Assignment - Acquisition Analytics

SUBMISSION

Submitted By:

1. Anshul Srivastava

<Abstract>

Objective

- To predict the probability of response of each prospect and target the ones most likely to respond to the next telemarketing campaign

Constraints

- Achieve 80% of total responders at the minimum possible cost

- 45,000 data points.

Analysis

- Exploratory Data Analysis
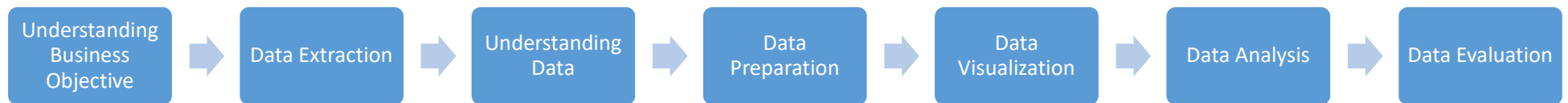
- Cost Analysis

- Building Response Model

Source

- Upgrad.com

<Problem solving methodology>

Use flow chart

| Understanding Business Objective | → | Data Extraction | → | Understanding Data | → | Data Preparation | → | Data Visualization | → | Data Analysis | → | Data Evaluation |

\<Exploratory Data Analysis\>

Objective

- To identify best suited variables for the prospect Identification

Approach

-   By identifying key variables and segmenting them into proper bucktes

Constraints

- Reach out to 80% of total responders in minimum cost.

Results

- the youngest and the eldest are the most likely to buy the investment product.

- student and retired have the highest response rates.

- prospective customers contacted through cellular medium are significantly ore likely to opt in to the investment scheme

- negative employment variation rate seems to be related to higher response rates

<Model Building (Response Model)>

Objective

- To identify best model out of Logistic Regression, Decision Tree and Random Forest for building response model.

Approach

- By looking at metrics like Accuracy, Sensitivity and Specificity of models, finding best model

Results

- **Logistic Regression** turns out to be the best model as it is simple and robust at times.
- With an accuracy of 72% and all other metrics like precision and recall at their best.
- Implementing this model as our response model for the acquisition of prospected customers.

Conclusion

- Even though Random forest was giving much higher accuracy but at the cost of over fitting.

Cost Analysis

decile	total	actual	cumresp	gain	cumlift	cost	cumcost
1	1236	569	569	41.17221	4.117221	6074	6074
2	1236	259	828	59.91317	2.995658	5437.8	11511.8
3	1235	117	945	68.37916	2.279305	5323.9	16835.7
4	1236	92	1037	75.03618	1.875904	5329.25	22164.95
5	1235	66	1103	79.81187	1.596237	5133.233	27298.18
6	1236	68	1171	84.73227	1.412205	4849.917	32148.1
7	1236	75	1246	90.15919	1.287988	5235.2	37383.3
8	1235	64	1310	94.79016	1.184877	4951.417	42334.72
9	1236	39	1349	97.61216	1.08458	5211.283	47546
10	1236	33	1382	100	1	5098.183	52644.18

- From the Dataframe attached above, to reach out to 80% of respondants we only have to contact 50% of the total test population.
- Since the cost of every decile is also predicted, we can predict cost for acquiring customers.
- 80% of respondant = 1103 from 1382 respondants
  50% of test population = 6178 from 12537 test population
  Actual population = 45000 approx.
  Reach out to 50% of Total Population = 22500

  Cost of reaching out to 50 % of test population = ₹ 27,300

  Cost of reaching out to Total population = (Cost of Reaching out to single Customer) * (50% Total Population)
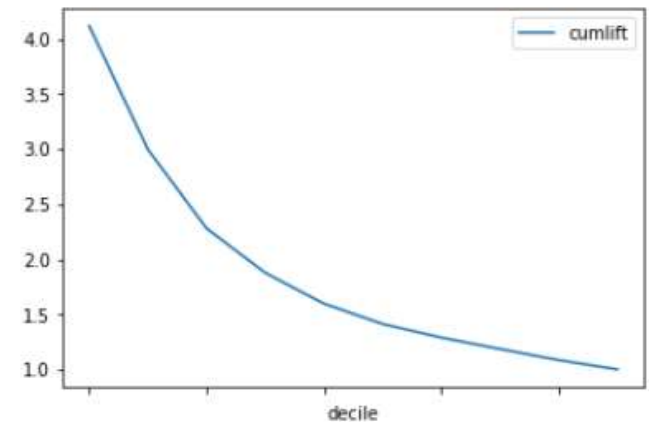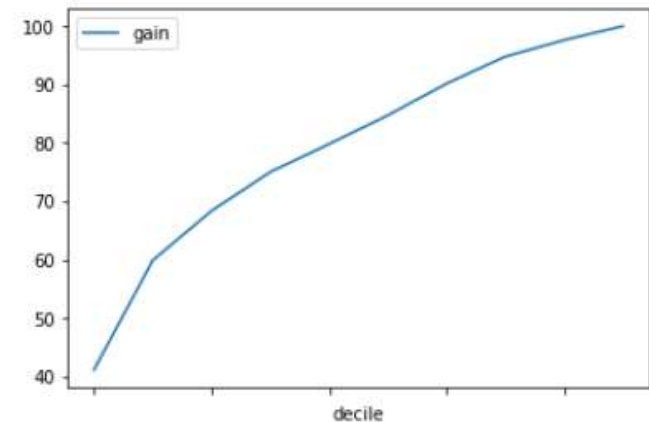
  $$= (27300/6178) * (22500)$$
  $$= ₹ 99,425.0$$
  $$= \sim ₹ 1 \text{ Lac}$$

Lift Chart & Gain Chart

Gain Chart – Only reaching out to 50% of the population we can get 80% of our respondents.

Lift Chart - Even at 50% of the population our model is 1.5 X times better than reaching the entire population.

\<Conclusions\>

With constraints of reaching out to 80% of the total respondents from the entire population along with minimum cost investment. Our recommendations are as follows:

- the youngest and the eldest are the most likely to buy the investment product.

- student and retired have the highest response rates.

- prospective customers contacted through cellular medium are significantly ore likely to opt in to the investment scheme

- negative employment variation rate seems to be related to higher response rates

- 50% of the entire population of 45000 approx. have to be targeted to reach out to 80% of total respondents.

- Average call duration will be around 255 seconds = ~6 minutes

- Cost of Acquisition = 1 Lac approx.