Heinz 95-845: Project Proposal

Sarah Cho jscho/jscho@andrew.cmu.edu
Heinz College
Carnegie Mellon University
Pittsburgh, PA, United States

Maggie Lu YAOL4/yaol4@andrew.cmu.edu
Heinz College
Carnegie Mellon University
Pittsburgh, PA, United States

Askari Shah ANDREWID/email@address.edu
Heinz College
Carnegie Mellon University
Pittsburgh, PA, United States

1. Proposal Details (10 points)

Please provide information for the following fields. Your proposal write-up should be no more than 2 pages.

1.1 What is your proposed analysis? What are the likely outcomes?

There have been studies linking low socio-economic status(SES) with the prevalence of depression, and in the United States, we see the difference in food choices becoming more and more pronounced among those who have high disposable income and those who are living below the poverty line. Even though there have been numerous studies proving the link between nutrition and physiological health, the link between food and mental health has only been recently discussed.

For our project, we would like to examine how SES and nutrition affect one's mental health. We aim to explore the relationships and interactions between an individual's SES and her nutritional intake. Additionally, we would like to compare SES and nutritional factors on their effectiveness as predictors for depression. We predict that depression can be linked to both SES and nutrition, but we are not certain whether the combination of both can be a better predictor than either one.

1.2 Why is your proposed analysis important? Maggie

Our analysis would be able to provide a holistic view on how SES and nutrition affect one's mental health. Most studies in this field have only been focusing on

1.3 How will your analysis contribute to existing work? Provide references, *e.g.*, see: ?.Sarah

1.4 Describe the data. If applicable, please also define Y outcome(s), U treatment, V covariates, and W population. Maggie

1.5 What evaluation measures are appropriate for the analysis? Which measures will you use?Askari

Appropriate measures for analysis include confusion matrix (accuracy, precision, recall), ROC curve, Precision/Recall curves and statistical analyses techniques (Confidence intervals of errors, hypothesis tests of models). We intend to use confusion matrix, ROC curves and Precision/Recall curves. Depending on our data and class distributions along with our proposed use case, we will pick the most suitable/relevant measures from the listed appropriate measures.

1.6 What study design, pre-processing, and machine learning methods do you intend to use? Justify that the analysis is of appropriate size for a course project.Maggie

dt, lr, bagging, boosting, k-nn, neural network

1.7 What are possible limitations of the study? Sarah

1.8 Who will your analytic pipeline? In one or two sentences, describe an example of its use. Sarah

References