

A dark blue background featuring a complex, glowing network graph composed of numerous small, semi-transparent blue spheres connected by thin white lines, creating a sense of depth and connectivity.

## LEAD SCORE CASE STUDY

# Group Members

- Kavita Verma
- Abhishek Viswakarma
- Vishal Varma

# Contents

- ❖ Problem Statement
- ❖ Business Objective
- ❖ Problem Approach
- ❖ EDA
- ❖ Model Evaluation
- ❖ Conclusion

## Problem Statement

- ❖ X Education, an educational organization, provides online classes designed specifically for working professionals. Every day, many professionals visit their website, searching for courses, and go through a form-filling process. After completing the form, the company classifies these individuals as potential leads.
- ❖ After getting these leads, the sales team reaches out through calls, emails, and other ways. Some leads get converted during this outreach, but most don't go any further.
- ❖ The typical lead conversion rate at X education is around 30%. This means that if they acquire 100 leads in a day, only about 30 of them are converted. To make this process more efficient, the company wants to identify the most promising leads, also known as Hot Leads.
- ❖ Finding these leads could improve the conversion rate. The sales team can now focus on communicating with potential customers instead of calling everyone.

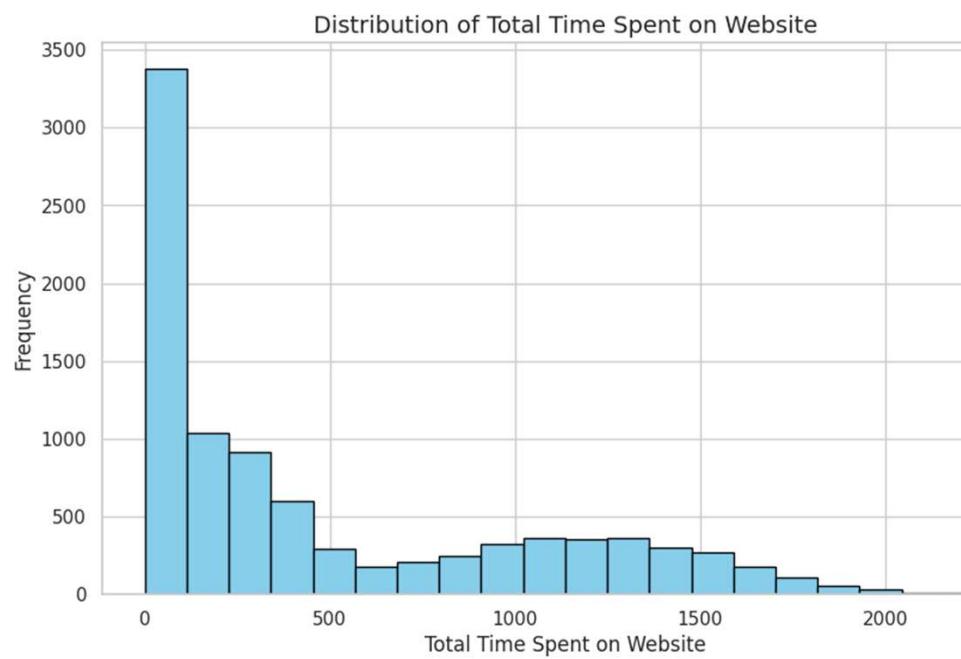
## Business Objective

- ❖ Lead X has requested us to develop a model capable of assigning a lead score ranging from 0 to 100 to each lead. This will enable them to distinguish hot leads and enhance their conversion rates accordingly.
- ❖ The CEO aims to attain a lead conversion rate of 80%.
- ❖ They require the model to accommodate future constraints such as peak time actions, optimizing manpower allocation, and post-achievement strategies once the target is reached.

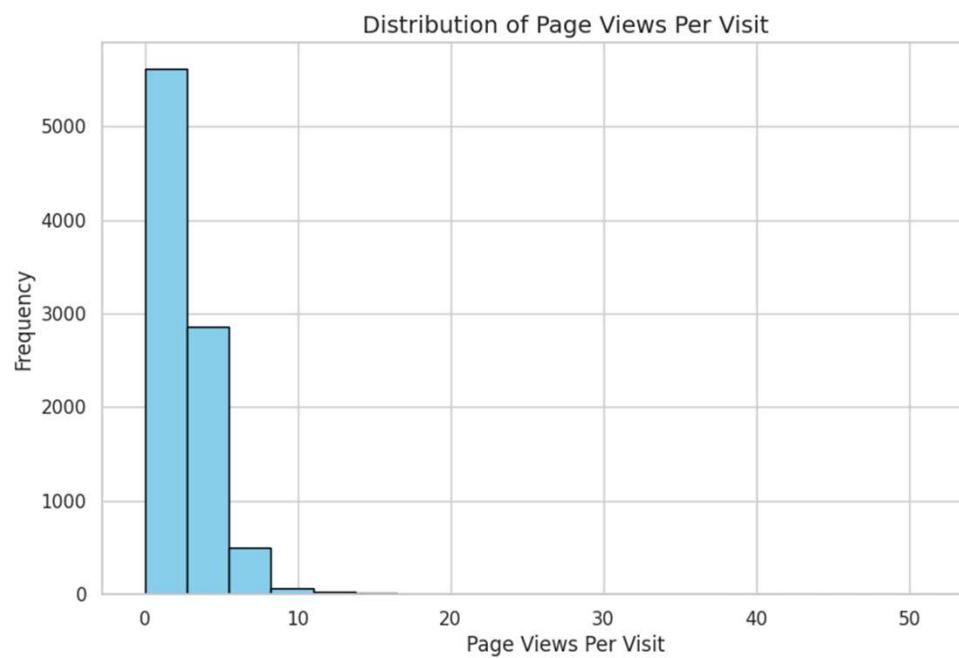
## Problem Approach

- ❖ Import leads data set and analyze
- ❖ Data cleaning and preparation
- ❖ EDA
- ❖ Categorical variable analysis
- ❖ Dummy variable creation
- ❖ Test Train split
- ❖ Scaling
- ❖ Model Building
- ❖ Model Evaluation

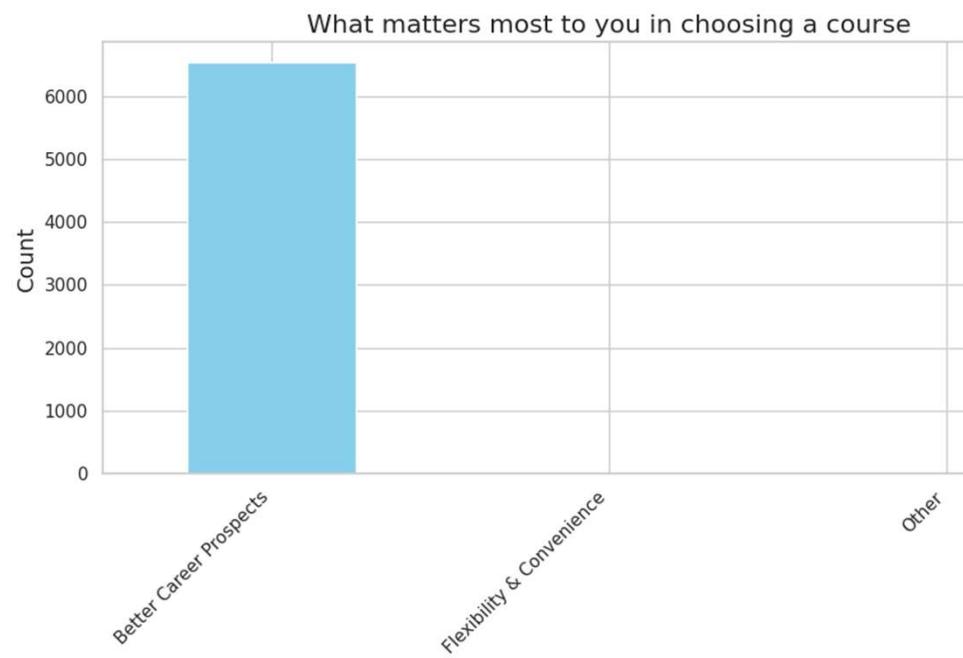
# EDA



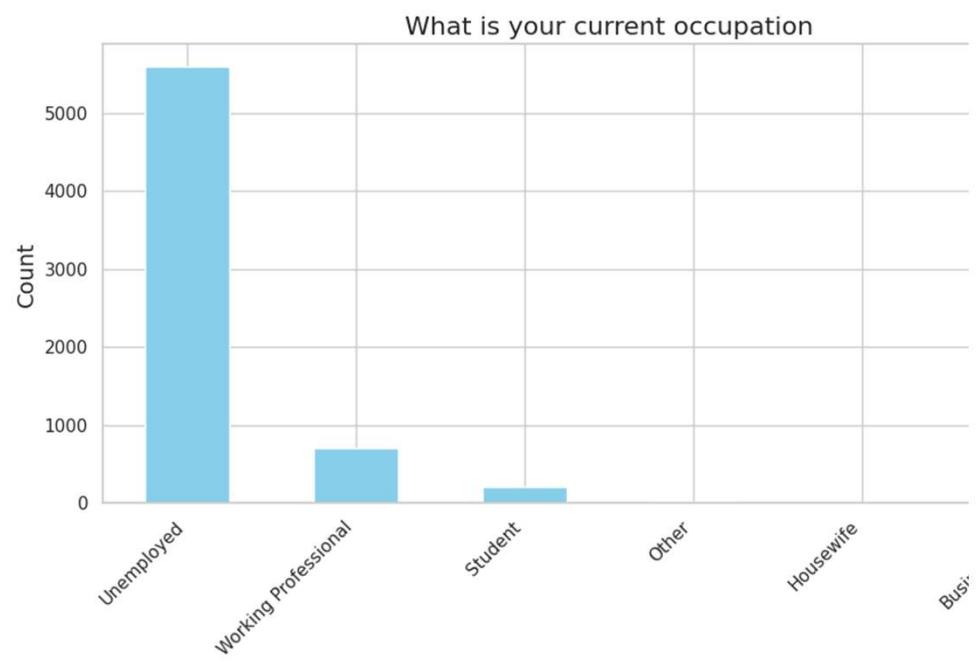
# EDA



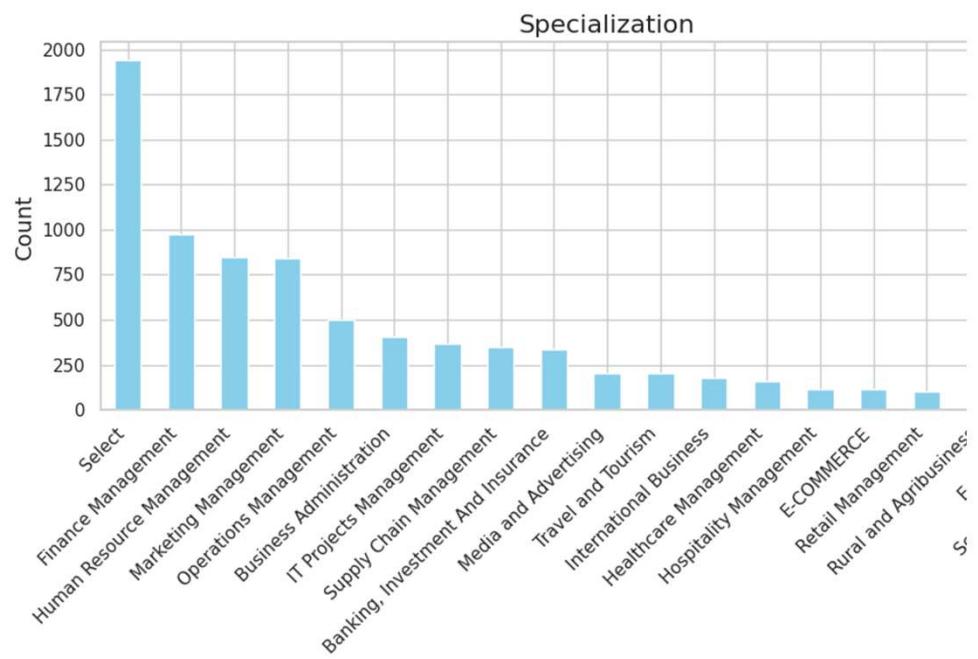
# EDA



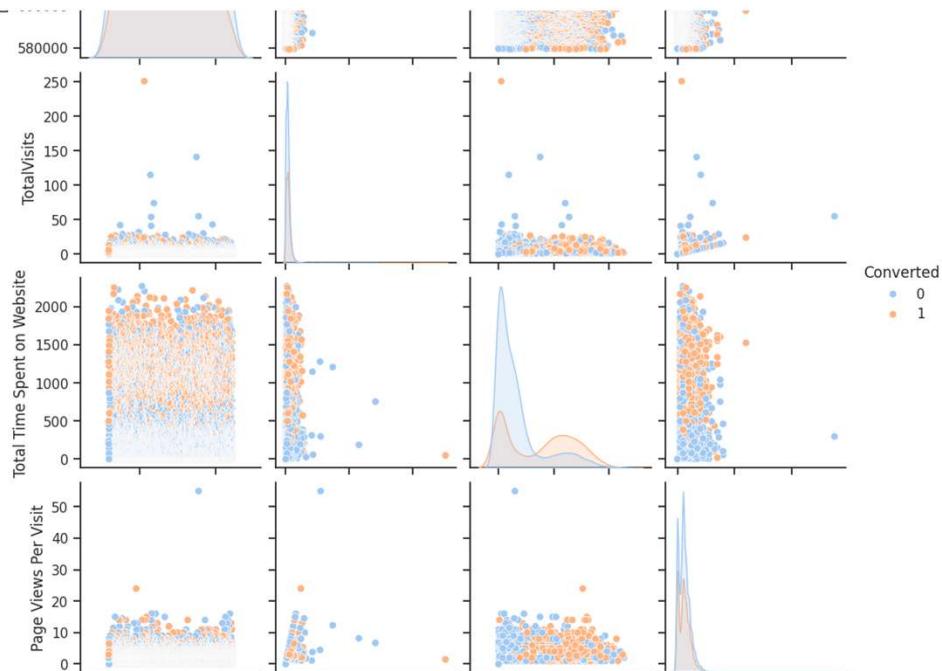
# EDA



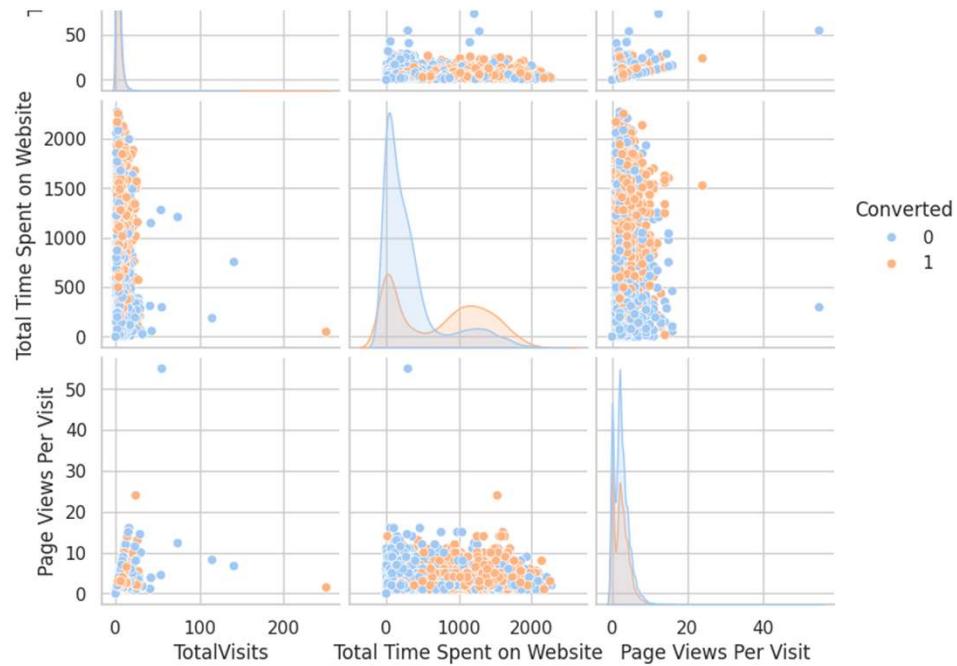
# EDA



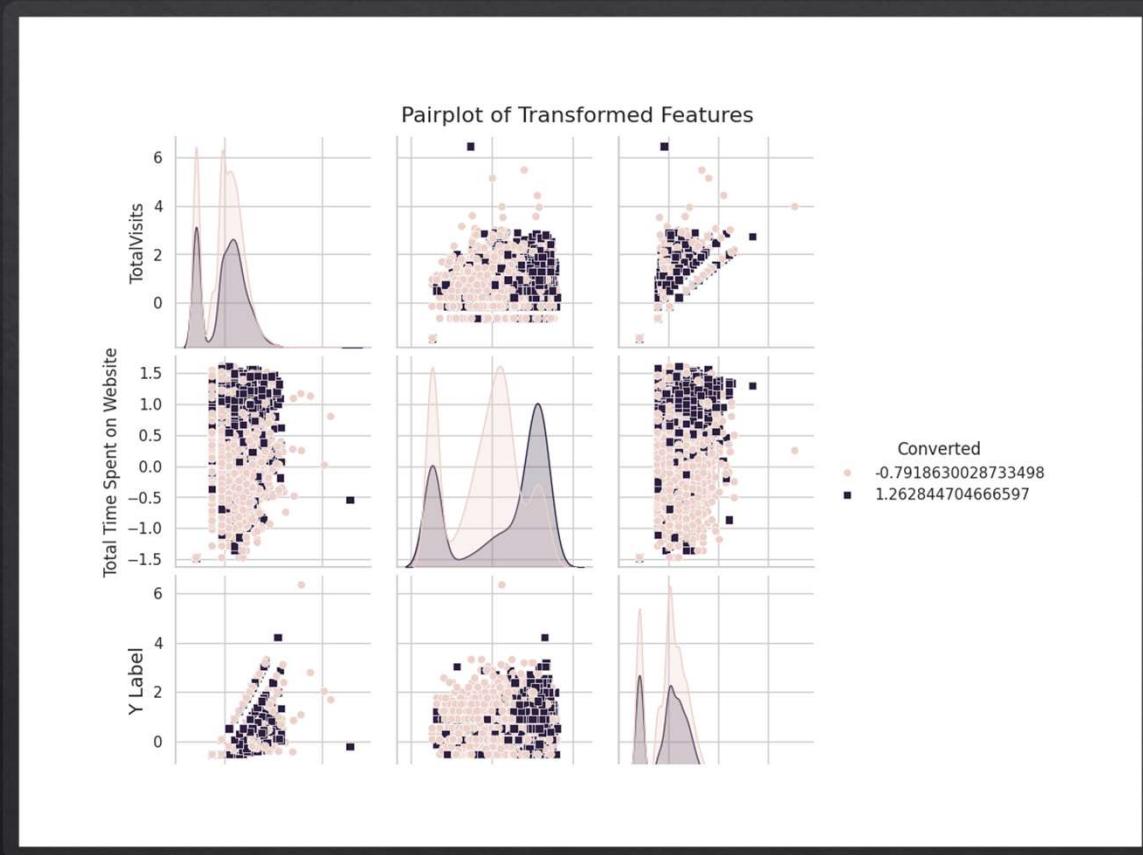
# EDA



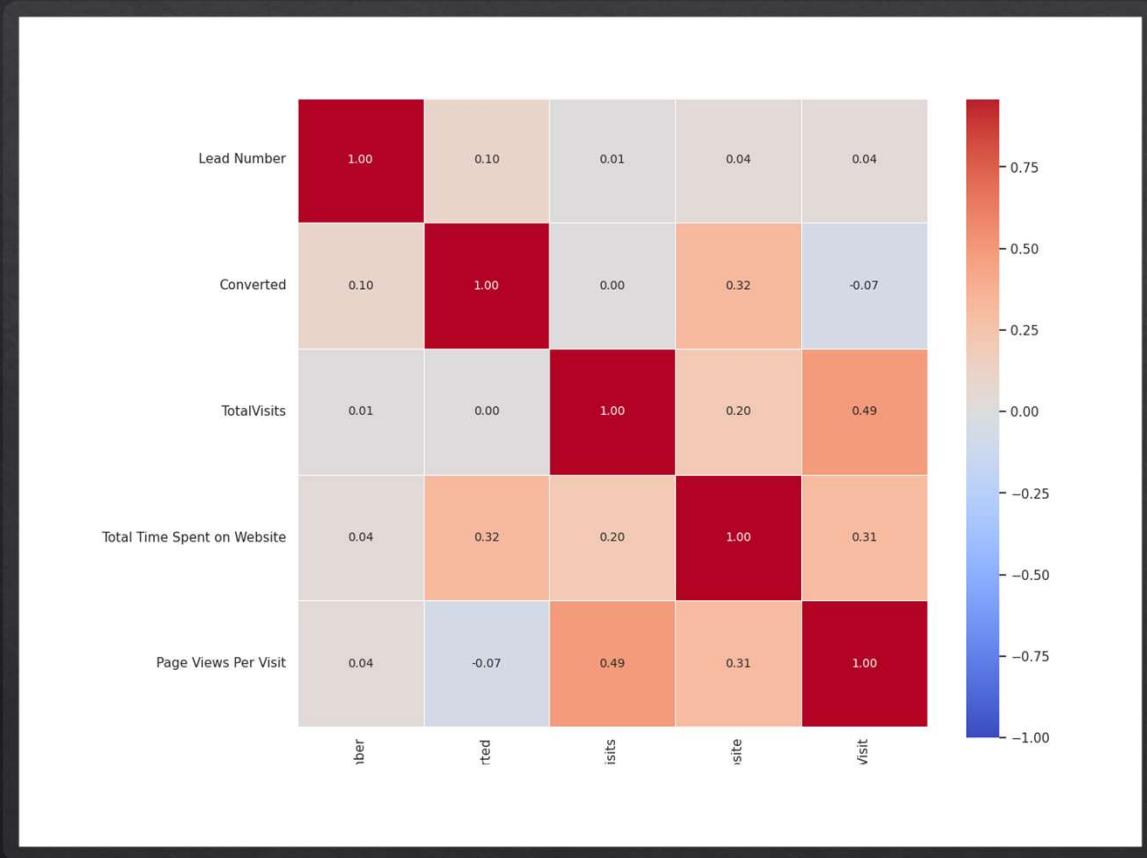
# EDA



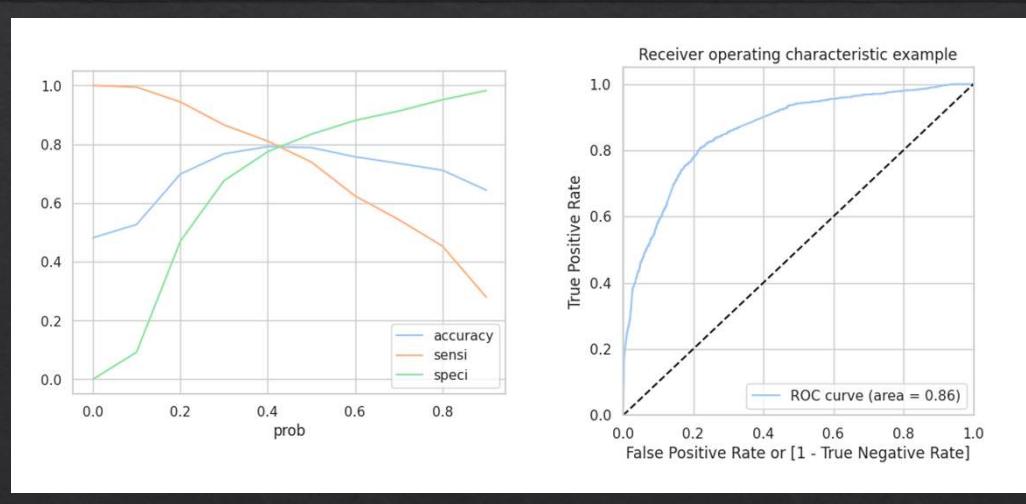
# EDA



# EDA



# Model Evaluation



With a tradeoff of 0.42 between precision and recall, it is deemed acceptable to classify any prospect lead with a conversion probability exceeding 42% as a hot lead.

## Conclusion

- ❖ Observing a conversion rate of 30-35% (close to average) for API and Landing Page submissions, contrasted with significantly lower rates for Lead Add forms and Lead imports, it is evident that greater emphasis should be placed on leads originating from API and Landing Page submissions.
- ❖ The most common last activity is email opened, while the highest conversion rate is associated with SMS sent. The majority of leads are unemployed, yet the maximum conversions are observed among working professionals.
- ❖ Leads who dedicate more time to browsing the website demonstrate a greater propensity for conversion.
- ❖ The maximum number of leads are generated by Google/direct traffic, while the highest conversion ratio is observed from references and the Welingak website.