

LEAD SCORE CASE STUDY

AGENDA

Problem Statement

Solution Methodology

Key Findings

Business Recommendations



PROBLEM STATEMENT



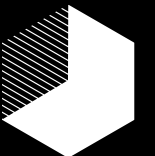
An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses.

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%.

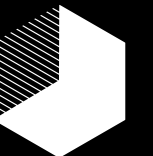
Now, although X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted. To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'. If 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. A typical lead conversion process can be represented using the following funnel:

Lead Conversion Process - Demonstrated as a funnel As you can see, there are a lot of leads generated in the initial stage (top) but only a few of them come out as paying customers from the bottom. In the middle stage, you need to nurture the potential leads well (i.e. educating the leads about the product, constantly communicating etc.) in order to get a higher lead conversion.

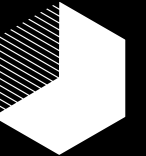
X Education has appointed you to help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company requires you to build a model wherein you need to assign a lead score to each of the leads such that the customers with a higher lead score have 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%.



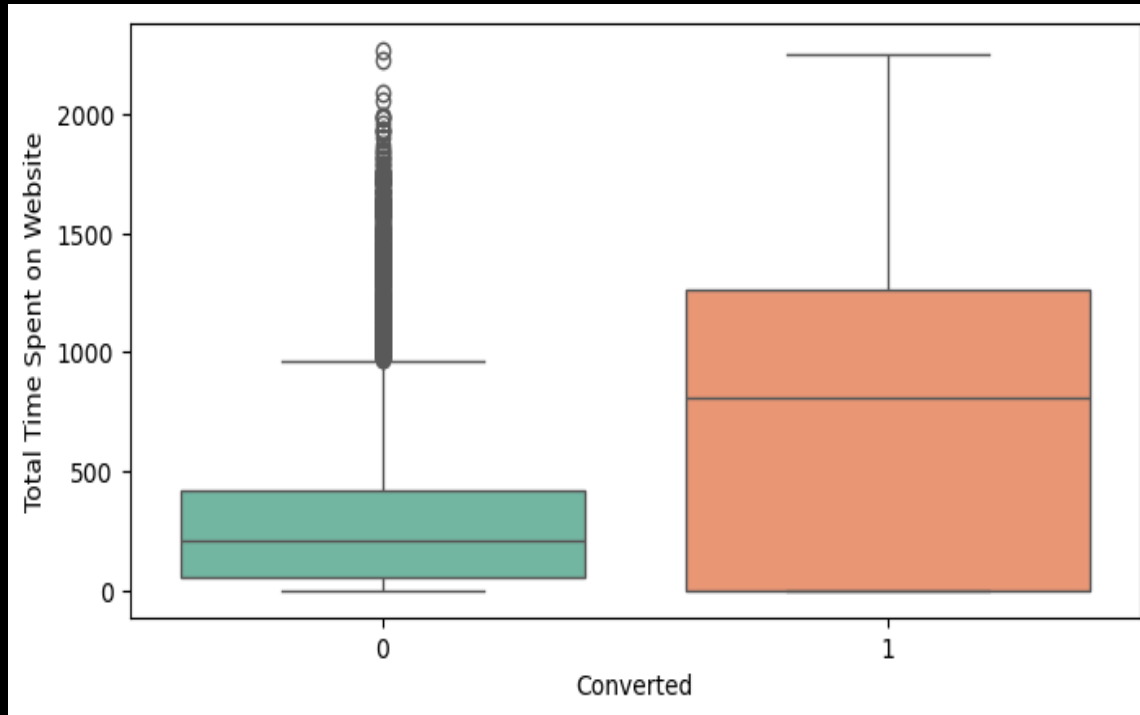
- Data cleaning and missing-value handling
- Binary encoding for yes/no columns
- Dropping extremely sparse columns
- Dummy encoding for multi-level categorical variables
- Splitting into train/test sets
- Using RFE + p-values/VIF for feature selection
- Checking confusion matrix, precision, recall, etc.



KEY FINDINGS

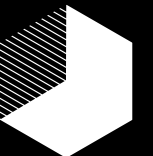


TOTAL TIME SPENT ON WEBSITE VS CONVERSION

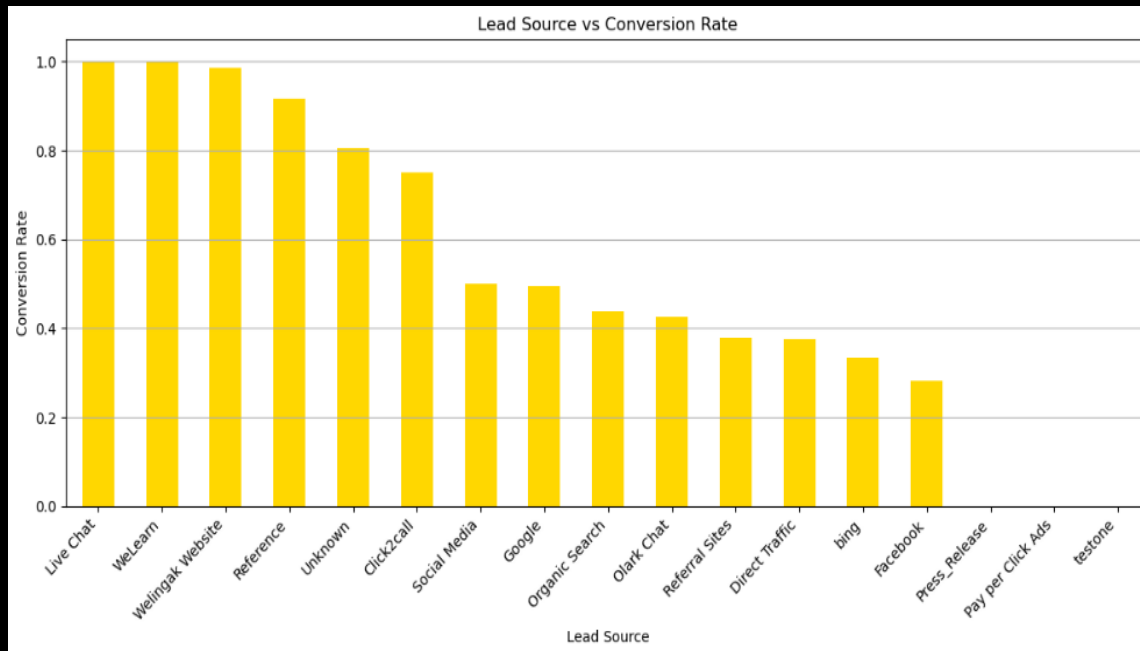


Leads who converted spent significantly more time on the website.

Higher engagement → higher likelihood of conversion.

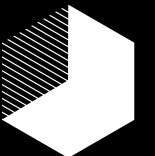


LEAD SOURCE VS CONVERSION RATE

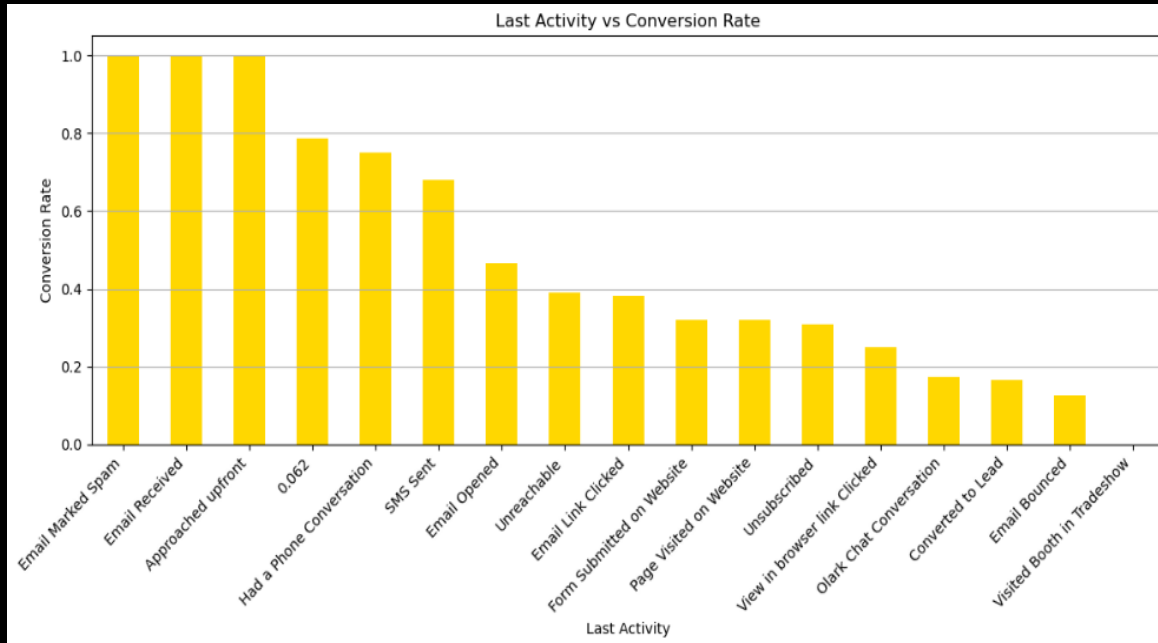


Lead sources like Olark Chat, Google, and Direct Traffic have the highest conversion rates.

Focus marketing spend on top-performing channels.

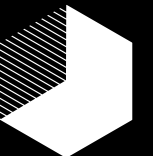


LAST ACTIVITY VS CONVERSION RATE

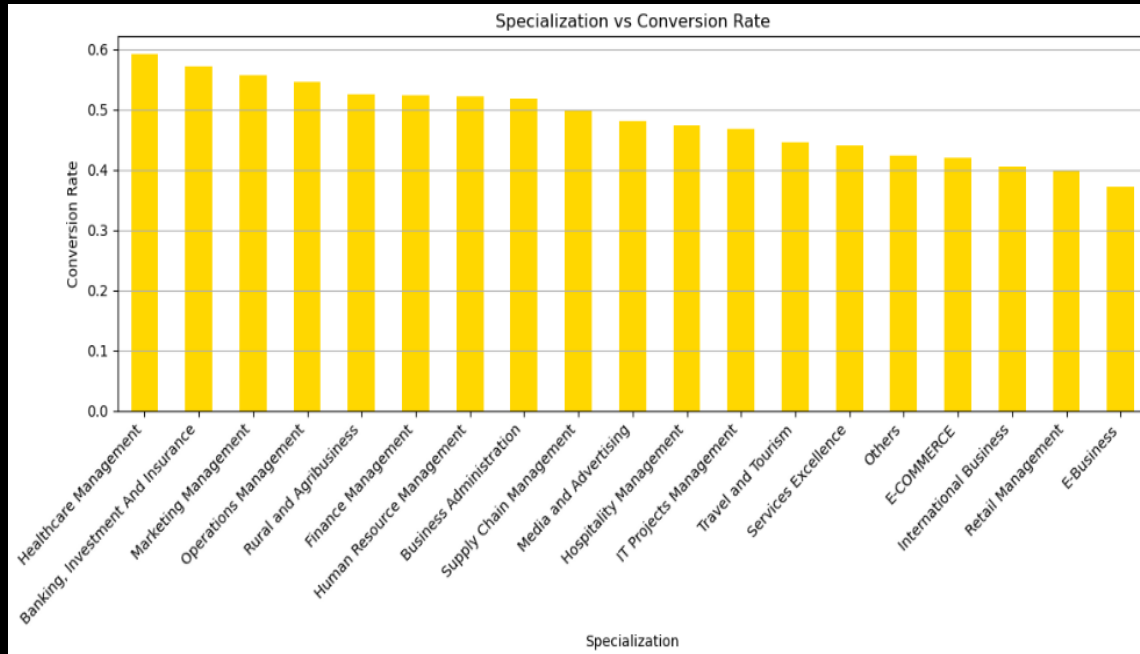


Leads who engaged via Olark Chat, Email Link Clicked, or Email Opened were more likely to convert.

Recent engagement is a strong conversion signal.

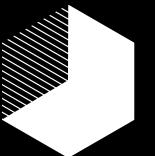


SPECIALIZATION VS CONVERSION RATE



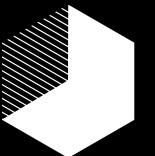
Specializations like Finance, HR, and IT Project Management show higher conversion rates.

Tailor offerings to target high-converting career paths.



BUSINESS RECOMMENDATION

1. FOCUS OUTREACH ON LEADS WITH HIGH ENGAGEMENT (TIME SPENT, REPEAT VISITS).
2. PRIORITIZE LEAD SOURCES LIKE OLARK CHAT, GOOGLE, AND DIRECT TRAFFIC.
3. TRAIN SALES TEAM TO RESPOND QUICKLY TO HIGH-INTENT ACTIONS (E.G., EMAIL OPENED, CHAT CONVERSATIONS).
4. TAILOR MESSAGING BASED ON SPECIALIZATION TO IMPROVE RELEVANCE AND CONVERSION.
5. USE LEAD SCORES TO SEGMENT AND DISTRIBUTE LEADS ACROSS SALES REPS EFFICIENTLY.
6. DURING INTERN-SUPPORTED PERIODS, LOWER PREDICTION THRESHOLD TO MAXIMIZE RECALL.
7. DURING HIGH-EFFICIENCY PHASES, RAISE THRESHOLD TO FOCUS ON HIGH-PRECISION CALLS.
8. LEVERAGE PASSIVE CHANNELS (EMAILS/SMS) FOR MEDIUM AND LOW-POTENTIAL LEADS.
9. MONITOR MODEL PERFORMANCE AND RECALIBRATE THRESHOLDS QUARTERLY.
10. ALLOCATE RESOURCES TOWARD CUSTOMER RETENTION, UPSELLING, AND REFERRAL CAMPAIGNS POST-CONVERSION.





THANK YOU

**ABHISHEK SINGH
ADALENE LYDIA
ADARSH KUMAR**