ADS\_PHASE2

Assessment of marginal workers in tamilnadu-a socioeconomic analysis

Algorithm:

**# Import necessary libraries**

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import accuracy\_score

import matplotlib.pyplot as plt

**# Load your dataset (replace 'your\_dataset.csv' with the actual dataset file path)**

data = pd.read\_csv('your\_dataset.csv')

**# Data Cleaning (if needed)**

# Example: Handle missing values

data.dropna(inplace=True)

**# Split the dataset into features (X) and target labels (y)**

X = data.drop('target', axis=1)

y = data['target']

**# Split the dataset into training and testing sets (80% train, 20% test)**

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

**# Train a machine learning model (Random Forest in this example)**

model = RandomForestClassifier()

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

**# Test the model by making predictions on the test set**

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

**# Calculate accuracy to evaluate model performance**

accuracy = accuracy\_score(y\_test, y\_pred)

print(f'Accuracy: {accuracy}')

**# Data Visualization (example)**

# Create a scatter plot of two features

plt.scatter(data['feature1'], data['feature2'])

plt.xlabel('Feature 1')

plt.ylabel('Feature 2')

plt.title('Scatter Plot of Features')

plt.show()