mathwork

Adam Ibrahim

September 2025

1 Introduction

2 MLOPS

Forward propagation seems to work by just executing the neural network with random weights and biases

$$Z = WX + b$$

Where Z is the neuron W is the weight matrix, X is the input vector, and b is the bias, for binary outputs eg (Right or Wrong) you use the sigmoid function to use as the activation function. ReLU is defined as

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Dont know if it's approximate or exact this is called the logistic function, sigmoid function etc... There's also the ReLU function or the Rectified Linear Unit

$$\text{ReLU}(x) = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases}$$

it's also important to note that there are more activation functions and I don't feel like writing them all out

3 MatrixOps

A matrix is a 2d array in cs or a 2d vector a matrix is denoted by uppercase letters so "A" is a matrix but "a" is not a matrix. You could also bold it to emphazise it but it's up to you.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

This is a 3x3 matrix, you can also have non-square matrices like

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

this is a 2x3 matrix, you can also have a mxn matrix where m is the number of rows and n is the number of columns. You can also have a 1xn matrix which is a row vector or a mx1 matrix which is a column vector. You can also have a 1x1 matrix which is just a scalar. With this you can relate this to a system of equations like

$$f(x) = \begin{cases} 2x + 3y = 6 \\ 4x + 5y = 10 \end{cases}$$

This can be represented as a matrix equation like

$$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$$

This is a coefficent matrix, where it's just the coefficents of the variables in the equations. You can also have an augmented matrix which is the coefficent matrix with the constants on the right side of the equations. This is called an augmented matrix because it's augmented with the constants.

$$\begin{pmatrix} 2 & 3 & | & 6 \\ 4 & 5 & | & 10 \end{pmatrix}$$

with these augmented matrices we can generalize any system of equations to an augmented or coefficent matrix. This is useful because we can use row operations to solve the system of equations. Row operations are just operations that we can do to the rows of the matrix to get a solution. You already used row operations in middle school when you did substitution and gaussian elimination A way to describe a solution for a matrix is to put it in row echelon form or reduced row echelon form. Row echelon form is when the leading coefficient of each row is to the right of the leading coefficient of the previous row. The leading coefficient is the first non-zero number in a row. It looks like

$$\begin{pmatrix} 1 & 2 & 3 & | & 6 \\ 0 & 1 & 4 & | & 5 \end{pmatrix}$$

This is in row echelon form

$$\begin{pmatrix} 1 & 0 & 0 & | & -14 \\ 0 & 1 & 0 & | & 5 \end{pmatrix}$$

This is in reduced row echelon form

To get to row echelon form you can use the following row operations:

- Swap the positions of two rows (interchange)

- Multiply a row by a non-zero scalar (scaling)
- Add or subtract a multiple of one row to another row (replacement)

Matrix multiplication is multiplying two matrices (I know shocking)

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 \end{pmatrix}$$