Lead Conversion Prediction Model: Analysis and Insights

- **Presented by**: Yashaswi Mahindrakar

- **Date**:30 Spetember 2024

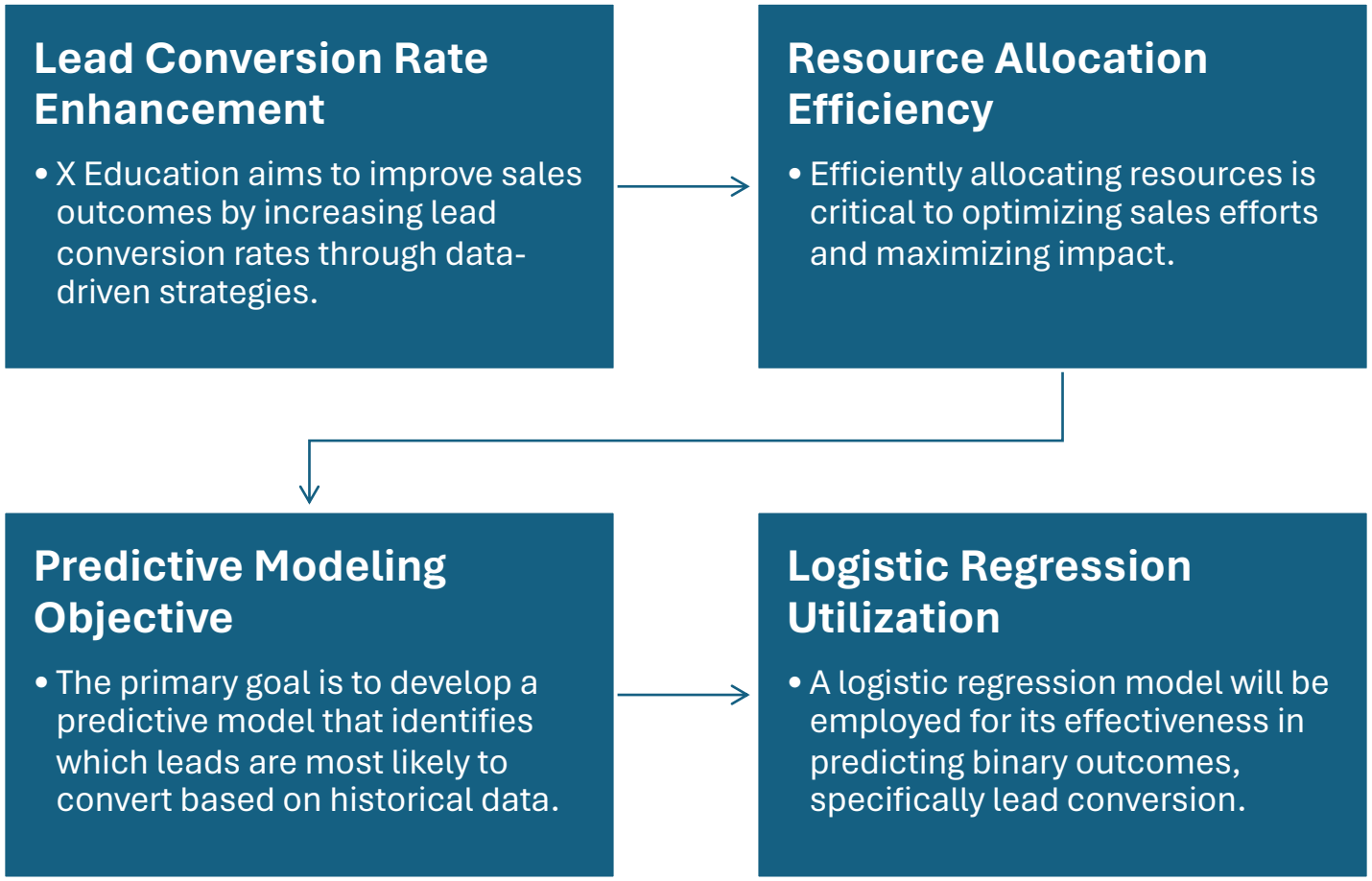
Problem Statement & Objective

Problem Statement:
X Education wants to enhance lead conversion rates and allocate resources efficiently to improve sales outcomes. The goal is to predict which leads are most likely to convert based on historical data.

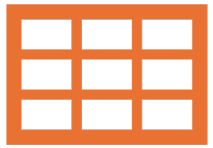
Objective:
Develop a **logistic regression model** to predict lead conversion and provide actionable insights to optimize resource allocation during sales efforts.

Understanding the Challenge

Enhancing Lead Conversion Rates through Predictive Modeling



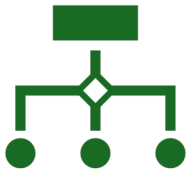
Approach Overview



Data Cleaning & Preprocessing

Handled missing values and performed categorical encoding.

Standardized numeric variables.



Model Selection

Implemented **logistic regression** due to its interpretability and suitability for binary classification.

Split data into training (70%) and testing (30%) sets for validation.

Model Evaluation

Key metrics: **Accuracy**, **Precision**, **Recall**, and **ROC-AUC**.

Key Technical Results

- Model Performance and Top Predictors of Conversion

	Model Metrics	Top Predictors
2	Accuracy	94%
3	ROC-AUC Score	0.98
4	Tags: Will revert after reading the email	Positive
	Lead Profile: Potential Lead	Positive
6	Lead Quality: Worst	Negative

Business Impact - Results in Simple Terms

High Conversion Potential Leads: Leads tagged with “Will revert after reading the email” and classified as “Potential Lead” are the most likely to convert. These leads should be prioritized in the sales process.

Low Conversion Potential Leads: Leads rated as "Worst" in **Lead Quality** should receive lower priority, as they are less likely to convert.

Actionable Insights - Intern Phase Strategy

Maximizing Conversions During the Intern Phase:

Focus on High-Probability Leads: Interns should prioritize calling leads predicted to have a high conversion probability (model outputs close to 1).

Segment Leads by Profile: Tag-based leads such as "Will revert after reading email" should be the first point of contact.

Insert flowchart outlining intern call strategy for high-probability leads

Variable	Coefficient
Tags_Will revert after reading the email	2.421309
Lead Profile_Potential Lead	1.489551
Lead Quality_Worst	-1.303787

Actionable Insights - Post-Target Strategy

Reducing Unnecessary Calls Once Target is Met:

- **Set a Threshold for Contact**: Only engage leads with a predicted probability of conversion above 80% (based on model predictions).
- **Use Non-Intrusive Methods**: For lower-probability leads, use emails or SMS to maintain communication without making phone calls.

Conclusion

EFFECTIVE MODEL: THE LOGISTIC REGRESSION MODEL HAS BEEN HIGHLY EFFECTIVE IN PREDICTING LEAD CONVERSION, WITH AN AUC OF 0.98.

ACTIONABLE INSIGHTS: STRATEGIC USE OF THE MODEL'S OUTPUTS CAN HELP ALLOCATE RESOURCES EFFICIENTLY, PARTICULARLY DURING THE INTERN PHASE OR AFTER REACHING TARGETS.

NEXT STEPS: FURTHER REFINEMENT OF THE MODEL COULD INCLUDE TESTING MORE ADVANCED TECHNIQUES LIKE DECISION TREES OR ENSEMBLE METHODS FOR IMPROVED PRECISION.