DETAILS OF THE MACHINE LEARNING CODE

DATA CLEANING:

- This study deals with the FSS designs producing a high transmission coefficient of -0.45 dB for a frequency range between 1 to 20 GHz so all the designs not showing such characteristics inside this frequency range are eliminated from the dataset.
- Only the starting and ending points of the frequency range for which a particular design shows a transmission coefficient higher than -0.45 dB, also known as frequency cutoffs, are pertinent to modeling the relationship and, hence isolated for further processing.
- Since one design can produce multiple pairs of cutoff frequencies, all the pairs may also be included. But including more than one cutoff pair is not useful for training because the MLP architecture can only model a relationship that has a fixed number of frequency cutoff pairs for every FSS design. So, in this study, for every design only the first frequency cutoff pair is included.

DATA SCALING:

- All the inputs and outputs are bought in the range of (-1, 1) by using the following mean normalization formula :

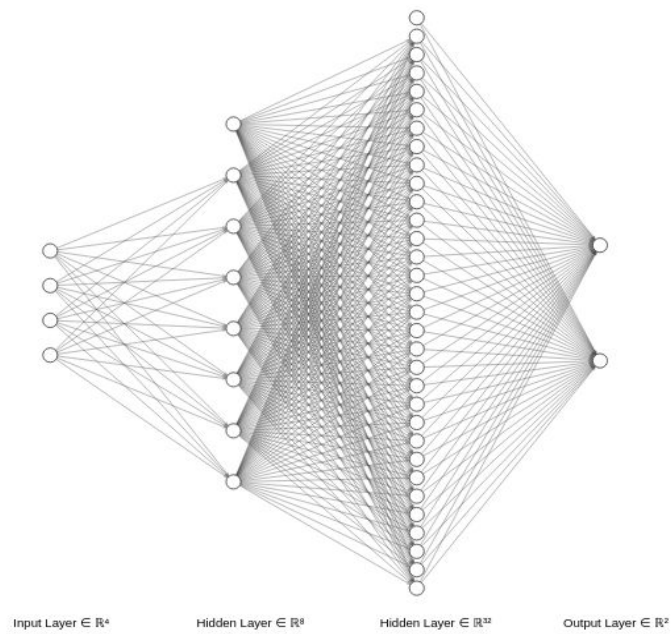
$$x = \frac{x' - \mu}{x_{max} - x_{min}}$$

Data scaling plays an important role in machine learning as unscaled data can lead to **EXPLODING GRADIENTS** causing the learning process to fail. Since, the range of tanh is also (-1, 1), scaling of outputs allows tanh to be used as an activation function in the output layer which helps in achieving convergence easily.

MULTI LAYERED PERCEPTRON ARCHITECTURE:

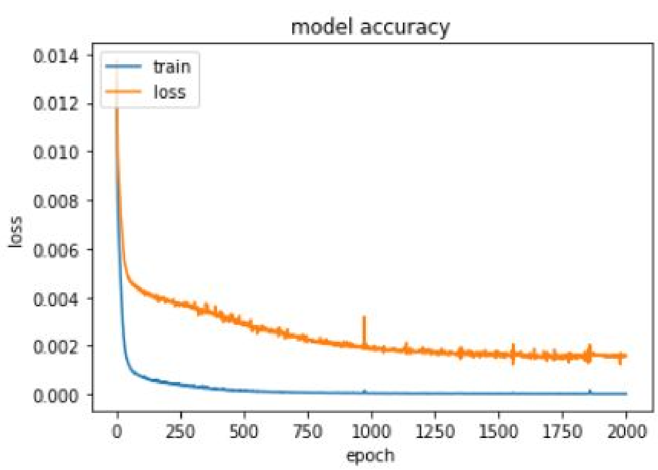
To model the relationship between FSS dimensions and cutoff frequencies a multi-layered perceptron design is used with following features:

- **INPUT LAYER => 4 NEURONS** [4 inputs dimensions]
- **FIRST HIDDEN LAYER => 8 NEURONS**
- **SECOND HIDDEN LAYER => 32 NEURONS**
- **OUTPUT LAYER => 2 NEURONS** [2 cutoff frequencies]



- The hidden layers implement the **RECTIFIED LINEAR ACTIVATION** function which because of its infinite range of all positive numbers, helps avoid **EASY SATURATION** of gradients.
- To capture **NON-LINEARITIES** in the relationship while avoiding saturation of the gradients the output layer uses the **HYPERBOLIC TANGENT ACTIVATION** function.
- The number of perceptrons in each layer is in powers of 2 because that helps the GPU to take advantage of optimizations related to efficiencies in working with powers of two.
- Number of hidden layers and the number of neurons per layer are taken such that the most optimum accuracy in the final solution can be achieved.

- The data is divided into training, validation and test data in the ratio **3:1:1**. Such a ratio is chosen so that a large amount of data is available for training the model but enough still remains to evaluate the accuracy of the model.
- **LOGARITHMIC HYPERBOLIC COSINE** loss function is used in this study because it is similar to 'mean squared error loss' for small values of input, but tends towards $abs(x) - log(2)$ for large values of input, which helps ignore occasional wildly incorrect predictions.
- **ADAM** optimizing algorithm is used for this model because it combines the best properties of the AdaGrad and RMSProp algorithms to provide an optimization algorithm that can handle sparse gradients on noisy problems. Hyperparameters for adam used in this model are as follows:
  - Learning Rate (α) = 0.001
  - Exponential Decay Rate for first moment (β_1) = 0.9
  - Exponential Decay Rate for second moment (β_2) = 0.999
  - Epsilon (ε) = 1e - 08
- A total of **2000** epochs are used to train the MLP because any number more than that would lead to overfitting. Learning curve for this model is as follows:



Loss function used to calculate loss in this graph is log-cosh and the loss is calculated on outputs that are scaled using the mean normalization method.

FREQUENCY LOOKUP TABLE:

- The reason for making a lookup table is two fold:
    - The relation from cutoff to dimension is hard to capture because it is highly complex
    - The Lookup table can accommodate all the possible combinations of dimensions with a very small step size and higher range compared to the input data and still remain feasible.
- A total of 41,62,091 combinations of dimensions are generated and fed into a trained MLP so that they can be mapped with cutoff frequencies and stored inside a lookup table.
- This lookup table is used to find an entry having minimum average absolute error for the frequencies provided by the user. One important benefit of using such a technique is that multiple entries can have same errors implying that multiple designs can give the similar results hence giving the user the freedom to select the best suited one.