Decision Tree Implementation Report

Team 7 Epitome
Qianwen Li
Bingyu Zhang
Amey Jain

[Link to Bit bucket]
https://bitbucket.org/epitomeTask7/c4.5-decision-tree

[Link to Video]
https://youtu.be/yF6RTzIORf8

[Work Division]
Qianwen Li and Bingyu Zhang worked mainly on building the tree using C4.5 algorithm and cross validation.
Amey Jain worked on Reduced Error Pruning.

[Cross Validation Results]
1. Product Selection:
Average accuracy of Product Selection in 10 fold Cross Validation is 84.5%.

2. Product Introduction:
Average accuracy of Product Introduction in 10 fold Cross Validation is 92.5%.

[Result]
1. Production Selection
Final Model:
Vacation <= 15
| property <= 1.9516: C2
| property >   1.9516
| | salary <= 20.23: C1
| | salary >   20.23: C3
| eCredit <= 26
| | Vacation <= 46: C4
| | Vacation >   46
| | | LifeStyle = spend << saving: C4
| | | LifeStyle = spend >   saving: C5
| | | LifeStyle = spend >> saving: C5
| eCredit > 26
| | property <= 5.7345
| | | salary <= 15.64: C1
| | | salary >   15.64
| | | | Vacation <= 23
| | | | | Type = student: C4
| | | | | Type = engineer: C1
| | | | Vacation >   23
| | | | | eCredit <= 41: C4
| | | | | eCredit >   41
| | | | | | eCredit <= 42: C1
| | | | | | eCredit >   42
| | | | | | | salary <= 20.83: C4
| | | | | | | salary >   20.83
| | | | | | | | salary <= 20.93: C1
| | | | | | | | salary >   20.93
| | | | | | | | | property <= 2.6248: C1
| | | | | | | | | property >   2.6248: C4
| | property >   5.7345: C1

Predicted Result:

Type	Life-Style	Vacation	eCredit	Salary	Property value	Label
Student	spend<saving	12	19	14.79	3.7697	C1
Student	spend>>saving	29	10	16.19	2.4839	C4
Student	spend<<saving	28	60	15.46	1.1885	C1
Engineer	spend>saving	15	41	21.26	1.4379	C2
Librarian	spend<saving	2	9	19.7207	0.6913	C2
Librarian	spend>saving	7	9	12.7098	1.4728	C2
Professor	spend>saving	5	10	20.883	1.3131	C2
Professor	spend<saving	3	15	16.5711	0.4792	C2
Student	spend<saving	9	71	25.7	2.0947	C3
Student	spend>saving	10	67	27.11	3.8391	C3
Doctor	spend>saving	7	229	30.61	7.0074	C3
Doctor	spend<saving	8	243	25.33	8.7276	C3
Professor	spend>saving	51	5	18.98	2.8944	C5
Doctor	spend>saving	34	51	19.9	3.9544	C4
Student	spend>>saving	39	40	19.3	3.8317	C4
Student	spend<<saving	36	57	19.61	4.888	C4
Professor	spend>>saving	34	30	20.91	2.4095	C4
Librarian	spend<<saving	48	35	20.15	2.4436	C4
Professor	spend>>saving	52	5	22.63	2.2115	C5
Engineer	spend>saving	50	17	32.59	1.2229	C5
Engineer	spend>>saving	50	15	21.78	2.0736	C5

2. Production Introduction
   Final Model:
   Advertisement_budget <= 1.49
   | Period <= 50
   | | Monthly_fee <=1.74: 1
   | | Monthly_fee > 1.74
   | | | Customer = student: 0
   | | | Customer = business: 1
   | | | Customer = doctor: 0
   | | | Customer= professional: 1
   | Period > 50
   | | Advertisement_budget <= 0.89: 0
   | | Advertisement_budget > 0.89
   | | Monthly_fee <= 3.73 : 1
   | | Monthly_fee >   3.73 : 0
   Advertisement_budget >   1.49 : 1

Predicted Result:

Service_Type	Customer	Monthly_Fee	Budget	Size	Promotion	Interest_Rate	Period	Label
Fund	Student	0.75	0.93	Small	Web&Email	1	5	1
Fund	Business	1.1	0.93	Small	Web&Email	1	65	1
Loan	Other	2.17	3.07	Small	Full	1	89	1
Mortgage	Business	1.2	1.17	Small	Web	4	10	1
CD	Business	1.2	1.09	Small	Web	0	26	1
Bank_Account	Professional	2.02	0.94	Large	None	3	15	1
Bank_Account	Doctor	4.11	1.07	Large	Web	1	20	0
Bank_Account	Student	4.08	0.98	Large	None	0	15	0
Loan	Business	14.17	4.83	Medium	Web	3	84	1
Loan	Professional	11.12	5.19	Large	Web	4	103	1
Mortgage	Professional	10.68	6.01	Large	Web	2	85	1
Mortgage	Doctor	12.99	5.21	Medium	Web	3	87	1
Mortgage	Business	13.65	3.71	Large	None	1	87	1
CD	Business	5.63	7.15	Medium	Web	1	88	1
CD	Business	6.46	3.69	Large	Web	1	98	1
Fund	Business	2.19	0.43	Small	Web	0	66	0
Fund	Business	2.45	0.45	Small	Web	0	74	0
Fund	Student	2.28	0.79	Small	Full	0	57	0
Fund	Other	1.97	0.82	Small	None	2	67	0
CD	Business	1.71	0.85	Small	Full	0	71	0
CD	Business	2.79	0.76	Small	Full	0	69	0
Loan	Professional	7.97	0.83	Large	Full	2	107	0
Loan	Professional	8.99	1.2	Large	Full	1	104	0
Loan	Business	7.4	1.3	Large	Full	1	115	0
Loan	Other	7.36	1.26	Large	Web&Email	1	98	0
Loan	Student	8.66	0.73	Medium	Web&Email	0	96	0
Mortgage	Other	6.57	0.81	Medium	Web&Email	0	88	0
Mortgage	Student	6.74	0.68	Medium	Web&Email	1	70	0
Mortgage	Other	5.49	1.01	Large	Full	2	80	0
Mortgage	Business	3.9	0.98	Large	Full	0	94	0
Loan	Professional	5.24	3.68	Small	Web&Email	4	44	1
Mortgage	Doctor	3.7	5.94	Small	Web	3	47	1
Fund	Other	4.81	6.35	Small	Web	5	15	1
CD	Doctor	3.78	4.9	Small	Full	4	53	1
Mortgage	Doctor	12.8	5.33	Large	Web	1	92	1
CD	Business	8.29	4.44	Medium	Full	0	86	1
CD	Business	1.87	0.74	Small	None	0	69	0
Loan	Other	8.46	1.14	Large	Web&Email	1	107	0
Mortgage	Professional	4.37	0.62	Large	Full	0	98	0
Loan	Other	4.32	4.97	Small	Full	4	46	1