ITEC 620

Fall 2022

**Assignment 2**

Due November 7, 5:15 PM

Please submit your assignment via Canvas, **as a single Word file**.

**For all questions that involve generating R output, you must include both the output itself and the command or sequence of commands used to generate it to receive credit. Include ONLY the commands used to generate that output; do not include entire scripts.**

**1. Association Rules**

Principal Investment Group offers their clients a variety of different stock and bond funds. They use a relatively “hands-off” approach, and let clients choose whatever mix of funds they’d like. They would like to analyze their clients’ choices of funds to see if any obvious patterns emerge, and to identify opportunities to help them improve their portfolios.

The csv file “Principal” contains information about client investments. Each row represents a client, and each column represents a stock or bond fund. A “1” in a cell indicates that the fund comprises at least 10% of that client’s portfolio.

**a)** Create a set of association rules, setting the minimum support to 0.03 and minimum confidence to 0.08. Sort the rules from highest confidence to lowest confidence, and show the resulting list of rules.

**b)** There are many association rules in the table with lift ratios that are well below 1. For example, the association rule: “If BondIndex, then StockIndex” has a lift ratio of 0.61. What does that reveal about the relationship between these two funds?

**2. Forecasting**

The csv file “PhonePlans” contains annual data from the International Telecommunication Union on the total number of land lines and mobile phone plans (in millions) from 2005-2021. The data are split between developed countries and developing countries.

**a)** Show a plot of the “LandLinesDeveloping” time series.

**b)** Create (single) exponential smoothing models for the “LandLinesDeveloping” and “MobilePlansDeveloping” time series, and for each model, show a plot that contains both the time series data and the forecasts made by the model.

**c)** Using the model from part b, predict the number of land lines and number of mobile plans in developing countries in 2022.

**d)** An analyst suggests that there is a simple way to improve the predictions made by these two models. Do you agree? Explain your answer briefly (1-2 sentences).