MULTIPLE REGRESSION PROJECT

Manasa Swetha Tiramareddy (U20933669)

Kiran Kumar Barnala(U74685418)

Aravind Dudam(U04440266)

(Equal contribution)

Citation:

<https://archive.ics.uci.edu/ml/datasets/concrete+compressive+strength>

I-Cheng Yeh, "Modeling of strength of high performance concrete using artificial neural networks," Cement and Concrete Research, Vol. 28, No. 12, pp. 1797-1808 (1998)

# Description of Variables:

Number of instances 1030  
Number of Attributes 9  
Attribute breakdown 8 quantitative input variables, and 1 quantitative output variable

Given are the variable name, variable type, the measurement unit and a brief description. The concrete compressive strength is the regression problem. The order of this listing corresponds to the order of numerals along the rows of the database.  
  
Name -- Data Type -- Measurement -- Description  
  
Cement (component 1) -- quantitative -- kg in a m3 mixture -- Input Variable  
Blast Furnace Slag (component 2) -- quantitative -- kg in a m3 mixture -- Input Variable  
Fly Ash (component 3) -- quantitative -- kg in a m3 mixture -- Input Variable  
Water (component 4) -- quantitative -- kg in a m3 mixture -- Input Variable  
Superplasticizer (component 5) -- quantitative -- kg in a m3 mixture -- Input Variable  
Coarse Aggregate (component 6) -- quantitative -- kg in a m3 mixture -- Input Variable  
Fine Aggregate (component 7) -- quantitative -- kg in a m3 mixture -- Input Variable  
Age -- quantitative -- Day (1~365) -- Input Variable  
Concrete compressive strength -- quantitative -- MPa -- Output Variable

A complete listing of the data used in your project if you have less than 100 cases. If more, a sample of 100 cases is sufficient.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cement | Blast Furnace Slag | Fly Ash | Water | Superplasticizer | Coarse Aggregate | Fine Aggregate | Age | Concrete compressive strength |
| 540.0 | 0.0 | 0.0 | 162.0 | 2.5 | 1040.0 | 676.0 | 28 | 79.99 |
| 540.0 | 0.0 | 0.0 | 162.0 | 2.5 | 1055.0 | 676.0 | 28 | 61.89 |
| 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 270 | 40.27 |
| 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 365 | 41.05 |
| 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 825.5 | 360 | 44.30 |
| 266.0 | 114.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 90 | 47.03 |
| 380.0 | 95.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 365 | 43.70 |
| 380.0 | 95.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 28 | 36.45 |
| 266.0 | 114.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 28 | 45.85 |
| 475.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 28 | 39.29 |
| 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 825.5 | 90 | 38.07 |
| 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 825.5 | 28 | 28.02 |
| 427.5 | 47.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 270 | 43.01 |
| 190.0 | 190.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 90 | 42.33 |
| 304.0 | 76.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 28 | 47.81 |
| 380.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 90 | 52.91 |
| 139.6 | 209.4 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 90 | 39.36 |
| 342.0 | 38.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 365 | 56.14 |
| 380.0 | 95.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 90 | 40.56 |
| 475.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 180 | 42.62 |
| 427.5 | 47.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 180 | 41.84 |
| 139.6 | 209.4 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 28 | 28.24 |
| 139.6 | 209.4 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 3 | 8.06 |
| 139.6 | 209.4 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 180 | 44.21 |
| 380.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 365 | 52.52 |
| 380.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 270 | 53.30 |
| 380.0 | 95.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 270 | 41.15 |
| 342.0 | 38.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 180 | 52.12 |
| 427.5 | 47.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 28 | 37.43 |
| 475.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 7 | 38.60 |
| 304.0 | 76.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 365 | 55.26 |
| 266.0 | 114.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 365 | 52.91 |
| 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 825.5 | 180 | 41.72 |
| 475.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 270 | 42.13 |
| 190.0 | 190.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 365 | 53.69 |
| 237.5 | 237.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 270 | 38.41 |
| 237.5 | 237.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 28 | 30.08 |
| 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 90 | 37.72 |
| 475.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 90 | 42.23 |
| 237.5 | 237.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 180 | 36.25 |
| 342.0 | 38.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 90 | 50.46 |
| 427.5 | 47.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 365 | 43.70 |
| 237.5 | 237.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 365 | 39.00 |
| 380.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 180 | 53.10 |
| 427.5 | 47.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 90 | 41.54 |
| 427.5 | 47.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 7 | 35.08 |
| 349.0 | 0.0 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 3 | 15.05 |
| 380.0 | 95.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 180 | 40.76 |
| 237.5 | 237.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 7 | 26.26 |
| 380.0 | 95.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 7 | 32.82 |
| 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 180 | 39.78 |
| 190.0 | 190.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 180 | 46.93 |
| 237.5 | 237.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 90 | 33.12 |
| 304.0 | 76.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 90 | 49.19 |
| 139.6 | 209.4 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 7 | 14.59 |
| 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 825.5 | 7 | 14.64 |
| 475.0 | 0.0 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 365 | 41.93 |
| 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 825.5 | 3 | 9.13 |
| 304.0 | 76.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 180 | 50.95 |
| 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 28 | 33.02 |
| 304.0 | 76.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 270 | 54.38 |
| 266.0 | 114.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 270 | 51.73 |
| 310.0 | 0.0 | 0.0 | 192.0 | 0.0 | 971.0 | 850.6 | 3 | 9.87 |
| 190.0 | 190.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 270 | 50.66 |
| 266.0 | 114.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 180 | 48.70 |
| 342.0 | 38.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 270 | 55.06 |
| 139.6 | 209.4 | 0.0 | 192.0 | 0.0 | 1047.0 | 806.9 | 360 | 44.70 |
| 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 594.0 | 7 | 30.28 |
| 190.0 | 190.0 | 0.0 | 228.0 | 0.0 | 932.0 | 670.0 | 28 | 40.86 |
| 485.0 | 0.0 | 0.0 | 146.0 | 0.0 | 1120.0 | 800.0 | 28 | 71.99 |
| 374.0 | 189.2 | 0.0 | 170.1 | 10.1 | 926.1 | 756.7 | 3 | 34.40 |
| 313.3 | 262.2 | 0.0 | 175.5 | 8.6 | 1046.9 | 611.8 | 3 | 28.80 |
| 425.0 | 106.3 | 0.0 | 153.5 | 16.5 | 852.1 | 887.1 | 3 | 33.40 |
| 425.0 | 106.3 | 0.0 | 151.4 | 18.6 | 936.0 | 803.7 | 3 | 36.30 |
| 375.0 | 93.8 | 0.0 | 126.6 | 23.4 | 852.1 | 992.6 | 3 | 29.00 |
| 475.0 | 118.8 | 0.0 | 181.1 | 8.9 | 852.1 | 781.5 | 3 | 37.80 |
| 469.0 | 117.2 | 0.0 | 137.8 | 32.2 | 852.1 | 840.5 | 3 | 40.20 |
| 425.0 | 106.3 | 0.0 | 153.5 | 16.5 | 852.1 | 887.1 | 3 | 33.40 |
| 388.6 | 97.1 | 0.0 | 157.9 | 12.1 | 852.1 | 925.7 | 3 | 28.10 |
| 531.3 | 0.0 | 0.0 | 141.8 | 28.2 | 852.1 | 893.7 | 3 | 41.30 |
| 425.0 | 106.3 | 0.0 | 153.5 | 16.5 | 852.1 | 887.1 | 3 | 33.40 |
| 318.8 | 212.5 | 0.0 | 155.7 | 14.3 | 852.1 | 880.4 | 3 | 25.20 |
| 401.8 | 94.7 | 0.0 | 147.4 | 11.4 | 946.8 | 852.1 | 3 | 41.10 |
| 362.6 | 189.0 | 0.0 | 164.9 | 11.6 | 944.7 | 755.8 | 3 | 35.30 |
| 323.7 | 282.8 | 0.0 | 183.8 | 10.3 | 942.7 | 659.9 | 3 | 28.30 |
| 379.5 | 151.2 | 0.0 | 153.9 | 15.9 | 1134.3 | 605.0 | 3 | 28.60 |
| 362.6 | 189.0 | 0.0 | 164.9 | 11.6 | 944.7 | 755.8 | 3 | 35.30 |
| 286.3 | 200.9 | 0.0 | 144.7 | 11.2 | 1004.6 | 803.7 | 3 | 24.40 |
| 362.6 | 189.0 | 0.0 | 164.9 | 11.6 | 944.7 | 755.8 | 3 | 35.30 |
| 439.0 | 177.0 | 0.0 | 186.0 | 11.1 | 884.9 | 707.9 | 3 | 39.30 |
| 389.9 | 189.0 | 0.0 | 145.9 | 22.0 | 944.7 | 755.8 | 3 | 40.60 |
| 362.6 | 189.0 | 0.0 | 164.9 | 11.6 | 944.7 | 755.8 | 3 | 35.30 |
| 337.9 | 189.0 | 0.0 | 174.9 | 9.5 | 944.7 | 755.8 | 3 | 24.10 |
| 374.0 | 189.2 | 0.0 | 170.1 | 10.1 | 926.1 | 756.7 | 7 | 46.20 |
| 313.3 | 262.2 | 0.0 | 175.5 | 8.6 | 1046.9 | 611.8 | 7 | 42.80 |
| 425.0 | 106.3 | 0.0 | 153.5 | 16.5 | 852.1 | 887.1 | 7 | 49.20 |
| 425.0 | 106.3 | 0.0 | 151.4 | 18.6 | 936.0 | 803.7 | 7 | 46.80 |
| 375.0 | 93.8 | 0.0 | 126.6 | 23.4 | 852.1 | 992.6 | 7 | 45.70 |
| 475.0 | 118.8 | 0.0 | 181.1 | 8.9 | 852.1 | 781.5 | 7 | 55.60 |

Regression analyses of the following model combinations:

**Code**:

rm(list = ls())

install.packages("rio")

install.packages("moments")

install.packages("data.table")

library(moments)

library(rio)

#importing data

data = import("~/Concrete\_Data (1).xls")

#subset data

data\_1 = data[sample(1:nrow(data),100),]

set.seed(20933669)

data\_1

colnames(data\_1)=tolower(make.names(colnames(data\_1)))

colnames(data\_1)

attach(data\_1)

#Structure of data

str(data\_1)

Text

Description automatically generated

#Scatterplot matrix of the continuous variables

cont\_data = data\_1[,c("cement","blast.furnace.slag","fly.ash",

"water","superplasticizer","coarse.aggregate","fine.aggregate","age","concrete.compressive.strength")]

plot(cont\_data,pch=19)

Table

Description automatically generated

#Correlation matrices of the continuous variables in numbers

library("corrplot")

corr\_matrices = cor(cont\_data)

corrplot(corr\_matrices,method="number")

Chart

Description automatically generated

#Correlation matrices of the continuous variables in numbers

corrplot(corr\_matrices,method="ellipse")

Chart, bubble chart

Description automatically generated

# Regression Analysis:

lrm=lm(concrete.compressive.strength~cement,data=cont\_data)

summary(lrm)

Text

Description automatically generated

The above simple regression model has one independent variable which is cement and the dependent variable is concrete compressive strength

Here regression equation is: concrete.compressive.strength = 0.09368\*cement + 8.52556

lrm\_1=lm(concrete.compressive.strength~age,data=cont\_data)

summary(lrm\_1)

Text

Description automatically generated

The above simple regression model has one independent variable which is age and the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength = 0.10892\*age + 30.63694

**Code**:

lrm\_2=lm(concrete.compressive.strength~fly.ash,data=cont\_data)

summary(lrm\_2)

Text

Description automatically generated

**Analysis**: The above simple regression model has one independent variable which is fly.ash and the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength = -0.04040\*fly.ash + 37.10847

lrm01=lm(concrete.compressive.strength~cement+age,data=cont\_data)

summary(lrm01)

Text

Description automatically generated

The above multiple regression model has independent variables cement and age where the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength =0.09277\*cement +0.10590\*age+ 4.52333

lrm001=lm(concrete.compressive.strength~age+fly.ash,data=cont\_data)

summary(lrm001)

Text

Description automatically generated

The above multiple regression model has independent variables age and fly.ash where the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength = -0.02310\*fly.ash +0.10474\*age 31.99912

**Code:**

lrm0001=lm(concrete.compressive.strength~cement+fly.ash,data=cont\_data)

summary(lrm0001)

Text

Description automatically generated

**Analysis**: The above multiple regression model has independent variables cement and fly.ash where the dependent variable is concrete compressive strength

**Analysis**: Here regression equation is : concrete.compressive.strength = 0.006421\*fly.ash +0.094608\*cement+ 7.931994

lrm1=lm(concrete.compressive.strength~cement+blast.furnace.slag+fly.ash,data=cont\_data)

summary(lrm1)

Text

Description automatically generated

The above multiple regression model has independent variables cement ,blast.furnace.slag and fly.ash where the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength = 0.7017\*fly.ash +0.12413\*cement+ 0.08883\*blast.furnace.slag-10.59857

**Code:**

lrm2=lm(concrete.compressive.strength~age+I(age^2),data=cont\_data)

summary(lrm2)

Text

Description automatically generated

**Analysis:** The above multiple regression model has independent variable age where the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength = 21.9317815+0.4773506\*age-0.0011773\*age^2

lrm2=lm(concrete.compressive.strength~blast.furnace.slag+I(blast.furnace.slag^2),data=cont\_data)

summary(lrm2)

Text

Description automatically generated

The above multiple regression model has independent variable blast.furnace.slag where the dependent variable is concrete compressive strength

Here regression equation is : concrete.compressive.strength = 32.6240164+0.1236214\*blast.furnace.slag-0.0005674\*blast.furnace.slag^2

**Code:**

lrm3=lm(concrete.compressive.strength~blast.furnace.slag+age+I(cement\*blast.furnace.slag),data=cont\_data)

summary(lrm3)

Text

Description automatically generated

**Analysis:**

The above multiple linear model has independent variables which are variables blast.furnace.slag , age and cement where all the variables have the significant effect on the concrete compressive strength where the p values are 1.20e-07,1.57e-05 and 5.84e-10 which are less than 0.05 hence we can reject the null hypothesis and also the value of Multiple R-squared is 0.4204(42%) which signifies the best fit model. Also the residual error is minimum in this respective model.

Here regression equation is : concrete.compressive.strength = 0.1055549\*age-0.1835900\*blast.furnace.slag+0.0008784\*I(cement\*blast.furnace.slag)+28.4579054

This model says that concrete strength starts changing from the intercept value i.e. 28.4579054 and increase by 0.10555 times of age and decrease by 0.1835900 times the blast furnace and increase by 0.0008 times of the multiplication of cement and blast furnace slag.

Hence from the above results we can say that the above model is best fit model compared to all the other models.

Regression assumptions for linearity, independence of errors, normality of errors, and equality of error variances (L.I.N.E.)

**Code:**

#Linearity

plot(cont\_data$concrete.compressive.strength, lrm3$fitted.values,pch=19,

main="Actual strength values and Fitted values")

abline(0,1,lwd=3,col="red")

Chart, scatter chart

Description automatically generated

**Analysis**:

From the above linearity plot it shows that the values are on the either side and close to the fitted values(predicted) and the line represents the values that are equal (actual = predicted). Hence there is a linear relationship.

#Normality

qqnorm(cont\_data$concrete.compressive.strength,pch=19,main="qq plot")

qqline(cont\_data$concrete.compressive.strength,col="red",lwd=3)

Chart, line chart

Description automatically generated

#OR

hist(lrm3$residuals,col="red",

main="Histogram of Residuals")

Chart, histogram

Description automatically generated

**Analysis**:

From the above plots we can say that the error value between the actual and predicted and they are normal and equally distributed over the abline.

#Equality

plot(cont\_data$concrete.compressive.strength,rstandard(lrm3),ylim=c(-4,4),

pch=19,main=" Equality Plot")

abline(0,0,col="red",lwd=3)

Chart, scatter chart

Description automatically generated

**Analysis**:

For the predicted values, the residual errors are both positive and negative which in turn gives us a low error while adding them all, the residuals are standardized and found to be on the either side of the line, hence we can say it is equally distributed.

The two types of prediction confidence intervals resulting from independent variable values

**Code**:

pdata=data.frame(cement=342,blast.furnace.slag=38,age=180)

pdata

predict(lrm3,pdata,interval="predict")

predict(lrm3,pdata,interval="confidence")

Text

Description automatically generated with medium confidence

**Analysis**:

From the above result for cement component quantitative value which is 342, blast furnace slag value which is 38 and the age value which is of 180, we get concrete compressive strength fit value is 51.89703 which is considered to be the best fit value. The concrete compressive strength value ranges from 44.6814 and 59.10993 which is the confidence interval range from the above result