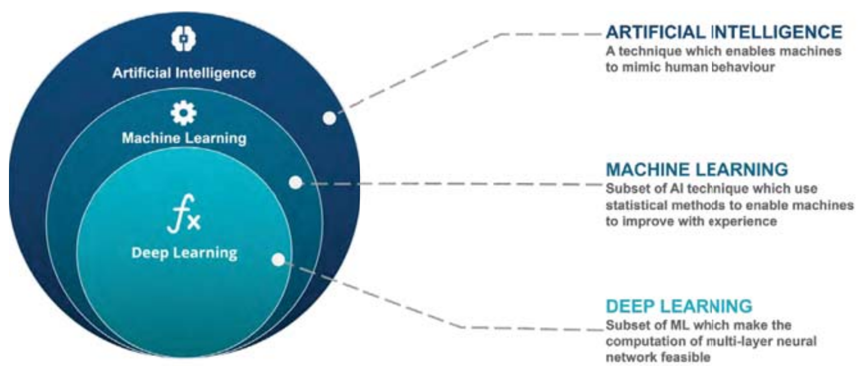
Machine learning is a subtopic of artificial intelligence.

Machine Learning describes the "algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead."

Artificial Intelligence, on the other hand, relates to the general ability of computer systems to make complex decisions and act autonomously, similar to how humans and other animals are capable of complex thought.

You may also have heard about **Deep Learning**. This is a specific type of machine learning that mimics the biological structure of a brain, which involves many layers of neurons each making a single decision or observation. This is where the term "neural networks" comes from.

We think this image best describes the relationship between these topics:



Click here to view the blog post where we got this image.
They also have a great YouTube video on this topic.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

ML and Artificial Intelligence helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

Machine Learning & Artificial Intelligence

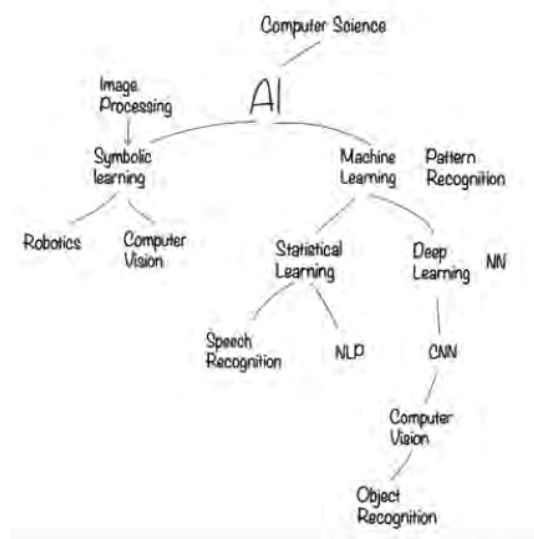
What is Artificial Intelligence?

Aritificial intelligence is a branch of computer science. The goal is to make systems that can function intelligently and independently. Typically, this means achieving rational decision making or action. but often we also characterize it as thinkig or acting in a way that a human might.

Making a system “intelligent” s a very broad goal and hence there are many subtopic in artififical intelligence. For example, how can we make computers see? How can we make computers hear? How can we make computers make decisions?

There are also different breadths of intelligence. Some AI systems may be intelligent at just a specific task. Other, more advanced projects may aim to make a computer more broadly intelligent, with skills and tools that can adapt to different circumstances.

One other way of looking at this is that the aim of AI is to take on a larger share of the cognitive load from the user. It is inefficient for the programmer to have to dictate every single task and subtask that a computer must do. If the computer is able to make some small decisions on its own then the user is not as bogged down in the intricacies of solving a problem, and can focus on the larger more abstract tasks.



The above diagram depicts some of the many subtopics of Artificial Intelligence. Click here to watch the full YouTube video.

What am I looking at?

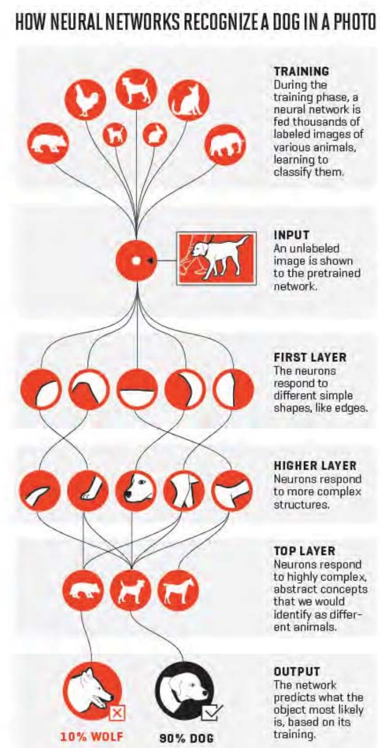
The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

ML and Artificial Intelligence helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

Machine Learning & Artificial Intelligence



Computer vision is a great example of deep learning. Click here to learn more.

What is Deep Learning?

Deep learning is a type of machine learning that is inspired by the human brain. Deep learning uses bery large and multi-layered networks to solve problems.

Different architectures of deep learning exist. For example, deep neural networks are neural networks with multiple layers, where each layer is composed of multiple “neurons” and each neuron executes a function based on the inputs it is provided. The human brain is composed of around 100 billion neurons. Deep belief networks and recurrent neural networks are also types of deep learning algorithms.

Deep learning helps break complex tasks into multiple steps or multiple layers of abstraction. Deep learning is also great for very diverese and unstructured data.

Computer visions systems, chat bots, virtual assistants, and translation technologies all use deep learning.

What am I looking at?

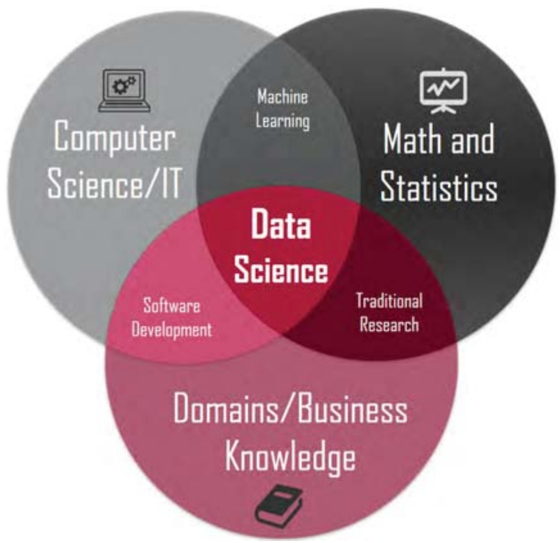
The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

ML and Artificial Intelligence helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

Machine Learning & Artificial Intelligence



Click here for a five-part series about the core concepts of data science, hosted by Towards Data Science.

What is Data Science?

Data science is a profession that has become very popular over the past few years. Data science is the combination of the math, statistics, and computer skills necessary to work with machine learning models, in addition to the business and domain knowledge necessary to use machine learning to solve real-world problems.

Data scientists typically consult for business owners, and start by understanding what kind of problems they have and are trying to solve. Data scientists then find and process data that they think are relevant to the problem, whether it be from online sources, collecting data manually, or other ways. Once the data is processed, the data scientist can then explore the data to get a feel for which parts of the data are most crucial. The exploratory works feeds into creating a preliminary model for the data. She then applies diverse machine modelling techniques to find the best performing model to solve the business problem. She then communicates and visualizes her findings back to the business owner.

We will dive deeper into the data science workflow in a future chapter.

Click here for a video about a day in the life of a data scientist.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

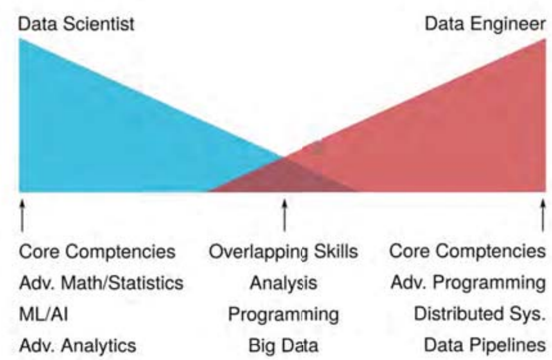
ML and Artificial Intelligence helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

Machine Learning & Artificial Intelligence

What is Deep Engineering?

Data engineers are the enablers of data scientists. They help build infrastructure like data pipelines and databases to feed data to data scientists.

A data scientist is only as agood as the data they have access to, so the role of the data engineer is very important. Data engineers work with servers, access logs, and distribution softwares to get a data scientist's models up and running, and to help them serve their results to the customer.



Click here to read more about the difference between data scientists and data engineers.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

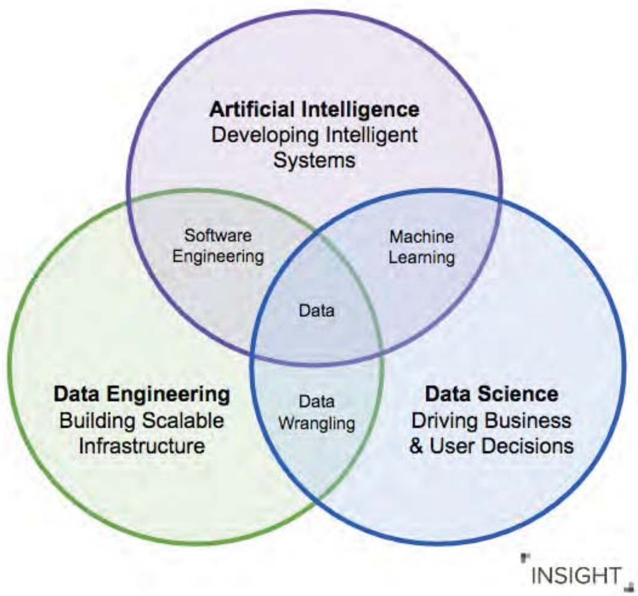
We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

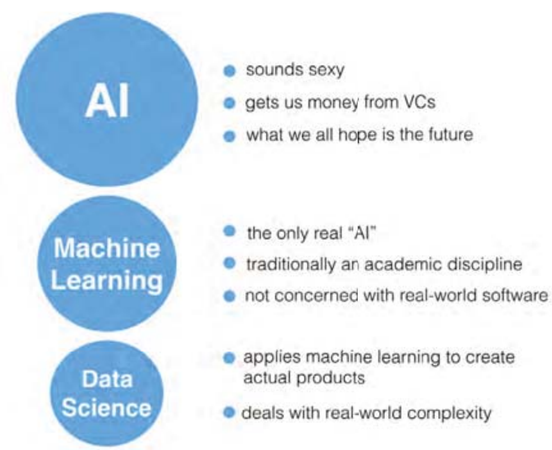
ML and Artificial Intelligence helps you place the specific role of machine learning inside the greater ecosystem of computer science, data science, and artificial intelligence.

Machine Learning & Artificial Intelligence

Here are two other ways to look at the relationship between these convepts.



Click here to read the full article on Medium.



Click here to read the full article on Medium.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

History of ML captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

History of Machine Learning


Carl Friedrich Gauss

BBC has a great online resource about the history of Machine Learning. Click here to view it.

Machine Learning begins with mathematics and statistics.

While math has been a developing subject for thousands of years, some notable accomplishments in what would become machine learning are as follows:

Least Squares (Carl Friedrich Gauss - 1795)
This theorem helps us quantify how well a model fits a given amount of data by summing the squares of all the differences between the predicted an actual values.

Bayes Theorem (Thomas Bayes - 1812)
This theorem is about conditional probability. In the spam filtering example, Bayes Theorem helps us compute the probability of “spam” given that the subject header included the word “free”.

Markov Chains (Andrey Markov - 1913)
This is a mathematical method to model a series of events, or the ability to flow from one state to another depending on some amount of probability. These help us model complex systems.

These are just three examples of the many different theorems and algorithms that comprise machine learning. In fact, the mathematics of machine learning is still evolving and growing every day.


Thomas Bayes


Andrey Markov

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

History of ML captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

History of Machine Learning

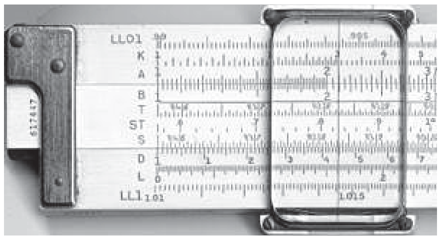
The next milestone is the invention of the computer and the ability to compute.

While mathematics and statistics are important and require a certain amount of sophistication for Machine Learning to be successful, there can be no Machine Learning without an equally sophisticated machine.

Early computers were analog, meaning that rather than representing a value with a combination of bits (as digital computers do), a value was represented in some directly measureable way, for example by a voltage level, a hydraulic pressure, or the position of a lever arm. Analog computers did not necessarily need to be powered. The slide rules (invented in the 1620s) is a great example of a mechanical analog computer. Electronic analog computers, which were the precursor to today's digital computers, were invented around 1902.

Unfortunately, analog computers were too limited in their capacity and ability to adapt to solving different types of problems. When the digital computer was invented around 1940, the analog computer quickly went obsolete.

One notable pioneer in digital computing is Alan Turing, a mathematician who devised a number of techniques to crack inercepted coded naval messages. His work enabled the Allies to defeat the Nazis in many crucial World War II engagements. Turing had a considerable involvement in the development of computer science, algorithms, and computation, all which helped create a model for what would become the general purpose computer.



Slide Rule (invented 1620s)

Alan Turing (1912 - 1954)

Konrad Zuse's Z3, the world's first fully automatic digital computer (1941)

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

History of ML captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

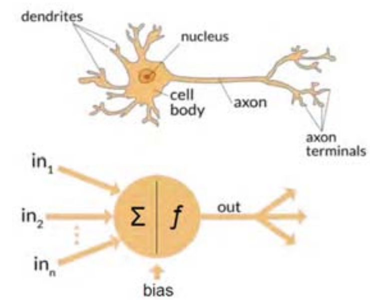
History of Machine Learning

Invention of the Neural Network

Neural Networks are a type of machine learning that model the human brain. Our brains operate with a network of neurons, each making a minute observation or decision. When combined in muliple layers, these decisions can be scaled up in complexity.

A great example of neural networks is computer vision, where objects can be detected in a series of pixels. A neural network is used in this case, where the first layer looks for borders and edges, the second layer combines borders and edges to make shapes, like a circle or square, and the third and final layer combines the shapes to make a complex object like a tree, person, or animal.

The neural network was conceptualized by Warren McCulloch and Walter Pitts in 1943.

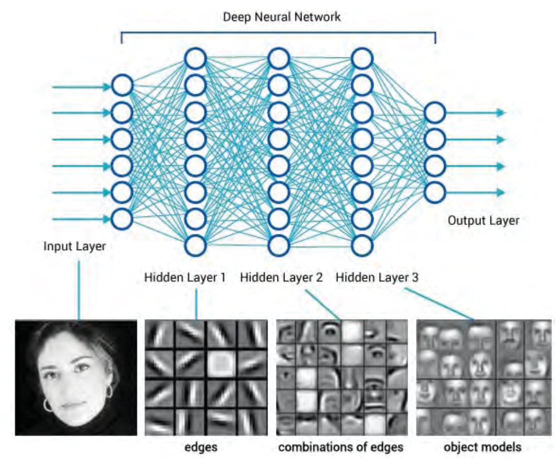


Click here to read how atificial and biological neural netowrks differ on Towards Data Science.

Artificial Intelligence “Winter”

Unfortunately, implementing a neural network at scale proved difficult due to the lack of computational power and flexibility. Other machine learning models too, were seeing great improvements in mathematical formation, coupled with great difficulty in implementation.

These frusterations led a “winter” in Artificial Intelligence in the late 1960s and continuing through until the late 1980s, where funding for machine learning research was hard to come by, enthusiasm was lacking, and there were little breakthroughs.



Click here to read a great Medium blog post about Computer Vision.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

History of ML captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

History of Machine Learning

Shift to Data-Oriented Approach

The first machine learning models focused on a knowledge-driven approach, meaning that the computer was meant to supply very high-level, human readable and symbolic knowledge. This is very difficult and was part of what led to the Artificial Intelligence "Winter".

However, as computational power began to improve drastically in the early 90s, a shift towards data-driven machine learning became apparent. In this new appoach, computers analyze large amounts of data and draw conclusions from the results. Rather than trying to solve problems algebraicly, which humans do because we cannot compute large amounts of data, computers can sort through many, many examples and combine the learning from each.

This led to a new focus on creating tools to solve certain "subsets" of problems rather than a single tool or machine to solve all problems.

One key culmination of this effort, and a turning point that rekindled hope in Machine Learning and the power of computing was when IBM's Deep Blue supercomputer beat Gary Kasparov, the reigning world champion, in a game of chess in 1997. While there is debate about whether Deep Blue actually represented an "intelligent" system, or just leverage brute-force computing, the result remained the same - the world now knew that computers can learn and computers can win.

Click here to read a great article about Deep Blue's win on PRI.org

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

History of ML captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

History of Machine Learning

Gesture recognition technology is powered by machine learning and helps us interact with computers in new way. Click here to learn more.

Machine learning is at the core of autonomous vehicles and self-driving systems that can make our cities safer. Click here to learn more.

The Internet: Making Machine Learning Main Stream

After DeepBlue, the sky was the limit. Throughout the 2000s, computing power continued to grow exponentially, meaning that harder and more complex problems were being solved, larger data sets were being analyzed, and more complex operations were being performed.

There was a constant push to find the limit of Machine Learning. How far could it go? Deep Blue's success was replaced in 2015 through 2017, when AlphaGo, a supercomputer built by Google, beat a series of professionals (including Ke Jie, the world's reigning champion) at Go, a game considered much more difficult and complex than chess.

Throughout the 2000s and 2010s, more and more resources became available through the internet. Anyone with access to the web became able to install their own . This crowd-sourcing perspective on problem solving has, as in many other industries, substatantially increased the productivity rate in which computer scientists can work on machine learning problems, in addition to improving literacy and transparency about machine learning's capabilities.

Cloud computing, which has become popular in the late 2010s, gives access to high powered clutered computing power accessible at any time and anywhere. This means that anyone can deploy an intensive machine learning algorithm or complex data analysis task, without needing to own any special equipment.

Now, as we approach 2020, machine learning is becoming ubiquitous. It is being used to improve business processes, to maximize investment returns, to get you to work faster and safer. Anything and everything that can be optimized can find value in a machine learning approach, and as you read this there is probably someone out there trying to make it happen.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Demystifying Machine Learning** section will give you an indication of what machine learning is capable of and how it fits into the data and computer science ecosystem.

History of ML captures the ancestry and origins of machine learning, and gives indication to why it has become so important and popular over the past several years.

History of Machine Learning

What's next? The singularity?

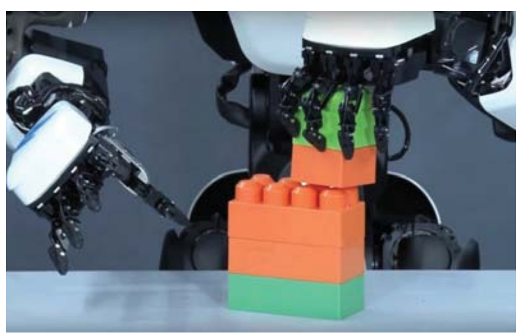
Many people believe that there will come a point where computers begin to learn on their own and we will no longer need to instruct them to do things. This is called the singularity - when computers will take on their own "life".

Shows like the "Jetsons" and "Futurama" depict an optimistic view of this phenomenon, where computers can become an intricat part of our lives, cooking us dinner and doing our laundry. They might even become a companion and friend.

Movies like "I, Robot", "Transcendence" and "Ex Machina" showcase a dystopian view of artificial intelligence. What happens if computer gain a sentience and turn against us?

While these are pop culture references and have little technical backing, the scientific viewpoint is also quite conflicted. Many scientists think it is plausible, though there are many limitations and obstacles yet to overcome. Notable, there is still a lack in the hardware and software necessary to support ultra-complex computing of this style. There is also little evidence to believe that computers would suddenly take on emotion complex enough to understand the difference between like and dislike.

When will the singluarity occur? Scientists are estimating some time in the 2030s or 2040s, if at all.



Click here for a great article on today's best humanoid robots.

Ray Kurzweil, Google's Director of Engineering, is prediciting that the singularity will happen by 2045. Click here to see his reasoning.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Machine Learning Mechanics

In this chapter, we will look under the hood of machine learning. To build trust in the process, we want you to understand at a high level what happens to the data inputted to an algorithm and how it becomes an insightful output.

Specifically, we will cover the following topics:

How does ML work?
This section goes through the underlying mechanics and architecture of machine learning in general.

Types of ML
This section breaks machine learning down into a few different categories and explains the purposes of each.

Data Science Workflow
This sections walks through a day in the life of a data scientist, to show the complete process of how machine learning is used to solve business problems.

Common ML Algorithms
This section walks through the intricacies and specific applications of several common machine learning algorithms.

Select a topic from the left to begin.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

How does ML work? goes through the underlying mechanics and architecture of machine learning in general

Demystifying Machine Learning
What is ML?
What Problems Can ML Solve?
ML and Artificial Intelligence
History of ML
Machine Learning Mechanics
How does ML work?
Types of ML
Data Science Workflow
Common ML Algorithms
Machine Learning Interpretability
What is Interpretability?
Local Interpretability
Global Interpretability
Interpretability Tools

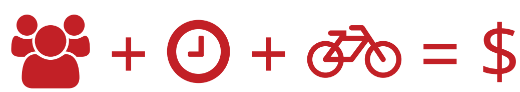
How does Machine Learning Work?

When you first hear of machine learning, it might come across as a black box that converts large amounts of data into powerful insights.

While this is not untrue, it is a very macroscopic view of machine learning and does not give any indication to how the insights are being generated. Let's dive one layer deeper by looking at each part of this process in detail. We will do this in the context of the bicycle shop example from the previous chapter.

Imagine that you own a store that sells bicycles. You want to create a machine learning model that quantifies your daily profits as function of various aspects of your business.

For example, you have data like the following:
• how many customers visit the shop
• how many hours your shop is open
• whether certain models are on sale, and the available stock of each bicycle type



You want to use your model to better understand the performance of certain bicycle models you have sold in the past and to help predict the performance of a new bike listed in the future.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

How does ML work? goes through the underlying mechanics and architecture of machine learning in general

How does Machine Learning Work?

What is data?

Data comes in many forms: large, small, dense, sparse, clean, messy. The first step in a typical data science workflow is to collect, join, and clean data relevant to the problem at hand.

The simplest input to a machine learning model is a simple table of “features” (columns) and “examples” (rows).

More complex data structures might use schemas like snowflake or star to combine information from multiple tables.

Whatever the case, data contains structured information about the problem you are trying to solve, where each row typically is an example or recorded occurrence.

Bicycles Sold	Date	Day of Week	Total Foot Traffic	Ongoing Sale	Opening Mountain Bike Stock	Opening Road Bike Stock	Opening Performance Bike Stock
27	June 13, 2019	Friday	121	Yes	27	47	32
36	June 14, 2019	Saturday	247	Yes	16	36	27
22	June 15, 2019	Sunday	198	Yes	0	28	15
14	June 16, 2019	Monday	73	No	30	47	34

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

How does ML work? goes through the underlying mechanics and architecture of machine learning in general

How does Machine Learning Work?

Do I have the right data?

It might not be clear at the beginning what type of data will provide the most value to solving the problem. In the bicycle shop example, what if the color of the bicycle has a significant impact on its popularity?

A good data scientist can help identify when important factors are omitted, however it can take an entire team of diverse backgrounds and subject matter experts to pinpoint exactly what is missing. This can be time consuming.

In our experience, it is typically best to work with what you have or what's most easily accessible and see how far you can get. When you have some results, you can evaluate whether the accuracy and performance fits your needs, and then go back to reiterate if need be.

How good is my data?

When collecting and cleaning data, there are also tradeoffs between accuracy, volume, time, and resources. You may think that large and dense data sets are the key to success, but that is not always true or necessary. Most projects will never have enough time and money to acquire the data that is needed to create a perfect model. That's life and data science is built to handle it.

There are many way to work around missing and flawed data. A core component being a data scientist is learning to work with messy data, and not get hung up in the pursuit of perfection. We encourage you to consider the cost versus reward value of the resources you are investing in a machine learning problem.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

How does ML work? goes through the underlying mechanics and architecture of machine learning in general

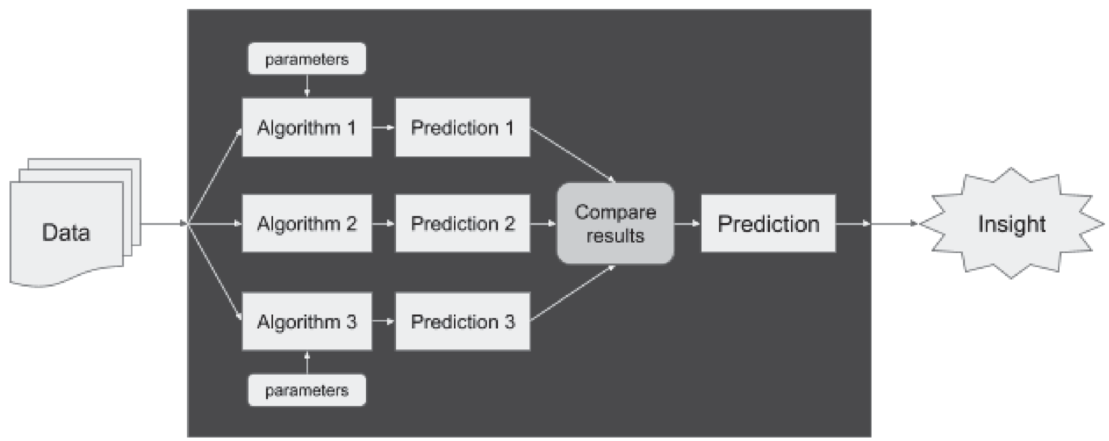
How does Machine Learning Work?

Once an appropriate data set has been created, the next step is to select a machine learning algorithm to create a model that represents the problem.

Again, there are many different data science algorithms, all based on a different combinations of techniques from calculus, statistics, and linear algebra. Each has its pros and cons and is best applied to a very particular type of problem. Fundamentally, though, they all result in the same thing: a model of reality.

Similar to how speed is distance over time and profit equals revenue minus costs, machine learning algorithm uses data from the past to create a “mathematical rule” (or set of rules) that approximates what is going on in your problem. For example, we might find that blue bikes are worth $10 more than red bikes, and bikes of Brand A are double the value of bikes of Brand B. More bikes are sold on Wednesdays than Tuesdays.

Once we have a approximate model of reality, we can use it to predict certain future outcomes by giving the mathematical rule information about the scenarios we expect to see. For example, what would my profit be on a typical week, given than I have 10 red Brand B bikes and 5 blue Brand A bikes available for purchase?



What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

How does ML work? goes through the underlying mechanics and architecture of machine learning in general

How does Machine Learning Work?

Which models should I test and how many?
How do I know know which model is best?

Data scientists typically test out more than one machine learning algorithm per problem, and compare the results to select what is most appropriate for the given situation.

There are many pros and cons to each model type, and some models can be tuned using tuning parameters in order to optimize for a certain result. The data scientist's role is to assess how well a model performs from a holistic view.

Once the models have been compared, the data scientist may choose to reiterate on the modelling process - going back to select new parameters or try a different algorithm - or they may continue with their selected model and use it to provide finalized predictions.

One common example of this is tuning to minimize false negatives or false positives. Imagine that you are in the medical field and are creating a machine learning model to predict whether a patient has a cancer. In this case, you would want to minimize false negatives - ie. avoid tagging patients with cancer as being healthy. You would rather follow up with many patients unnecessarily than risk missing a patient in need.

On the other hand, consider that you are an insurance broker looking to flag cases that are likely to be fraudulent. The investigation process for a fraud case is quite costly and you know there are likely more fraudulent cases than you can dedicate resources to. In this case, you would want to minimize the number of false positives - ie. "false alarms" that claim a case is fraudulent when really it is not. You would only want to flag the cases that you are very very certain are fraudulent in order to best spend your resources.

Tuning for false positives versus false negatives, and evaluating the importance of each, is a great example of the role of a data scientist in comparing different data science models.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

How does ML work? goes through the underlying mechanics and architecture of machine learning in general

How does Machine Learning Work?

What do I do once I have a model that works?

Congrats! The hard part is over.

Now that you have a working model, it is time to apply it to your business. This typically means using the model to generate predictions.

In the bike shop example, you could use the model to predict next week's profits based on the amount and type of bicycles you have in your show room. This would help you manage your finances better and potentially open up capital for new projects like a marketing campaign or employee retreat.

You could also use the model to simulate the success of different bicycle models. For example, if you are considering adding a new model to your inventory, you could estimate its success based on past performance of bicycles with similar features. You might find that blue bicycles sell much better than red bikes, and so if the bike you are looking to add is blue, that would be ideal.

As you can see, this step in the process begins to err into the world of business and strategy as opposed to strictly science and computation. Data science and machine learning are extremely helpful tools for information costs and benefits, but (at least at present), they cannot make decisions for you as they only see a subset of all information. The human element and domain experience are a vital and necessary component of successfully applying machine learning to business.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Types of ML breaks machine learning down into a few different categories and explains the purposes of each.

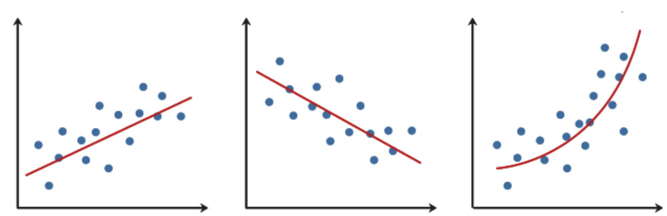
Types of Machine Learning

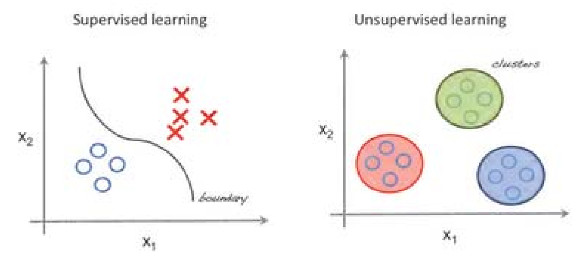
Machine Learning algoithms have a lot of variation and differences in how they are logically structured and which types of problems they are best stuited for. Hence, data scientists use a few different classification schemes to subset the large number of available algorithms.

We will talk about these classifications in detail over the next few slides.

The first classification is based on supervision, or the way in which a model learns. Examples are supervised learning, unsupervised learning, semisupervised learning, and reinforcement learning.

The second classification is based on what the model is used for. Examples are classification, clustering, dimensionality reduction and regression, though there are more.





What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Types of ML breaks machine learning down into a few different categories and explains the purposes of each.

Types of Machine Learning

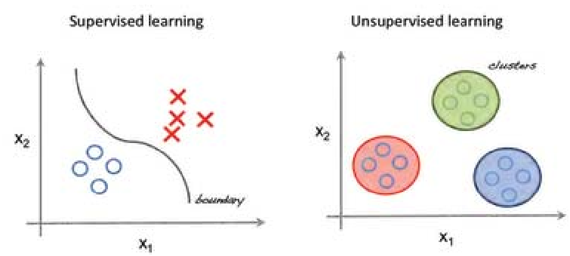
Supervised and Unsupervised Learning

This classification scheme is based on the way in which an algorithm learns. You can think of this as a measure of how much "ground truth" data the algorithm has to work with.

The majority of machine learning algorithms in use today are **supervised** in nature. In supervised learning, the algorithm is given a large set of "ground truth data" and creates a model that maps input to output.

MNIST's digit recognition corpus is a great example of a supervised learning problem. The algorithm is given a large set of images with a corresponding label. The algorithm outputs a model with the ability to label new images.

Unsupervised learning refers to algorithms that do not have any labeled or ground truth data. Rather, they typically try to cluster or sel-organize the data in some specified way such that a classification scheme becomes apparent to the user. Unsupervised learning is looking for commonalities and differences between each piece of data.



This image is a great visualization of the two schemes.
If your looking for a tutorial on Unsupervised Learning, click here.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Types of ML breaks machine learning down into a few different categories and explains the purposes of each.

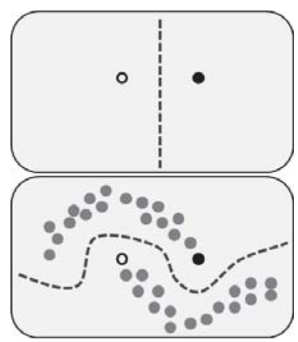
Demystifying Machine Learning
What is ML?
What Problems Can ML Solve?
ML and Artificial Intelligence
History of ML
Machine Learning Mechanics
How does ML work?
Types of ML
Data Science Workflow
Common ML Algorithms
Machine Learning Interpretability
What is Interpretability?
Local Interpretability
Global Interpretability
Interpretability Tools

Types of Machine Learning

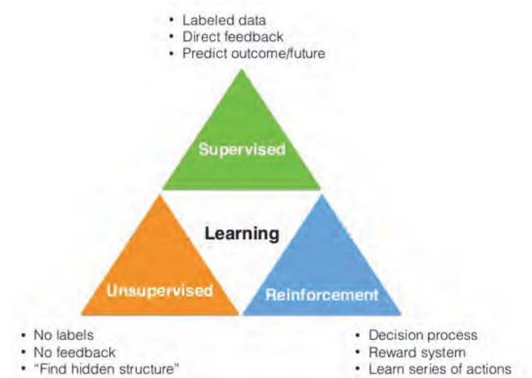
There are two other supervision types that also fit in this spectrum.

Semisupervised Learning falls between supervised and unsupervised learning. Typically the algorithms are given a small amount of labeled data and a large amount of unlabeled data. A combination of both supervised and unsupervised techniques are leveraged in tandem. Semisupervised algorithms are beneficial because they are often more accurate than unsupervised algorithms and the dat acquisition is less costly (creating labels can be very difficult and time consuming).

Reinforcement Learning is a bit different than the supervision schemes we have discussed. The goal of reinforcement learning is to inventivize or penalize certain actions that the algorithm is taking in order to model an optimal behavior. Reinforcement algorithms are not given any ground truth data, but instead learn from experience or trial-and-error.



This image is a great depiction of the power of semisupervised learning. In the first image, a small number of labeled data is known. In the second, unlabled data is added and the resulting policy changes drastically. Click here to learn more.



The above image is a good summary of supervision. If you're looking for more information about Reinforcement Learning, click here to read a great blog post, which includes information about AlphaGo's reinforcement algorithms.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

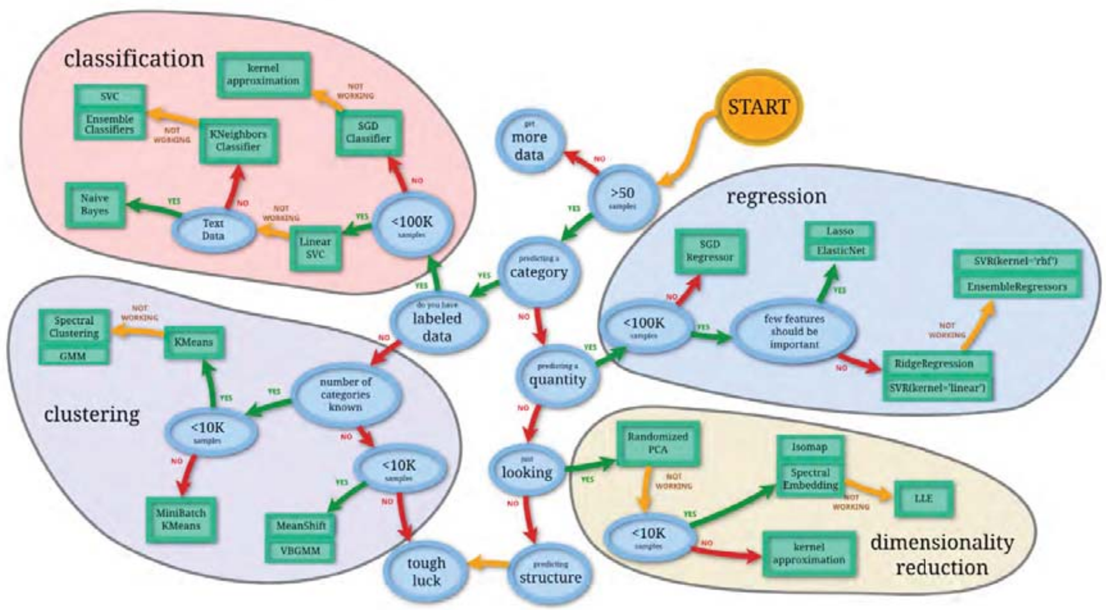
The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Types of ML breaks machine learning down into a few different categories and explains the purposes of each.

Demystifying Machine Learning
What is ML?
What Problems Can ML Solve?
ML and Artificial Intelligence
History of ML
Machine Learning Mechanics
How does ML work?
Types of ML
Data Science Workflow
Common ML Algorithms
Machine Learning Interpretability
What is Interpretability?
Local Interpretability
Global Interpretability
Interpretability Tools

Types of Machine Learning

The next type of classification scheme we want to talk about relates to the task the algorithm is trying to accomplish. We think scikit-learn, a popular Python library for machine learning, has the best visualization for this.



Click here to view this on scikit-learn's website.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

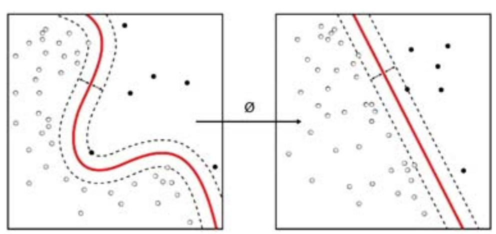
The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Types of ML breaks machine learning down into a few different categories and explains the purposes of each.

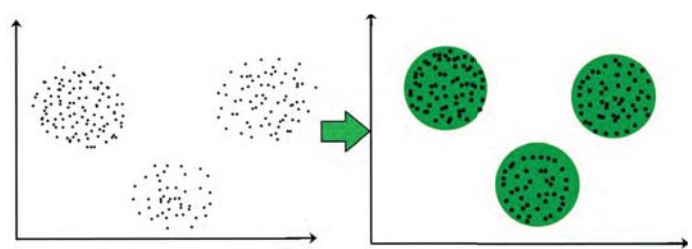
Types of Machine Learning

Classification algorithms are used to identify which of a set of categories a new observation belongs to. Simple classification problems are binary, meaning they have only two categories. The spam email filtering example uses classification to filter emails into either the spam or non-spam (sometimes called ham) categories. Other classification problems can have many categories. Adding categories, however, typically requires a more sophisticated and complex algorithm. Classification algorithms are typically supervised.

Clustering algorithms identify groups of data that share similarites. The goal is to fit each data point into a specific cluster. Clustering is typically unsupervised, so you can think of it as the algorithm creating categories for the data based on what it sees rather than what it is told. Clustering algorithms are often used in social media schemes - for example, sites like LinkedIn might recommend new connections to you based off of other users it identifies as being in or near your cluster.



Classification algorithms are often visualized by segregating data with lines. Click here to learn more.



Clustering algorithms are often visualized by segregating data with ellipses. Click here to learn more.

What am I looking at?

The Machine Learning Primer is a flipbook of concepts and information to introduce you to machine learning.

We want you to walk away from this primer with increased trust and understanding in machine learning.

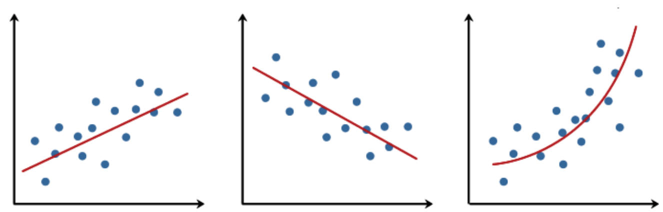
The **Machine Learning Mechanics** section will look under the hood of machine learning to give you an understanding of what happens to the data inputted to an algorithm and how it becomes an insightful output.

Types of ML breaks machine learning down into a few different categories and explains the purposes of each.

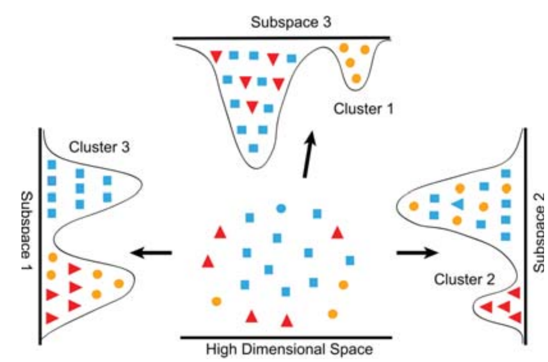
Types of Machine Learning

Regression algorithms estimate the numerical relationship among variables. In the bike shop example, being able to quantify how much additional revenue adding a new bike model to the inventory would bring is an example of regression. Simple regression algorithms are linear in nature. More complex algorithms can take on various polynomial and multivariate shapes.

Dimensionality Reduction is a method to reduce or combine the amount of variables that an algorithm needs to consider. This can be important when you have a dataset with many, many features where some of the features many be highly correlated. Demographic data is a great candidate for dimensionality reduction, because, for example, areas of poverty may be highly correlated with elevated mortality rates and elevated crime. Having all three of these features may negatively impact the efficacy of machine learning algorithms. Dimensionality reduction is a way to find highly correlated features and combine or reduce them in some way.



Clustering algorithms are often visualized by segregating data with ellipses. Click here to learn more.



By reducing the dimensionality of a dataset in multiple different ways, we can look at the data from different perspectives. Combining these results together can give powerful insights. Click here to learn more.