**Data Mining Techniques:**

|  |
| --- |
| * **Clustering : k-means :**   K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K.  The Κ-means clustering algorithm uses iterative refinement to produce a final result. The algorithm inputs are the number of clusters Κ and the data set. The data set is a collection of features for each data point. The algorithms starts with initial estimates for the Κ centroids, which can either be randomly generated or randomly selected from the data set.   * **Classification : gradient Boosting classifier:**   Gradient boosting is a machine learning technique for regression  and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision tree. It builds the model in a stage-wise fashion like other boosting methods do, and it generalizes them by allowing optimization of an arbitrary differentiable loss function. |

**Results:**

The result of prediction shows how the global warming is affected by consumption of electricity and the degree of temperatures as main reasons and that’s base on the accuracy of our classification model which is equal 88%.as shown in figure 1 which agrees with our hypothesis.

Our research question was (Predict the effects of global warming on global temperature, carbon dioxide levels and annual electricity generation)

And the answer is: the global warming dramatically increases the consumption of electricity and the degree of temperatures but the carbon dioxide increases slightly but steadily, which would produce noticeable effects in the future.

Figure 1



**A screenshot of a cell phone

Description automatically generatedScreenshots:**

Out [Dataset]

A close up of a map

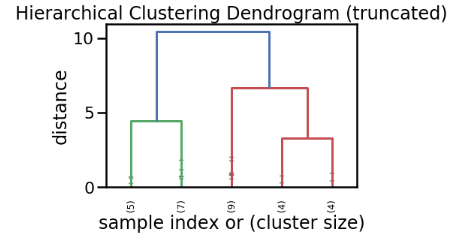
Description automatically generated

Out[clustering using K-means Techniques]

A screenshot of a computer

Description automatically generated

Out[the result extracting from graph using k-means. The cluster initiate by five groups]

****

Out[hierarchical clustering for the dataset]

**A screenshot of a cell phone

Description automatically generated**

Out[Assigning the clusters and plotting the observations as per hierarchical clustering]

**A screenshot of a cell phone

Description automatically generated**

Out [using clustering and then add it to dataset. That means the dataset clustered into five groups that found it above]

**A close up of a map

Description automatically generated**

Out[the graph shows how to the **electricity generation and temperature are increased dramatically by x-axis**

**but the carbon dioxide is** **rising slightly**]