**Experiment – 5: Demonstrate Decision tree algorithm for Regression problem.**

**Source Code:**

import numpy as np

import pandas as pd

from sklearn.tree import DecisionTreeRegressor

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

from sklearn.metrics import mean\_squared\_error, r2\_score

import pandas as pd

import matplotlib.pyplot as plt

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

data = pd.read\_csv('/content/Salary\_Data.csv')

# Assuming 'YearsExperience' is your independent variable and 'Salary' is your dependent variable

# Replace these with your actual column names if different

X = data[['YearsExperience']]

y = data['Salary']

# Split the data into training and testing sets

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

# Create and train the decision tree regression model

model = DecisionTreeRegressor(random\_state=42) # You can adjust hyperparameters here

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

# Make predictions on the test set

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

# Evaluate the model

mse = mean\_squared\_error(y\_test, y\_pred)

r2 = r2\_score(y\_test, y\_pred)

print(f"Mean Squared Error: {mse}")

print(f"R-squared: {r2}")

# Plot the results

plt.figure(figsize=(8, 6))

plt.scatter(X\_test, y\_test, color='blue', label='Actual')

plt.plot(X\_test, y\_pred, color='red', label='Predicted')

plt.xlabel('Years of Experience')

plt.ylabel('Salary')

plt.title('Decision Tree Regression')

plt.legend()

plt.show()

**Output:**

Mean Squared Error: 101047709.83333333

R-squared: 0.8021759743735162

