Midterm Project

Submission type: Individual
Deadline: November 1st, 2020 (end-of-the-day)

This midterm project replaces midterm exam.

1. <u>***Task Description:***</u> Given train dataset "train_data.csv", find the best model for the below task
   A. **Input Features**: Morphological phenotypes
      i. Colum D-BU (ST***CV) : Cortical Volume of 70 brain regions
      ii. Colum BV-EM (ST***TA) : Average Thickness of 70 brain regions
      iii. Beware of missing features
   B. **Task (3-logit Regression)**: Predict the cognitive assessment scores of subjects for test dataset "sample_evaluation_data.csv"
      i. Colum A (ADAS11): Alzheimer's Disease Assessment Scale (11 questions version)
      ii. Colum B (ADAS13): Alzheimer's Disease Assessment Scale (13 questions version)
      iii. Colum C (MMSE): Mini-Mental State Examination

2. **Project Requirement**
   A. Perform Bayesian linear regression based on the given train dataset "train_data.csv"
      i. Any form of basis function is allowed
   B. Do Bayesian model selection
   C. Analyze and discuss your models and results in markdown cells
   D. No use of library for Bayesian inference is allowed
   E. TAs will assess your model with a test dataset, and grade your project accordingly

3. **Implementation Requirement**
   A. Use the Google Colab (https://colab.research.google.com/)
   B. (**IMPORTANT**) Only submit the .ipynb file and model weights file
      i. In the first markdown cell, write your name, ID, and the link to your Colab

4. **Grading Factors**
   A. Evaluation scores on test
      i. Ground truth of test dataset will not be available to students
      ii. Scores will be ranked and used as a reference for grading
   B. Comparison, analysis, discussion etc.
      i. Write in English
      ii. Be thorough and precise

TA's note: Beware of plagiarism (we use proprietary plagiarism checkers in addition to BlackBoard plagiarism checker). Ask the TAs if you have any questions.