Cognition-based Smart Computing: Foundations, Frameworks and Applications

Editors: Dr. Sangita Chaudhari, Dr. Pallavi Vijay Chavan, Dr. Parikshit Mahalle, Dr. Pranav Pawar

Can Neural Networks Dream? Applying Crick and Mitchison theory of Reorganization to Multi Layer Perceptrons for reduction of Catastrophic Interference

Author 1	Author 2	Author 3
Aatmaj Kaustubh Mhatre	Ashwini Dalvi	Irfan Siddavatam
Dept. Computer Engineering, K.J. Somaiya College of Engineering, Vidyavihar,Mumbai, Maharashtra India	Dept. Information Technology, K.J. Somaiya College of Engineering, Vidyavihar,Mumbai, Maharashtra India	Dept. Information Technology, K.J. Somaiya College of Engineering, Vidyavihar,Mumbai, Maharashtra India
aatmaj.mhatre@gmail.com	ashwinidalvi@somaiya.edu	irfansiddavatam@somaiya.edu
https://orcid.org/0009-0006-0795-0713	https://orcid.org/0000-0001-9015-457X	https://orcid.org/0000-0002-0858-3154

ABSTRACT

According to Crick and Mitchison's theory of reorganisation, the reverse learning mechanism is responsible for preventing the brain from getting overloaded. This mechanism occurs in REM sleep and is closely associated with dreams. In this study, the reverse learning mechanism is applied to artificial neural networks and its effects on catastrophic forgetting are analysed. The performance of the network with respect to catastrophic forgetting before and after the application of reverse learning is compared.

INDEX TERMS

Reverse Learning, Catastrophic Forgetting, Life long learning, Computational Neuroscience

1 Introduction and Background

Dreams are one of the most complex processes that occur in our brain. There are many theories explaining the reason behind the occurrence of dreams. One such theory that explains why dreams occur is the 'Crick and Mitchison theory'. This theory tries to explain the underlying mechanisms occurring in the brain during dreams. According to this theory, a mechanism known as 'reverse learning' occurs in our brains while we dream. This explains the bizarre nature of dreams and the reason why dreams cannot be remembered. According to the theory [8], dreams that occur in REM sleep [1] are an attempt at removing unnecessary information (parasitic modes) from the brain [28]. This theory postulates the process of removing unwanted modes - reverse learning [2]- as one of the underlying mechanisms that cause dreams in REM sleep [21] [7]. According to Crick and Mitchison, the goal of the reverse learning mechanism is preventing the brain from getting overloaded with too many memories, as they say - 'We dream in order to forget' [9] [29] [24]. Counterintuitive to the common intuition that REM sleep strengthens the connections between brain neurons, the reverse learning mechanism actually aims at weakening the connections. During this mechanism, random signals are given to the brain. These random signals are known as PGO waves [19] [30] [42] [3]. The connections of neurons of the brain that respond to these waves

[1]One of the four phases of sleep in mammals and birds is rapid eye movement (REM sleep), which is defined by the fast and random movement of the eyes. [43]. REM sleep is postulated as one factor responsible for reducing the brain size [7]

[2]Not to be confused with 'Reversal learning'

[3]Ponto-geniculo-occipital waves (PGO waves) are distinctive waveforms that resemble the ran-

are dampened down. This is just the opposite of the learning process, where the connections between neurons which fire together are strengthened. The theory also postulates that this mechanism serves to reduce network 'obsessions'[4]. If the Crick and Mitchison theory of dreams is so critical for biological brains, can it be applied in artificial neural networks? Can artificial brains dream too? The objective of this study is to apply Crick and Mitchison's theory of Reorganization in artificial brains by making neural networks dream and observing its positive effects if any.

According to this theory, the overload of the brain is reduced by reverse learning. Similar to the overloading of natural brains, artificial neural networks also get overloaded. One of the areas of application where the overloading of neural networks causes problems is lifelong learning. The ability of neural networks to continuously learn and update their decisions based on new data is called lifelong learning [35]. A Neural network capable of lifelong learning can learn even when data of different orientations are provided sequentially to the model. When new types of data are fed into the model, it should be able to learn the new data, as well as remember the old data without requiring retraining from scratch. Applications of AI in areas like natural language processing or self-driving vehicles where the nature of the data changes rapidly require lifelong learning. The same models are required to adapt to dynamic environments without exhaustive re-training. However lifelong learning is a challenge for artificial neural networks. Artificial neural networks currently in practice cannot learn continuously for a long period of time. These networks tend to 'forget' at a very rapid rate.

The reason behind this rapid forgetting is known as Catastrophic interference or Catastrophic forgetting. Catastrophic interference is a phenomenon that prevents neural networks from lifelong sequential learning [32] [40]. This effect arises when the new learning data interferes with old network connections and causes the network to drastically forget what it had learnt earlier. For example, if a model is first trained on data A, and then on data B, the backpropagation algorithm caused the weights of A to be erased. This is because while training for data B, the model focuses on minimizing the error rate of B only, without concern for any of the previous data. This erasing occurs at a rapid rate, hence the term catastrophic. When the old data is erased, the model requires retraining on the old data, leading to a huge loss of performance.

While continual learning remains a challenge for neural networks, biological brains are capable of continually acquiring knowledge without forgetting previous knowledge [33]. They can easily adapt to new environments without needing to

dom signals occurring in REM sleep, which according to Crick and Mitchison Theory act as candidates for activating random neurons in the brain. [21]

[4]According to the theory, obsession is a state where the model always predicts only a particular set of values [9]

retrain. On learning new facts, biological brains are capable of learning them at the same time remembering the old facts. Forgetting, if any is very small in size and is not as drastic as that of artificial brains. One of the factors that prevent dreaming in biological brains is REM sleep. What if REM sleep is introduced in ANNs? Can dreaming be a solution to this problem in artificial brains as well? Recent studies have shown how sleep can be effective against catastrophic forgetting in spiking neural networks [38] [16]. Our study takes inspiration from this idea. Application of Crick and Mitchison's dream theory (reverse learning) in artificial networks does reduce the rate of catastrophic forgetting.

In our research, the reverse learning mechanism is applied to multi-layer perceptions (MLP). The model is used to train the MNIST handwritten digit database [41] for optical character recognition. To simulate sequential learning, the entire training dataset is split into two parts. The model is first trained sequentially first on the first part and then on the second. The rate of catastrophic forgetting is measured across a number of samples. To identify if dream sleep has any effect on catastrophic forgetting, the same process is repeated over. But this time, after the model is trained on the first part, dream sleep (reverse learning) is introduced. Rapidly alternating sessions of a few samples of learning are followed by many noise signals (sleep). Various cases are discussed in the results and discussions section. Performance of the model with sleep and without sleep are compared by measuring accuracy drop across sample size. Accordingly, this result concludes that dreaming does help reduce the rate of catastrophic forgetting.

The main contributions of the study are listed below -

1. By applying the Crick and Mitchison reorganization theory, dream sleep - reverse learning algorithm as an effective method of reducing catastrophic forgetting is presented. The reverse learning mechanism for artificial neural networks is discussed in detail. Its implementation is compatible with the current training methods has been provided.

2. Reverse learning has been applied to the MNIST database. The corresponding reduction in catastrophic forgetting has been measured using accuracy graphs. The results have been analyzed.

3. Various parameters for the reverse learning mechanism have been discussed.

2 Literature survey

Catastrophic interference is the tendency of neural networks to forget previously learnt data at sudden and steep rates. Causes and prevention of Catastrophic interference have been extensively studied after its discovery by McCloskey and

Cohen in 1989 [32]. Some of the significant studies include works by French [11] and Kemkar [26]. Methods of prevention of Catastrophic interference were proposed by French [12] and Kirkpatrick [27]. A recent study by Parisi on life-long learning has also attracted attention[35]. Goodfellow [18] investigated the extent to which the catastrophic forgetting problem occurred in neural networks for established and new algorithms. Andrew Robins [36] tried to overcome this problem using the rehearsal methods, which use previously learned data again for reducing catastrophic interference. Research by Tyler Hayes [22] proposes a brain-inspired approach REMIND which replays compressed memories. According to the authors, REMIND's replay is analogous to the replay that occurs in the brain during waking hours and it would be interesting to explore how to effectively create a variant that utilizes sleep/wake cycles (REM sleep) for replay. Moreover, they state that this could be particularly advantageous for an agent who is mostly engaged in online learning during specific hours and offline consolidation during other hours. This REM sleep replay is something similar to what is proposed in our study.

Mandivarapu [31] provides a new method for training data - Self-Nets which claims to incorporate new data with little retraining, and minimal forgetting, without storing prior training data. However, Self-nets are novel concepts and thus do not provide compatibility with old techniques. Our study uses the backpropagation algorithm to implement the dream sleep algorithm, hence not requiring the implementation of any new methods in pre-existing implementations.

Further recent studies have mapped sleep with catastrophic interference [2] [4]. Tadros developed a sleep algorithm for spiking neural networks for the reduction of catastrophic interference [38]. Experiments related to random noise images using pseudo-rehearsal on SNN were conducted by Antonov [5]. Gonzalez tried to simulate sleep in SNN for improvement of continual learning by combining consolidation of new memory traces with re-consolidation of old memory traces [17]. They tried to use Slow wave sleeps to reduce catastrophic interference. Golden simulated sleep in a much similar manner as our research - by providing random noise to SNN neurons but reinforcing old knowledge without using the old data [15]. Watkins used sinusoidally modulated noise as an analogue to sleep [39]. A few studies on neural networks and dreams have also been presented. Fachechi tried to simulate dreams on hopfeild nets for increasing storage capacity [10]. However, none of these studies presents reverse learning as an effective medium against catastrophic forgetting. The majority of these studies have used some kind of 'reinforcing' to improve the model by reusing the old data. However, these methods will not be useful in the condition that the old data cannot be stored in any form, compressed or otherwise. On the other hand, the Dream sleep algorithm presented in our study does not utilize old training data in any format.

The reverse learning mechanism was first proposed by Crick and Mitchison

[8] [7] [9]. They simulated the mechanism on artificial neural networks for research on biological neural networks. Studies like these have used ANNs to prove concepts of biological neural networks [3]. Christos used hop-field neural networks to prove that Crick and Mitchison's theory reduces overloading [6]. All these studies were applications of artificial neural networks as a technique for simulating biological brain systems. However, a study which applies their theory to artificial neural networks or aspects of AI/ ML could not be found. In this study, the application of their theory has been done to try to reduce catastrophic interference.

3 Dreaming network

3.1 Problem statement

Consider two separate categories of data A and B. A data set containing of data belonging to either category is split into corresponding two datasets A' and B'. Example A is odd digits while B is even digits. Then Dataset A' contains samples of images of odd digits while Dataset B' contains samples of even digits. A simple multilayer perceptron model is trained on the data. The aim of the model is to classify an image that belongs to which of the following categories is formed by $A \cup B$. (Example digits)

The model is first provided with dataset A and then with dataset B. The assumption here is that the worst-case scenario is present. The following assumptions are made -

1. The model cannot store any training data
2. Sets A and B have no element in common $A \cap B = \emptyset$
3. Catagories of data in set A and set B have no element in common, that is Sets A' and B' have no element in common $A' \cap B' = \emptyset$
4. Data in set A is provided sequentially followed by data in set B

3.1.1 Catastrophic forgetting

Since the storage of the data is not permitted, the method for training is training on dataset A followed by training on dataset B. However such a sequential form of learning on two datasets (A followed by B) results in catastrophic forgetting. When the model is trained on dataset A. It forms associations related to categories of data in A. Thereafter when it is trained on dataset B, it forms associations related to categories of dataset B. While forming these associations, it erases the associations formed earlier. Hence it forgets the knowledge gained when trained on A. this can be inferred from the research by [32]. In their case, they have used the addition of one as A and the addition of 2 as B.

3.2 Reverse Learning Dream Sleep algorithm

The reverse learning mechanism makes the model 'forget' connections. The forward propagation and the backpropagation algorithms, that form the learning algorithm give invalid inputs and process the network. This increases the neuron connections on valid firings and reduces the connections on invalid firings. The reverse learning mechanism acts just the opposite to the forward propagation method. It provides the model with random noise as input. Then it processes the model by reducing any connections that fire together. The reverse learning mechanism is based on the essence proposed by the Crick and Mitchison theory [7]. The random noise given to the network is an analogue to the PGO waves that occur in our brain.

According to them, the process of reverse learning is dampening of physical neuron connections in the brain by PGO waves. During sleep, PGO waves acting as a random noise are fed into the biological neural network. Then the connection strengths between the neurons that fire in response to this random noise are reduced.

A similar reverse learning process can be implemented for ANN. Firstly, provide random values as input to input neurons. (Range of the values depending on the model) Then forward propagate the values in the network through all layers. Reduce the weights of the neurons that show a high value. These steps can be easily translated to use methods that are already in practice, that is the standard forward learning algorithms - forward propagation and backward propagation. This is because backpropagation has the same effect of reducing the weights if the expected values of all output neurons are made 0. This can be achieved by providing a label that is not present in either of the sets to the model. The methodology for reverse learning is explained alrogithm 1. How the entire dream sleep algorithm works in explained in algorithm 3.

In reverse learning, unnecessary connections are reduced but not completely erased. We can simulate the Reverse Learning mechanism on Artificial Neural Networks using a set of simple procedures, as proposed in the re-organization theory [9] and modified for application in multi layer perceptrons.

1) Provide random noise as input to the model.

2) Set the label to k where k is a label which is not present in the training data and corresponds to the expected activation of all output neurons to 0. (Example in this case label can be '10' or 'A')

3) Forward propagate the model with the random values.

4) Set the learning rate to sleep learning rate. Backward propagate the model.

Here steps 2 and 4 have the same effect as adjusting the synapses between neurons to reduce the association slightly. This is because by setting the expected answer (label) to something the model cannot predict, (Example setting label '10' or 'A' when a model can predict only numbers from 0 to 9) the expected activation values of the output layer are set to 0. When the model predicts any of valid predictions for the noise, that value must be reduced. Example if the model predicts 8, it has an obsession for value '8' or neurons associated with value '8' are overloaded. The connections of all such neurons must be reduced. This reduction is automatically handled by the backpropagation algorithm, as the expected values are set to 0.

This algorithm 3 summarized the reverse learning procedure.

Algorithm 1 Reverse Learning

Input: Model trained on a Dataset D. Model is designed for total possible labels (classes) set S with total dimensionality d. Model accepts numerical input values in range (a, b). The model is a classifier model that predicts the probability of the each class in S.

Output: Model with single iteration of reverse learning performed on it. (Model with 1 dream cycle introduced on it.)

Parameters: Rate of forgetting α

Steps:
1. Generate a d dimensional datapoint with random values in the range (a, b)
2. Assign a label l to the datapoint such that $l \notin S$
3. Feed forward the datapoint in the model
4. Backpropagate the model with learning rate = α

The below section represents the mathematical representation of this algorithm.

3.3 Mathematical representation of Reverse Learning

The reverse learning mechanism can be applied to an artificial neural network. The reverse learning mechanism is defined using mathematical notations for a multilayer perceptron that trains using the forward and backward approaches.

3.3.1 Definations

$X_{i,j}$ = Activation of i^{th} neuron in layer j.

We assume that the activation has range from -1 to 1 except for input and output neurons, that is $X_{i,j} \in (-1,1)(\forall l > j > 0)$
$W_{i,j,n,m}$ = Weight between the i^{th} neuron in layer j an
α = Rate of forgetting (analogous to learning rate)
l = number of layers
n_j = number of neurons in layer j
$b_{i,j}$ = bias corresponding to i^{th} neuron in layer j
$\delta_{i,j}$ = Error corresponding to i^{th} neuron in layer j
Let h represent a strictly monotonically increasing rectification function in the range 0,1

3.3.2 Forward pass

The forward pass works as follows:- Provide random values as input to the input neurons. The range of the values depends on the model. Forward propagate the values in the network through all layers.

For a general case of a network operating in the range 0-1, the random inputs can be represented as

$$X_{i,0} = \textit{Random input between 0 and 1}$$

Feeding these random numbers into the network and applying forward propagation (weighted sum of neuron outputs) the following expression is obtained:

$$X_{i,j} = h\left(\sum_{k=0}^{n_j} (W_{i,j,k,l-1} \cdot X_{k,l-1}) + b_{i,j})\right) \quad (1)$$

3.3.3 Backward pass

Reduce the connections that fire to the random noise. This can be done by the delta learning rule with no output as the expected output. That is by implementing standard backpropagation procedure by setting the expected values of all output neurons to 0.

This results in the following expression for the updated weights

$$W^{+}_{i,j,k,j-1} = W_{i,j,k,j-1} + \alpha \cdot \delta_{i,j} \cdot X_{k,j-1} \quad (2)$$

Where the error function is given by:

$$\delta_{i,j} = X_{i,j}\left(1 - X_{i,j}\right)\sum_{k=0}^{n_j}\left(\delta_{k,j+1}\cdot W_{i,j,k,j+1}\right) \tag{3}$$

$$\delta_{i,l} = -\left(X_{i,l}\right)^2\left(1 - X_{i,l}\right) \tag{4}$$

3.4 Training using Dream sleep

We are considering a case of a continuously learning model, we have an initial dataset 'A' on which the model is trained. After the model is trained, we have to retrain it on data points that appear as time progresses. Intermixing of a few parts of A and many parts of B can be done. However in this study, we are considering an worst case scenario where the model doe not return to 'A' data after it is trained, and is trained only by 'B' data. We are also considering the worst case where A and B data are mutually exhaustive, that is, type of data that occurs in B does not appear in A and data in A does not appear in B.

By introducing dream cycles while training the dataset B, the rate of catastrophic forgetting can be reduced. The methodology is explained in the flowchart 1. The detailed algorithm is in algorithm 3. The process introduces k samples of sleep after every n samples of training of dataset B.

If the model would have been trained without dream cycles, then the algorithm would have been a traditional training algorithm explained in 2

The dream sleep algorithm works differently. The first part A is trained first as traditional method suggests[5]. It is however followed by rapid alternating sessions of B and sleep. The sequential order of A-B is followed, once 'A' is trained, data of A is completely excluded from the model. [6]

The training algorithm with sleep can be expressed as is as follows.

1) Train data-set A.

2) Train n elements (here images) from data set B.

3) Run sleep algorithm k times.

4) Repeat steps 2 and 3 until data set B is exhausted.

The training algorithm can be summarized in algorithm 3

3.4.1 Train to Sleep ratio

After n successive training of datapoints, k iterations of sleep are performed. The ratio n:k can be identified as one parameter which effects the performance of the

[5]It is observed that introducing sleep between A does not have any significant effect.

[6]This idea can be extended to three different data sets A, B, C while intermixing sleep between B and C training.

model. The value of batch-size also plays a significant role on the performance, especially when value of n is comparable with the batch-size. On experimental basis it is found that training of 2 samples followed by sleep for 20 cycles is an optimal solution, however 20 samples followed by 20 sleep cycles is more efficient. Reducing the value of n and increasing the value of k increases the amount of processing required. [7]

Algorithm 2 Training without dream sleep

Input: Datasets A, Set of sequentially available datapoints B, classifier model M. Dataset A is available only in the initial stage. B is available as the points emerge. Neither A nor B can be stored. A and B have mutually exhaustive classes.
Output: Trained model that can classify only B.
Parameters: Model architecture
Steps:
1. Train on Dataset A
2. Train on Dataset B
3. end

Algorithm 3 Training with dream sleep

Input: Datasets A, Set of sequentially available datapoints B, classifier model M. Dataset A is available only in the initial stage. B is available as the points emerge. Neither A nor B can be stored. A and B have mutually exhaustive classes.
Output: Trained model that can classify both A and B.
Parameters: Number of training points to be considered before sleep : n
Number of sleep cycles: k
Rate of forgetting: α

Steps:
1. Train on A
2. **While** samples of B remaining:
    a. Train n samples of B
    b. **For** k times:
        Perform Reverse Learning on M with rate α
3. end

---

[7] Increasing sleep beyond 20 has little effect on improving the performance of the model, but increases the processing required.

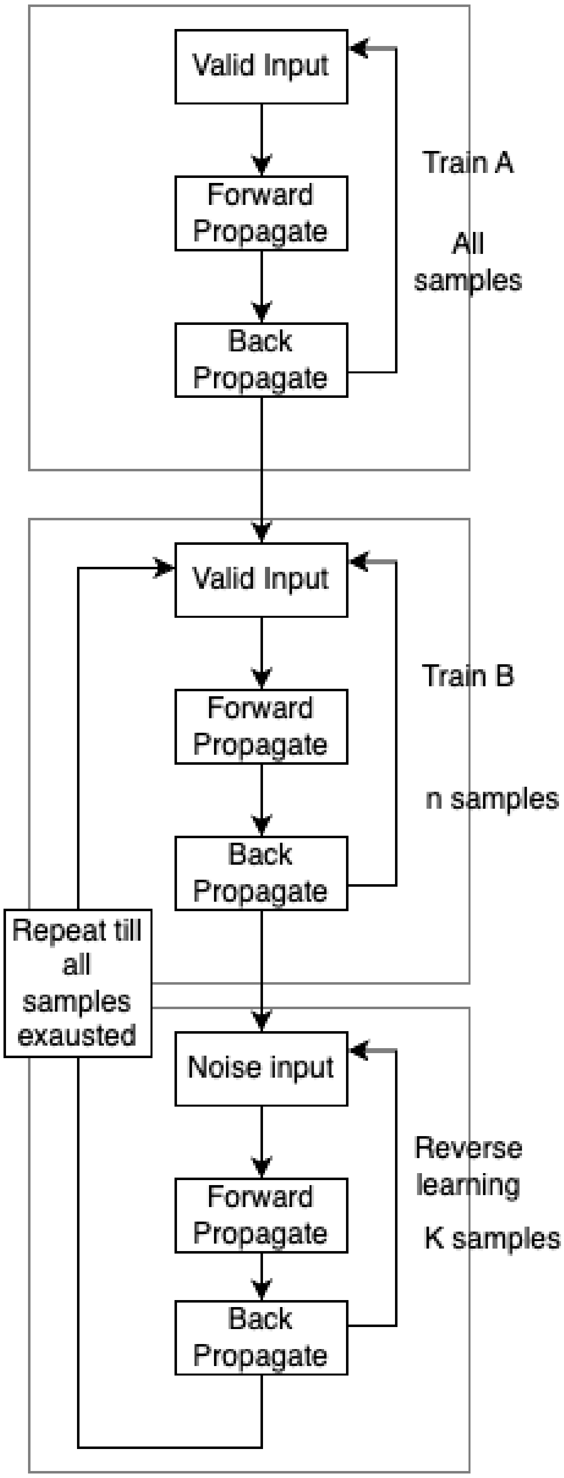

Figure 1: Dream sleep reverse learning methodology flowchart

3.5 Relation with Regularization

Considering the principle behind the dream theory, that dreaming prevents overfitting of the brain, the dream sleep algorithm can be interpreted as a form of regularization. Regularization is a method employed in machine learning and deep learning to avoid overfitting and enhance the model's ability to generalize. It improves the performance of the model by reducing the variance in the model. It makes the model simpler. The dream cycles that are introduced have a similar effect of making the model simpler by removing the unwanted connections in the neurons. Due to the regularization effect of the dream cycles, the overfitting of the model on the available training data that is available is reduced, so that it can accommodate new training data that is made available at a later stage.

3.5.1 Weight decay

In weight decay regularization techniques, weights are reduced and taken towards zero. A neural network featuring smaller weight matrices results in more straightforward models that significantly minimize the issue of overfitting [25].

Weight decay works by adding a penalty term to the cost function.

$\delta'_{i,j} = \delta_{i,j} + \lambda * \omega(i, j)$

As lambda increases, weights decrease. Where omega is a penalty term that is a function of the trainable model parameters [34]. The two most common techniques of weight decay are L1 and L2 regularization. Both L1 and L2 regularization work towards reducing the weights of the neural network. L1 regularization (LASSO) reduces the weights by changing the cost function to force the weights to zero. L2 regularization, also referred to as Ridge regression, is a technique that imposes a penalty on model parameters by adding a term to the loss function. This penalty term causes the model's weights to be "decayed" or pulled towards zero, effectively shrinking their magnitude. This decay is only partial, allowing for some non-zero values that still contribute to the model's predictions.

However, the dream sleep algorithm does not work in the same manner. Rather than reducing the absolute value of all weights, it reduces only those weights that respond to the stimulus of random input. Weights in the hidden layer may increase or decrease depending on the situation. This means that the importance of the weights that are responsible for producing a valid output when the input itself is invalid is reduced drastically. While the importance of other weights will not change much.

For the last layer, we can prove that all the weights will be taken towards a negative value.

Since

$$W^{+}_{i,j,k,l-1} = W_{i,j,k,l-1} + \alpha \cdot \delta_{i,l} \cdot X_{k,l-1} \quad (5)$$

$$W^{+}_{i,j,k,l-1} = W_{i,j,k,l-1} + \alpha \cdot - (X_{i,l})^2 (1 - X_{i,l}) \cdot X_{k,l-1} \quad (6)$$

Since $X_{i,j}$ is in range (0,1), all weights will reduce. For hidden layers, the value of the weights depends on the weights of the next layer, thus depending on their values they might increase or decrease.

While training, the neural networks work on the principle - ”neurons that fire together wire together”. During dream cycles, the principle is that ”connections between neurons that fire together are reduced” This is different from weight decay regularization where the absolute value of the weights is reduced and taken to 0.

4 Experimental Setup

In this section, the dreaming methodology has been applied on the MNIST digit database. Detailed account on the implementation has been given.

4.1 Data set

MNIST handwritten data set [41] was used for training the data. This data set consists of handwritten digits 0-9 in grey-scale 28x28 pixel image format. 50000 samples of training data and 9500 different samples of testing data were used for this study. This is the standard MNIST database [41].

4.2 Model Architecture

The target accuracy we aim for in the architecture is 85-90% Although various methods of increasing the accuracy beyond 95% are in practice, some of them including complex networks like CNN [20] [14], the goal is study is not to increase the accuracy but to apply dream theory in neural network for catastrophic forgetting. One of the main reasons for the decreased accuracy thought the study is due to the fact that every image is trained only once (for 1 epoch only) to simulate dynamic environments, especially those in which data cannot be stored. [8] [9]

4.2.1 Neural Network

The supervised neural network used for training the data consists of 784 input neurons, 10 output neurons. Every output neuron corresponds to digit from 0-9.

[8] in applications where data can be stored, interleaved learning or other methods already in practise can be used in a much effective manner [12] [32] [13]

[9] Similar results have been obtained for multi-epoch training as well.

The network consists of two hidden layers of 16 neurons each. Stochastic gradient descent is performed with batch-size of 4 samples. Weight and bias learning rate was set to 1. The weights were initialized with small random numbers. For simulation of dynamic environments, every image was trained for one epoch only. Sigmoid activation function was used. More details of the model and algorithms used can be found at the digital sources [37] and [23]

Variations of the model and the effect it has on catastrophic forgetting can be seen in fig TODO.

The model is implemented in Java by using ANN4j package [1].

5 Results and Discussions

This section will include the experimental results of the theory presented. The dream sleep algorithm (Algorithm 3 was is applied on the above described neural network and dataset. Then the results were compared with the without dream algorithm (Algorithm 2) The results of various experiments performed on hyper-parameters and network configurations of the proposed dream sleep algorithm will be explained. The reduction in the rate of catastrophic interference will be presented in a visual manner. This section justifies how incorporating dream cycles into the training process helps to catastrophic interference.

5.1 Performance of dream sleep algorithm

The following results can be seen through the experiments conducted -

5.1.1 Increase in Accuracy

1) The introduction of Dream Sleep Cycles improves the overall Accuracy.
2) In both the cases (with and without sleep) accuracy declines as more samples of class B are trained. However, Accuracy without sleep is lower than accuracy with sleep.

5.1.2 No effect on Recall of Class B

1) Dream cycles have no effect on the recall of Class B, which means that the model does get trained on class B.

5.1.3 Decrease in Catastrophic interference

1) Recall of Class A with sleep is with greater than recall without sleep, which means that the rate of forgetting is reduced.

Table 1: Improvement in accuracy at various Training Levels and dream to sleep ratios

Training Level	**Without Sleep**	**With Sleep 2:20**	**With Sleep 20:20**	**With Sleep 200:200**
10%	41.23157895	80.08421053	55.13684211	55.31578947
25%	32.72631579	66.14736842	50.31578947	57.68421053
50%	26.74736842	55.50526316	45.15789474	54.67368421
100%	24.18947368	45.07368421	41.18947368	49.53684211

2) The overall accuracy also proves that the model forgets at a slower rate when dream cycles are introduced.

5.2 Results for a single digit.

Consider case where training for nine digits was performed and after which one digit was trained. The results for few digits are shown as samples in this section. One thing to note here is that the performances of the model across digits varies with respect to its characteristics in the data-set. These characteristics include the occurrence of every digit.

Number of samples of each digit are from 0 to maximum possible training samples at intervals of 100 samples of B.

5.2.1 For digit 1

We train all digits except 1 first, and then train digit 1.

A={0,2,3,4,5,6,7,8,9}

B={1}

Accuracy graphs for different configurations of sleep ratio are given in figure 2. The x axis represents the number of samples from 0% to 100% while the y acis represents the accuracy of the model in percentage. Vaiours train to noise ratios are color coded. Example 2:20 represents 2 samples followed by 20 noise images (dream sleep) is colored blue. It can be seen that Catastrophic interference is reduced for different configurations of dream sleep. The red curve represents the performance without sleep. We see that the accuracy without sleep is 25% while with sleep it greater than is 50%. The results at 100% training are summarized in table 1 -

If the model is made to sleep in busts, then the increment in accuracy at each bust can be clearly seen. Fig. 3 shows one such result. After every 2000 samples, 2000 cycles of dream sleep were introduced. The effect at each spike is clearly

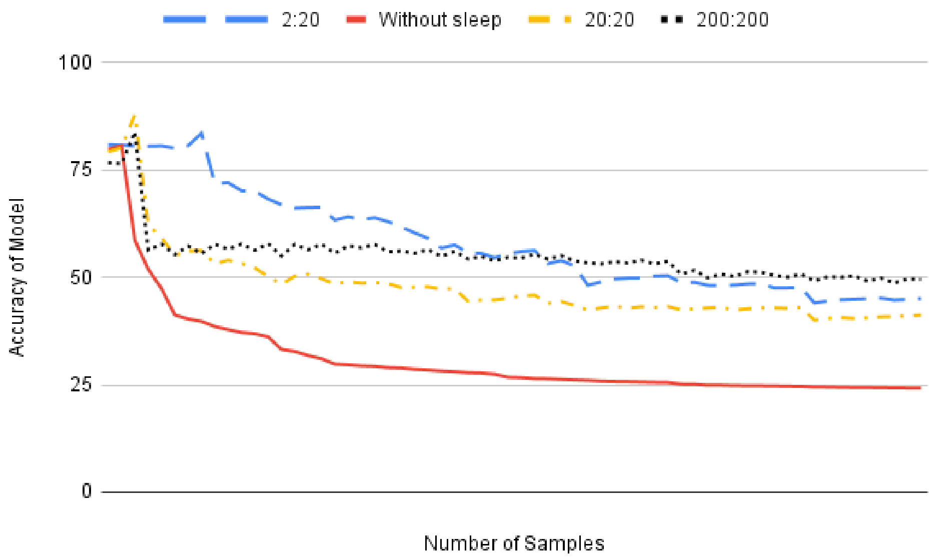

Figure 2: Accuracy vs number of samples for various train to noise ratios.

visible. This further proves that the accuracy is increased by the dream sleep cycles.

Fig. 4 shows a detailed version of model performance for the 2:20 configuration. The individual recall graphs for two digits, 1 and 4 are shown. There is no effect on the recall of class B (digit 1) while recall of class A (digit 4) is higher. at 100% training, Recall of 4 without sleep is 10.72% while with sleep is 22.96%.

Fig. 5 shows the model performance for a different configuration. This is a denser model with three hidden layers of 32, 16, 16 neurons each. This further shows improvement in accuracy. The red curve shows the accuracy with sleep while the blue curve shows without sleep. Fact that the curve of with sleep is higher than without sleep shows how the rate of the catastrophic interference is reduced. At 100% training, the accuracy without sleep is 22.23157895 while with sleep is 44.38947368. This is 2x improvement in the accuracy.

5.2.2 For digit 0

One thing to note is that the improvement depends on the nature of the dataset and the configuration of the model. Hence, there are varying results in various models and datasets. However, one general observation is that towards the middle of the graph, the accuracy with sleep is greater than without sleep. Hence the results obtained for digit 0 are different than that if digit 1.

Graph for digit 0 can be seen in Fig. 6.

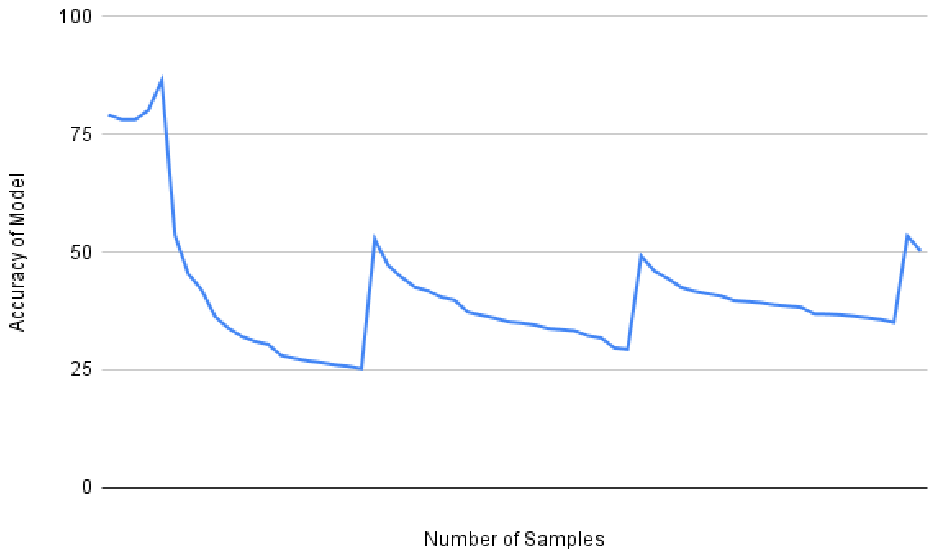

Figure 3: Accuracy vs number of samples. Effect of sleep bursts for digit 1 is seen clearly. Wherever dream cycles are introduced, a spike in accuracy is seen.

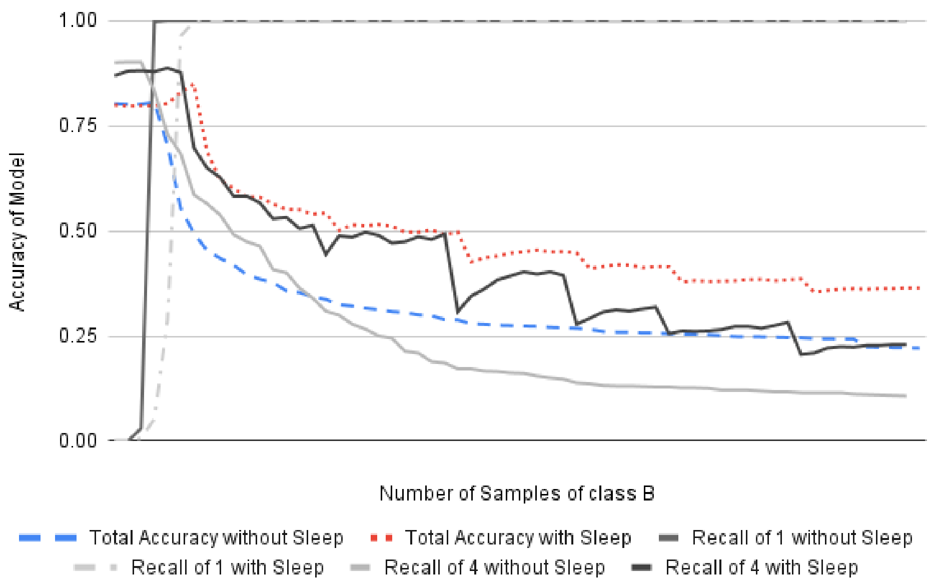

Figure 4: Accuracy and Recall vs number of samples for various configurations. This graph shows how the recall of class A and class B is affected.

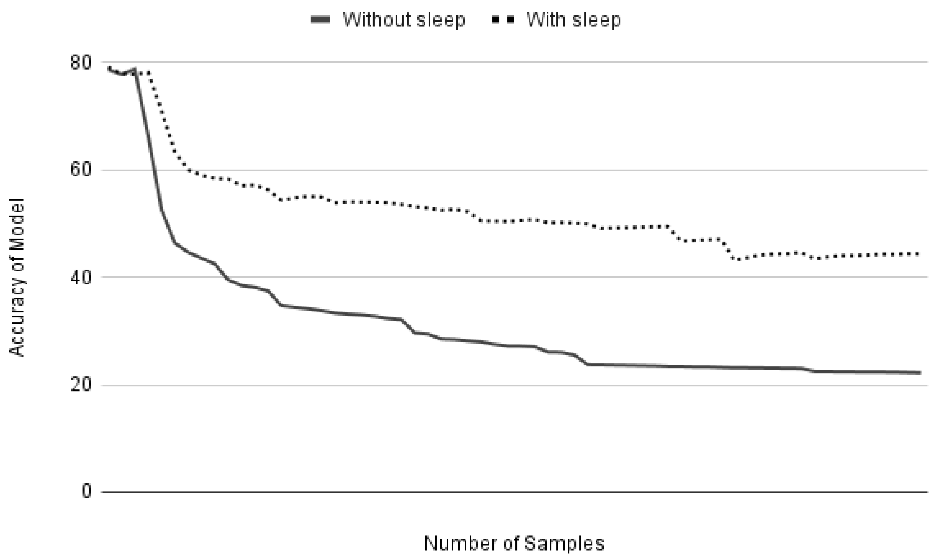

Figure 5: Model performance for digit 1 in a different configuration - denser network (784,32,16,16,10).

A={1,2,3,4,5,6,7,8,9}
B={0}
We can see for some configurations, the accuracy increase of about 10% (1.4x improvement) is obtained by introducing dream sleep cycles.

5.2.3 Fashion Dataset

Fig. 7 shows the reduction in catastrophic interference on the Fashion MNIST dataset. This dataset contains images of fashion products from 10 categories in standard MNIST format [44]. All other parameters of the model remain the same.
A= {T-shirt/top, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, Ankle boot}
B={Trouser}
Accuracy is increased from 25.63% to 40.96%.

5.3 Performance for two digits

The effects of catastrophic forgetting when set B contains two digits are shown for two combinations.

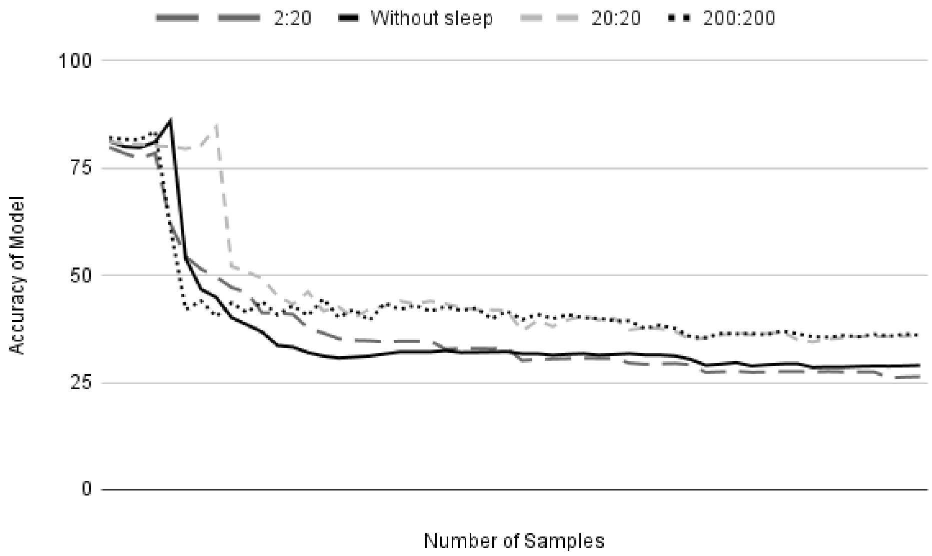

Figure 6: Accuracy vs number of samples for digit 0 show 10% improvement in accuracy

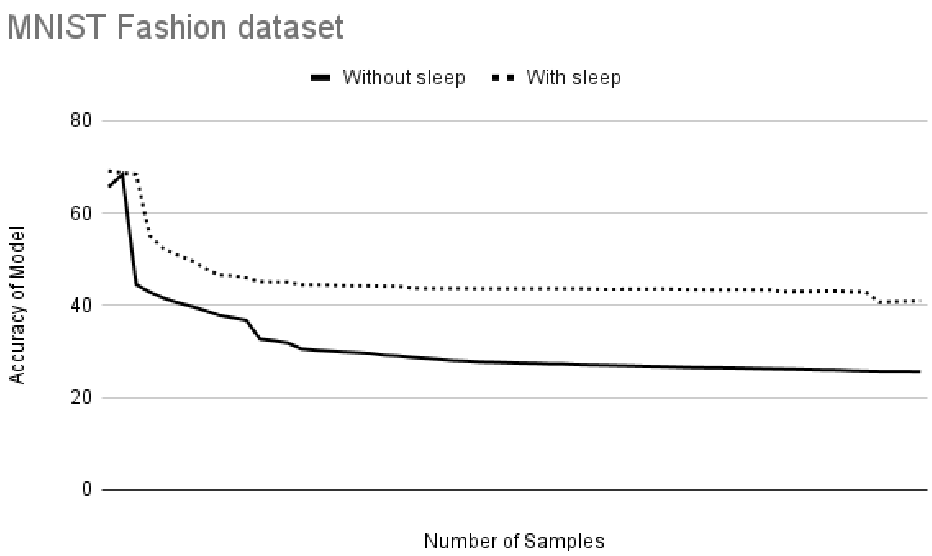

Figure 7: Model performance for on MNIST fashion Dataset.

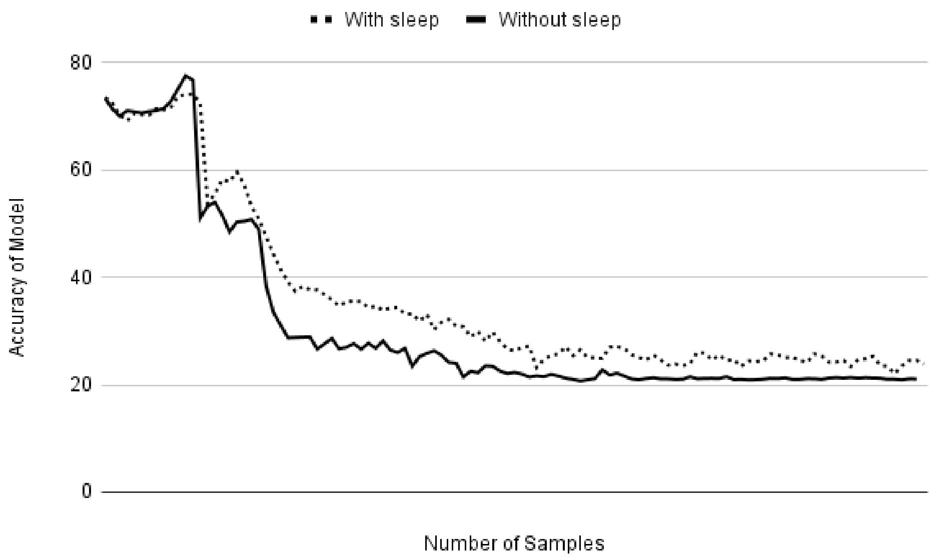

Figure 8: Model performance for two simultaneous training of digit 1 and 5

5.3.1 1 and 5

Fig. 8 shows improvement of the accuracy. In the middle section of the graph, accuracy is increased, although it tapers down towards the end. This means that the rate of catastrophic interference is slightly lower with sleep than without sleep.

A={0,2,3,4,6,7,8,9}

B={1,5}

5.3.2 0 and 5

Fig. 9 shows the results The accuracy is always greater with sleep than without sleep. At 100% training, Accuracy is improved from 19.37% to 26.34%.

A={1,2,3,4,6,7,8,9}

B={0,5}

5.4 Sequential learning for three sets

For further insight into how sleep performs for sequential learning, the data set is divided into 3 sets instead of two. Training of subset A is followed by subset B followed by the last subset C. No overlapping of data is present.

A={0,2,3,4,6,7,8,9}

B={1}

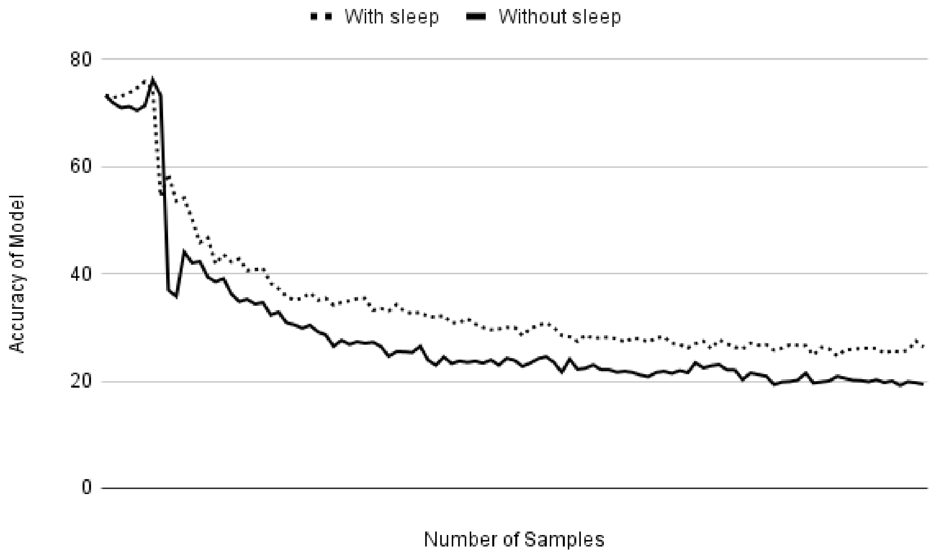

Figure 9: Model performance for two simultaneous training of digit 0 and 5

Table 2: Comparison of performance of models with and without sleep.

Parameter	Without Sleep	With Sleep
Accuracy at 50%samples of B	Accuracy drops catastrophically at a point	Accuracy drops but is greater than accuracy without sleep
Accuracy at 100%training	Accuracy at 25%	Accuracy at 25% to 50% (Upto 2x improvement)
Recall of Class B	High	High
Recall of Class A	Drops along with accuracy	Drops along with accuracy but at times is greater than without sleep
Catastrophic forgetting	Higher	Lower

C={5}

Three conditions are evaluated. First, no sleep cycles are introduced. Then sleep cycles are introduced between both B and C. The third case is where the sleep cycle is introduced only in the middle subset B. From the graph obtained in Fig. 10, it is visible how dream sleep increases performance.

5.5 Comparison of models with and without sleep

Models with sleep show up to 2x accuracy measurement. This shows that the rate of catastrophic interference is significantly reduced. The performance of the models with and without sleep is summarized in table 2.

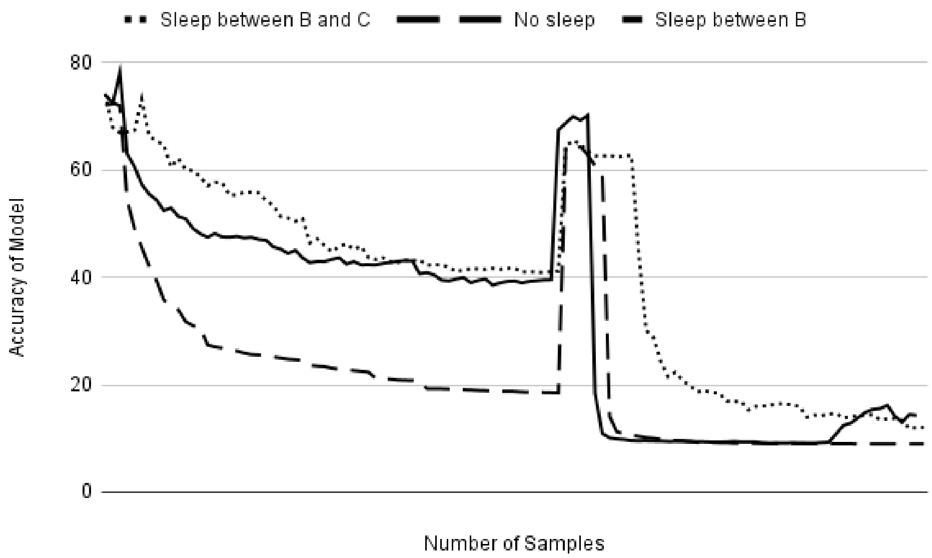

Figure 10: Sequential learning of A followed by B followed by C.

6 Conclusion

In this study, MLPs are used for training the data. various cases and parameters of dream sleep are discussed. The rate of Catastrophic forgetting is reduced to a certain extent by the use of reverse learning dream sleep theory. However, this is subject to various hyper-parameters and data characteristics. Based on the results from this study, this idea can be expanded to more complex forms of networks like CNN, RNN and even those networks which closely mimic the structure of the brain, like hop-field nets or spiking neural networks.

The proposed work presents a novel dream sleep algorithm for multilayer perceptrons. This algorithm is the application of the Crick and Mitchiston theory of reorganization in machine learning models. By introducing the proposed 'dream cycles' in between training cycles, the rate of catastrophic interference (forgetting) was found to be reduced. This is because of the fact that such a dreaming prevents the brain from becoming overloaded, in other words prevents the model from overfitting on the training data set. This is done by reducing the unwanted connections between neurons thus acting as a regularization for the model. This study discusses various aspects of this algorithm, its working, the intuition behind this algorithm, various parameters associated with it and its relation with currently used regularization methods for multilayer perceptrons. Based on the experimental findings, it is evident that the proposed work shows promising potential.

References

[1] Aatmaj-zephyr/ann4j: Artificial neural networks for java this package provides object oriented neural networks for making explainable networks. object oriented network structure is helpful for observing each and every element the model. this package is developed for xai research and development. https://github.com/Aatmaj-Zephyr/ANN4j/tree/main. (Accessed on 12/11/2022).

[2] Ai uses artificial sleep to learn new task without forgetting the last — new scientist. https://www.newscientist.com/article/2346597-ai-uses-artificial-sleep-to-learn-new-task-without-forgetting-the-last/. (Accessed on 12/11/2022).

[3] How artificial neural networks paved the way for a dramatic new theory of dreams — discover magazine. https://www.discovermagazine.com/mind/how-artificial-neural-networks-paved-the-way-for-a-dramatic-new-theory-of. (Accessed on 12/11/2022).

[4] Researchers made an ai whose performance increases if they let it sleep and dream : Sciencealert. https://www.sciencealert.com/neural-networks-performance-increases-if-they-re-allowed-to-sleep-and-dream. (Accessed on 12/11/2022).

[5] Dmitry I Antonov. Noise data impact on catastrophic interference in spiking neural networks. In *AIP Conference Proceedings*, volume 2466, page 070003. AIP Publishing LLC, 2022.

[6] George A Christos. Investigation of the crick-mitchison reverse-learning dream sleep hypothesis in a dynamical setting. *Neural Networks*, 9(3):427–434, 1996.

[7] Francis Crick. Neural networks and rem sleep. *Bioscience reports*, 8(6):531–535, 1988.

[8] Francis Crick and Graeme Mitchison. The function of dream sleep. *Nature*, 304(5922):111–114, 1983.

[9] Francis Crick and Graeme Mitchison. Rem sleep and neural nets. *Behavioural brain research*, 69(1-2):147–155, 1995.

[10] Alberto Fachechi, Elena Agliari, and Adriano Barra. Dreaming neural networks: forgetting spurious memories and reinforcing pure ones. *Neural Networks*, 112:24–40, 2019.

[11] Robert M French. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences*, 3(4):128–135, 1999.

[12] Robert M French. Dynamically constraining connectionist networks to produce distributed, orthogonal representations to reduce catastrophic interference. In *Proceedings of the sixteenth annual conference of the cognitive science society*, pages 335–340. Routledge, 2019.

[13] galoosh33 (https://stats.stackexchange.com/users/119749/galoosh33). How to avoid 39;catastrophic forgetting39;? Cross Validated. URL:https://stats.stackexchange.com/q/314712 (version: 2017-11-24).

[14] Dong-yuan Ge, Xi-fan Yao, Wen-jiang Xiang, Xue-jun Wen, and En-chen Liu. Design of high accuracy detector for mnist handwritten digit recognition based on convolutional neural network. In *2019 12th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pages 658–662, 2019.

[15] Ryan Golden, Jean Erik Delanois, Pavel Sanda, and Maxim Bazhenov. Sleep prevents catastrophic forgetting in spiking neural networks by forming joint synaptic weight representations. *bioRxiv*, page 688622, 2020.

[16] Ryan Golden, Jean Erik Delanois, Pavel Sanda, and Maxim Bazhenov. Sleep prevents catastrophic forgetting in spiking neural networks by forming a joint synaptic weight representation. *PLOS Computational Biology*, 18(11):e1010628, 2022.

[17] Oscar C González, Yury Sokolov, Giri P Krishnan, Jean Erik Delanois, and Maxim Bazhenov. Can sleep protect memories from catastrophic forgetting? *Elife*, 9, 2020.

[18] Ian J Goodfellow, Mehdi Mirza, Da Xiao, Aaron Courville, and Yoshua Bengio. An empirical investigation of catastrophic forgetting in gradient-based neural networks. *arXiv preprint arXiv:1312.6211*, 2013.

[19] Jarrod A Gott, David TJ Liley, and J Allan Hobson. Towards a functional understanding of pgo waves. *Frontiers in human neuroscience*, 11:89, 2017.

[20] Jay Gupta. Going beyond 99% — mnist handwritten digits recognition — by jay gupta — towards data science. https://towardsdatascience.com/going-beyond-99-mnist-handwritten-digits-recognition-cfff96337392. (Accessed on 12/12/2022).

[21] A Hassard. Reverse learning and the physiological basis of eye movement desensitization. *Medical hypotheses*, 47(4):277–282, 1996.

[22] Tyler L Hayes, Kushal Kafle, Robik Shrestha, Manoj Acharya, and Christopher Kanan. Remind your neural network to prevent catastrophic forgetting. In *European conference on computer vision*, pages 466–483. Springer, 2020.

[23] hmkcode. Backpropagation step by step. https://hmkcode.com/ai/backpropagation-step-by-step/. (Accessed on 12/11/2022).

[24] Erik Hoel. The overfitted brain: Dreams evolved to assist generalization. *Patterns*, 2(5):100244, 2021.

[25] Shubham Jain. Regularization in deep learning with python code. https://www.analyticsvidhya.com/blog/2018/04/fundamentals-deep-learning-regularization-techniques/, Feb 2024. (Accessed on 03/24/2024).

[26] Ronald Kemker, Marc McClure, Angelina Abitino, Tyler Hayes, and Christopher Kanan. Measuring catastrophic forgetting in neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[27] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13):3521–3526, 2017.

[28] Johannes Lehtonen. The theory of crick and mitchison concerning the function of dream sleep: A psychoanalytic comment. 1988.

[29] Wei Li, Lei Ma, Guang Yang, and Wen-Biao Gan. Rem sleep selectively prunes and maintains new synapses in development and learning. *Nature neuroscience*, 20(3):427–437, 2017.

[30] Andrew S Lim, Andres M Lozano, Elena Moro, Clement Hamani, William D Hutchison, Jonathan O Dostrovsky, Anthony E Lang, Richard A Wennberg, and Brian J Murray. Characterization of rem-sleep associated ponto-geniculo-occipital waves in the human pons. *Sleep*, 30(7):823–827, 2007.

[31] Jaya Krishna Mandivarapu, Blake Camp, and Rolando Estrada. Self-net: Lifelong learning via continual self-modeling. *Frontiers in artificial intelligence*, 3:19, 2020.

[32] Michael McCloskey and Neal J Cohen. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, pages 109–165. Elsevier, 1989.

[33] Aliaksei Mikhailiuk. Deep lifelong learning — drawing inspiration from the human brain — by aliaksei mikhailiuk — towards data science. https://towardsdatascience.com/deep-lifelong-learning-drawing-inspiration-from-the-human-brain-c4518a2f4fb9. (Accessed on 12/12/2022).

[34] Reza Moradi, Reza Berangi, and Behrouz Minaei. A survey of regularization strategies for deep models. *Artificial Intelligence Review*, 53(6):3947–3986, 2020.

[35] German I Parisi, Ronald Kemker, Jose L Part, Christopher Kanan, and Stefan Wermter. Continual lifelong learning with neural networks: A review. *Neural Networks*, 113:54–71, 2019.

[36] Anthony Robins. Catastrophic forgetting, rehearsal and pseudorehearsal. *Connection Science*, 7(2):123–146, 1995.

[37] Grant Sanderson. 3blue1brown - but what is a neural network? https://www.3blue1brown.com/lessons/neural-networks. (Accessed on 12/11/2022).

[38] Timothy Tadros, Giri Krishnan, Ramyaa Ramyaa, and Maxim Bazhenov. Biologically inspired sleep algorithm for reducing catastrophic forgetting in neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 13933–13934, 2020.

[39] Yijing Watkins, Edward Kim, Andrew Sornborger, and Garrett T Kenyon. Using sinusoidally-modulated noise as a surrogate for slow-wave sleep to accomplish stable unsupervised dictionary learning in a spike-based sparse coding model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 360–361, 2020.

[40] Wikipedia. Catastrophic interference - wikipedia. https://en.wikipedia.org/wiki/Catastrophic_interferencecite_note-Ratcliff1990-2. (Accessed on 12/11/2022).

[41] Wikipedia. Mnist database - wikipedia. https://en.wikipedia.org/wiki/MNIST_databasecite_note-1. (Accessed on 12/11/2022).

[42] Wikipedia. Pgo waves - wikipedia. https://en.wikipedia.org/wiki/PGO_wavescite_note-:0-1. (Accessed on 12/11/2022).

[43] Wikipedia. Rapid eye movement sleep - wikipedia. https://en.wikipedia.org/wiki/Rapid_eye_movement_sleep. (Accessed on 12/11/2022).

[44] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017.