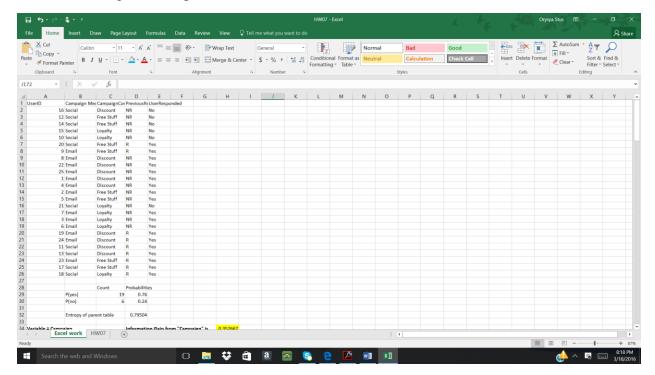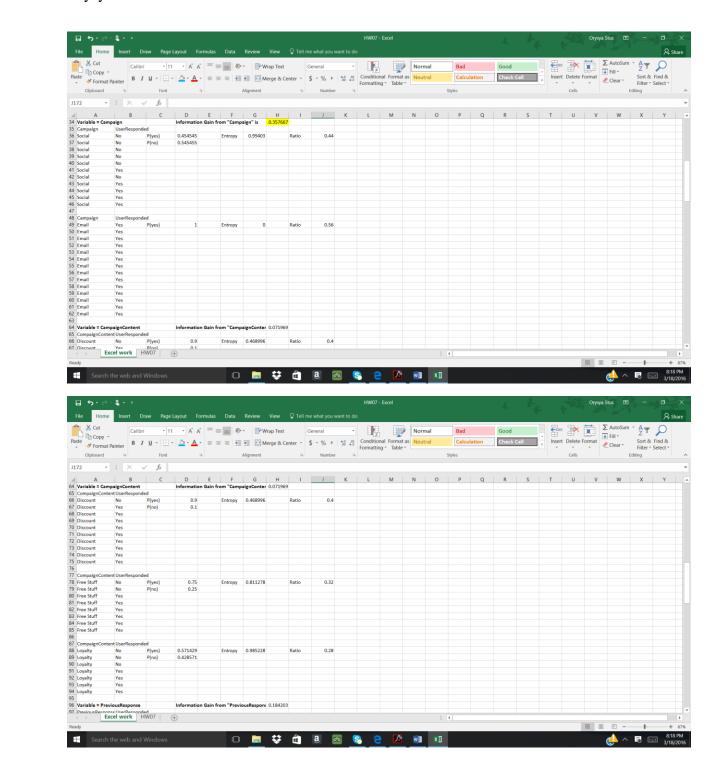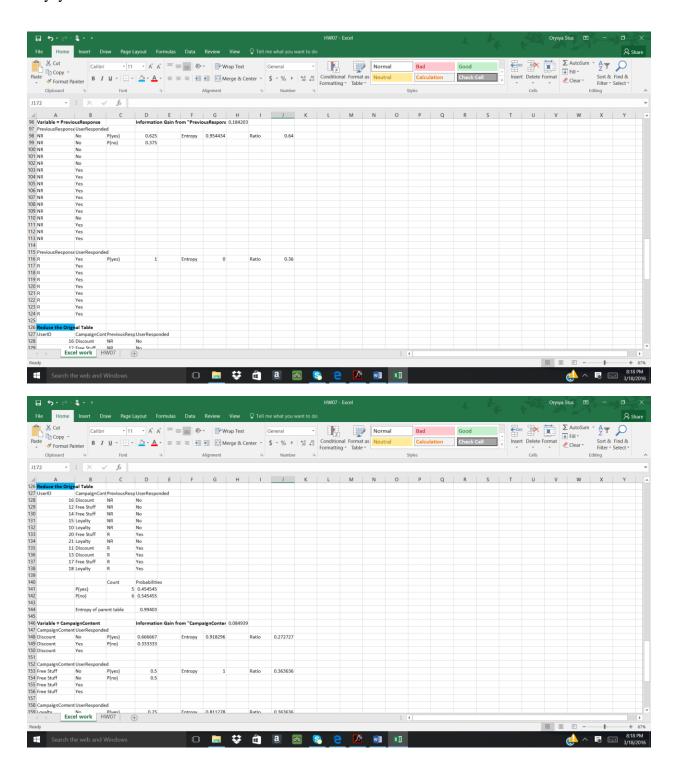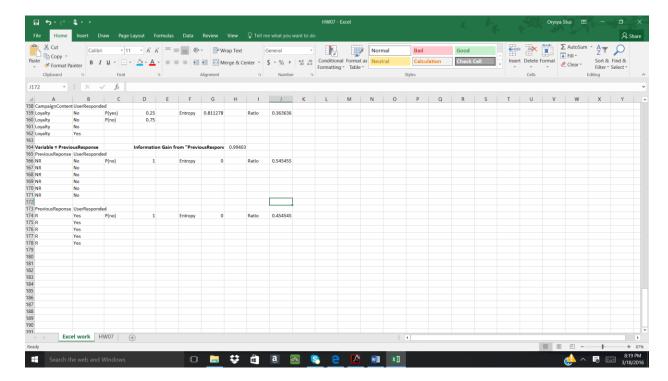Orysya Stus

UC San Diego Extension
Advanced Web Analytics:
Harnessing the Predictive Power
Winter 2016
Homework#7

1. The following was completed in Excel in order to achieve the decision tree:

Orysya Stus

	A	B	C	D	E	F	G	H	I	J
34	Variable = Campaign			Information Gain from "Campaign" is			0.357667		
35	Campaign	UserResponded								
36	Social	No	P(yes)	0.454545		Entropy	0.99403		Ratio	0.44
37	Social	No	P(no)	0.545455						
38	Social	No								
39	Social	No								
40	Social	No								
41	Social	Yes								
42	Social	No								
43	Social	Yes								
44	Social	Yes								
45	Social	Yes								
46	Social	Yes								
47										
48	Campaign	UserResponded								
49	Email	Yes	P(yes)	1		Entropy	0		Ratio	0.56
50	Email	Yes								
51	Email	Yes								
52	Email	Yes								
53	Email	Yes								
54	Email	Yes								
55	Email	Yes								
56	Email	Yes								
57	Email	Yes								
58	Email	Yes								
59	Email	Yes								
60	Email	Yes								
61	Email	Yes								
62	Email	Yes								
63										
64	Variable = CampaignContent			Information Gain from "CampaignConter			0.071969		
65	CompaignContent	UserResponded								
66	Discount	No	P(yes)	0.9		Entropy	0.468996		Ratio	0.4
67	Discount	Yes	P(no)	0.1						

	A	B	C	D	E	F	G	H	I	J
64	Variable = CampaignContent			Information Gain from "CampaignConter			0.071969		
65	CompaignContent	UserResponded								
66	Discount	No	P(yes)	0.9		Entropy	0.468996		Ratio	0.4
67	Discount	Yes	P(no)	0.1						
68	Discount	Yes								
69	Discount	Yes								
70	Discount	Yes								
71	Discount	Yes								
72	Discount	Yes								
73	Discount	Yes								
74	Discount	Yes								
75	Discount	Yes								
76										
77	CompaignContent	UserResponded								
78	Free Stuff	No	P(yes)	0.75		Entropy	0.811278		Ratio	0.32
79	Free Stuff	No	P(no)	0.25						
80	Free Stuff	Yes								
81	Free Stuff	Yes								
82	Free Stuff	Yes								
83	Free Stuff	Yes								
84	Free Stuff	Yes								
85	Free Stuff	Yes								
86										
87	CompaignContent	UserResponded								
88	Loyalty	No	P(yes)	0.571429		Entropy	0.985228		Ratio	0.28
89	Loyalty	No	P(no)	0.428571						
90	Loyalty	No								
91	Loyalty	Yes								
92	Loyalty	Yes								
93	Loyalty	Yes								
94	Loyalty	Yes								
95										
96	Variable = PreviousResponse			Information Gain from "PreviousRespon:			0.184203		
97	PreviousResponse	UserResponded								

Orysya Stus

Screenshot 1

	A	B	C	D	E	F	G	H	I	J
96	Variable = PreviousResponse			Information Gain from "PreviousRespon				0.184203		
97	PreviousResponse	UserResponded								
98	NR	No	P(yes)	0.625		Entropy	0.954434		Ratio	0.64
99	NR	No	P(no)	0.375						
100	NR	No								
101	NR	No								
102	NR	No								
103	NR	Yes								
104	NR	Yes								
105	NR	Yes								
106	NR	Yes								
107	NR	Yes								
108	NR	Yes								
109	NR	Yes								
110	NR	No								
111	NR	Yes								
112	NR	Yes								
113	NR	Yes								
114										
115	PreviousResponse	UserResponded								
116	R	Yes	P(yes)	1		Entropy	0		Ratio	0.36
117	R	Yes								
118	R	Yes								
119	R	Yes								
120	R	Yes								
121	R	Yes								
122	R	Yes								
123	R	Yes								
124	R	Yes								
125										
126	**Reduce the Original Table**									
127	UserID	CampaignCont	PreviousResp	UserResponded						
128	16	Discount	NR	No						
129	17	Free Stuff	NR	No						

Screenshot 2

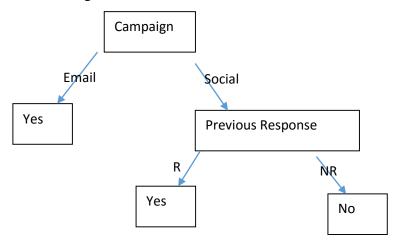|	A	B	C	D	E	F	G	H	I	J
126	**Reduce the Original Table**									
127	UserID	CampaignCont	PreviousResp	UserResponded						
128	16	Discount	NR	No						
129	12	Free Stuff	NR	No						
130	14	Free Stuff	NR	No						
131	15	Loyalty	NR	No						
132	10	Loyalty	NR	No						
133	20	Free Stuff	R	Yes						
134	21	Loyalty	NR	No						
135	11	Discount	R	Yes						
136	13	Discount	R	Yes						
137	17	Free Stuff	R	Yes						
138	18	Loyalty	R	Yes						
139										
140			Count	Probabilities						
141		P(yes)	5	0.454545						
142		P(no)	6	0.545455						
143										
144		Entropy of parent table		0.99403						
145										
146	**Variable = CampaignContent**			Information Gain from "CampaignConter				0.084939		
147	CampaignContent	UserResponded								
148	Discount	No	P(yes)	0.666667		Entropy	0.918296		Ratio	0.272727
149	Discount	Yes	P(no)	0.333333						
150	Discount	Yes								
151										
152	CampaignContent	UserResponded								
153	Free Stuff	No	P(yes)	0.5		Entropy	1		Ratio	0.363636
154	Free Stuff	No	P(no)	0.5						
155	Free Stuff	Yes								
156	Free Stuff	Yes								
157										
158	CampaignContent	UserResponded								
159	Loyalty	No	P(yes)	0.25		Entropy	0.811278		Ratio	0.363636

Orysya Stus

Decision Tree generated:

Prediction considering if: Campaign Medium = Social, Campaign Content = Loyalty, Previous Response = R
-Yes for "UserResponded"

Orysya Stus

Validation in KNIME: