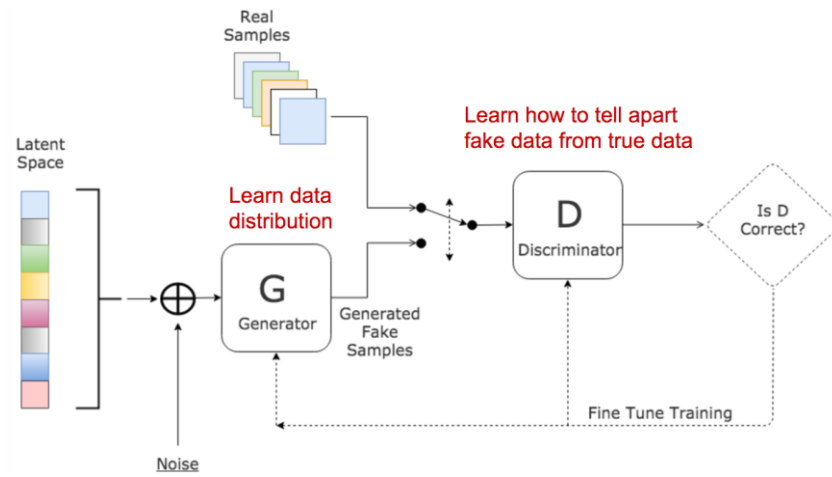


Sliced Wasserstein Distance for Generative Adversarial Networks

– Applications in eHealth and Bioinformatics –

Project Description: Generative Adversarial Networks (GANs) are becoming popular and successful at modeling data distributions from given samples. And Wasserstein distances are increasingly used in a wide variety of applications in machine learning, such as eHealth and bioinformatics, as they take into account spatial information, and can be defined between continuous and discrete data distributions. These factors (among others) have made Wasserstein distances particularly popular in defining objectives for generative modeling. In this project, we consider the sliced Wasserstein distances which may be estimated efficiently through one-dimensional projections. The developed methods will target non-stationary data such as time-series health data, and group anomaly detection for point clouds.



Target dataset. Developed algorithms will be tested on open-source datasets, which will be made available during the summer internship. Also applications in computational biology will turn around detecting a cluster from single RNAseq datasets to see if it is a real cluster or a noise-cluster (aberrant cluster).

Learning Opportunities. Students will enhance their programming skills in Python and acquire new knowledge in generative modeling; data distributions; anomaly detection; Optimal Transport, etc.

Duties/Activities: The interns will run and test different instances of the machine learning code on real data. Some code in python will be provided.

Required Skills: Python and/or Matlab

Preferred Intern Academic Level: B.Sc. 3rd year, and/or B.Sc. 4th year.

Expected Team Size: 2 students

Mentors: Dr. Abdelkader Baggag <abaggag@hbku.edu.qa>
Abdulaziz Al-Homaid <abalthomaid@hbku.edu.qa>