

Lead Conversion Analysis and Recommendations

Problem Statement

Lead Conversion Analysis and Recommendations

- The goal is to predict the likelihood of lead conversion using a logistic regression model and assign a lead score between 0 and 100 to each lead.
- The analysis aims to identify key factors influencing lead conversion and provide actionable recommendations to improve conversion rates.
- This will help the sales team prioritize leads and allocate resources more effectively.

Data Preprocessing

Lead Conversion Analysis and Recommendations

- Created dataframe to create a comprehensive dataset.
- Dropped irrelevant columns such as 'Prospect ID', 'Lead Number', 'Lead Origin', 'Lead Source', 'Do Not Email', 'Do Not Call', 'Last Activity', 'Country', 'City', and 'Last Notable Activity' to focus on relevant features.
- Handled missing values by filling them with a placeholder value 'Unknown' to ensure the model can handle missing data without errors.
- Converted categorical variables into dummy/indicator variables to enable the logistic regression model to process categorical data effectively.

Model Building

Lead Conversion Analysis and Recommendations

- Defined feature variables (X) and target variable (y). The target variable 'Converted' indicates whether a lead was converted (1) or not (0).
- Split the dataset into training (80%) and testing (20%) sets to evaluate the model's performance on unseen data.
- Trained a logistic regression model on the training set. Logistic regression is chosen for its simplicity and effectiveness in binary classification tasks.

Model Evaluation

Lead Conversion Analysis and Recommendations

- Made predictions on the test set to evaluate the model's performance.
- Calculated the accuracy of the model, which is a common metric to evaluate classification models. The model achieved an accuracy of approximately 94% on the test set.
- Printed the classification report, which provides detailed metrics such as precision, recall, and F1-score. These metrics help understand the model's performance in terms of correctly identifying converted and non-converted leads.

Lead Scoring

Lead Conversion Analysis and Recommendations

- Predicted probabilities for each lead in the dataset. The probabilities indicate the likelihood of each lead converting.
- Added lead scores to the original dataframe. The lead scores are calculated by multiplying the predicted probabilities by 100.
- Saved the dataframe with lead scores to a new CSV file. This allows the sales team to access and use the lead scores for prioritizing leads.

Key Variables Influencing Conversion

Lead Conversion Analysis and Recommendations

- Top Three Variables:
 - Total Time Spent on Website: Leads who spend more time on the website are more likely to convert as they show higher engagement and interest.
 - Page Views Per Visit: Leads who view more pages per visit are more engaged and have a higher probability of conversion.
 - Total Visits: Leads with a higher number of visits are more likely to convert as they show consistent interest.
- Top Three Categorical/Dummy Variables:
 - Lead Source_Reference: Leads that come through references have a higher probability of conversion. Focusing on referral programs can help increase conversions.
 - What is your current occupation_Working Professional: Leads who are working professionals are more likely to convert. Tailoring marketing strategies to target working professionals can improve conversion rates.
 - Specialization_Human Resource Management: Leads with a specialization in Human Resource Management have a higher probability of conversion. Creating targeted content and offers for this specialization can enhance conversions.

Recommendations for Aggressive Lead Conversion

Lead Conversion Analysis and Recommendations

- **Prioritize High-Scoring Leads:** Use the lead scores generated by the model to prioritize leads with the highest scores. These leads have the highest probability of conversion and should be contacted first.
- **Allocate Interns to High-Scoring Leads:** Assign the interns to focus on calling and following up with the high-scoring leads. This ensures that the most promising leads receive immediate attention.
- **Personalized Communication:** Train interns to use personalized communication strategies when contacting leads. Addressing leads by their name and referring to their specific interests can increase the chances of conversion.
- **Follow-Up Schedule:** Implement a structured follow-up schedule to ensure that leads are contacted multiple times if necessary. Persistence can lead to higher conversion rates.

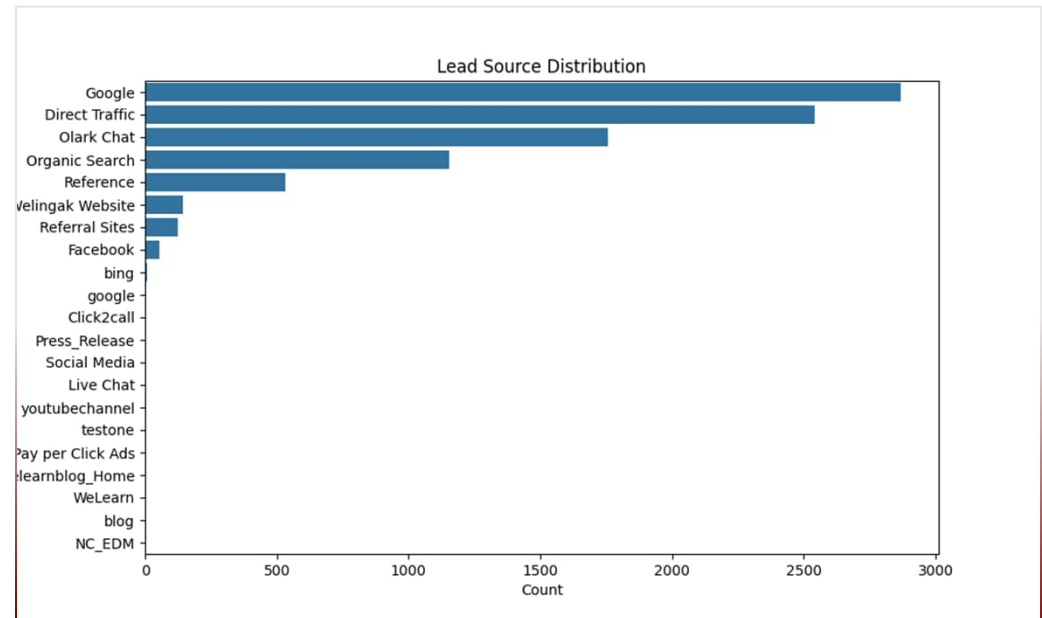
Recommendations for Minimizing Useless Phone Calls

Lead Conversion Analysis and Recommendations

- **Focus on Low-Scoring Leads:** Use the lead scores to identify leads with lower probabilities of conversion. These leads can be deprioritized for phone calls during this period.
- **Automated Communication:** Implement automated email or SMS campaigns to engage with low-scoring leads. This allows the sales team to maintain contact without making phone calls.
- **Content Marketing:** Develop and share valuable content (e.g., blog posts, webinars, e-books) that can nurture low-scoring leads over time. This can help keep leads engaged without direct phone calls.
- **Data Analysis:** Use this period to analyze data and identify trends or patterns in lead behavior. This can provide insights for future marketing strategies and improve overall lead conversion processes.

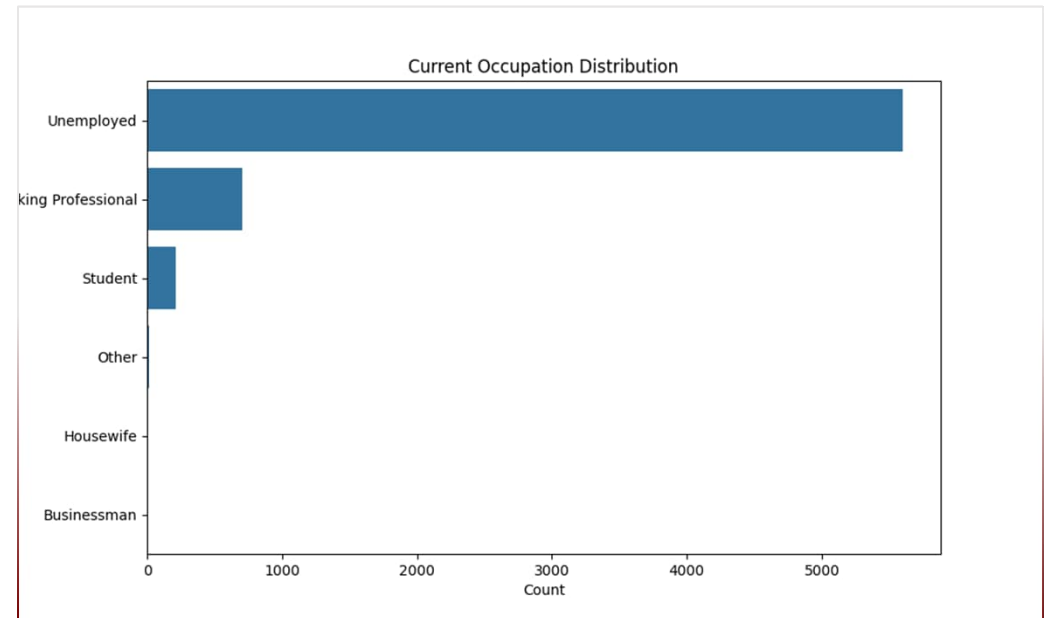
- The distribution of lead sources helps understand which sources are most effective in generating leads.

Lead Source Distribution



- The distribution of current occupations helps identify which occupations are more likely to convert.

Current Occupation Distribution



- The distribution of specializations helps identify which specializations are more likely to convert.

Specialization Distribution

