City University Of London

MSc in Artificial Intelligence

Project Report

2021/22

Ablation Study on Faster-RCNN for Hepatocellular Carcinoma Detection

Benjamin Dwumah

Supervised by: Dr. Alex Ter-Sarkisov

December 21, 2022

**Declaration**

By submitting this work, I declare that this work is entirely my own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the assessment instructions and any other relevant programme and module documentation. In submitting this work I acknowledge that I have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. I also acknowledge that this work will be subject to a variety of checks for academic misconduct.

Signed: Benjamin Dwumah

**ABSTRACT**

Hepatocellular carcinoma (HCC) is one of the most prevalent causes of cancer incidences and deaths. Despite many years of research and the creation of new medical interventions, patients with HCC continue to have poor treatment outcomes. Patients with HCC suffer from unfulfilled concerns like risk prediction, individualised treatments, accurate prognosis and early diagnosis. In recent years, there has been a massive growth in Artificial Intelligence (AI) applications in medical research, and the field of HCC is no exception. Deep learning algorithms are among the most advanced AI-based machine learning algorithms for processing and analysing complicated multimodal data, from routine diagnostic factors to high-resolution medical images. In this research project, I present my experiment results and review for early diagnosis of HCC using deep learning techniques, specifically Computer Vision. I have done detailed experiments on the object detection model Faster Region-Based Convolutional Neural Network (Faster R-CNN) for detecting HCC. I experimented with different backbones for the Faster R-CNN model. I concluded that the backbone plays a significant role in the Faster-RCNN architecture for good accuracy results and performance. The codebase is available at: <https://github.com/Ben74x/Indiv_Proj>

**Keywords:** Hepatocellular carcinoma (HCC), Artificial intelligence, Deep learning, Computer Vision, Object Detection.

**1. Introduction**

Hepatocellular carcinoma (HCC) is an aggressive primary liver cancer that develops in the setting of chronic parenchymal liver diseases and is among the top causes of cancer incidence and mortality worldwide (Bray et al., 2018; Yang et al., 2019). While the hardship of having HCC has decreased with effective antiviral therapy, HCC cases associated with metabolic syndrome are expected to rise further due to the significant increase in the commonnes of non-alcoholic fatty liver disease (NAFLD) in the general population (Stepanova et al., 2017).

Decades of research in HCC have resulted in the development of a screening protocol, non-invasive imaging-based diagnostic modalities, and different treatment modalities, including surgical, locoregional and systemic therapies (Llovet et al., 2021; Yang and Heimbach, 2020). Nevertheless, the results of patients with HCC continue to be poor. There are areas of critical unmet demand in early detection, accurate prognostication, risk prediction, and individualized treatments.

A lot of health data is produced by HCC patients. While this is exciting for researchers, guaranteeing that such large amounts of data are converted into actionable insights can be difficult. Artificial intelligence (AI) is considered to be capable of synthesising and analysing multimodal data with extraordinary levels of accuracy, and the use of AI to several areas of medicine, including hepatology, has grown rapidly in recent years (Ahn et al., 2021). For a wide range of tasks and clinical applications, such as image classification, detection and segmentation, etc, AI-based concepts offer a variety of techniques. Recent developments in AI, specifically in the field of medical image analysis, provide a vast array of automated tools for obtaining precise measurements of biomarkers, exposing delicate features, categorising tissue characteristics, and conducting radiomics for in-depth analysis of raw imaging data. The introduction of deep learning techniques have made the AI revolution of the past ten years conceivable. This research analyses how the object detection model model, Faster R-CNN, performs in the detection of HCC lesions.

The main question for this research study is: *“Can customized Faster R-CNN models detect HCC lesions better than the base model?”*

**1.1 Objectives, Project Product, and Beneficiaries**

In this project, my goal is to research, understand, and experiment the Faster R-CNN model to see if it is appropriate for use in the detection of HCC by contrasting the base model with modified versions. The project's final product is a fully operational Faster-RCNN model that has been trained on data from liver cancer patients with HCC and can be used in real-world scenarios.

The following are the project deliverables:

* Extensive research on the Faster R-CNN model for HCC lesion detection
* Results of using the top-performing model to find HCC lesions in ultrasound scans

By incorporating the top performing algorithm in this project into their applications, researchers from many fields of study, as well as medical professionals working in the field of cancer, can gain from this project. Additionally, this project offers thorough information concerning the effectiveness of the Faster RCNN algorithm and its contributing factors to the performance of HCC detection, which can aid in future research. Furthermore, anyone with a foundation in machine learning will also gain from developing their knowledge of deep learning, particularly computer vision.

**1.2 Structure of the Project Report**

In Section 2, I discuss Artificial Intelligence and its applications in the health sector. I also detail machine learning, its types and limitations. I then discuss deep learning and give a brief overview of the computer vision and its applications in the health sector. Finally, I describe the concept of object detection and talk about the Faster R-CNN algorithm discussing all the essential parameters.

In Section 3, I discuss the method of experiments involving the data for the project, the model, other external tools and evaluation metrics.

In Section 4, my findings from the experiments are presented along with graphs, diagrams and figures.

In Section 5 and Section 6, I introduce the project results and reevaluate the research question with regard to the experiments conducted. I then discuss the results and conclude my project report.

**2. Context**

**2.1 Artificial Intelligence**

**2.1.1. Definition**

In a 2004 study, John McCarthy gave the following definition of artificial intelligence (AI), despite the fact that there have been numerous other definitions over the past few decades, "It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable" (McCarthy 2004, p. 2).

Turing, known as the "Father of Computer Science," posed the question, "Can machines think?" in this paper. From there, he proposed a test which is now popularly known as the "Turing Test". In the test, a human interrogator attempts to differentiate between a computer and a text response from a human. While this test has been heavily scrutinised since its publication, it remains an essential part of the history of AI in addition to an ongoing concept within philosophy due to its use of linguistic ideas.

In recent times, a book by Peter Novig and Stuart Russell, Artificial Intelligence: A Modern Approach has become one of the leading learning material in the study of AI. In it, they explore four potential objectives or definitions of AI, differentiating between computer systems based on their reasoning and thinking vs acting:

* Human approach:
  + Systems that think like humans
  + Systems that act like humans
* Ideal approach:
  + Systems that think rationally
  + Systems that act rationally

Comparing Alan Turing’s definition to Peter Novig and Stuart Russell own, his definition would fall under the human approach which is systems that act like humans.

In its basic form, AI can be defined as an area incoporating robust datasets and computer science to solve problems. It has the subfields of machine learning and deep learning, which are widely discussed in the context of artificial intelligence.

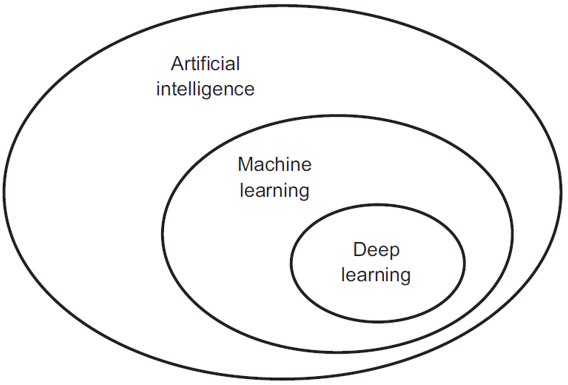


Fig. 1: Aritificial Intelligence and its subfields

AI is grouped into two types which are Weak AI and Strong AI. Weak AI, known as Narrow AI, has been trained and focused on performing specific tasks. The majority of AI applications today are driven by weak AI. The type of AI is anything but weak; it powers many applications like self-driving cars, voice assistants, personalized marketing, facial recognition systems, and gamified therapy. Stong AI comprises Artificial Super Intelligence (ASI) and Artificial General Intelligence (AGI). A computer with an intellect comparable to humans, a self-aware awareness, and the capacity to learn, reason, and make plans for the future would be said to have AGI, also known as general AI. ASI, also known as Superintelligence, would be more intelligent and capable than the human brain. Even though there are now no real-world applications for strong AI and it is only theoretical, experts in the field of artificial intelligence are continuously studying its potential.

**2.1.2. AI APPLICATIONS IN HEALTHCARE**

Healthcare delivery may change as a result of AI. It can boost output and care delivery efficiency, enabling healthcare systems to serve more people with more effective treatment. AI can assist healthcare professionals in having a better experience, permitting them to spend more time providing direct patient care and lowering burnout. One of the biggest success stories in our time is healthcare. Life expectancy has increased globally due to significant advancements in medical technology. However, as people live longer, healthcare systems must contend with expanding patient demand, rising expenditures, and a stretched-thin staff. Population ageing, shifting patient expectations, a change in lifestyle preferences, and the endless cycle of innovation are just a few of the inescapable drivers that drive demand. The effects of an ageing population stand out among these. One-fourth of the people in North America and Europe will be over 65 by 2050, which means the healthcare systems will be required to cope with more patients with complicated demands. It is expensive to manage these patients, and systems must change from a philosophy of periodical care to one that is considerably more proactive and centred on long-term patient care.

The expenditure on healthcare is growing. Healthcare systems will have difficulty staying sustainable unless significant structural and innovative changes are made. Health systems require a larger workforce as well. However, while the world economy could generate 40 million new healthcare jobs by 2030, the World Health Organization predicts a 9.9 million physician, nurse, and midwife shortage over the same time period. Not only must we attract, train, and sustain more medical professionals, but we must also ensure that their time is spent where it adds the most value, which is caring for patients. AI, which is based on automation, has the potential to revolutionise healthcare and assist in addressing some of the issues raised above. AI can improve care outcomes as well as the efficiency and effectiveness of care delivery. It can also improve healthcare practitioners' daily lives by allowing them to spend more time caring for patients, increasing staff morale and retention. It can even help bring life-saving treatment methods to market faster. Simultaneously, concerns have been expressed about the influence AI may have on patient populations, professionals, and health systems, as well as the risks involved; there are ethical debates about how AI should be used.

According to Spatharou, Hieronimus and Jenkins (2020), a growing number of governments have set forth goals for AI in healthcare, and several are making significant investments in the field. Venture capital (VC) expenditure for the top 50 companies in healthcare-related AI reached $8.5 billion, and huge tech companies, startups, pharmaceutical and medical device companies are all involved in the developing AI healthcare ecosystem.

**2.2. Machine Learning**

**2.2.1 Definition**

Machine learning is a subset of artificial intelligence. Machine learning focuses on data driven learning, whereas artificial intelligence focuses on general intelligent behaviour. Machine learning is the process by which a computer learns from data. The program takes in data, and from the data, the program learns. In the process, the program creates a model, which can then be used to make predictions about future data. The predicted values are usually, but not always, probabilistic in nature. That is, the model is a mathematical model which gives probabilities for events in the data. However, the model is only sometimes probabilistic. For example, decision trees are not probabilistic, but they are a model of the data.

A more detailed definition of machine learning would be “A computer program is said to learn from experience E with respect to some class T and performance P, if its performance at tasks in T, as measured by P, improves with experience E” (Mitchell 1997, p. 2).

Machine learning is a vast and vital field of study. The practical significance of machine learning cannot be overstated. Many of the applications we use every day are machine learning applications. Email apps, for example, use machine learning to distinguish between spam and legitimate emails. It is also used in search engines to prioritise web pages, in speech recognition applications to recognise speech, and in image recognition applications to recognise images.

**2.2.2 Types of Machine Learning**

Machine learning algorithms can solve a wide range of problems. Machine learning algorithms are currently trained using four distinguished methods: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.

Machine learning algorithms can be trained in a variety of ways, each with its own benefits and drawbacks, like any other method. We must first examine the types of data that each type of machine learning consumes in order to comprehend the benefits and drawbacks of each type. Labeled data and unlabeled data are the two types of data used in machine learning. Although labelled data has both the input and output parameters in a properly machine-readable form, labelling the data initially takes a significant amount of human effort. In unlabelled data, only one or none of the parameters are present in machine-readable form. This eliminates the need for human work but calls for more complex solutions. Furthermore, certain machine learning algorithms have extremely specific applications; yet, the four primary approaches are still in use today.

Supervised learning is the most prevalent kind of machine learning. In supervised learning, the algorithm analyses labelled data. In other words, the data has some marking, and for each input, the appropriate output is given to the application. Numerous applications include supervised learning. Email applications, for instance, employ supervised learning to categorise emails as spam or not. Applications for image recognition identify images using supervised learning. Applications for speech recognition classify sounds using supervised learning. Constructing a model that links inputs to outputs is the aim of supervised learning.

Learning from unlabeled data is referred to as unsupervised learning. The algorithm gathers information and makes learning from it. However, the algorithm is not informed of the proper output. The algorithm understands relationships between data points in an abstract way; human input is not necessary. Unsupervised learning methods are flexible because of the development of these hidden structures. Unsupervised learning algorithms can modify their underlying structures dynamically to respond to the data rather than using a predetermined and stated problem statement. This provides more post-deployment development than supervised learning techniques. There are numerous uses for unsupervised learning. Unsupervised learning may be used by an email application, for instance, to group spam emails.

Machine learning techniques such as semi-supervised learning allow the algorithm to learn from labelled and unlabeled input. The use of semi-supervised learning is widespread. Semi-supervised learning, for instance, might be used by an email application to categorise emails as spam or not. The algorithm may use unlabeled data to identify spam email clusters. Then, it might classify emails as spam or not using labelled data. Better models can be learned via semi-supervised learning than with supervised learning alone.

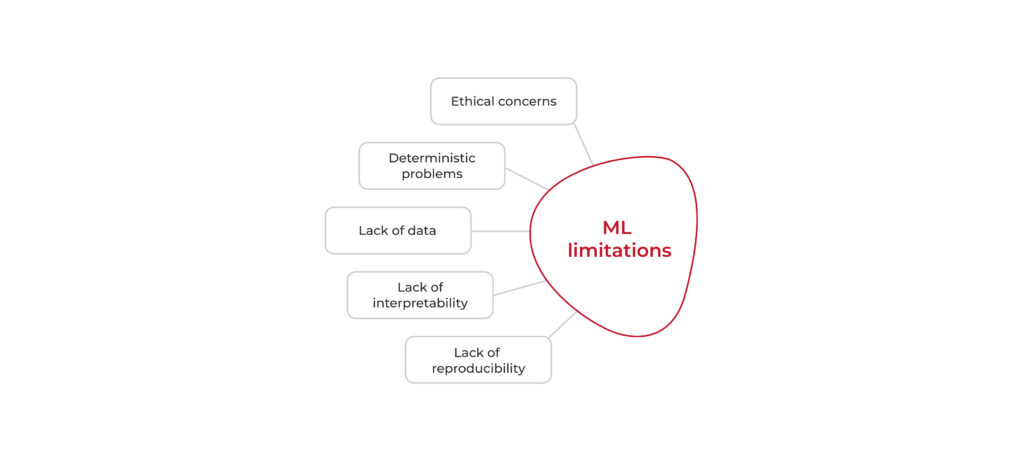
Reinforcement learning is a subset of machine learning in which the algorithm learns through "trial and error" experiments. The algorithm starts with an initial state and then acts. The algorithm is then rewarded or punished for its actions. The process is then repeated, beginning with the initial state until the algorithm has explored the space of possible actions. Reinforcement learning is used in a wide range of applications. For example, a chess programme might use reinforcement learning to teach itself how to play the game. The goal of reinforcement learning is to learn a policy that determines each state's best course of action.

**2.2.3 Machine Learning Limitations**

Developers of machine learning systems have been able to make these systems think more like humans in recent years, executing difficult tasks and coming to choices after doing in-depth analysis. Robots doing a variety of professions for humans is not a sci-fi movie premise anymore; it is a reality. However, there are still a number of machine learning algorithms' limitations, despite the advancements ML developing teams have achieved in this area.

Although machine learning is incredibly helpful for many applications, it isn't always the best option. Implementing machine learning applications isn't always essential, it's not always a good idea, and it sometimes even makes things worse. The world has been greatly touched by ML. According to Harari (2015) in his famous book HOMO DEUS, we are gradually moving towards an ideology known as dataism, which means individuals will embrace data and algorithms more than their own personal convictions. There has been cases already on this ideology on how people trust GPS instructions to take them to their destinations. People have occasionally hit roadblocks after blindly following a navigation device's recommendations without once consulting a map or the people around them.

When developing a project that needs to process a lot of data, machine learning offers a creative way of accomplishing it. But one thing we must take note of is that, there are crucial factors that should be taken into account before deciding to use ML as a tool to build applications. One needs to be aware of the potential drawbacks and limitations of this formidable technology before putting it into use. There are five basic areas in which ML issues can be categorized as seen in the figure below.

Fig. 2: Limitations of Machine Learning

* **Ethical Concerns**: There are many benefits in using machine learning algorithms. The use of these algorithms to automate tasks, evaluate vast volumes of data, and make complex decisions has helped humanity a lot. However, these algorithms does have certain disadvantages. Bias can exist in these algorithms at any stage of development and bias cannot be completely elimanted because these algoithms are created and trained by humans. Many ethical issues are still unresolved. Taking self-driving cars as an example, who is responsible if something goes wrong. In the event of an accident who should be blamed – the driver, the automobile company or the software. One thing is certain: ML cannot independently decide on challenging moral or ethical issues. Measures have been put in place to takcle this issue with the introduction of explainable AI. However, we will need to develop a a complete system in the near future to address ethical issues with machine learning technology.
* **Deterministic Problems**: Machine learning is an advanced tool which is used in a variety of fields. For example, in the meteorology field, machine learning algorithms can assist in the calibration and correction of sensors to measure pressure, temperature and humidity for the environment. Also, models can be built to simulate atmospheric emissions to predict pollution. Although machine learning algorithms are able to do such complex tasks, they are unable to comprehend the physics of a weather system. Machine learning models are capable of making these predictions but the computations of intermediate fields like density may result in negative values that defy the laws of physics. Machine learning is not capable of understanding the cause and effect relationships. Although, a neural network can connect input and output data, it cannot determine why there are related.
* **Lack of data**: Given the complex arhictectures of machine learning algorithms, they usually need a lot of data for training in order to function well. The amount of data needed by ML models increases with its complex architectures. Some could choose to reuse data when training these algorithms but this never produce satisfactory outcomes. Furthermore, the absence of good data is a further issue. This is distinct from merely lacking data. Consider a scenario in which your neural network needs additional data and you provide it with a sufficient amount of low-quality input. The accuracy of the model may be severely hampered as a result.
* **Lack of interpretability**: Interpretability is a significant issue with machine learning algorithms. Consider the situation where you are building a model for a financial company to identify fraudulent transactions. The model has to be able to defend how it categorizes transactions in this situtation. For a task like this, a machine learning application might do well in terms of accuracy and responsiveness, but it might not be able to prove its outcomes. If machine learning techniques are to be used in practice, it is crucial that they become interpretable.
* **Lack of reproducibility**: A complex and explanding problem in machine learning called lack of reprocibility is made worse by lack of model testing procedures and transparency in code. New models are created and are quickly implemented in practical applications. Nevertheless, despite the fact that the models are created to incorporate the most recent scientific advancements, they might not function in actual situations. Various industries and professions can use reproducibility to use the same model and find solutions to issues more quickly. Safety, dependability, and the capacity to spot bias can all be impacted by a lack of reproducibility.

There is no denying that AI has provided humans with a number of exciting new opportunities. Some have, however, also come to believe that machine learning applications are capable of resolving any issue facing mankind. It is optimal for machine learning systems to be used on tasks that would normally be completed by humans. If you don't ask it to be imaginative, intuitive, or utilize common sense, it can do well. Machine learning applications are capable of learning from concrete data quiet effectively, but they lack the human capacity to comprehend the world and how it works. An ML application, for instance, can be trained to know what a cup looks like but it is unable to comprehend that the cup contains coffee. Machine learning algorithms make people's life better and their work more efficient, but it cannot completely replace them because it is not capable of performing many responsibilities. While ML has several benefits, there are also some drawbacks.

**2.3 Deep Learning**

**2.3.1 Definition**

Deep learning is a subset of machine learning concerned with creating and implementing algorithms and models capable of processing and analyzing vast volumes of data. Before deep learning, we depended on classic machine learning techniques such as logistic regression, support vector machine (SVM), Bayes classifier, decision trees, etc. These classic techniques are also referred to as flat algorithms. Normally, a preprocessing stage called feature extraction is required when using these classic techniques. These traditional machine learning techniques can employ the depiction of the given raw data created by feature extraction to complete a task. For instance, we can now divide the data into classes. Feature extraction is typically a difficult process that calls for in-depth understanding of the problem area. For best results, this preprocessing layer needs to be modified, examined, and improved across numerous iterations. In deep learning, the feature extraction stage is not needed. The layers have the ability to independently learn the underlying depiction of the raw data.

Understanding that deep learning is fueled by enormous amounts of data is essential to understanding why it has gained such popularity. The evolution of big data has opened up several possibilities for deep learning advancements. According to an article by Garling, 2015, Andrew NG explained AI in these terms: “*I think AI is akin to building a rocket ship. You need a huge engine and a lot of fuel. If you have a large engine and a tiny amount of fuel, you won’t make it to orbit. If you have a tiny engine and a ton of fuel, you can’t even lift off. To build a rocket you need a huge engine and a lot of fuel.*”

Using deep learning to summarize this one can say that deep learning algortihms are the rocket and data is the fuel.

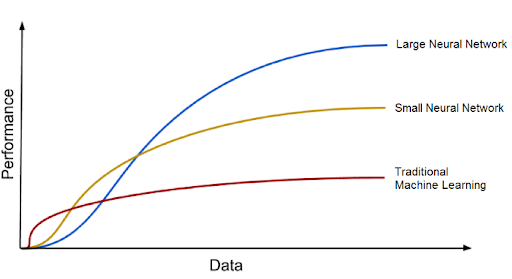


Fig. 3: Deep learning models against data

Deep learning methods continually examine data using a predetermined logical structure in an effort to reach conclusions that are comparable to those reached by humans. This is done by using neural networks, which is a multi-layered architecture of algorithms.

**2.3.2 Origin of Neural Networks**

A brain’s biological neurons serve as a model for artificial neural networks (ANNs). In simplified terms, artificial neural networks mimic several fundamental biological neural network functions. In order to draw comparisons between natural and artificial neural networks, let's first examine biological neural networks. A biological neural network is made up of many neurons.

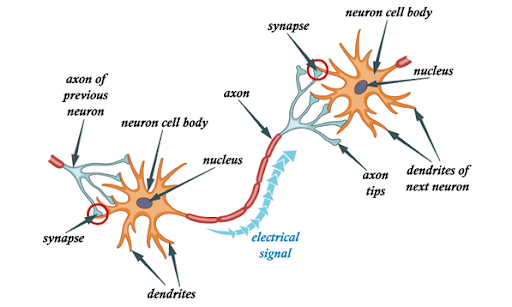
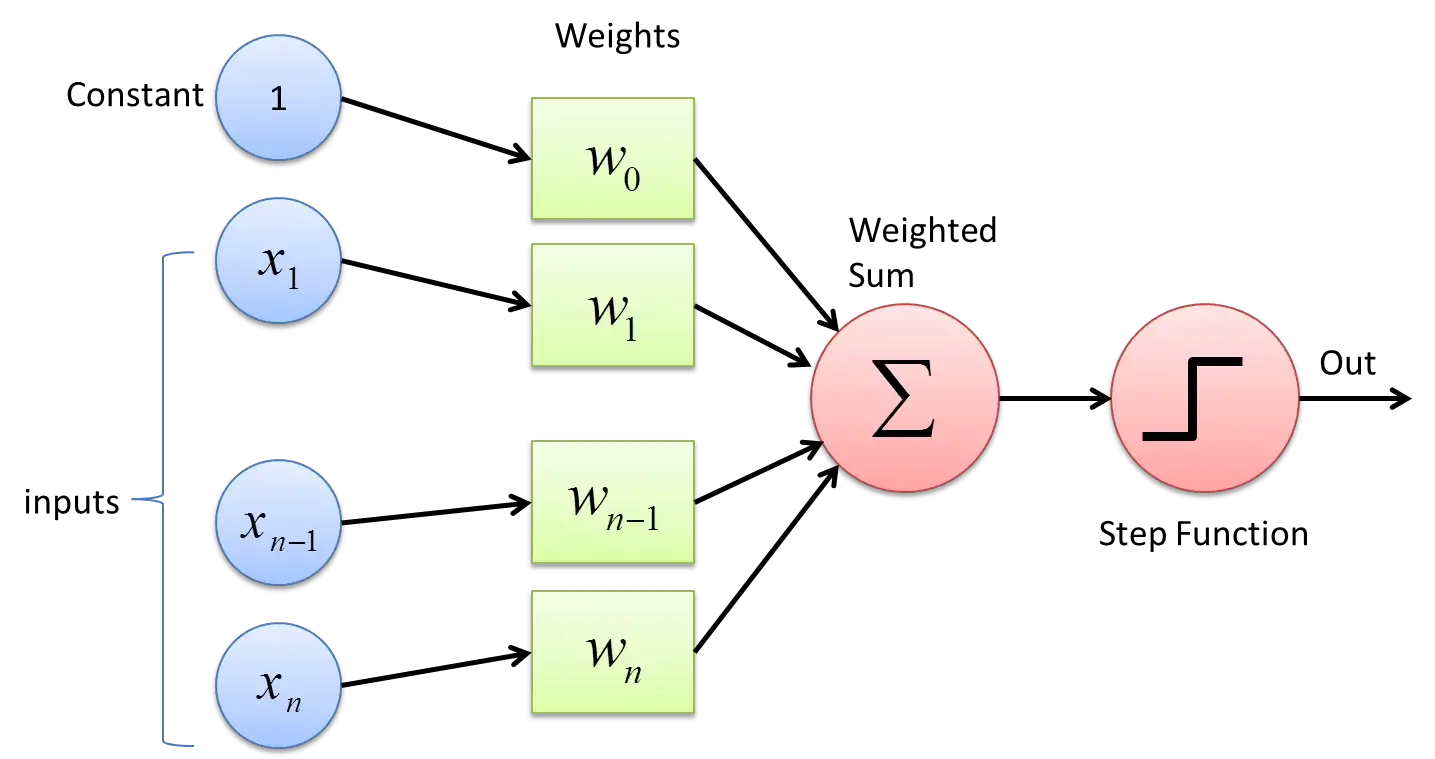
****

Fig. 4: A biological neural network model

A neuron consists of a cell body, dendrites and an axon (Woodruff, 2018). Dendrites are thin structures that protrude from the cell body. Axon is also a cellular extension that emerges from the cell body. The majority of neurons send impulses via the axon and pick up signals with the dendrites. Signals travel from one neuron’s axon to another neuron’s dendrite at the large percentage of synapses. Since voltage gradients are kept in place in the membranes of all neurons, all neurons are electronically excitable. The neuron produces an electrochemical pulse known as an action potential if the voltage fluctuates by a significant enough amount over a brief period of time. Rapidly moving down the axon, this potential triggers synaptic connections.

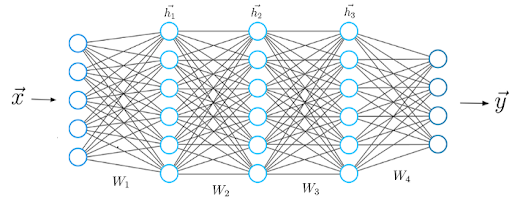
**2.3.3 Perceptron**

A Perceptron is a supervised learning algorithm for binary classifiers (DeepAI, 2019). Binary classifiers determine whether an input, which is often represented by a collection of vectors, falls into a certain category. In summary, a perceptron is a single-layer neural network which consists of input values, weights and bias, net sum, an an activation function (DeepAI, 2019). Perceptron works by multiplying all of the input values by their weights. The weighted sum is then calculated by adding the sum of all these multiplied values. The output of the perceptron is then generated by applying the weighted sum to the activation function. The vital job of the activation function is to make sure that the output is transferred between necessary values like (0,1) or (-1,1).

Fig. 5: Perceptron

**2.3.4 Artificial Neural Networks**

A neural network typically consists of a group of linked nodes. These nodes are known as neurons. These synthetic neurons are based in part on the actual brain cells as mentioned in 2.3.2. In deep learning, a neuron is nothing more than a graphichal depiction of numerical values. For a biological neuron, an axon is any link between two neurons. In deep learing, these connections are called weights, which are numerical values, serve to represent the link between the artificial neurons. The design of the human brain served as the inspiration for the neural network's architecture. We can train neural networks to recognize patterns and categorize various types of information, just like how our brains do it naturally. The probability of detecting and producing a proper result is increased by using the different layers of neural networks as a kind of filter that operates from coarse to fine. Like this, the human brain functions. The brain seeks to make comparisons with familiar objects whenever we get new information. Deep neural networks also make use of the same idea.

Fig. 5: Aritificial neural network

**2.3.5 Feedforward Networks**

A feed forward neural network is a type of ANN in which there is no recurrence in the connections between the nodes. Since input is only processed in one channel, the feed forward model is the most straightforward type of neural network. Although the data may flow via several buried nodes, it always proceeds forward and never backward. A single layer perceptron is a common example of a feed-forward neural network in its most basic configuration. A number of inputs are introduced into the layer in this model and multiplied by the weights. The weighted input values are then summed together to produce a total. The value produced is frequently 1, and if the sum of the values is below the cutoff, the output value is -1. The threshold is typically set at zero. In classification tasks, the single layer perceptron is a crucial feed forward neural network model. Single layer perceptrons can also contain some features of machine learning. The neural network may compare the outputs of its nodes with the desired values using a property known as the delta rule, which enables the network to train its weights to create output values that are more accurate. This learning and training procedure results in a gradient descent. Although the procedure for changing weights in multi-layered perceptrons is almost comparable, it is more formally known as back-propagation. In these circumstances, the network's hidden layers are each changed in accordance with the output values generated by the final layer.

From Fig. 6, we can see the architecture of a feedforward neural network. The input layer is supplied with input ***x***, which is data from which the neural network learns. The neurons in the input layer is equal to the number of elements in the vector ***x***. To simplify it, every input neuron corresponds to a single vector element. The output layer, the final layer, generates a vector ***y*** that represents the outcome of the neural network. The values of the neurons in the output layer are represented by the elements in this vector. The network must carry out specific mathematical computations in the layers in between input and output layers in order to generate a prediction vector ***y***. These are called hidden layers.

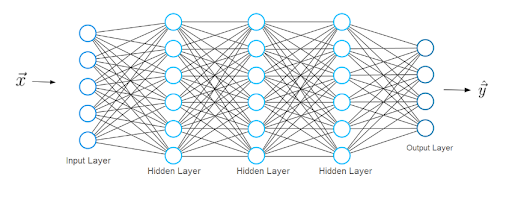


Fig. 6: Feedforward neural network

We can properly investigate the process of learning now that we have a better understanding of neural network architecture. Let's go cautiously. The neural network generates a prediction vector, that we'll refer to as ***h***, for an input feature vector called ***x***.

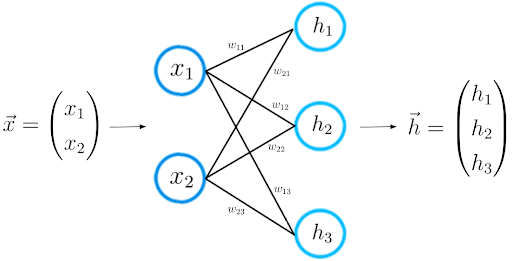


Fig. 7: Structure of forward propagation

This process is also known as step forward propagation. We calculate the dot product between the input vector ***x*** and the weight matrix ***W***, which connects the two neuron layers. The outcome vector, that we refer to as ***z***, is produced by this dot product.