Owing to the exponential growth in many scientific fields, much of today’s innovative applications having large scale datasets challenge the natural intelligence of human brain which is the most intelligent system in the world. Learning new patterns in large scale datasets manually, in a fast and intelligent manner, is beyond the capacity or patience of any human being. To address this issue, researchers developed the concept of Neural Network (NN).

Among the existing NN architectures, Multilayer Feedforward Neural Network (MFNN) with a single hidden layer is applied widely for solving a wide range of supervised nonlinear pattern classification tasks. The Back Propagation Network (BPN) is the most popular supervised training algorithm that is used to train MFNN. The larger the network, the more connections it has, the longer it takes to train a new problem. Even on relatively simple problems, BPN algorithm requires the complete set of training samples to be presented in all the epochs and also consumes many epochs for training and hence takes long time. If a large scale dataset is rendered for training, then the BPN training algorithm may take much longer time to train.

Basically, the performance of any training algorithm is determined by its training speed and accuracy. Many researchers have concentrated on improving the performance of the training algorithms in terms of training speed and accuracy. All the existing methods are focused directly or indirectly on tuning the network’s training parameters and also domain-specific knowledge. As the existing techniques employ all the input samples in the training dataset to the network for classification at each and every single epoch till the training terminates, it consumes more time.
This research proposes a simple and new fast training algorithm, namely Adaptive Skipping Training (AST) algorithm, to improve the training speed of NN without affecting the accuracy through Stochastic Sample Selection ($S^3$) method. The core idea is when an input pattern is categorized perfectly by the network, it will not be presented again for the subsequent $n$ epochs. Only the patterns that are not categorized perfectly will be presented again for the next epoch. By this, the total number of input samples employed by the network for training at each epoch gets reduced. Thereby, reducing the total amount of training dataset will reduce the training time. Based on the concept of AST algorithm, different variations of training algorithms, namely Linear Adaptive Skipping Training (LAST), Exponential Adaptive Skipping Training (EAST), Half of Threshold (HOT) Adaptive Skipping Training and Rapid Adaptive Skipping Training (RAST) Algorithm and Constant Adaptive Skipping Training (CAST), have been developed. They can be incorporated into any supervised training algorithms and also into an incremental learning method to improve the training speed.

The performance of the proposed algorithms is proved empirically and statistically by comparing its results with that of the existing techniques in terms of training samples, training time and accuracy. In order to estimate the performance of the proposed training algorithms precisely, the experiments are repeated with different dataset sizes and learning rates. In order to avoid overfitting, the fivefold cross-validation method has been used for the above experiments and the results are analyzed. All AST algorithms perform well with better accuracy, irrespective of the size of datasets.