Northeastern University
College of Science

College of Science,
Northeastern University
7 Speare Pl 13970-WVG,
Boston MA, 02115
TEL: 201-230-5271
E-mail: berman.ed@northeastern.edu
URL: https://bermanator.studio/

March 7, 2023

Northeastern University
360 Huntington Avenue
Boston, MA 02115

Dear Application Committee for the Ascent Award,

I hope the following exposition finds you well and adequately details my proposed research project on the Point Spread Function (PSF).

Purpose and Significance

When cosmologists try to take pictures of space, objects in the foreground distort the light that stars are emitting from the background. As such, our measurements are not accurate portraits of the stars. Stars are examples of what Astronomers call point sources, and so the aptly named Point Spread Function (PSF) is a mathematical model that quantifies exactly how the light is being distorted. My goal is to build a better model of the PSF so that we may reverse engineer its effects to get more accurate data of our stars. This is useful both for Cosmologists but also all scientists taking measurements from cameras in space that are prone to these lensing affects [1].

Objectives and Methods; Outcomes, Evaluation, and Dissemination

We are currently using a python library called PIFF to do our PSF modeling, and in doing so, we are using its functionality beyond its original purpose [2]. PIFF assumes that the distortion of light is random, which we know not to be the case. As such, my **objective** is to write my own library to replace it.

The only potential challenges in this is the lack of documentation on PIFF. The lack of documentation places more responsibility on me to reverse engineer it in the effort to figure out what shortcomings it has. Figuring out these shortcomings will guide me in creating features I would like to implement in my own library.

I already have all of the resources I need to accomplish this project. I have sufficient computing resources with access to discovery. Additionally, I have been given access to both simulated data and images taken by the James Webb Space Telescope by the collaboration that Professor McCleary and I are apart of. Moreover, Professor McCleary has been given a grant to build her own computer lab in the Dana Research Center, which will be complete and available for me to access during my project should I need resources out of the scope of the discovery cluster.

This project will take me 7 weeks and take place during the Summer 1 term. During the first week I will set out on creating my own metrics to assess how well my program models the PSF. This is something I can do with my existing data from the get-go and something that can guide me as my project evolves. In the second and third weeks I hope to be finished with a prototype of my program. This will have a complete outline of everything I want it to do, but will need to be tuned for some more specific design choices. This code will run on simulated data where we can know what the PSF should be and use our metrics to see how well we did. We will continue to use these metrics throughout the project to evaluate our sucess, specifically, we will use a collection of statistics called rho statisitcs which are standard in the field. In the third and fourth weeks I will use my code on real data to create PSF models that are useful to cosmologists. Then, in the fifth and sixth weeks I will continually tune some of the finer design choices and discuss my progress with the collaboration to see where these is room for improvement. I will spend my last week writing up my results. If my results prove to be better than PIFFs I hope to publish my work. It would at least qualify for the Journal of Open Source Software, but has potential to get published in a more reputable peer reviewed journal.

About the Learner

I have over 6 months of relevant research experience from a summer internship and self developed research course last fall related to Machine Learning on image data. I am currently taking a PhD-Level Class: Math 7223 Riemannian Optimization. This course is allowing me to build up the machinery to tackle optimization problems on smooth manifolds using the Manopt Library, which has implementations in Julia, Matlab, and Python, all languages that play nicely with our current ecosystem. Using Riemannian Optimization techniques in the context of PSF parameter estimation would be a novel approach to the literature and I am actively preparing to implement them. Specifically, there exists a structure known as the Shear Manifold that I plan to do my optimization on [3]. Moreover, the precedent is set for using these techniques on image data, as it is commonplace in robotics and computer vision. I expect this project to move me up the research ladder from a researcher who can help out in running tests to a researcher who can write his own software to be published. My mentor has been working in this area since her PhD and to date has been instrumental on making sure I do not get stuck for too long while also giving me the freedom to learn.