

COMP5511 Artificial Intelligence Concepts - Assignment 2 -

Important Notes

1. Write your report using your preferred word editor (maximum 15 pages). On top of the first page, provide your name and matriculation number.
2. Students are recommended use Python with Pytorch to solve the tasks given.
3. The solution and report should be the results of each individual work.
4. The report together with the codes should be submitted in a zip file (MatricNumber.zip) to LEARN@PolyU (<https://learn.polyu.edu.hk/ultra/course>) under "COMP5511->Assessments->Assignment 2" before the due date of **11:59PM on 4 December 2022 (Sunday)**. No late submission will be accepted.

Assignment Description

Task 1: Compared with simple MLP, deep neural networks process data in a more sophisticated way. Therefore, they are more suitable for solving some more complex tasks. However, DNNs always suffer from the problem of gradient exploding or vanishing.

In this part, you are required to solve an image classification task CIFAR-10 with a deep learning model ResNet [1] which maintains the feature representation of shallow models via identity mapping. Besides, to demonstrate the effectiveness of the identity mapping of ResNet, you are suggested to compare the performance of the ResNet-18 network architecture with and without shortcut connections, i.e., identity mapping. The detailed code of standard ResNet-18 can refer to [2]. You should modify the model file to remove the shortcut connection, finetune the hyperparameter, and train the modified network. The performance of the two networks should be compared according to the accuracy and training convergence curve.

The dataset CIFAR-10 regarding this classification task is available on the website [3]. You can refer to Pytorch to implement the deep learning model.

Task 2: Real-world applications often suffer from imbalanced problems [4], where the class of interest is relatively rare as compared with other class(es). One of the simple yet effective methods to alleviate imbalanced problems is to use data augmentation. In this task, you are required to implement a data-augmentation-based deep learning model to solve an imbalanced face mask detection task [5], where the number of samples in "with_mask" class is much greater than that of samples in "without_mask" class.

Firstly, a deep learning model (e.g., AlexNet, VGG, GoogleNet, etc.) should be implemented to evaluate the performance under the test data without using any data-augmentation methods.

Furthermore, you can arbitrarily use any one or several data-augmentation methods [6], such as rotation, and flip, to enlarge the training samples of the minority class to alleviate the imbalanced problem and then use the deep learning model to evaluate the performance under the test data.

In this task, you should use two metrics, accuracy and G-Mean [4] to evaluate the performance. The dataset regarding this classification task is available on the website [5]. You can refer to Pytorch to implement the deep learning model and the data-augmentation methods [6].

Report

The report should try to include the following details:

Task 1:

- A short description of the gradient exploding and vanishing problem.
- A short description of the experimental settings, such as the hyper-parameters of the implemented model, the architecture of the modified network, etc.
- Experimental results (accuracy and convergence curve) of both models.
- A detailed analysis of the effectiveness of the identity mapping to avoid gradient exploding and vanishing.
- Any other interesting information that you think is pertinent.

Task 2:

- A short description of the imbalanced problem and the implemented data-augmentation methods.
- A short description of the experimental settings, such as the hyper-parameters of the implemented model, the number of samples of minority class and majority class in the train/validation/test data, etc.
- Experimental results (accuracy and G-Mean) of the model without data augmentation and with data augmentation.
- What is the difference between accuracy and G-Mean, and which one is suitable for evaluating the performance of the imbalanced problem?
- Any other interesting information that you think is pertinent.

Notes

- If you do not have enough computing resources to support the tasks, you can manually reduce the number of samples in the dataset and clarify it in the report.
- If you have any questions regarding Task 1, please contact Ms. Chen Xinyi (xinyi-97.chen@connect.polyu.hk).
- If you have any questions regarding Task 2, please contact Mr. Wang Zhenzhong (zhenzhong16.wang@connect.polyu.hk).

References

- [1] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [2] <https://github.com/pytorch/vision/blob/main/torchvision/models/resnet.py>

- [3] <https://www.cs.toronto.edu/~kriz/cifar.html>
- [4] Y. Tang, Y. -Q. Zhang, N. V. Chawla and S. Krasser, "SVMs Modeling for Highly Imbalanced Classification," in *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 1, pp. 281-288, Feb. 2009.
- [5] https://github.com/zhongpolyu/face_mask_detection
- [6] <https://pytorch.org/vision/main/transforms.html>