**Lead Scoring 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%.

Now, although X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted. To make this process more efficient, the company wishes to identify the most potential leads, also known as ‘Hot Leads’. If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.

**Goals of the Case Study**

• To build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads.

• A higher score would mean that the lead is hot, i.e. is most likely to convert whereas a lower score would mean that the lead is cold and will mostly not get converted.

**Algorithm:**

1. Reading the dataset

2. Data Cleaning

3. Exploratory Data Analysis

4. Data Preparation

5. Splitting of Train and Test set

6. Scaling of features

7. Model Building Using Statsmodel and RFE

8. Model Evaluation for Train set

9. Model Evaluation for Test set

10. Results

**Recommendations:**

• High lead conversation rates are possible for "Lead Origin\_Lead Add Form"

• The company should focus on the customers who had a last activity of phone conversation

• Working professional has a higher rate of lead conversation rate

• The company should concentrate on sending SMS to customers to convert them into leads

• Lead sources like "Welingak Website" and "Olark Chat" to be focused

• The company should send emails to convert into leads

• The customers who spent time on the website have a higher lead conversion rate. The company should know how to convert customers into leads when they enter into website

• The customers whose specializations are "Rural and Agribusiness" and "Financial management" are likely to be converted into leads.