**Ex2 Report**

**Part 1:**

תמונה שמכילה טקסט, צילום מסך, שחור ולבן

התיאור נוצר באופן אוטומטיIn this part I was asked to visualize the data. Using matplotlib, I plotted 4 examples per class from train.csv:

**Part 2:**

In this part I trained a logistic regression classifier on the Fashion-MNIST dataset. I used 0.8 of the data for training, and the remaining 0.2 for validation.

I tried every combination of the following values:

* Batch size: 32, 64, 128.
* Learning rate: 0.001, 0.01, 0.1.
* Regularization coefficient: 0.001, 0.01, 0.1.

These are some of the results:

תמונה שמכילה טקסט, צילום מסך, תרשים, מספר

התיאור נוצר באופן אוטומטיBest model:

תמונה שמכילה טקסט, צילום מסך, קו, עלילה

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, תרשים, קו

התיאור נוצר באופן אוטומטי

Here are my insights:

* Lower batch size might lead to slower convergence, but also to better performance on unfamiliar data.
* Higher learning rate might lead to faster convergence but can also skip optimal solutions.
* If the regularization coefficient is too low, the model would learn the training data too closely. If it’s too high, the model would be too simple to capture the required patterns.

**Part 3:**

In this part I trained a NN model with one hidden layer on the dataset.

I tried 30 different combinations of the following values:

* Batch size: 32, 64, 128.
* Learning rate: 0.001, 0.01, 0.1.
* Regularization coefficient: 0.001, 0.01, 0.1.
* Hidden layer size: 50, 100, 200.
* Activation function: ReLU, tanh, sigmoid.
* Dropout rate: 0, 0.5.

These are some of the results:

Best model:

תמונה שמכילה טקסט, צילום מסך, קו, תרשים

התיאור נוצר באופן אוטומטי

Other models:תמונה שמכילה טקסט, תרשים, קו, צילום מסך

התיאור נוצר באופן אוטומטי תמונה שמכילה טקסט, צילום מסך, תרשים, קו

התיאור נוצר באופן אוטומטי תמונה שמכילה טקסט, קו, צילום מסך, תרשים

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, גופן, קו

התיאור נוצר באופן אוטומטי

Here are my insights:

* Lower hidden layer size, as well as lower regularization coefficient might result in underfitting, while higher values can result in overfitting.
* Dropout can help prevent overfitting, but it might also slow down convergence.
* I continued to see that higher learning rate and higher batch size might lead to faster convergence but might also skip optimal solutions and lead to worse generalization.

Instructions on how to run the code:

Each part of the assignment has its own .py file (part1.py, part2.py, part3.py). To run any of them you can open a CLI, navigate to the project’s directory and run ‘python parti.py’.

Note: When running the code of part 2 or part 3, you will be asked to provide the number of epochs that should be performed in the training as input.