Mitchell Kolb
CPTS 315
Course Project Proposal
October 23th, 2022
WSU

1.) Data Mining Task: What is your data mining task? This task could be a series of exploratory questions that you want to investigate or analyze. What is your motivation behind choosing this task for your project?

I am choosing to work on the Divorce Predictors Data Set from the UC Invine Machine Learning Repository. I want to work on this dataset because the effort that was put into this dataset is very thorough and I think I can draw some conclusions about why couples get a divorce. This data has a lot of questions so i can use my knowledge about phycology to match questions together to help with reducing the amount of data.

2.) Dataset: What is the source of your data? Provide a link to your data source if you acquired it online.

The data from first observation is kind of basic because it's a series of questions with answers 1-4. 1 being bad/low to 4 being excellent/high. This data set is 170 users big which is kind of small compared to other datasets i've seen where they can range to tens of thousands. The interesting part about this data is that each user was asked 52 questions, so their is lots of data per user to go through. The link to my dataset is below.

3.) Methodology: How will you solve the data mining task? You should have some idea of the algorithms or software tools you plan to investigate. Please feel free to use existing data mining and machine learning tool kits (e.g., Weka, Scikit-Learn) as needed for your project.

My thoughts on how I will go about solving this task is to first get a general idea on the scores that people submitted. I would like to see what questions are the lowest, highest, most deviated from. Then I will try to consolidate the questions that are asked into into groups which will be used to determine types of people, wants, and goals to see if their is a relationship. I think that the Apache Mahout toolkit could be useful for this project because it focuses on clustering classification of data.

4.) Final product: What will be the outcome of this project? How will you measure the success of your course project? Will this project help you explore or learn something new?

My hope for the outcome of this project is to use the questions that were asked to determine what people think predicts divorces in couples and how those thoughts translate to actual research that is completed by licensed professionals.

Links:
https://archive-beta.ics.uci.edu/ml/datasets/divorce+predictors+data+set