# Weather Prediction using Python
(Ideal Playing Conditions)

Department of Chemical Engineering & Technology
Indian Institute Of Technology(IIT-BHU)-Varanasi

INTRODUCTION

A cricket game demands ideal weather conditions and systematic analysis of weather pattern can be helpful in deciding venues for cricket matches.

Our algorithm trains on a dataset consisting of various weather parameters and predicts whether a cricket match can be held on a particular day. The core concept of our project is Gaussian naïve bayes theorem which helps in predicting whether play is possible, given certain parameters. It is essentially a statistical model and the probability(likelihood to be more precise) is calculated using the Gaussian curve.

Bayes Theorem
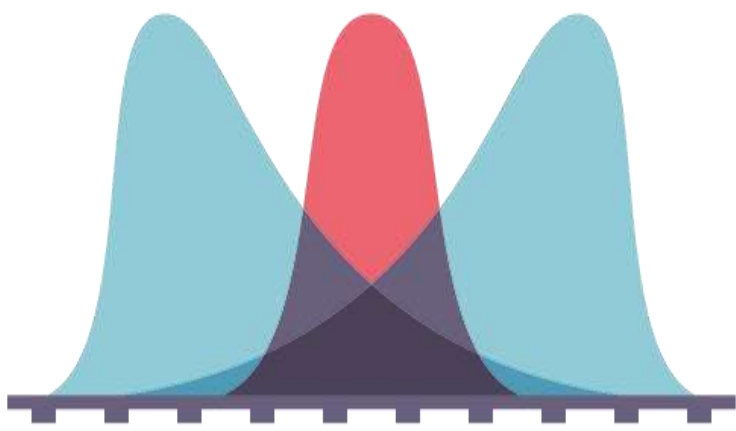
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Gaussian curve analysis

WORKING OF MODEL

1.1) READING AND PROCESSING THE DATASET

The dataset was imported from a link as a csv file from which was further read and processed pandas and matplotlib libraries of python.

```
In [1]: import pandas as pd
```

```
In [2]: df= pd.read_csv(r"C:\Users\HP\Downloads\new_dataset.csv")
        df
```

Reference - https://drive.google.com/file/d/1l4TXw85dfPRjNc3Yelm60QD6-9b_u_rN/view?usp=sharing

1.2) TESTING DATA

Creating a new data frame by transforming string dataset into numbers.

```
In [7]: #Creating the new dataframe
        inputs['outlook_n']= outlook_at.fit_transform(inputs['outlook'])
        inputs['Temp_n']= outlook_at.fit_transform(inputs['temp'])
        inputs['Hum_n']= outlook_at.fit_transform(inputs['humidity'])
        inputs['win_n']= outlook_at.fit_transform(inputs['windy'])
        inputs
```

Out[7]:

	outlook	temp	humidity	windy	outlook_n	Temp_n	Hum_n	win_n
0	rainy	hot	high	f	2	1	0	0
1	rainy	hot	high	t	5	1	0	1
2	overcast	hot	high	f	4	1	0	0
3	sunny	mild	high	f	6	3	0	0

1.3) FORECASTING

- The model forecasted that whether match should be played or not..
- We converted the string data set into numerical.
- Forecasting was done and its accuracy was predicted using gaussian naive bayes algorithm.

```
In [9]: #Applying the Gaussian naivebayes
        classifier = GaussianNB()
        classifier.fit(inputs_n,target)
Out[9]: GaussianNB()

In [10]: #85% accuracy
         classifier.score(inputs_n,target)
Out[10]: 0.8571428571428571

In [11]: #Prediction
         classifier.predict([[1,0,1,1]])
Out[11]: array(['yes'], dtype='<U4')

In [12]: df
Out[12]:
```
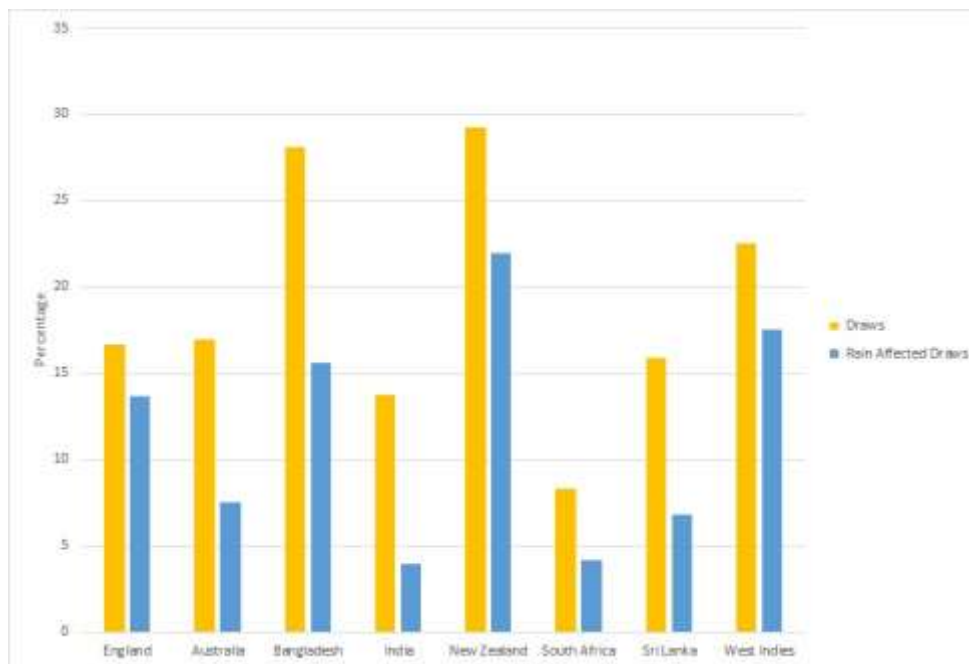
	outlook	temp	humidity	windy	play
0	rainy	hot	high	f	no
1	rainy	hot	high	t	no
2	overcast	hot	high	f	yes
3	sunny	mild	high	f	yes
4	sunny	cold	normal	f	yes
5	sunny	cold	normal	t	no
6	overcast	cold	normal	t	yes
7	rainy	mild	high	f	yes

RAIN AFFECTED DRAWS

RELEVANCE

- This project can help organizing committee of a tournament in choosing better venues.
- A lot of teams are put in a disadvantageous situation due to abandoned matches. This can be easily prevented.
- The organizing committee, broadcasters as well as spectators lose money due to rain abandoned matches.
- This can be applied to other sports with minor modifications in code. For example in track and field sports.

CONCLUSION

- The model took input of the dataset consisting of the parameters such as temperature, humidity, outlook and wind conditions.
- Based on this dataset, the model made a statistical prediction of whether a cricket match is possible on a particular date and venue

REFERENCES

1) https://youtu.be/Gv9_4yMHFhI

2) https://youtu.be/mtbJbDwqWLE

3)https://www.udemy.com/course/pythonforbeginners/
4) https://www.udemy.com/course/data-science-linear-regression-in-python/
5)Probability And Statistics in Engineering by **WILLIAM W.HINES**

CONTRIBUTOR

Name-Harshit Sharma

Roll No. -21045058

Email- harshit.sharma.che21@itbhu.ac.in