ML 7 - Normalization By Virat Tiwari

December 1, 2023

1 NORMALIZATION

In normalization we scale down the values of dataset between the range of 0 to 1 , whatever data is it , we just scale that data in between 0 to 1

[1]:
```
# we import "seaborn as sns"  for getting or importing dataset

import seaborn as sns
```

[4]:
```
# Here we load the data ansd store that data in df

df=sns.load_dataset("taxis")
```

[6]:
```
# . herad ( ) function is used for accessing intial 5 datapoints

df.head()
```

[6]:

	pickup	dropoff	passengers	distance	fare	tip	tolls	total	color	payment	pickup_zone	dropoff_zone	pickup_borough	dropoff_borough
0	2019-03-23 20:21:09	2019-03-23 20:27:24	1	1.60	7.0	2.15	0.0	12.95	yellow	credit card	Lenox Hill West	UN/Turtle Bay South	Manhattan	Manhattan
1	2019-03-04 16:11:55	2019-03-04 16:19:00	1	0.79	5.0	0.00	0.0	9.30	yellow	cash	Upper West Side South	Upper West Side South	Manhattan	Manhattan
2	2019-03-27 17:53:01	2019-03-27 18:00:25	1	1.37	7.5	2.36	0.0	14.16	yellow	credit card	Alphabet City	West Village	Manhattan	Manhattan
3	2019-03-10 01:23:59	2019-03-10 01:49:51	1	7.70	27.0	6.15	0.0	36.95	yellow	credit card	Hudson Sq	Yorkville West	Manhattan	Manhattan
4	2019-03-30 13:27:42	2019-03-30 13:37:14	3	2.16	9.0	1.10	0.0	13.40	yellow	credit card	Midtown East	Yorkville West	Manhattan	Manhattan

[14]:
```
# Here we done the "normalization"
# for doing the normalization we should first import MInMazScaler through the␣
 ↪sklearn.preprocessing

from sklearn.preprocessing import MinMaxScaler
```

[15]:
```
# We store the MinMaxScaler in min_max variable

min_max=MinMaxScaler()
```

[17]:
```
# Here we take 3 features for fit un MinMaxScaler

min_max.fit(df[["distance","fare","tip"]])
```

[17]:
```
MinMaxScaler()
```

[18]:
```
# Finally we transform the data into the normalization

min_max.transform(df[["distance","fare","tip"]])
```

[18]:
```
array([[0.04359673, 0.04026846, 0.06475904],
       [0.02152589, 0.02684564, 0.        ],
       [0.0373297 , 0.04362416, 0.07108434],
       …,
       [0.11280654, 0.10067114, 0.        ],
       [0.03051771, 0.03355705, 0.        ],
       [0.10490463, 0.09395973, 0.10120482]])
```

[19]:
```
min_max.transform([[1,3,4]])
```

```
/opt/conda/lib/python3.10/site-packages/sklearn/base.py:409: UserWarning: X does
not have valid feature names, but MinMaxScaler was fitted with feature names
  warnings.warn(
```

[19]:
```
array([[0.02724796, 0.01342282, 0.12048193]])
```

THANK YOU SO MUCH !!

YOURS VIRAT TIWARI :)