

SP3

**1 Q1 (ISLR 3.4)**

**2 Q2 (ISLR 2.1)**

**3 Q3 (ISLR 2.1)**

**4 Q4 (ISLR 2.5)**

**5 Q5**

This question should be answered using the `Carseats` data set in the ISLR package. You may use `help(Carseats)` to learn more about the data set.

**5.0.1 a.**

Split the data into a testing and training set.

**5.0.2 b.**

Train a linear regression model and evaluate on the test data.

**5.0.3 c.**

Fit a KNN regression on the training data and evaluate on the test data using  $K = 10$ .

**5.0.4 d.**

Use cross-validation to choose an appropriate value for  $K$ .

**6 Q6**

**6.0.1 a.**

Generate  $n = 1000$  values  $X_1$  and  $X_2$  from a uniform distribution between  $-1$  and  $1$ . Generate  $Y = X_1 X_2 + \varepsilon$  where  $\varepsilon$  are i.i.d from a  $N(0, .1)$  distribution.

**6.0.2 b.**

Split the data into testing and training and fit a KNN regression model using  $K = 10$ . Calculate the predictions on the test data.

**6.0.3 c.**

Using the same split fit a linear regression model and calculate the predictions using the testing data.

**6.0.4 d.**

Plot the true values for  $Y$  in the test set (on the  $x$ -axis) against the predictions for the two models (on the  $y$ -axis). Which model does better? Why?