**ALY6040 90248 Data Mining Applications SEC 01 Summer 2023 CPS [BOS-D-HY]**

**Module 33 Assignment — Technique Practice**

A black and white logo

Description automatically generated

**NORTHEASTERN UNIVERSITY**

**College of Professional Studies, Boston, MA, 02215.**

**Submitted by**

Hanchan Xu  [xu.hanc@northeastern.edu](mailto:xu.hanc@northeastern.edu)

Zihan Ma [ma.zihan1@northeastern.edu](mailto:ma.zihan1@northeastern.edu)

**Instructor**

 Prof. Kasun Samarasinghe

**Date**

07/26/2023

**Locally Inspired** **XN Project: Pros & Cons**

**Part 1**

**Introduction**

The provided code makes use of R and various libraries for the analysis and prediction of mushroom edibility based on various characteristics. The analysis is divided into two parts, first focusing on the entire mushroom data set, and then focusing on the "gill-size" characteristic.

**Data Preparation**

The data for this analysis was imported from an Excel file ("mushrooms.xlsx") using the readxl library. The data contains 8124 observations and 23 characteristics, including mushroom class, cap shape, cap surface, cap color, bruises, odor, and many others. All characteristics are recorded as categorical variables. The class variable, which can either be "p" (poisonous) or "e" (edible), is our target variable for classification.

**Data**

**Data**

**Data**

**Data**

**Data**

**Part 2**

**Data**

**Data**

**Data**

**Data**

**Data**

**Data**

**Data**

**Data**

**Data**

**Data**

**References:**