SIT307_SIT720 - Machine Learning - Week 1: Introduction to machine learning

1. 1.1 Welcome to SIT307 and SIT720
2. 1.2 Learning with Deakin
3. 1.3 Essential skills and concepts
4. 1.4 Defining Machine Learning
5. 1.5 Real-world applications of machine learning
6. 1.6 Machine Learning steps
7. 1.7 Supervised learning overview
8. 1.8 Unsupervised learning overview
9. 1.9 Reinforcement learning
10. 1.10 Model evaluation and selection
11. 1.11 Mathematics and ML
12. 1.12 Vectors and its operations
13. 1.13 Distances between vectors
14. 1.14 Matrix algebra
15. 1.15 Some special matrices
16. 1.16 Let's get started
17. 1.17 Setting up Python
18. 1.18 Types of variables
19. 1.19 Branching and decisions
20. 1.20 Iterations
21. 1.21 Function
22. 1.22 Python modules and packages
23. 1.23 Using the NumPy Package
24. 1.24 Matrix addition and subtraction
25. 1.25 Matrix multiplication
26. 1.26 Matrix division
27. 1.27 Useful matrix manipulation code in Python
28. 1.28 Pass activity
29. 1.29 Weekly wrap up
30. data

Welcome to SIT 307 and SIT720 Machine Learning.

Machine Learning is a complicated and challenging field. The techniques and methods you learn here will open a world of opportunities in the ever-expanding Machine Learning field.

This Masters level Unit moves quickly through basic Python programming, reminders on linear algebra, matrices, probability mathematics and statistics, as well as introducing practical examples of Machine Learning. Assessments involve coding real world examples in the programming language, Python.

In weeks 7-10 you will extend your knowledge into more complicated algorithms and models of machine learning using Python.

Unit Learning Goals

By the end of the Unit you should be able to:

- Demonstrate knowledge and understanding of fundamental concepts of machine learning methods and their application in solving real-world problems.
- Explain different design decisions that need to be taken to build machine learning based solutions.
- Analyse, evaluate and compare the quality (generalisation, performance etc.) of machine learning based solutions that are designed for a specific problem.
- Effectively communicate the design, development and performance of a machine learning model to technical and non-technical audiences.

Meet the team

Check out the Unit Staff page (go to Resources>Unit information>Unit staff) to see all the details of your Unit Chair, Campus Co-ordinator and Tutors are for this unit.

Key learning resources

The unit guide (go to Resources>Unit information>Unit guide) gives a full overview of the 10 weeks and the assessments. Please read it carefully.

The Deakin University Library reference list (go to Resources>Reading list) gives you access to all the unit readings.

Your fellow students, your educators and the IT Help Hub are all great resources too.

Communication

I will post important announcements through the home page, though the discussion forum Discussion board or via your Deakin email. Please check regularly.

Please use the weekly discussion forums to discuss the content. Don't forget you can help and learn from each other.

Support

This discussion board has specialised forums, familiarise yourself with them:

- **Questions for the Unit Chair** – to post questions related to all matters regarding the admin and teaching of this unit. The answers to these questions may help fellow students
- **Student Discussion** - This Student Forum allows you to get to know fellow students and to discuss unit content and support each other in your learning.
- **Ask your Librarian** – Please use this discussion area to ask your Liaison Librarian any questions you have about using library resources for your assessments such as searching databases, referencing, using EndNote and finding and evaluating information. The answers to these questions may also help fellow students.
- **Assessment Forum** – to discuss your questions and answers about assessments.
- **Tech Support Discussion area** – questions for technical issues
- Weekly Discussion Forum - to discuss the weekly content

The educational team will respond to all questions as soon as they can. BOOKMARK it now!

The School of IT has introduced a IT Help Hub to help students with mathematics, programming and any other course related issues. Please make use of this resource. **Online support sessions are available** and there is a drop-in service on campus.

Note: any external links in this unit will take you to third-party websites, which may ask for your personal details. Please read our privacy policy for more information.

https://pixabay.com/images/id-3370832/

Learning with Deakin

Learning with Deakin

Deakin's units are based on a few, but important, learning principles. To get the most out of this unit, we encourage you to keep in mind the following.

Our learning philosophy

Deakin University is a pioneer in innovative digital learning. Several key principles underpin all our units, which are designed to:

- challenge you with new and interesting ideas
- prompt you to connect and engage with other learners and Deakin educators
- test your knowledge and understanding against clearly defined learning outcomes.

Get involved

Getting the most out of your learning experience with Deakin begins by getting involved.

One of the key ways to optimise your learning is to participate in the tasks set by your educators and engage in topic-related conversations with other learners.

Learning through conversation

In this unit, you're invited to comment on topics, discuss issues and ask and/or answer questions. Some of the ways you can do this include:

- Attending classes and workshops where you will have an opportunity to ask questions and engage in interactive discussions with your fellow students and your teachers

Using the discussion forums to:

- add a post ("start a new thread") to begin or continue a conversation by building on what others are saying
- using "reply" to ask a question or make a comment that indicates you're interested in what someone else has said and encourages them (or others) to expand on their ideas
- making a comment that links, compares or contrasts different themes in the conversation
- disagreeing with a comment in a curious, constructive and compassionate manner

Contributing to your own and other's learning

You'll notice that at the end of each page in this week, there's a chance for you to make a contribution. We call it 'Activity'. It's more than just making a comment. It's what brings the course alive for everyone. Our educators bring their years of research and expertise to the units we put together. And every learner also brings their life experience to the unit. And it's the meeting of these two things that turns abstract ideas into real knowledge and skills.

We talk about the three C's that guide these long exchanges with one another. Be curious, be constructive, and be compassionate. Let's unpack each of these for a minute. Being curious means always being ready to ask the next question, being open to finding out something new and being prepared to have your own ideas challenged. Being constructive is the other side of that. It's about carefully stimulating other learner's curiosity, not closing anyone down with a quick or harsh comment, being careful of the way you respond to each other, which brings us to compassion. How do we stand in another's shoes? How do we really see from their perspective?

At a time when people are building walls and closing down borders, we believe that the power of education is in opening up our world so that we can see each other more clearly and live with each other more peacefully. So thanks for joining us. Now it's over to you let's have some fun learning together.

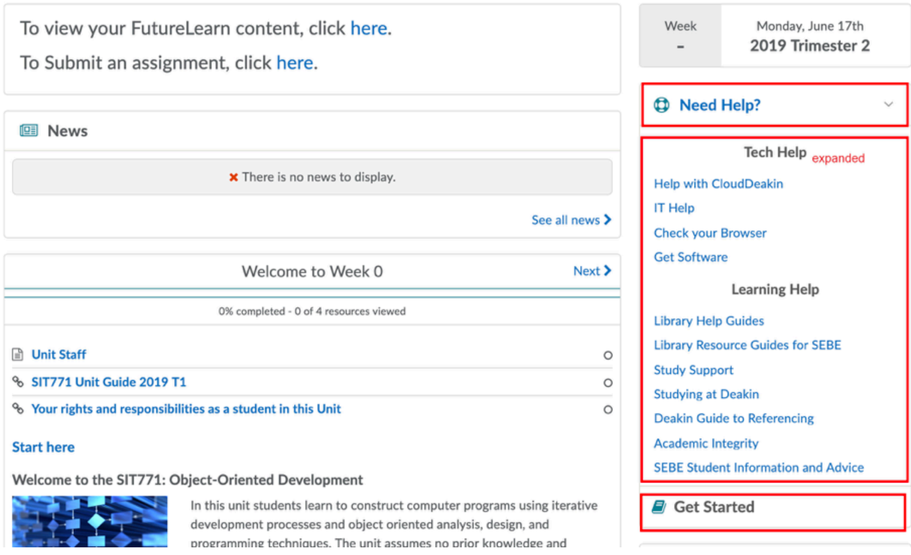
Activity

1. Practise getting involved by reflecting on the following and using the discussion forum to join the conversation:

- Why do you think getting involved is an important part of study?
- How will you get involved?
- What barriers might you encounter to getting involved and what are your tips for overcoming these?

2. You can get help with using Cloud Deakin by accessing the "Need Help"widget on the Home Page (Home at the top of the screen) and the "Get Started" link: (see both in the image below)

Explore the help links on the Home Page (Home at the top of the screen), and take note of the help available to you as a Deakin student.



© Getty Images

Essential skills and concepts

We're going to leap in and make this first week a big week.

You will become familiar with machine learning terminology, refresh your knowledge of **vectors**, **matrices** and **probability** and install tools to build some simple **Python programs**.

Learning Goals

By the end of week 2 you will be able to:

- discover real world applications of machine learning
- explore the different elements and types of machine learning algorithms
- review the basic concepts and skills necessary to understand building blocks of machine learning algorithms.

Scope and level

Students with a wide range of skill levels are participating in this Unit. Everyone has very different backgrounds and skills. If you have experience in a certain area in this cross-disciplinary field then take the opportunity to extend your knowledge with the extra resources offered or with personal exploration. You will learn a lot from explaining concepts to other students in the comments. There's no better way to test your understanding.

Python 3

Whether your coding skills are rusty or you're a coding guru, you will need to start learning Python immediately. Your assessments require Python so it's important to jump in, look at some programs and learn what they're doing. Python is a clear and easy language to learn. If you have issues with learning Python programming, you can access Helphub for a one to one session. In addition, you can ask questions to your educators during the workshop and cloud seminars. You can also post your queries in the discussion forums for educators to address and share your problem with peers to solve through the forum.

If you've coded in Python before, you will have a head start. Take a moment to help any students who are inexperienced coders. You will consolidate your own learning by explaining concepts to someone else.

Beginners

If you are an inexperienced coder, it's a good idea to get started on learning Python. We have collected a range of resources (go to Resources>Reading list) for you. Try one of the beginner items on the Resources page. "Head first Python" is a visual learning online text course (so install Anaconda first) while Datacamp is an interactive real time learning method. Browse through and select one that works for you.

What is an algorithm?

Previous students have asked about this strange word.

An algorithm is a set of instructions.

Computers need precise instructions to perform tasks. An computer algorithm is written like teaching a child to make a cake. If a recipe is: '1 cup plain flour, 1 cup SR flour, pinch of salt, 2 eggs lightly beaten, cook mod oven 30 minutes' there are many assumptions.

A computer algorithm would require something more like:

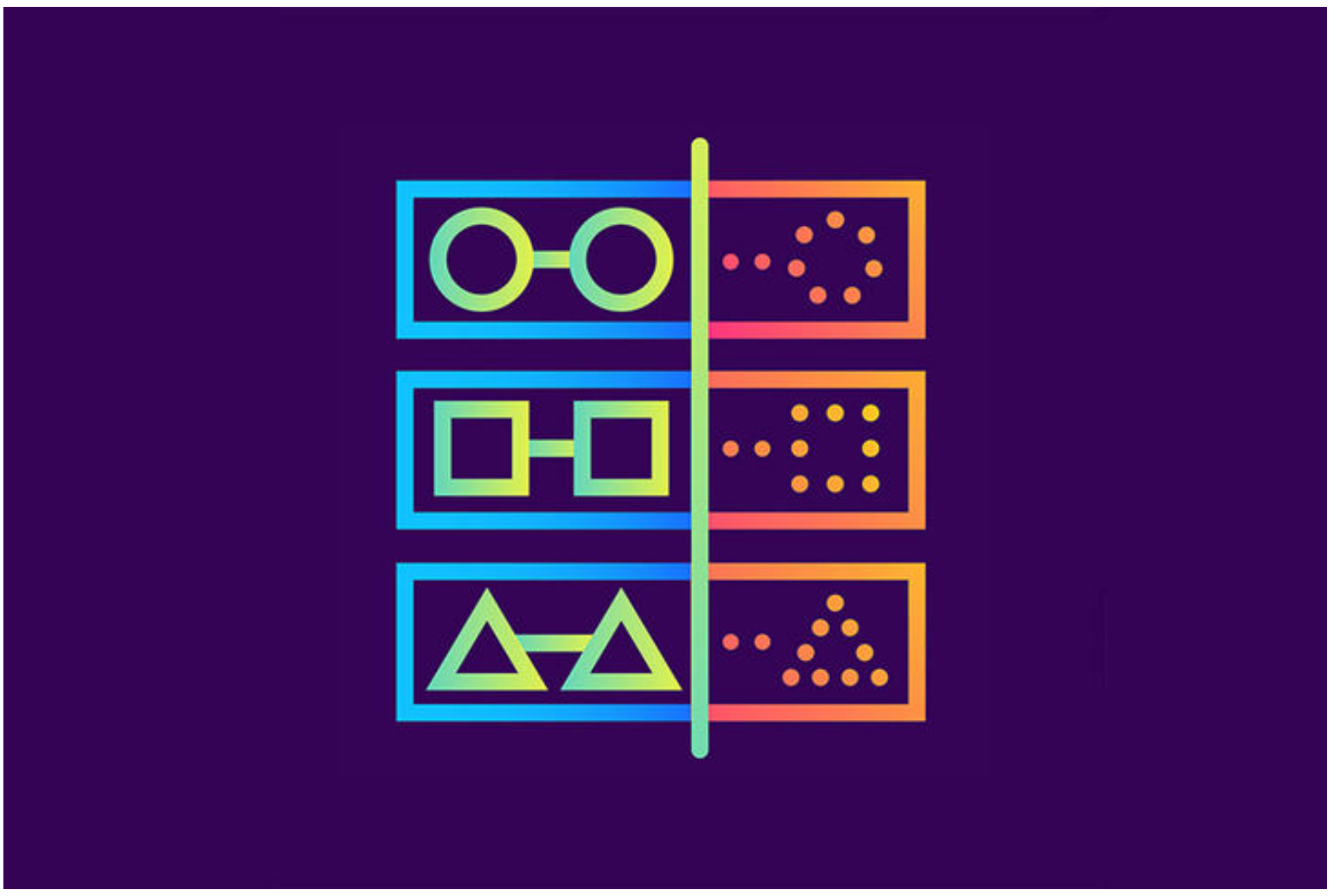
```
select working area near clean fresh water, a domestic oven and cooking tools
activate the oven at temperature setting 180 degrees C
select one 'good bowl' 30cm diameter, condition 'clean', type 'mixing', made of material 'steel'
  if 'good bowl' available, go to place bowl
    if no 'good bowl' available, pause cooking (5 minutes) and load package 'washing up'
                        if 'washing up' fails to load, call 'stop cooking', wait for further instructions.

place bowl securely on a level surface convenient to your height
select one mixing implement
        select a spoon of type 'cooking' made of material 'wood' featuring a handle in excess of 25cm in length
        else if you have no such spoon, select a tool of type 'spatula' with a total length of not less than 25cm
    if you have a silicone spatula select it
                else if no silicone spatula,
                then if you have a spatula of material 'flexible food-grade plastic' select it
else… and so on through every little step
'stop cooking' function
        deactivate oven
        put away tools
        buy cake at shop
end
```

Coding is a different way of thinking. Luckily there are many ready made modules we can use instead of writing all our own code. We will use NumPy, SciPy and Matplotlib amongst others.

Activity

How much programming have you done? If you an experienced coder, share any tips for getting started that you found helpful. Why it was helpful. Share your experience with your fellow students.



© Getty Images

Defining Machine Learning

What makes Machine Learning (ML) different from a simple computer program?

The 1959 definition below makes Machine Learning (ML) seem like magic:

> "Field of study that gives computers the ability to learn without being explicitly programmed,"
> (Samuel 1959)

As we explore further you will see ML is a set of tools to derive meaning from data. For now, here is another definition that describes ML in mathematical terms.

Machine Learning as an equation

> "A computer program is said to learn from experience *E* with respect to some class of tasks *T* and performance measure *P*, if its performance at tasks in *T*, as measured by *P*, improves with experience *E*,"
> (Mitchell 1997, p. 2)

This is an elaborated, perhaps overly complicated definition of how a computer program learns. It learns only when performance improves with experience. So, what is experience? It is the data that we provide for the machine to process.

Experience *E*, is the information the machine processes using the tasks *T*, to get closer to what you want to produce, the performance measure *P*. If the machine continues to perform the tasks but produces rubbish then it hasn't learned.

A machine can be said to learn if it can perform tasks on the information you supply, and react to the data to get results that get closer to a useful result over time.

Consider how humans learn and make decisions. We analyse data, find patterns and use those patterns to make decisions or to predict the outcome of future events. For example, if you buy a car, you might look up reviews, check the features of different models, think about how you want to use the car, take some test drives and then make a choice.

Too much data

So why do we need ML? It is all about the volume of data available to us. There is so much information that we cannot make sense of it. It's as though you have to choose the perfect vehicle for every person at Deakin. It would be a massive task to know what every person wants and needs and to add all the details of various models of vehicle in order to make good decisions.

Automated systems (ML) will learn from the data we give them and, importantly, respond to changes in the data they are given to produce useful results.

Activity

How would you define machine learning? Do a bit more reading and come up with your own definition.

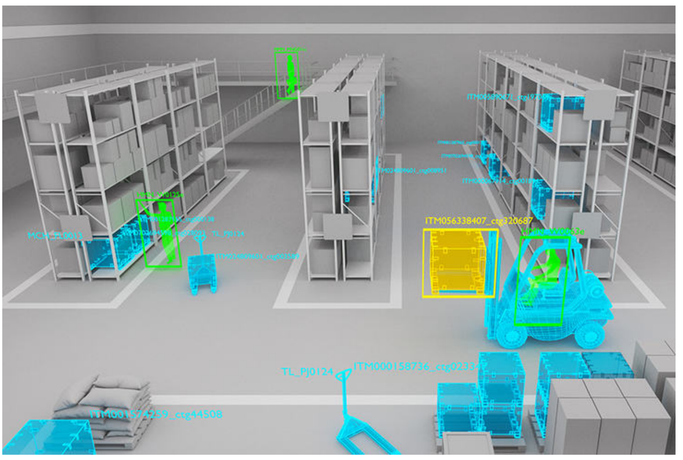
Share your response in the discussion forum

Readings

Samuel, A 1959 'Some Studies in Machine Learning Using the Game of Checkers', *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210–229.

Mitchell, T (1997), *Machine Learning*, McGraw Hill.

Wikipedia n.d., Machine learning, viewed 29 June 2018.



© Getty Images

Real-world applications of machine learning

Machine learning involves the use of computer algorithms to learn how to perform different tasks.

Let's look at a few examples where machine learning is being applied.

Robotics

Robots are going to be our companions in the future. They will do our cleaning, cooking and vacuuming. They'll do our reading, schedule tasks, remind us to take our medications and play our favourite song at just the right time and volume. Machine learning is a fundamental part of enabling robots to perform these activities.

Some examples of machine learning applications in robotics are creating walking and jumping patterns for humanoid robots and tuning the way a prosthetic leg moves in relation to the other leg of a one-legged patient.

Simultaneous Localization and Mapping (SLAM) uses data to find routes for rescue robots. Check out this amazing demonstration of real-time mapping using the pulsing light of LIDAR.

Computer vision

Board Games

Board games are one of the oldest applications of machine learning. There are many successful algorithms which play checkers and board games such as 'Go'.

In March 2016; AlphaGo, the board-game-playing AI from Google's DeepMind beat Korean Go Champion Lee Sedol 4-1. AlphaGo uses deep neural networks and Monte Carlo tree search.

Watch this 3 minute trailer for the AlphaGo documentary. The win is a defining moment in AI similar to when the computer, Deep Blue, beat Garry Kasparov at chess in 1996.

This is an additional video, hosted on YouTube.

Voice Recognition

Speech recognition and machine learning have experienced waves of major innovations through recent history. It has benefited from advances in deep learning, as well as big data and is widely used in a variety of applications.

Technologies such as Siri and Google Home are examples of this. Siri uses speech recognizer, natural language processing and text-to-speech techniques.

The Australian Government Tax Office now holds millions of voice prints to identify Australians.

> Their voice is compared with a stored voiceprint which captures more than 140 unique physical and behavioural characteristics of a person such as length of the vocal tract and nasal passage, the size and shape of the larynx, pitch, cadence and accent. (Nott, 2016)

How would you go at recognising people by voice? How about a million people? This is where machine learning excels.

Digit Recognition

Digit Recognition is the task of reading in the images of handwritten numbers and letters and outputting its machine-encoded equivalent. Machine Learning methods (SVM and Deep Learning) have hit >99% accuracy for this task.

Other examples are detection of numberplates, printed numbers on bills, handwriting and CAPTCHA.

Other interesting applications

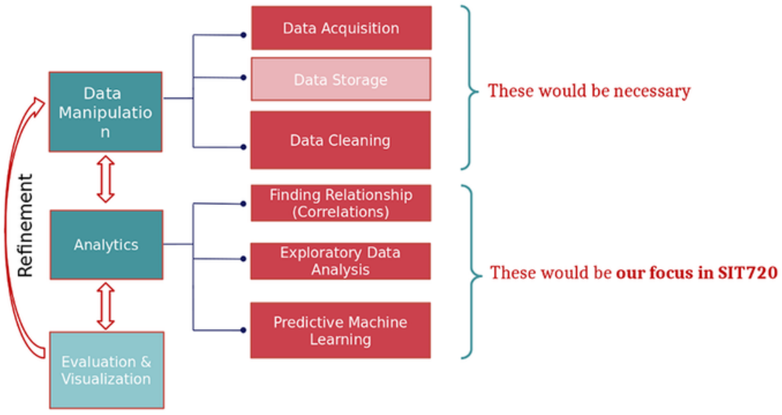
A wide variety of machine learning applications are related to:

- healthcare analytics: diagnosis and prognosis
- stock market prediction
- business analytics
- facial recognition

Activity

What do you think of these examples? Find another problem or application where machine learning could be used.

Share your discoveries in the discussion forum and respond to at least one other comment.



© Getty Images

Machine Learning steps

Have you ever thought about how you learn to make decisions?

Let's assume that we have been asked to research a problem and make a decision. You must have access to information (data) from which you can learn. Once you have data, what do you do with it? Most people analyse the information to find patterns and relationships. You make a decision based on your work. Finally, you evaluate the model. Did you make a good decision or could it have been better?

Let's look at each of these stages in more detail and relate them to ML.

Step 1: Data Manipulation

This is a process of data preparation. ML usually uses the largest sets of data available.

The first step in data manipulation is *Data Acquisition*. Data acquisition is the process of sampling information that illustrates real world physical conditions with a predefined measurement. Using our car example, you might measure engine size, number of doors, size of tyres etc.

The data acquired should be reliable for converting into digital numeric values that can be manipulated by a computer. Number of car doors is easy; style is less easy to define numerically.

Once the data is properly stored, any redundant, noisy, unusable parts of it should be trimmed. We call this procedure *Data Cleaning*. Data cleaning is a major step as real-world datasets are highly affected by noise, redundancy and missing values. We might delete any three-wheeled cars because they're so unusual.

Step 2: Analytics

The second main step in machine learning is *Analytics*. Analytics mainly involves *finding relationships and correlations* in the prepared data in order to design an accurate model based on that input data.

In addition, *Exploratory Data Analysis* is an approach for analysing datasets in order to summarise their main characteristics or features. Many exploratory data analysis methods use visual illustration of data, based on different features. Things like graphs, charts and tables make data easier to understand.

Finally *Predictive Machine Learning* is the last stage of Analytics. It uses a variety of statistical techniques such as predictive modelling in order to build a classifier or intelligent system for decision making. We will come back to these ideas over the next few weeks.

Step 3: Evaluation and Visualisation

The final result of Analytics is an intelligent system or model. As the last step of ML design, we have to evaluate the performance of the system. "Did I choose the right car?"

If the quality and performance of the intelligent system does not achieve a satisfactory outcome, the *Refinement* procedure is required and another round of data manipulation and analytics becomes necessary. Again, we will come back to these ideas as you move through the Unit.

In this Unit we mainly focus on the *Analytics* step. However it will be necessary to use parts of the *Data Manipulation* step to build a model.

Activity

If you would like another explanation of the scope and complexity possible and the methods used to make the machine learn, watch this 10 minute video from Android Authority.

It goes over some of the concepts we've covered so far, and introduces others.

Tip: if you watch it at twice the speed, it takes half as long!

This is an additional video, hosted on YouTube.

Does this change your own ideas about ML? Student Discussion your thoughts.
