

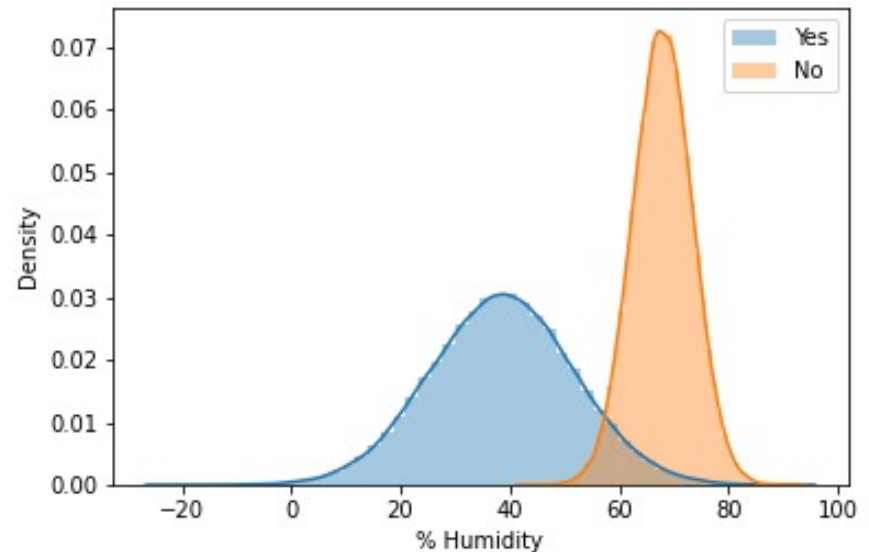
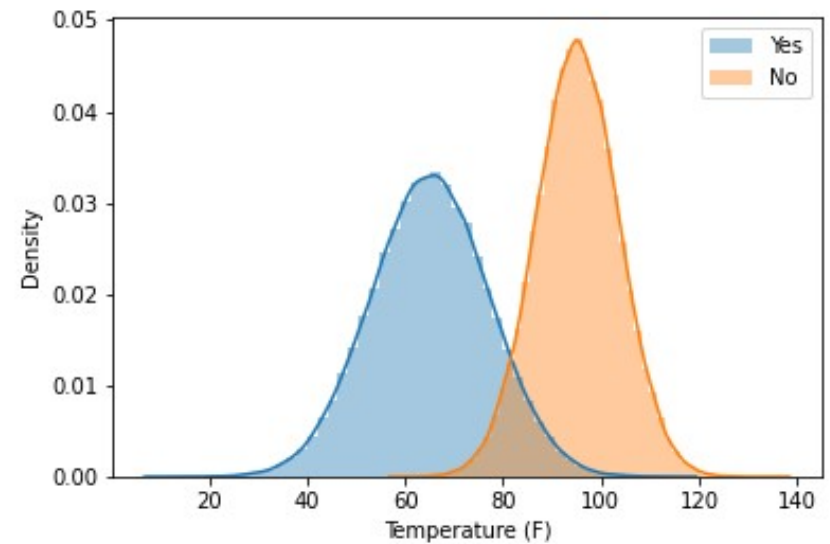
Naive Bayes Classification

Gaussian Naive Bayes

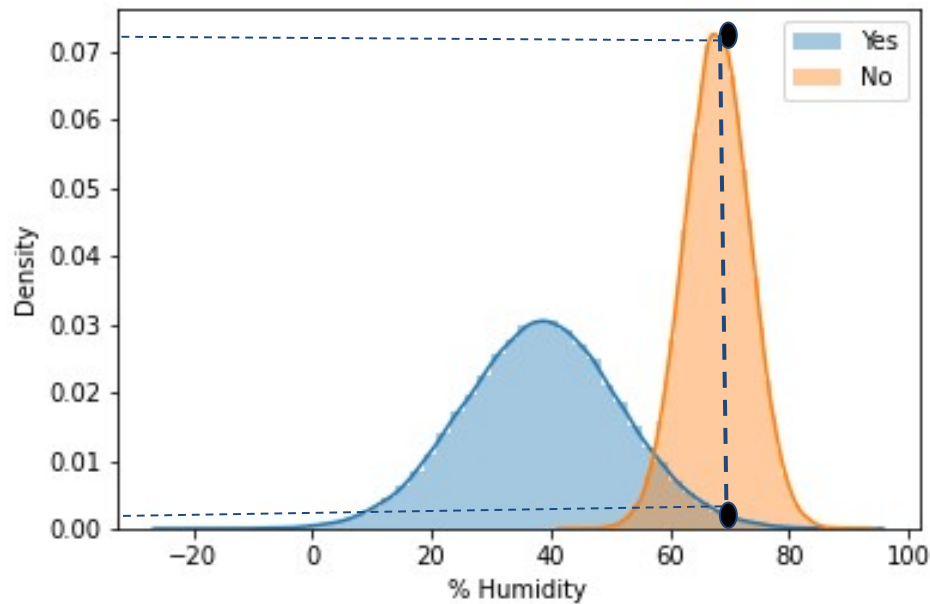
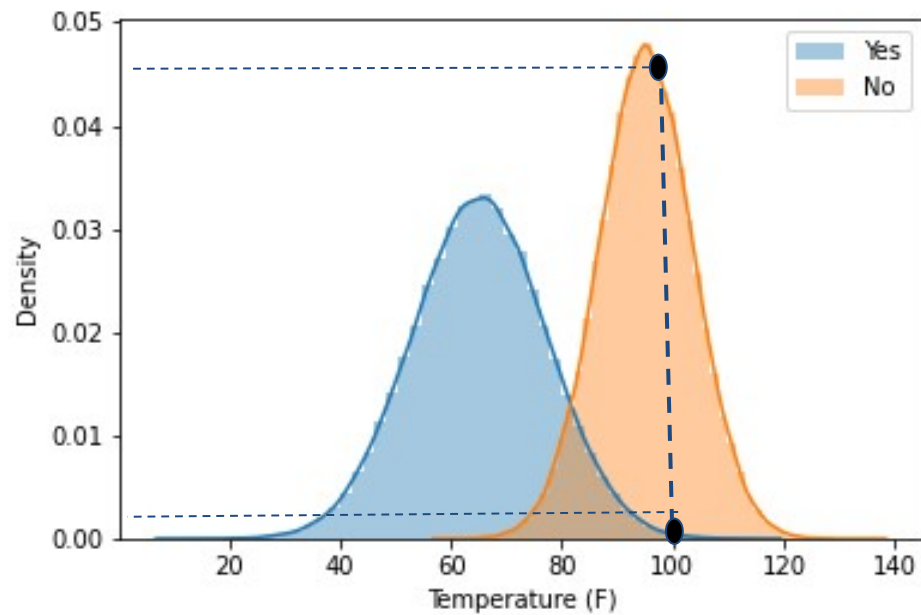
Gaussian Naive Bayes is a type of Naive Bayes classifier that assumes that your **predictors follow a normal (Gaussian) distribution**.

It can be used when your predictors are continuous.

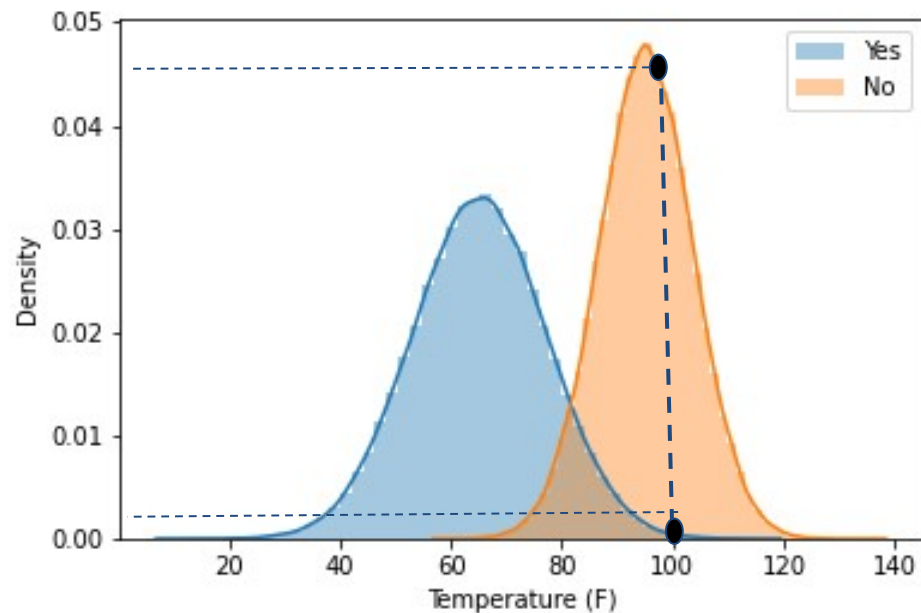
Temperature	Humidity	Play
54	40	Yes
63	42	Yes
81	60	No
72	28	Yes
68	70	Yes
59	40	Yes
83	43	Yes
99	67	No
66	20	Yes
42	35	Yes
103	70	No
98	75	No
80	30	Yes



Will we play tennis if it is 100 degrees with 70% humidity?

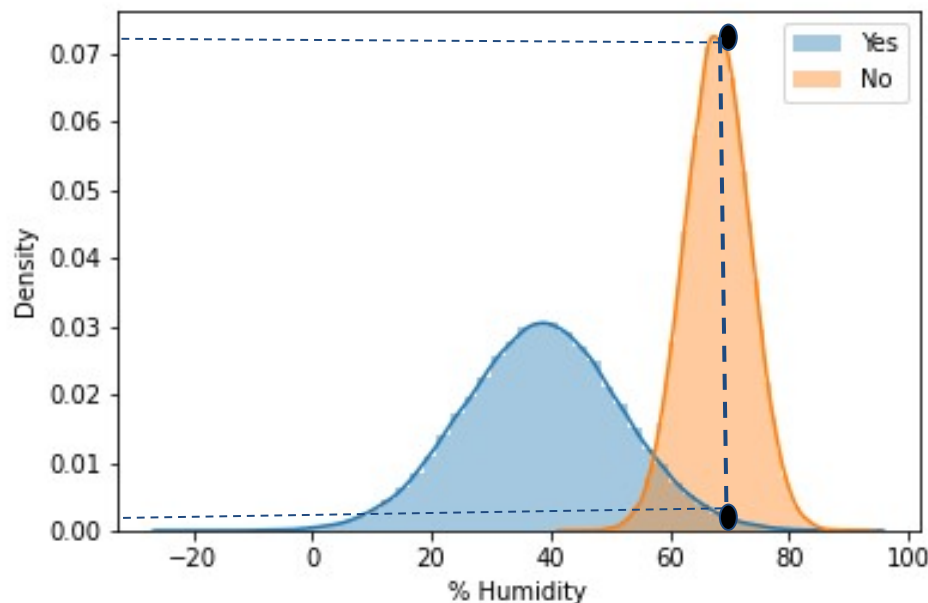


Will we play tennis if it is 100 degrees with 70% humidity?



$$\begin{aligned}
 P(\text{Yes} \mid 100, 70) &= P(100 \mid \text{Yes}) * P(70 \mid \text{Yes}) * P(\text{Yes}) \\
 &= 0.002 * 0.003 * 0.69 \\
 &= 4.15e^{-6}
 \end{aligned}$$

$$\begin{aligned}
 P(\text{No} \mid 100, 70) &= P(100 \mid \text{No}) * P(70 \mid \text{No}) * P(\text{No}) \\
 &= 0.045 * 0.08 * 0.30 \\
 &= 1080e^{-6}
 \end{aligned}$$



Will we play tennis if it is 100 degrees with 70% humidity?

Naive Bayes Pros & Cons

Pros

- Easy & fast
- Can handle multi-class predictions
- Performs well when assumption of independence holds
- Is fairly robust even if independence assumption doesn't hold
- Works well with large datasets

Cons

- Can run into issues if a category was not observed in training data set
- Posterior probabilities are not true probabilities
- May not perform as well as other more complex machine learning methods