AI-DRIVEN EXPLORATION AND PREDICTION OF COMPANY REGISTRATION TRENDS WITH REGISTRAR OF COMPANIES (ROC)

Due Date : 22/09/2023
Completion Date: 20/10/2023

Project Title: ROC Company Analysis

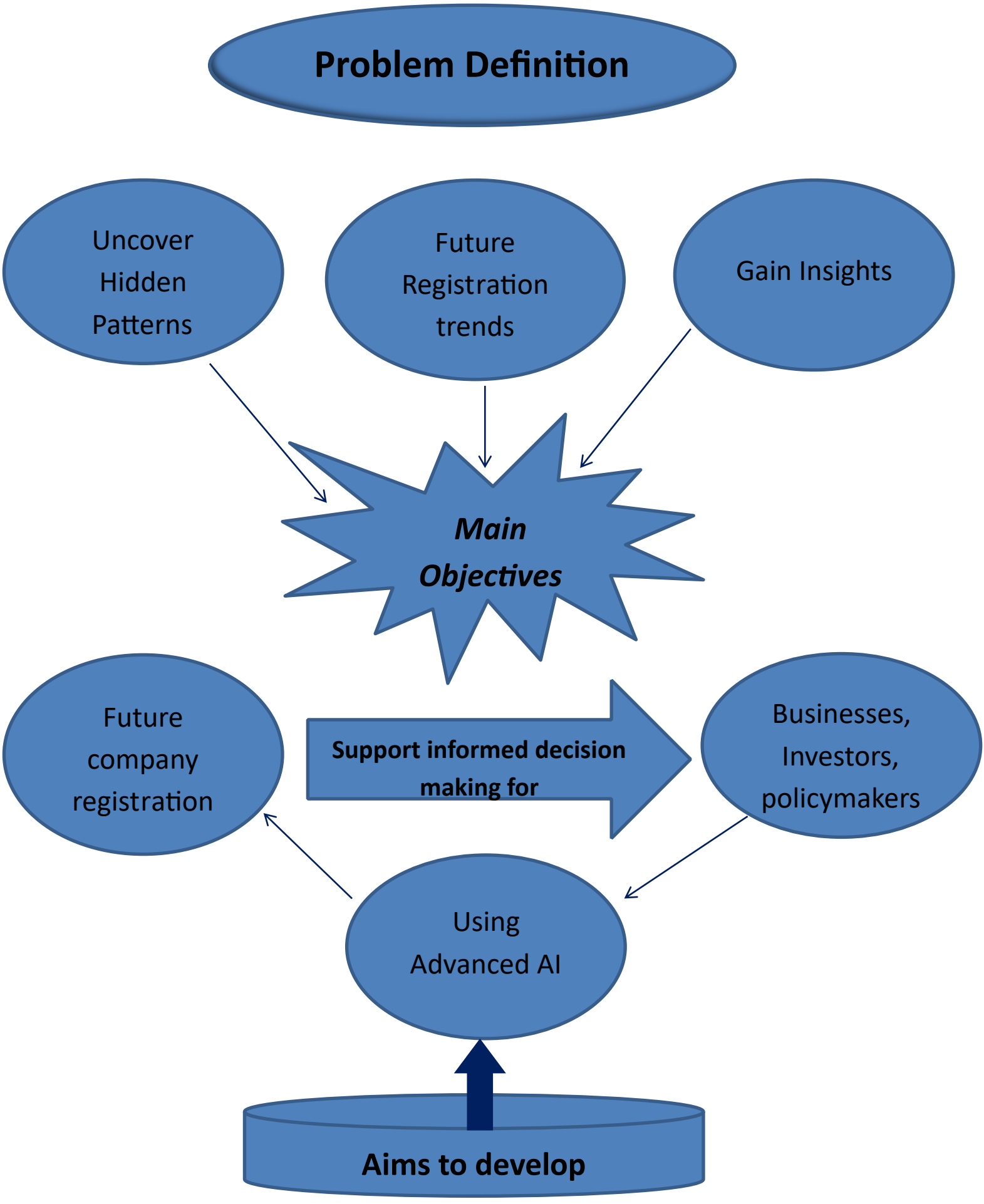
Project Summary:

In this document, we can clearly see about above project in this document. And discuss about problem definition, scope of project, vision of project, phase of projects, detail about team members, what will we do, how it will be complete, requirements, problem analysis and conclusion.

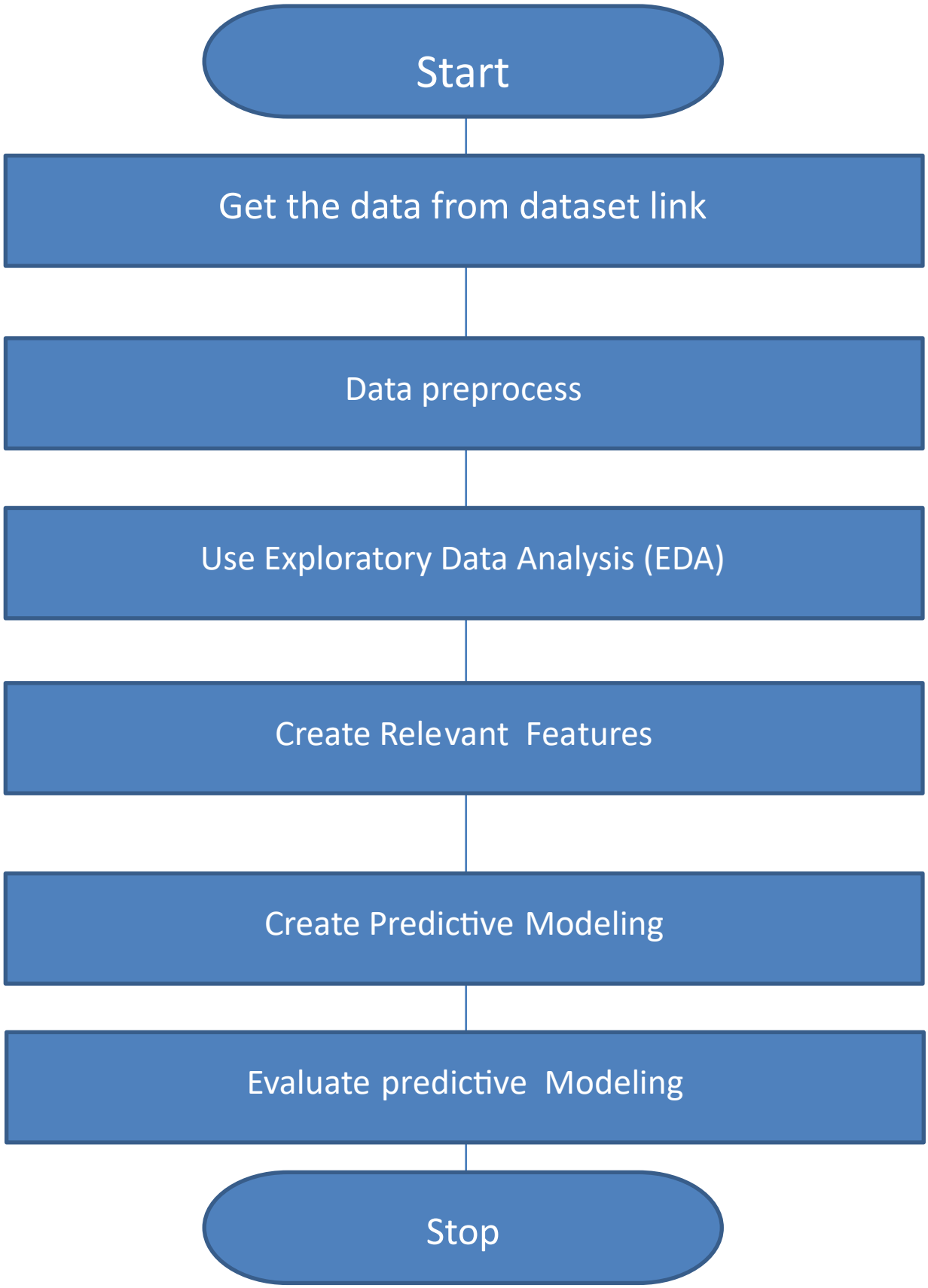
The AI-driven analysis aims to uncover hidden patterns, discover valuable insights into the company landscape, and forecast future registration trends. By applying cutting-edge AI algorithms, the study seeks to identify unique characteristics and relationships among registered companies, enabling a more sophisticated understanding of the business ecosystem in Tamil Nadu.

Development Platform: Google Colab Jupyter notebook

Dataset Link: https://tn.data.gov.in/resource/company-masterdata-tamil-nadu-upto-28th-february-2019


Problem Definition
Uncover Hidden Patterns
Future Registration trends
Gain Insights
Main Objectives
Future company registration
Support informed decision making for
Businesses, Investors, policymakers
Using Advanced AI
Aims to develop

Design Thinking:



Missing Value Analysis:

Next, i will cleanup the dataset which is the important part of data science. Missing data can lead to wrong statistics during Modeling and predictions.

Feature Engineering:

Till now, i explored the dataset, did missing value corrections and data visualization. Next, i have started feature engineering. Feature engineering is useful to improve the performance of machine learning algorithms and is often considered as applied machine learning.

Outlier Detection:

In this part i removed all the records outlined in dataset. Outliers impacts Model accuracy. I used *Tukey Method* used for outlier detection.

Modeling:

In this sections, i tried different models and compare the accuracy for each. Then, i performed Hyperparameter Tuning on Models that has high accuracy.

Before i split the dataset i need to transform the data into quantile using sklearn Preprocess.

Data Splitting:

Next, i split data in test and train dataset. Train dataset will be used in Model training and evaluation and test dataset will be used in prediction. Before i predict the test data, i performed cross validation for various models.

Prediction:

Till now, i worked on EDA, Feature Engineering, Cross Validation of Models, and Hyperparameter Tuning and find the best working Model for my dataset. Next, I did prediction from my test dataset and storing the result in CSV.

Conclusion:

Certainly, here concluding the project documentation for AIdriven exploration and prediction of company registration trends with registrar of companies, we summarize the key finding , insights and implication derived from our AI-driven exploration and prediction trends are achieved their own and future prediction will executed successful when above all done.