

Technical Architecture :

Project Design Phase-II
Technology Stack
(Architecture & Stack)

Date	28 October 2023
Team ID	NM2023TMID03093
Project Name	Quantitative Analysis of Candidates in 2019 Lok Sabha Election

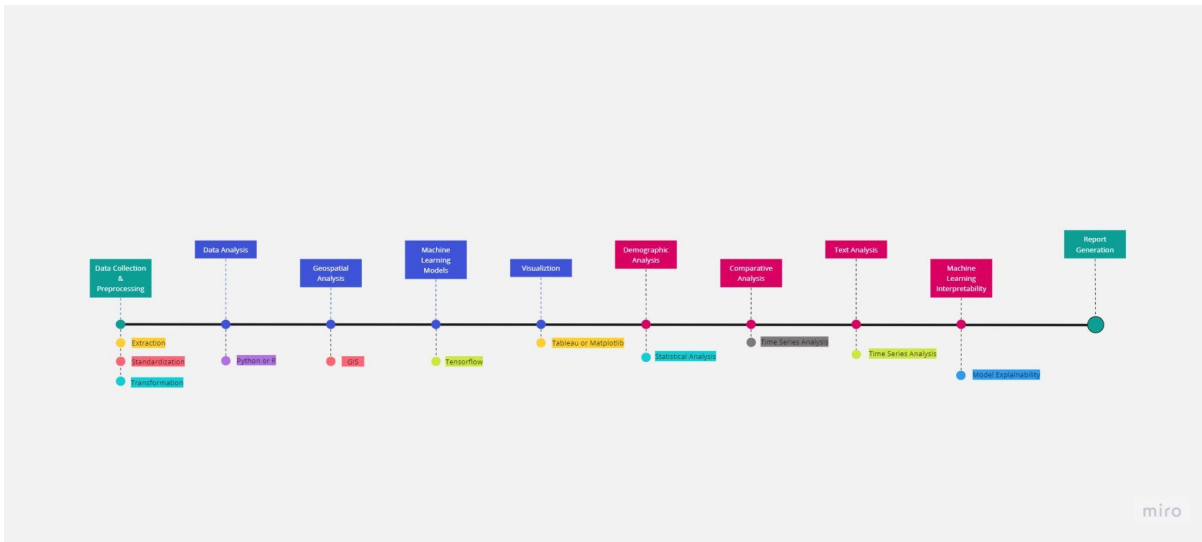


Table 1 : Components and its Description and Technologies

S.NO	COMPONENTS	DESCRIPTION
1	Data Collection	Gathering election data, candidate profiles, and voting results from various sources, including government records, election commissions, and news outlets. Web scraping tools, data extraction software, and APIs to access official election data.

2	Data Cleaning and Preprocessing	Refining the collected data to remove duplicates, handle missing values, and standardise formats for consistency. Data cleaning libraries in Python (e.g., pandas), and scripting languages for data transformation.
3	Data Analysis	Examining the data to identify trends, patterns, and insights, such as seat distribution, vote share, and regional variations. Data analysis tools like R, Python (e.g., with libraries such as Matplotlib and Seaborn), and statistical software for advanced analysis.
4	Geospatial Analysis	Analysing the geographic distribution of seats and vote shares to understand regional political dynamics. Geographic Information Systems (GIS) tools, geospatial libraries like GeoPandas, and mapping software.
5	Machine Learning Models	Building predictive models to understand the factors that influenced election outcomes and to forecast results. Python with machine learning libraries (e.g., scikit-learn), and deep learning frameworks like TensorFlow for more complex analyses.
6	Visualization	Creating informative charts, graphs, and maps to visually represent the election data, making it more accessible to a wider audience. Data visualization tools like Tableau, Matplotlib, and D3.js for interactive visualizations.
7	Demographic Analysis	Exploring how demographic factors such as caste, religion, and income levels correlate with voting patterns. Statistical analysis tools and demographic data sources, including census data.

8	Comparative Analysis	Comparing the 2019 election data with previous elections to identify trends and shifts in political dynamics. Data differencing techniques, time-series analysis, and data visualization for comparisons.
9	Text Analysis	Analyzing news articles, social media sentiment, and political speeches to understand how key issues influenced the election. Natural Language Processing (NLP) tools, sentiment analysis libraries, and text mining techniques.
10	Machine Learning Interpretability	Examining machine learning models to interpret the factors that played a crucial role in election outcomes. Tools for feature importance analysis and model explainability (e.g., SHAP values, LIME).
11	Report Generation	Compiling findings and insights into a comprehensive report or presentation for decision-makers, researchers, or the public. Document preparation software like Microsoft Word or LaTeX, and presentation software like PowerPoint or LaTeX Beamer.