Lead Score – Case Study

Problem Statement :- An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses.

The company markets its courses on several websites and search engines like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos. When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals. Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education is around 30%.

X Education has appointed you to help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company requires you to build a model wherein you need to assign a lead score to each of the leads such that the customers with a higher lead score have a higher conversion chance and the customers with a lower lead score have a lower conversion chance. The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

Steps followed :-

- Data has been analysed by removing un-necessary data along with replacing with existing data.
- Post that dummy variables have been created for the categorical columns and included them in the dataset.
- Further data has been split for the Train-Test datasets.
- Using RFE & Logistic model building, we checked the p values & VIF to check their correlation.
- Model has been built has below mentioned final conclusion has been made.

Parameters	Test Data	Train Data	Delta (Test - Train)
Accuracy	90.92%	90.81%	0.11%
Sensitivity	91.41%	92.05%	-0.64%
Specificity	90.62%	90.10%	0.52%