LEAD SCORE CASE STUDY

BY

BALAJI RAMAN

SHARHEDHA RAGHAVAN

Assumptions

- Columns substituted with common variable 'others' to create meaningful categories

- Many nan columns have been retained with 'no comments ' or ' no data' where meaningful

- Constant value columns like mostly yes or no for almost 90 % or records have been removed

TOP PARAMETERS - FIRST 8

const	-0.7052	0.083	-8.461	0.000	-0.869	-0.542
Total_Time_Spent_on_ Website	0.8774	0.049	17.933	0.000	0.782	0.973
Specialization_Travel and Tourism	-1.1729	0.369	-3.181	0.001	-1.895	-0.450
Lead_Source_Welingak Website	3.0108	0.785	3.834	0.000	1.472	4.550
Lead_Origin_Lead Add Form	1.2264	0.304	4.038	0.000	0.631	1.822
What_is_your_current _occupation_OTHERS	-1.4676	0.109	-13.521	0.000	-1.680	-1.255
Tags_Already a student	-4.4102	0.718	-6.143	0.000	-5.817	-3.003
Tags_Closed by Horizzon	5.6754	1.013	5.600	0.000	3.689	7.662
Tags_Interested in full time MBA	-3.1390	0.735	-4.273	0.000	-4.579	-1.699

TOP PARAMETERS –NEXT 7

Tags_Interested in other courses	-3.0078	0.335	-8.990	0.000	-3.664	-2.352
Tags_Not doing further education	-4.1833	1.014	-4.125	0.000	-6.171	-2.196
Tags_Ringing	-4.2554	0.238	-17.866	0.000	-4.722	-3.789
Tags_Will revert after reading the email	3.3057	0.177	18.658	0.000	2.958	3.653
Tags_switched off	-4.4398	0.528	-8.414	0.000	-5.474	-3.406
Last_Notable_Activity_OTHERS	1.1614	0.603	1.926	0.054	-0.021	2.343
Last_Notable_Activity_SMS Sent	2.1102	0.120	17.548	0.000	1.874	2.346

Vif values to check correlation

	FEATURES	VIF
3	Lead_Origin_Lead Add Form	1.73
11	Tags_Will revert after reading the email	1.54
14	Last_Notable_Activity_SMS Sent	1.46
2	Lead_Source_Welingak Website	1.26
6	Tags_Closed by Horizzon	1.24
0	Total_Time_Spent_on_Website	1.19
10	Tags_Ringing	1.09
4	What_is_your_current_occupation_OTHERS	1.05
12	Tags_switched off	1.03
1	Specialization_Travel and Tourism	1.02
13	Last_Notable_Activity_OTHERS	1.02
5	Tags_Already a student	1.01
8	Tags_Interested in other courses	1.01
7	Tags_Interested in full time MBA	1.00
9	Tags_Not doing further education	1.00

Roc Curve – area 0.97
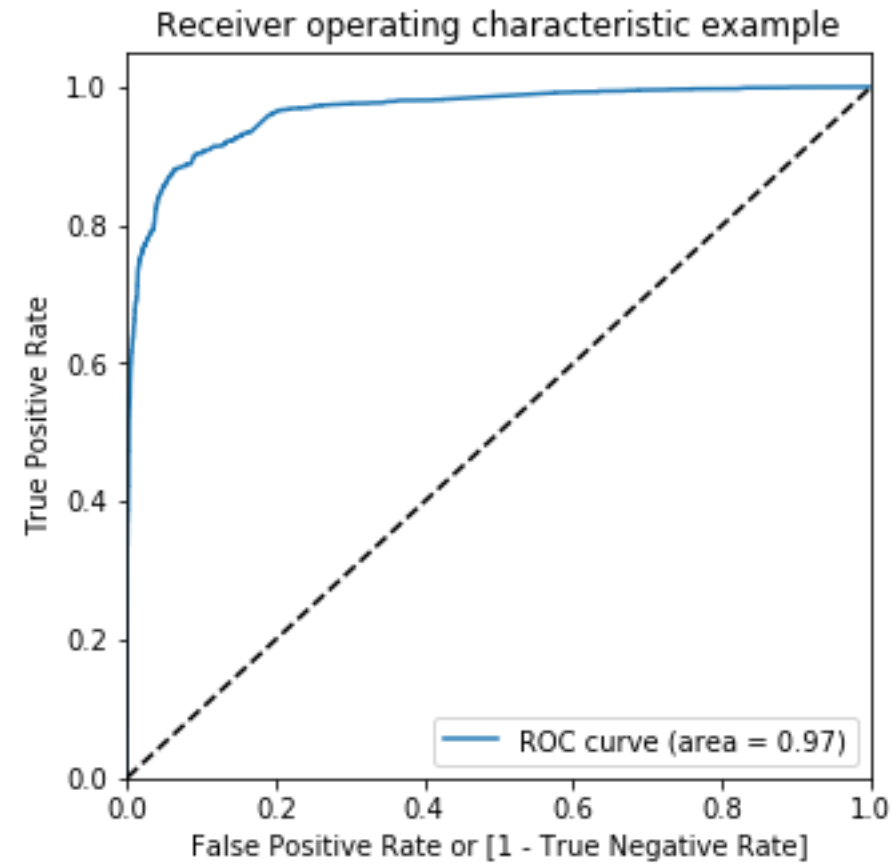
Cut off value determination – 0.3
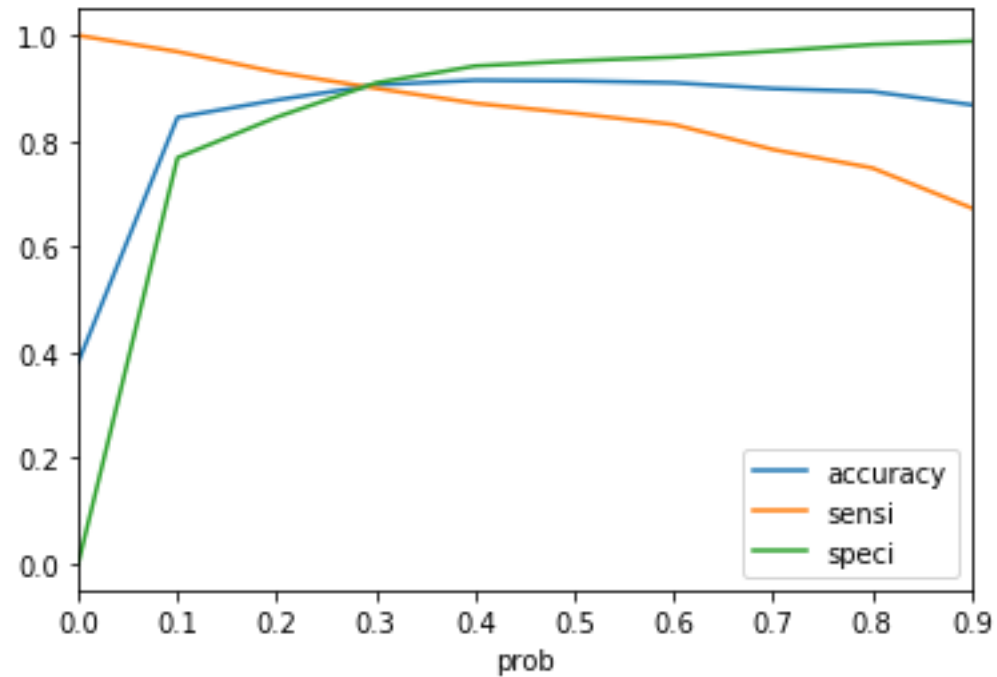
train set sensitivity specificity

- # Let's see the sensitivity of our logistic regression model
- TP / float(TP+FN)
- 0.8998377939983779
- # Let us calculate specificity
- TN / float(TN+FP)
- 0.9100449775112444

Test set – sensitivity and specificity

- # Let's see the sensitivity of our logistic regression model
- TP / float(TP+FN)
- 0.9187214611872146
- # Let us calculate specificity
- TN / float(TN+FP)
- 0.9051878354203936

Accuracy metrics

- train accuracy
- 0.9061533704390847

- Test accuracy
- 0.9105339105339105

CONFUSION MATRIX

- TRAIN
- array([[3642, 360],
-        [247, 2219]], dtype=int64)

- TEST
- array([[1518, 159],
-        [89, 1006]], dtype=int64)

Final Conclusions

- ## Positively correlated
  - **Total_Time_Spent_on_Website = 0.8774**
  - **Lead_Source_Welingak Website   3.0108**
  - **Lead_Origin_Lead Add Form   1.2264**
  - **Last_Notable_Activity_SMS Sent   2.1102**
  - **Last_Notable_Activity_OTHERS   1.1614**
  - **Tags_Will revert after reading the email   3.3057**
  - **Tags_Closed by Horizzon   5.6754**

- ## Negatively Correlated
  - **What_is_your_current_occupation_OTHERS -1.4676**
  - **Tags_Already a student   -4.4102**
  - **Specialization_Travel and Tourism   -1.1729**
  - **Tags_switched off   -4.4398**
  - **Tags_Ringing   -4.2554**
  - **Tags_Not doing further education   -4.1833**
  - **Tags_Interested in other courses   -3.0078**
  - **Tags_Interested in full time MBA   -3.1390**

Recommendations

- **The leads who fall in following category have to be avoided as they don't improve the business.**
  - **The course is <u>not</u> preferred by students , full time mba aspirants**
  - **Travel and tourism people do not benefit from these course offerings**
  - **People interested in other courses don't come back**
  - **Uninterested people don't pick calls or respond**
  - **People with no proper occupation status do not enrol**
- **The following categories result in <u>very high conversion rates</u> . Business should ficus on the following leads.**
  - **leads who respond on SMS and email responses take the course**
  - **People who spend reasonable time on the course webpage should be contacted more if they do not have other negative points affecting their decisions.**
  - **Business should focus on following leads to improve business.**
    - **Lead_Source_Welingak Website**
    - **Tags_Closed by Horizzon**
    - **Tags_Will revert after reading the email.**