Domas Budrys – Assignment 6 CSCI5080

Question 1 (Time spent: 3h 30min):

a)

Eyes -because it has the highest information gain across all attributes.

b)

$Info(D) = -\frac{3}{6}log_2\left(\frac{3}{6}\right) - \frac{3}{6}log_2\left(\frac{3}{6}\right) = 1$

$Info_{height}(D) = \frac{3}{6}\left(-\frac{1}{3}log_2\left(\frac{1}{3}\right) - \frac{2}{3}log_2\left(\frac{2}{3}\right)\right) + \frac{3}{6}\left(-\frac{2}{3}log_2\left(\frac{2}{3}\right) - \frac{1}{3}log_2\left(\frac{1}{3}\right)\right) = .9183\ bits$

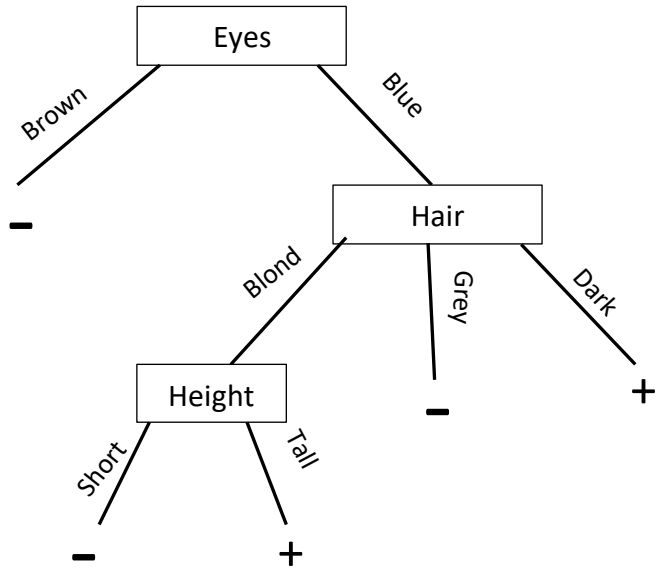
$\boldsymbol{Gain(height) = 1 - .9123 = .0817\ bits}$

$Info_{hair}(D) = \frac{3}{6}\left(-\frac{2}{3}log_2\left(\frac{2}{3}\right) - \frac{1}{3}log_2\left(\frac{1}{3}\right)\right) + \frac{2}{6}\left(-\frac{1}{2}log_2\left(\frac{1}{2}\right) - \frac{1}{2}log_2\left(\frac{1}{2}\right)\right) + \frac{1}{6}\left(-\frac{1}{1}log_2\left(\frac{1}{1}\right)\right) = .7925\ bits$

$\boldsymbol{Gain(hair) = 1 - .7925 = .2075\ bits}$

$Info_{eyes}(D) = \frac{4}{6}\left(-\frac{3}{4}log_2\left(\frac{3}{4}\right) - \frac{1}{4}log_2\left(\frac{1}{4}\right)\right) + \frac{2}{6}\left(-\frac{2}{2}log_2\left(\frac{2}{2}\right)\right) = .5409$

$\boldsymbol{Gain(eyes) = 1 - .5409 = .4591\ bits}$



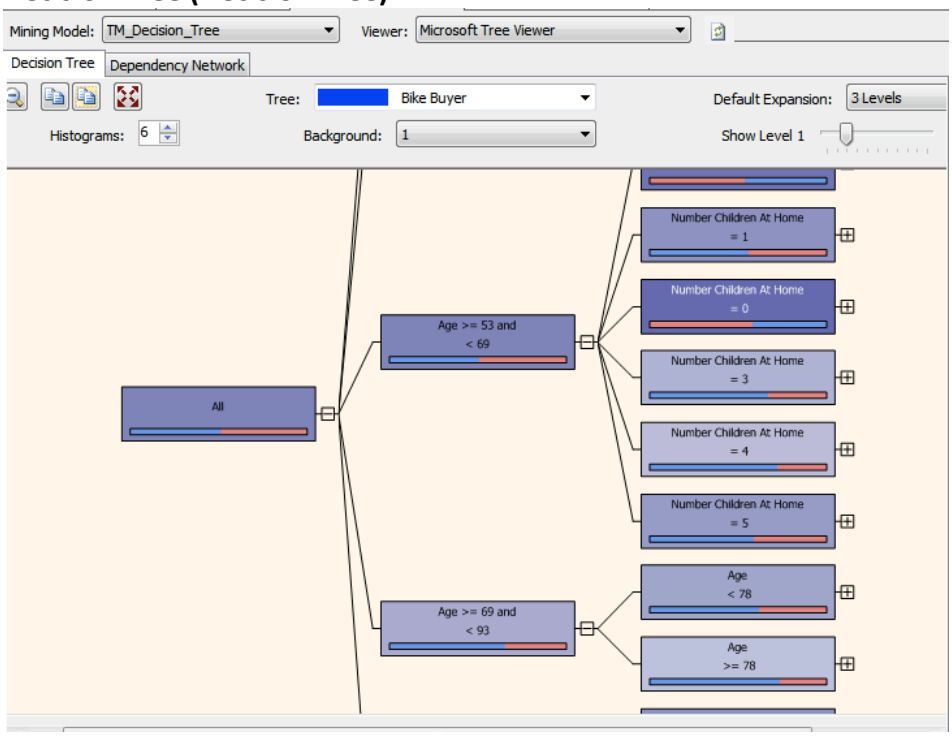
c)

Example	height	hair	eyes	Class
*X*7	short	gray	brown	–
*X*8	tall	dark	brown	–

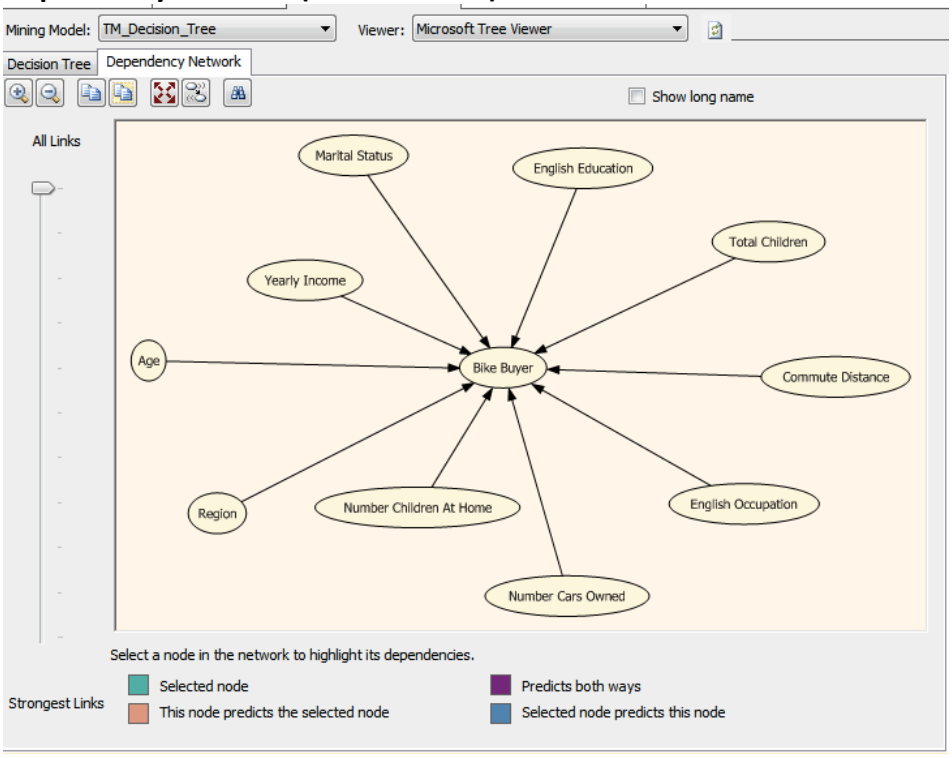
Question 2 (Time spent: 3h):

(2):

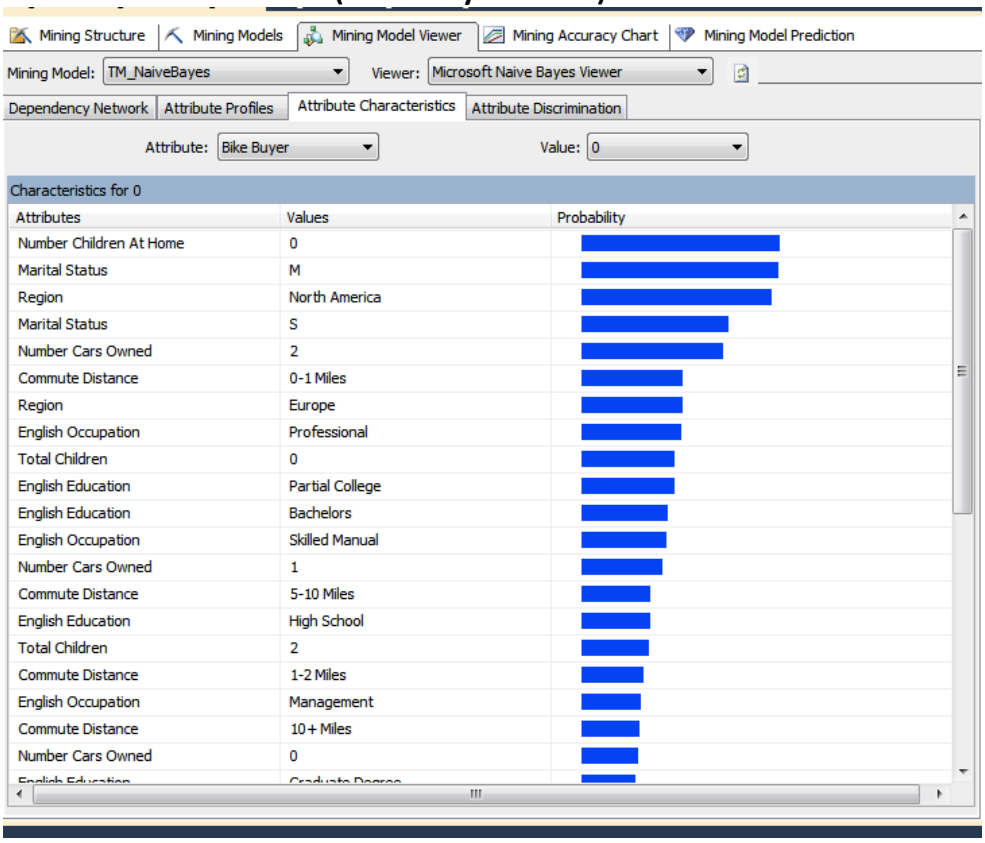
Decision Tree (Decision Tree)



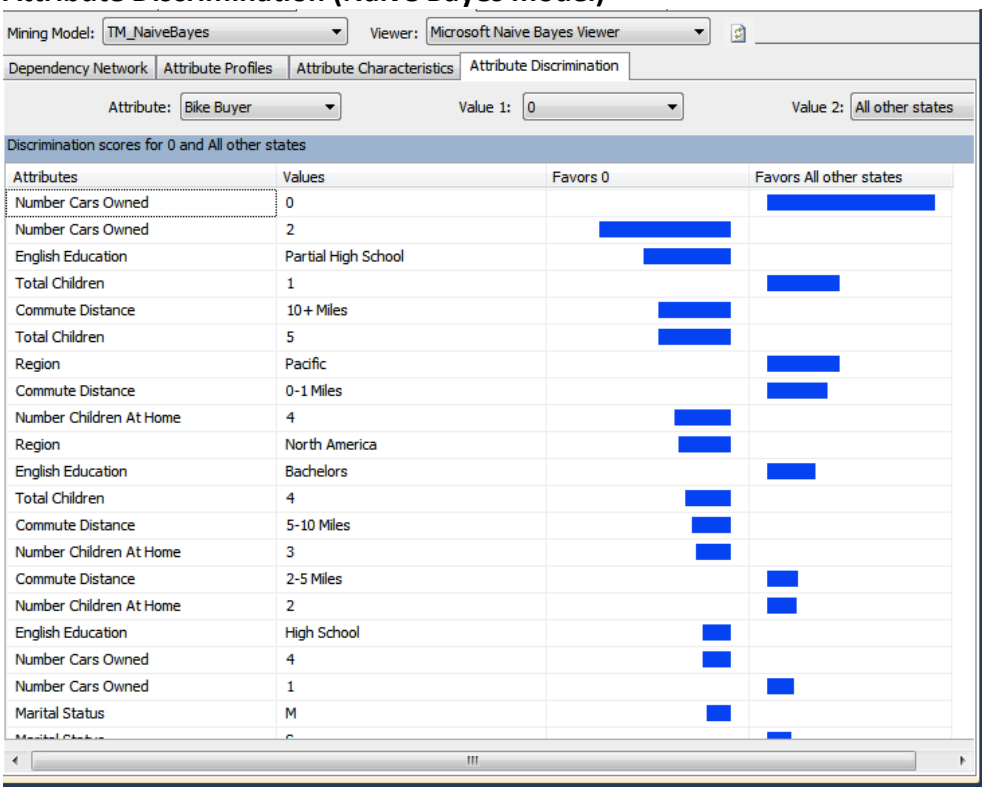
Dependency Network (Decision Tree)



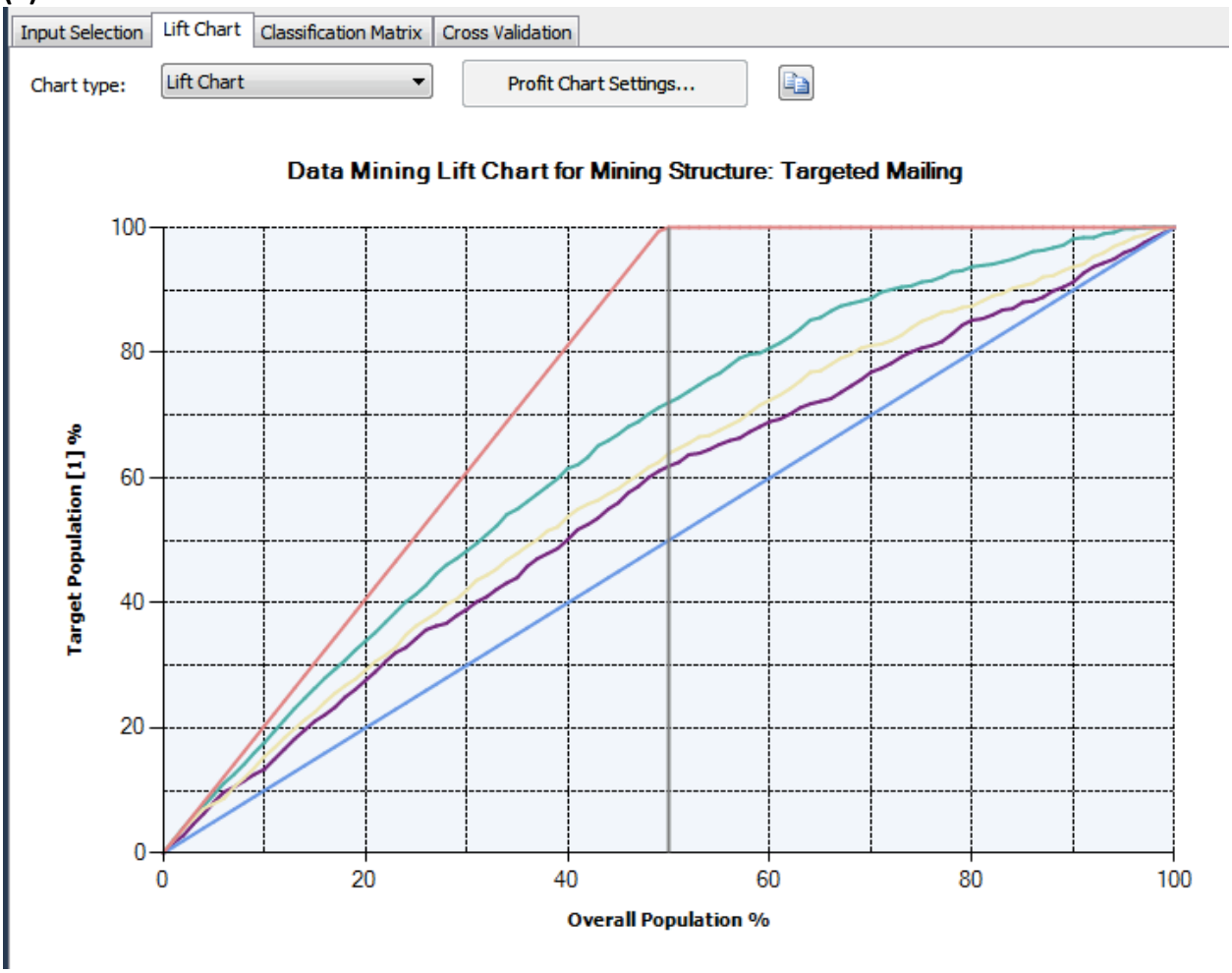
Attribute Characteristics (Naïve Bayes model)



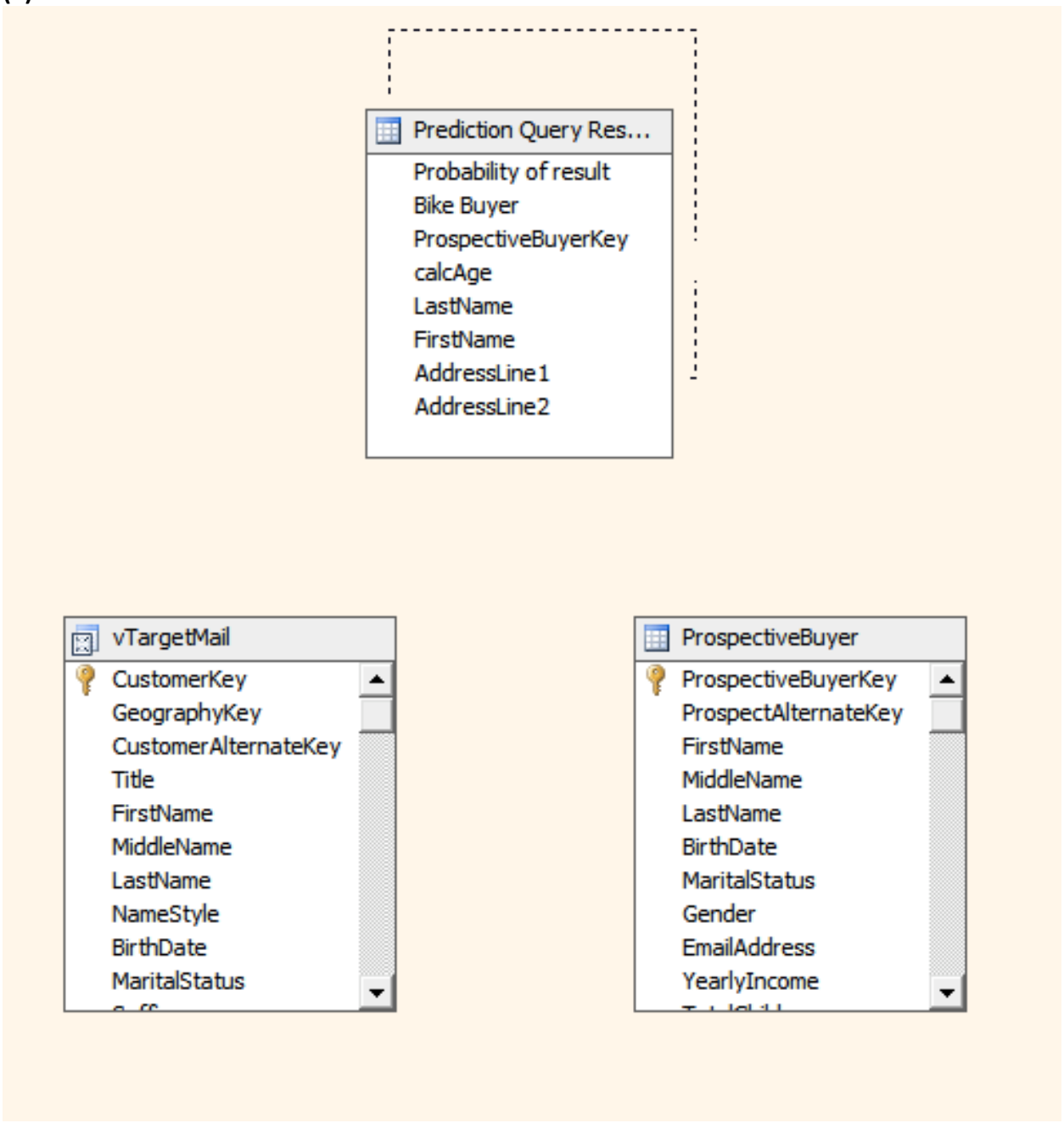
Attribute Discrimination (Naïve Bayes model)



(3):


Input Selection
Lift Chart
Classification Matrix
Cross Validation
Chart type:
Lift Chart
Profit Chart Settings...
Data Mining Lift Chart for Mining Structure: Targeted Mailing
Target Population [1] %
Overall Population %
0
20
40
60
80
100

(4)


Prediction Query Res...
Probability of result
Bike Buyer
ProspectiveBuyerKey
calcAge
LastName
FirstName
AddressLine1
AddressLine2
vTargetMail
CustomerKey
GeographyKey
CustomerAlternateKey
Title
FirstName
MiddleName
LastName
NameStyle
BirthDate
MaritalStatus
ProspectiveBuyer
ProspectiveBuyerKey
ProspectAlternateKey
FirstName
MiddleName
LastName
BirthDate
MaritalStatus
Gender
EmailAddress
YearlyIncome