Instituto Superior Técnico
IST-2024 - Deep Learning

Report - Deep Learning - HW1

Students:
BRUNO Elie
GAUGES Maira

Professor:
FIGUEIREDO Mario

December 9, 2024

Contents

1 Question I

2 Question II

2.1 Logistic Regression

2.1.1 Tuning Learning Rate

For this experiment, logistic regression was trained with a batch size of 32, L2 regularization set to 0.01, and a learning rate from the range $\{10^{-5}, 10^{-3}, 0.1\}$. The training was conducted for 100 epochs.

Table 1: Logistic Regression: Tuning Learning Rate

BS	LR	L2	Val Acc	Test Acc
32	10^{-5}	0.01	0.4694	0.4623
32	10^{-3}	0.01	0.5264	0.5247
32	0.1	0.01	0.3889	0.3807

The best performance was achieved with a learning rate of 10^{-3}, resulting in the highest validation and test accuracy (table 1). The training and validation losses for different learning rates are shown in fig. 1, and the validation accuracies are presented in fig. 2.

2.2 Feedforward Neural Network (MLP)

2.2.1 Default Hyperparameters vs. Batch Size = 512

For this experiment, the default hyperparameters were used, and the batch size was varied between 64 (default) and 512.

Table 2: Feedforward Neural Network: Default Hyperparameters vs. Batch Size = 512

BS	HS	L	Drop	Val Acc	Test Acc
64	200	2	0.3	0.6068	0.6057
512	200	2	0.3	0.5028	0.5200

The results (table 2) demonstrate that a smaller batch size (64) yields better performance. The corresponding training and validation losses are presented in fig. 3, and the validation accuracies are in fig. 4.

2.2.2 Effect of Dropout

The effect of different dropout values $\{0.01, 0.25, 0.5\}$ was explored.

Table 3: Feedforward Neural Network: Effect of Dropout

Model	Drop	HS	L	Val Acc	Test Acc
MLP	0.01	200	2	0.5741	0.5713
MLP	0.25	200	2	0.6054	0.6000
MLP	0.5	200	2	0.6026	0.5963

The results (table 3) indicate that dropout values around 0.25–0.3 achieve the best balance between regularization and performance. Training and validation losses for different dropout values are shown in fig. 5, and the validation accuracies are presented in fig. 6.

2.2.3 Effect of Momentum

Momentum values $\{0.0, 0.9\}$ were tested with a batch size of 1024.

Table 4: Feedforward Neural Network: Effect of Momentum

Model	Mom	BS	Val Acc	Test Acc
MLP	0.0	1024	0.4701	0.4883
MLP	0.9	1024	0.5933	0.6033

The results (table 4) demonstrate that using momentum significantly improves validation and test accuracy. Training and validation losses for different momentum values are shown in fig. 7, and the validation accuracies are presented in fig. 8.

2.3 Discussion

The results show:

- The optimal learning rate for logistic regression was 10^{-3} (table 1).
- Increasing the batch size reduced training time but negatively impacted accuracy (table 2).
- Dropout values around 0.25 to 0.3 balanced regularization and model performance (table 3).
- Momentum significantly improved accuracy for large batch sizes (table 4).

3 Question III

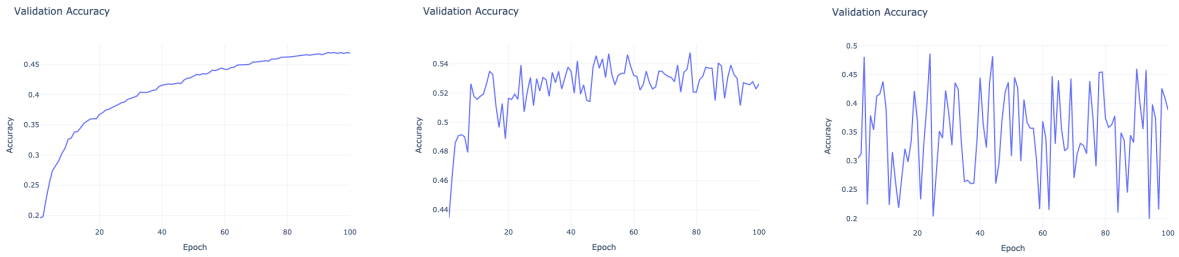
4 Appendix: Graphs and Visualizations

4.1 Logistic Regression: Training and Validation Loss

(a) LR = 10^{-5} (b) LR = 10^{-3} (c) LR = 0.1

Figure 1: Logistic Regression: Training and Validation Loss for different Learning Rates.

4.2 Logistic Regression: Validation Accuracy



(a) LR = 10^{-5} (b) LR = 10^{-3} (c) LR = 0.1

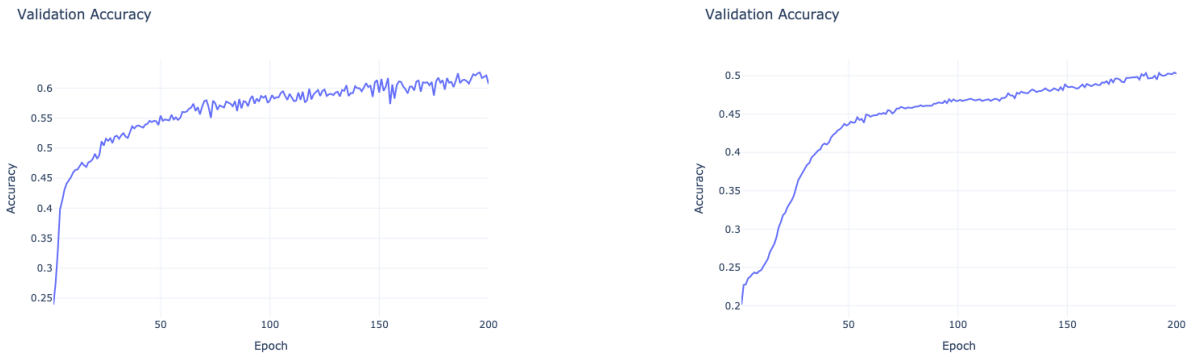
Figure 2: Logistic Regression: Validation Accuracy for different Learning Rates.

4.2.1 Default Hyperparameters vs. Batch Size = 512

The effect of varying the batch size between the default (64) and a larger value (512) was explored.

(a) Batch Size = 64 (b) Batch Size = 512

Figure 3: MLP: Training and Validation Loss for different Batch Sizes.



(a) Batch Size = 64

(b) Batch Size = 512

Figure 4: MLP: Validation Accuracy for different Batch Sizes.

4.3 MLP: Training and Validation Loss (Effect of Dropout)

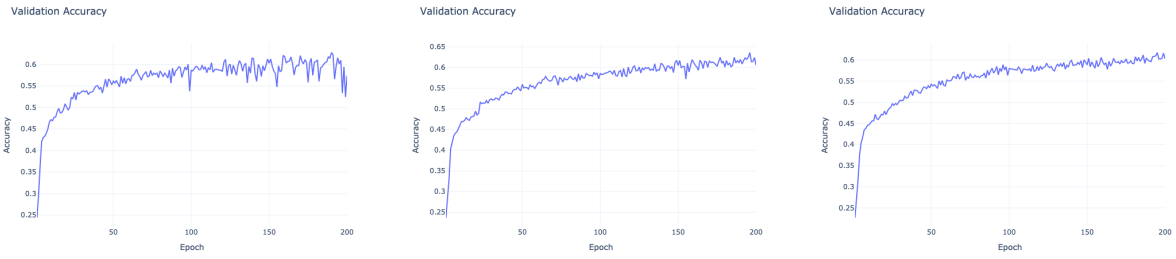
(a) Dropout = 0.01

(b) Dropout = 0.25

(c) Dropout = 0.5

Figure 5: MLP: Training and Validation Loss for different Dropout values.

4.4 MLP: Validation Accuracy (Effect of Dropout)



(a) Dropout = 0.01

(b) Dropout = 0.25

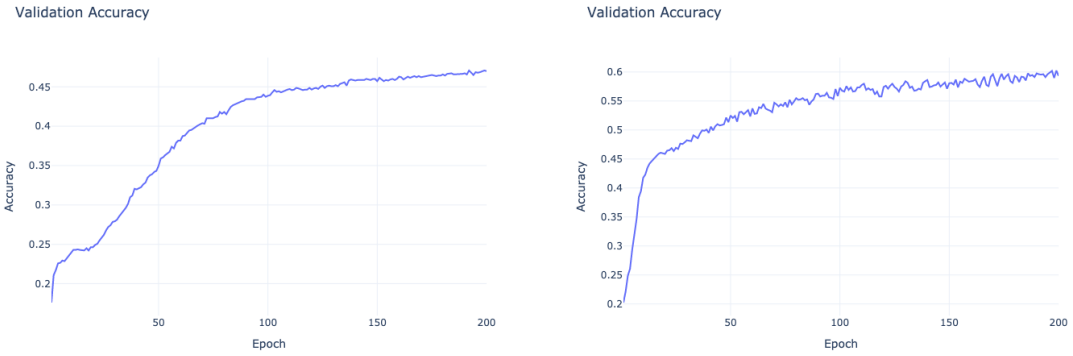
(c) Dropout = 0.5

Figure 6: MLP: Validation Accuracy for different Dropout values.

4.5 MLP: Effect of Momentum

(a) Momentum = 0.0 (b) Momentum = 0.9

Figure 7: MLP: Training and Validation Loss for different Momentum values.



(a) Momentum = 0.0 (b) Momentum = 0.9

Figure 8: MLP: Validation Accuracy for different Momentum values.