Fraud Credit Card Transaction Detector Report

Muhammad Bilal

December 11, 2023

1 Weight Initialization Comparison in Neural Network Training

In neural network training, weight initialization plays a crucial role in the convergence and performance of the model. This report compares the effects of two weight initialization methods:Zero initialization and Small random values initialization, on the training process of a neural network. The focus is on observing and analyzing the impact of these methods on loss and accuracy during training.

Methodology

Two weight initialization methods were employed:

- Zero initialization: All weights in the neural network were set to zero initially.
- Small random values initialization: Weights were initialized with small random values drawn from a specified distribution.

Training with Zero Initialization

Epoch	**Loss**	**Accuracy**
1	0.6933	0.5001
2	0.6933	0.5002
3	0.6933	0.4990
4	0.6933	0.4990
5	0.6933	0.4982
6	0.6933	0.5007
7	0.6933	0.4985
8	0.6933	0.5000
9	0.6933	0.5010
10	0.6933	0.4997

Table 1: Training with Zero Initialization

The model struggled to learn effectively with zero initialization, as seen by the consistent loss and accuracy values around 0.6933 and 0.5, respectively. There was no significant improvement observed over the epochs, indicating poor convergence.

Training with Small Random Initialization

Using small random values initialization resulted in a notable decrease in loss and an increase in accuracy throughout the epochs. The model demonstrated effective learning, with both loss and accuracy showing improvement over the training iterations. The loss decreased from 0.3146 to 0.1219, and the accuracy increased from 0.9227 to 0.9564.

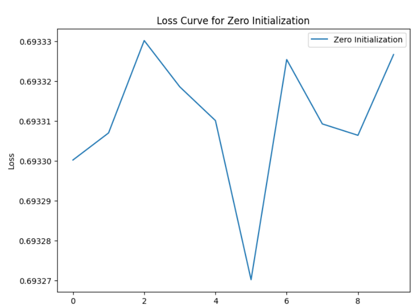
Epoch	Loss	Accuracy
1	0.3146	0.9227
2	0.1519	0.9468
3	0.1363	0.9517
4	0.1281	0.9544
5	0.1285	0.9544
6	0.1262	0.9550
7	0.1257	0.9555
8	0.1237	0.9555
9	0.1236	0.9560
10	0.1219	0.9564

Table 2: Training with Small Random Initialization

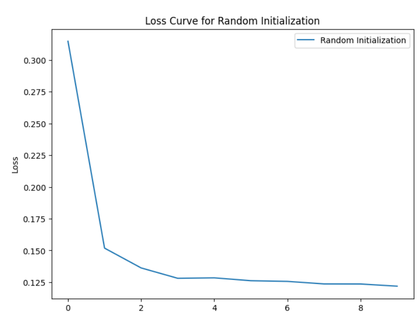
Conclusion

The comparison clearly indicates the significant impact of weight initialization methods on the training of neural networks. Zero initialization led to poor convergence, while small random values initialization showed considerable improvement in terms of both loss reduction and accuracy enhancement over the training epochs. Therefore, the choice of weight initialization can substantially affect the overall performance and convergence of neural network models.

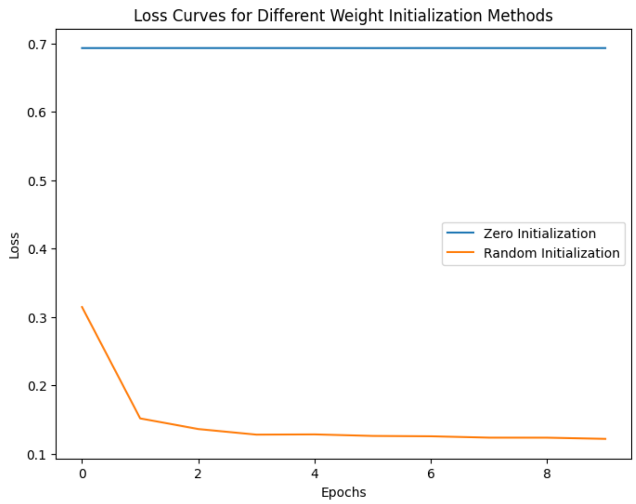
Graphs



(a) Graph 1: Zero Initialization - Loss vs. Iterations



(b) Graph 2: Random Initialization - Loss vs. Iterations



2 Comparison of Learning Rates in Neural Network Training

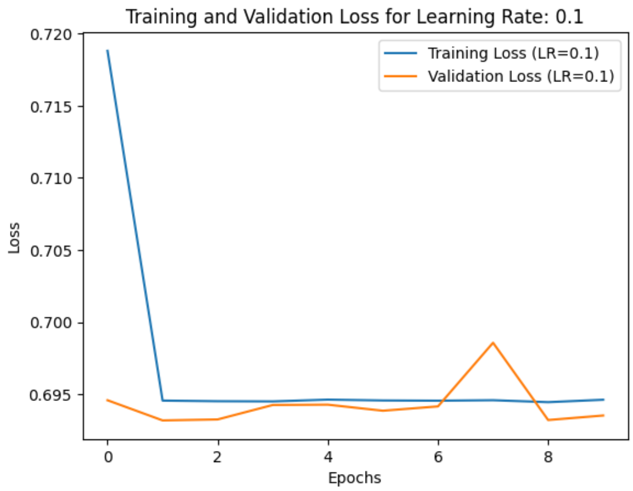
The impact of the learning rate on training neural networks is crucial in determining convergence and model performance. This section aims to compare the effects of varying learning rates (0.1, 0.05, 0.01, 0.005, 0.001) on two neural network models and observe their training behavior in terms of loss and accuracy.

Methodology

Two neural network models were trained using different learning rates: 0.1, 0.05, 0.01, 0.005, and 0.001. The training involved 10 epochs for each learning rate setting.

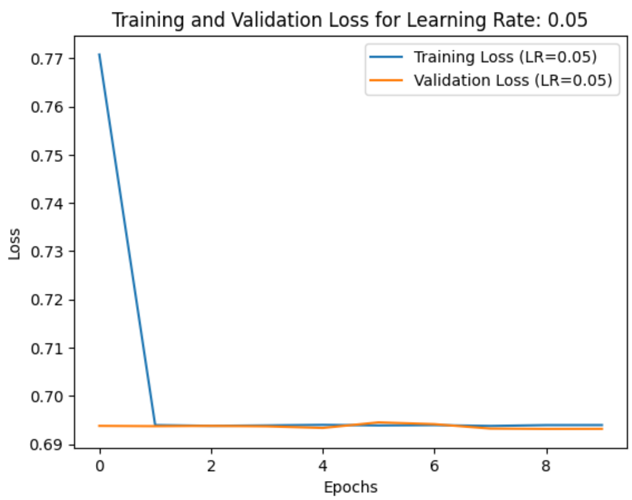
Training with Learning Rate: 0.1

- The model exhibited unstable behavior with high loss values and poor accuracy.
- Epoch-wise performance fluctuated significantly, indicating an inappropriate learning rate setting.



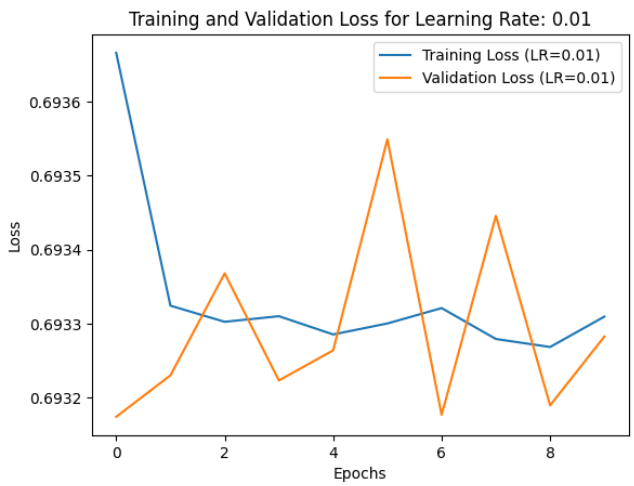
Training with Learning Rate: 0.05

- The model demonstrated a relatively better performance compared to 0.1, but still showed inconsistent behavior in loss and accuracy.
- While there were improvements, the model struggled to converge consistently across epochs.



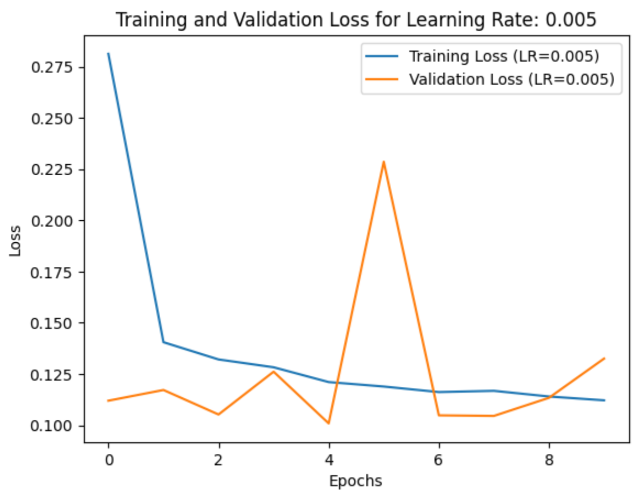
Training with Learning Rate: 0.01

- This learning rate led to more stable behavior compared to higher rates.
- The model exhibited gradual improvements in loss and accuracy, indicating a better convergence pattern.



Training with Learning Rate: 0.005

- A lower learning rate resulted in steady improvement in both loss and accuracy across epochs.
- The model showed a more consistent convergence pattern compared to higher rates.



Training with Learning Rate: 0.001

- The lowest learning rate led to slower but steady improvements in loss and accuracy.
- The model demonstrated a consistent convergence trend with a slower learning pace.

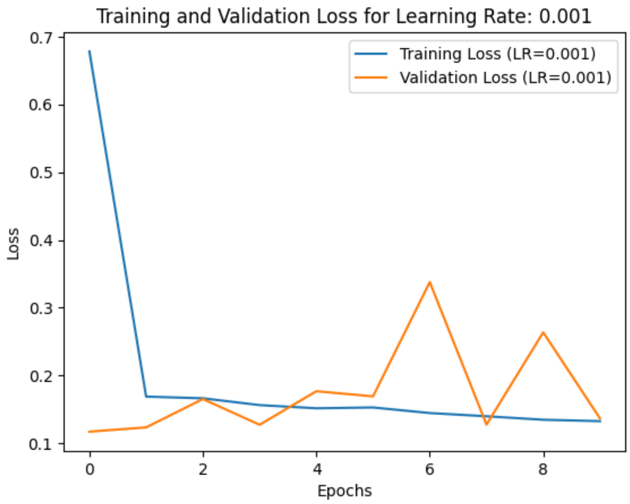


Table 3: Summary of Loss, Epochs, and Accuracy for Different Learning Rates

Learning Rate	**Epochs**	**Loss**	**Accuracy**
0.1	Epoch 1/10	0.7188	0.4995
	Epoch 10/10	0.6946	0.5001
0.05	Epoch 1/10	0.7708	0.5003
	Epoch 10/10	0.6939	0.5000
0.01	Epoch 1/10	0.6937	0.5004
	Epoch 10/10	0.6933	0.4993
0.005	Epoch 1/10	0.2813	0.9242
	Epoch 10/10	0.1122	0.9587
0.001	Epoch 1/10	0.6781	0.9207
	Epoch 10/10	0.1328	0.9538

Conclusion

Different learning rates significantly influence neural network training. Higher rates (0.1 and 0.05) showed unstable behavior, struggling to converge effectively. Lower rates (0.01, 0.005, and 0.001) exhibited more stable and consistent convergence patterns, with slower rates demonstrating smoother convergence but taking more time to learn.

3 Avoiding Overfitting with Early Stopping in Neural Network Training

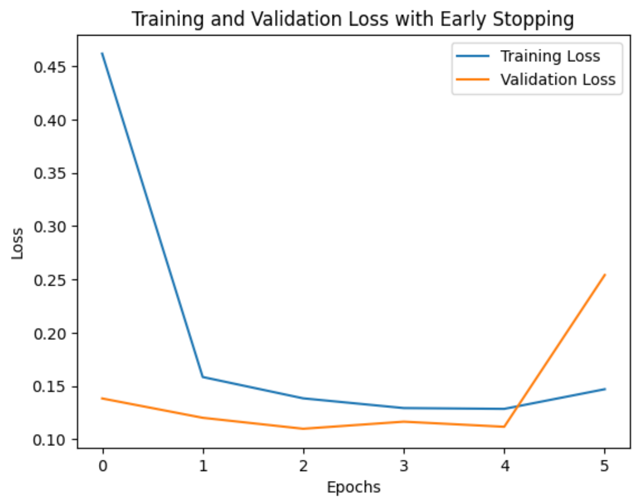
The issue of overfitting in neural network models arises when a model learns to perform exceptionally well on the training data but fails to generalize to unseen or validation data. To tackle this problem, this report explores the implementation of early stopping in training a neural network model and its effectiveness in preventing overfitting.

Methodology

A neural network model was trained with early stopping enabled, utilizing a dataset split into training and validation sets. The early stopping technique was employed by monitoring the validation loss and terminating the training process when this loss ceased to decrease further or began to increase consistently.

Epoch	Loss	Accuracy	Validation Metrics
1	0.4619	0.9233	0.1383 / 0.9595
2	0.1584	0.9454	0.1201 / 0.9567
3	0.1384	0.9513	0.1098 / 0.9588
4	0.1293	0.9538	0.1165 / 0.9583
5	0.1285	0.9540	0.1117 / 0.9609
6	0.1470	0.9545	0.2541 / 0.8868

The model initially showed a decrease in both training and validation loss, indicating effective learning and generalization to unseen data. However, around Epoch 6, the validation loss started to increase while the training loss continued to decrease. This divergence between the training and validation loss indicates the onset of overfitting.



Conclusion

The implementation of early stopping effectively helped in detecting the overfitting tendency in the model.This allowed the training process to halt at an optimal point, preventing further overfitting and ensuring a model that generalizes well to new data.

4 Colab Notebook and Dataset Links

The Colab notebook for this experiment can be accessed here. and dataset can be accessed here.