Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

Import Packages and Connect to Drive

In [1]:
```
!python --version
```

```
Python 3.9.13
```

In [2]:
```
import os
os.getcwd()
```

Out[2]:
```
'/home/wsuser/work'
```

In [6]:
```
import os
import IPython.display as ipd
import matplotlib.pyplot as plt#graph plotting
import numpy as np#Array or Multi dimi. arrays
import pandas as pd#Reading files and sving it as tablur format
import time
from collections import Counter, OrderedDict#[1,2,3,3]
import cv2#opencv - python - Image processing- short form -cv2
#from google.colab.patches import cv2_imshow#Is used for printing the images
import random#Generating random variables or values
#import seaborn as sns#graph plotting
import json
import joblib
import pickle
from ibm_watson_machine_learning import APIClient
```

In [7]:
```
import sklearn
print('The scikit-learn version is {}.'.format(sklearn.__version__))
```

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search | EN | 28°C Haze | 10:33 18-11-2022

Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

Error

In [7]:
```
import sklearn
print('The scikit-learn version is {}.'.format(sklearn.__version__))
```

```
The scikit-learn version is 1.0.2.
```

In [8]:
```
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import train_test_split
from sklearn import metrics
from sklearn.utils import class_weight
from sklearn.utils.class_weight import compute_class_weight
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import cross_val_score
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.pipeline import make_pipeline
from sklearn.linear_model import Ridge, Lasso, ElasticNet
from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor
from sklearn.model_selection import GridSearchCV
from sklearn.exceptions import NotFittedError
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import StratifiedKFold
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import SGDClassifier
from sklearn import metrics
```

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search

EN 28°C Haze 10:33 18-11-2022

Data Loading and Preprocessing

In [10]:
```
import os, types
import pandas as pd
from botocore.client import Config
import ibm_boto3

def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
cos_client = ibm_boto3.client(service_name='s3',
    ibm_api_key_id='ge0l8EdyQxhVAqykYTv3jPIjXaicgiET0P6yZji-ubLg',
    ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.eu.cloud-object-storage.appdomain.cloud')
```

In [11]:
```
bucket = 'rainfall-donotdelete-pr-eoii6hshulobg1'
object_key = 'rainfall.csv'

body = cos_client.get_object(Bucket=bucket,Key=object_key)['Body']

# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __iter__, body )

df = pd.read_csv(body)
df
```

Out[11]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	...	Humidity3pm	Pressure9am	Pressure3pm	Cloud9
0	2008-12-01	Albury	13.4	22.9	0.6	NaN	NaN	W	44.0	W	...	22.0	1007.7	1007.1	

Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

Out[11]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	...	Humidity3pm	Pressure9am	Pressure3pm	Cloud9
0	2008-12-01	Albury	13.4	22.9	0.6	NaN	NaN	W	44.0	W	...	22.0	1007.7	1007.1	
1	2008-12-02	Albury	7.4	25.1	0.0	NaN	NaN	WNW	44.0	NNW	...	25.0	1010.6	1007.8	N
2	2008-12-03	Albury	12.9	25.7	0.0	NaN	NaN	WSW	46.0	W	...	30.0	1007.6	1008.7	N
3	2008-12-04	Albury	9.2	28.0	0.0	NaN	NaN	NE	24.0	SE	...	16.0	1017.6	1012.8	N
4	2008-12-05	Albury	17.5	32.3	1.0	NaN	NaN	W	41.0	ENE	...	33.0	1010.8	1006.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
142188	2017-06-20	Uluru	3.5	21.8	0.0	NaN	NaN	E	31.0	ESE	...	27.0	1024.7	1021.2	N
142189	2017-06-21	Uluru	2.8	23.4	0.0	NaN	NaN	E	31.0	SE	...	24.0	1024.6	1020.3	N
142190	2017-06-22	Uluru	3.6	25.3	0.0	NaN	NaN	NNW	22.0	SE	...	21.0	1023.5	1019.1	N
142191	2017-06-23	Uluru	5.4	26.9	0.0	NaN	NaN	N	37.0	SE	...	24.0	1021.0	1016.8	N
142192	2017-06-24	Uluru	7.8	27.0	0.0	NaN	NaN	SE	28.0	SSE	...	24.0	1019.4	1016.5	

142193 rows × 24 columns

In [12]: df = df.dropna()

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search

EN 28°C Haze 10:33 18-11-2022

Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

Error

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

Out[12]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	...	Humidity3pm	Pressure9am	Pressure3pm	Cloud9
5939	2009-01-01	Cobar	17.9	35.2	0.0	12.0	12.3	SSW	48.0	ENE	...	13.0	1006.3	1004.4	
5940	2009-01-02	Cobar	18.4	28.9	0.0	14.8	13.0	S	37.0	SSE	...	8.0	1012.9	1012.1	
5942	2009-01-04	Cobar	19.4	37.6	0.0	10.8	10.6	NNE	46.0	NNE	...	22.0	1012.3	1009.2	
5943	2009-01-05	Cobar	21.9	38.4	0.0	11.4	12.2	WNW	31.0	WNW	...	22.0	1012.7	1009.1	
5944	2009-01-06	Cobar	24.2	41.0	0.0	11.2	8.4	WNW	35.0	NW	...	15.0	1010.7	1007.4	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
139108	2017-06-20	Darwin	19.3	33.4	0.0	6.0	11.0	ENE	35.0	SE	...	32.0	1013.9	1010.5	
139109	2017-06-21	Darwin	21.2	32.6	0.0	7.6	8.6	E	37.0	SE	...	28.0	1014.6	1011.2	
139110	2017-06-22	Darwin	20.7	32.8	0.0	5.6	11.0	E	33.0	E	...	23.0	1015.3	1011.8	
139111	2017-06-23	Darwin	19.5	31.8	0.0	6.2	10.6	ESE	26.0	SE	...	58.0	1014.9	1010.7	
139112	2017-06-24	Darwin	20.2	31.7	0.0	5.6	10.7	ENE	30.0	ENE	...	32.0	1013.9	1009.7	

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search

EN 28°C Haze 10:34 18-11-2022

Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

Error

139112	06-24	Darwin	20.2	31.7	0.0	5.6	10.7	ENE	30.0	ENE	...	32.0	1013.9	1009.7

56420 rows × 24 columns

In [13]:

```
for i in df.columns:
  print(i, df[i].dtype)
  print(df[i].value_counts())
  print("################")
```

```
Date object
2010-11-16    25
2009-03-09    25
2009-03-28    25
2009-03-26    25
2010-05-17    25
              ..
2008-03-26     1
2008-03-25     1
2008-03-24     1
2008-03-23     1
2008-01-02     1
Name: Date, Length: 3416, dtype: int64
################
Location object
Darwin              3062
Perth               3025
Brisbane            2953
MelbourneAirport    2929
PerthAirport        2913
SydneyAirport       2870
Watsonia            2730
Mildura             2594
MountGambier        2465
NorfolkIsland       2464
```

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search

EN 28°C Haze 10:34 18-11-2022

```
Cobar               534
Name: Location, dtype: int64
################
MinTemp float64
 13.0    349
 9.6     342
 8.5     333
 9.7     331
 12.5    331
        ...
 31.4      1
-5.3       1
-3.8       1
-5.8       1
-6.7       1
Name: MinTemp, Length: 348, dtype: int64
################
MaxTemp float64
19.6    303
18.2    300
19.4    300
20.1    298
23.5    294
       ...
45.4      1
47.3      1
45.8      1
6.3       1
7.5       1
Name: MaxTemp, Length: 395, dtype: int64
################
Rainfall float64
0.0      36709
0.2       3009
0.4       1593
0.6       1087
0.8        834
```

Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

```
14.2       2
14.5       1
Name: Sunshine, Length: 145, dtype: int64
################
WindGustDir object
E      4516
N      4210
W      4161
SW     4052
ENE    4028
SE     3930
SSW    3898
WSW    3791
S      3636
ESE    3312
SSE    3295
NE     3185
WNW    2989
NW     2612
NNE    2516
NNW    2289
Name: WindGustDir, dtype: int64
################
WindGustSpeed float64
35.0     3965
39.0     3884
37.0     3727
31.0     3582
33.0     3580
         ...
113.0       2
122.0       1
9.0         1
124.0       1
111.0       1
Name: WindGustSpeed, Length: 61, dtype: int64
################
```

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search | EN | 28°C Haze | 10:34 18-11-2022

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['Date'] = pd.to_datetime(df['Date'],
```

Out[19]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	...	Humidity3pm	Pressure9am	Pressure3pm	Cloud9
5939	2009-01-01	4	17.9	35.2	0.0	12.0	12.3	11	48.0	1	...	13.0	1006.3	1004.4	
5940	2009-01-02	4	18.4	28.9	0.0	14.8	13.0	8	37.0	10	...	8.0	1012.9	1012.1	
5942	2009-01-04	4	19.4	37.6	0.0	10.8	10.6	5	46.0	5	...	22.0	1012.3	1009.2	
5943	2009-01-05	4	21.9	38.4	0.0	11.4	12.2	14	31.0	14	...	22.0	1012.7	1009.1	
5944	2009-01-06	4	24.2	41.0	0.0	11.2	8.4	14	35.0	7	...	15.0	1010.7	1007.4	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
139108	2017-06-20	6	19.3	33.4	0.0	6.0	11.0	1	35.0	9	...	32.0	1013.9	1010.5	
139109	2017-06-21	6	21.2	32.6	0.0	7.6	8.6	0	37.0	9	...	28.0	1014.6	1011.2	
139110	2017-06-22	6	20.7	32.8	0.0	5.6	11.0	0	33.0	0	...	23.0	1015.3	1011.8	
139111	2017-06-23	6	19.5	31.8	0.0	6.2	10.6	2	26.0	9	...	58.0	1014.9	1010.7	
139112	2017-06-24	6	20.2	31.7	0.0	5.6	10.7	1	30.0	1	...	32.0	1013.9	1009.7	

56420 rows × 24 columns

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['Date'] = pd.to_datetime(df['Date'],
```

Out[19]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	...	Humidity3pm	Pressure9am	Pressure3pm	Cloud9
5939	2009-01-01	4	17.9	35.2	0.0	12.0	12.3	11	48.0	1	...	13.0	1006.3	1004.4	
5940	2009-01-02	4	18.4	28.9	0.0	14.8	13.0	8	37.0	10	...	8.0	1012.9	1012.1	
5942	2009-01-04	4	19.4	37.6	0.0	10.8	10.6	5	46.0	5	...	22.0	1012.3	1009.2	
5943	2009-01-05	4	21.9	38.4	0.0	11.4	12.2	14	31.0	14	...	22.0	1012.7	1009.1	
5944	2009-01-06	4	24.2	41.0	0.0	11.2	8.4	14	35.0	7	...	15.0	1010.7	1007.4	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
139108	2017-06-20	6	19.3	33.4	0.0	6.0	11.0	1	35.0	9	...	32.0	1013.9	1010.5	
139109	2017-06-21	6	21.2	32.6	0.0	7.6	8.6	0	37.0	9	...	28.0	1014.6	1011.2	
139110	2017-06-22	6	20.7	32.8	0.0	5.6	11.0	0	33.0	0	...	23.0	1015.3	1011.8	
139111	2017-06-23	6	19.5	31.8	0.0	6.2	10.6	2	26.0	9	...	58.0	1014.9	1010.7	
139112	2017-06-24	6	20.2	31.7	0.0	5.6	10.7	1	30.0	1	...	32.0	1013.9	1009.7	

56420 rows × 24 columns

56420 rows × 24 columns

In [20]:
```
for col in df.columns:
  print(col, df[col].dtype)
```

```
Date datetime64[ns]
Location int64
MinTemp float64
MaxTemp float64
Rainfall float64
Evaporation float64
Sunshine float64
WindGustDir int64
WindGustSpeed float64
WindDir9am int64
WindDir3pm int64
WindSpeed9am float64
WindSpeed3pm float64
Humidity9am float64
Humidity3pm float64
Pressure9am float64
Pressure3pm float64
Cloud9am float64
Cloud3pm float64
Temp9am float64
Temp3pm float64
RainToday int64
RISK_MM float64
RainTomorrow int64
```

In [21]:
```
df['year'] = df['Date'].dt.year
df['month'] = df['Date'].dt.month
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['day'] = df['Date'].dt.day
```

Out[21]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	...	Cloud9am	Cloud3pm	Temp9am	Temp3pm	Rain
5939	2009-01-01	4	17.9	35.2	0.0	12.0	12.3	11	48.0	1	...	2.0	5.0	26.6	33.4	
5940	2009-01-02	4	18.4	28.9	0.0	14.8	13.0	8	37.0	10	...	1.0	1.0	20.3	27.0	
5942	2009-01-04	4	19.4	37.6	0.0	10.8	10.6	5	46.0	5	...	1.0	6.0	28.7	34.9	
5943	2009-01-05	4	21.9	38.4	0.0	11.4	12.2	14	31.0	14	...	1.0	5.0	29.1	35.6	
5944	2009-01-06	4	24.2	41.0	0.0	11.2	8.4	14	35.0	7	...	1.0	6.0	33.6	37.6	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
139108	2017-06-20	6	19.3	33.4	0.0	6.0	11.0	1	35.0	9	...	0.0	1.0	24.5	32.3	
139109	2017-06-21	6	21.2	32.6	0.0	7.6	8.6	0	37.0	9	...	7.0	0.0	24.8	32.0	
139110	2017-06-22	6	20.7	32.8	0.0	5.6	11.0	0	33.0	0	...	0.0	0.0	24.8	32.1	
139111	2017-06-23	6	19.5	31.8	0.0	6.2	10.6	2	26.0	9	...	1.0	1.0	24.8	29.2	
139112	2017-06-24	6	20.2	31.7	0.0	5.6	10.7	1	30.0	1	...	6.0	5.0	25.4	31.0	

56420 rows × 27 columns

56420 rows × 27 columns

In [22]:
```
del df['Date']
```

In [23]:
```
df
```

Out[23]:

	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	WindDir3pm	...	Cloud9am	Cloud3pm	Temp9am	Temp3pr
5939	4	17.9	35.2	0.0	12.0	12.3	11	48.0	1	12	...	2.0	5.0	26.6	33.
5940	4	18.4	28.9	0.0	14.8	13.0	8	37.0	10	10	...	1.0	1.0	20.3	27.
5942	4	19.4	37.6	0.0	10.8	10.6	5	46.0	5	6	...	1.0	6.0	28.7	34.
5943	4	21.9	38.4	0.0	11.4	12.2	14	31.0	14	15	...	1.0	5.0	29.1	35.
5944	4	24.2	41.0	0.0	11.2	8.4	14	35.0	7	14	...	1.0	6.0	33.6	37.
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
139108	6	19.3	33.4	0.0	6.0	11.0	1	35.0	9	4	...	0.0	1.0	24.5	32.
139109	6	21.2	32.6	0.0	7.6	8.6	0	37.0	9	9	...	7.0	0.0	24.8	32.
139110	6	20.7	32.8	0.0	5.6	11.0	0	33.0	0	13	...	0.0	0.0	24.8	32.
139111	6	19.5	31.8	0.0	6.2	10.6	2	26.0	9	6	...	1.0	1.0	24.8	29.
139112	6	20.2	31.7	0.0	5.6	10.7	1	30.0	1	6	...	6.0	5.0	25.4	31.

56420 rows × 26 columns

Welcome to Project! Delight | IBM | IBM-Project-17869-165967 | IBM-Project-45258-166072 | Download file | iLovePDF

github.com/IBM-EPBL/IBM-Project-17869-1659676805/blob/main/MODEL%20BUILDING/model%20elevation.ipynb

Error

In [24]:
```
for col in df.columns:
  print(col, df[col].dtype)
```

```
Location int64
MinTemp float64
MaxTemp float64
Rainfall float64
Evaporation float64
Sunshine float64
WindGustDir int64
WindGustSpeed float64
WindDir9am int64
WindDir3pm int64
WindSpeed9am float64
WindSpeed3pm float64
Humidity9am float64
Humidity3pm float64
Pressure9am float64
Pressure3pm float64
Cloud9am float64
Cloud3pm float64
Temp9am float64
Temp3pm float64
RainToday int64
RISK_MM float64
RainTomorrow int64
year int64
month int64
day int64
```

In [25]:
```
X = df.drop('RainTomorrow', axis=1)
Y = df['RainTomorrow']
```

In [26]:
```
X
```

Screenshot (3)_pag....jpg | Screenshot (3)_pag....jpg | Screenshot (3) (1).pdf | Screenshot (3).pdf | Show all

Type here to search

EN 28°C Haze 10:35 18-11-2022