

Lead Scoring Model for X Education

This presentation dives into a comprehensive lead scoring model designed to optimize sales performance at X Education. We'll explore the business problem, the technical approach, key findings, and actionable insights for maximizing lead conversion.

 by Y Shivam Kumar



Business Problem

Sales Optimization

X Education aims to enhance its sales process by identifying high-potential leads for conversion.

Lead Scoring Model

The objective is to develop a model that assigns a probability score to each lead, indicating their likelihood of conversion.

Sales Team Prioritization

This model will empower the sales team to prioritize leads effectively and focus their efforts on those most likely to convert.

Solution Approach

1 Data Preprocessing

Cleaning and preparing the data for model training.

3 Feature Importance Analysis

Identifying key variables influencing conversion.

2 Model Training

Utilizing Logistic Regression to predict conversion probability.

4 Lead Scoring System

Assigning scores between 0-100 to rank leads based on their conversion potential.

How to create your Lead Scoring model



Data Preprocessing

Handling Missing Values

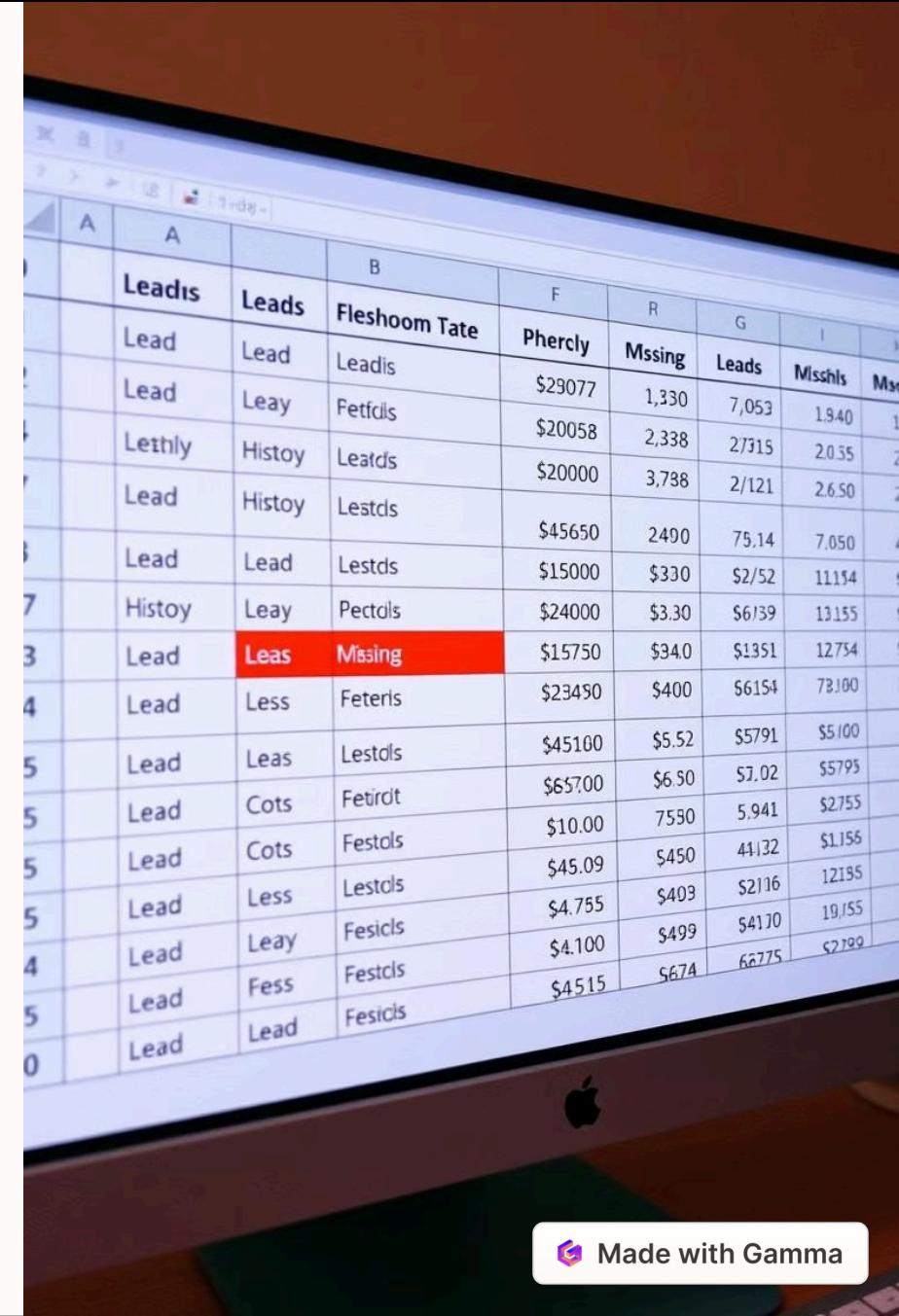
Dropping columns with high missing values and imputing categorical values with 'Unknown' for a comprehensive approach.

Encoding Categorical Variables

Applying one-hot encoding to transform categorical features into numerical values for model compatibility.

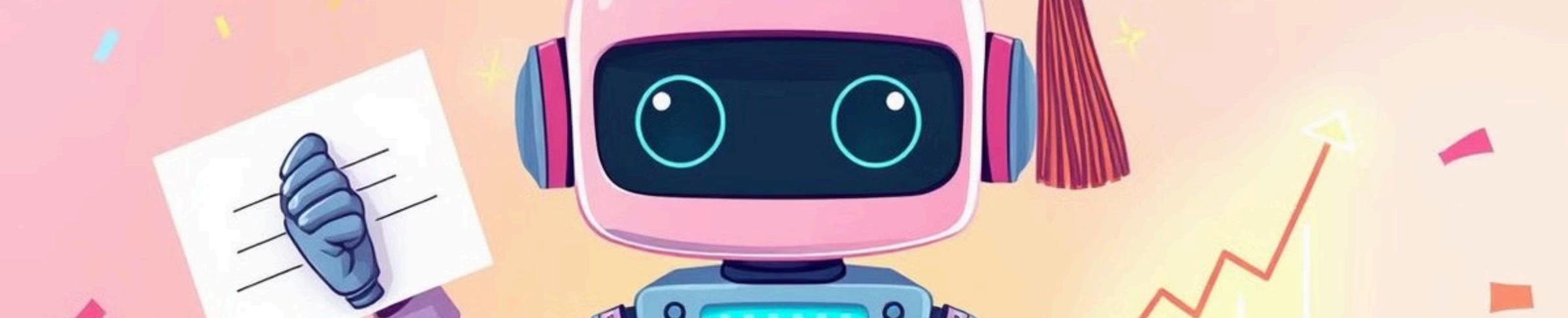
Scaling Numerical Variables

Standardizing numerical features for better model performance and convergence, ensuring all variables are on a similar scale.



A screenshot of a Microsoft Excel spreadsheet titled '3-day'. The data is organized into several columns:

	A	A	B	F	R	G	I	J	K	L	M
	Leadis	Leads	Fleshoom Tate	Phercly	Mssing	Leads	Misslhs	Miss	Miss	Miss	Miss
1	Lead	Lead	Leadis	\$29077	1,330	7,053	19.40	1	1	1	1
2	Lead	Leay	Fetfdls	\$20058	2,338	27315	20.55	2	2	2	2
3	Lethly	Histoy	Leatds	\$20000	3,738	2/121	26.50	3	3	3	3
4	Lead	Histoy	Lestds	\$45650	2490	75.14	7.050	4	4	4	4
5	Lead	Lead	Lestds	\$15000	\$330	\$2/52	11154	5	5	5	5
6	Histoy	Leay	Pectols	\$24000	\$3.30	\$6/39	13155	6	6	6	6
7	Lead	Leas	Missing	\$15750	\$94.0	\$1351	12754	7	7	7	7
8	Lead	Less	Feterls	\$23490	\$400	\$6154	73100	8	8	8	8
9	Lead	Leas	Lestols	\$45100	\$5.52	\$5791	\$5100	9	9	9	9
10	Lead	Cots	Fetirdt	\$65700	\$6.90	\$7.02	\$5795	10	10	10	10
11	Lead	Cots	Festols	\$10.00	7550	5.941	\$2755	11	11	11	11
12	Lead	Less	Lestols	\$45.09	\$450	41132	\$1155	12	12	12	12
13	Lead	Leay	Fesicls	\$4.755	\$409	\$2116	12135	13	13	13	13
14	Lead	Fess	Festcls	\$4.100	\$499	\$4170	19/55	14	14	14	14
15	Lead	Lead	Festcls	\$4515	\$674	\$6775	\$2799	15	15	15	15
16	Lead							16	16	16	16



Model Training - Logistic Regression



Training

Trained a Logistic Regression model with 80% training and 20% testing split for comprehensive model evaluation.



Scaling

Used Standard Scaling for feature normalization, ensuring all features have similar ranges and preventing biased model training.

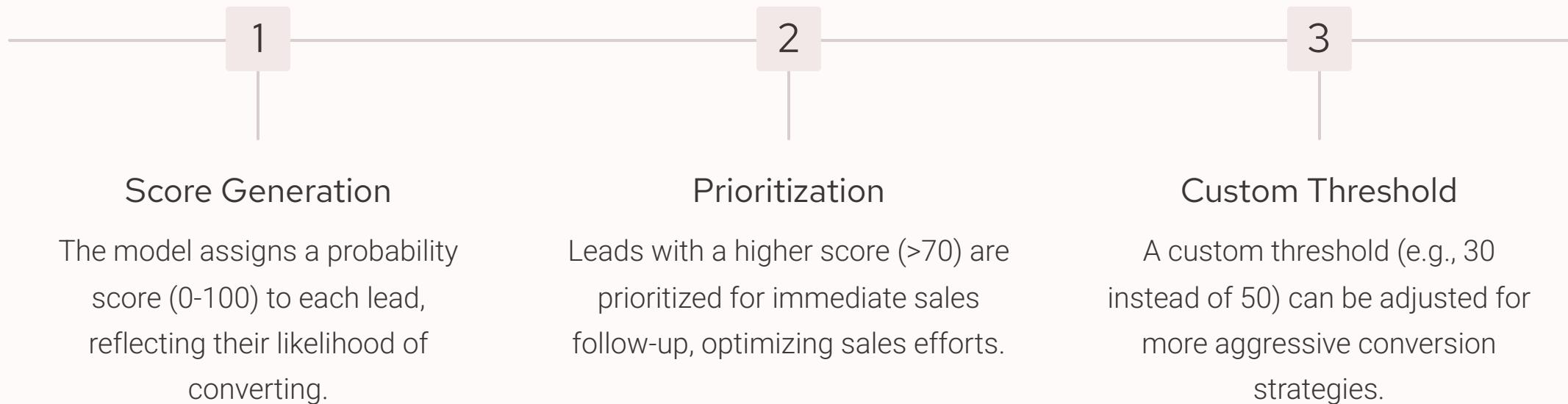


Evaluation

Evaluated model performance using accuracy, precision, recall, and the confusion matrix to assess its effectiveness.

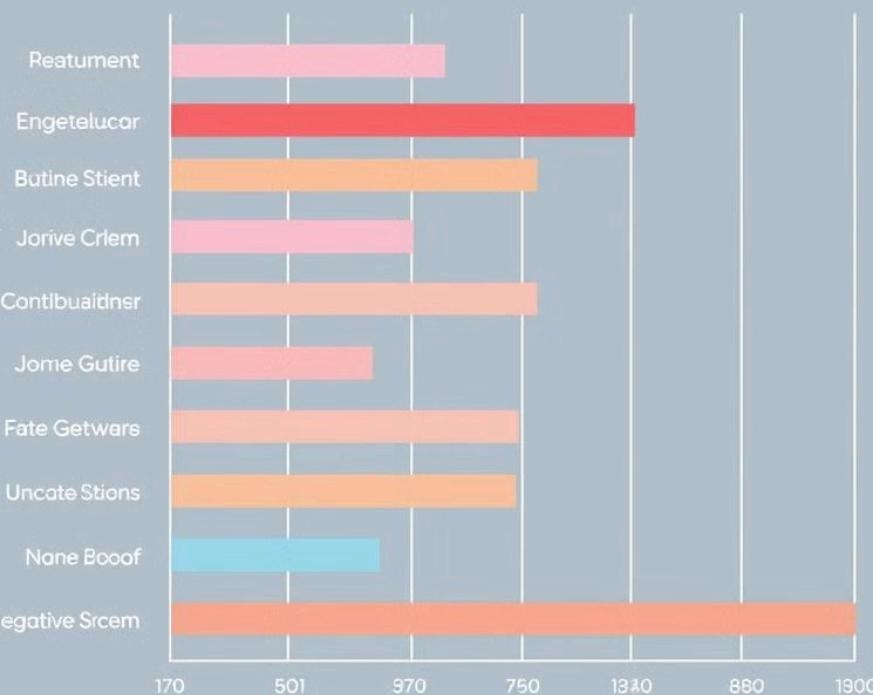
2000
2000
2000

Lead Scoring System



Top 10 Influencing Features by Impact!

The top 10 influencing features have significant impact when included in the score.



Style fun and ingeate imprent retions ser feature:
ocer are pin on lend lead and influencts.

Key Findings - Feature Importance

Positive Influencers

1

Tags_Will revert after reading the email, Tags_Unknown, Tags_Closed by Horizon, indicating potential customer engagement and positive interactions.

Negative Influencers

2

Lead Source_Reference, Tags_switched off, Specialization_Unknown, suggesting potential disengagement or lack of interest.

Business Strategy: Aggressive Lead Conversion



Business Strategy: Reducing Unnecessary Calls

1

Call

Leads with a score above 90.

2

Use

Email and SMS follow-ups instead of calls.

3

Set

A higher threshold before calling.

4

De-prioritize

Leads with wrong numbers or missing info.

Next Steps & Improvements

1

Fine-tune

Adjust classification thresholds.

2

Apply

SMOTE for class imbalance.

3

Use

RFE to refine key predictors.

4

Implement

Real-time lead scoring.

