Name: Daniel Li
Faculty Supervisor: Aleksander Madry
Direct Supervisor: Harshay Shah
Term: 2022 Summer
Jun 15, 2022

Towards Understanding Feature Transfer in Transfer Learning

Project Overview
Provide an explanation/background of your UROP project that includes with whom you are conducting research.

One straightforward and common measure of transfer learning performance is downstream task accuracy. This measure could be misleading, however. There are two hidden factors that could affect feature transfer: task similarity and shortcuts. Task similarity helps with the transfer of learned features while shortcuts hamper the reuse of learned features. One main goal of my UROP project is to develop a quantitative method to measure task similarity, shortcut abuse, and their effect on feature reuse. At the same time, I will investigate in various transfer learning techniques and try to come up with one that increases feature transfer by controlling task similarity and/or preventing reliance on shortcuts.

I am doing the research with MadryLab under direct supervision and guidance of Harshay Shah.

Personal Responsibilities & Goals
Describe your planned role in the project. Be as specific as you can about your personal research duties/responsibilities, expected deliverables, and research goals you hope to accomplish by the end of term.

My role in the project is to carry out all the experiments. This includes reading relevant papers (to have better understanding of the topic), writing the code for the experiments, and reporting the results of the experiments.

Personal Statement
Briefly state why you are interested in this UROP and explain what you hope to gain from it.

The concept of transfer learning sounds very fascinating to me as it saves resources and improves efficiency when training new models. However, we don't fully understand its underlying mechanism yet, and it could provide misleading results. I hope that, by doing this UROP, I will have a much better understanding of transfer learning and enhance my ability to conduct machine learning experiments.