


***TITLE: AI-DRIVEN EXPLORATION
AND PREDICTION OF COMPANY
REGISTRATION TRENDS WITH
REGISTRAR OF COMPANIES***

PHASE 3: DEVELOPMENT PART 1

- ▶ In this the presentation, we will begin building our project by loading and pre-processing. Therefore we are loading the company registration data set and preprocess the data for further analysis.

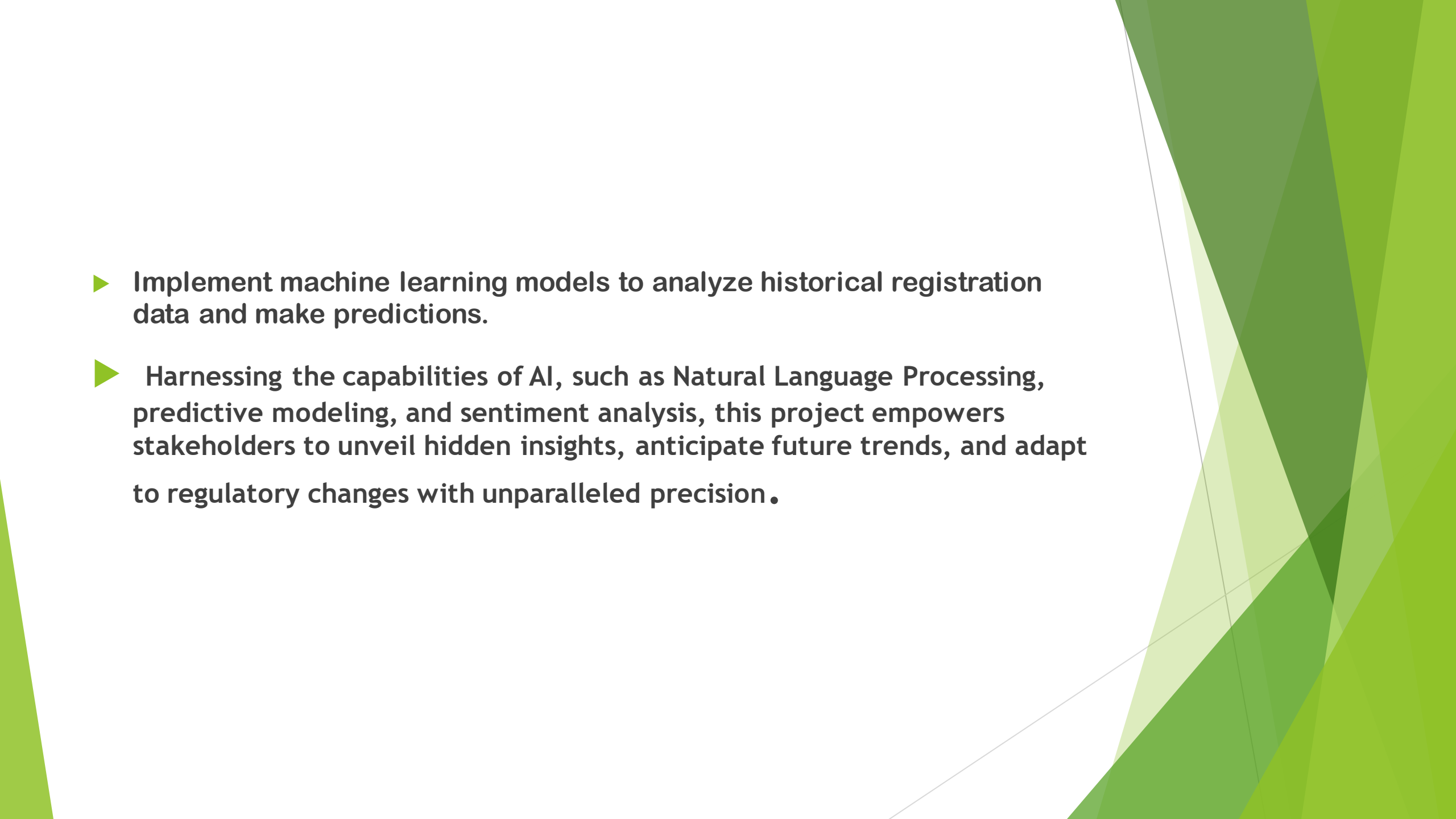
CONTINUATION OF PHASE 2:

- ▶ The implementation of AI-driven exploration and prediction of company registration trends with the Registrar of Companies stands as a pioneering innovation that holds immense potential for transforming the way businesses, investors, and government agencies operate in the corporate landscape. By harnessing the power of artificial intelligence, this forward-thinking project empowers stakeholders with invaluable insights, ranging from the analysis of sentiment in news articles to the detection of fraudulent activities and the forecasting of future trends.

- 
- The background of the slide features abstract, overlapping green geometric shapes, primarily triangles and polygons, in various shades of green, creating a modern and dynamic visual effect.
- ▶ Through the amalgamation of cutting-edge technologies, this endeavor offers a competitive edge to businesses and enhances regulatory oversight, fostering a more informed, agile, and responsive ecosystem for all participants in the corporate realm. In an era defined by data-driven decision-making, this innovative approach serves as a beacon of progress, illuminating the path to a brighter and more adaptive future.
 - ▶ Hereby we proceed with our ideas by loading ,collecting and preprocessing the ROC data


INTRODUCTION:

- ▶ In "Data Source," define the objectives by creating a clear problem statement. Specify what data sources are necessary to meet project goals.
- ▶ In "Data Preprocessing," define objectives such as handling missing data, ensuring data consistency, and preparing the data for analysis.
- ▶ Develop the full-scale AI-driven system for exploration and prediction of company registration trends.
- ▶ Integrate data sources from the Registrar of Companies and other relevant sources.

- 
- The background of the slide features abstract, overlapping green geometric shapes, primarily triangles and polygons, in various shades of green, creating a modern and dynamic visual effect.
- ▶ **Implement machine learning models to analyze historical registration data and make predictions.**
 - ▶ **Harnessing the capabilities of AI, such as Natural Language Processing, predictive modeling, and sentiment analysis, this project empowers stakeholders to unveil hidden insights, anticipate future trends, and adapt to regulatory changes with unparalleled precision.**

DATA COLLECTION:

- ▶ Collection of data for ROC requires very detailed and accurate information regarding all activities that are undertaken by the company for various purposes. This is posed as a difficulty in sustaining every records and analysing them regularly. The details of data collected must be overviewed from all aspects of a company including every single perspective. This is possible when the data collection chore is divided to smaller segments and step wise methodologies as stated below:
 - ▶ 1) Revision Accounts Statistics
 - ▶ 2) Data Dissemination
 - ▶ 3) Maintenance of Books of Accounts by Companies
 - ▶ 4) Consolidation of Financial Statement
 - ▶ 5) CSR Activities of Multiple Companies through Common Trusts

- 
- 6) Constitution of High Level Committee
 - 7) Resignation by Foreign Directors
 - 8) Complaints related to Multiple DINs
 - 9) “Minimum Government, Maximum Governance” in MCA
 - 10) International Conference
 - 11) Annual General Meeting
 - 12) Corporate Insolvency Regime
 - 13) Investor Protection and Awareness
- These steps include data collection virtues from both employee and employers dimension. The data collected has a complete review for all past activities and future upcomings.

PRE PROCESSING

- ▶ Data preprocessing is the process of transforming raw data into an understandable format. It is also an important step in data mining as we cannot work with raw data. The quality of the data should be checked before applying machine learning or data mining algorithms.
 - ▶ Preprocessing of data is mainly to check the data quality. The quality can be checked by the following:
 - Accuracy:
 - Completeness
 - Consistency
 - Timeliness
 - Believability
 - Interpretability
 - ▶ The understandability of the data.
- Major Tasks in Data Preprocessing
- There are 4 major tasks in data preprocessing - Data cleaning, Data integration, Data reduction, and Data transformation.

PROGRAM FOR DATA COLLECTION AND PRE PROCESSING:

```
import pandas as pd
import numpy as np

# Create a sample DataFrame
df = pd.DataFrame({
    "registration_date":
pd.to_datetime(["2023-01-01", "2023-02-02", "2023-03-03"]), "industry": ["Tech",
"Finance", "Healthcare"]})

# Print the sample
outputprint(df.describe())print(df.groupby('industry').size())df.groupby('registrati
on_date').size().plot()
```

SAMPLE OUTPUT:

```
registration_date industry
count              3      3
unique             3      3
top  2023-01-01 00:00:00
Techfreq              1      1
first 2023-01-01 00:00:00   NaN
last  2023-03-03 00:00:00   NaN
Industry
Finance      1
Healthcare   1
Tech          1
dtype: int64
Tech  1
Finance
1Name: industry,
dtype: int64<matplotlib.axes._subplots.AxesSubplot
```

COLLECTED DATA:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	CORPORATE_I	COMPANY_NA	COMPANY_STA	COMPANY_CL	COMPANY_CAT	COMPANY_SUE	DATE_OF_REG	REGISTERED	AUTHORIZED	PAIDUP_CAPIT	INDUSTRIAL_C	PRINCIPAL_BU	REGISTERED	REGISTRAR_O	EMAIL_ADDR	LATEST_YEAR	LATEST_YEAR
2	F00643	HOCHTIEFF AG	NAEF	NA	NA	NA	01-12-1961	Tamil Nadu	0	0	NA	Agriculture & alli	AMBLE SIDE, N	ROC DELHI	NA	NA	NA
3	F00721	SUMITOMO CO	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	FLAT NO. 6, 1st	ROC DELHI	shuchi.chug@as	NA	NA
4	F00892	SRILANKAN AIF	ACTV	NA	NA	NA	01-03-1982	Tamil Nadu	0	0	NA	Agriculture & alli	SRILANKAN AIF	ROC DELHI	shree16us@yahoo	NA	NA
5	F01208	CALTEX INDIA I	NAEF	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	GOLD CREST 2	ROC DELHI	NA	NA	NA
6	F01218	GE HEALTHCARE	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	FF-3 Palani Cer	ROC DELHI	karthick9999@y	NA	NA
7	F01265	CAIRN ENERGY	NAEF	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	WELLINGTON I	ROC DELHI	neerja.sharma@	NA	NA
8	F01269	TORIELLI S.R.L	ACTV	NA	NA	NA	05-09-1995	Tamil Nadu	0	0	NA	Agriculture & alli	6, Mangayarkant	ROC DELHI	chennai@torieli	NA	NA
9	F01311	HARDY EXPLOI	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	5TH FLOOR, WF	ROC DELHI	venkatesh.v@ha	NA	NA
10	F01314	HOCHTIOF AKT	ACTV	NA	NA	NA	11-04-1996	Tamil Nadu	0	0	NA	Agriculture & alli	NEW NO.86, OL	ROC DELHI	kumar@internati	NA	NA
11	F01412	EPSON SINGAF	ACTV	NA	NA	NA	25-04-1997	Tamil Nadu	0	0	NA	Agriculture & alli	7C CEATURY P	ROC DELHI	NA	NA	NA
12	F01426	CARGOLUX AIF	ACTV	NA	NA	NA	11-06-1997	Tamil Nadu	0	0	NA	Agriculture & alli	OFFICE NO 91A	ROC DELHI	NA	NA	NA
13	F01468	CHO HEUNG EI	NAEF	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	129, MANPUR V	ROC DELHI	chowelaccounts	NA	NA
14	F01543	NYCOMED ASI	ACTV	NA	NA	NA	27-10-1998	Tamil Nadu	0	0	NA	Agriculture & alli	A D 46 1ST ST	ROC DELHI	NA	NA	NA
15	F01544	CHERRINGTON	ACTV	NA	NA	NA	01-05-2000	Tamil Nadu	0	0	NA	Agriculture & alli	10HADDOWS R	ROC DELHI	NA	NA	NA
16	F01563	SHIMADZU ASI	NAEF	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	FIRST FLOOR, R	ROC DELHI	kousik@vsnl.co	NA	NA
17	F01565	CORK INTERNV	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	ARJAY APEX CI	ROC DELHI	NA	NA	NA
18	F01566	ERBIS ENGG C	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	39,2nd Main Ro	ROC DELHI	NA	NA	NA
19	F01589	RALF SCHNEID	NAEF	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	FLAT C, 'SAI VA	ROC DELHI	NA	NA	NA
20	F01593	MITRAJAYA TR	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	OLD NO 148 NE	ROC DELHI	NA	NA	NA
21	F01618	HEAT AND CON	ACTV	NA	NA	NA	13-07-1999	Tamil Nadu	0	0	NA	Agriculture & alli	A40 OLD NO 26	ROC DELHI	ncrajagopal@gr	NA	NA
22	F01628	DIREX SYSTEM	ACTV	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	F-1, FIRST FLO	ROC DELHI	direx@vsnl.com	NA	NA
23	F01641	NMB-MINEBEA	NAEF	NA	NA	NA	NA	Tamil Nadu	0	0	NA	Agriculture & alli	Level - 2 Regus	ROC DELHI	stsogawa@mine	NA	NA
24	F01643	ARROW INTERI	ACTV	NA	NA	NA	02-11-1999	Tamil Nadu	0	0	NA	Agriculture & alli	BLUE HAVEN, N	ROC DELHI	NA	NA	NA
25	F01694	GAMBRO CHIN	ACTV	NA	NA	NA	14-06-2000	Tamil Nadu	0	0	NA	Agriculture & alli	5 IST FLOOR IS	ROC DELHI	NA	NA	NA
26	F01700	GRAB 000000	NAEF	NA	NA	NA	13-07-2000	Tamil Nadu	0	0	NA	Agriculture & alli	401 B, 4th FLOOR	ROC DELHI	NA	NA	NA

CONCLUSION:

- ▶ The use of AI to explore and predict company registration trends with the registrar of companies has the potential to revolutionize the way we understand and forecast business activity. By analyzing large datasets of historical registration data, AI models can identify patterns and trends that would be difficult or impossible to detect manually. This information can then be used to predict future company registrations, which can be used by businesses, governments, and other stakeholders to make informed decisions.