

**Regression Assignment**

[support@intellipaat.com](mailto:support@intellipaat.com)

+91-7022374614

US: 1-800-216-8930 (Toll Free)

**Instructions for the Assignment**

1. Use a single file in Jupyter Notebook or Goggle collab for the entire assignment.
2. Download the Assignment file (File> Download > Download as .ipynb), compress it in a Zip folder and submit through LMS only.
3. Use the given dataset (**Concrete Compressive Strength.csv**) for this Assignment.
4. Submit the Assignment by 11th February 2022 without fail.
5. This Assignment is an Exam kind and is considered for the evaluation and certification. No support is provided by our technical team.
6. If you have any doubts how to work on this assignment (No technical support), please drop an email to [support@intellipaat.com](mailto:support@intellipaat.com)

**FINAL ASSIGNMENT**

**Problem Statement**

Predicting the compressive strength of concrete given its composition and its age.

**Variable description**

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| Cement | Cement (kg in a m3 mixture) |
| BFSlag | Blast Furnace Slag – non-metallic co-product produced during a furnace operation(kg in a m^3 mixture) |
| FlyAsh | Fly Ash – Ash produced from burning coal (kg in a m3 mixture) |
| Water | Water  (kg in a m3 mixture) |
| Superplasticizer | Superplasticizer – Additive for preventing aggregate formation (kg in a m^3 mixture) |
| CoarseAggregate | Coarse Aggregate -  typically gravel   (kg in a m3 mixture) |
| FineAggregate | Fine Aggregate – crushed stone or sand between 9.5 mm and 75 µm (kg in a m3 mixture) |
| Age | Age (day) |
| Ccstrength | Concrete compressive strength(MPa, megapascals) |

1. The total number of missing values within the dataset.
2. Which column / feature requires correction in the type of value they hold?
3. After imputation of nulls with mean what is the average value of the compressive strength in concrete?
4. The feature that has a moderately strong relationship with compressive strength in concrete is?
5. Standardize the dataset using standardscaler(), split the dataset into train and test of proportions 70:30 and set the random state to 1. Build a Linear Regression Model on the data and the resulting r-squared value is between which range?