|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ISQA/CSCI 4890 – Homework 4** | | | | |
|  | **Question** | **Requirement** | **Due Date** | **Note** |
|  | **Task 1:**  Use the R software (Weka or other software) to run question #1 and #4 (pg 93). Print out the results:  For Question #1, use the first 1000 records as a training set and the second 1000 records as the testing set (you will find the Insurance Fraud data set on Canvas)  For Question #4, use the first 600 records as a training set and the last 50 as the testing set (you will find the Loan Application data set on the Canvas) | Please turn in work file to yshi@unomaha.edu. | 11/09/2020 18:00 |  |
|  | **Task 2:**  Use the Logistic Regression (R, Weka or other software) to run question #5 (pg 116) and print the results, including the Confusion Matrix of the test.  Note:  Step 1: Transform the first 2000 records of Fraud Raw Data Sets according to #5  Step 2: Use the first 1000 records as a training set and the second 1000 records as the testing set. |  |  |  |
|  |  |  |  |  |