MIDDLE SEMESTER EXAMINATION REPORT
INTRODUCTION TO DATA SCIENCE

Group 3

Luis Pandiagan	2108541051
Ragil Mona Bianca	2108541052
Sariwati Sagala	2108541055
Mikael Triatama	2108541058
Sinsin Triana Q	2108541077
Cokorda Istri Sintya Dwi Mirah C	2208541047
Ni Wayan Rusniati	2208541071

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

UDAYANA UNIVERSITY

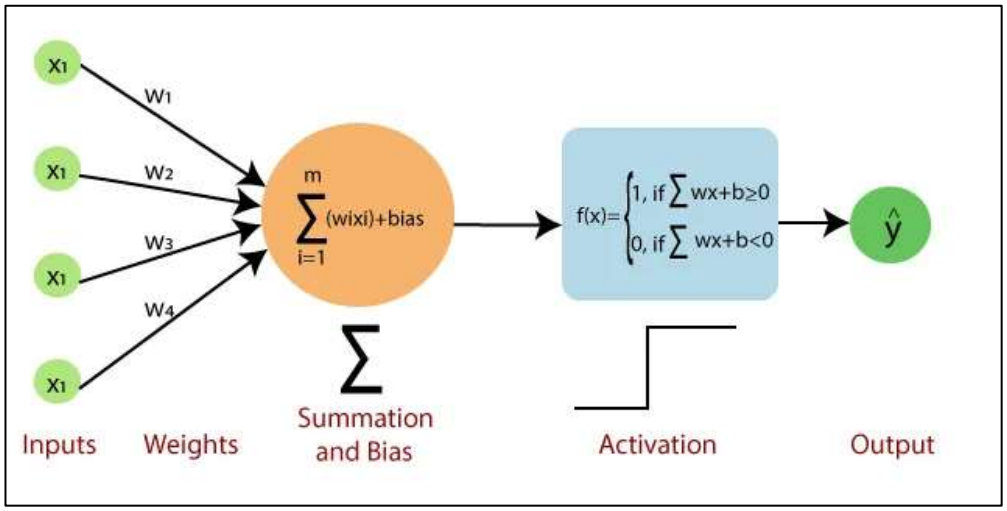
2023

Number 7

Look at Price in the Diamond.xlsx file. If Price = f(Carat, Cut, Color, Clarity), create a perceptron architecture to estimate Price. Determine the predictor/explanatory attributes that have the greatest and least influence on Price.

Answer

Preceptor architecture for estimating Price.



Explanation:

X1 = input Carat value with W_1 weight

X2 = input Cut value with W_2 weight

X3 = input Color value with W_3 weight

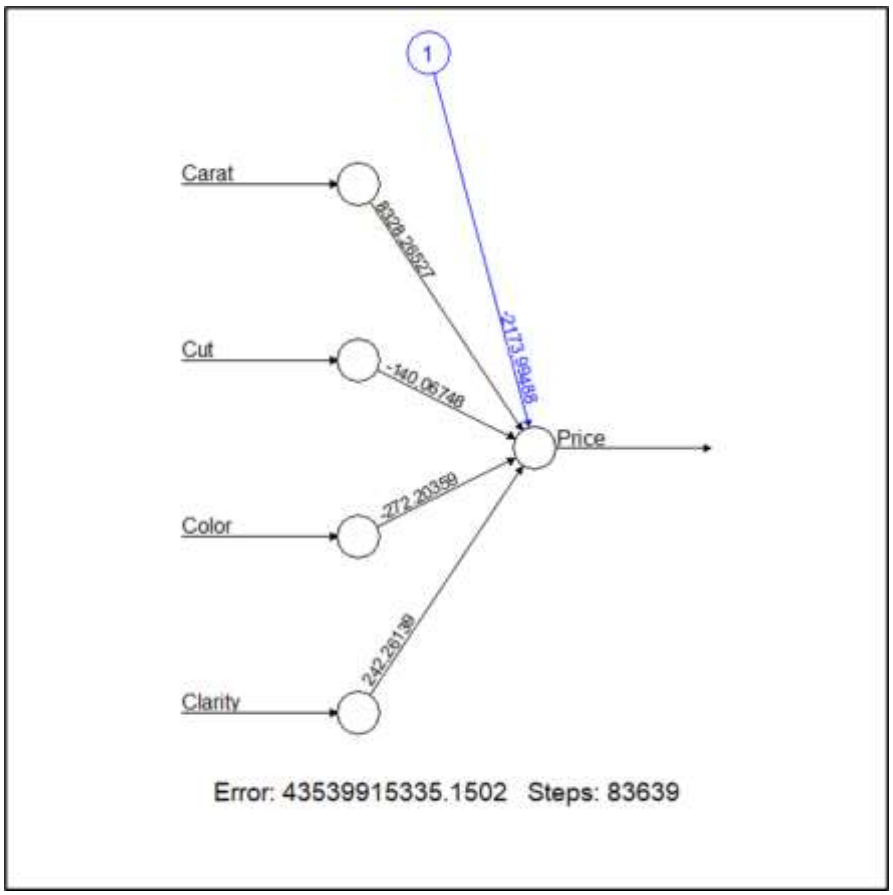
X4 = input Clarity value with W_4 weight

b = bias value

f = activation function

y = Price value output

By using the help of R-studio software obtained:



From the plot above it is known that the inputs from the layers are carat, cut, color, and clarity. Then the input goes into the hidden layer which is used as much as 1 to extract and remember useful features. Furthermore, the extract results are in the form of input patterns which are then used to predict network results or output layer values. In this case the output produced by Price. The error is the loss value obtained when training, which is 43539915335. Then steps are the number of iterations, so the total of all iterations is 83639 so that the weight results are as shown in the picture.

Then, to determine the predictor/explanatory attributes that have the greatest and smallest influence on Price also using the help of R-studio software, the following output is produced:

```
Call:
lm(formula = Price ~ Carat + Cut + Color + Clarity, data = data)
Residuals:
     Min       1Q   Median       3Q      Max
-20623.8   -663.5   -121.6    515.8  10801.8

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -2185.894     22.739  -96.13   <2e-16 ***
Carat        8323.820     13.962  596.17   <2e-16 ***
Cut          -137.449      4.715  -29.15   <2e-16 ***
Color        -270.836      3.751  -72.20   <2e-16 ***
Clarity       244.067      3.371   72.40   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1412 on 53935 degrees of freedom
Multiple R-squared:  0.8747,     Adjusted R-squared:  0.8747
F-statistic: 9.412e+04 on 4 and 53935 DF,  p-value: < 2.2e-16

> # Menampilkan koefisien
> coefficients <- coef(model)
> print(coefficients)
(Intercept)       Carat         Cut       Color     Clarity
 -2185.8935   8323.8198   -137.4494   -270.8356    244.0672

> # Menampilkan pengaruh terbesar dan terkecil
> max_coefficient <- names(coefficients)[which.max(coefficients)]
> min_coefficient <- names(coefficients)[which.min(coefficients)]

> print(paste("Atribut dengan pengaruh terbesar:", max_coefficient))
[1] "Atribut dengan pengaruh terbesar: Carat"

> print(paste("Atribut dengan pengaruh terkecil:", min_coefficient))
[1] "Atribut dengan pengaruh terkecil: (Intercept)"
```

It is obtained from the output results that the regression model for determining Price using the Carat, Cut, Color and Clarity predictors is:

```
Price = -2185.8935 + 8323.8198 Carat - 137.4494 Cut - 270.8356 Color +
244.0672 Clarity
```

The attribute with the greatest or smallest effect can be determined by looking at the value of the coefficients in the regression model.

The coefficient with the smallest value or it can be said that the predictor attribute that has the smallest effect is Color with the smallest coefficient value -270.8356.

Then the coefficient with the largest value or it can be said that the predictor attribute that has the greatest influence is Carat with the largest coefficient value, namely 8323.8198.