DL1_Gedächtnisprotokoll

EYES ONLY ACCESS

SCP Foundation

EYES ONLY ACCESS

TOP SECRET

HANDLE VIA INDICATED CONTROLS

TUCOM - GLADIUS - PROCTOR - WARDEN

THE FOLLOWING DOCUMENT CONTAINS CLASSIFIED INTELLIGENCE AND IS INTENDED FOR PERSONNEL WITH LEVEL 4 SECURITY CLEARANCE OR HIGHER ONLY. PERSONNEL GRANTED ACCESS TO THIS DOCUMENT ARE HELD RESPONSIBLE FOR THE SECURITY AND DISPOSAL OF THE MATERIALS THEREIN. UNAUTHORIZED REPRODUCTION AND/OR DISSEMINATION OF THIS DOCUMENT IS EXPRESSLY FORBIDDEN AND PUNISHABLE BY TERMINATION. FAILURE TO COMPLY WITH THIS PROTOCOL WILL RESULT IN SEVERE DISCIPLINARY REPERCUSSIONS.

FILE ID	Incident 096-1-A
OBJECT CLASSIFICATION	☐ SAFE ☒ EUCLID ☐ KETER ☐ _____

INTSEC 202-81-14 XOD

Of course the following content may contain errors. If you find any, please contact me :) Sorry for the mistakes in advance

1. Exercise (8 * 2.5 = 20)

Multiple Choice questions about all topics

- What is the meaning of **condition number**?
  - a) The **steepness** of the minimum
  - b) The curvature of the minimum
- Properties of **log-cosh** (classification or regression, robust to mislabeled data or not)
- How many weights and biases does `torch.nn.Conv2d(20, 50, 3)` have?
- Using `nn.MaxPool2d(2,2)` on a **64 × 64** input results in **( )%** of inputs not being used.
- $\sum_{i=1}^{d} |y_i - t_i|$ **means** what loss function sums up along the dimentions?
- Properties of **ResNets**
- Properties of **U-Net** (skip connection shares (spatial or temporary) location)

2. Exercise (5+5+5+5=20)

Given **satellite-captured** images of size **200 × 200 pixels** depicting a certain region's vegetation, the image can be divided into **10 × 10 patches**, each of size **20 × 20 pixels**. Each patch has an associated label, resulting in a total of **100 labels**.

1. Draw a model architecture diagram for training this classification task and select an appropriate model.

2. In real-world scenarios, sensor data often contains numerous NaN (missing) values. Please provide strategies to handle this issue.

3. If the model is applied to an interaction layer, we want to check the model outputs at any time, what potential issues might arise? Provide possible solutions.

(in einsicht a tutor mentioned that we should make minor adjustments to the architecture rather than replacing the entire architectur. I used YOLO in this case, but he considered it transfer learning, for which I only received 1 point.)

4. The model exhibits very low accuracy, but the issue is not related to the architecture, loss function, or hyperparameters. Name one possible cause and propose corresponding solution.

3. Exercise (Optimization) (20)

Very similar to exercise 12
i coun't remember the error function. sry about that. but the eigenvalue of Hessian matrix I calculated in 3.b is: $\lambda_i' = 1 + \alpha \cdot \lambda_i^{-1}$

a. Show that the Hessian matrix is:

$$H(w) = I + \alpha \cdot \Sigma^{-1}$$

b. Condition number

Given that $\lambda_1 > \lambda_2 > \cdots > \lambda_d$ are the eigenvalues of Σ,
Hint: The eigenvalues of Σ^{-1} satisfy:$\lambda_1^{-1} < \lambda_2^{-1} < \cdots < \lambda_d^{-1}$
Show that the condition number is given by:

$$c = \frac{1/\alpha + 1/\lambda_d}{1/\alpha + 1/\lambda_1}$$

c. The condition number in this problem is approximately: $\frac{\min(\alpha,\lambda_d)}{\min(\alpha,\lambda_1)}$, How does α influence the model's optimization?

4. Exercise (RNNs) (20)

similar to RNN exercise

5. Exercise (Programming) (20)

We aim to fine-tune the bias term rather than the weights.

$$\min_b \frac{1}{N} \sum_{i=1}^{N} \max(0, f(X_i, b) - T_i)$$

(Note: This formula may contain errors)

- **Shape of X : $N \times 1$
- **Shape of T : $N \times 1$
- **Iterations**: 1000

Question 1:

```
def finetune(X,T,b model):
        optim = torch.nn.SGD((b), lr=0.004)
        # Space for code



        return b
```

Question 2: we do not want b to be complex. How should we modify the formula and code to achieve this?

question 3: What should be done to improve the model's generalization to noise?

(I added a Dropout layer here, but tutor said that this is an early-stage adjustment, and we should take that into account.)