Report 1

Understanding Deep Learning:

Breaking Down the Black Box and Shedding an Intuitive Light

Author: Faizan Khalid Mohsin

Course: CHL5250HY - Special Topics in Biostatistics

Professor: Wendy Lou.

March 24, 2021

Table of Contents

[Abstract 3](#_Toc68953424)

[Introduction 4](#_Toc68953425)

[Methods 8](#_Toc68953426)

[Theoretical Framework of Neural Networks 8](#_Toc68953427)

[Simplest Neural Network: Simple Feedforward Shallow Neural Network 8](#_Toc68953428)

[Statistics, Probability, and Information Theory in Neural Networks. 10](#_Toc68953429)

[Neural Networks Structure, Terminology, and Mathematical Representation. 10](#_Toc68953430)

[Regression losses 15](#_Toc68953431)

[### Metrics to evaluate the performance of an algorithm 16](#_Toc68953432)

[### Optimizer 17](#_Toc68953433)

[Activation Functions 17](#_Toc68953434)

[Gradient Descent & Stochastic Gradient Descent 20](#_Toc68953435)

[Back Propagation 22](#_Toc68953436)

[Hyperparameters of neural networks 22](#_Toc68953437)

[Regularization: Methods to Improve Neural Network Performance 24](#_Toc68953438)

[The Idea of Regularization: 24](#_Toc68953439)

[Parameter Norm Penalties 24](#_Toc68953440)

[Ensemble Methods 25](#_Toc68953441)

[Data Augmentation 26](#_Toc68953442)

[Semi-Supervised Learning: 26](#_Toc68953443)

[Dropout 27](#_Toc68953444)

[# Application of Neural Networks with Applications 28](#_Toc68953445)

[## Setting up the software. 28](#_Toc68953446)

[### Installing Python 28](#_Toc68953447)

[### Installing Tensorflow 28](#_Toc68953448)

[### Installing Keras 28](#_Toc68953449)

[## Neural Networks to Predict Popularity of the Ted Talk. 29](#_Toc68953450)

[### Data 29](#_Toc68953451)

[### Univariable Analysis 29](#_Toc68953452)

[### Multivariable Analysis 29](#_Toc68953453)

[### List of issues in Neural Network Application (also applicable to machine learning). 29](#_Toc68953454)

[Creating Images to Represent Neural Network Architectures. 30](#_Toc68953455)

[# Results and Insights: Deep Neural Network 31](#_Toc68953456)

[Discussion and Conclusion 34](#_Toc68953457)

[References 34](#_Toc68953458)

[Appendix 35](#_Toc68953459)

[Definitions and Further Explanations. 35](#_Toc68953460)

[Tensor: 35](#_Toc68953461)

[Derivations: 35](#_Toc68953462)

[Deriving deep learning loss functions using Maximum Likelihood method. 35](#_Toc68953463)

[SDG Expectation 35](#_Toc68953464)

[Derivation showing that the L2 penalty causes the weights to shrink. 36](#_Toc68953465)

[Potential Research Idea 37](#_Toc68953466)

[List of pandas dataframe methods. 37](#_Toc68953467)

[## Questions 39](#_Toc68953468)

# Abstract

**Goal:** We aim to answer why deep learning is one of the most successful and popular AI algorithms. We present in depth exactly what deep learning is by introducing the mathematical foundation and framework on which it is built, as well as provide an intuitive understanding of how deep learning and its different components work. Finally, we demonstrate and explain the several advantages of deep learning. **Overview:** Deep learning is a type of neural network; to be precise, it is a deep neural network. Most components of neural networks can be represented as vectors and matrices, and many calculations can be represented as matrix multiplications or by taking the expectation of different mathematical objects and operators. **Methods:** We discuss the different important components of neural networks: artificial neurons, layers, perceptrons, activation functions, cost functions, backpropagation, (stochastic) gradient descent, and minibatches. **Advantages:** (1) The performance of deep neural networks consistently continues to improve as more and more data is provided for it to train on. This is in contrast with many statistical and other machine learning models whose performance plateaus after a particular data (sample) size is reached and the performance is, in a sense, capped by some upper theoretical bound. (2) Further, it is one of the most versatile algorithms capable of tackling many different problems, as well as a very wide variety of data types such as sound data, image data, natural language, and structured data. (3) Lastly, one of the most amazing properties of deep neural networks is the ability to self-select features - and to do so directly from the raw data it is provided. **Conclusion:** These collective advantages found in deep learning is what overwhelmingly makes it the most successful and powerful artificial intelligence algorithm.

# Introduction

What is deep learning?

The short answer is that it is a type of neural network, more specifically a deep neural network. Neural networks (NN) are a type of Machine Learning algorithm and machine learning (ML) itself is a subfield of artificial intelligence. Machine learning is divided into three main branches: supervised learning, unsupervised learning, and reinforcement learning. Deep learning (DL) is a special approach in machine learning which covers all three branches and seeks also to expand them to address other problems in artificial intelligence which are not usually included in machine learning such as knowledge representation, reasoning, and planning among others (Skansi, 2018, pg 51). In this paper, we will only cover the supervised learning branch of deep learning.

Deep learning employs four main fields: Mathematics, Probability and Statistics, Information Theory, and Computer Science (Programming) (Skansi, 2018, pg 41). The idea of a neural network originates from biology and neuroscience mimicking how computation and thought processes of the brain work. The biological neural network consists of a population of interconnected neurons.

The neuron is a cell that can be electrically excitable and is capable of communicating with other cells and neurons. The neuron consists of the cell-body (soma), - which contains the nucellus; dendrites, - which receive signals (neurotransmissions and electrochemical stimulation) from other cells; and lastly, a single axon, - which transmits signals to other cells and can extend as far as 1 meter in humans. See Fig. 1 (Top) illustrating the different components of a biological neural in detail. The specialized connections through which the neurons are connected, and signals are transmitted are called synapses.

This idea of receiving signals, processing them, and further transmitting them is adopted into the concept of artificial neurons through inputs and outputs as illustrated in Fig. 1 (Bottom) which also shows the parallels between a biological and artificial neuron.



![ Fig.1: Comparing a biological neuron (Top) with artificial neuron (Bottom) and their components.](images\_graphresults/biological\_and\_ai\_neuron\_image.png)

This establishes the concept of an artificial neuron.

These artificial neurons are connected to form an artificial neural network. Unlike the biological neural network, which is very complex with many intricate connections, in artificial neural networks, the neurons are connected in a set pattern.

In an artificial neural network, the neurons are organized in layers. Each layer has several neurons. The neurons within each layer are not connect to one another. They are only connected to neurons in the layer before or after it.

The first layer in the artificial neural network serves a very special function: it takes in inputs and is hence called the input layer. There are as many neurons as there are different types of inputs. The last layer of the NN is called the output layer as it produces the final desired outcomes of interest. The number of neurons in the output layer depends on the types and number of outcomes desired. This can be very simple such as a binary outcome indicating if the input image contains a cat or not, thus will only have one neuron in the output layer which will produce either a 0 or a 1.

The layers in between the input and output layers are called hidden layers. They are the layers that perform the main calculations including feature selection of the data and patterns recognition to best predict the outcomes, or best match the pattern of the data provided to the NN. There can be any number of hidden layers in a NN.

To summarize, a population of artificial neurons connected to one another through three main types of layer:s the input layer, the hidden layers, and the output layer is what forms the basic structure of an artificial neural network and is inspired from the biological NN structure of the brain. See Fig. 2 below.



![ Fig. 2: Artificial Neural Network.](images\_graphresults/Artificial\_neural\_network\_image.png)