

feature-scaling

October 4, 2024

```
[11]: import pandas as pd
      from scipy import stats
      import numpy as np
      df=pd.read_csv("/content/bmi.csv")
      df
```

```
[11]:
```

	Gender	Height	Weight	Index
0	Male	174	96	4
1	Male	189	87	2
2	Female	185	110	4
3	Female	195	104	3
4	Male	149	61	3
..
495	Female	150	153	5
496	Female	184	121	4
497	Female	141	136	5
498	Male	150	95	5
499	Male	173	131	5

[500 rows x 4 columns]

```
[12]: df.dropna()
```

```
[12]:
```

	Gender	Height	Weight	Index
0	Male	174	96	4
1	Male	189	87	2
2	Female	185	110	4
3	Female	195	104	3
4	Male	149	61	3
..
495	Female	150	153	5
496	Female	184	121	4
497	Female	141	136	5
498	Male	150	95	5
499	Male	173	131	5

[500 rows x 4 columns]

```
[13]: max_val=np.max(np.abs(df['Height']))
max_val
```

```
[13]: 199
```

```
[14]: max_val1=np.max(np.abs(df['Weight']))
max_val1
```

```
[14]: 160
```

```
[15]: from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
df[['Height','Weight']]=sc.fit_transform(df[['Height','Weight']])
df.head(10)
```

```
[15]:
```

	Gender	Height	Weight	Index
0	Male	0.247939	-0.309117	4
1	Male	1.164872	-0.587322	2
2	Female	0.920357	0.123647	4
3	Female	1.531645	-0.061823	3
4	Male	-1.280283	-1.391027	3
5	Male	1.164872	-0.061823	3
6	Male	-1.402541	-0.432764	5
7	Male	-0.974639	0.154559	5
8	Male	0.247939	-0.494587	3
9	Female	-0.057706	-0.092735	4

```
[16]: from sklearn.preprocessing import MinMaxScaler
scaler=MinMaxScaler()
df[['Height','Weight']]=scaler.fit_transform(df[['Height','Weight']])
```

```
[17]: df.head(10)
```

```
[17]:
```

	Gender	Height	Weight	Index
0	Male	0.576271	0.418182	4
1	Male	0.830508	0.336364	2
2	Female	0.762712	0.545455	4
3	Female	0.932203	0.490909	3
4	Male	0.152542	0.100000	3
5	Male	0.830508	0.490909	3
6	Male	0.118644	0.381818	5
7	Male	0.237288	0.554545	5
8	Male	0.576271	0.363636	3
9	Female	0.491525	0.481818	4

```
[21]: df2=pd.read_csv("/content/bmi.csv")
```

```
[22]: from sklearn.preprocessing import Normalizer
scaler=Normalizer()
df2[['Height','Weight']]=scaler.fit_transform(df2[['Height','Weight']])
df2
```

```
[22]:
```

	Gender	Height	Weight	Index
0	Male	0.875578	0.483077	4
1	Male	0.908381	0.418144	2
2	Female	0.859536	0.511075	4
3	Female	0.882353	0.470588	3
4	Male	0.925448	0.378875	3
..
495	Female	0.700071	0.714073	5
496	Female	0.835527	0.549450	4
497	Female	0.719753	0.694230	5
498	Male	0.844819	0.535052	5
499	Male	0.797227	0.603680	5

[500 rows x 4 columns]

```
[23]: df3=pd.read_csv("/content/bmi.csv")
```

```
[24]: from sklearn.preprocessing import MaxAbsScaler
scaler=MaxAbsScaler()
df3[['Height','Weight']]=scaler.fit_transform(df3[['Height','Weight']])
df3
```

```
[24]:
```

	Gender	Height	Weight	Index
0	Male	0.874372	0.60000	4
1	Male	0.949749	0.54375	2
2	Female	0.929648	0.68750	4
3	Female	0.979899	0.65000	3
4	Male	0.748744	0.38125	3
..
495	Female	0.753769	0.95625	5
496	Female	0.924623	0.75625	4
497	Female	0.708543	0.85000	5
498	Male	0.753769	0.59375	5
499	Male	0.869347	0.81875	5

[500 rows x 4 columns]

```
[25]: df4=pd.read_csv("/content/bmi.csv")
```

```
[26]: from sklearn.preprocessing import RobustScaler
scaler=RobustScaler()
df4[['Height','Weight']]=scaler.fit_transform(df4[['Height','Weight']])
```

```
df4
```

```
[26]:
```

	Gender	Height	Weight	Index
0	Male	0.125000	-0.178571	4
1	Male	0.660714	-0.339286	2
2	Female	0.517857	0.071429	4
3	Female	0.875000	-0.035714	3
4	Male	-0.767857	-0.803571	3
..
495	Female	-0.732143	0.839286	5
496	Female	0.482143	0.267857	4
497	Female	-1.053571	0.535714	5
498	Male	-0.732143	-0.196429	5
499	Male	0.089286	0.446429	5

```
[500 rows x 4 columns]
```