

## da-ex-12-linear-regression

June 4, 2023

```
[61]: from pandas import read_csv
      from sklearn.model_selection import train_test_split
```

```
[62]: def load_dataset(fname):
      data=read_csv(fname)
      data=data.dropna(axis=1)
      dataset=data.values
      X=dataset[:, :-1]
      y=dataset[:, -1]
      return X,y
```

```
[63]: X.shape
```

```
[63]: (1465, 3)
```

```
[64]: y.shape
```

```
[64]: (1465,)
```

```
[65]: X,y=load_dataset('/content/drive/MyDrive/Project/water_potability.csv')
      X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.
      ↪33,random_state=1)
      print('Train',X_train.shape,y_train.shape)
      print('Test',X_test.shape,y_test.shape)
```

```
Train (2194, 6) (2194,)
Test (1082, 6) (1082,)
```

```
[66]: from sklearn.linear_model import LinearRegression
```

```
[67]: model=LinearRegression()
      model.fit(X_train,y_train)
```

```
[67]: LinearRegression()
```

```
[68]: rsq=model.score(X_train,y_train)
      print("R-square value",rsq)
```

R-square value 0.004077994896011683

```
[69]: print('w0',model.intercept_)  
      print('w1',model.coef_)
```

```
w0 0.40899213085523084  
w1 [-4.82225198e-04  2.12930240e-06  3.80394421e-03 -6.73621758e-05  
    -2.95479968e-03  1.69575575e-02]
```

```
[70]: ypred=model.predict(X_test)
```

```
[71]: from sklearn.metrics import mean_squared_error,mean_absolute_error  
      import math
```

```
[72]: MSE=mean_squared_error(y_test,ypred)  
      RMSE=math.sqrt(MSE)  
      print('Mean squared Error',MSE)  
      print('Root Mean squared Error',RMSE)
```

```
Mean squared Error 0.24140257122617667  
Root Mean squared Error 0.4913273564805614
```

```
[73]: MAE=mean_absolute_error(y_test,ypred)  
      print('Mean Absolute Error',MAE)
```

```
Mean Absolute Error 0.4770692383332409
```