**AI Methods Coursework Part 2**

**Owen Read - F025172**

Introduction

The following report is going to describe and explain the methods I followed before, during and after programming the river flow neural network.

Chosen Language

For the development of the Neural network, I used java for the main program and python for the graphs plotting the rate at which the network learns. The reason I chose java was because of its fast runtime, meaning I can run a maximum of 9999999 epochs (due to matplotlib constraints) in a realistic time. For example, up to 10000 epochs takes less than a second to run. Python was my choice of graph plotting languages as the library matplotlib makes plotting graphs very easy and intuitive.

Data pre-processing

The data set that was given came with some interesting outliers. Some presented were impossible values for the data that we are using such as negative Daily flow/ Daily rainfall. Another outlier that was presented were alphabetic entries such as ‘#’. To deal with those I combined a multiple thing such as:

* Removing the whole column
* Calculating the mean of the 2 values either side of where the outlier was (x + y / 2)

The reason I chose to go with these methods is because they continue to make the data legitimate for the Neural net to train from. If the value was numeric but is wrong the mean was taken as it could have been the user creating the data set accidentally pressing the value. The reason I wanted to delete values that had alphabetic characters as there was not a general idea to go from, meaning the original value has no relation to the dataset.

Splitting the dataset between 60/20/20?