Prediction of FRP (Fire Radiative Power) using Linear Regression

By

Mayukh Jain **(0801CA201036)**

submitted to

Prof. Puja Gupta Ma'am

SGSITS INDORE

Solving Linear Regression

1. Analytical Solution:
    a. Train MSE loss: 12986.91797729304
    b. Dev MSE loss: 21784.585730789662
2. Gradient Descent:
    a. Train MSE loss: 24204.37107862648
    b. Dev MSE loss: 37364.82621404062

Hyperparameters used for Analytical Solution:

- C ((L2 Regularization Constant): 1e-5

Hyperparameters used for Gradient Descent:

- Learning Rate: 0.001
- C (L2 Regularization Constant): 1e-5
- Batch Size: 16
- Max steps: 60000
- Eval steps: 1000

Calculation of gradient:

$$Gradient = \frac{-2}{N}[X(Y - \hat{Y})] + 2\lambda w$$

Gradient Descent Stopping Criteria

Step 0: Save the model parameters every time performance on the development set improves

Step 1: When the training completes, return these parameters instead of the parameters in the last epoch.

With early stopping: **Gradient Descent Dev loss:** 37437.23253705786

Without early stopping: **Gradient Descent Dev loss:** 37507.35806383286

Effect of Regularization

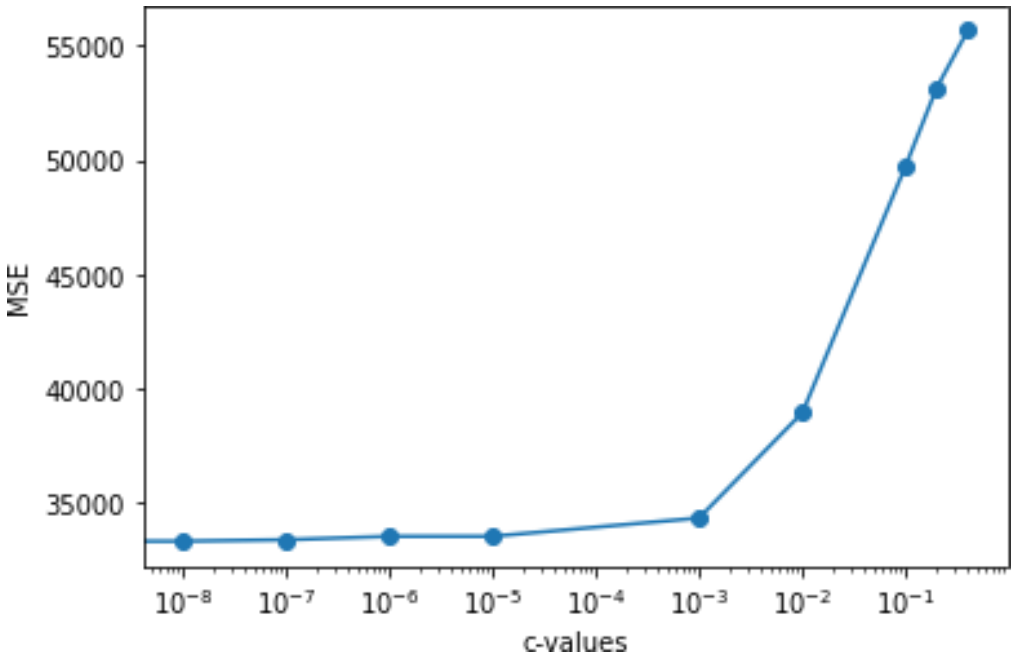

Figure 1: Effect of regularization parameter C on MSE

Figure 1 represents variation of MSE with respect to C (L2 Regularization Constant). By observation, MSE increases for big values of C. We have chosen C value as 1e-5 as it is giving less MSE and also giving sufficient attention to regularization.

As C values increases, we got high MSE since we are giving more attention to summation of weights.

Basis Functions

- We have used Basis functions for attributes: “bright_t31”, “confidence”

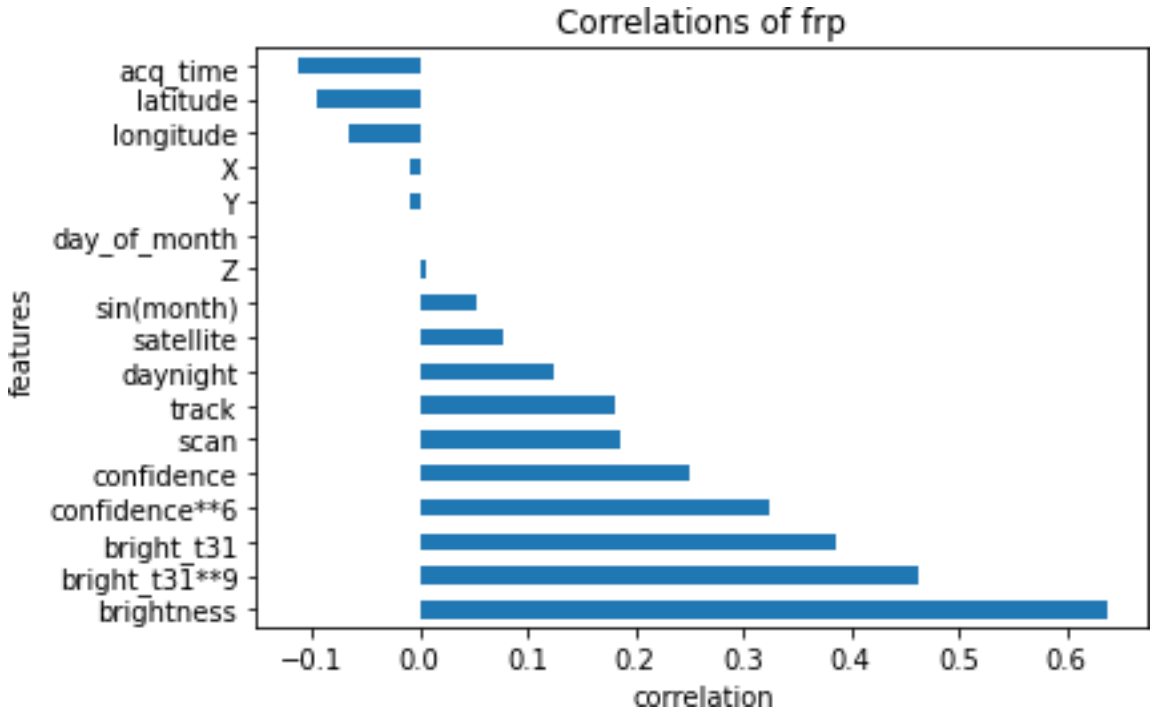

Figure 2: Correlations of FRP and Features

- Due to the high correlation, we used “$(bright\_t31)^9$” and “$(Confidence)^6$”, and for brightness, it had very less effect after using the basis function.
- We also included an extra feature “sine(month)” because of periodicity of the trigonometric function sine.
- Similarly, we added many other extra features which decreased our dev loss.

With basis function: **Analytical Dev loss:** 21784.585730789662

Without basis function: **Analytical Dev loss:** 37100.57336023834

Training Plots:

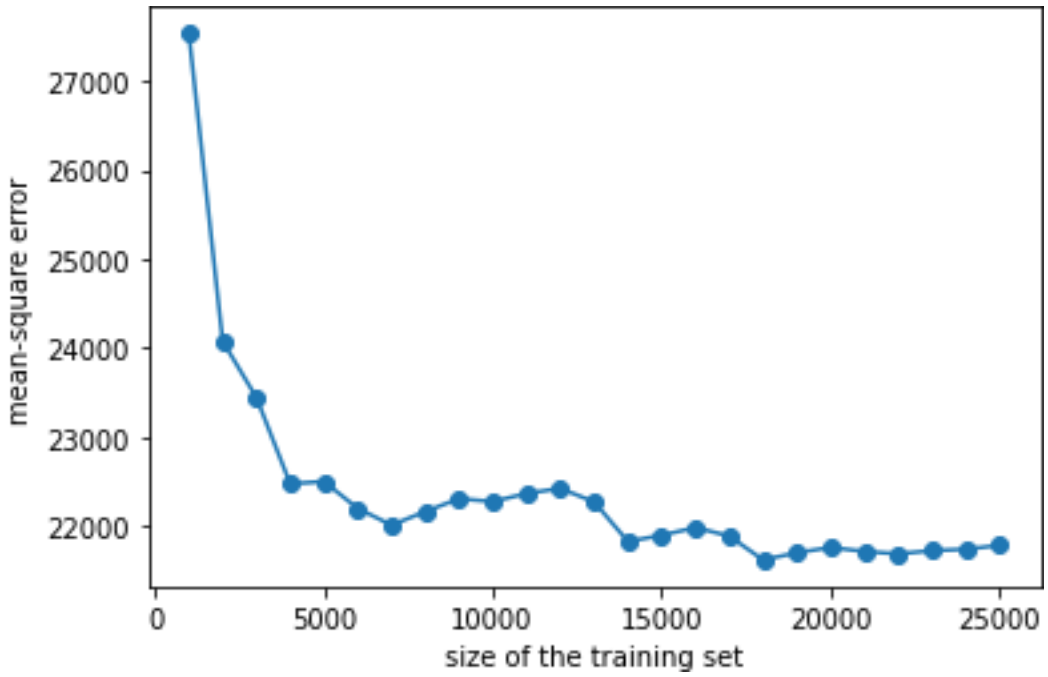

Figure 3: MSE vs Size of Training set

Figure 3 represents variation of MSE with respect to size of training set. From this plot, we can infer that for training size of 18000, MSE turns out to be the least!

Feature Importance

Most Important Feature: “brightness”

Least important Features: "version", "instrument"

Here is the plot representing the correlation of FRP with the corresponding features:

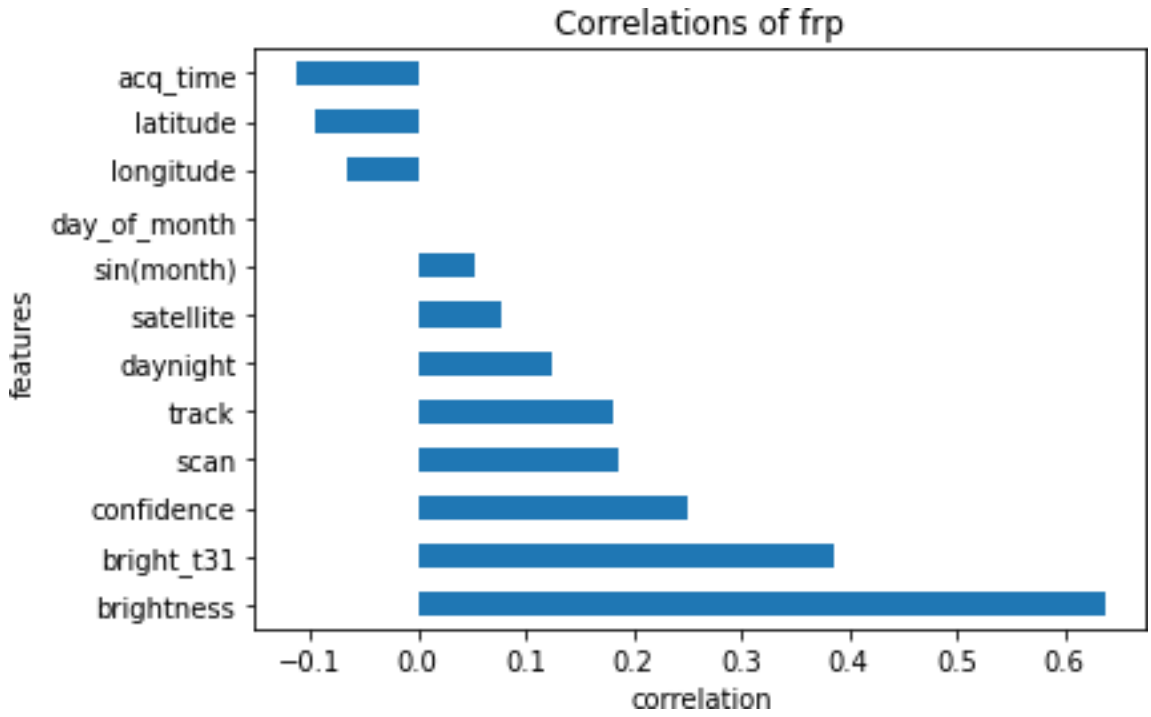

Figure 4: Correlations of FRP with features

Here, “frp” has highest correlation with “brightness”, while it has no correlation with “version” and “instrument”. Moreover, version and instrument both have single value.

Hence, it can be concluded that “brightness” is the most important feature, while “version” and “instrument” are the least important features because their values remain constant for whole dataset.