ECE 720 Assignment Hands-on Project 3

Assignment 3 due on March 18, 2022, 11:59 PM MST.

Late Submission Policy: 3 free late days, then **10%** off per day late.

1 Introduction

In the last assignment, we've explored the quantitative analysis and automated testing method on deep learning models, specifically, convolutional neural networks (CNNs). In this assignment, we will continue exploring the quantitative analysis and testing on stateful deep learning models—Recurrent Neural Networks. Before working on the assignment, it is highly recommended to read one related work: DeepStellar [1]. Note that basic knowledge about Python programming is required for finishing this assignment.

2 Environment Setup

In this assignment, we **highly recommend** you to use Google Colaboratory, so that you don't have to install anything on your own PC or laptop. The DL framework used in this assignment is PyTorch.

To set up the environment for assignment 3, simply upload the given `HW3.ipynb` file to Google Colab, then run the notebook according to the given.

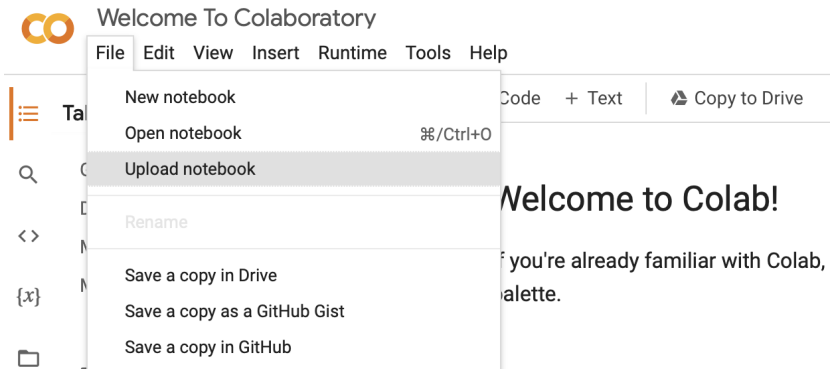

Figure 1: Upload a notebook to Google Colab

If you want to set up the environment on your own computers, here are the prerequisites (only these settings have been tested).

Python	≥ 3.7
PyTorch	1.10.2
torchtext	0.11.2
scikit-learn	1.0.2
nltk	3.7

Table 1: Prerequisites

3 Assignment Objectives

Please follow the detailed requirements of each question in the notebook.

1. [7 points] Implement state abstraction—a key component in DeepStellar.
   - 1.1 [4 points] Build state abstraction and transition model based on training data.
   - 1.2 [3 points] Based on your implementation, which state is most frequently visited? (3pt)
2. [4 points] Implement a function to obtain the corresponding trace* (state transition) in the abstracted model of a given text.
3. [4 points] Implement the metrics for measuring state-based trace similarity and transition-based trace similarity. (2pt for each metric)
4. [4 points] Use DeepStellar to analyze adversarial attack.
   - 4.1 Output traces of original data and attacked data. (1pt)
   - 4.2 Draw a figure to visualize each trace. (1pt)
   - 4.3 Calculate their state-based trace similarity and transition-based trace similarity based on the defined functions in 3. (1pt)
   - 4.4 Analyze the difference between original data and attacked data: give a brief explanation on why the model's prediction result is incorrect on the attacked data. (1pt)
5. [1 points] Brief discussion on the open question: how to further improve the state abstraction method?

4 Submission Guidelines

You need to submit this assignment as a zip file (.zip) containing: 1) notebook (keep all output cells), and 2) project report via eClass. **In the project report, please screenshot all TODO code blocks.** The zip file's name should be `[First name]_[Student ID]_asg3.zip` . Please keep the exact same file structure as the following. For example,

```
zhijie_1234567_asg3.zip
├── HW3.ipynb
└── report.pdf
```

Please note that questions regard submission should first be directed to the TA.

References

[1] Du, X., Xie, X., Li, Y., Ma, L., Liu, Y., and Zhao, J. Deepstellar: Model-based quantitative analysis of stateful deep learning systems. In *Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (New York, NY, USA, 2019), ESEC/FSE 2019, Association for Computing Machinery, p. 477–487.

ACADEMIC INTEGRITY

Students at the University of Alberta must read and follow, in its entirety, the

Code of Student Behaviour

Failure to know the code is not an acceptable excuse for breaking the code.

The University of Alberta is committed to the highest standards of academic integrity and honesty. Students are expected to be familiar with these standards regarding academic honesty and to uphold the policies of the University in this respect. Students are particularly urged to familiarize themselves with the provisions of the Code of Student Behaviour (on the University Governance website) and avoid any behaviour which could potentially result in suspicions of cheating, plagiarism, misrepresentation of facts and/or participation in an offence. Academic dishonesty is a serious offence and can result in suspension or expulsion from the University.

Engineering students studying in the province of Alberta should also follow the

Code of Ethics

by The Association of Professional Engineers and Geoscientists of Alberta (APEGA).

The Code of Student Behaviour should not be too hard to follow. Listen to your instructor, be a good person, and do your own work, as this will lead you toward a path to success. Failure to follow the code can result in a grade of ‘F’ for the course, a transcript remark, suspension, and even expulsion from the university.

"Integrity is doing the right thing, even when no one is watching"
C. S. Lewis

Engineering at Alberta