

Linear Regression Algorithm

Libraries

```
In [38]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

Import dataset

```
In [39]: dataset = pd.read_csv("F:/Data set/insurance.csv")
```

Encoding and Clean the data

```
In [40]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1338 entries, 0 to 1337
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   age        1338 non-null    int64
 1   sex        1338 non-null    object
 2   bmi        1338 non-null    float64
 3   children   1338 non-null    int64
 4   smoker     1338 non-null    object
 5   region     1338 non-null    object
 6   charges    1338 non-null    float64
dtypes: float64(2), int64(2), object(3)
memory usage: 73.3+ KB
```

```
In [41]: ct=ColumnTransformer(transformers=[('encoder',OneHotEncoder(),[1])],remainder='passthrough')
dataset.sex=np.array(ct.fit_transform(dataset))

ct=ColumnTransformer(transformers=[('encoder',OneHotEncoder(),[4])],remainder='passthrough')
dataset.smoker=np.array(ct.fit_transform(dataset))
```

Dataset locating

```
In [42]: x=dataset.iloc[:,0:5].values
y=dataset.iloc[:, -1].values
```

Splitting the dataset into the Training set and Test set

```
In [43]: x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.30,random_state=0)
```

Training the Linear Regression model on the training set

```
In [44]: regressor = LinearRegression()
regressor.fit(x_train,y_train)
```

```
Out[44]: LinearRegression()
```

Predicting the test set results

```
In [45]: y_pred = regressor.predict(x_test)
```

Score values

```
In [46]: regressor.score(x_train,y_train)
```

```
Out[46]: 0.730045447921334
```

```
In [47]: regressor.score(x_test,y_test)
```

```
Out[47]: 0.789479034986701
```

New observation data

```
In [48]: regressor.predict([[19,1,27,0,0]])
```

```
Out[48]: array([24928.25623771])
```