ML4SCI

Proposal : Google Summer of Code 2021

Personal Information

- **Name:** Harsh Pandey
- **Telephone:** +91-8840241988
- **E-mail:** harsh.pandey.cd.phy19@itbhu.ac.in
- **Country of Residence:** India
- **Timezone:** IST (India)
- **Typical Working Hours:** 11:30-13:30, 15:00-19:00, 22:00-05:00 (IST)
  6:00-8:00, 9:30-11:30, 16:30-23:30 (UTC)
- **Language:** English

Project Proposal

Project Title

Domain Adaptation for Decoding Dark Matter with Strong Gravitational Lensing

Overview

Strong gravitational lensing has always been a promising flight to better understand the Dark matter and its underlying nature. Deep learning methods have the potential to accurately identify images containing substructure, and differentiate WIMP particle dark matter from other well motivated models, including vortex substructure of dark matter condensates and superfluids.

This project focuses on further development of the DeepLense pipeline that combines state-of-the art of Deep learning models with strong lensing simulations initially based on PyAutoLens for strong gravitational lens modeling.

The focus of this project is on domain adaptation to account for the differences in the modeling and available real observational data.

I strongly believe that this project will provide better performance to the current thought underlying the understanding of Dark-matter.

I have been working on open source for almost one year now. It is always great when to see your code being used and appreciated by people. Until now I have worked on individual projects, in our campus open source group and a few open source organizations, also the current project of mine covers the modelling of gravitational lenses using “Lenstronomy” library and I am doing a data science research intern that includes working with a millions of Data and running it on complex Deep learning models.

Benefits to Community

This open source community has always been benefited with GSoC but the benefit here does not only restrict to the organization itself but also reaches to all working non-working people in the field of physics especially in the field of Cosmology, Dark matter and Gravitational lensing. The project confines an idea to understand Dark matter in a more easy way with the help of Computer vision and Deep learning.

Previous projects and courses related to project

Projects

- **Gravitational Lensing under Dr. Prasun Dutta(Associate Prof. IIT BHU) (01/2021 - 04/2021)**

  Time delay and amount of deflection faced by light while passing near a gravitational lens.

  Studying microlensing-induced stellar variability in the LMC (EROS, MACHO) as well as in the Galactic bulge (DUO, MACHO, OGLE).

- **COVID-19 Face Mask detector (02/2021 - 03/2021)**

  Used CNN, Convolutional Neural Network for the project and trained the model with help of libraries like Pandas, Numpy, Scikit

  Learn, Pathlib, PIL, Open-CV and many more. Model is 99.73% accurate and identifies the face with a mask or without a mask.

- **Random celebrity classifier (12/2020 - 12/2020)**

The model to predict the probability of faces with using Deep-learning and python's libraries i.e. Pandas, Numpy, OpenCV, Matplotlib and Sci-kit learn.

Deployed the model on AWS using flask and other basic web development tools.

- **Portfolio Website (02/2021 - 02/2021)**

  Built a portfolio website using HTML, CSS, Bootstrap and Django.

Courses

- Solar & Plasma Physics
- Quantum Physics
- Classical physics & Electromagnetism
- Basics of Electronics and Electrical Engineering
- Condensed matter physics
- Computational physics
- Natural Language Processing
- Deep Learning
- Mathematical methods

Deliverables

Milestones

1. Investigation:

Includes understanding the documentation of PyAutoLens more deeply with simulating more complex models of lenses and then reading both the papers more thoroughly for the better understanding of all terminology and CNN approaches used earlier in the project.

A. Getting familiar with the code.

Timeline period: 13th April - 17 May

2. Coding & Documentation:

A. Community bonding. **17 May - 7 June 2021**
B. Understanding DeepLense functionality and improving interface to PyAutoLens. **7 June - 16 August 2021**

Engagements during Summer

I have no engagements during summer, I'll be staying back at home for most of it. I have mentioned my typical working hours above and on an average will be able to spend 40 hours per week on the project.

Work/Internship/Academic Experience

- **Data-Science Intern: SkyQuest Technology**

  Scraped the news and press releases of top 1000 pharmaceutical companies. Build the custom sentiment analysis of the model using that data, with the help of NLP and spacy and pytorch.

- **Campus Activities**

  I am pursuing Bachelor of Technology in Engineering physics from Indian Institute of Technology, BHU Varanasi. I am currently in my sophomore year.

  Bagged First Position in Supernova, an event that included Case-studies, Paper-presentations, Quizzes, Observational Astronomy in India's second largest Tech-fest i.e. Technex, IIT BHU.