

LEAD SCORE ASSIGNMENT

Group Members:

Soumyadip Bhattacharyya

Shubham Chourasia

Anshul Srivastava

Bhagyashree Nolkha

Abstract

Objective

- Assigning a lead score on the basis of the user behaviour

Analysis Step

- Exploratory Data Analysis & Outliers handling
- Logistic Regression

Source

- UpGrad's Lead Data

Problem Solving Methodology

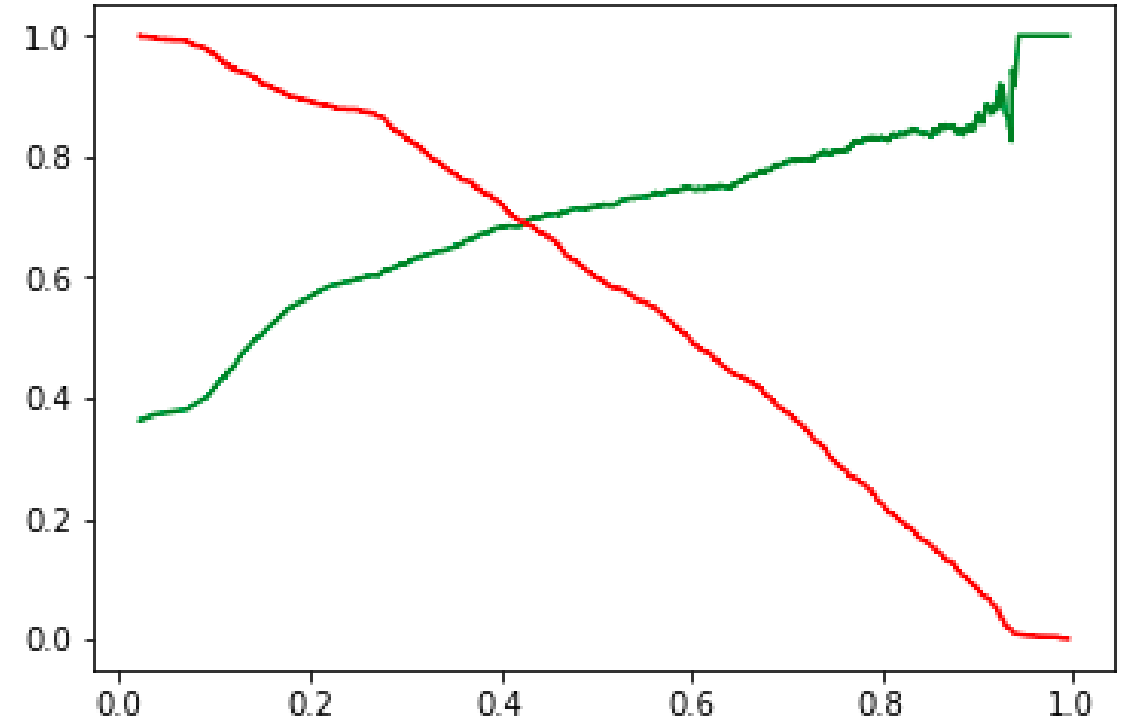
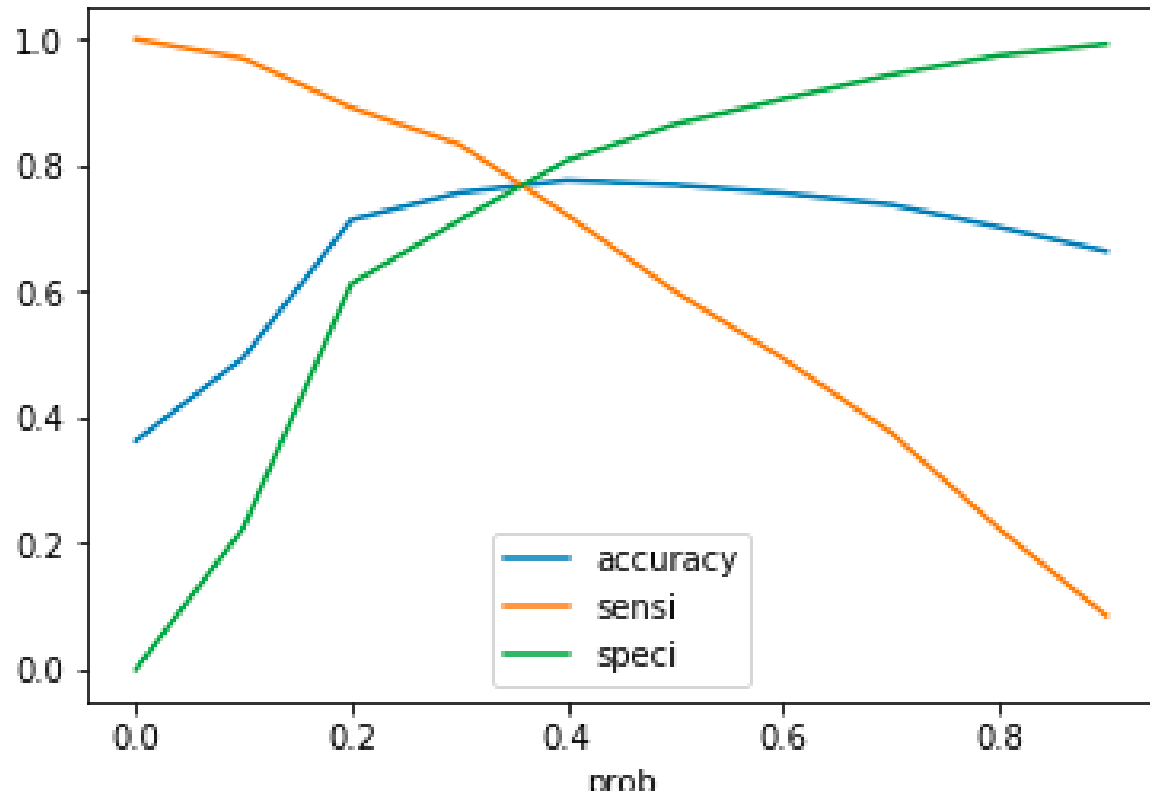


Analysis Steps

Analysis steps taken are as follows :

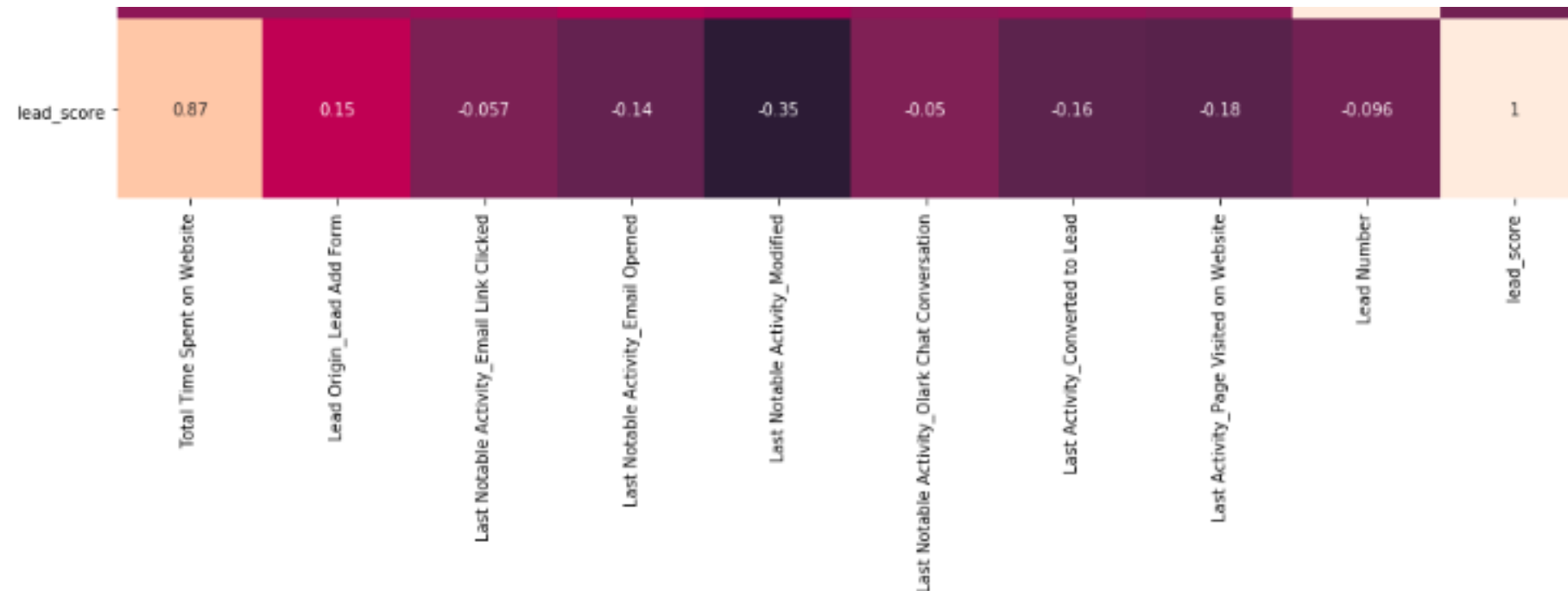
- Importing Dataset and inspecting the dataset
- Exploratory data analysis (**EDA**) to understand data, handle all NA values, outliers etc.
- Creating Dummy variables
- **Correlation matrix** to look for multicollinearity
- Handling Outliers by choosing only data points between **5-95 percentile**
- **Test-Train split** of the dataset
- **Scaling Features** using Standard Scaler method
- **Model Building (Logistic Regression)** using stats models
- **Features Selection using RFE** top **15 features** were selected
- Dropping features with p-values higher than **0.05** & VIFs greater than **5**
- Re-running model until a stable solution is reached
- Plotting ROC curve to choose optimum cut-off for conversion probability
- Calculation of Specificity, sensitivity, accuracy, precision & recall. Optimising on precision to be above 80%
- Final precision attained is **0.81** on test set

Choosing 0.75 probability cut-off in order to get above 80% precision



Correlation Snapshot for Lead Score

- Total Time Spent, Lead Origin & Last Notable Activity are major features responsible for conversion
- **Total Time spent** has a strong positive correlation (**+0.87**) with Leads score
- Amongst Last Notable Activity columns values, **Modified** has a moderate negative correlation with lead score



Recommendation

- Higher Time Spent on website shows higher intent for conversion hence, focus should be on providing as much as possible relevant information on the website & improving User experience
- Optimisation of digital campaigns on users who have opened **Lead Add Form** in order to convert users who haven't filled the lead through remarketing campaigns
- Lead Notable Activity being Modified drives down conversion intent hence, sales team need to work on getting & meeting all requirements in one go

Thank You!