

Lab 04

1. Consider binary classification problem where we want to predict whether a student will pass or fail based on their study hours; the logistic regression model has been trained and the learned parameters are $a_0 = -5$ (intercept) and $a_1 = 0.8$ (coefficient for study hours).

1. Write the logistic regression equation for problem. The logistic regression equation for predicting the probability of passing based on study hours is given; $a_0 = -5$, $a_1 = 0.8$

Substitute values into equation

$$P(\text{pass} | x) = \frac{1}{1 + e^{-(-5 + 0.8x)}}$$

2. Calculate the probability that student who studies for 7 hours will pass.

Student who studies for $x = 7$ hours.

$$P(\text{pass} | 7) = \frac{1}{1 + e^{-(-5 + 0.8(7))}} = 0.6457 \approx 64.5\%$$

3. Determine the predicted class for student based on threshold of 0.5

If $P(\text{pass} | x) \geq 0.5$ predicted class = pass
If $P(\text{pass} | x) < 0.5$ predicted is fail

for $x = 7$ hours

$$P(\text{pass} | x) = \frac{1}{1 + e^{-(-5 + 0.8(7))}} = 0.6457 \geq 0.5$$

~~∴ Predicted class is pass~~

II Consider $z = [2, 1, 0]$ for three classes. Apply softmax fun to find the probability value of 3 classes.

Softmax formula:

$$z_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad \text{where } k \text{ is no. of classes}$$

$$e^{z_1} = e^2 \approx 7.389 \quad e^{z_2} = e^1 \approx 2.718$$

$$e^{z_3} = e^0 = 1$$

$$\sum_{j=1}^k e^{z_j} = e^2 + e^1 + e^0 \approx 7.389 + 2.718 + 1 = 11.107$$

$$\rightarrow \text{Softmax}(z_1) = \frac{e^{z_1}}{\sum_{j=1}^k e^{z_j}} = \frac{7.389}{11.107} \approx 0.665$$

$$z_2 = \frac{2.718}{11.107} \approx 0.245$$

$$z_3 = \frac{1}{11.107} \approx 0.090$$

Thus the softmax probabilities for 3 classes are approx 0.665, 0.245 & 0.090.

Lab 05

1. After building the logistic regression model, write the answer for the following questions.

a) The key variables impacting the employee retention are:

→ Satisfaction lower: Lower satisfaction increases attrition.

→ Time spent in company: Employee with 51 years tend to leave.

→ Salary: Low salaries lead to higher turnover.

→ No. of projects & avg monthly hours: Very high or low values affect retention.