Data Warehouse and Mining

Mini project report

Regression algorithm

Made by:

- **Surya Pratap – 8367**
- **Rachel Jose – 8361**
- **Nolita Rego – 8363**

Under the guidance of: **Mrs. Sujata Deshmukh**

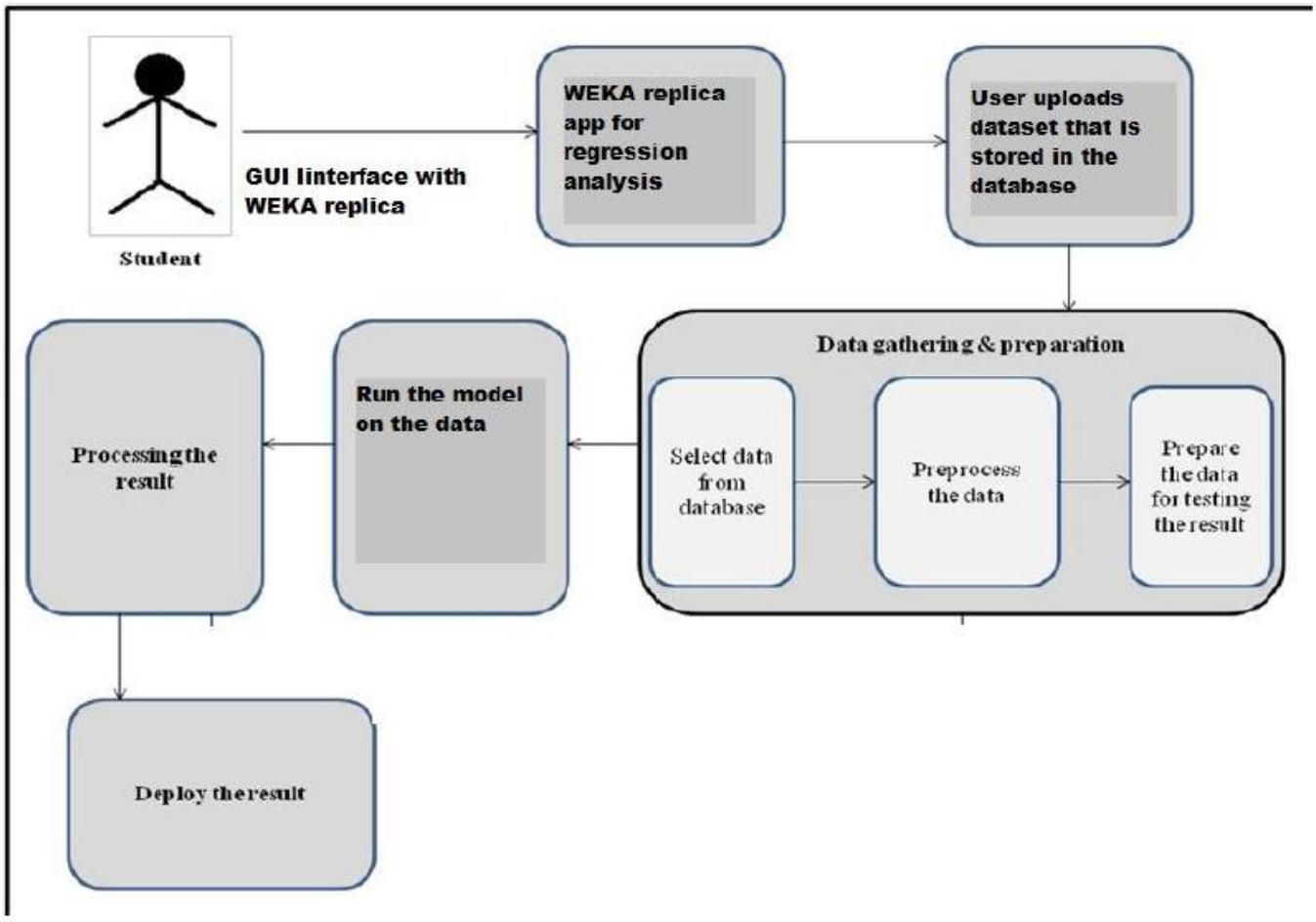
- **Problem Definition: -**

  The system takes in any dataset (for regression model) and evaluated the dataset of Linear Regression model after analyzing the dataset and displays the resulting graphs and other details. And finally gives the accuracy percentage of the model on that particular dataset.

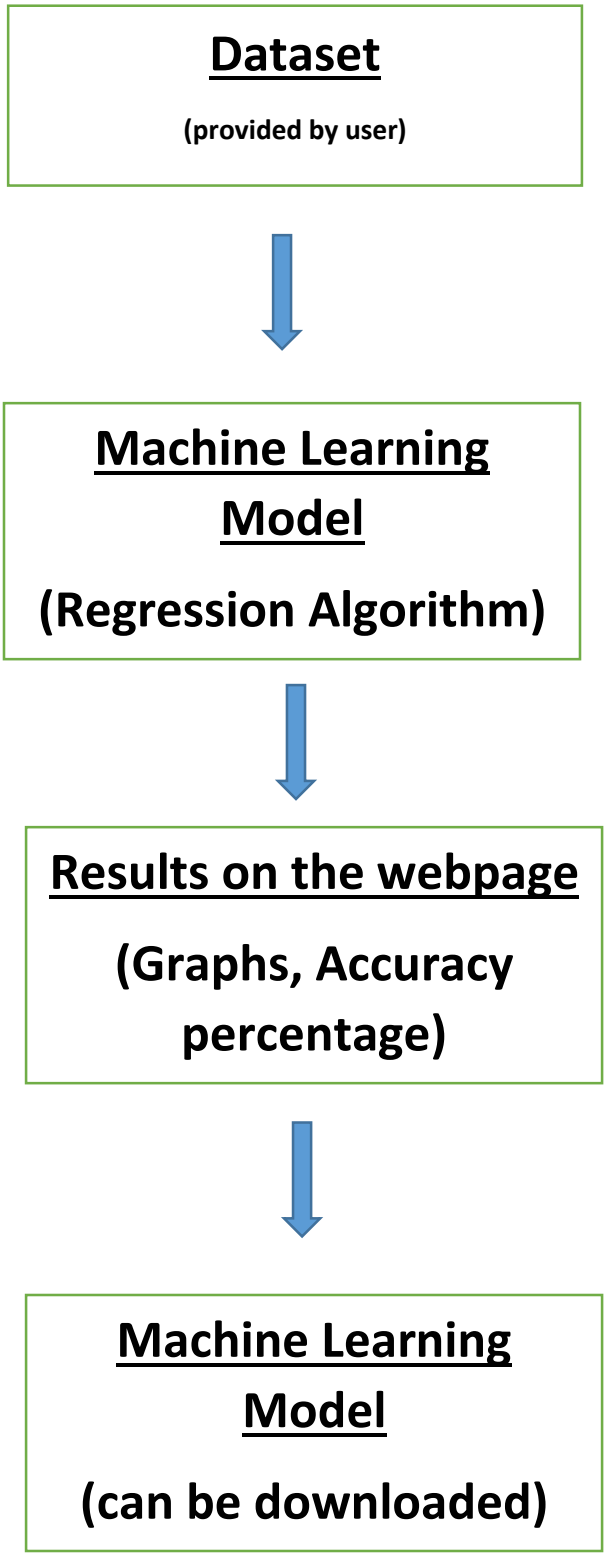
- **Requirements: -**

  - **Dataset** – The dataset must be .csv file and the output variable to be predicted must be at the last column of the dataset table.
  - **Python libraries** – pandas, numpy, matplotlib, seaborn, scikit-learn, flask, werkzeug, os, pathlib, joblib and basic libraries

- **Architecture Diagram: -**



- **Flow Diagram: -**

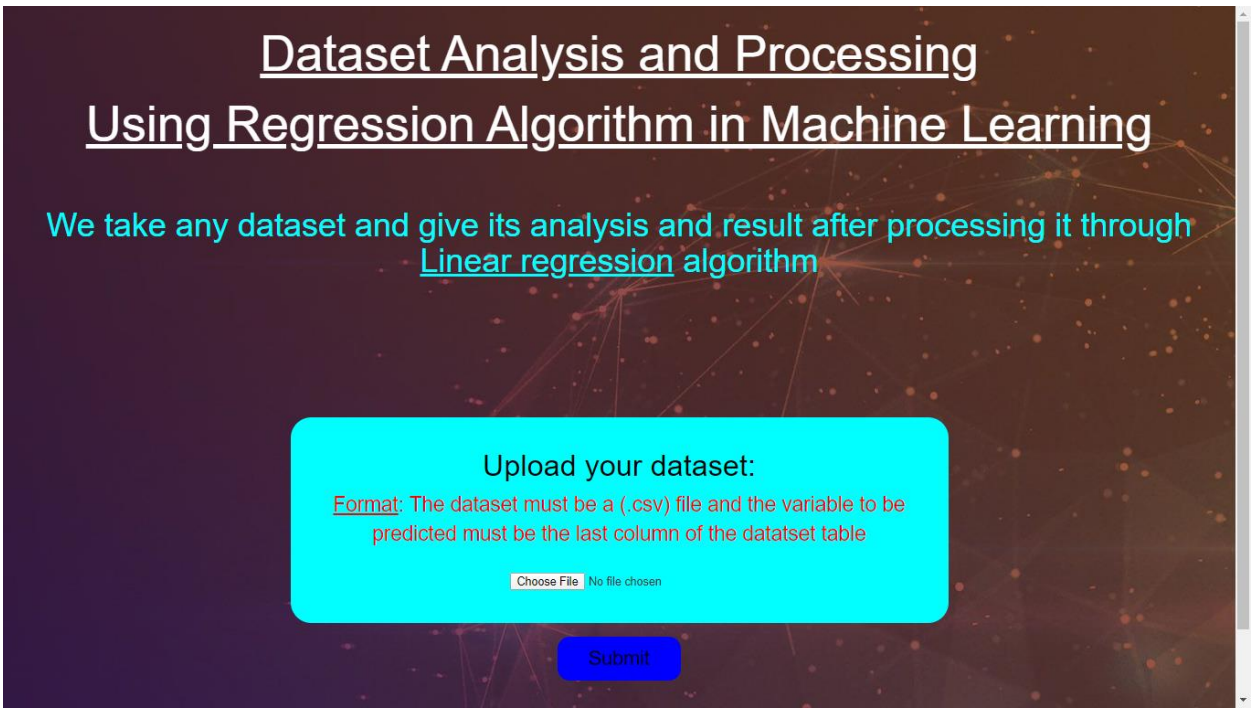


- **Code: - Uploaded in the .zip file on Moodle. Open the main.py file and run using the server.**

- **Screenshots: -**

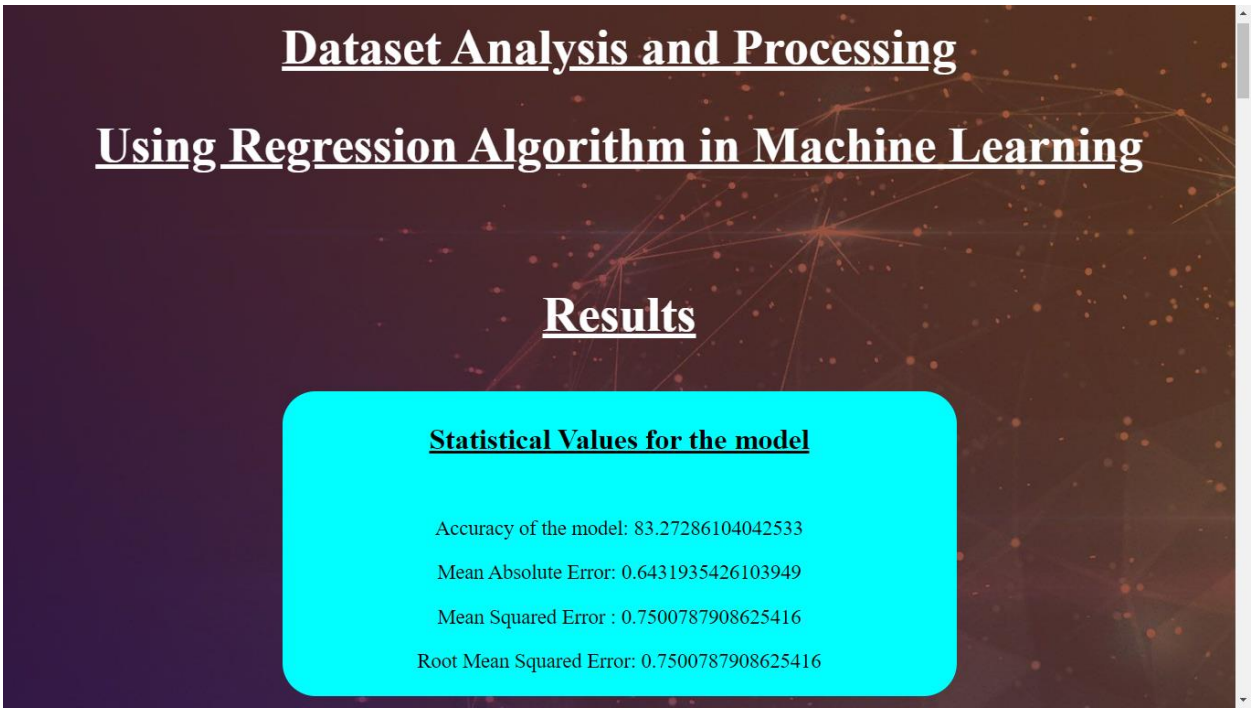
  - **Main page**

User uploads the dataset.



○ Results page

The Model runs on the given dataset and gives the results in three forms- Statistical values, Dataset Analysis and Results of Linear Regression.

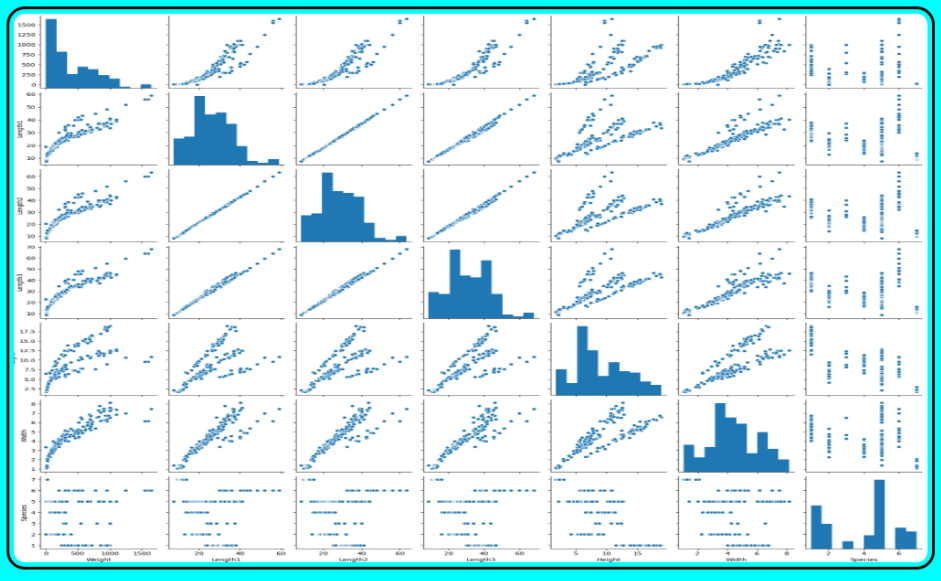


Dataset analysis

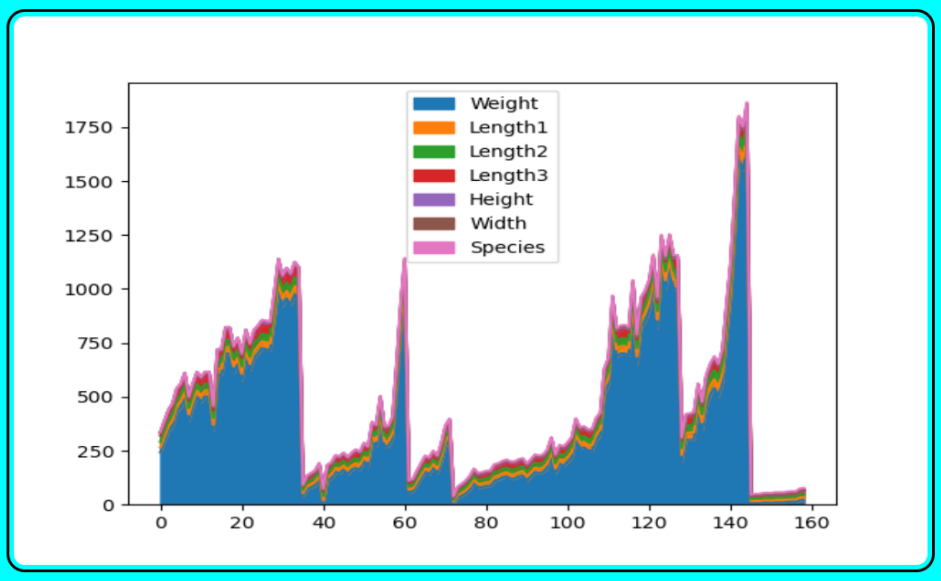
Weights Table

	Coefficient
Weight	0.002754156746377741
Length1	0.5350458769783998
Length2	0.9119974428142997
Length3	-1.2937220137104026
Height	-0.011263283804232742
Width	-0.8992161348433028

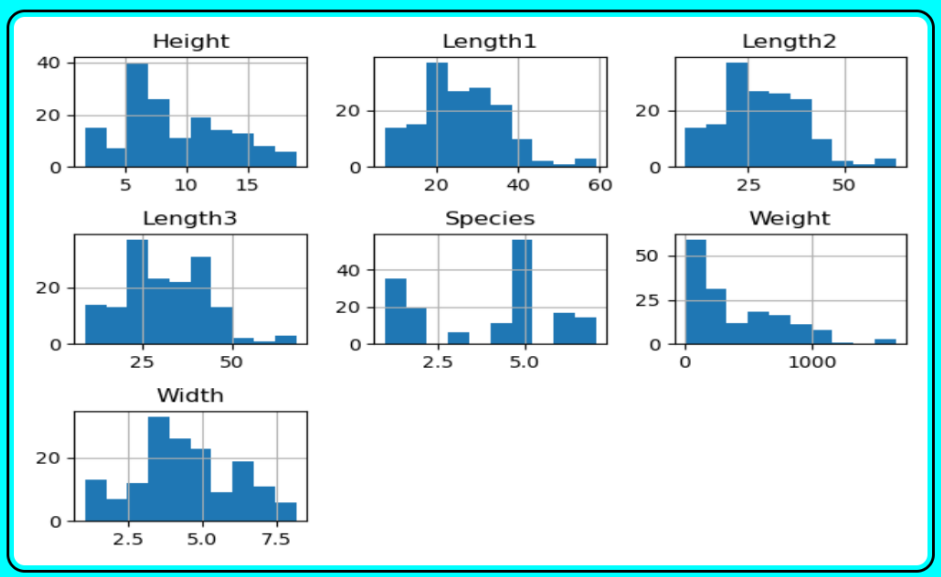
Pair Plots



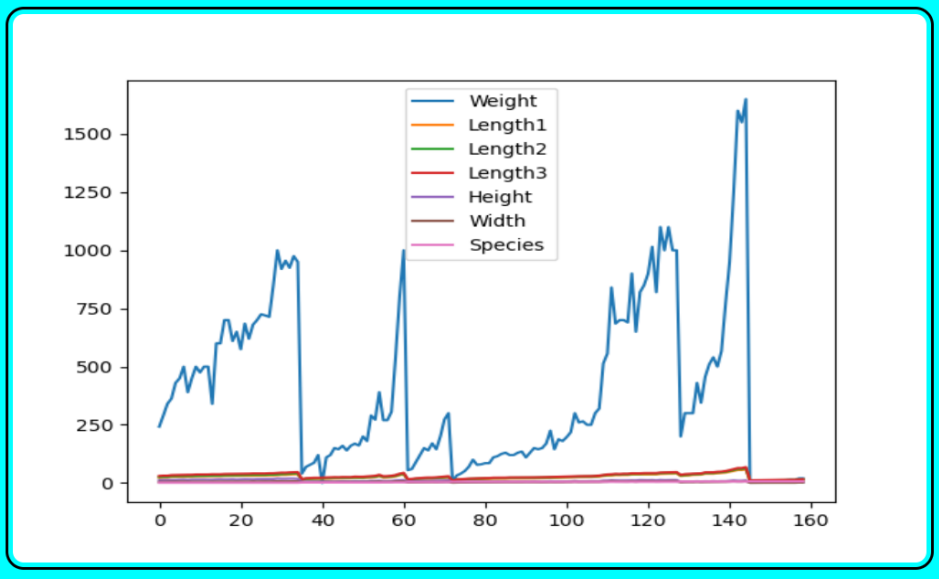
Area Plot



Histogram

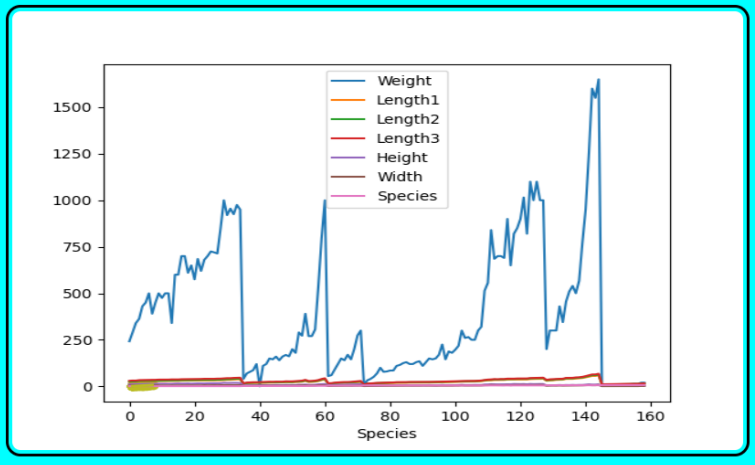


Line Plot

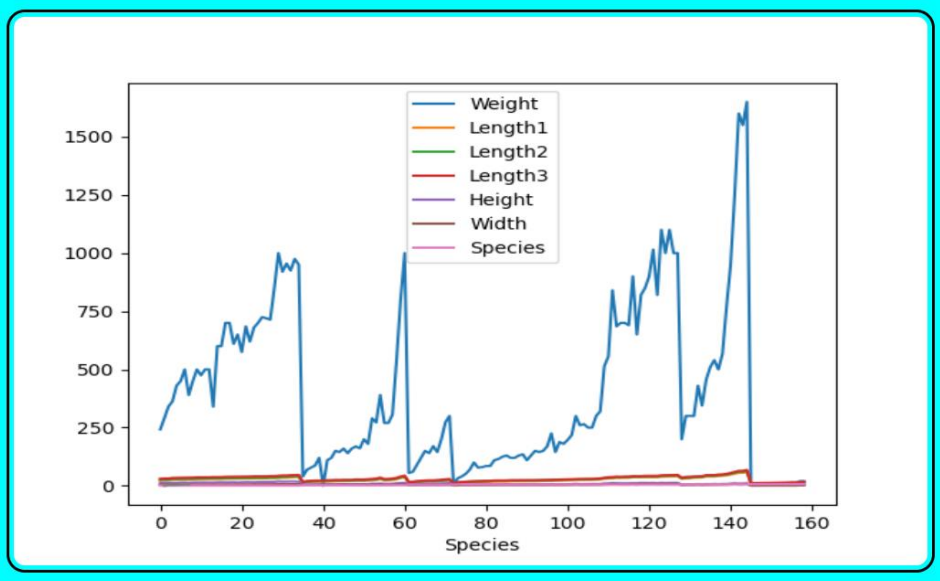


Results of Linear Regression

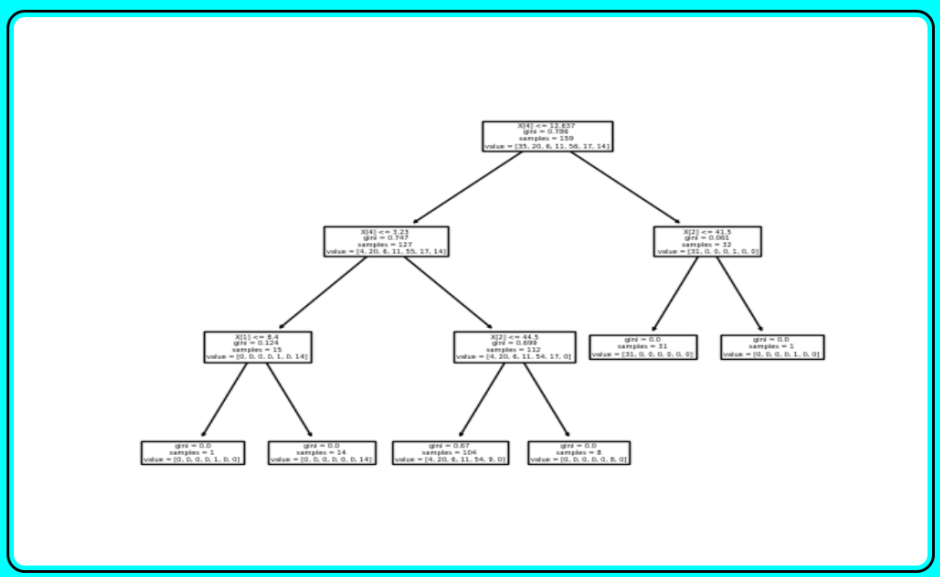
Regression Plot



Line Plot

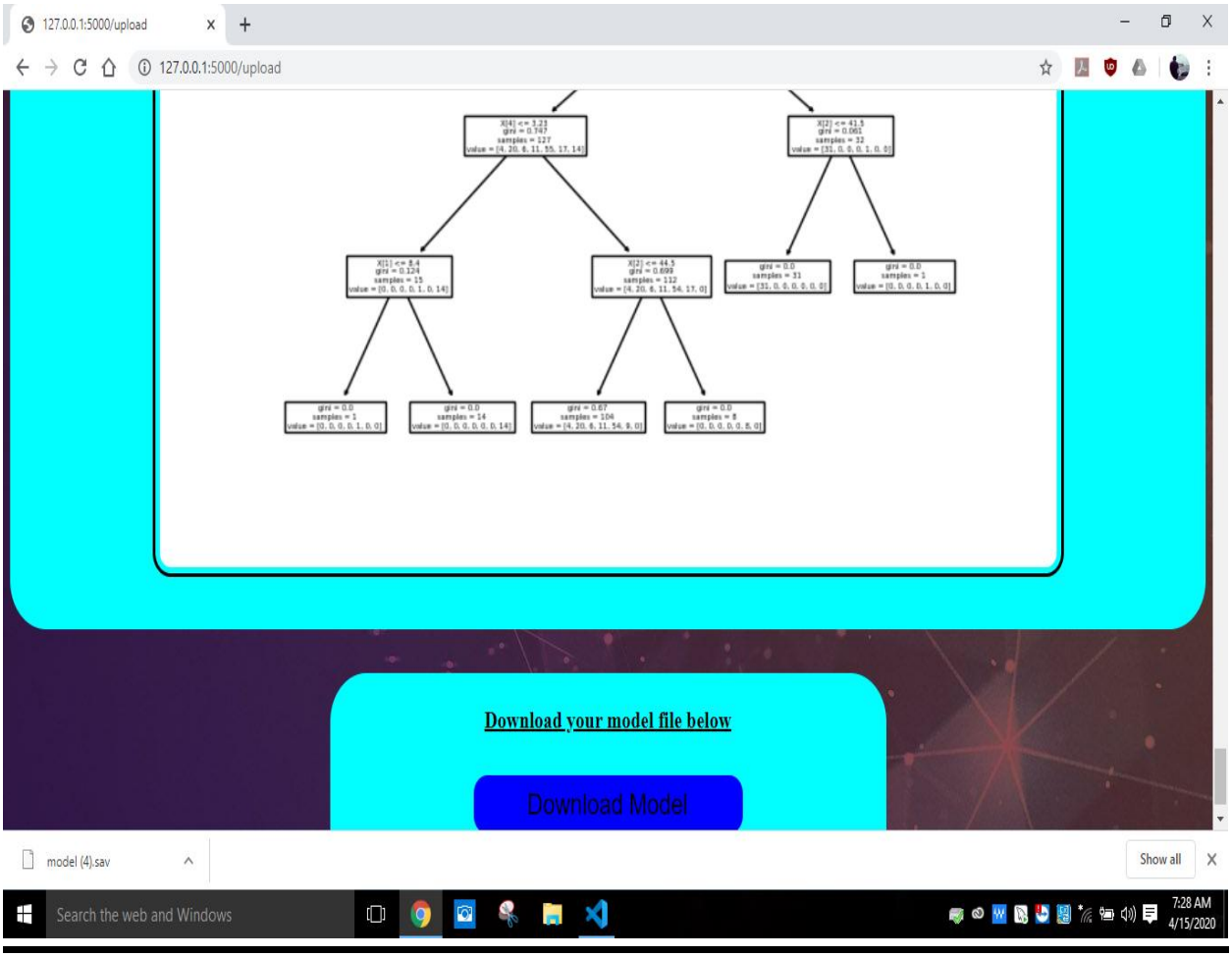


Decision Tree



○ Download Option

There is an option for the user to download the model (. sav) and save it in his system for personal use.



- **<u>Conclusion</u>: -** We have successfully made a Linear Regression model that takes in any dataset (in the mentioned format) and displays the results in three forms – Statistical Analysis, Dataset Analysis and Results of Linear Regression on that model. Graphs, and values are displayed for the dataset as results. We have also provided an option to save the model.

Thank You!!