Week 6 in-class activities/Lab

Harshith Gudapati
BJ84775

Task1a: Interpreting Logistic Regression model

Given a logistic regression model

$$\ln\left(\frac{p}{1-p}\right) = -3 + 0.8 \times \text{Hours\_Studied} + 1.5 \times \text{Review\_Session}$$

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

a. 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?
Answer – Based on the code given, the answers are as follows:-

```
log_odds_0 = log_odds(hours_studied, 0)
probability_0 = logistic_function(log_odds_0)
print('Log_odds for 0_rev_ses =',log_odds_0)
print('odds for 0_rev_ses =',np.exp(log_odds_0))
print('Pass_likelihood for 0_rev_ses=',probability_0)
```

```
Log_odds for 0_rev_ses = -1.4
odds for 0_rev_ses = 0.2465969639416065
Pass_likelihood for 0_rev_ses= 0.19781611144141825
```

b. If Thomas goes to the review session, what is the updated 1) log_odds, (2) odds, and (3) likelihood of passing the exam?
Answer – Based on the code given, the answers are as follows:-

```
log_odds_1 = log_odds(hours_studied, 1)
probability_1 = logistic_function(log_odds_1)
print('Log_odds for 1_rev_ses =',log_odds_1)
print('odds for 1_rev_ses =',np.exp(log_odds_1))
print('pass_likelihood for 1_rev_ses =',probability_1)
```

```
Log_odds for 1_rev_ses = 0.10000000000000009
odds for 1_rev_ses = 1.1051709180756477
pass_likelihood for 1_rev_ses = 0.52497918747894
```

c. If Thomas studied more or less hours, would the answer change?
   - **Answer – In the equation the coefficient of the hours_studies is 0.8. so, whenever there is an increase in the study hours the log odds will increase finally causing the probability to increase.**
   - **The model shows positive correlation between study hours and passing**

d. How would you interpret the coefficient of review_session (1.5) from the above experiment?
   - **Answer – When the student attends the review session it increases the log odds of passing by 1.5.**

- **When u caluculate the odds, attending the review session makes passing about 4.5 time more likely than not attenging.**

e. Using similar reasoning, how would you interpret the coefficient of hours_studied (0.8)
Answer – As the coefficient is 0.8 for hours_studied. Increase in the hours studied increases the log odds of passing by 0.8 and make passing about 2.2 times more likely.

f. How would you interpret the intercept?
   - **Answer – The intercept value give is -3, this represents the log odds of passing when both hours_studied and review session are zero.**
   - **Which mean when thomas did not study and does not attend the session. The log odds of passing would be -3.**
   - **Which means he is 5% likely to pass.**

g. For someone who studied 8 hours, would you recommend him/her to attend the review session?
Answer – we can know this by calculating the likelihood.

```
Log_odds for 0_rev_ses = 3.4000000000000004
odds for 0_rev_ses = 29.964100047397025
Pass_likelihood for 0_rev_ses= 0.9677045353015495
```

```
Log_odds for 1_rev_ses = 4.9
odds for 1_rev_ses = 134.28977968493552
pass_likelihood for 1_rev_ses = 0.9926084586557181
```

As u can see here when the student does not attend the session the likelihood to pass Is 96.7% and when he attends the session the likelihood is 99.2%. The difference between the probabilities is less and the hours studies is already high. The student may not need to attend the session to pass the exam.

h. What type of students seems to benefit most from the review session?
Answer – from my interpretation of the model, the students whose hours of study are low and for the students whose hours are between 2-5 (medium) may benefit from attending the review session.