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

**(RoC)**

**INTRODUCTION:**

In an era marked by relentless technological advancement, the Registrar ofCompanies (RoC) plays a pivotal role in regulating and overseeing the corporate landscape of any nation. The process of company registration is not merely a bureaucratic formality; it is a reflection of the economic pulse of a country. To harness the immense potential hidden within the vast troves of data maintained by the RoC, the advent of Artificial Intelligence (AI) promises a transformative journey into the realm of exploration and prediction of company registration trends.

AI-driven exploration and prediction of company registration trends entail the utilization of cutting-edge machine learning algorithms, data analysis techniques, and natural language processing to decipher the intricate patterns and dynamics within the RoC's extensive repository of corporate records. This innovative approach empowers businesses, policymakers, investors, and analysts to gain unprecedented insights into the evolving landscape of entrepreneurship, economic activity, and regulatory compliance.

**INTRODUCTION:**

**OBJECTIVES:**

Clearly define your objectives, such as understanding industry trends, forecasting demand for certain services, or identifying potential market opportunities.

**DATA COLLECTION:**

Gather historical company registration data from the RoC, which includes information about new company registrations, business categories, geographical locations, and other relevant variables. You may need to work with the RoC or other data providers to access this data.

**DATA PREPROCESSING:**

Prepare the data by cleaning, transforming, and structuring it for analysis. This might involve dealing with missing values, duplicates, and ensuring data consistency.

**FEATURE ENGINEERING:**

Create relevant features from the data, such as time-based variables (month, quarter, year), geographical regions, and industry categories. These features will be used as inputs to your AI model.

**AI MODEL SELECTION:**

Choose the appropriate AI models for your analysis. Depending on the complexity of the data and your objectives, this could include linear regression, decision trees, random forests, neural networks, or more advanced models like LSTM or Transformers.

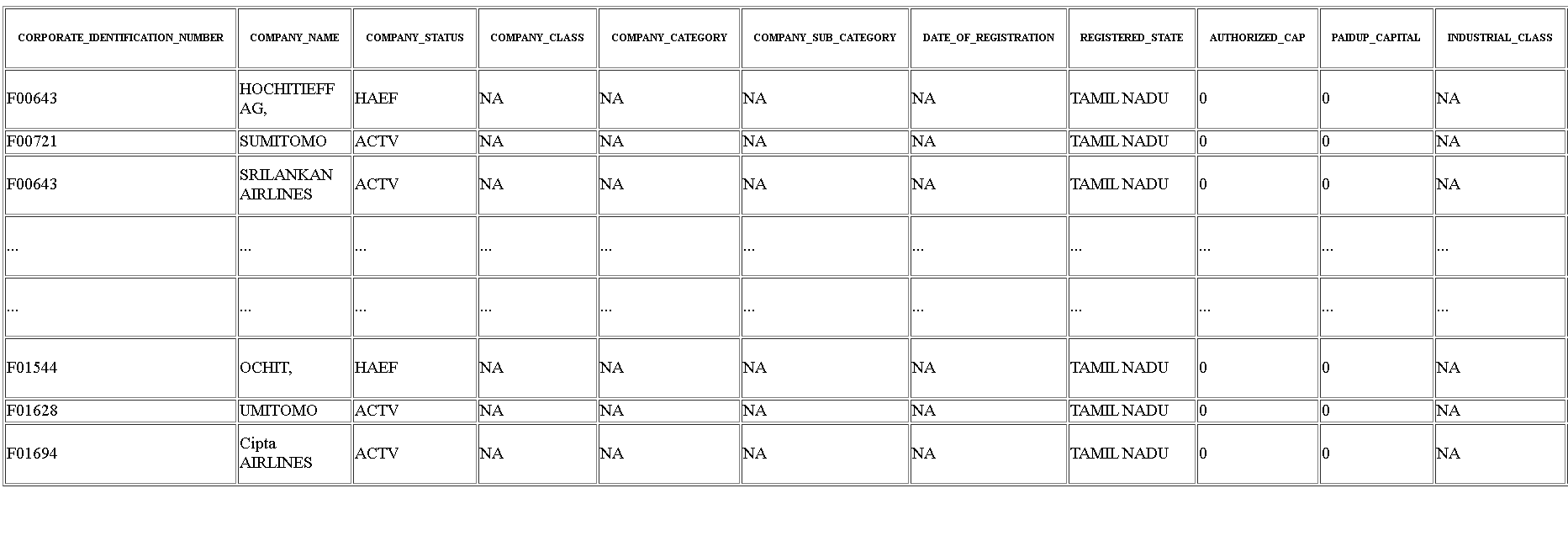
**VISUALIZATION:**

Create visualizations to communicate your findings effectively. Use charts, graphs, and dashboards to present the trends and forecasts.

**MODEL EVALUATION:**

Assess the performance of your AI model using appropriate metrics (e.g., Mean Absolute Error, Root Mean Squared Error) and ensure it meets your objectives.

**GIVEN DATASET:**

****

**STEP TO FOLLOW:**

**1.IMPORT THE LIBRARY:**

By Importing the Libraries like:

* Numpy
* Matplotlib
* Seaborn
* Pandas

**Program:**

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

**2.LOAD THE DATASET:**

Execute the necessary code or functions to read the dataset and load it into memory. This step makes the data available for analysis.

**Program:**

# Load your dataset

data = pd.read\_csv('roc.csv', encoding='latin1', low\_memory=False)

**3.EXPLORATORY DATA ANALYSIS(EDA):**

* Perform initial data exploration to understand the dataset's structure and content.
* Use visualization libraries like Matplotlib and Seaborn to create charts and graphs to identify trends and patterns.

**Program:**

RED = "\033[91m"

GREEN = "\033[92m"

YELLOW = "\033[93m"

BLUE = "\033[94m"

RESET = "\033[0m"

print(BLUE + "\nDATA CLEANING" + RESET)

missing\_values = df.isnull().sum()

print(GREEN + "Missing Values : " + RESET)

print(missing\_values)

**REMOVING DUPLICATE VALUE:**

To remove duplicate values from a list in a program, you can use various programming languages and approaches. Here's a general algorithm to remove duplicates from a list in Python.

**PROGRAM:**

mean\_fill = df.fillna(df.mean())

df.fillna(mean\_fill, inplace=True)

duplicate\_values = df.duplicated().sum()

print(GREEN + "Duplicate Values : " + RESET)

print(duplicate\_values)

df.drop\_duplicates(inplace=True)

**LOADING AND PREPROCESSING:**

**Loading:**

# Import necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

**Preprocessing:**

# Load your dataset

data = pd.read\_csv('roc.csv', encoding='latin1', low\_memory=False)

**Clean the Data:**

print(BLUE + "\nDATA CLEANING" + RESET)

**PROGRAM:**

# Import necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load your dataset

data = pd.read\_csv('roc.csv', encoding='latin1', low\_memory=False)

# Summary statistics

summary\_stats = data.describe()

# Data distribution (Histogram)

plt.figure(figsize=(10, 6))

sns.histplot(data['COMPANY\_NAME'], bins=20, kde=True)

plt.title('Distribution of Registration Dates')

plt.xlabel('Registration Date')

plt.ylabel('Frequency')

plt.show()

**DATA ANALYSIS:**

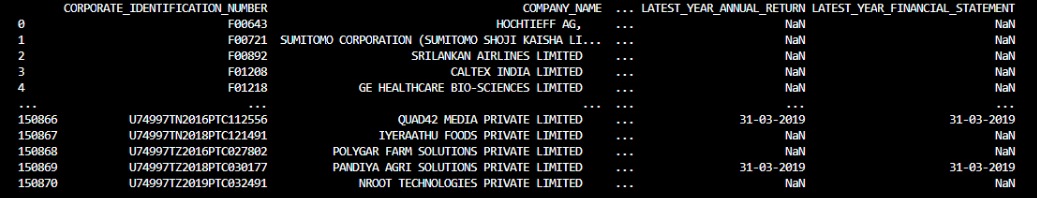
print(BLUE + "\nDATA ANALYSIS" + RESET)

summary\_stats = df.describe()

print(GREEN + "Summary Statistics : " + RESET)

print(summary\_stats)

**OUTPUT:**

****

**MODEL ACCURACY:**

print(BLUE + "\nMODELLING" + RESET);

X = df.drop("Outcome", axis=1);

y = df["Outcome"];

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state= 769);

scaler = StandardScaler();

X\_train = scaler.fit\_transform(X\_train);

X\_test = scaler.transform(X\_test);

model = svm.SVC(kernel="linear");

model.fit(X\_train, y\_train);

y\_pred = model.predict(X\_test);

accuracy = model.score(X\_test, y\_test);

print(GREEN + "Model Accuracy : " + RESET);

print(accuracy);

**OUTPUT:**

MODELLING…..

Model Accuracy: 0. 7337662337662337

**SAVED A NEW FILE:**

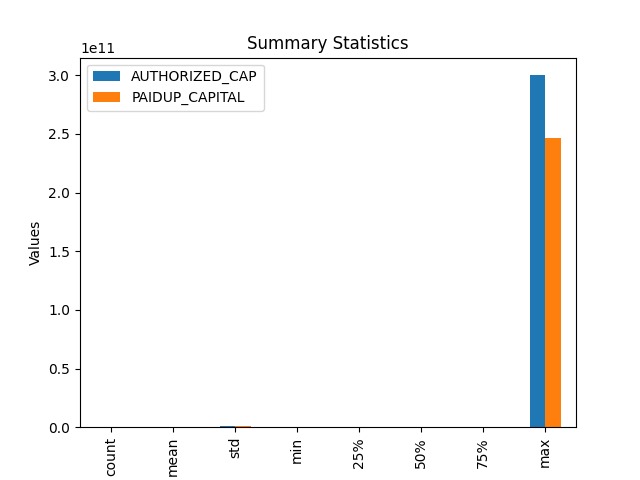
df.to\_csv("clean\_diabetes.csv", index=False)

print(BLUE + "\nDATA SAVING" + RESET)

print(GREEN + "Data Cleaned and Saved !" + RESET)

print("\n")

**OUTPUT:**

****

**CONCLUSION:**

In conclusion, the application of AI-driven exploration and prediction of company registration trends with the Registrar of Companies (RoC) represents a significant advancement in the field of business analytics and regulatory compliance. This innovative approach harnesses the power of artificial intelligence and data analysis to provide valuable insights and predictions related to company registrations. the integration of AI into the exploration and prediction of company registration trends with the RoC has the potential to revolutionize the way businesses are regulated and how they strategize for the future. With responsible and ethical deployment, this technology can contribute to a more transparent, efficient, and dynamic business environment.