Project: Linear Regression to Predict Energy Output of a Power Plant

Project Statement:

The environment plays an important role in supply of energy throughout the world. The energy output in a power plant is highly dependent on the environmental changes that occur at the location. If the energy outputs are predicted accurately, then the energy suppliers can collaborate with other sources to minimize the cost and get efficient production of energy. In this project, we are fitting a Linear Regression formula, that can be used to predict the Energy Output at a Power Plant by comparing the relation between various factors such as Ambient Temperature, Ambient Pressure, Vacuum, Relative Humidity. The analytics are performed on a dataset obtained from a Combined Cycle Power Plan (CCPP). The model build gives a very accurate prediction of Energy Output for any newly supplied dataset with the same or similar dataset. Finally, a graph will be plotted which shows the closeness of the predicted energy output with the actual energy output in the dataset.

Team Members

1. Chinmayee Tapaskar
2. Anushree Ankola
3. Eha Sah
4. Raksha Srinivas

Installation Guide

The Technologies used and to be Installed

* Analytical Engine – R, RStudio
* Front End – Rshiny [R Library]
* Back End Server – MySQL

The Language used: R Programming

Features planned to deliver

* Build a Front-End UI for any user to upload dataset of his/her choice.
* Connect the UI to the Database – MySQL.
* Read dataset from the UI.
* Store the dataset in tabular form in the database.
* Retrieve the dataset information to RStudio.
* Perform analytics on dataset and form a model to fit Linear Regression formula to predict Energy Output.
* Plot a graph to visualize the accuracy of prediction.
* Build a notification panel to view real time notification about Power Energy over the world

Features delivered successfully

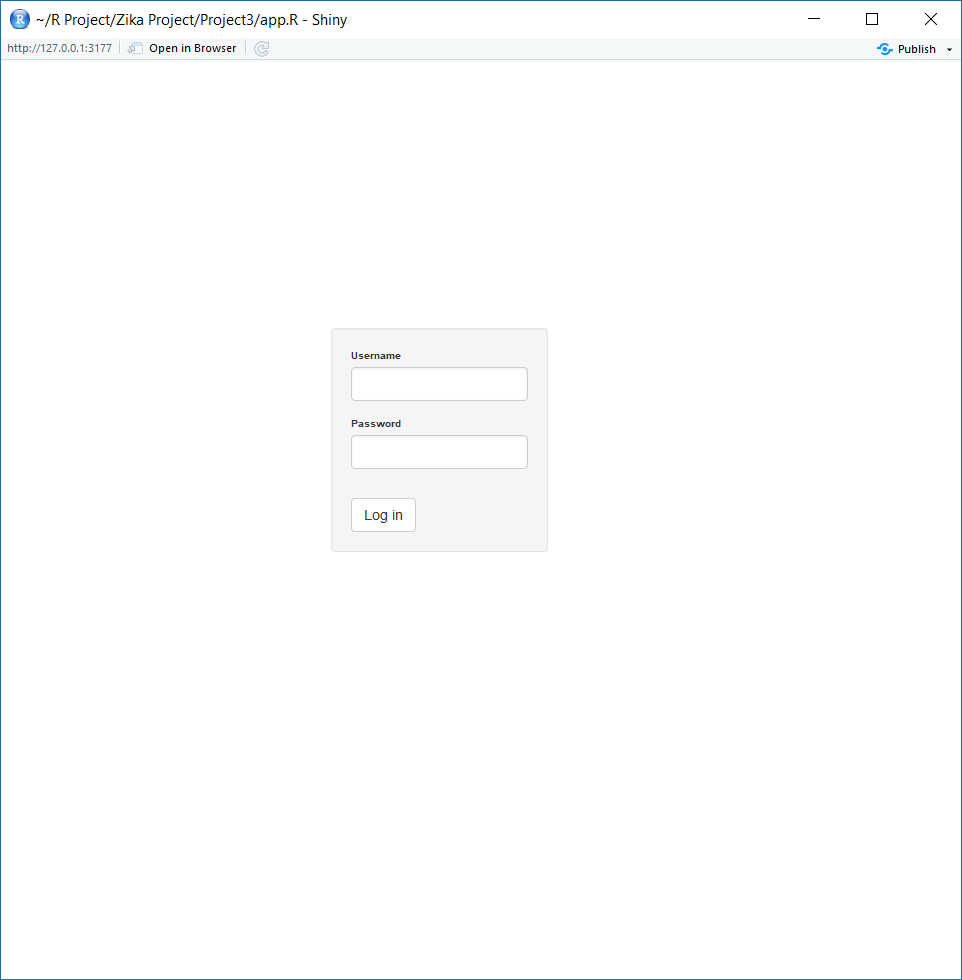
* Build a Front-End UI for any user to upload dataset of his/her choice.
* Connect the UI to the Database – MySQL.
* Read dataset from the UI.
* Store the dataset in tabular form in the database.
* Retrieve the dataset information to RStudio.
* Perform analytics on dataset and form a model to fit Linear Regression formula to predict Energy Output.
* Plot a graph to visualize the accuracy of prediction.

Known Issues

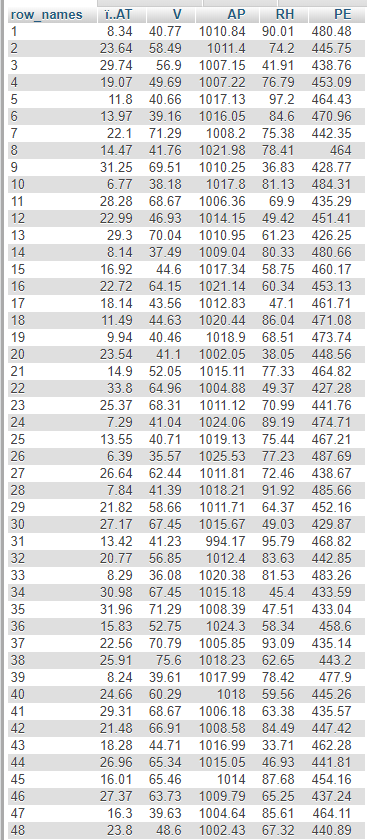
* Team was new to R and took time to understand the features and functionalities of R, hence could not deliver the last feature listed.
* Team was a novice in the usage of R Shiny package.
* Learning a new language called R from root was a challenge
* Team tried to connect various NoSQL databases with the RStudio was futile.
* Performing classification algorithms on another dataset was unproductive.
* Finding the right algorithm to predict/forecast energy output was a challenge but due to dedicated efforts, team was able to come up with regression.
* Due to the previous dataset constraints, the team had to choose a different dataset to perform analytics and predict the output.

User Guide:

1. Install R from <https://www.r-project.org/>
2. Install RStudio from <https://www.rstudio.com>
3. Run RStudio after Installation
4. Run the RScript app.r
5. View the UI of the Project



1. In the login page, enter the user id – adbms1 and Password – admin1 and log-in to the site
2. Upload any dataset to predict the Energy Output



1. View the graph of how accurately the model predicts the Energy Output of the dataset

