**Week 1: 05/23/2022-05/27/2022**

(Note: the work log for one week should be within one page.)

Goals/Objectives:

(Note: this needs to be filled in the beginning of the week, no later than Tue):

* Read Papers in Literature Folder
* Ask Dr. Le what are tier-1,2, and 3 journals and conferences.
* Low priority: Download python t-SNE library
* Low priority: Implementation of alternative dimension reduction methods.

Accomplishments:

(Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* I was able to at least read both papers.
* I asked Dr. Le the difference between rank/tier 1,2, and 3 journals and conferences were.
* Downloaded and tested the tsne-cuda library.

Difficulty and issues:

(Not finished goals/objectives, TODO; Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Read about the Kullback-Leibler Convergence to better understand the papers.
* Learn how to compute the gradient of a function in preparation for next week’s tasks.

Detailed Summary: I was able to read both of the papers before my meeting with Dr. Le and Sanuj Kumar and learned that I still have to learn about the Kullback-Leibler Convergence to find out why it is so important to keep this property in the cost function. After some back and forth, Sanuj Kumar suggested splitting up the cost function to be able to assign separate weights to the cost of the points of interest and the rest of the points. Once most of it was laid out, I am tasked with trying to compute the gradient of this new function so we can implement it into code. After realizing how much was changed from t-SNE and how much we back-tracked, the project has shifted its focus from improving t-SNE to improving SNE such that it has a focus on points of interest. I also got to ask Dr. Le about the conference rankings and he explained to me the H-Index for conferences and authors.

**Week 2: 05/30/2022-06/3/2022**

(Note: the work log for one week should be within one page.)

Goals/Objectives:

(Note: this needs to be filled in the beginning of the week, no later than Tue):

* Get Gradient of the new cost function
* Low priority: Fix python t-SNE library error
* Low priority: Implementation of alternative dimension reduction methods.

Accomplishments:

(Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* I was able to get some of the gradient with the help of a paper I found.
* Learned a valuable lesson about my attitude in approaching problems.

Difficulty and issues:

(Not finished goals/objectives, TODO; Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Had trouble finding the gradient, I will be more in touch with Sanuj & Dr. Le to resolve it by the next meeting.

Detailed Summary: It was a slower week overall as I tried to review how to get the gradient, then finding the paper that helped me figure some parts out. Other than that, Dr. Le mentioned how my way of viewing things was wrong and I plan to change it. I progressed a little into getting the python t-SNE working on Google Colab, but I encountered a couple issues and will take a look into them next week.

**Week 3: 06/6/2022-06/10/2022**

(Note: the work log for one week should be within one page.)

Goals/Objectives:

(Note: this needs to be filled in the beginning of the week, no later than Tue):

* Get Gradient of the new cost function
* Low priority: Fix python t-SNE library error
* Low priority: Implementation of alternative dimension reduction methods.

Accomplishments:

(Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Got the gradient and wrote a .tex file for it.
* Found the Github Code describing the LAPS and GAPS methods
* Started repository to start combining the different codes.

Difficulty and issues:

(Not finished goals/objectives, TODO; Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Start looking at implementing the new gradient.
* Fix python t-SNE library error.

Detailed Summary: With the gradient complete, we can start implementing it into the pytorch code for dimension reduction and eventually see how the new cost function performs.

**Week 4: 06/13/2022-06/17/2022**

(Note: the work log for one week should be within one page.)

Goals/Objectives:

(Note: this needs to be filled in the beginning of the week, no later than Tue):

* Fix minor issue with gradient
* Implement gradient in new

Accomplishments:

(Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Fixed Minor Issue with gradient
* Gradient was implemented in torch

Difficulty and issues:

(Not finished goals/objectives, TODO; Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Unsure about the validity of implementation.

Detailed Summary: Implementation was in fact not correct, torch implementation needs to be either fixed or scrapped for now.

**Week 5: 06/20/2022-06/24/2022**

(Note: the work log for one week should be within one page.)

Goals/Objectives:

(Note: this needs to be filled in the beginning of the week, no later than Tue):

* Implement Gradient
* Start running tests

Accomplishments:

(Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Implemented gradient through python normally.
* Had a few figures to show with Breast\_Cancer dataset showing the promise of fSNE.

Difficulty and issues:

(Not finished goals/objectives, TODO; Note: this needs to be filled in at the end of the week, Friday or Saturday or Sunday)

* Still a little unsure about the validity of implementation.

Detailed Summary: fSNE was slightly outperforming tSNE when running on default parameters.

**Week 6: 06/27/2022-07/1/2022**

Goals/Objectives:

* Review Gradient
* Implement fSNE using parallelization

Accomplishments:

* Gradient should be finalized
* Extremely basic implementation of fSNE in parallel was developed

Difficulty and issues:

* Gradient more difficult than anticipated

Detailed Summary: After further review by Dr. Le, the gradient was incorrect. Dr. Le gave us the gradient he derived and told us to focus on implementation for now. I will try to reach his same derivation once this is done. Parallel code is running faster than normal python implementation, but it is not within expectations. When it normally takes 30 minutes in python for loops, it takes 12 minutes with parallelization with ten cores.

**Week 7: 07/4/2022-07/8/2022**

Goals/Objectives:

* Speed up implementation of parallel fSNE
* Test on big datasets

Accomplishments:

* Parallel fSNE code optimized, now running within expectations.

Difficulty and issues:

* Parallelization speed was not caused by bad batching, but by repeated duplication of variables.

Detailed Summary: Slow speed in parallelization was caused by a quirk unique to python’s multiprocessing library. Instead of passing references when setting processes in parallel, multiprocessing copies the variables to each method. Once shared variables were implemented, what normally took 30 minutes with for loops, now takes around 3 minutes with 10 cores.

**Week 8: 07/11/2022-07/15/2022**

Goals/Objectives:

* Test on big datasets

Accomplishments:

* Discovered a couple errors with code and fixed them.

Difficulty and issues:

* Running large datasets requires a lot of RAM whether running on GPU or CPU and it is more than colab can handle.

Detailed Summary:

**Week 9: 07/18/2022-07/22/2022**

Goals/Objectives:

* Test on big datasets

Accomplishments:

* Previous error with Google Colab not being able to handle datasets has disappeared, inexplicably. Likely received a weaker GPU when trying to run it last time.

Difficulty and issues:

* LAPS score is different with cuda TSNE and sklearn TSNE, sklearn is much better and it beats fSNE every time.

Detailed Summary:

**Week 10: 07/25/2022-07/29/2022**

Goals/Objectives:

* Test on big datasets
* Test on other types of datasets

Accomplishments:

Difficulty and issues:

Detailed Summary: