

CUSTOMER INSURANCE CLAIM PREDICTION PROJECT

I. Introduction

Stakeholders:

Primary Stakeholder: Insurance Company

Problem: Efficiently assess the risk associated with each customer to optimize insurance pricing and reduce claim payouts.

Secondary Stakeholders: Marketing Team, Underwriting Team

Problem: Marketing team wants to target specific demographics effectively. Underwriting team needs to assess risk accurately.

II. Data Overview

Data Source:

The dataset was obtained from Kaggle and includes comprehensive customer information.

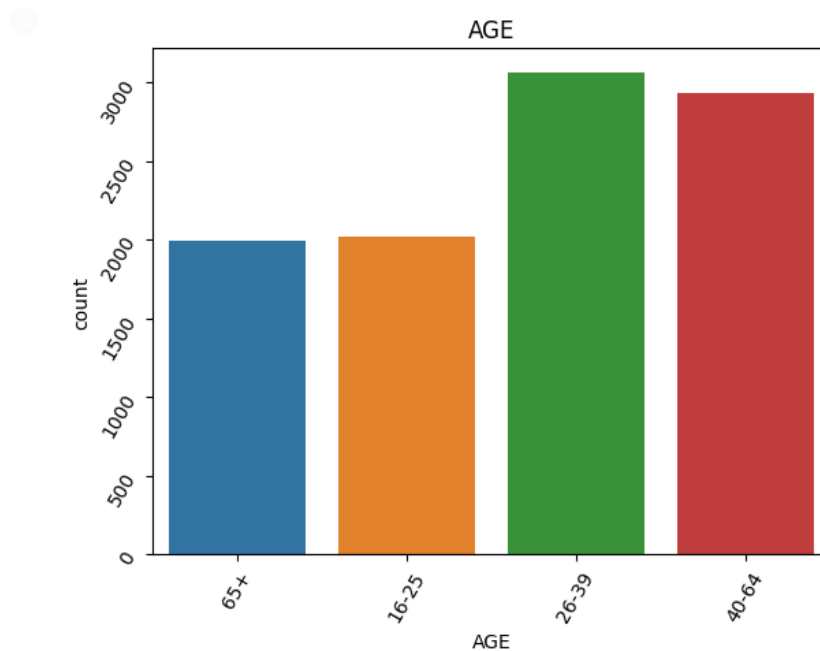
Data Attributes:

The dataset comprises various customer attributes such as gender, race, credit score, vehicle ownership, and more.

We have a diverse set of attributes to explore for predicting claim likelihood.

III. Key-Finding 1 - Age Distribution

Visualization 1



Finding:

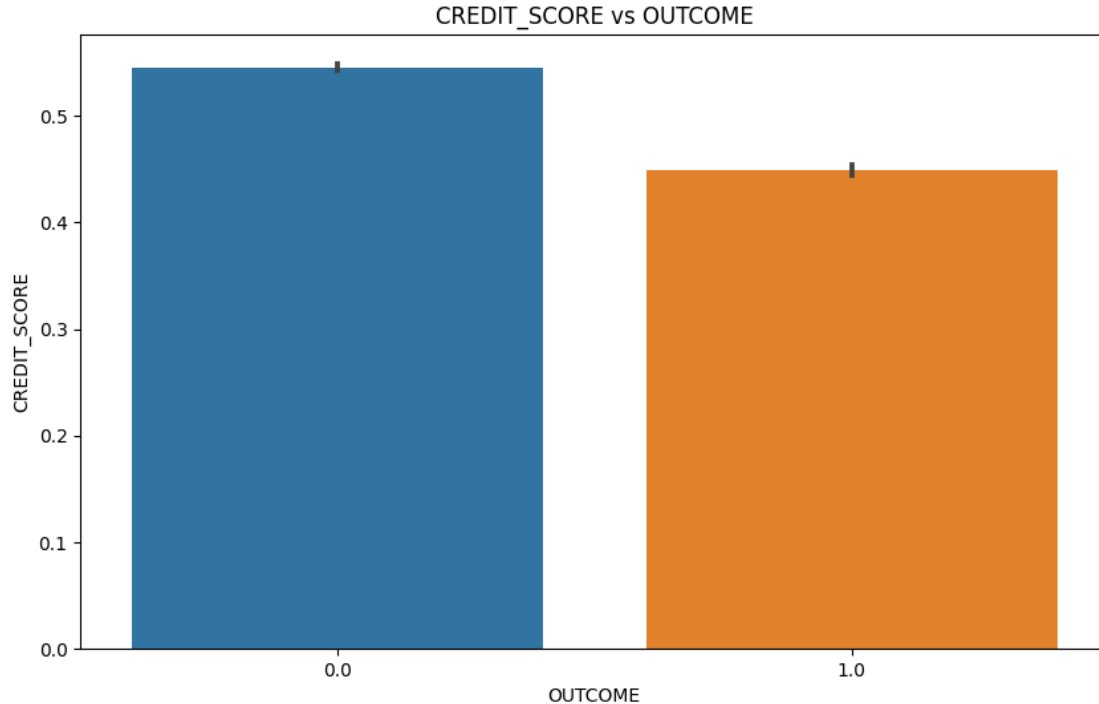
The visualization displays the distribution of customer ages, revealing that the majority of customers fall into the '26-39' and '40-64' age groups.

Implication for Stakeholder (Insurance Company):

By tailoring insurance offers to cater specifically to these age demographics, the company can increase customer engagement and potentially reduce claim frequency.

IV. Key-Finding 2 - Credit Score vs. outcome

Visualization 2



Finding:

The chart illustrates the correlation between credit scores and the likelihood of filing an insurance claim, showing that customers with lower credit scores are more likely to file claims.

Implication for Stakeholder (Underwriting Team):

The underwriting team should consider credit score as a significant risk factor when determining insurance premiums and coverage. This can lead to more accurate risk assessments and optimized pricing.

v. Model Strengths and Limitations

Strengths:

The predictive model exhibits strong performance with an accuracy of 87%, high precision (90%), recall (91%), and an F1-Score of 91%.

The ROC-AUC score of 85% indicates that the model effectively separates positive and negative cases.

Limitations:

The model's performance is limited by the quality and completeness of the data. Incomplete or inaccurate data may impact predictions.

The model may not capture all relevant risk factors, and there could be external factors affecting claim likelihood.

VI. Model Interpretation (Classification Problem)

Classification Metric: False Positive (FP) and False Negative (FN) Rates.

False Positives:

Occur when the model predicts a claim, but the customer doesn't actually file one. This could lead to offering insurance to customers who are less likely to claim, impacting profitability.

False Negatives:

Happen when the model predicts no claim, but the customer ends up filing one. This could result in denying insurance to customers who could have been profitable, affecting customer satisfaction.

Class Balance: It's essential to consider class balance in the evaluation to minimize the impact of false positives and false negatives on the stakeholder.

VII. Final Recommendations

Recommendations for Insurance Company:

Targeted Marketing:

Tailor marketing campaigns for '26-39' and '40-64' age groups to increase engagement and conversions.

Underwriting Guidelines:

Incorporate credit scores as a key risk factor when determining insurance premiums and coverage.

Claim Prevention:

Engage high-risk customers with preventive measures or resources to reduce claims frequency.

Continuous Monitoring:

Regularly update and retrain the predictive model with new data to ensure its accuracy and relevance in assessing customer risk.

Slide 8: Conclusion and Questions

Conclusion:

Implementing these recommendations will enhance risk management, refine customer targeting, and improve overall business performance for the Insurance Company.

Questions?

Feel free to customize the content, add more details, or adjust the presentation to better suit your project's specifics and the needs of your stakeholders. This presentation aims to provide a comprehensive overview of your project and its impact on the stakeholders.