Over the past few decades, neural networks modelling has been considered as one of the most powerful tools, and its ability to handle a huge amount of data made it very popular in the literature. Having deep hidden layers in the models has recently become an interest that has started to surpass classical methods performance in many fields, especially in pattern recognition (Mohammed and Al-Zawi, 2017).

Deep learning models have become popular in the bioinformatics field. Singh et al. (2016) used a unified CNN framework that automatically learns combinatorial interactions among histone modification marks to predict gene expression. Qi et al. (2012) used a deep multilayer perceptron (MLP) architecture with multitask learning to perform sequence-based protein structure prediction.

By using neural networks, the idea of setting up a lightly parameterised function shaped by human can be forgotten. Instead, it allows us to set up a highly parameterised function that is very flexible and will be conveniently shaped during the learning phase. To put it simply, a deep learning model automatically learns complex functions that map inputs to outputs and rules out the need to use hand-crafted features (Singh et al., 2016). Since the input data of this project has a very large feature dimension, neural network is considered as one of the suitable approaches.

The idea of neural network came from the most interesting organ in the human body, the brain. The human brain is made up of billions of basic units called neurons. Figure 1 illustrates the basic neuron unit. The neuron is made up of dendrites, a cell body and an axon connecting to axon terminals. The brain will receive information or inputs which are then transferred into the cell body through dendrites. The cell body works as the processing unit, where all the learnt information is then transferred into outputs and passed down by the axon. The muscles then receive the outputs from the axon terminals for actions. McCulloch and Pitts first studied this concept in 1943 to form a mathematical model (Bakar and Tahir, 2009).

Figure 2 shows a one hidden layer feed forward network with inputs x\_1,…,x\_i, and output y\_k. Each input has its own synaptic weight. The weights are then passed onto the hidden layer, which consists of several hidden neurons. A weighted summation of the inputs is performed by each neuron and then it passes a nonlinear activation function.

In our case, the input will be a matrix of the gene expression of whole blood and the output will be a matrix of the gene expression prediction of another tissue. The visualisation of the expected neural network model is shown in Figure 3. Note that the number of hidden layers is yet to be decided, and it is for visualisation purpose only. Same thing applies for the number of input and output nodes. The number of input and output nodes in the real model will depend on the number of shared genes between whole blood and the other tissue.

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8308186>

[DeepChrome: deep-learning for predicting gene expression from histone modifications | Bioinformatics | Oxford Academic (oup.com)](https://academic.oup.com/bioinformatics/article/32/17/i639/2450757)

<https://pdfs.semanticscholar.org/d77e/c7f520c433d22885408dca375169771c0b9c.pdf>

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032235>