

# **CS550: Massive Data Mining and Learning**

## **Homework 2**

Due 11:59pm Monday, March 21, 2022

# Submission Instructions

**Assignment Submission:** Include a signed agreement to the Honor Code with this assignment. Assignments are due at 11:59pm. All students must submit their homework via Canvas. Students can typeset or scan their homework. These questions require thought but do not require long answers. Please be as concise as possible. You should submit your answers as a writeup in PDF format, for those questions that require coding, write your code for a question in a single source code file, and name the file as the question number (e.g., question\_1.java or question\_1.py), finally, put your PDF answer file and all the code files in a folder named as your Name and NetID (i.e., Firstname-Lastname-NetID.pdf), compress the folder as a zip file (e.g., Firstname-Lastname-NetID.zip), and submit the zip file via Canvas. For the answer writeup PDF file, we have provided both a word template and a latex template for you, after you finished the writing, save the file as a PDF file, and submit both the original file (word or latex) and the PDF file.

**Late Policy:** The homework is due on 3/21 (Monday) at 11:59pm. We will release the solutions of the homework on Canvas on 3/25 (Friday) 11:59pm. If your homework is submitted to Canvas before 3/21 11:59pm, there will no late penalty. If you submit to Canvas after 3/21 11:59pm and before 3/25 11:59pm (i.e., before we release the solution), your score will be penalized by  $0.9^k$ , where  $k$  is the number of days of late submission. For example, if you submitted on 3/24, and your original score is 80, then your final score will be  $80 \times 0.9^3 = 58.32$  for  $24 - 21 = 3$  days of late submission. If you submit to Canvas after 3/25 11:59pm (i.e., after we release the solution), then you will earn no score for the homework.

**Honor Code:** Students may discuss the homework problems with peers. However, each student must write down their solutions independently to show they understand the solution well enough in order to reconstruct it by themselves. Students should clearly mention the names of all the other students who were part of their discussion group. Using code or solutions directly obtained from the web or others is considered an honor code violation. We check all the submissions for plagiarism. We take the honor code seriously and expect students to do the same.

Discussion Group (People with whom you discussed ideas used in your answers):

On-line or hardcopy documents used as part of your answers:

I acknowledge and accept the Honor Code.

(Signed)\_\_\_\_\_

If you are not printing this document out, please type your initials above.

Answer to Question 1(a)

Answer to Question 1(b)

Answer to Question 1(c)

Answer to Question 1(d)

Answer to Question 1(e)(a)

Answer to Question 1(e)(b)

Answer to Question 1(e)(c)

Answer to Question 1(e)(d)

Answer to Question 2(a)

Answer to Question 2(b)

Answer to Question 2(c)(a)

Answer to Question 2(c)(b)

Answer to Question 3(a)

Answer to Question 3(b)

Answer to Question 4(a)

Answer to Question 4(b)