**Week 6 in-class activities/Lab**

**Task1a: Interpreting Logistic Regression model**

Given a logistic regression model



Answer the following questions:

(you may use the provided “logistic regression” notebook and AI assistant.)

1. Thomas studied for two hours and did not attend the review session. What is his (1) log odds, (2) odds, and (3) likelihood of passing the exam?
2. If Thomas goes to the review session, what is the updated 1) log\_odds, (2) odds, and (3) likelihood of passing the exam?
3. If Thomas studied more or less hours, would the answer change?
4. How would you interpret the coefficient of review\_session (1.5) from the above experiment?
5. Using similar reasoning, how would you interpret the coefficient of hours\_studied (0.8)
6. How would you interpret the intercept?
7. For someone who studied 8 hours, would you recommend him/her to attend the review session?
8. What type of students seems to benefit most from the review session?

**Task 1b: Build a logistic regression model**

Using the dataset “student\_data.csv,” write code to (1) create a visualization of the data, (2) fit a model using logistic regression, (3) output model coefficients and performance metrics such as accuracy and AUC and ROC; **NOTE: For this exercise, you will train and test on the same given dataset, instead of doing train/test split. Make sure you give the correct GPT prompt.**

**Task 2: Understanding and Prevent Overfitting in the context of SVM**

Write code to fit a Support Vector Machine model using (1) linear kernel and (2) RBF kernel. For the RBF kernel, use grid search to find the best gamma parameter using k-fold cross-validation.

Submission: 1(a) writeup in a doc; 1(b) and (2) Python Notetook uploaded to GitHub and submit a link to Blackboard; link to chatGPT log