Lead Score Case Study

Team Members

Aafia Arham

Member

Kaustav Ghosh

Member

Sumit kumar

Member

Problem statement

- X Education sells online courses to industry professionals.
- 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.

Business Objective:

- X education wants to know most promising leads.
- For that they want to build a Model which identifies the hot leads.
- Deployment of the model for the future use.

EDA

Lead Origin

- Leads are mostly coming from "Landing Page Submission" and "API" but leads comming from them has a lesser coversion rate.
- "Lead Import" has least leads.
- Highest lead conversion are from leads that are sourced from "Lead Add Form".

Recommendation

- Company should focus on improving the conversion rates of Landing Page Submission and API leads while maximising efforts on Lead Add Form leads. At the same time, they should explore whether Lead Import can be improved or is not a viable source.

Lead source

- Most of the leads are sourced from either "Google" or "Direct Traffic"
- There are similar text like google and Google. So binned these 2 data together.
- Conversion rate from "Reference" and "Welingak Website" is highest

Recommendation

- Company should focus on improving the quality of leads from Google & Direct Traffic, standardizing lead source data, and investing more in high-converting sources like Reference and Welingak Website to boost the overall lead conversion rate.

Do not email

- Majority of data is for "No"
- The customers that have "Do not Email" have better conversion rate

Recommendation

- Company should reduce reliance on emails for high-converting "Do Not Email" leads and instead engage them through calls, SMS, and other direct communication channels. At the same time, it should refine its email strategy for the remaining leads to improve engagement and conversion rates.

City

- Most leads are from Mumbai. But highest conversion rate is from "Thane & Outskirts"

Recommendation

- Company should focus marketing efforts on "Thane & Outskirts" to maximize conversions while maintaining presence in Mumbai.

Total Visits

- There are lot of outliers in the data. We have capped the outliers to 95% value for analysis.
- Median of converted and not converted is same and hence nothing can be concluded from TotalVisits

Recommendation

- Company may Focus on other behavioral metrics (e.g., "Total Time Spent on Website") which may have a stronger correlation with conversion.

Total Time Spent on Website

- Leads spending more time in website are more likely for a conversion. Directly proportional.

Recommendation

- Since leads that spend more time on the website are more likely to convert, this feature is a strong predictor of lead conversion.
- Company may use session analytics to identify which pages/content hold user attention and optimize them further for conversions.

Page Views Per Visit

- Outliers are processed
- Median of converted and non-converted are same , no inference can be drawn.

Recommendation

- Since the median values for converted and non-converted leads are the same, this feature does not provide a clear distinction between potential and non-potential leads for the company.

Last Activity

- Customers that have "SMS Sent" as "Last Activity" have better conversion rate.
- Most customer have "Opened Email" as "Last Activity"

Recommendation

- Company can prioritise leads with "SMS Sent" as the last activity, as they have a higher conversion rate, and implement targeted SMS follow-ups for leads who have only opened emails.

Specialisation

- Focus for conversion should be on the Specialisation with high conversion rate.

Recommendation

- Company should prioritise targeting and engagement strategies for specialisations with higher conversion rates to improve overall lead conversion efficiency.

What is your current occupation

- Housewife can be considered as Unemployed.

Recommendation

- Company should classify "Housewife" under "Unemployed" to simplify analysis and improve model accuracy.

These features are not recommended for the company to use for lead conversion since these columns are dropped due to low variance, with only "No" present in the column.

- Search
- Magazine
- Newspaper Article
- X Education Forums
- Newspaper
- Country
- Do not call
- Digital Advertisement
- Through Recommendations
- Receive More Updates About Our Courses
- Tags
- Update me on Supply Chain Content
- Get updates on DM Content
- I agree to pay the amount through cheque

Model Building

- Building model by removing the variable whose p-value is greater than 0.05 and via is greater than 5.
- Over all accuracy 80%

Model Evaluation

- Finding optimal cut off points, where, the optimal cut off is around 0.35

Conclusion

The key factors influencing the company's potential lead conversions are:

- Lead Source: Higher conversion rates observed from leads sourced through "Welingak Website" and "Olark Chat."
- Occupation: "Working Professionals" demonstrate a stronger likelihood of conversion.
- Engagement: Leads who spend more time on the website are more likely to convert.
- Last Activity: "SMS Sent" as the last recorded activity correlates with higher conversion rates.

Leads that are unlikely to convert into customers for the company.

- Last Activity: Leads whose last activity was "Olark Chat Conversation" show lower conversion rates.
- Lead Origin: Leads originating from "Landing Page Submission" have a lower probability of conversion.
- Specialisation: Leads categorised under "Others" show weaker conversion potential.
- Communication Preference: Leads who selected "Do Not Email" as "Yes" are less likely to convert.

Thank you