

بسم الله الرحمن الرحيم

سری چهارم تمرین درس تحلیل و طراحی آماری فرایندها

Instructions:

All of previous instructions + 20% reduce score for late answers.

Explain about all of your plots.

You can use "trainControl" function for cross validation in R. caret package

You can use sample for test and train.

Use set.seed(1) in those questions are necessary.

1. In Using PimaIndiansDiabetes:

a) Use 5-fold cross-validation to compare:

GLM: $y \sim \text{glucose} + \text{mass} + \text{age}$

GAM: $y \sim s(\text{glucose}) + s(\text{mass}) + s(\text{age})$

Explain which model is your preference and why?

b) Plot the smooth functions of Gam and interpret.

2. From Survival package data of lung,

a) Response: time (survival time)

Predict with: age, ph.ecog, sex

using LM.

b) By plot check your linearity and if it needed use gam.

c) Use spline if needed.

d) As a whole and from your theoretical concepts, you prefer gam or spline? Then explain about it in your model.

3. In library mlbench, data BreastCancer:

Use 2:10 columns as your covariates. Convert outcome to numeric (Benign = 0, Malignant = 1)

b) Fit logistic regression using:

Cl.thickness + Cell.size + Bare.nuclei. calculate OR and interpret it completely.

c) Check nonlinearity and if needed Fit GAM with splines

d) Evaluate accuracy

e) In gam Plot probability vs each predictor.(what we talk about in class)

4. Data epil, in MASS library:

A clinical study measured seizure counts for epilepsy patients before and after treatment. Use trt, age and base(baseline seizure count) as your predictors and Fit model based on 70% train and 30% test.

b) interpret confidence interval and your coefficients, and Rmse.

c) Plot predicted seizure counts vs age, with color = treatment.

5. In data of cats from MASS library:

Predict Hwt: heart weight by Bwt: body weight and Sex.

Use 80% train and 20% test.

Consider the linearity by visual and if it's not use gam.

By rmse and aic, interpret your selection.

Plot predicted Hwt vs Bwt for both models and color is another variable (sex).