



FACULTY MENTOR

Unpingco, Jose

PROJECT TITLE

Python Package for Multivariate Imputation by Chained Equations

PROJECT DESCRIPTION

Multivariate imputation is used for a wide variety of missing data problems in machine learning and data privacy. This method specifies the model by a set of conditional densities for every incomplete data variable. Although these method has been implemented for SAS, R, Stata, and S-PLUS, there is only fragmentary work for the Python language. The purpose of the project is to implement these methods as an open source, freely available Python module focused on data privacy via synthetic data imputation.

INTERNS NEEDED

2 MS Students

PREREQUISITES

ECE 143, 225A, 275A, 275B, 109