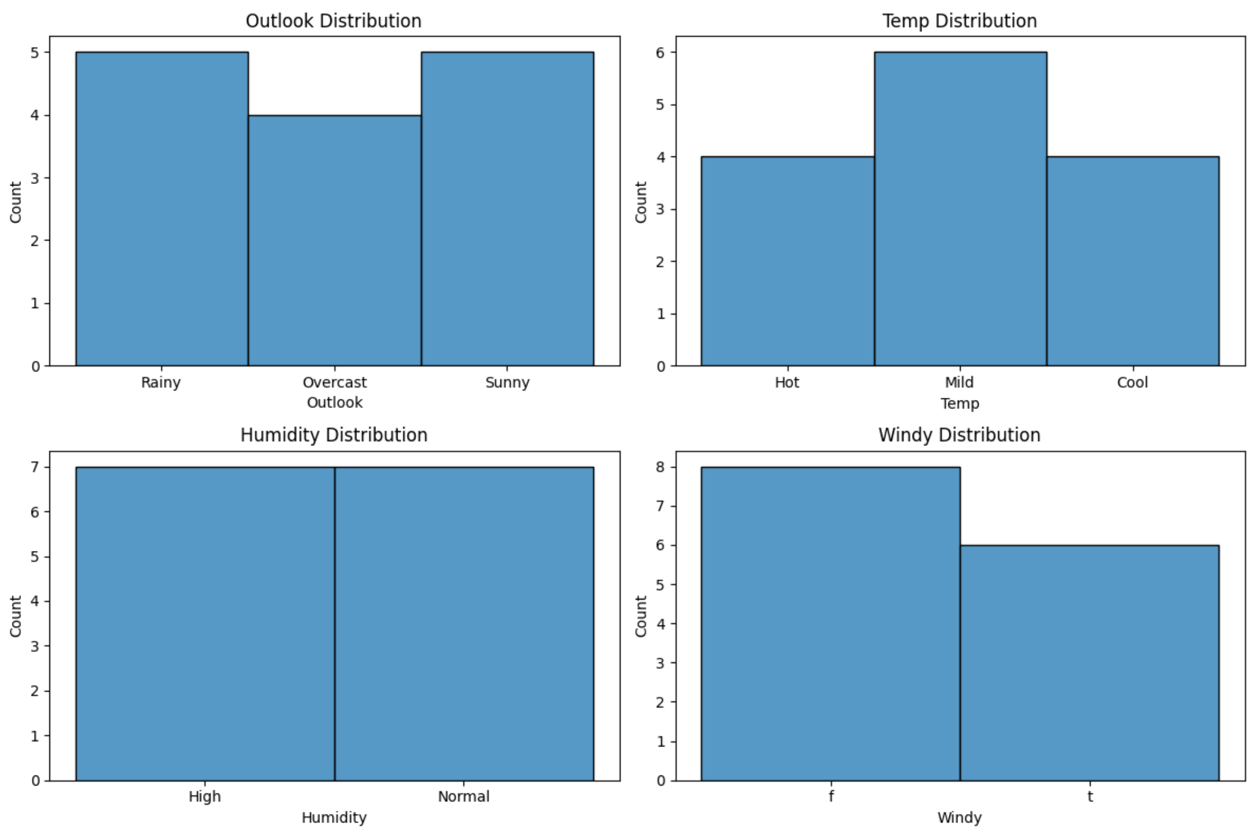
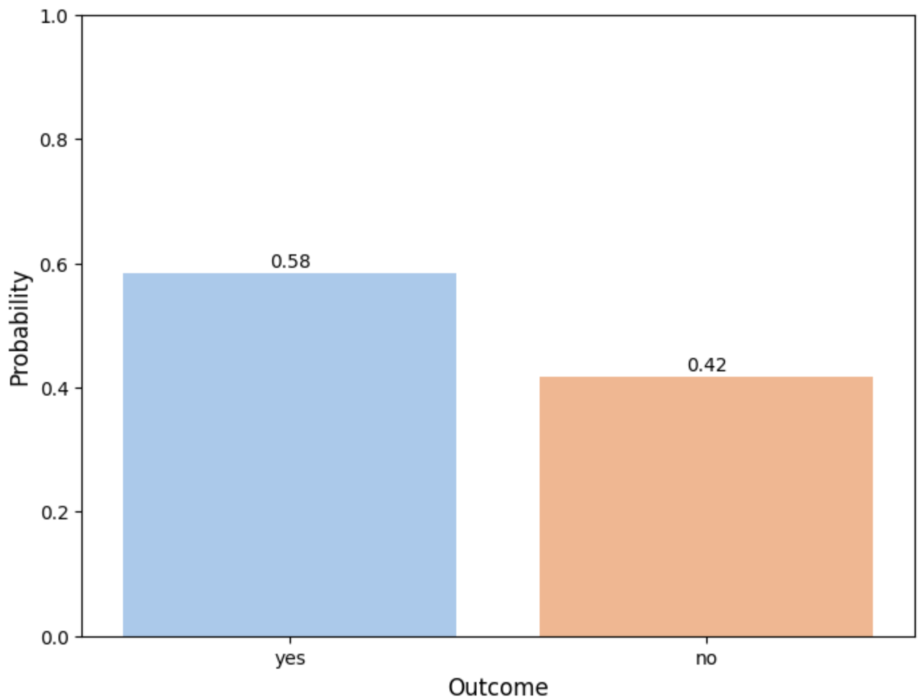
2 Naive Bayes

2.1 Data Pre-Processing & Visualisation

Data Distribution Plots after CATEGORICAL ENCODING



2.2 Prior Probabilities



2.3 Likelihood Probabilities

OUTLOOK	P(yes)	P(no)
Overcast	0.428571	0.000000
Rainy	0.285714	0.600000
Sunny	0.285714	0.400000

Temp	P(yes)	P(no)
Cool	0.428571	0.200000
Hot	0.285714	0.400000
Mild	0.285714	0.400000

Humidity	P(yes)	P(no)
High	0.285714	0.800000
Normal	0.714286	0.200000

Windy	P(yes)	P(no)
f	0.714286	0.400000
t	0.285714	0.600000

2.4 Making Predictions based on proportional probabilities

Test-Case	P(yes)	P(no)	Prediction
1	0.024295432458697763	0.005333333333333335	yes
2	0.005830903790087463	0.0	yes

Accuracy = 1.0

2.5 Likelihood Probabilities after Laplace smoothing

1. α = **0.01**

OUTLOOK	P(yes)	P(no)
Overcast	0.428165	0.001988
Rainy	0.285917	0.598410
Sunny	0.285917	0.399602

Temp	P(yes)	P(no)
Cool	0.428165	0.200795
Hot	0.285917	0.399602
Mild	0.285917	0.399602

Humidity	P(yes)	P(no)
High	0.286325	0.798805
Normal	0.713675	0.201195

Windy	P(yes)	P(no)
f	0.713675	0.400398
t	0.286325	0.599602

Making Predictions based on proportional probabilities after Laplace smoothing

Test-Case	P(yes)	P(no)	Prediction
1	0.02428843241892867	0.005359877943368312	yes
2	0.005854456555063937	0.00015854490113897748	yes

Accuracy = 1.0

2. $\alpha = \mathbf{0.1}$

OUTLOOK	P(yes)	P(no)
Overcast	0.424658	0.018868
Rainy	0.287671	0.584906
Sunny	0.287671	0.396226

Temp	P(yes)	P(no)
Cool	0.424658	0.207547
Hot	0.287671	0.396226
Mild	0.287671	0.396226

Humidity	P(yes)	P(no)
High	0.291667	0.788462
Normal	0.708333	0.211538

Windy	P(yes)	P(no)
f	0.708333	0.403846
t	0.291667	0.596154

Making Predictions based on proportional probabilities after Laplace smoothing

Test-Case	P(yes)	P(no)	Prediction
1	0.02422060697754426	0.005588315294667817	yes
2	0.006062135771772901	0.001464182382915439	yes

Accuracy = 1.0

3. $\alpha = \mathbf{1}$

OUTLOOK	P(yes)	P(no)
Overcast	0.400000	0.125000
Rainy	0.300000	0.500000
Sunny	0.300000	0.375000

Temp	P(yes)	P(no)
Cool	0.400000	0.250000
Hot	0.300000	0.375000
Mild	0.300000	0.375000

Humidity	P(yes)	P(no)
High	0.333333	0.714286
Normal	0.666667	0.285714

Windy	P(yes)	P(no)
f	0.666667	0.428571
t	0.333333	0.571429

Making Predictions based on proportional probabilities after Laplace smoothing

Test-Case	P(yes)	P(no)	Prediction
1	0.023333333333333334	0.007174744897959183	yes
2	0.0077777777777777776	0.007971938775510204	no

Accuracy = 0.5

4. $\alpha = \mathbf{10}$

OUTLOOK	P(yes)	P(no)
Overcast	0.351351	0.285714
Rainy	0.324324	0.371429
Sunny	0.324324	0.342857

Temp	P(yes)	P(no)
Cool	0.351351	0.314286
Hot	0.324324	0.342857
Mild	0.324324	0.342857

Humidity	P(yes)	P(no)
High	0.444444	0.560000
Normal	0.555556	0.440000

Windy	P(yes)	P(no)
f	0.555556	0.480000
t	0.444444	0.520000

Making Predictions based on proportional probabilities after Laplace smoothing

Test-Case	P(yes)	P(no)	Prediction
1	0.018937856775694616	0.010344489795918368	yes
2	0.0131302473644816	0.011885714285714286	yes

Accuracy = 1.0

5. α = **100**

OUTLOOK	P(yes)	P(no)
Overcast	0.335505	0.327869
Rainy	0.332248	0.337705
Sunny	0.332248	0.334426

Temp	P(yes)	P(no)
Cool	0.335505	0.331148
Hot	0.332248	0.334426
Mild	0.332248	0.334426

Humidity	P(yes)	P(no)
High	0.492754	0.507317
Normal	0.507246	0.492683

Windy	P(yes)	P(no)
f	0.507246	0.497561
t	0.492754	0.502439

Making Predictions based on proportional probabilities after Laplace smoothing

Test-Case	P(yes)	P(no)	Prediction
1	0.016568313127134977	0.01142360680677749	yes
2	0.015788362604415428	0.01164533786645278	yes

Accuracy = 1.0

2.6 observed differences in predictions

- We see that probability of not playing while OUTLOOK is Overcast is zero
- Assigning zero probability means ruling out unseen events from consideration, which is not be desired
- After Laplace smoothing with smoothing parameter alpha we overcome this problem
- Hence probability of not playing in Test-Case 2 is not 0 anymore and makes more sense
- We see a regular trend not matter what alpha value we have taken except for the case where alpha = 1
- Here the probability of not playing overtakes that of playing in test case 2 and hence accuracy drops down from 1 to 0.5
- Later on (alpha more than 1) the overtaking doesn't occur
- hence hyper-parameter tuning has also been done