**Coimbatore Institute of Technology**

**Pre-Assessment Test – Curnue**

**Task-02**

**Name : Suruthi R**

**Roll no : 1832054**

**M.Sc Data Science**

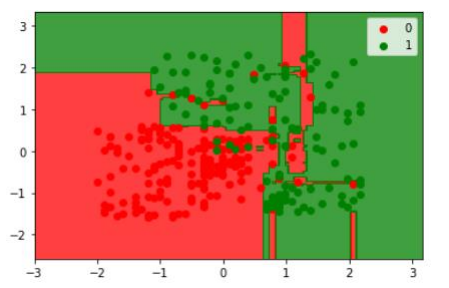
**Ques : SD03Q012**

**Abstract:**

Social network advertising, also social media targeting, is a group of terms that are used to describe forms of [online advertising](https://en.wikipedia.org/wiki/Online_advertising) that focus on [social networking services](https://en.wikipedia.org/wiki/Social_networking_service). One of the major benefits of this type of advertising is that advertisers can take advantage of the users' demographic information and target their ads appropriately. Social network advertising refers to the process of matching social network users to target groups that have been specified by the advertiser.

**Problem Statement:**

The problem statement is to understand the given dataset of Social\_Network\_Ads.csv and try to find the best suitable ML algorithm and write the code in python for algorithm from scratch and try to achieve the below output plot.



**About Dataset:**

The dataset ‘Social\_Network\_Ads.xlsx’ consists of 4 different columns as follows:

* ‘User ID’ – Unique user ID of the people.
* ‘Gender’ – Categorized as Male and Female.
* ‘Age’ – Age of the person in years.
* ‘Estimated Salary’ – The estimated earnings of the person.
* ‘Purchased’ – Whether the person purchased the product under the influence of the Ad, categorized as 1 – Purchased and 0 – Not Purchased.

Here, the column ‘Purchased’ acts as the target variable i.e., the variable to be predicted.

**Tools and Technologies used:**

Google Colaboratory (Colab) - Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. Language Used: Python 3

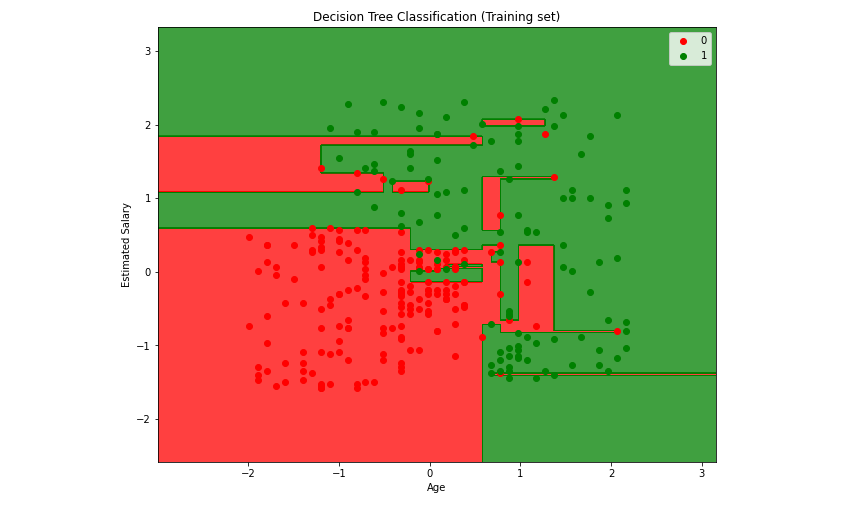
Tools used:

* NumPy
* matplotlib
* pandas
* sklearn

**About Algorithm used:**

Decision tree is the most powerful and popular tool for classification and prediction. A Decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label. Decision tree builds regression or classification models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes.

**Output:**



Accuracy Score: 93%