Python Module for Missing Data Imputation

Stretch Goal – Complete 2 Deep Learning Implementation for Missing Data (in PyTorch)

Mandatory Goal – Complete Denoising Autoencoders (DAEs) based model (*Gondara and Wang (2017))*

Detailed Plan

Start Date: June 17th, 2019

Weeks 1,2: Functional Implementation of *Gondara's* DAE Model

Weeks 3,4: Testing on Datasets mentioned in the paper (comparison will be done to MICE performance)

Weeks 5,6: Testing on other Datasets

Weeks 7,8: Model tuning and Code clean-up (if performance is great, if not work on improving it)

Weeks 9,10: Implementation of GAN based model

End Date: August 23rd, 2019

Remarks: Time gained, if any, in the initial implementation (DAEs) or in any steps, will be used in working on the GAN implementation.

Particulars

Git Repo - https://github.com/ambareeshsrja16/Python-Module-for-Missing-Data-Imputation

Primary Resource: MIDA: Multiple Imputation using Denoising Autoencoders

Progress Updates:

1. Weekly reports citing challenges faced, solutions for the same and progress made will be mailed to PI (Dr. Jose Unpingco) compulsorily on the Friday of every working week.
2. Presentations (short summary of the work) will be made compulsorily after Week 5 and Week 10