Machine Learning Practise
Playground Series - Season 3, Episode 9

Ben Colquhoun

7/3/23

Outline of the Problem

The task here is to predict the strength of concrete, given an instance of the concrete. This task was found as a Kaggle competition -"Playground Series - Season 3, Episode 9", joined 7 days into the 2 week duration.

Exploratory Data Analysis

Initial Features

Inital investigation of the data provided the following features from the the training set, with expanded explanations with reference to a Kaggle discussion of Phong Nyugen [1].

- id
  - Id of the instance
  - Integer
  - Unimportant feature for comparing strength of concrete

- CementComponent
  - Amount of cement added to concrete
  - Float
  - More cement improves strength of the concrete, though too much may cause it to be brittle. Important feature.

- BlastFurnaceSlag
  - Amount of slag added to concrete
  - Float

- Slag can be used as a cement substitute in creating concrete. Important feature.

- FlyAshComponent

  - Amount of fly ash added to concrete
  - Float
  - Fly ash can be used as a cement substitute in creating concrete. Important feature.

- WaterComponent

  - Amount of water added to concrete
  - Float
  - Water is the binding agent of the other components. Important feature.

- SuperplasticizerComponent

  - Amount of superplasticizer added to concrete
  - Float
  - Chemical additive to improve workability without changing water content. Reasonably important feature.

- CoarseAggregateComponent

  - Amount of coarse aggregate added to concrete
  - Float
  - Gravel/crushed stone added to cement for structure. Reasonably important feature.

- FineAggregateComponent

  - Amount of fine aggregate added to concrete
  - Float
  - Sand added to cement for structure. Reasonably important feature.

- AgeInDays

  - Days since concrete poured
  - Integer
  - Longer drying times increases the strength up to a point dependant on other factors. Important feature when combined with other factors.

References

[1] Nyugen, P. (2023, March). *Detailed feature description and feature engineering by ChatGPT*. Retrieved March 7, 2023, from https://www.kaggle.com/competitions/playground-series-s3e9/discussion/391066