

PROJECT

Problem statement



Data extraction



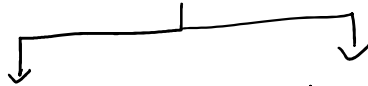
EDA



Preprocessing



what if unlabelled data?



clustering

Supportive column



evaluation ⇒

Problem statement:

We were trying to build a model that will find out if customers are happy or not with the insurance product. If not then what are the reasons behind it.

Project → combination of 2 things

- ① Sentiment Analysis
- ② Root cause Analysis

Analysis

we are having client with chain of retailers (ex: lifestyle, Shopper stops). They observed that number of customers are visiting their store but very less number of people buying products from them.

Reasons : Rates \uparrow

Bad customer service

we had a customer they were famous for their main course initially. Recently they added snacks & desserts so they found out that sales of main course has been drastically decreased."

Domain \rightarrow Banking \rightarrow ATM services
internet services

Retail industry

Human Resources domain

Hospitality domain

Health domain

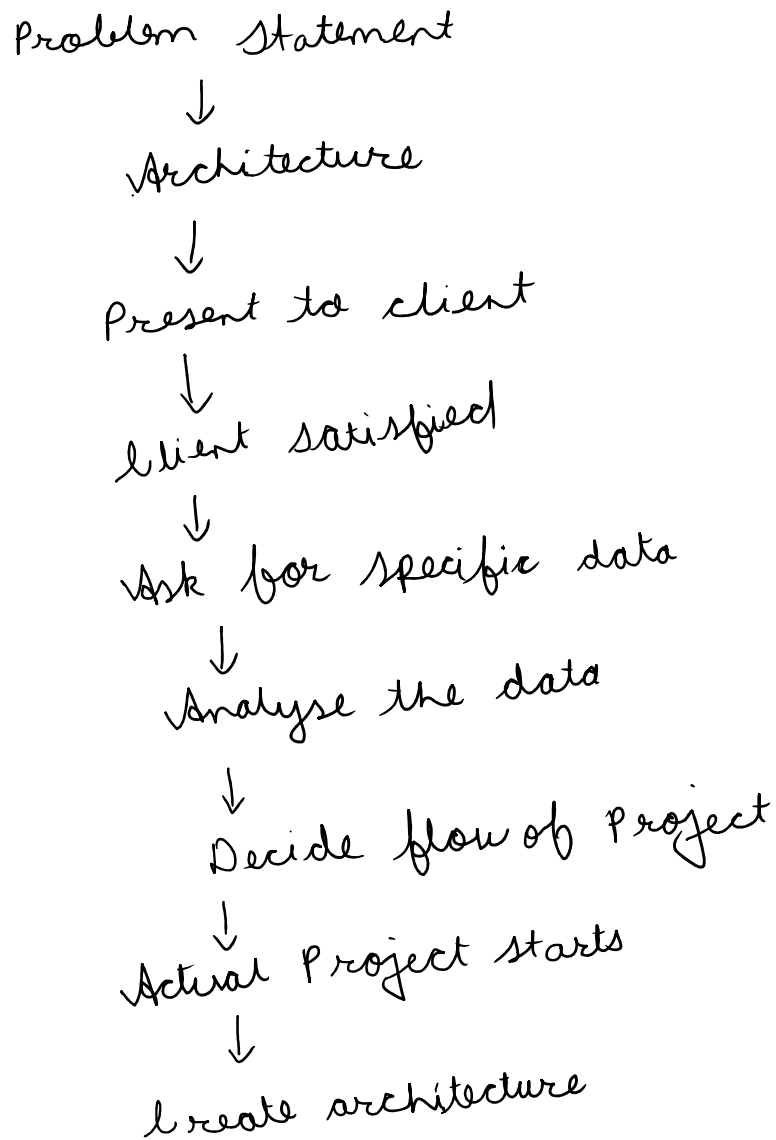
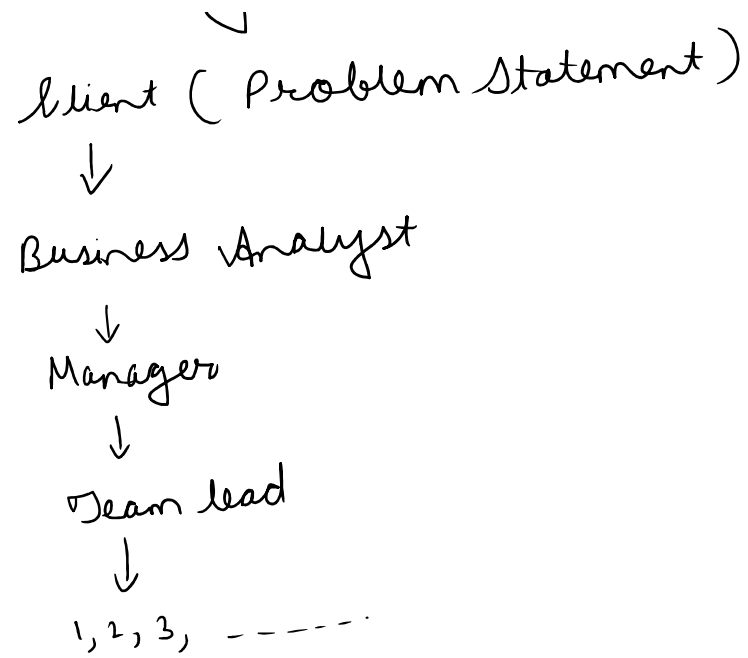
Supply chain management

Telecom domain

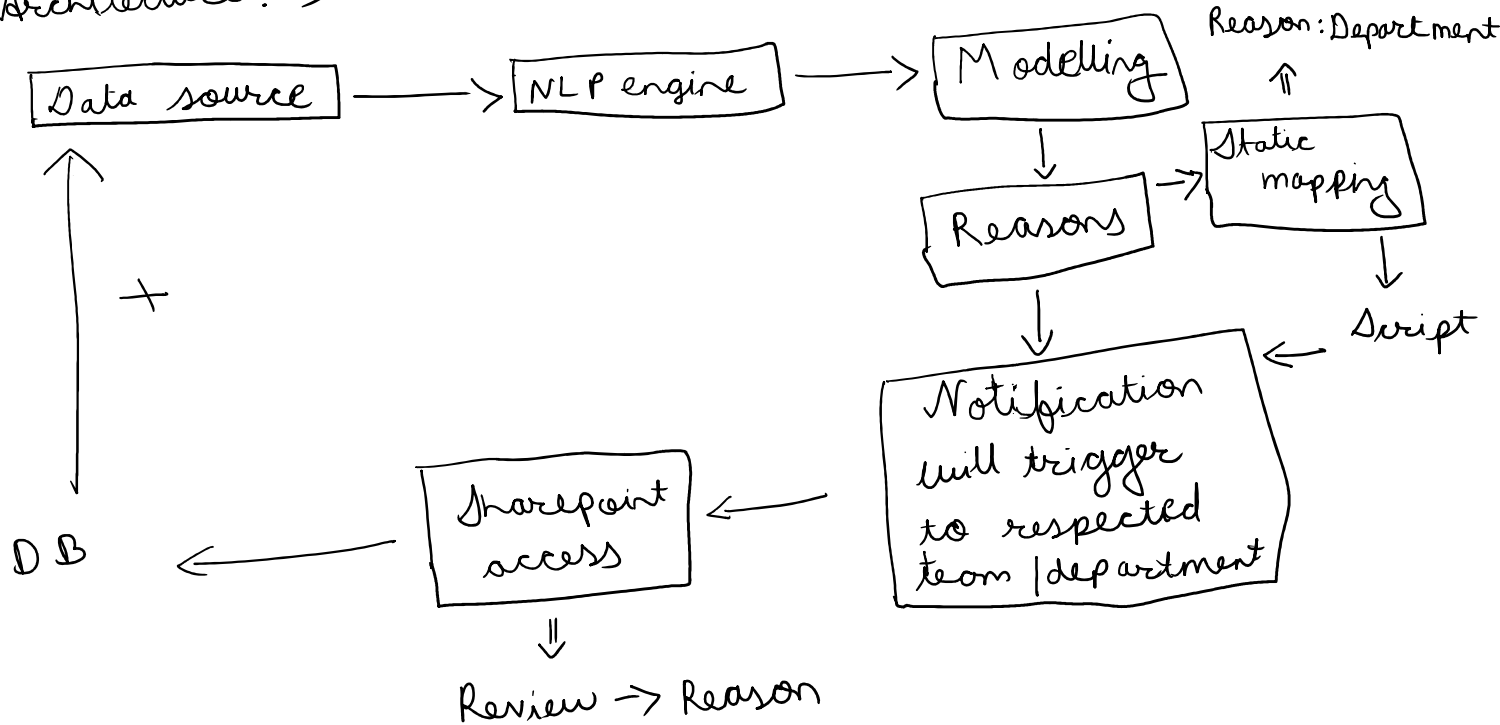
Types of Projects

- ① In house project
- ② Client based Project

Client (Problem Statement)



Architecture: →



Data:

- ① Quality
- ② Quantity
- ③ No continuous flow of data
- ④ exact data for our use case is not available

⇓
Public data / options

⇓
Feedback form / Google forms / surveys

Voice calls

Q What problems you faced while extracting data?

Private data / Secure data

NDA → Non Disclosure agreement

Q What is the difference between structured data & unstructured data?

Q Quantity of data

Around 1 GB

Around 10 lac rows

Data cycles ⇒ Monthly
yearly
quarterly

⇓
Data

⇓
Initial Preprocessing

⇓
Language detection

Language translation

⇓

EDA

→ Ngrams

Word cloud

Keywords extraction