1.

Null hypothesis: coefficient(TV) = coefficient(Radio) = coefficient(newspaper) = 0

We can reject the null hypothesis, and come to a conclusion that sales is related to advertising budgets for TV and radio rather than newspaper.

2. (still needs to be proved)

KNN classifier aims at classifying an observation to a class, which is an unsupervised learning method; KNN regression wants to predict a y-value based on the y-value of the n nearest observations around the target observation. It is a supervised learning method.

3.

Y = 50 +20\*GPA +0.07\*IQ+35\*Gender (1 if female, 0 if male) +0.01\*GPA\*IQ-10\*GPA\*Gender

a)

Only iii is correct. Because when the value of IQ and GPA are fixed, the difference between y(female) and y(male) would be 50+35\*1-10\* GPA\*1-( 50+35\*0-10\* GPA\*0) = 35-10\* GPA. In this sense, if GPA is high enough (> 3.5), then the value of above formula (35-10\* GPA) would be negative, thus giving male a higher income.

b)

y = 50 +20\*GPA +0.07\*IQ+35\*Gender (1 if female, 0 if male) +0.01\*GPA\*IQ-10\*GPA\*Gender

= 50+20\*4+0.07\*110+35\*1+0.01\*4\*110-10\*4\*1 = 50+80+7.7+35+4.4-40 = 137.1

c)

We cannot conclude on the interaction effect just basing on the coefficient value. If the standard deviation is extremely small, which leads to a big p-value, the existence of interaction effect could be significant.

4.

a)

The cubic model has a smaller training RSS, because the cubic model is more flexible so it could provide a tighter fit.

b)

Linear model would give a smaller test RSS. Because it better mimics the true relationship.

c)

In this case, a cubic model would give a smaller training RSS.

d)

It depends. If the true relationship is closer to a linear one, the linear model would give smaller RSS, otherwise the cubic model will give a smaller test RSS.