Lead Score – Case Study
X Education

upGrad
upGrad & IIITB | Data Science Program - Feb 2023

By – Kowshik Kr Sarma

Top three variables contributing most towards conversion probability:

01

TAGS - WILL REVERT AFTER READING THE EMAIL

COEFFICIENT - 1.571615

02

TOTAL TIME SPENT ON WEBSITE

COEFFICIENT - 1.084354

03

TAGS - CLOSED BY HORIZZON

COEFFICIENT - 1.057205

Top Three Categorical Variables:

01

TAGS - RINGING

02

TAGS - WILL REVERT AFTER READING THE EMAIL

03

TAGS - CLOSED BY HORIZZON
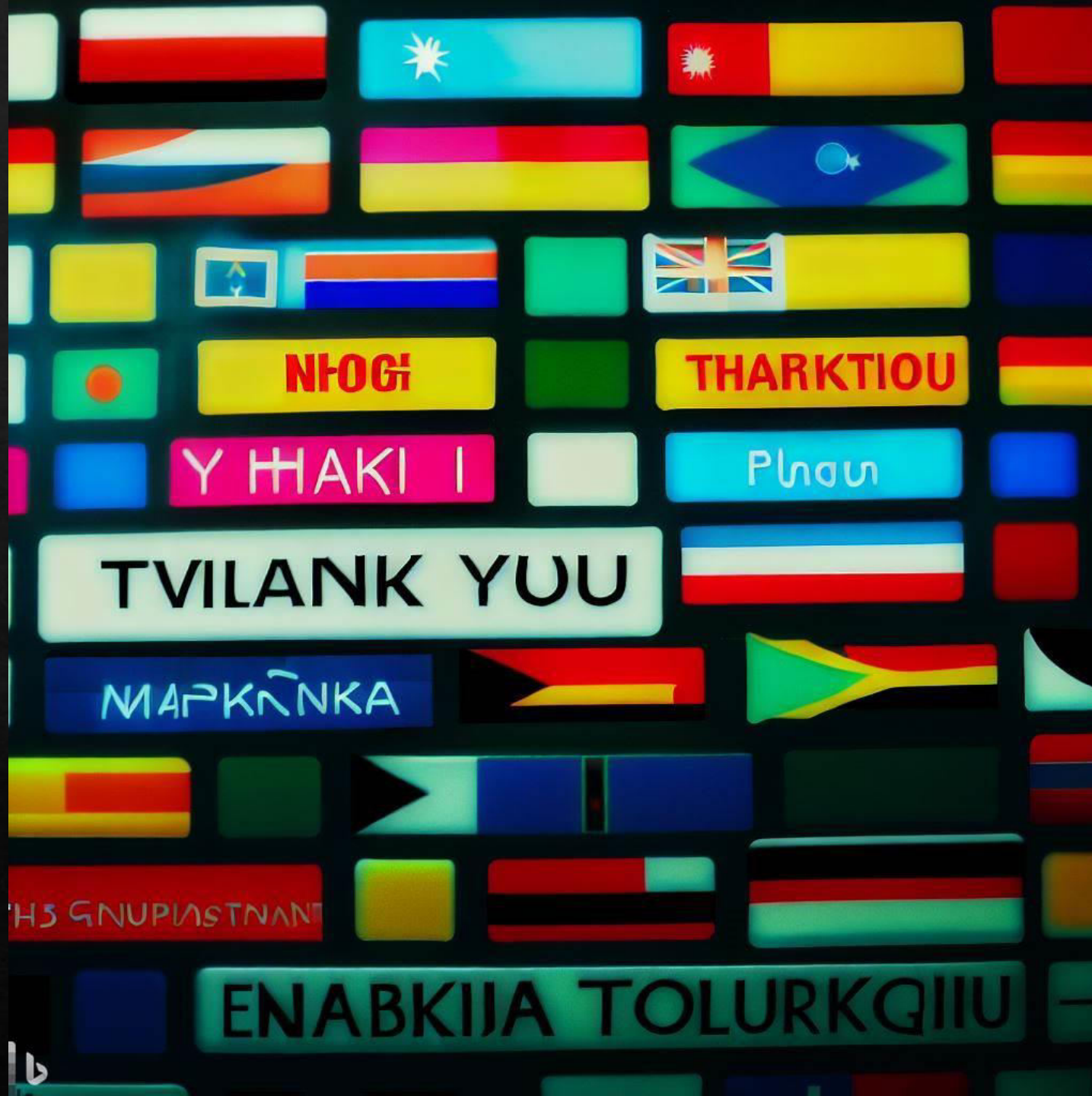
Receiver Operating Characteristic (ROC)

➤ ROC curve (area = 0.98)

➤ The model is demonstrating strong discrimination power.

➤ It's achieving a high True Positive Rate (sensitivity/recall) while keeping the False Positive Rate low.

➤ The model's predictions are well-separated, leading to minimal overlap between the two classes in terms of predicted probabilities
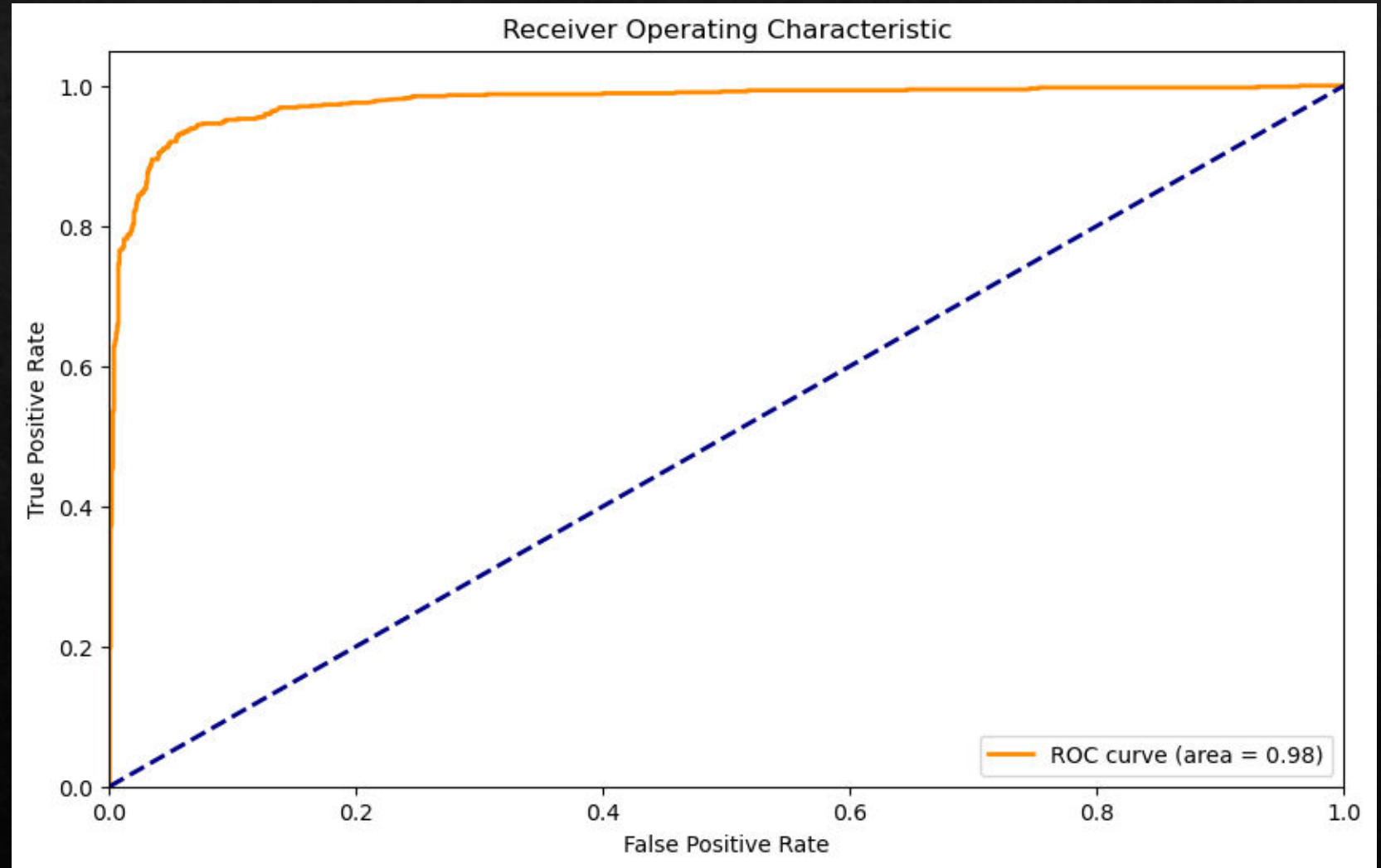
Output (snippet)

Prospect_ID	Lead_Score	Converted
4608	97.79624582	1
7935	0.678279578	0
4043	0.398070434	0
7821	0.22339053	0
856	2.522891727	0
927	22.09845072	0
318	99.22088393	1
1018	4.475900695	0
8151	0.712206052	0
1570	98.16123414	1
8086	99.462831	1
7689	97.96287725	1
5076	99.55825878	1
8752	27.41324594	0
2825	0.081860457	0
1840	99.80022809	1
6157	98.01345939	1
509	0.038404073	0
47	3.102671717	0
620	0.207957511	0
8030	97.39499307	1

Prospect_ID	Lead_Score	Converted
321	0.666229391	0
131	1.635762917	0
2485	99.80812828	1
5003	0.769902735	0
2880	2.082545917	0
3721	99.88872388	1
8606	0.159697135	0
3087	86.61689742	0
3383	49.77301651	0
6940	0.10892133	0
1393	96.98989186	1
80	2.535404072	0
8068	94.6734305	1
76	96.77268981	1
1688	0.827405163	0
5552	0.052332616	0
960	0.10892133	0
1369	32.36887066	1
4099	2.881123487	0
8435	33.8883259	1
119	0.033993343	0