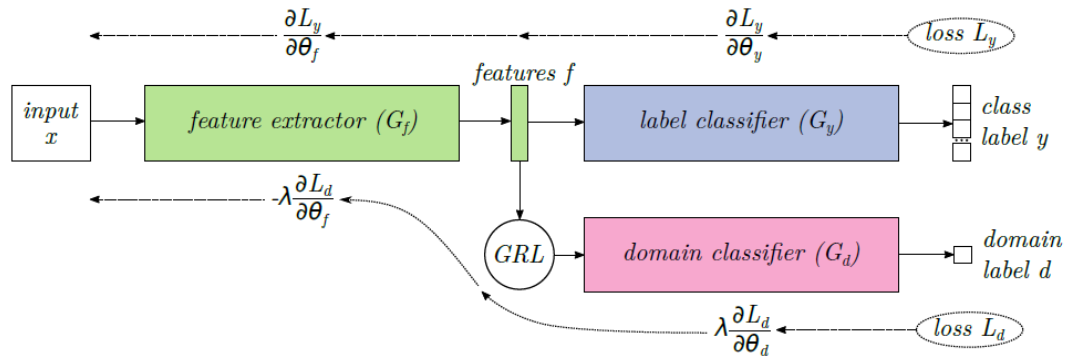


Domain Adaptation with Deep Learning for Natural Language Processing

Project Description. As machine learning is becoming a commodity, their increase usage is being restricted by data availability. While a model is trained and tested on a dataset that comes from the same distribution, relevant data might challenge the system. This is typical when the new data is of the same feature space, but it comes from a different distribution. Domain Adaptation is when we apply an algorithm which is trained in one or more “source domains” to a different (but somehow related) “target domain”.



In this project, we are interested in an NLP application to study domain adaptation for Part of Speech Tagging from Modern Standard Arabic to Dialectal Arabic using deep neural networks. The aim is to minimize the discrepancy between source (MSA) and target domains (Dialects). Extensive testing will be done on common datasets used in the field of domain adaptation.

Project Type: Research

Duties/Activities. The student will run a code in Python on domain adversarial neural networks and fine-tune the hyperparameters to reproduce existing results.

Required Skills: Python.

Preferred Intern Academic Level: Senior B.Sc. or MS student.

Learning Opportunities. Students will learn new concepts in domain adversarial neural networks, and the use of the Wasserstein distance (and its variants) for unsupervised domain adaptation. They will enhance their programming skills in Python.

Expected Team Size: 2 students

Mentors: Dr. Ahmed Abdelali – (contact) <aabdelali@hbku.edu.qa>
Dr. Abdelkader Baggag <abaggag@hbku.edu.qa>