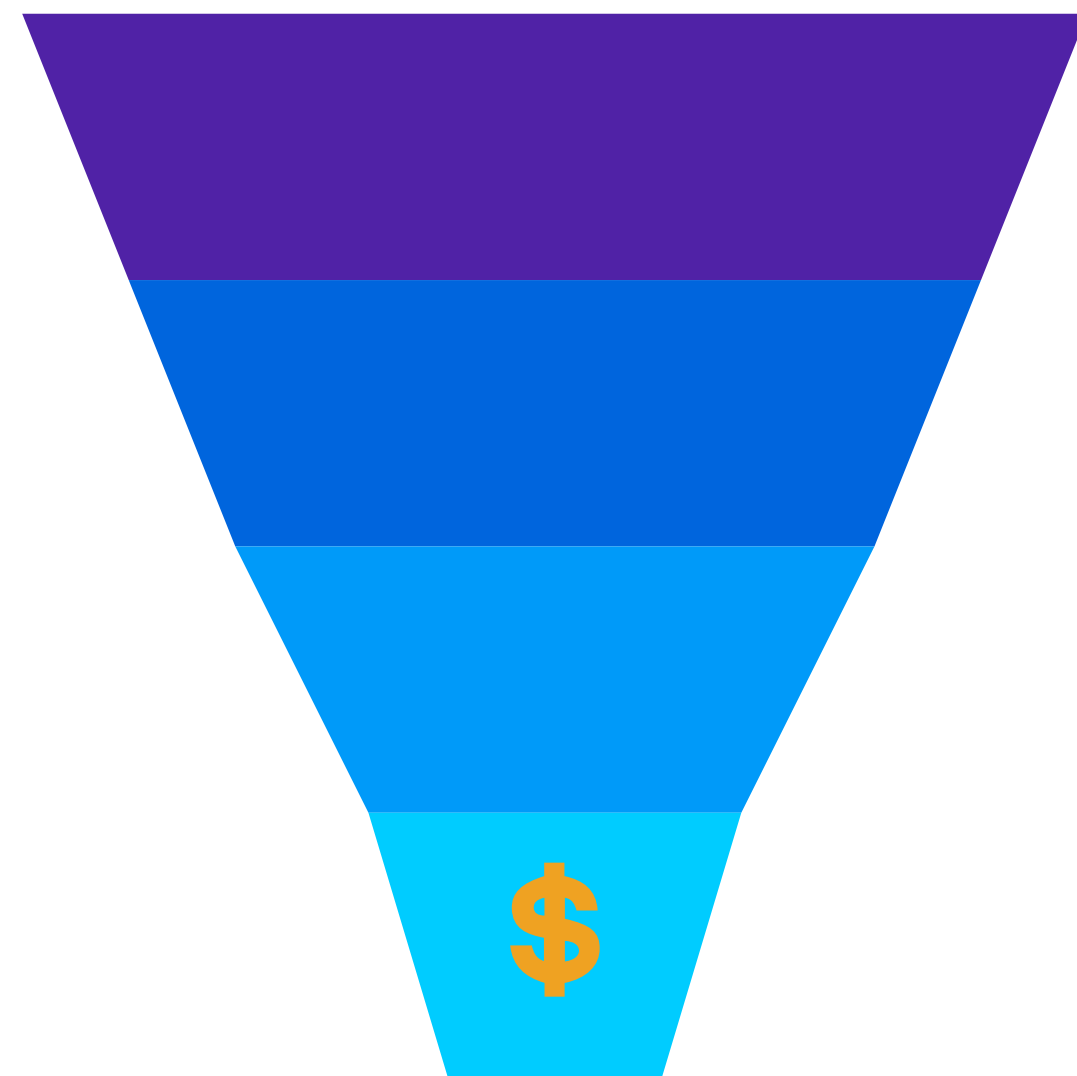
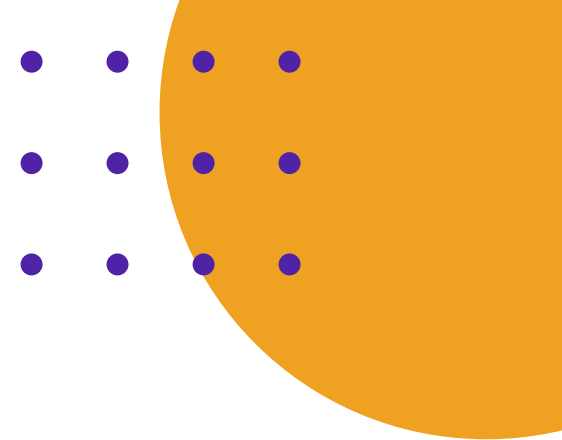
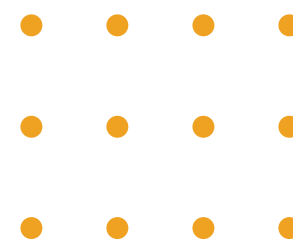




Sales Forecasting

Sunny, Harsha, Dinesh, Pranjal





Problem Statement

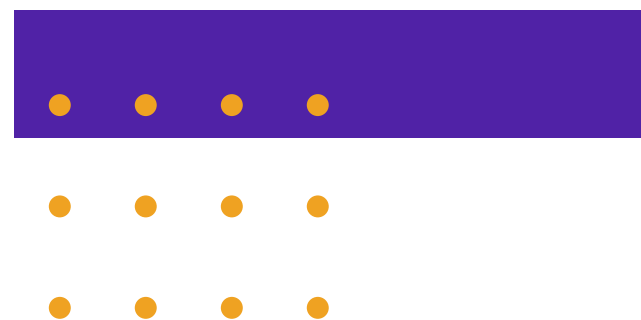
Are we on track to meet our quarterly sales target?

Objective

Predict the deal outcome early in the quarter to have better revenue predictability and enable insights for timely decision making

Potential Challenges

- Data threading across stages of the sales funnel
- Data volume - will this require distributed ML?
- Feature selection given the potentially wide range of available attributes
- Learn the Art of Data Storytelling - present insights in a meaningful and convincing way





Team Members and Tentative Responsibilities

Sunny Mattas : Data Preparation and Model Building
Harsha Bompalli Mutt : Feature Selection and Model Tuning
Dinesh Kumar : EDA and Visualization
Pranjal : Model Tuning and Results Interpretation

Inspiration

Many companies have been missing their financial targets (one of the many reasons for recent layoffs in the software industry).

Based on our general readings, we realized that companies need more predictability of their revenues and better understand the controllables to prevent revenue leakages.



Activities Involved

Data preparation & Data Cleaning

- 6 quarters of sales pipeline data
- All open opportunities as of any given week of the quarter is the starting dataset
- Measurement of the outcome of these opportunities by the end of the quarter
- Data cleanup removing bad data



Exploratory Data Analysis

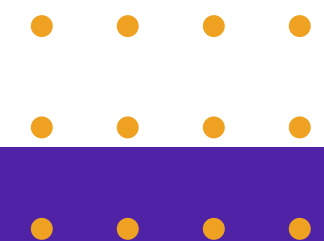
- Visualization
- Examining Relationships
- Examining distributions
- Outlier Detection and Treatment

Feature Engineering

- Correlation Analysis
- Feature Selection
- Define binary Class Variable as opportunity Won Vs Not Won (Lost + Open at the end of the quarter)

Model Development

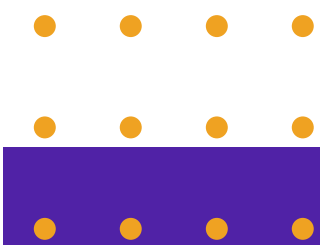
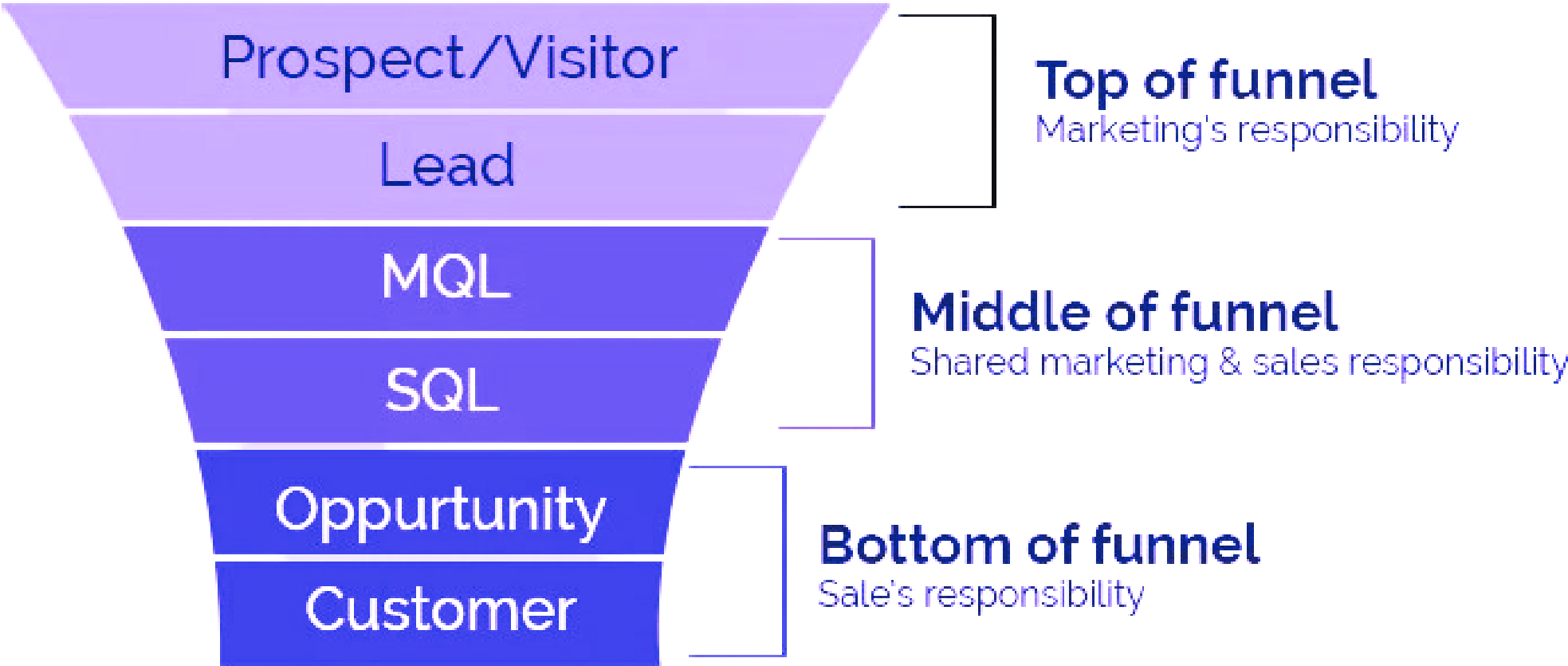
- Divide dataset into training and testing datasets - 5 quarters of training data and most recent quarter of testing data
- Predict the probability of winning opportunity by the end of the quarter
- XGboost



Sales Pipeline Overview

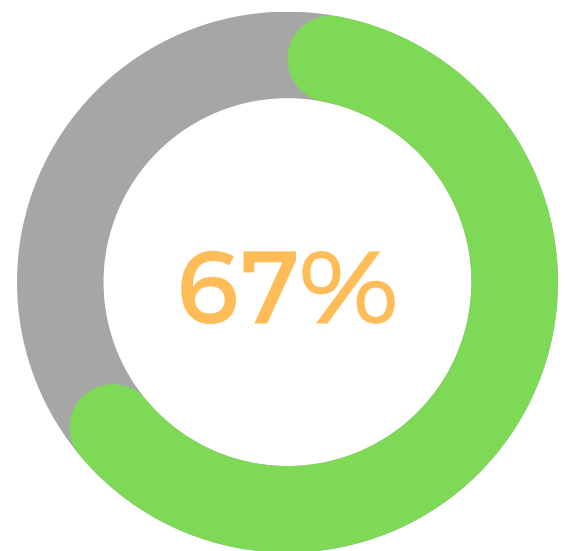


MARKETING & SALES FUNNEL

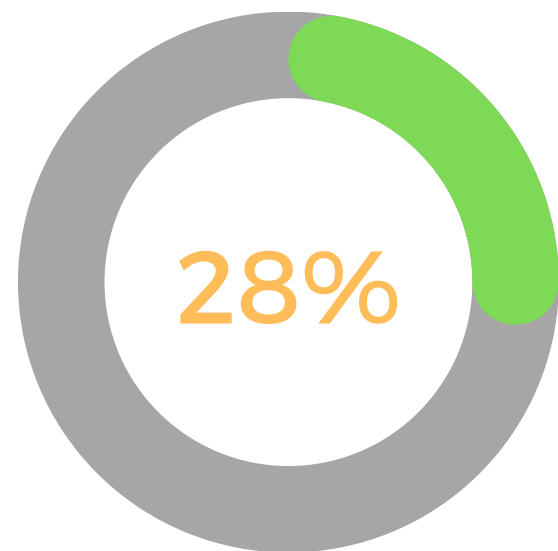


Key Findings

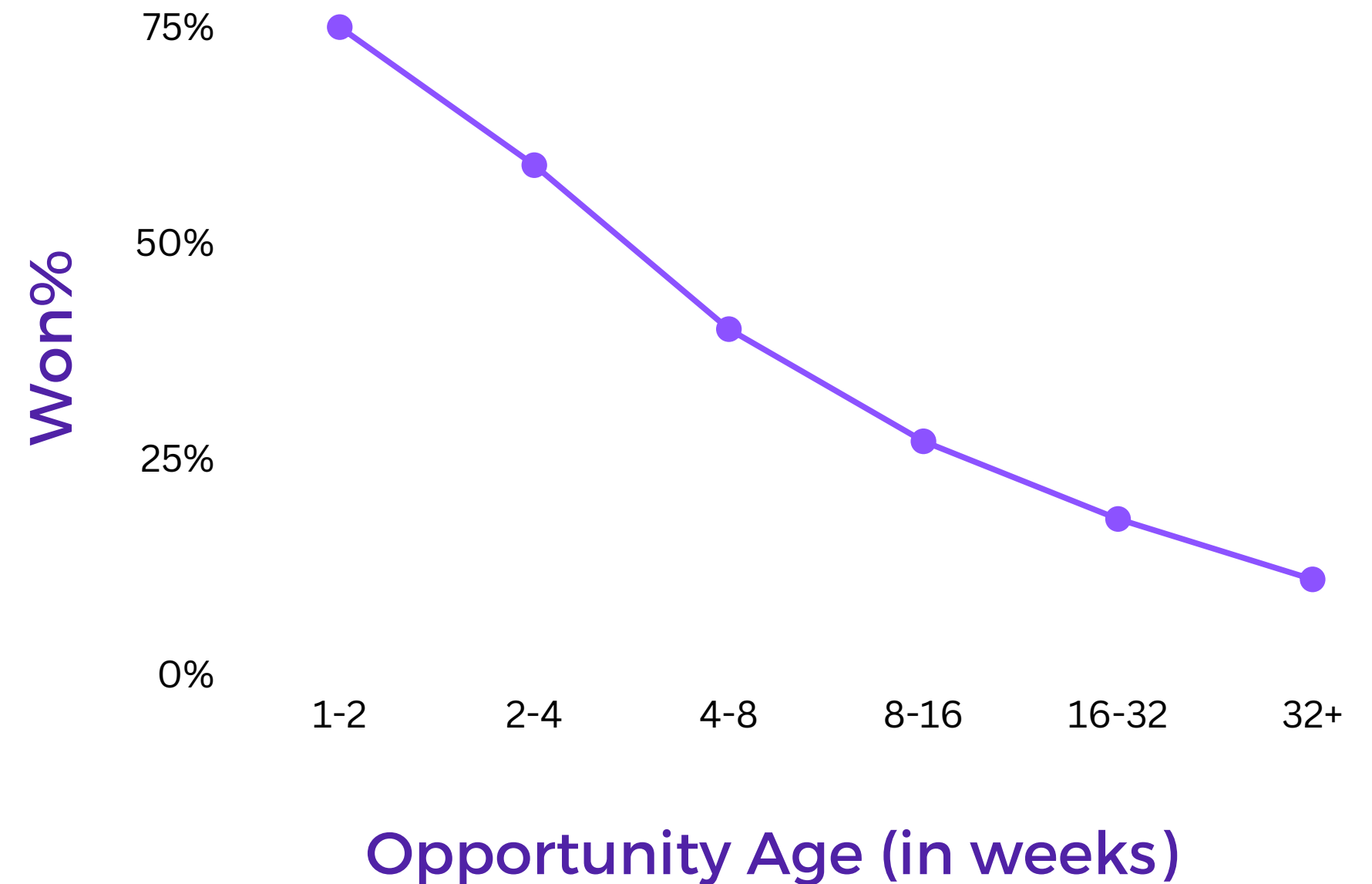
There is **2.5X** higher probability for an opportunity to be WON if closed in less than 4 weeks



Won%
age <= 4weeks

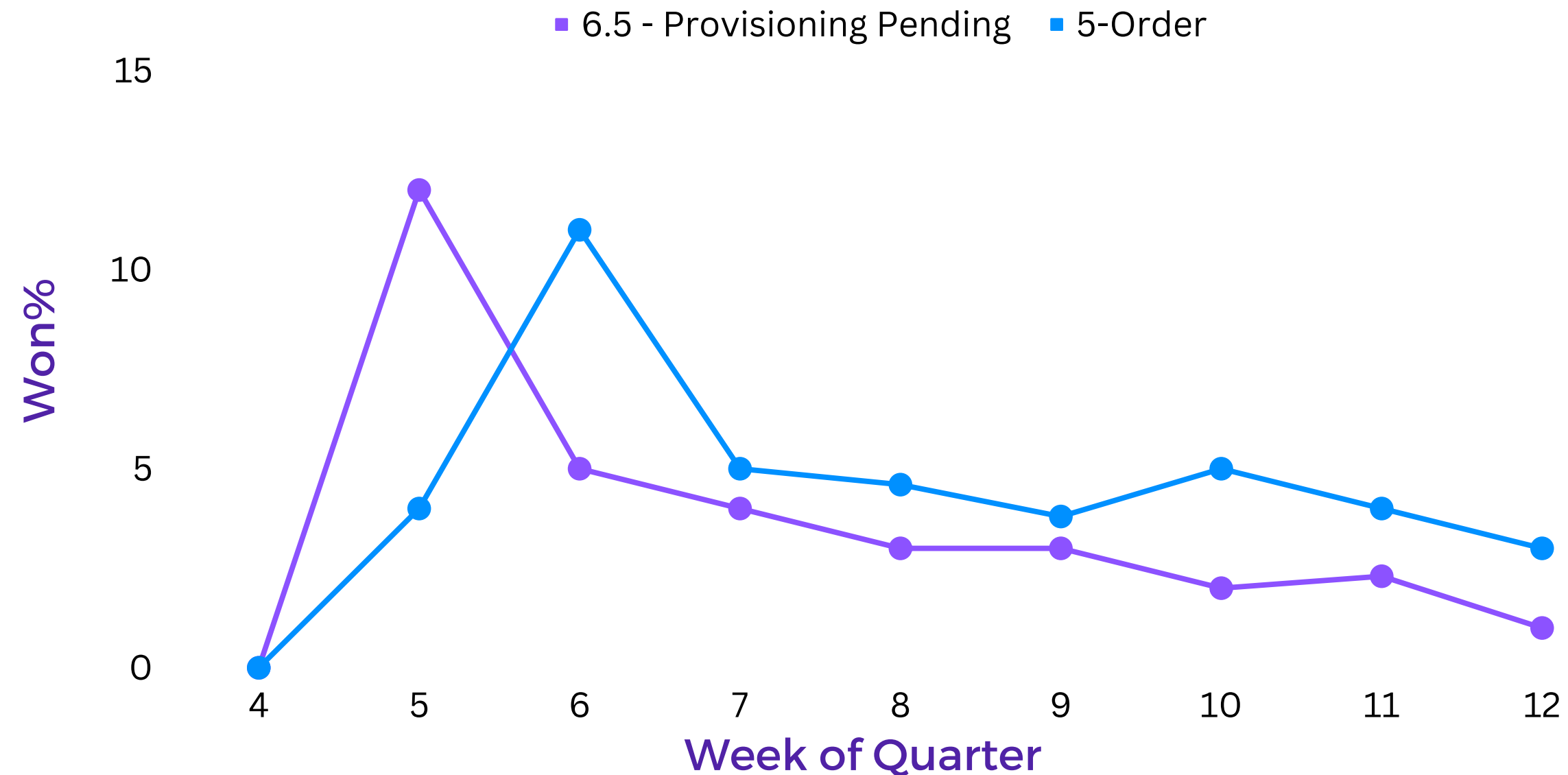


Won%
age <= 4weeks



Key Findings

Probability of an opportunity to be **WON** decreases if it stays in the same stage (especially after stage 4-negotiate) **longer than ~2 weeks**

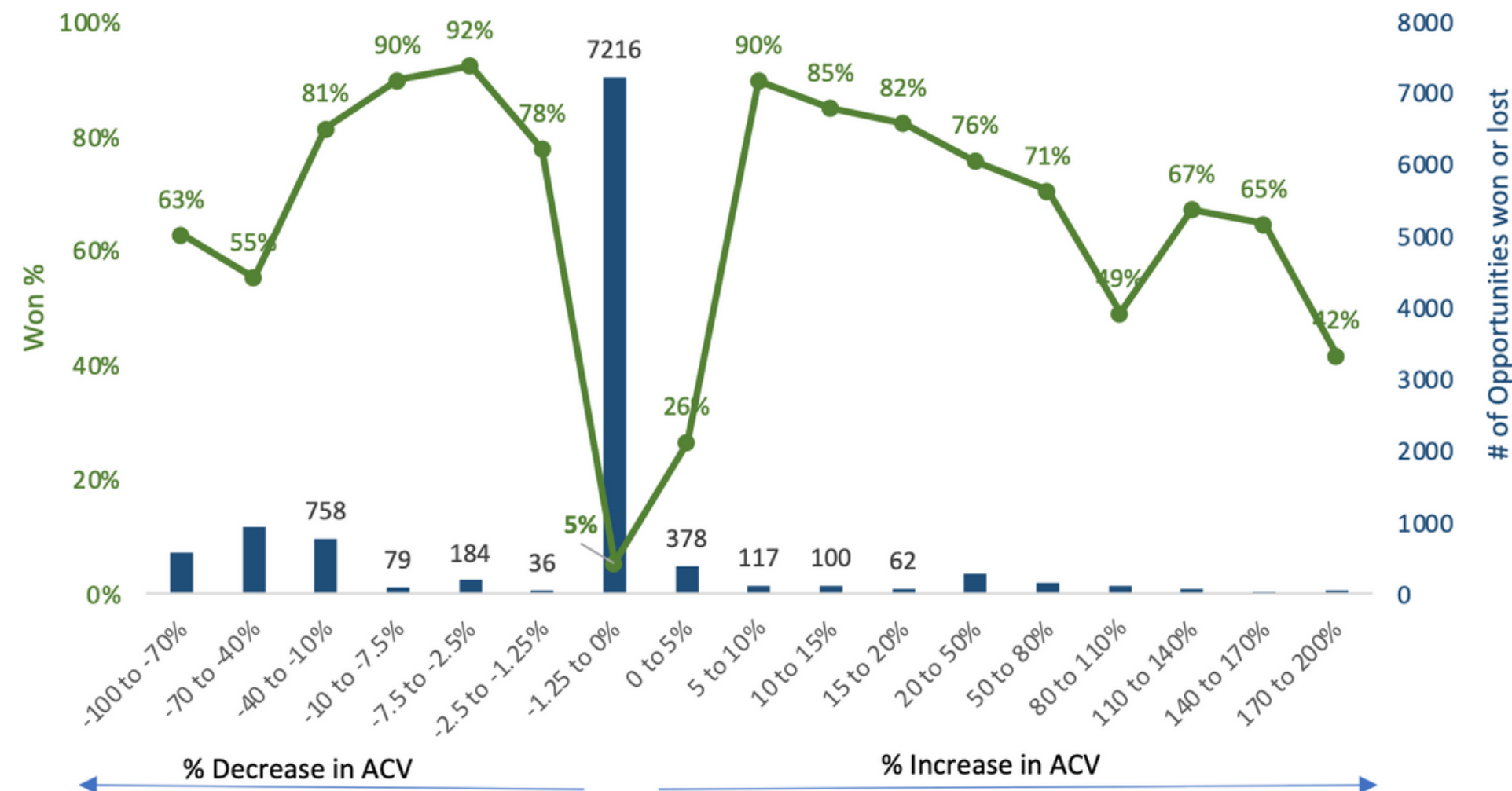


Won % for opportunities open in respective stage as of week 3

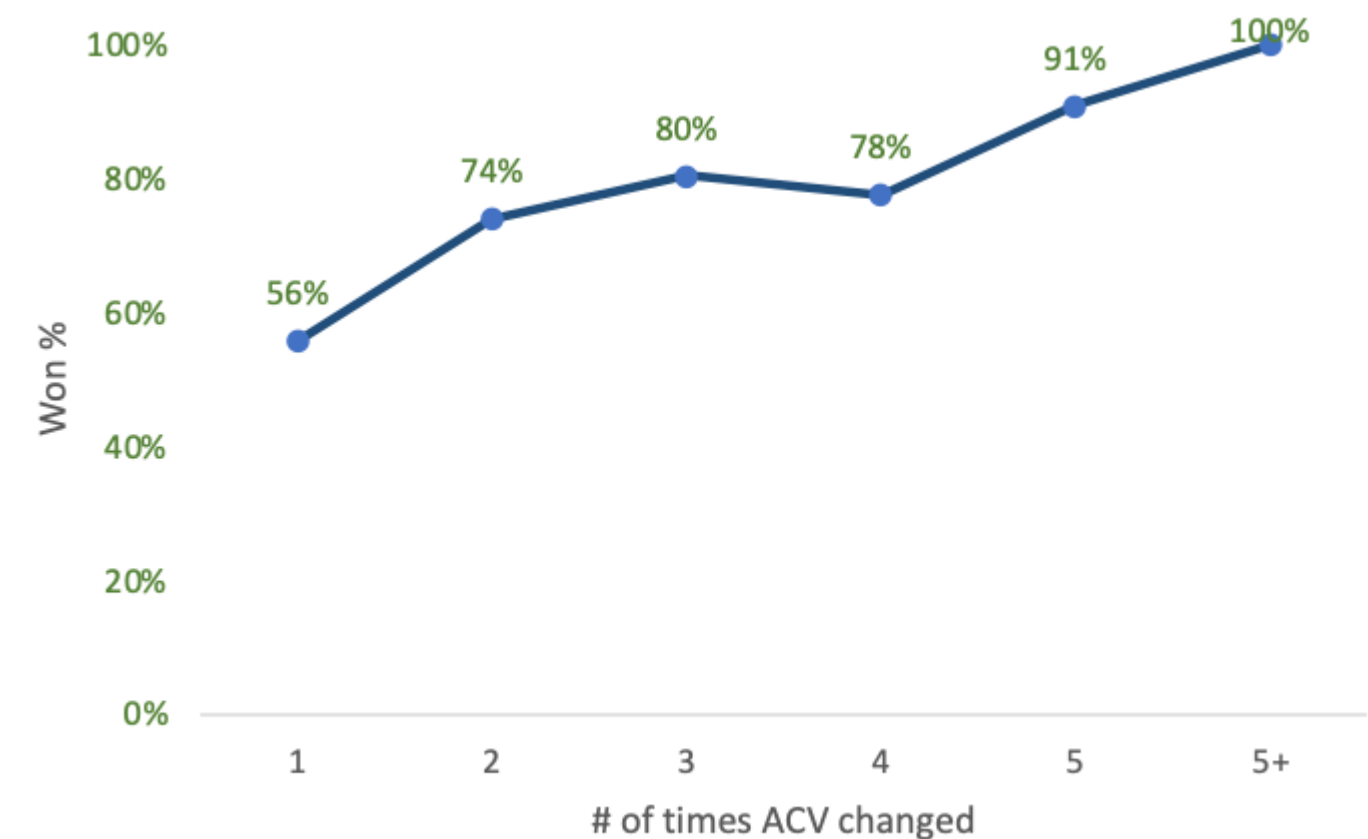
Key Findings

Active engagement with customers results in a Win-Win Scenario!!

- On an average, 85% opportunities are won if ACV was either increased or decreased within a range of 5-40%.
- Only 5% of the opportunities are won when ACV remained unchanged.



- An opportunity has 31% more chance of winning when it undergoes more than 2 negotiations during its lifecycle



Key Findings

EXISTING base matters while we pursue NEW logos

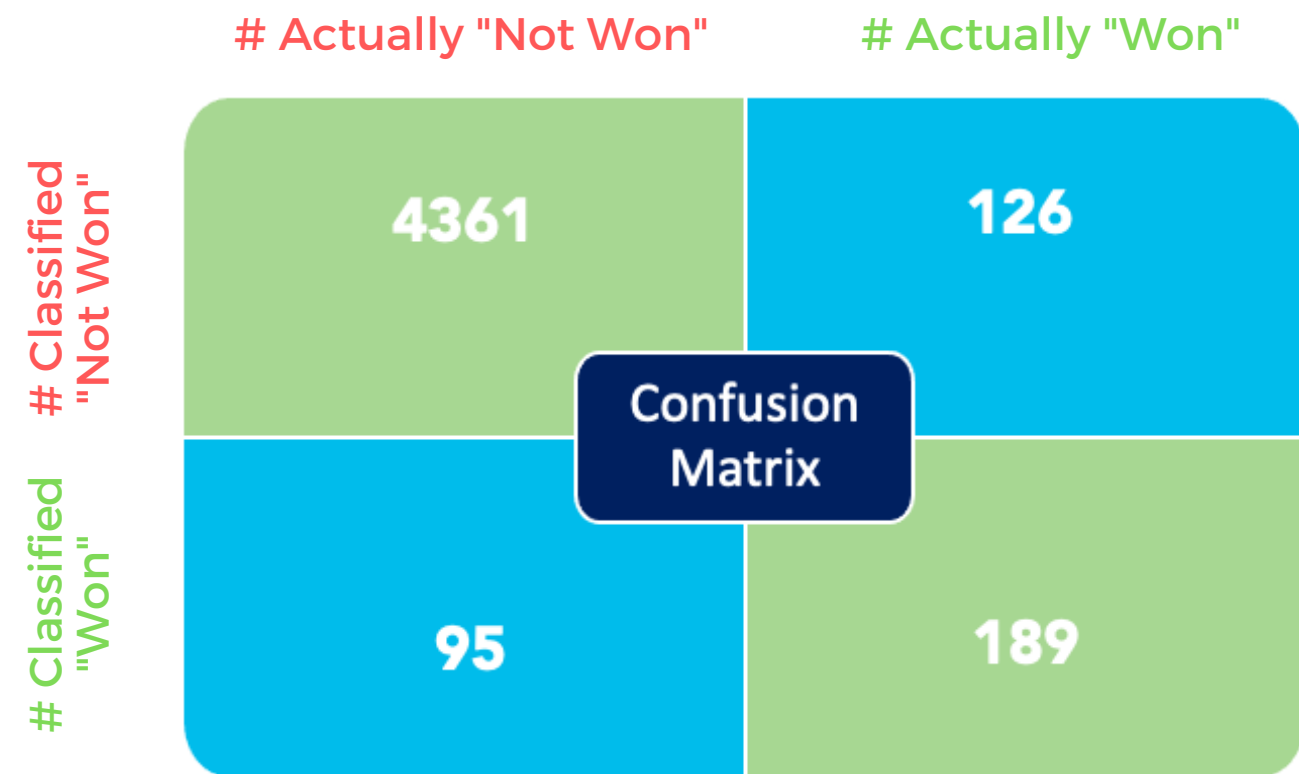
- Opportunities for EXISTING customers has **2.5X** higher win-rate compared to NEW customers
- Opportunities with "Product" as part of ELA has 2X higher win-rate than those where it was sold a-la carte



VS

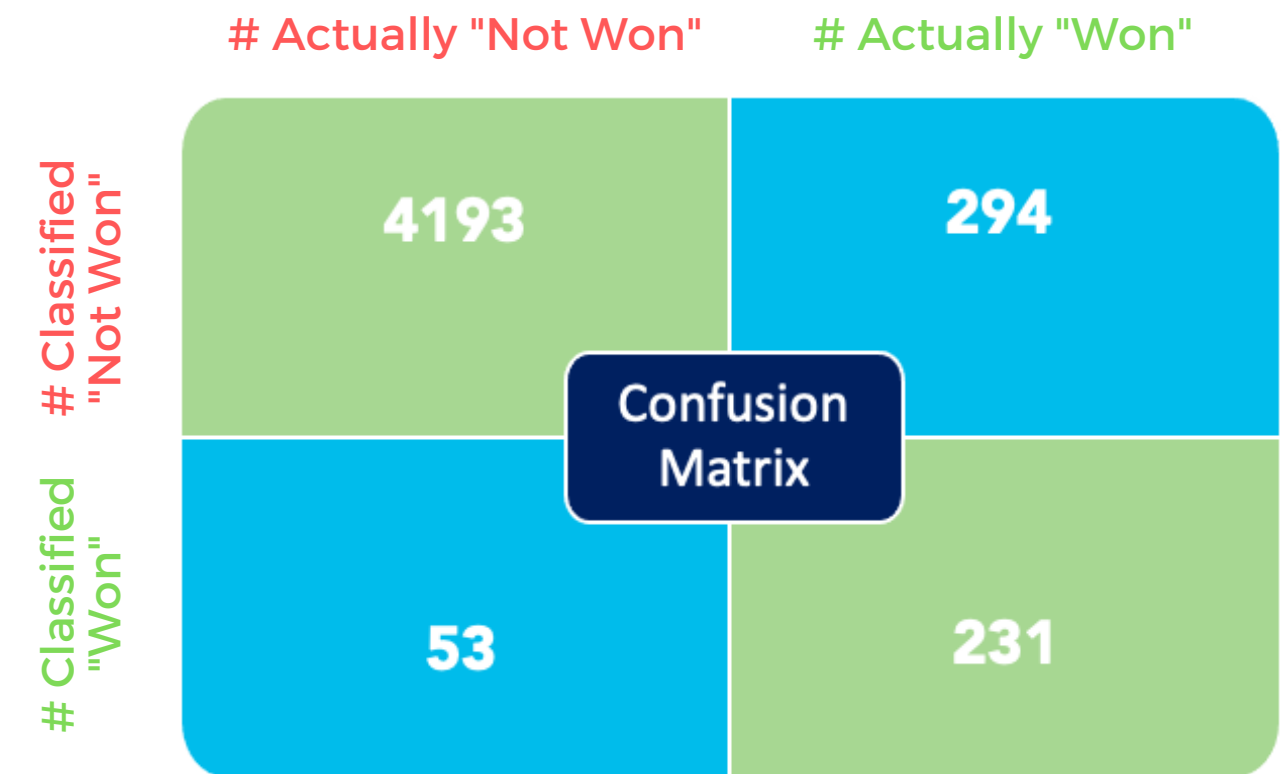


Preliminary Model Results for XGBOOST



Confusion Matrix based on a classification threshold that maximizes f1 score on the test data

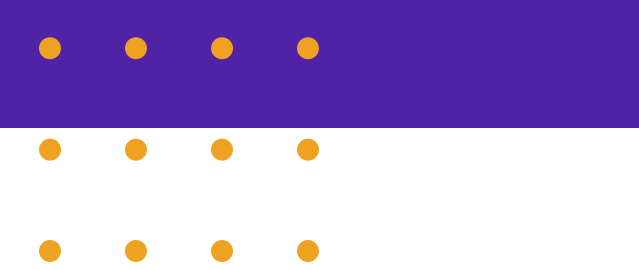
Model Accuracy Score **95.4%**



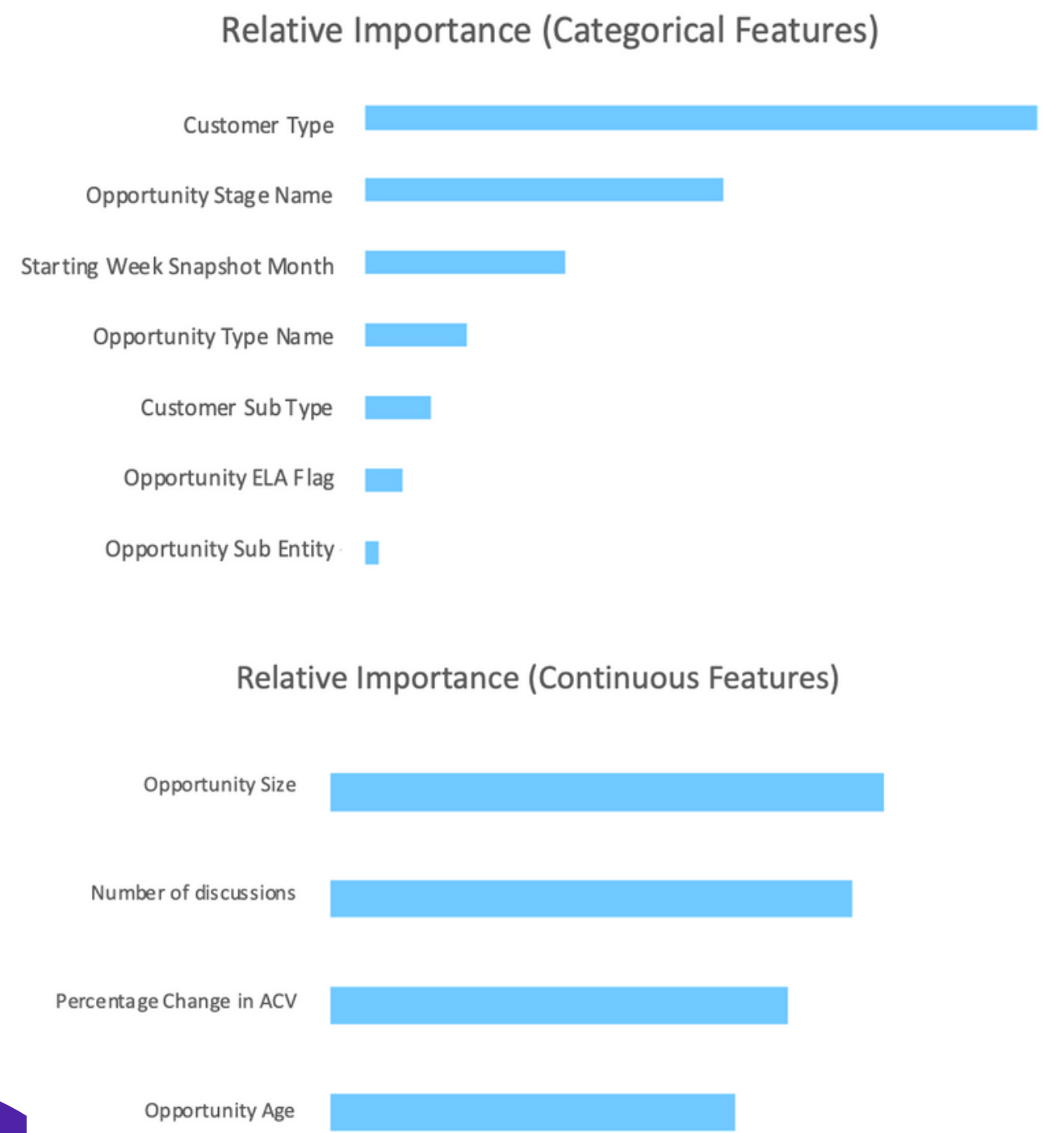
Confusion Matrix based on a classification threshold that maximizes maximizes TPR on test data

Model Accuracy Score **92.7%**

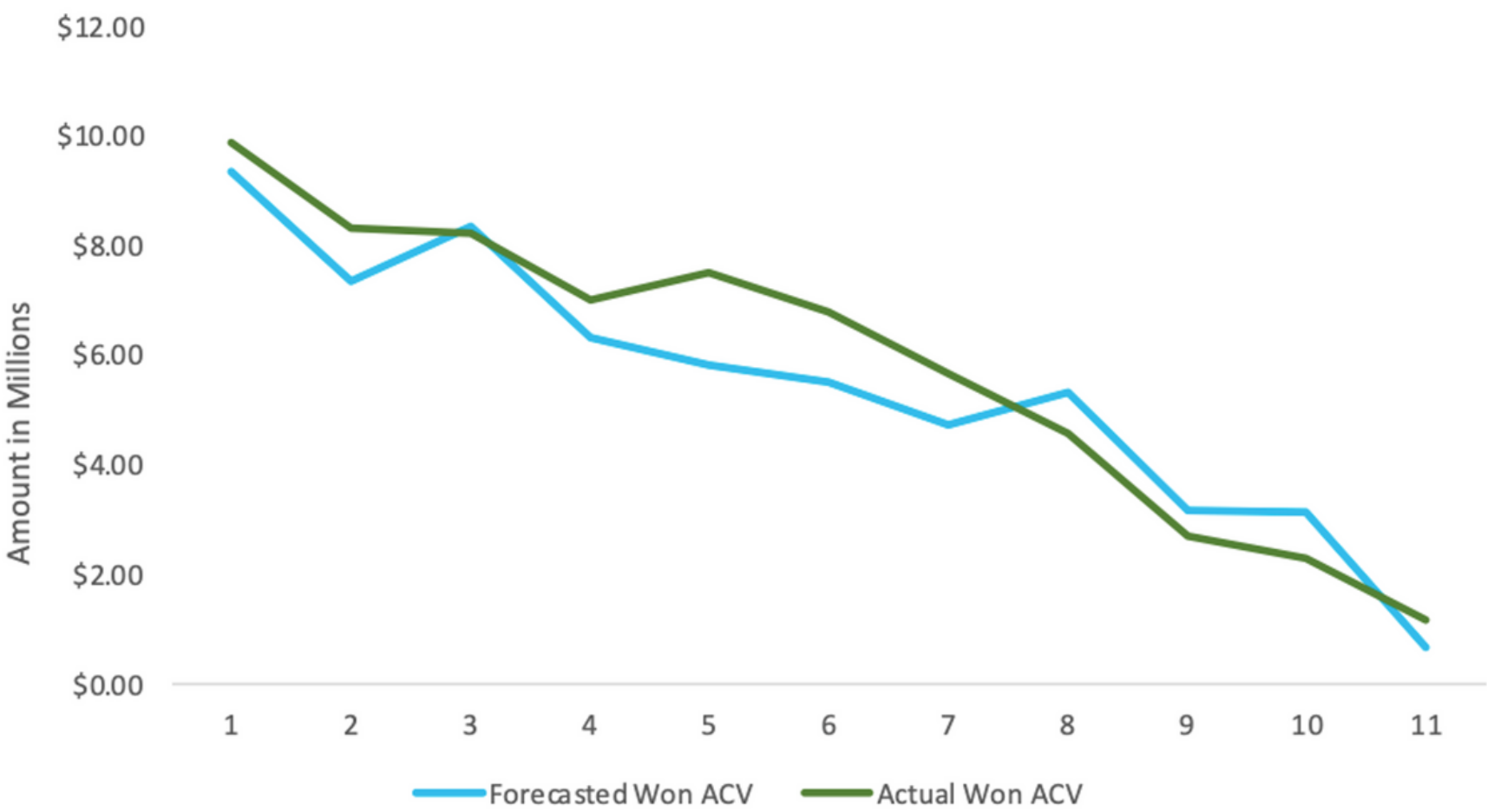
Preliminary Model Results for XGBOOST



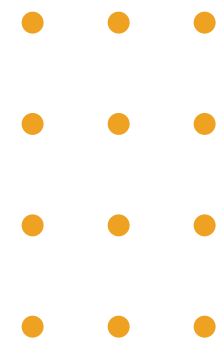
Relative Importance of the Features



Total ACV of opportunities forecasted to be Won on Testing Dataset



Model scored as of each starting week in Q1 2023



Thank You

