Brijesh verma
SUID: 826257344

# CIS 662: Intro to Machine Learning and Algorithms

# HW-4

## 1. Problem Statement:

2. For the same problem as HW3, i.e., predicting 2022 citations based on all the 2017-2021 citations.
   Try a 1-hidden layer **neural network (5-3-1 architecture)** using the backpropagation algorithm.
3. Begin by making design decisions: node functions, data normalization, output interpretation, optimizer choice, etc.
   Play with different values of the learning rate to see what works best for this problem.
4. Evaluate and compare with the results of HW3.
5. Your submission should include:
   1. code and predictions.
   2. report_HW4.pdf (explaining your results and your conclusions)

## 2. Results:

Three different optimizers were tested in this experiment: Stochastic Gradient Descent (SGD), Adam, and RMSprop. Optimizers are responsible for adjusting the weights of the neural network during training to minimize the loss. Each optimizer uses a specific learning rate, which controls the step size during weight updates.

A key aspect of the experiment was varying the learning rate for each optimizer. Learning rate plays a crucial role in the training process as it determines the step size for weight updates during gradient descent. A high learning rate may lead to overshooting the minimum loss, while a low learning rate may result in slow convergence or getting stuck in local minima. Here are the results of the experiments:

**1. SGD Optimizer**

- Learning Rates: [0.1, 0.01, 0.001, 0.0001]
- Batch Sizes: [4, 8, 16, 32]
- Number of Epochs: [50, 100]

The SGD optimizer produced a wide range of test loss values across different learning rates and batch sizes. Notably, a learning rate of 0.01 and batch size of 8 produced relatively low test loss, indicating better performance. A learning rate of 0.1 with a batch size of 4 also showed good results, but this combination may benefit from further tuning.

**2. Adam Optimizer**

- Learning Rates: [0.1, 0.01, 0.001, 0.0001]
- Batch Sizes: [4, 8, 16, 32]

- Number of Epochs: [50, 100]

The Adam optimizer consistently yielded lower test loss values compared to SGD. It demonstrated robustness to different learning rates and batch sizes. Notably, a learning rate of 0.1 with a batch size of 4 produced the lowest test loss, indicating that this combination is well-suited for this problem.

**3. RMSprop Optimizer**

- Learning Rates: [0.1, 0.01, 0.001, 0.0001]
- Batch Sizes: [4, 8, 16, 32]
- Number of Epochs: [50, 100]

The RMSprop optimizer showed results similar to Adam, with consistent performance across different learning rates and batch sizes. A learning rate of 0.1 with a batch size of 4 produced the lowest test loss, similar to Adam.

**3. Conclusion**

## 3. Conclusion:

In this experiment, the choice of optimizer had a significant impact on the model's performance, with Adam and RMSprop outperforming SGD. A learning rate of 0.1 performed well with both Adam and RMSprop optimizers, indicating that it was suitable for this dataset. The batch size also had an effect on the test loss, with smaller batch sizes tending to yield better results.

**In HW-3 the Nearest Neighbor strategy has the lowest MAE.** The Nearest Neighbour strategy involves assigning test data points to the nearest neighbour in the training set, considering the cluster label. The MAE for this strategy was 213.44 which is much higher than the MAE obtained by using a neural network.

Overall, the best combination for this problem appears to be using the Adam or RMSprop optimizer with a learning rate of 0.1 and a batch size of 4. However, it's important to note that hyperparameter tuning is an iterative process, and further experimentation may be required to fine-tune the model and improve performance. Additionally, other hyperparameters, such as the number of neurons in the hidden layer, could also be explored to enhance the model's predictive capabilities.

Brijesh verma
SUID: 826257344

**References:**

https://chat.openai.com/,
https://keras.io
https://machinelearningmastery.com/start-here/#timeseries