Optimizing Marketing Lead Scoring:

A Data-Driven Approach

Submitted by:
Sankalp Chaudhary (EPGDM - DS 55 Batch)
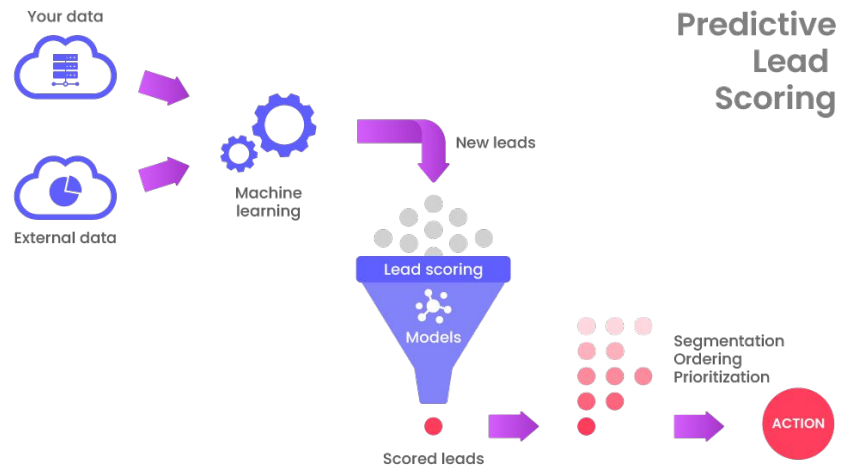Manisha Sanghmitra (EPGDM - DS 55 Batch
Sanjay Mishra (EPGDM - DS 55 Batch)

01

Table Of Content

Table Of Content

Predictive Lead Scoring

Introduction

- **About X Education**

X Education, a leading online education provider, aims to empower industry professionals by offering high-quality online courses. Its mission is to bridge the knowledge gap and enhance skills, catering to professionals seeking continuous learning opportunities. With a user-friendly platform and diverse course offerings, X Education serves as a hub for individuals aspiring to excel in their careers.

- **Importance of Lead Scoring**

Lead scoring plays a pivotal role in revolutionizing X Education's lead management strategy. By employing data-driven techniques, we can effectively prioritize potential customers based on their likelihood to convert. This strategic approach optimizes our sales efforts, allowing us to focus on nurturing promising leads. Lead scoring not only streamlines the communication process but also significantly increases the chances of converting leads into paying customers.
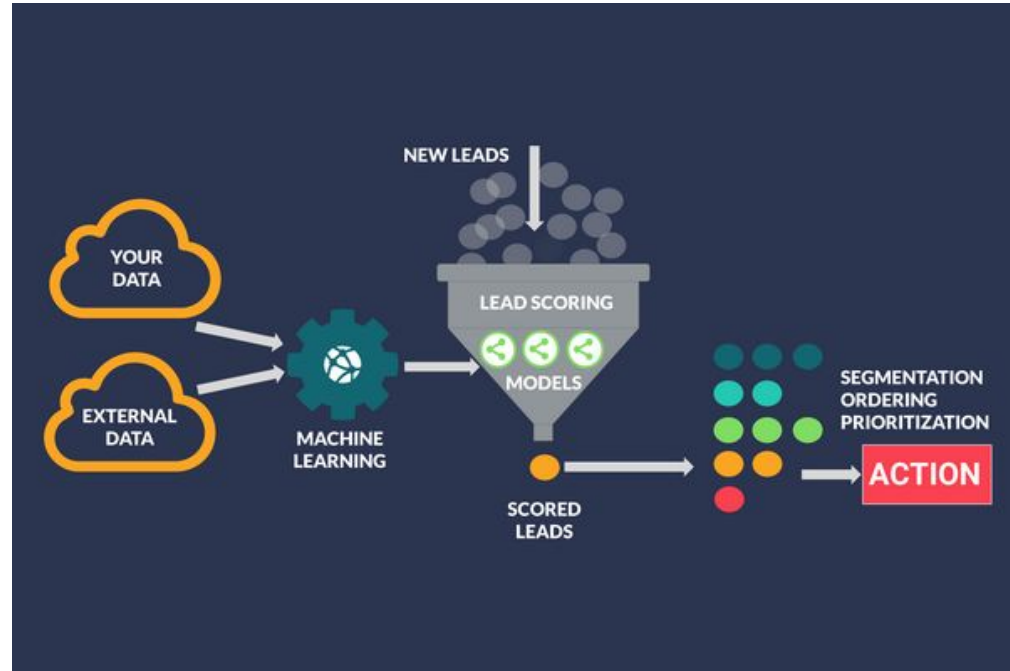
02 Business Understanding

- **Business Objectives:**
  - Increase lead conversion from 30% to 80%.
  - Personalize communication for a seamless customer journey.
  - Utilize data insights for targeted marketing.
- **Challenges:**
  - Managing a high volume of leads efficiently.
  - Enhancing personalization in nurturing strategies.
  - Gaining data-driven insights for tailored marketing.
  - Competing in a dynamic and competitive market.
- **Strategies:**
  - Shift towards personalized, trust-building interactions.
 - Investment in advanced technology and employee training.

03

- **Data Sources**
  - Shape of the dataframe (9240, 37)
- **Data Cleaning Steps:**
  - Dropped all the columns which had null values >3000 as they will not be consequential to the analysis
  - 'Select' as an entry is equivalent to value no entered or missing value so it was treated accordingly. 'Lead profile + 'How did you…'removed
  - Dropped City and Country as they will create multiple non-explicable categories.
- **Feature Selection Factors**
  - **Relevance to Lead Conversion:** The chosen features, such as 'TotalVisits,' 'Total Time Spent on Website,' and 'Page Views Per Visit,' directly reflect user engagement on the website, indicating genuine interest and increasing the likelihood of lead conversion.
  - **Informational Value:** Features like 'Lead Source,' 'Specialization,' and 'Last Notable Activity' provide valuable context about a lead's source, interests, and recent interactions.
  - **Simplicity and Interpretability**
  - **Data availability & Completeness**
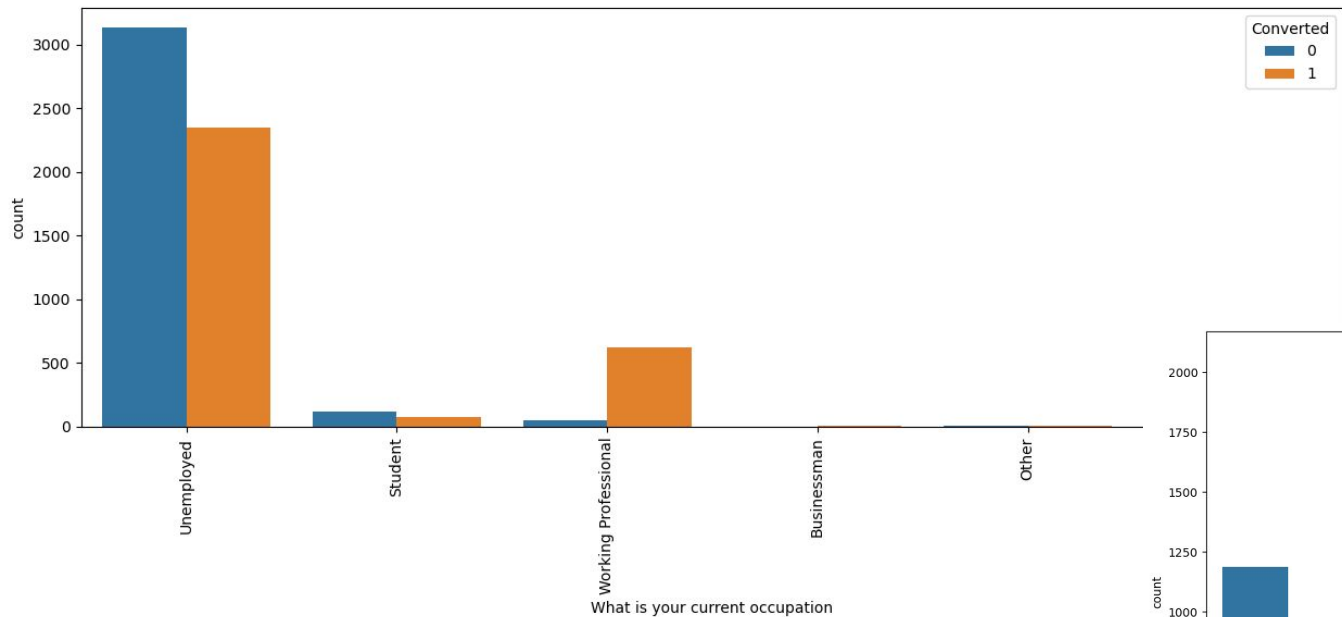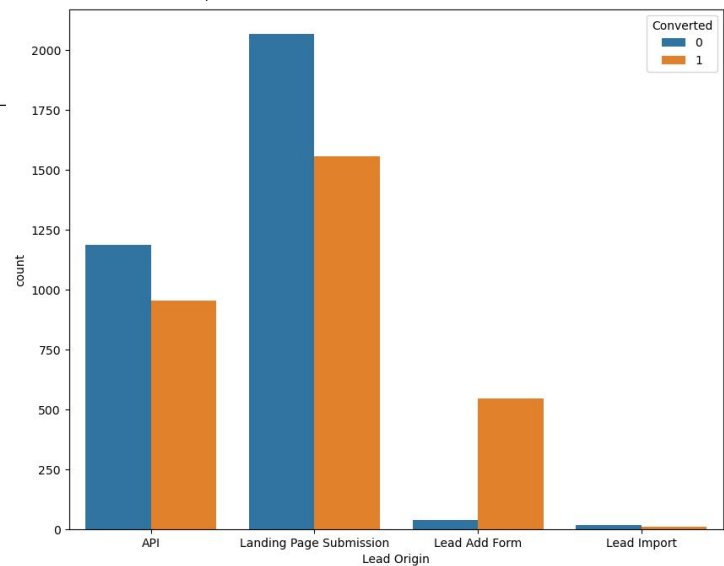
Exploratory Data Analysis
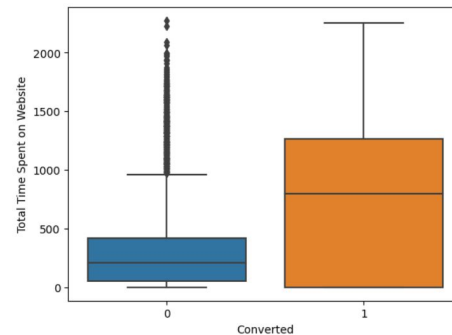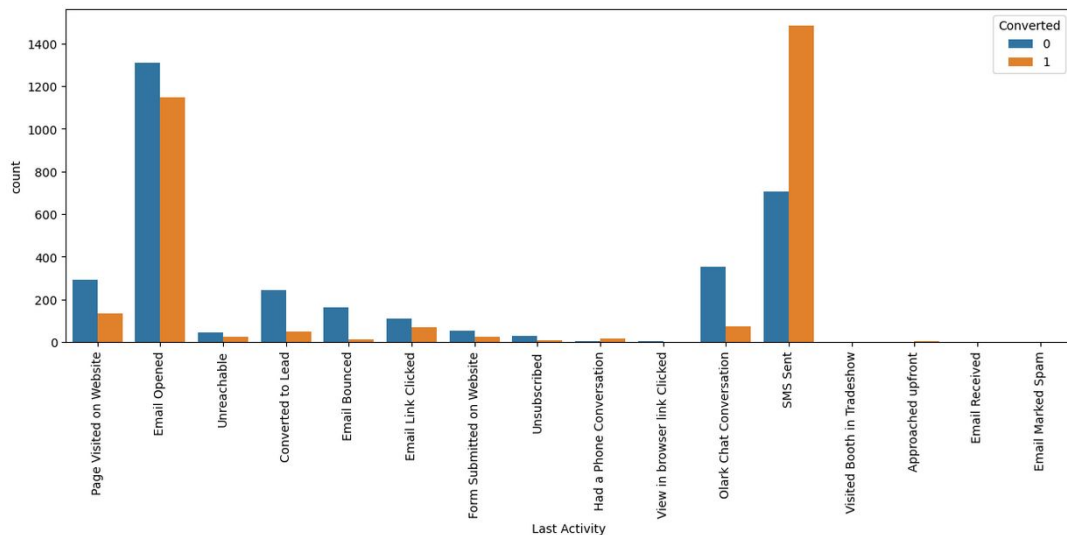
People who convert spend at least 3x time more on the Website

Plotting 'Last Activity' vs converted and non-converted leads. We can see the SMS is one of the key methods to aim for conversion

Current occupation vs Conversion. We can see that working professionals are more likely to convert

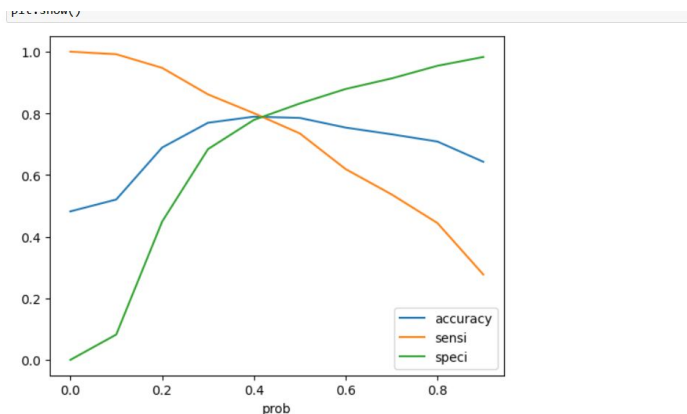Lead Add Form is one of the most successful avenues of leadorgin

Model Development

## Final Model Training Parameters

	Features	VIF
2	Lead Origin_Landing Page Submission	2.15
1	Total Time Spent on Website	1.93
8	Last Activity_SMS Sent	1.55
0	TotalVisits	1.52
3	Lead Origin_Lead Add Form	1.51
5	Lead Source_Welingak Website	1.32
9	What is your current occupation_Working Profes...	1.20
4	Lead Source_Olark Chat	1.06
6	Last Activity_Email Bounced	1.04
7	Last Activity_Had a Phone Conversation	1.01
10	Last Notable Activity_Unreachable	1.01

As you can see that around 0.42, you get the optimal values of the three metrics. So let's choose 0.42 as our cutoff now.

Model Evaluation

- Model Accuracy =0.79
- Model Precision = 0.80
- Model Recall = 0.73
- Specificity = 0.79
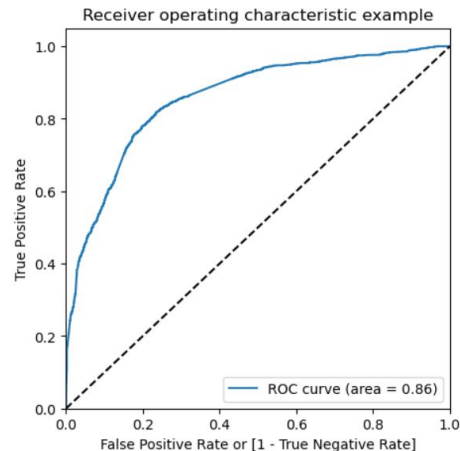
In the context of lead scoring, the model demonstrates promising performance. With an accuracy of 0.79, it correctly identifies leads nearly 80% of the time. A precision of 0.80 means that 80% of the leads predicted as promising are genuinely interested, reducing wasted resources. A recall of 0.73 signifies capturing 73% of all truly interested leads, crucial for not missing potential opportunities. The specificity of 0.79 indicates a high ability to correctly filter out non-promising leads. Balancing precision and recall is vital in lead scoring, ensuring efficient resource allocation and maximizing conversion rates. Further fine-tuning could optimize its efficacy.

Results & Insights

Characteristic Of High Conversion Leads

Total Time Spent on Website (Coefficient = 4.4600): Leads spending more time on the website are approximately 4.46 times more likely to get converted.

Lead Origin_Lead Add Form (Coefficient = 3.9776): Leads generated through the 'Lead Add Form' exhibit a nearly 3.98 times higher chance of conversion. Focus on optimizing interactions over this form

What is your current occupation_Working Professional (Coefficient = 2.6139) Tailoring marketing strategies and content specifically for working professionals

Maximizing Conversion Probability

Lead Source_Olark Chat (Coefficient = 1.2521):. *Focusing on engaging potential leads through chat interactions can boost conversion rates effectively.*

Last Activity_SMS Sent (Coefficient = 1.1030): *Implementing targeted SMS campaigns can enhance lead engagement and conversion opportunities.*

Last Activity_Had a Phone Conversation (Coefficient = 2.7601): *Leads with the last activity as 'Had a Phone Conversation' show a substantial 2.76 times higher chance of conversion. Prioritizing follow-ups with leads engaged in phone conversations can yield successful conversions.s.*

Business Impact: Pickup in Revenue

- *Previous Conversion Rate:* 38.5% -

Model Accuracy: 79% - *Model Precision:* 80%*

Increased Conversion Rate: With the new lead scoring model, the conversion rate is expected to significantly improve due to the higher accuracy and precision, atleast reaching to 50%. (Let's consider a moderate increase)

2. Effect on Revenue:

 Assumptions and Calculations:

- *Average Lead Value:* $Z (estimated revenue generated per converted lead)*

- *Number of Leads:* N (Assumed constant)*

- *Previous Revenue:* $X = N * 0.385 * Z*

- *New Revenue:* $Y = N * 0.45 * Z*

The enhanced lead scoring model not only leads to improved conversion rates but also positively impacts revenue, providing a substantial return on investment. This increased revenue can be reinvested in marketing, sales initiatives, and customer experience enhancements, fostering long-term business sustainability and growth.

Challenges & Limitations

- **Data Limitations**
  - Upon the inspection of the dataset we realized that number of good predictor variables had more than 30% null values and Select. (equivalent to missing values)
  - We had to open the the model to include more dummy variables because initially it was overfitting
- **Model Limitations**
  - The logistic regression model was able to achieve and accuracy of 80%
  - We would need to leverage other techniques like Decision Trees or Random Forest to improve the model as well as to not compromise on interpretability.

09 Next Steps & Recommendations

Model Upgradation & Change Management

Model Refinement

Continuous Feedback Loops

Exploring Advanced Algorithms

Incorporating New Data Sources

+

Continuous Monitoring

Dynamic Lead Scoring

Real-time Monitoring

Predictive analytics tio forecast targets

+

Collaborating With The Sales Team

Regular Feedback Sessions

Training & Collaboration

Shared Goals

09 Next Steps & Recommendations

Strategy for Aggressive Lead Conversion (Intern Hiring Period)

During the intern hiring period, X Education aims for aggressive lead conversion, utilizing the increased workforce efficiently:

1. **Precision-Centric Approach:**
   - Prioritize leads with the highest conversion probabilities based on the model's predictions.
   - Focus phone calls exclusively on these high-probability leads, ensuring a precision-oriented approach to maximize conversion rates.
2. **Comprehensive Follow-Up:**
   - Engage in thorough follow-up conversations with potential leads, addressing their queries and concerns promptly.
   - Utilize interns' availability to make comprehensive phone calls, ensuring no potential lead is left unattended.
3. **Personalized Communication:**
   - Tailor communication based on lead interests and interactions, making the conversation highly personalized.
   - Leverage interns' time to gather additional insights during phone calls, enabling customized offerings and solutions.
4. **Data-Driven Decision Making:**
   - Utilize real-time data analytics to identify patterns in successful conversions.
   - Adapt the communication strategy based on data-driven insights, optimizing the conversion process continuously.

Next Steps & Recommendations

Resource Optimization (Early Target Achievement)

- **Automated Engagement Sequences:**
  a. Implement automated email and SMS sequences tailored to lead preferences.
  b. Utilize intelligent automation tools to nurture leads without direct phone calls, minimizing intrusion while maintaining engagement.
- **Content-Driven Engagement:**
  a. Create engaging and informative content such as newsletters, webinars, or educational resources.
  b. Disseminate this content through non-intrusive channels, allowing leads to engage voluntarily without active solicitation.
- **Social Media and Online Platforms:**
  a. Leverage social media platforms and online forums to engage with leads passively.
  b. Share valuable content and participate in discussions, encouraging leads to express interest organically.
- **Selective and Purposeful Calls:**
  a. Reserve phone calls for critical interactions, such as closing deals or addressing specific lead inquiries.
  b. Ensure each call is purposeful and addresses a significant need, avoiding routine or generic conversations.
- **Continuous Monitoring and Optimization:**
  a. Monitor lead engagement metrics closely, identifying channels with the highest response rates.
  b. Continuously optimize the engagement strategy based on the effectiveness of different communication channels, focusing on the ones yielding the best results.

FIN.

Reachout to us

sankalp16@micamail.in
sanghamitra.manisha@gmail.com
sanjaymishra.sm2734@gmail.com