Michigan State University **Jan 27, 2023**
CMSE381 - Data Science **Dr. Xie**

CMSE381 - Quiz 2

I will adhere to the Spartan Code of Honor in completing this assignment.

Signed: _____ Print Name: _____

1. We are training a linear model to predict sales from TV, radio, and newspaper advertising. We get this output from our code.

	Coefficient	Std. Error	t-statistic	p-value
Intercept	6.7502	0.248	27.23	< 0.0001
TV	0.0191	0.002	12.70	< 0.0001
radio	0.0289	0.009	3.24	0.0014
TV×radio	0.0011	0.000	20.73	< 0.0001

What is the equation of the learned model?

$Y_{Sale} = 6.7502 + 0.0191 X_{TV} + 0.0289 X_{radio} + 0.0011 X_{TV} X_{radio}$

2. We want to build a model using a student's home state (Michigan, Ohio, or North Carolina) to predict his/her GPA. Write down your model.

$Y_{GPA} = \beta_0 + \beta_1 X_{Ohio} + \beta_2 X_{Michigan} + \epsilon$

$X_{Ohio} = \begin{cases} 1 & \text{if from Ohio} \\ 0 & \text{o.w.} \end{cases}$ $X_{Michigan} = \begin{cases} 1 & \text{if from M.} \\ 0 & \text{o.w.} \end{cases}$

3. Using a set of training data, we fit a model $Y = \beta_0 + \beta_1 X + \epsilon$ and obtain $\hat{\beta}_0 = 3.1$, $SE(\hat{\beta}_0) = 1.2$ $\hat{\beta}_1 = 2.4$, $SE(\hat{\beta}_1) = 1.1$, and $R^2 = 0.73$. What is the percentage of variance that can not be explained by this linear model?

$1 - R^2 = 1 - 0.73 = 0.27$