UC San Diego Extension

Advanced Web Analytics: Harnessing the Predictive Power

Winter 2016

Homework#5

Date Given: Feb 29, 2016 Due Date: March 6, 2016

==========

1. Suppose you roll a pair of dice and compute the sum of the numbers you get (Vegas - Craps game).
   - What is the probability of getting a 7 as the sum?
   - What are the odds in favor of getting a 7 as the sum?

2. Following data is captured by Google Analytics of a few visitors.

Page views	Average Session Length	Session Count	Converted
6	18.61	23	0
7	10.1	17	1
7	13.19	14	1
7	14.46	22	1
7	15.28	20	1
7	15.56	16	1
7	9.98	24	1
8	13.41	10	1
8	19.89	24	0
8	14.79	24	0
9	8.08	12	1
9	9.5	21	1
9	14.44	18	0
10	8.44	13	1
10	8.98	25	1
10	9.12	21	1
10	9.64	24	1
11	18.11	15	0
11	17.15	22	1
14	19.39	17	1

Use Logistic Regression method and Excel to predict if a visitor with following characteristics will convert or not (cut-off probability = 0.5)

- Page Views = 13
- Average Session Length = 16.8
- Session Count = 24

Verify your answer using KNIME software.

Answer for problem 2:

Excel

beta-0	9.7163953
beta-1	0.3168432
beta-2	-0.6041295
beta-3	-0.1190833

Prediction						
Page views	Avg Session Length	Session Count	Converted	e Term	Probability = eTerm/(1+eTerm)	Convert
13	16.8	24		2.288696	0.695928108	1

==========

R

```
Coefficients:
     (Intercept)         Pageviews  AvgSessionLength      SessionCount
          9.7158            0.3168           -0.6041           -0.1191

Prediction
  Pageviews AvgSessionLength SessionCount Converted      pred cutoff
1        13             16.8           24        NA 0.6959309    0.5
```

==========

KNIME