|  |
| --- |
| # -\*- coding: utf-8 -\*- |
|  | """ |
|  | Created on Thu Aug 29 04:47:53 2019 |
|  |  |
|  | @author: Rammohan |
|  | """ |
|  |  |
|  | #import required libraries |
|  | import numpy as np |
|  | import pandas as pd |
|  | import matplotlib.pyplot as plt |
|  | import seaborn as sns |
|  |  |
|  | #Read the dataset/Data Preprocessing |
|  | data=pd.read\_csv('concrete\_data.csv') |
|  | data |
|  |  |
|  | data.head() |
|  |  |
|  | #checking null values in the dataset |
|  | data.isnull().any() |
|  |  |
|  | x=data.iloc[:,:8] |
|  | x=data.iloc[:,:8].values |
|  | x |
|  |  |
|  | x.ndim |
|  |  |
|  | y=data.iloc[:,-1:] |
|  | y |
|  | y= data.iloc[:,-1:].values |
|  | y |
|  | #split the data intlo train and test |
|  | from sklearn.model\_selection import train\_test\_split |
|  | x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2,random\_state=0) |
|  |  |
|  | x\_test |
|  | #plot the variance between parameters |
|  | plt.hist((data.concrete\_compressive\_strength)) |
|  |  |
|  | data.hist() |
|  | #correlation between the variables |
|  | data.corr() |
|  | sns.heatmap(data) |
|  | #model |
|  | from sklearn.linear\_model import LinearRegression |
|  | mr=LinearRegression() |
|  | mr.fit(x\_train,y\_train) |
|  |  |
|  | y\_predict =mr.predict(x\_test) |
|  | y\_predict |
|  |  |
|  |  |
|  | import pickle |
|  | pickle.dump(mr,open('strength.pkl','wb')) |
|  | model=pickle.load(open('strength.pkl','rb')) |