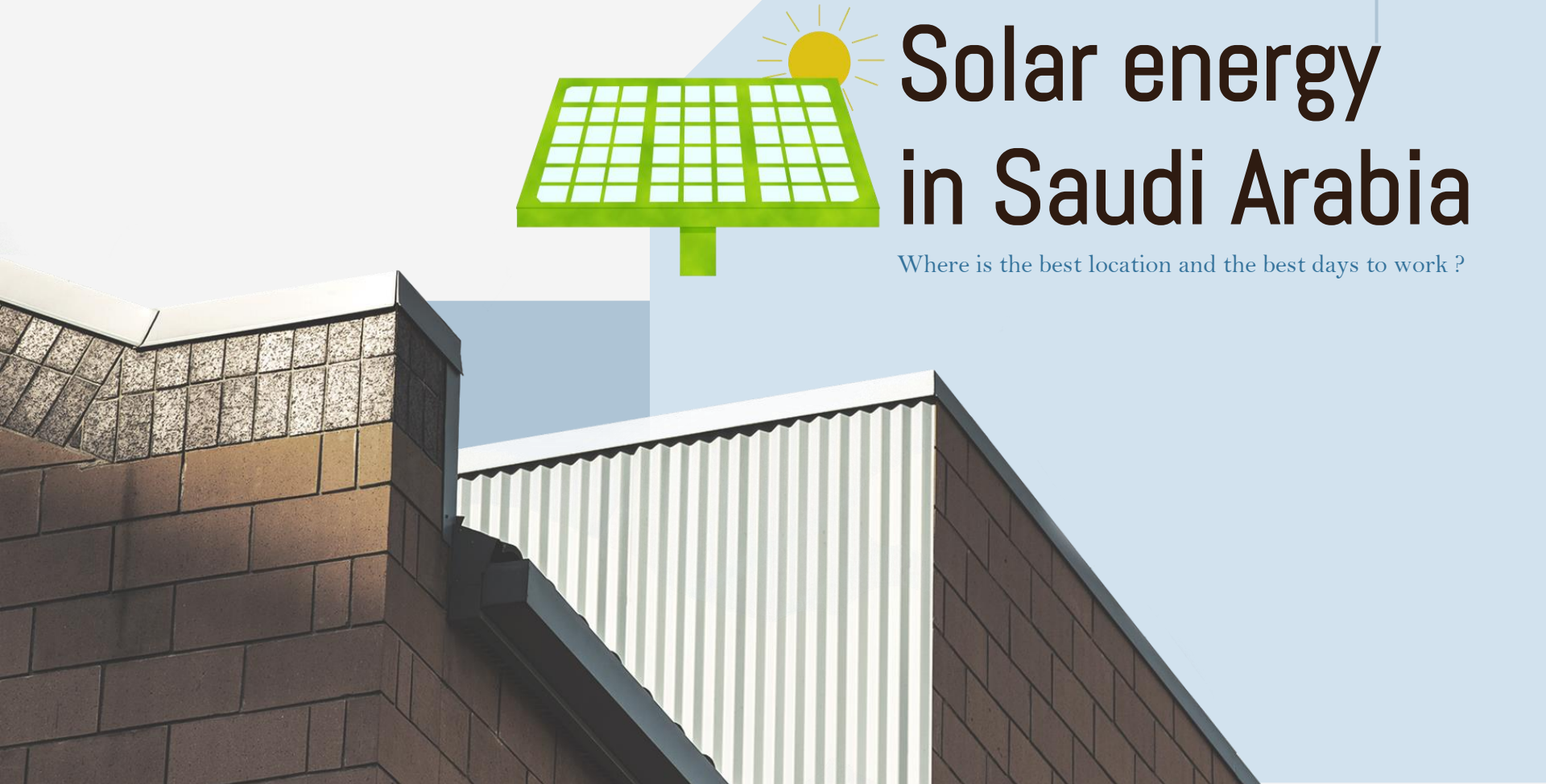


Solar energy in Saudi Arabia

Where is the best location and the best days to work ?

METHODOLOGY

01

INTRODUCTION

The Sun is an extremely powerful energy source, and sunlight is by far the largest source of energy received by Earth, but its intensity at Earth's surface is actually quite low. This is essentially because of the enormous radial spreading of radiation from the distant Sun.

With energy costs and consumption rising dramatically, the need for renewable energy has become urgent. In line with the Kingdom's Vision 2030 plan to diversify energy sources.

02

Load Data

02

Load Data

1

Source

Data Set From Kaggle, which includes the weather history of Saudi Arabia from 2017 -2019

2

Features AND Rows

Consist of 15 columns and 249023 rows

3

Target

Our target is a Weather, Sunny Or Not Sunny

FEATURES

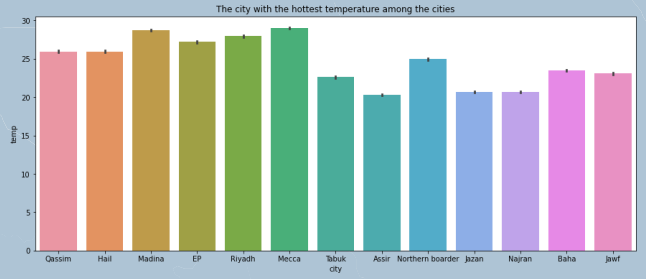
Field Name	Description
City	Name of cities in Saudi Arabia
Date	recorded date
Time	recorded time
Year	recorded date - year
Month	recorded date - month
Day	recorded date - day
Hour	recorded time - hour
Minute	recorded time - minute
Weather	recorded weather description (clear - sunny - Sunny - Cloudy - Rainy - Windy)
Temp	a meteorological code for upper air soundings
Wind	wind speed
Humidity	umidity is the amount of water vapor in the air
Barometer	used to measure air pressure in a certain environment
visibility	the quality or state of being visible, the degree of clearness

03 EDA

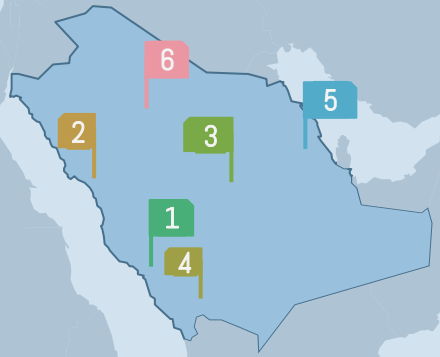
Clean Data


humidity 17
barometer 72
Full null by mean
We have Some Null
01
Drop unnamed: 0
Drop ID Colum
03
02
We Have Duplicates Rows
Drop columns: date, time

The best location we have to solar energy



1
Mecca
8.17%
2
Madina
8.14%
3
Riyadh
8.14%



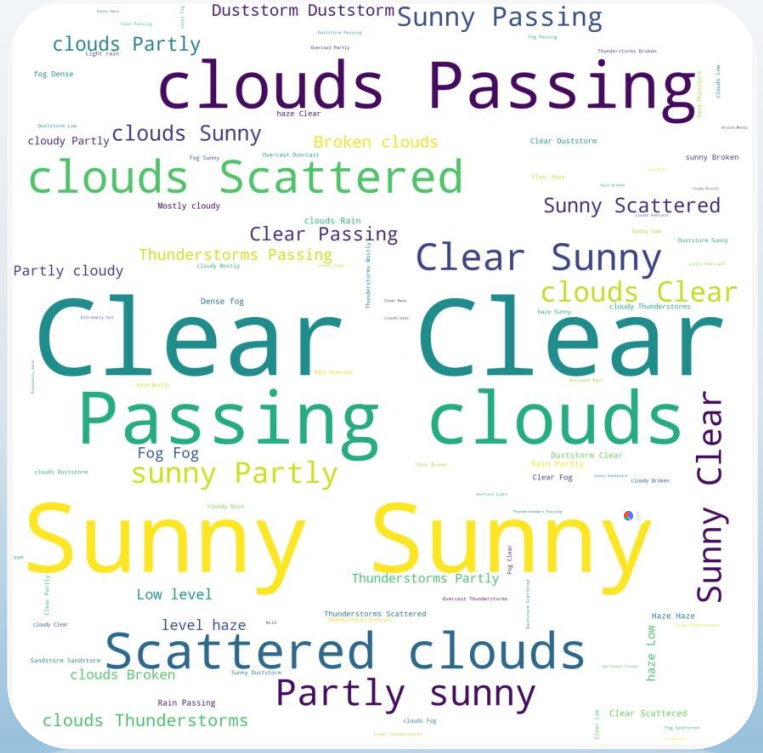
04 Feature Engineering

- Dummy to X & Replace y to 0 – 1

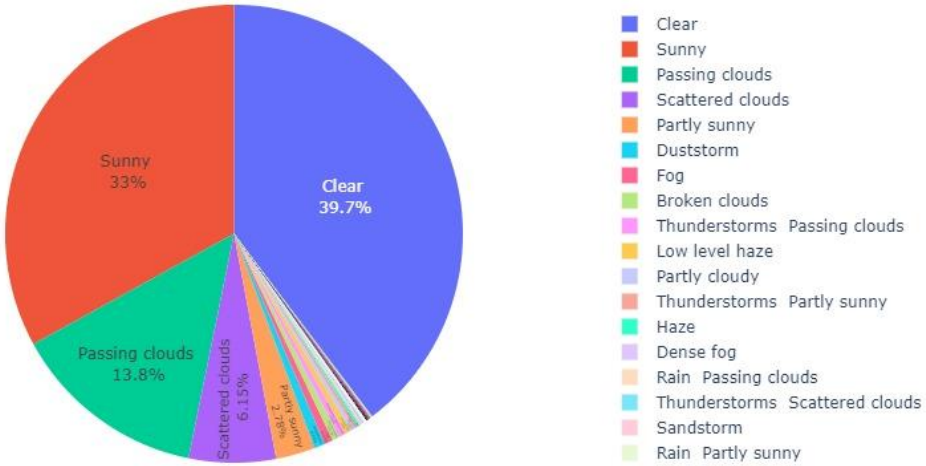
- Replace weather because is divided into 81 names

In Feature ‘ weather ‘
We Did Collecting Weather
Names And Limiting Them To

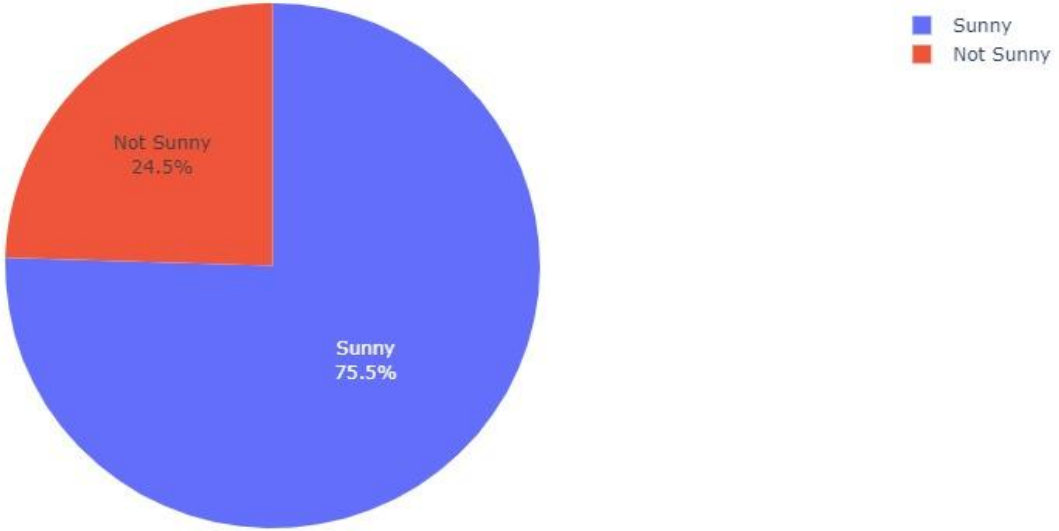
- Sunny
- Not Sunny



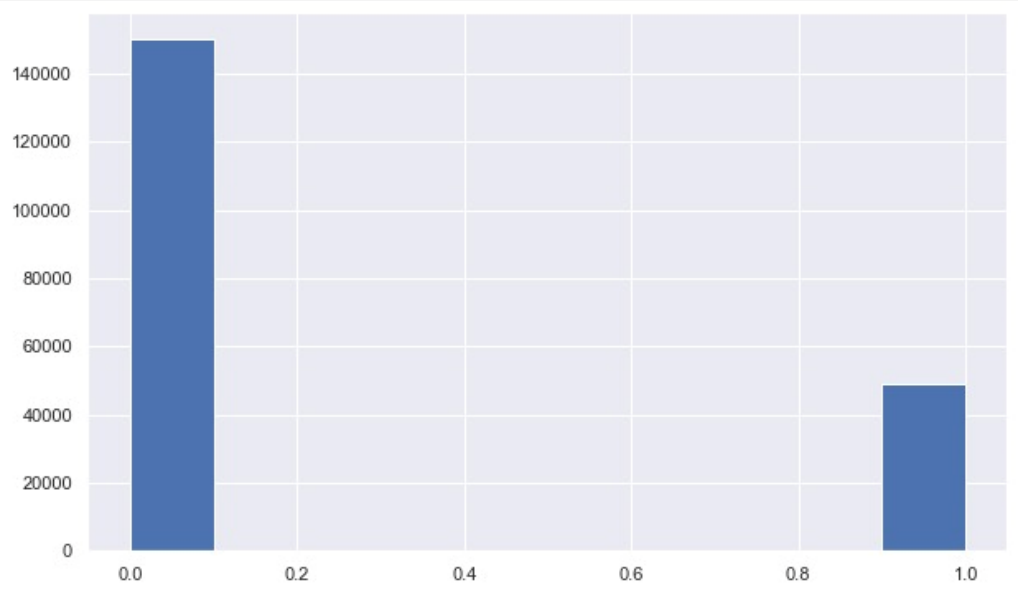
Weather Before Replace



Weather After Replace



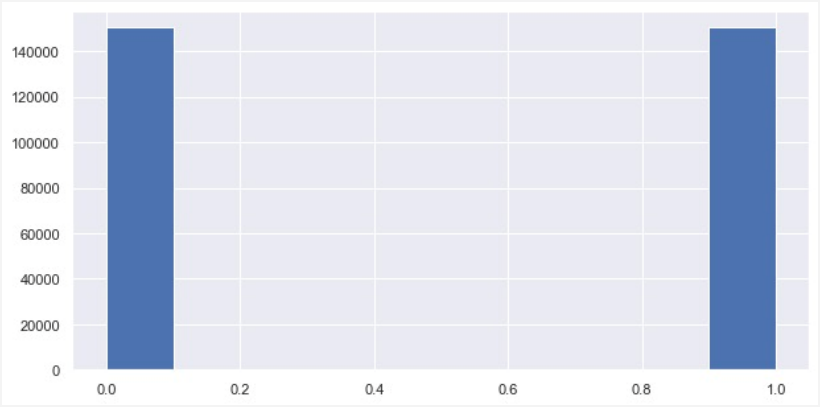
Data is imbalance



explain

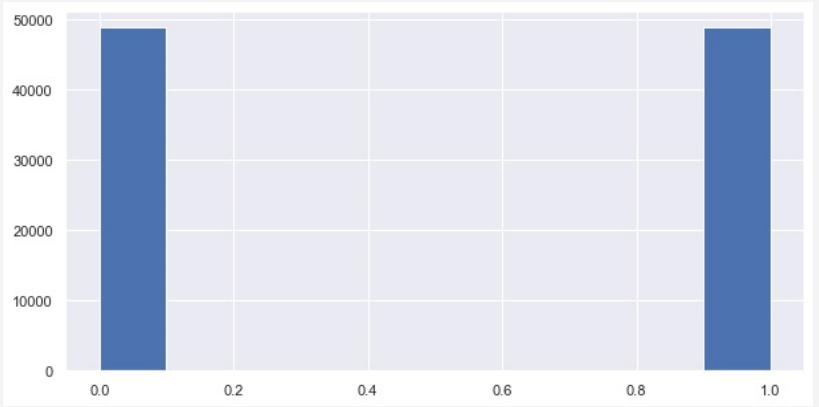
0 is Sunny
1 is Not Sunny

Over and under sampling



Over Sampling

Train: 0.9271
Test : 0.9014
F(1): 0.9004
Precision :0.9014
Recall : 0.8994



Under Sampling

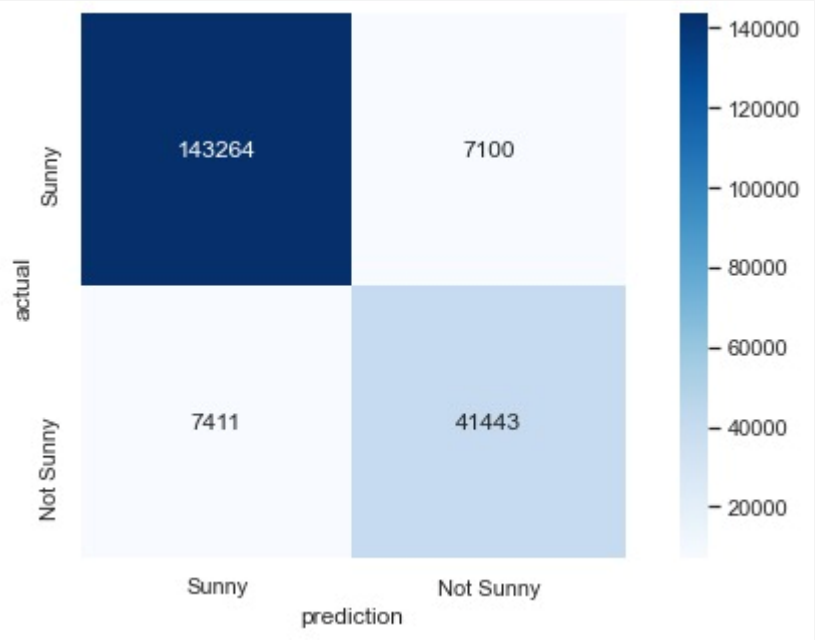
Train: 0.7547
Test : 0.5000
F(1): 0.8981
Precision : 0.8783
Recall : 0.9269

05 Experiments

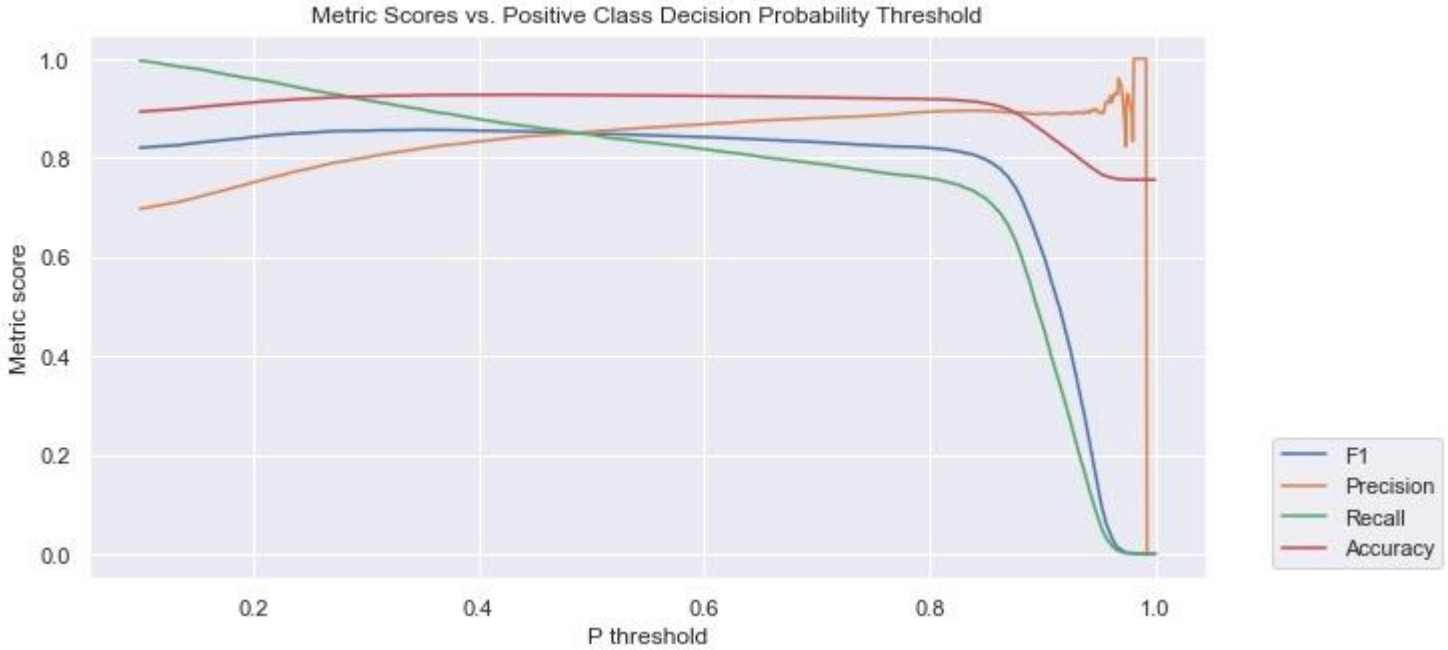
Experiments

models	train	test	F(1)	precision	Recall
Baseline - LR	0.9271	0.9266	0.9004	0.9014	0.8994
Standar Scaler	0.9291	0.9291	0.9037	0.9048	0.9027
MinMax scaler	0.9292	0.9295	0.9035	0.9048	0.9023
MinMax scaler - GSh	0.9291	0.9025	0.9037	0.9048	0.9027
Decision tree	0.9267	0.9266	0.1964	0.6221	0.5000
Decision tree - GSh	0.9941	0.9569	0.1964	0.6221	0.5000
Random Forest	1.0000	0.9688	0.9575	0.9538	0.9613
Random Forest - GSh	0.9999	0.9696	0.9592	0.9553	0.9632

Convution Matrix for Baseline Models



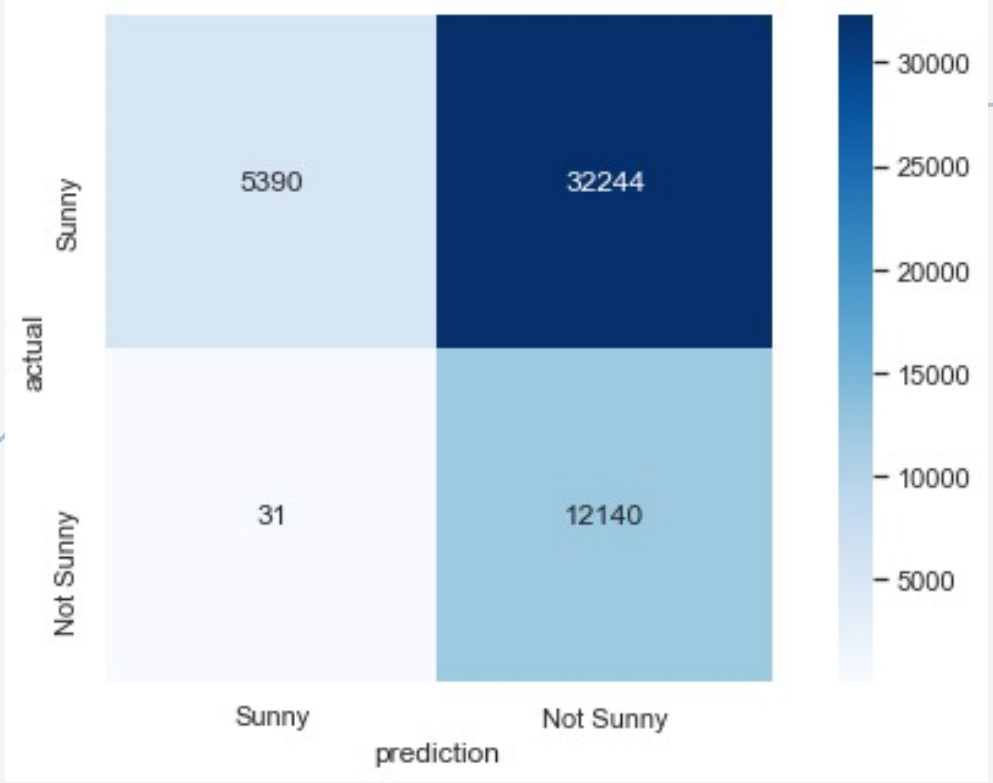
ROC for Baseline Model



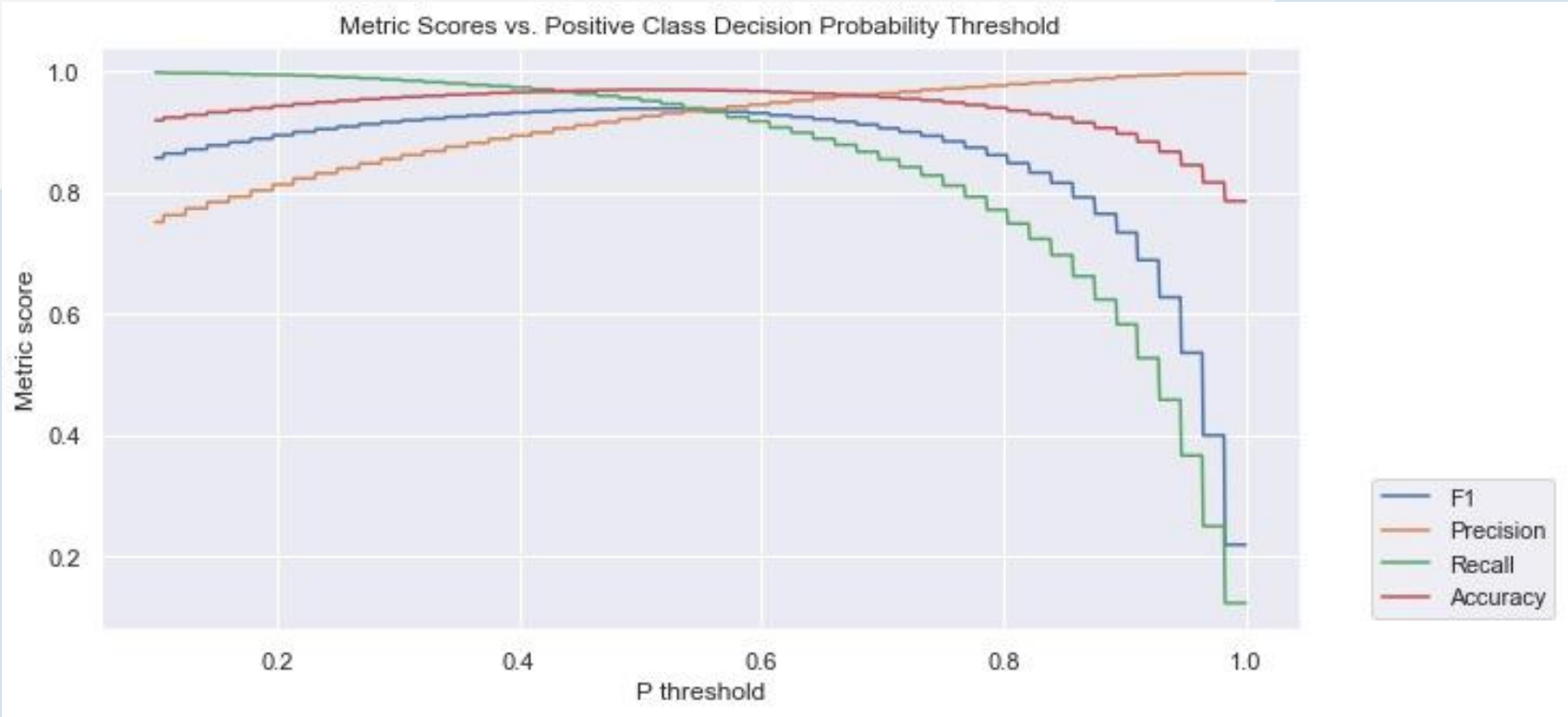
Random Forest

train	test	F(1)	precision	Recall
1.0000	0.9688	0.9575	0.9538	0.9613

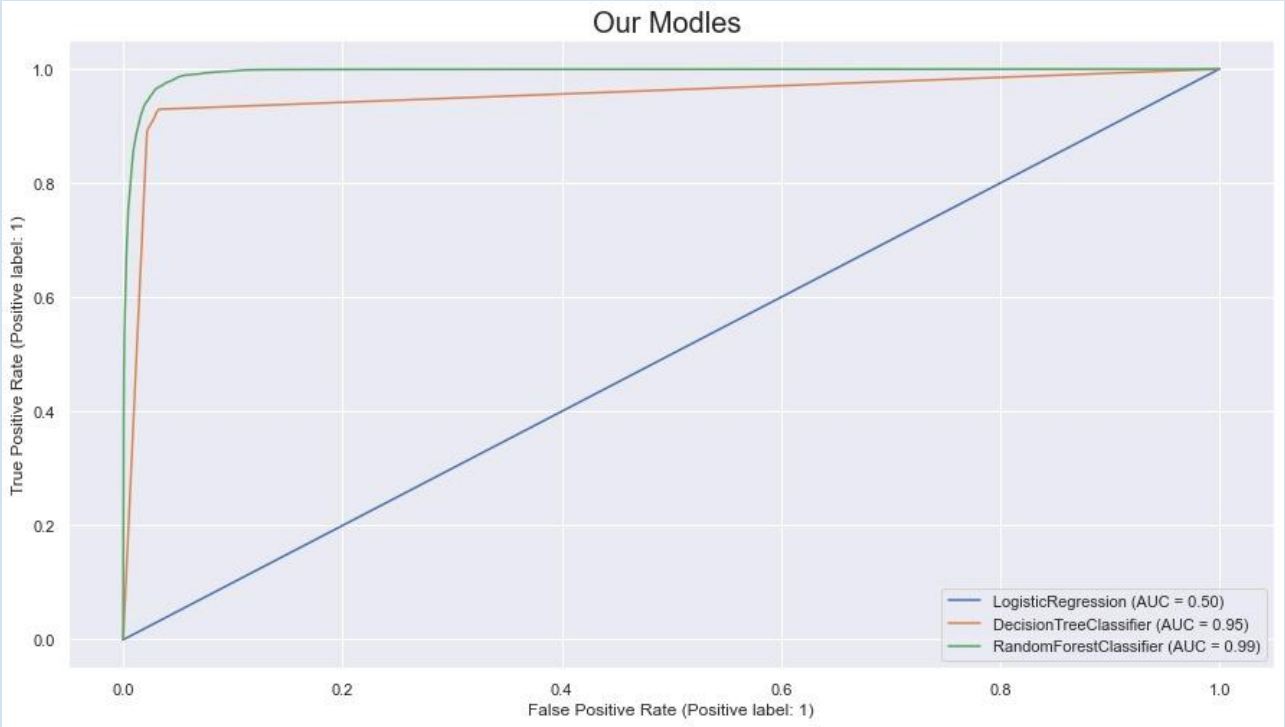
Convution Matrix Of random forest



Random Forest



ROC for some our models



Conclusion

- We found the best location to work with soler energy.
- The best model we have is random forest .

THANKS

for watching ..

Rawan Alshehri
Khaild Alharbi
Abdulrhman Alqurashi