Predicting Compressive Strength of Concrete using Machine learning

Authors

- **Andrew Bushnell**
  · andrewb7777
  Department of CEE, University of Illinois

- **Kanchan Kulhalli**
  · Kanchan-uiuc
  Department of CEE, University of Illinois

- **Vikram Gadge**
  · vgadge2
  Department of CEE, University of Illinois

Abstract

Using a data set covering the concrete compressive strength of a variety of different mixture components, we are going to create a machine learning program that will be able to predict when concrete failures will occur, which components and the combinations of these components will work best based on strength requirements of certain structures, and predict maximum allowable loads that can be achieved based on the mixtures

Using these future trends, we will we be able to reach certain conclusions on the future improvements, designs, and materials that should be used in certain structures that we will be able to present to individuals in the fields that use these structures. They then can use these recommendations in their future projects. We will be using the data set of “Concrete Compressive Strength” which was obtained using Kaggle.com [1]. The data comes in the form of an excel file and We will compile all of the data into specific tables and use them to create the future trends we stated above.

We will be creating new tables and figures that will be of comparisons of when the concrete fails vs the concrete material, strength of concrete vs water to cement ratio, concrete composition vs concrete strength, max allowable loads vs concrete material, max allowable loads vs concrete permutations. In future we will be adding cost of components as new dimension and check out what’s the best and minimal combination to make it cost effective and compare the cost and strength graph.

We intend to use Julia to compile these new tables using machine learning tools that can be used to predict permutations, concrete to water ratios etc, that are not specifically included within the data set so we can accurately predict these unknown values that can then be used to run theoretical tests in real life construction project scenarios.

Exploratory Data Analysis

The data set is composed of nine columns of data that state the following information: Fly Ash component, Water component, Superplasticizer, Coarse Aggregate, Age, and Concrete Compressive

Strength. These columns have the following units of measurements: kg in m^3 mixture, kg in m^3 mixture, kg in m^3 mixture, kg in m^3 mixture, kg in m^3 mixture, kg in m^3 mixture, kg in m^3 mixture, days, MPa megapascals. The excel data set has a total of 1030 rows of this data. We found a few discrepancies in the data set and we decided to clean the data before doing any exploratory analysis. The below section describes it in detail.

Data Cleaning

The Dataset that we selected comprised of rows that were repeating multiple times. We could only learn about this when we began with asking questions and trying to code them out. So we go back and clean the data set by using “Unique” fuction, after which the rows reduced from 1030 to 1005.

The second challenge we were faced with was that the values of ingredients in the concrete were same for multiple rows but only the Concrete compressive strengths were varying i.e the input columns with same values generated different output. So we took a mean of those values and combined them into a single row.

Description of the Dataset

Before we set out to ask questions, we generated the below table to get a deeper understanding of the columns in our dataset.

	variable	min	mean	median	max
	Symbol	Real	Float64	Float64	Real
1	:cement(kg per m3)	102.0	276.873	259.95	540.0
2	:blast_furnace_slag(kg per m3)	1.0	73.0007	20.0	359.4
3	:fly_ash(kg per m3)	1.0	55.6028	1.0	200.1
4	:water(kg per m3)	121.8	182.368	185.7	247.0
5	:superplasticizer(kg per m3)	1.0	6.34415	6.0	32.2
6	:coarse_aggregate(kg per m3)	801.0	974.597	968.0	1145.0
7	:fine_aggregate(kg per m3)	594.0	773.081	780.0	992.6
8	:age(days)	1	46.1663	28.0	365
9	:concrete_compressive_strength(MPa)	2.33	35.119	33.73	82.6

Several studies independly have shown that concrete stregth development is determined not only by the water cement ratio, but that it is also influenced by the content of other concrete ingredients as well. Hence we tried to look into effects of various other ingredients on the compressive strength of concrete.In the below sections, we describe our columns in details and ask pertinent questions on our dataset.

Water and Cement

The Abrams’ water-to-cement ratio (w/c) pronouncement of 1918 has been described as the most useful and significant advancement in the history of concrete technology. His most important formulation was the inverse proportionality between the w/c ratio and the strength of concrete. The generally accepted Abrams rule is a formulation of the observation that an increase in the w/c decreases

To check whether the Abrams’ law holds true or false. We are comparing the w/c ratio with the compressive strength of the concrete.

wc_ratio = select(df, [“cement”, “water”, “concrete_compressive_strength”])

df.wc_ratio .= df.water ./ df.cement

wc_ratio2 = select(df, [“wc_ratio”, “water”, “cement”,“concrete_compressive_strength”])

Blast Furnace Slag

Fly Ash

Coarse Aggregate and Fine aggregate

Superplasticizer

Age

Concrete Compressive Strength

Predictive Modeling

References

1. **Concrete Compressive Strength** https://www.kaggle.com/datasets/sinamhd9/concrete-comprehensive-strength