ANALYTIC AVENGERS

CUSTOMER LIFETIME VALUE

CONNECTING ALL OF US WITH THE WORLD

ABOUT US

Analytic Avengers is a dynamic company consisting of six experienced data scientists dedicated to revolutionizing business strategies. Specializing in market segmentation and precise ad targeting, Analytic Avengers employs advanced data analytics techniques to unlock invaluable insights for businesses. Leveraging their expertise in Customer Lifetime Value (CLV) predictions, the team empowers clients to optimize marketing efforts, tailor product offerings, and enhance customer experiences. By deciphering CLV, Analytic Avengers enables businesses to identify high-value customers, predict future revenue streams, and implement targeted retention strategies. With a keen focus on driving growth and maximizing ROI, Analytic Avengers is poised to propel businesses towards sustainable success in the competitive landscape of today's market.

TABLE OF CONTENTS

THE TEAM

Brenda Ngigi

Winnie Nzuve

Brian Gikonyo

Vincent Kiplang’at

Beatrice Adhiambo

Cynthia Gitonga

BUSINESS UNDERSTANDING

- CLV prediction is critical for subscription-based businesses.
- Accurately estimate long-term customer value for strategic decision-making.

PROBLEM STATEMENT

Objective: Develop a Predictive Model for Customer Lifetime Value (CLV):

- Critical for subscription-based businesses.
- Aims to provide key benefits:

a) Resource Allocation
b) Marketing Strategy Optimization
c) Customer Retention Programs

DATA

DATA FROM KAGGLE

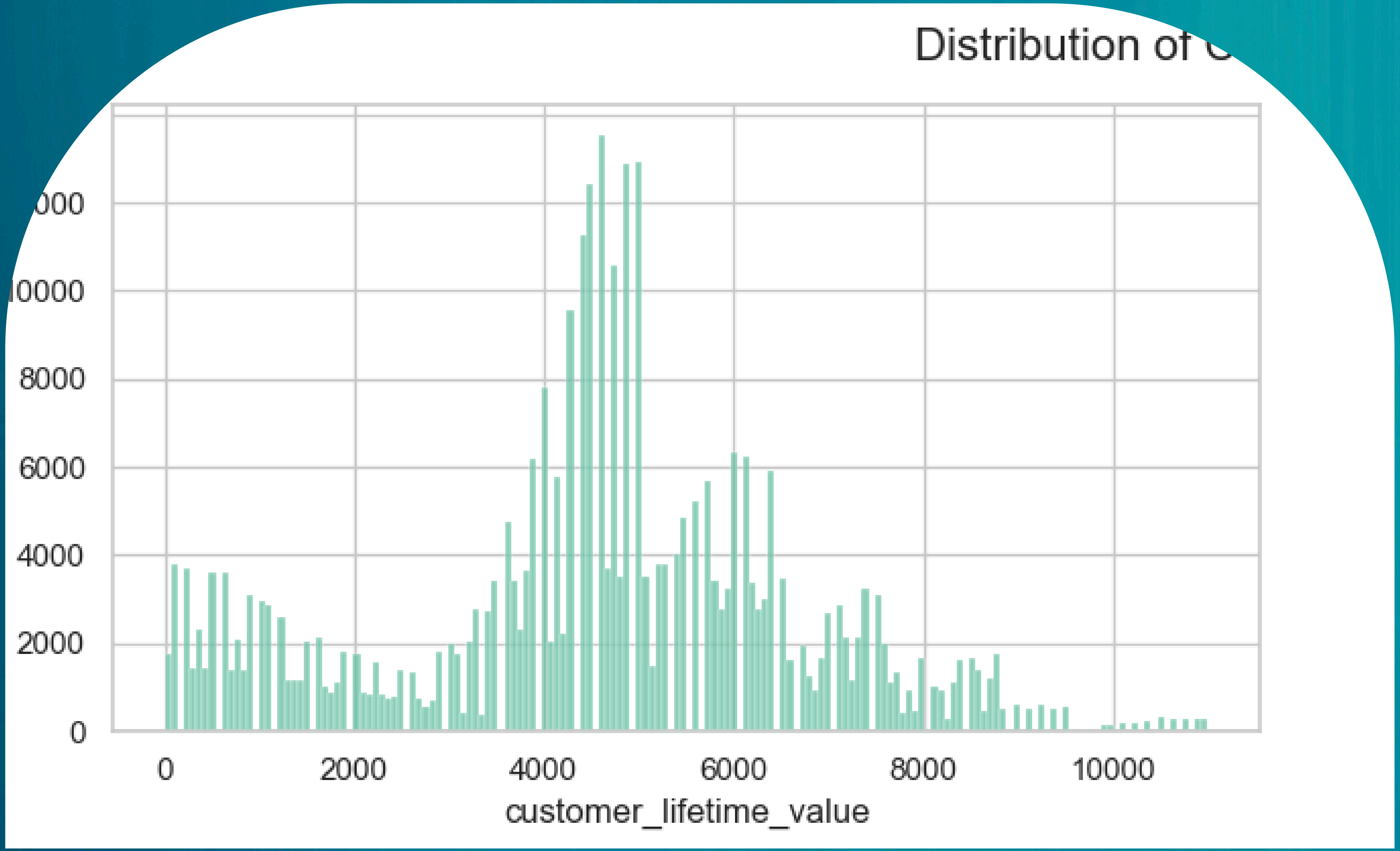
customer cases

customer info

customer id

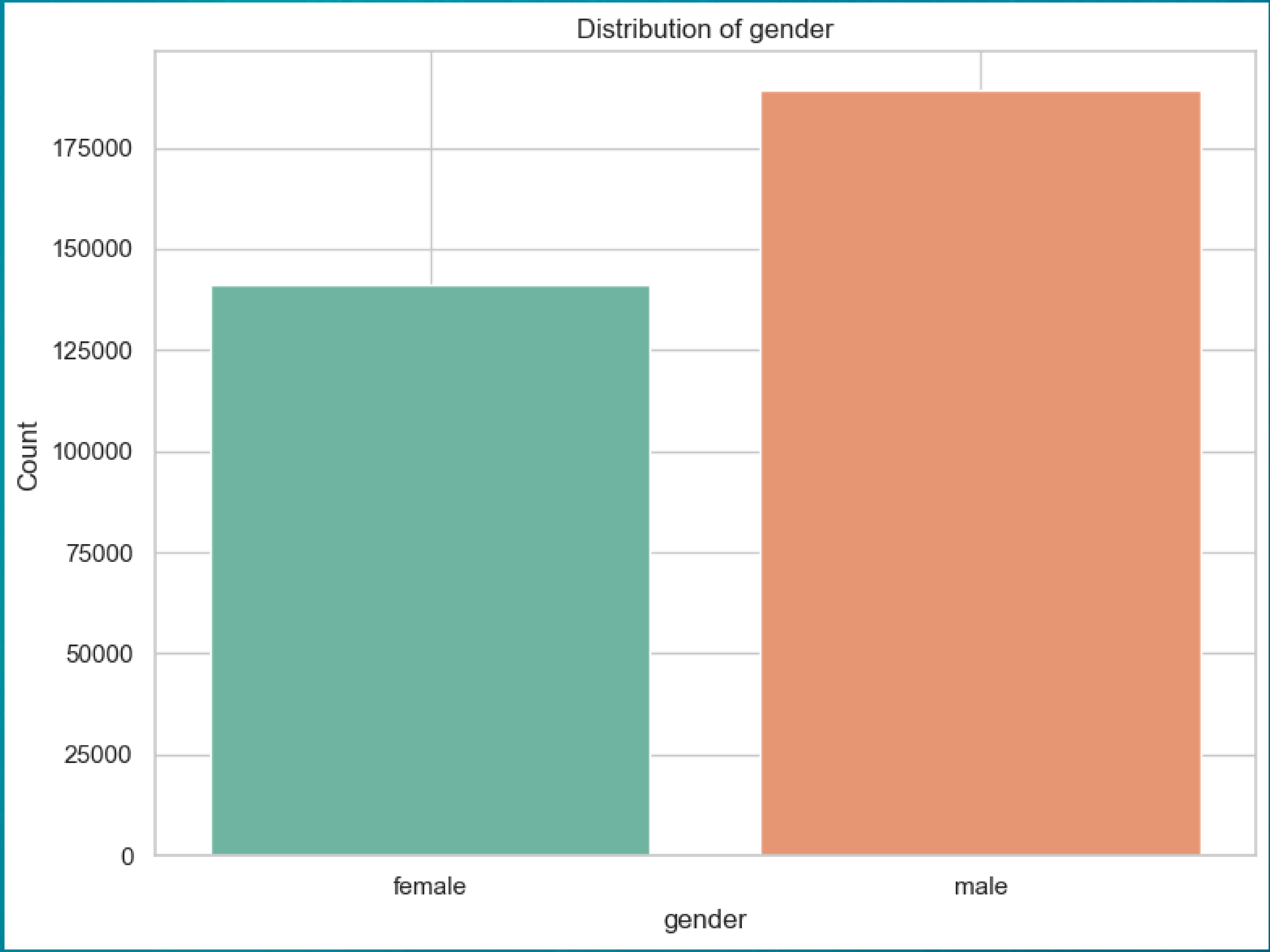
product info

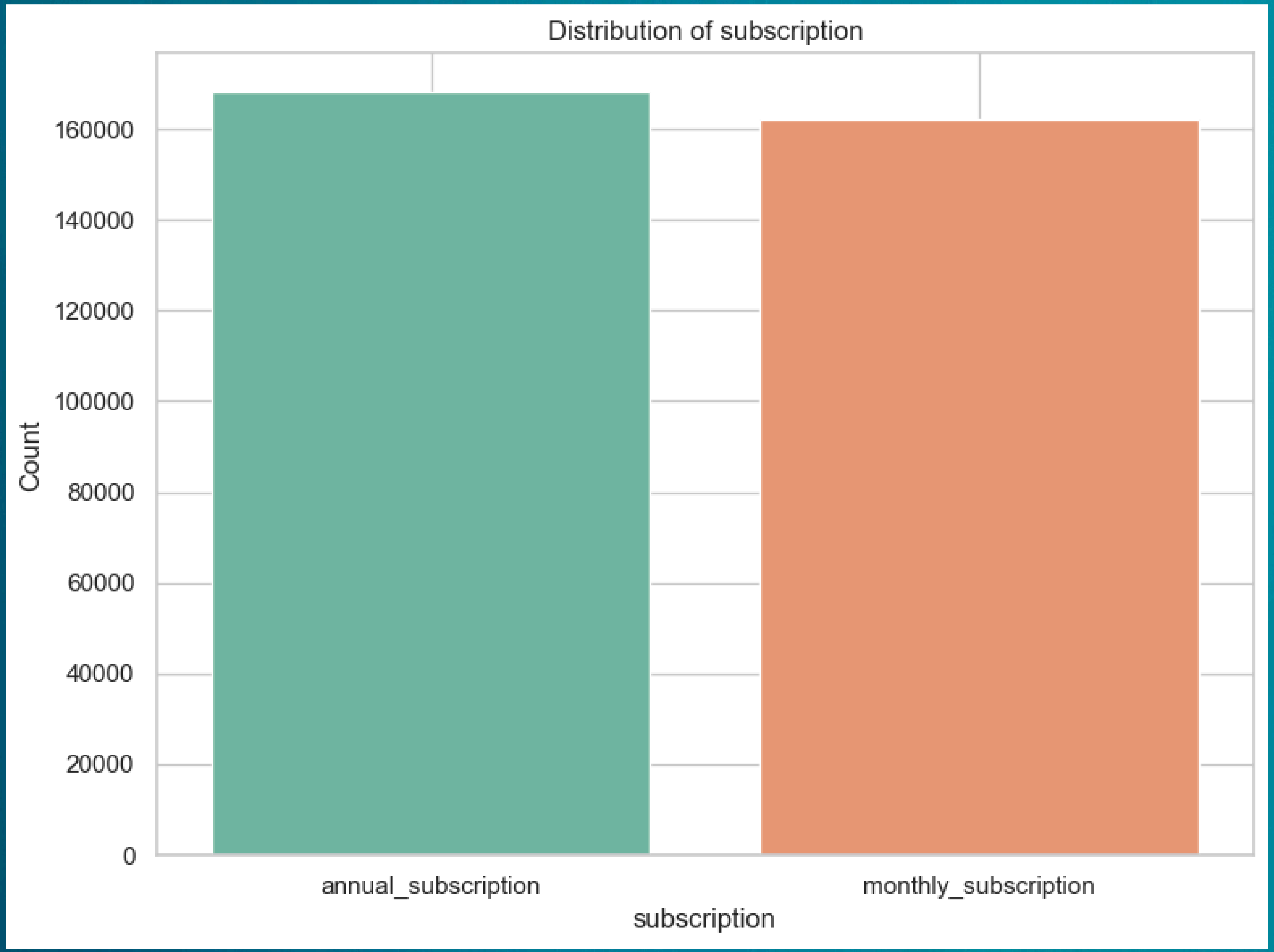
DATA ANALYSIS



Customer Lifetime Value (CLV) follows a normal distribution, with the majority of values clustering around a mean between $4000 to $6000. This distribution provides insight into the typical range of CLV values within the customer base.

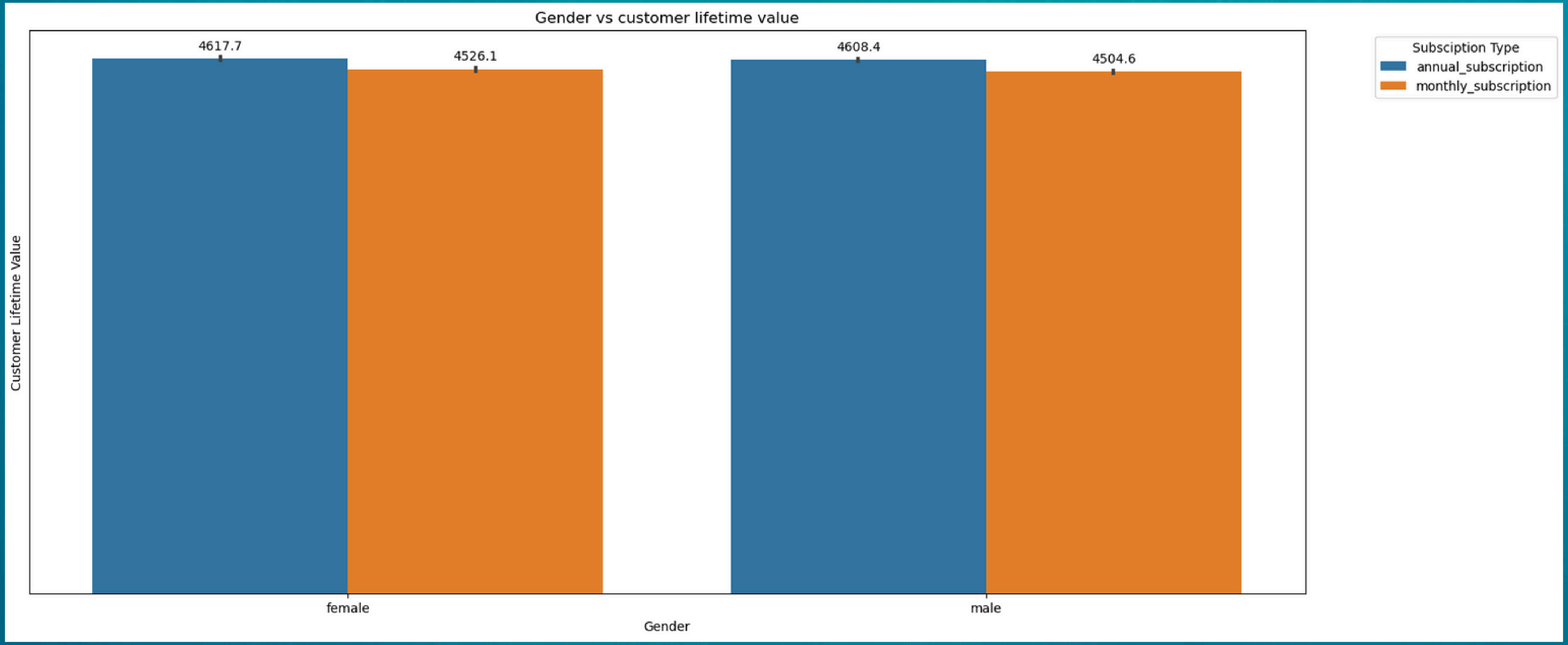
In our dataset, we observe a higher proportion of female customers compared to male customers. This demographic insight highlights the prevalence of female representation within our customer base





The bar chart displays the breakdown of subscription types, revealing a nearly equal proportion between annual and monthly subscriptions. This parity suggests a balanced preference among customers for both subscription durations, indicating a diverse customer base with varied subscription preferences.

The bar chart displays the breakdown of subscription types, revealing a nearly equal proportion between annual and monthly subscriptions. This parity suggests a balanced preference among customers for both subscription durations, indicating a diverse customer base with varied subscription preferences.



7

MODELLING

- Explored multiple models: Linear Regression, XGBoost, Random Forest, and Neural Network.
- Conducted cross-validation to evaluate model performance.
- Identified XGBoost as the top performer.
- Selected XGBoost for hyperparameter tuning due to its superior performance.

MODEL DEPLOYMENT

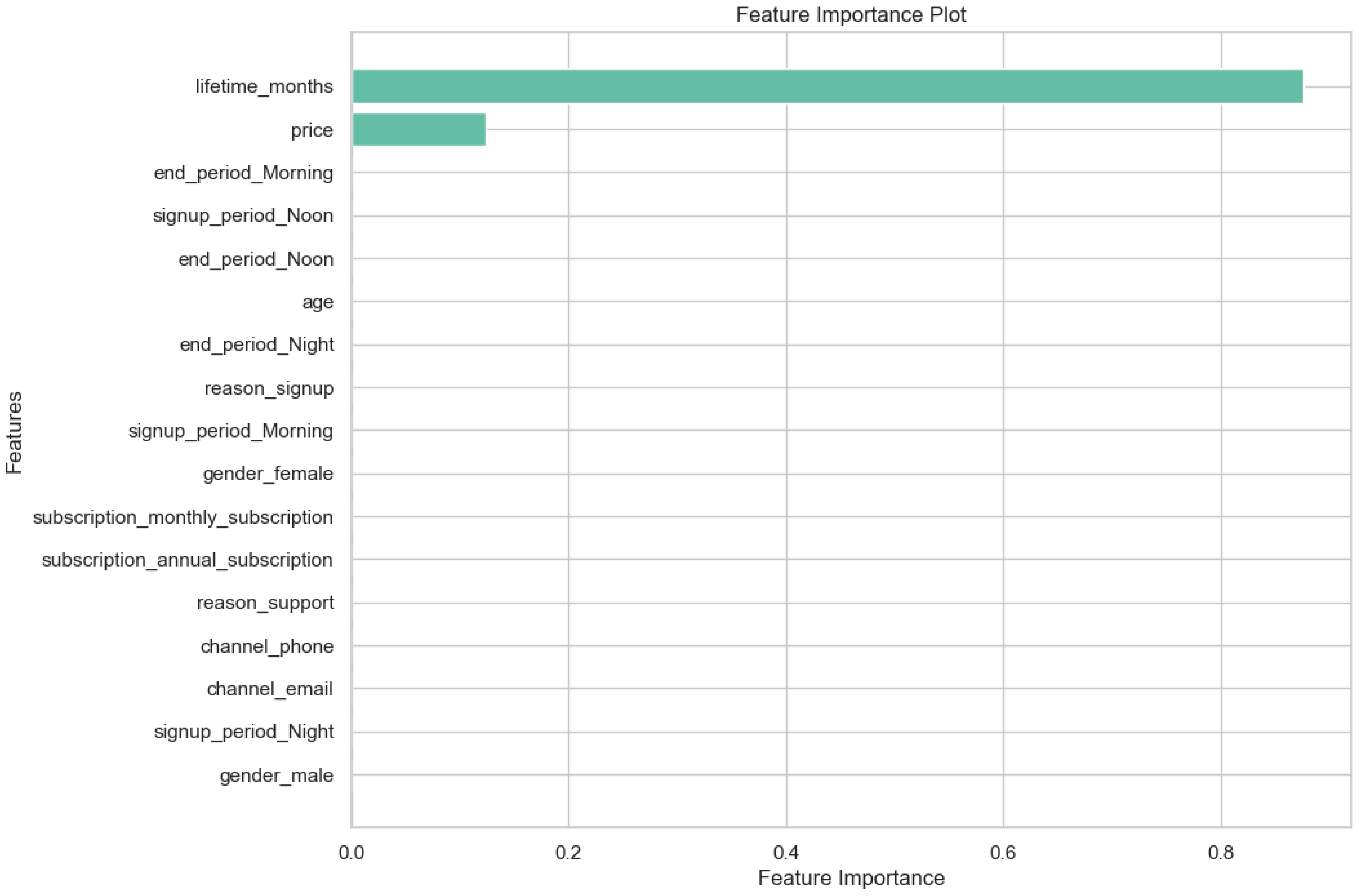
- Our model was deployed through a user-friendly website interface.
- The deployed model is based on XGBoost, hyperparameter-tuned for optimal performance.
- Users can input their data and receive real-time predictions.
- The website imports the tuned XGBoost model to predict unseen data.
- Predictions include an RMSE result, indicating prediction accuracy.
- Users receive insights on data improvements needed for better predictions.
- This interactive platform enhances decision-making and business strategies.

- Utilized RMSE, MAE, and R-squared metrics for evaluation.
- XGBoost outperformed other models due to its robustness in handling complex relationships and large datasets.
- RMSE (Root Mean Squared Error) measures the average deviation of predicted values from actual values, with lower values indicating better model fit.
- MAE (Mean Absolute Error) quantifies the average absolute difference between predicted and actual values.
- R-squared (Coefficient of Determination) assesses the proportion of variance in the dependent variable explained by the independent variables, with values closer to 1 indicating better model performance.
- XGBoost demonstrated superior performance with an RMSE of 0.0117, MAE of 0.0081, and R-squared of 0.9999, highlighting its accuracy in predicting customer lifetime value. Conversely, the Random Forest model achieved perfect scores due to overfitting, while the Neural Network model exhibited higher errors and lower explanatory power.

CONCLUSIONS

RECOMMENDATIONS

LIMITATIONS



FEATURE IMPORTANCE

Conclusions

- Developed a robust predictive model for estimating Customer Lifetime Value (CLV) in subscription-based businesses.
- Leveraged advanced machine learning techniques including data preprocessing, feature engineering, model selection, and hyperparameter tuning.
- Explored various models including Linear Regression, XGBoost, Random Forest, and Neural Network, with XGBoost emerging as the top performer.
- Achieved high accuracy with XGBoost model, exhibiting impressive metrics such as RMSE and MAE.
- Identified key drivers of CLV, including lifetime months and subscription price, emphasizing the importance of customer retention and pricing strategies.
- Deployed the model in a user-friendly website interface for predicting CLV based on user input, facilitating strategic decision-making for businesses.

Recommendation

- Implement targeted marketing strategies: Utilize insights from the CLV model to tailor marketing campaigns towards high-value customer segments, maximizing conversion rates and revenue generation.
- Enhance customer retention efforts: Develop personalized retention programs focusing on extending customer lifetimes by addressing churn risk factors and enhancing overall customer satisfaction.
- Optimize pricing strategies: Leverage the correlation between subscription price and CLV to adjust pricing tiers, ensuring alignment with customer value perceptions while maximizing long-term profitability.
- Continuously monitor and refine the CLV model: Regularly update the model with new data to adapt to evolving customer dynamics and market conditions, ensuring its effectiveness in driving strategic decision-making.
- Invest in customer experience initiatives: Prioritize investments in enhancing the overall customer experience to foster long-term loyalty and increase CLV, ultimately driving sustainable business growth.

Limitations

- Limited Models Used: The CLV model's performance is constrained by the selection of machine learning algorithms considered during the modeling process. Additional algorithms or ensemble methods might offer different insights and improve predictive accuracy.
- Model Generalization: While the CLV model may perform well on the training and test datasets, its performance on unseen data or in different market contexts may vary. Generalizing the model to diverse customer segments and market conditions requires careful validation and testing.
- Interpretability: Complex machine learning models like XGBoost may lack interpretability, making it challenging to understand the underlying factors driving CLV predictions. Clear interpretation of model outputs is essential for effective decision-making.

THANK YOU

connecting us with the world

Victoria Plaza

hello@analyticavengers.com

www.analyticavengers.com