Introduction to Deep Learning M2177.0043 May 8, 2025

Due on May 15, 2025

Seoul National University

Homework 6

INSTRUCTIONS

- The homework is due at 9:00am on May 15, 2025. Anything that is received after that time will be considered to be late and we do not receive late homeworks. We do however ignore your lowest homework grade.
- Homeworks need to be submitted electronically on ETL. Only PDF generated from LaTex is accepted.
- Make sure you prepare the answers to each question separately. This helps us dispatch the problems to different graders.
- Collaboration on solving the homework is allowed. Discussions are encouraged but you should think about the problems on your own.
- If you do collaborate with someone or use a book or website, you are expected to write up your solution independently. That is, close the book and all of your notes before starting to write up your solution.

1 Setup [0 points]

1. In this homework, we will build and experiment with whitebox attacks and blackbox attacks. You must use Google Colab, which provides free GPUs.
2. Upload hw6 files to your Google Drive.
3. Ensure you are periodically saving your notebook (File → Save) so that you don't lose your progress if you step away from the assignment and the Colab VM disconnects.
4. Once you have completed all Colab notebooks except collect_submission.ipynb, open collect_submission.ipynb in Colab and execute the notebook cells. This notebook/script will:
   - Generate a zip file of your code (.py and .ipynb) called hw6.zip.
   - Convert all notebooks into a single PDF file.
5. Submit the resulting PDF and the zip file to ETL.

2 White box attacks [50 points]

Follow and complete `white-box-attack.ipynb`.

3 Black box attacks [50 points]

Follow and complete `black-box-attack.ipynb`.