Clara

Anqi Chi SID:460204008

22/10/2018

LDA

Table 1: Coefficients of Predictors in LDA

	Estimated Coefficients
BMR	0.001249
Energy Intake	-0.434204
Total mins spent sedentary	-0.000046
Sex	2.093857
Whether currently on a diet	0.363241
Carbohydrate diet	-0.005032
Fat diet	0.001558
Protein diet	0.002124

Comment:

The above output suggests the following interpretations for each of the variables.

- People who have higher BMR are more likely to be obese.
- Lower EIBMR1(Energy intake) increases obesity probability (since the coefficient -0.434 is negative).
- Lower ADTOTSE(Total mins spent sitting or lying down) increases obesity probability.
- Sex = 1 for males. So females are more likely to be obese.
- People on a diet are more likely to be obese compared with people not on a diet
- People who have high fat or high protein diet type are more likely to be obese
- People who have high carbon diet type decrease the probability of obesity

CART

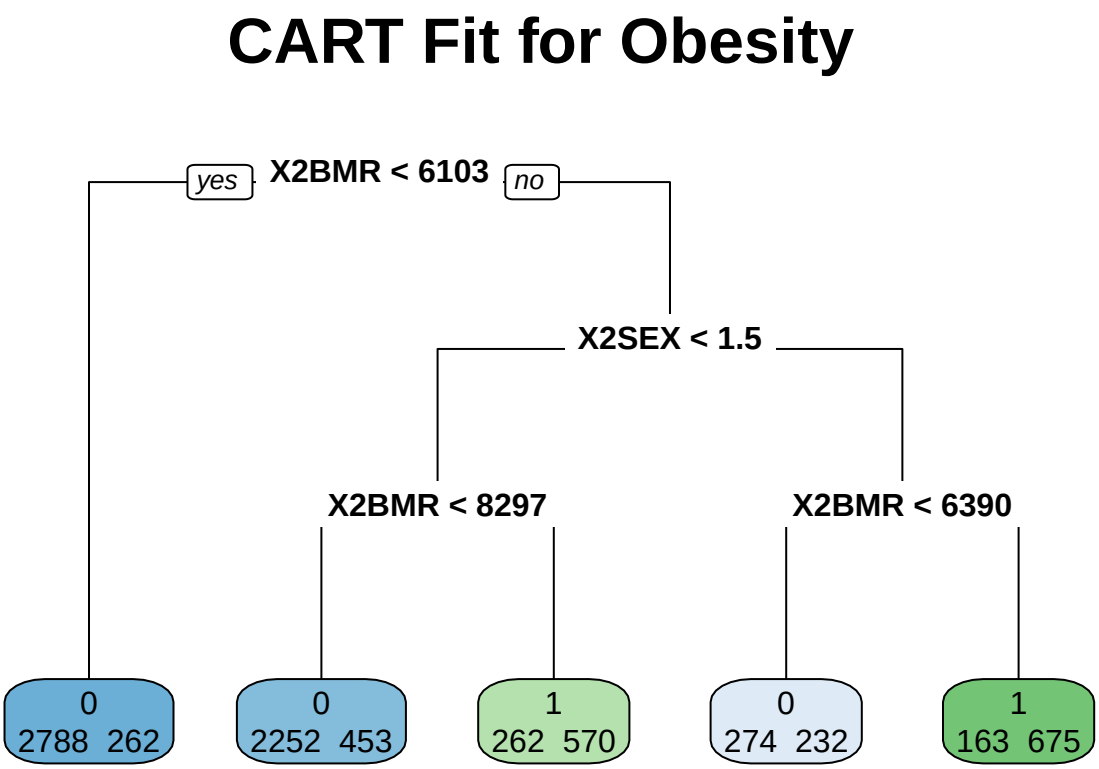

Table 2: Obesity Rate for Male with BMR greater than 6103

	6103 < BMR < 8297	BMR > 8297
estimated obesity rate	6.7%	68.5%

Table 3: Obesity Rate for Female with BMR greater than 6103

	6103 < BMR < 6390	BMR > 6390
estimated obesity rate	45.8%	80.1%

Comment:

- 91.4 percent of people(including male and female) whose BMR are less than 6103 are normal (not obese).
- 83.3 percent of male whose BMR are between 6103 and 8297 are normal.
- 68.5 percent of male whose BMR are greater than 8297 are obese.
- 54.2 percent of female whose BMR are between 6103 and 6390 are normal.
- 80.1 percent of female whose BMR are greater than 6390 are obese.

Table 4: Estimated Coefficients of the Logistic Regression

	Variable	Estimated Coefficient
(Intercept)	1	-17.770903
BMR	x_1	0.001879
EIBMR1	x_2	-0.711322
ADTOTSE	x_3	-6e-05
SEX	x_4	3.422641
BDYMSQ04	x_5	0.380154
CHOPER1	x_6	-0.010135
FATPER1	x_7	0.001507
PROPER1	x_8	-0.000998

Logistic regression

Comment:

The full model has the coefficeints for BMR, EIBMR1, ADTOTSE, SEX, BDYMSQ04, CHOPER1 as statisticlly signifficantly different from zero at the 0.05 level. The fitted model is

$$\begin{aligned}\text{logit}(p) = &- 17.770903 + 0.001879x_1 - 0.711322x_2 - 0.00006x_3 + 3.422641x4 \\ &+ 0.380154x_5 - 0.010135x_6 + 0.001507x_7\breve{}0.000998x_8\end{aligned}$$

where p is the probability of obesity.

where $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ represents BMR, EIBMR1, ADTOTSE, SEX, BDYMSQ04, CHOPER1, FATPER1, PROPER1 respectively.

summary

Table 5: Summary of the CV Errors

Methods	Errors
LDA	18.232
CART	17.551
Logistic Regression	17.778

The CV error for LDA is 18.232 percent. The CV error for CART is 17.551 percent. The CV error for glm is 17.778 percent