ASSIGNMENT - 5

Naive Bayes classifier - Probability Estimation:

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

P (PlayTennis = Yes)   =   9/14
P (PlayTennis = No)   =   5/14

OUTLOOK:
P (Outlook=Sunny | PlayTennis=Yes)   =   2/9
P (Outlook=Sunny | PlayTennis=No)   =   3/5
P (Outlook=Overcast | PlayTennis=Yes)   =   4/9
P (Outlook=Overcast | PlayTennis=No)   =   0/5
P (Outlook=Rain | PlayTennis=Yes)   =   3/9
P (Outlook=Rain | PlayTennis=No)   =   2/5

TEMPERATURE:
P (Temperature=Hot | PlayTennis=Yes)   =   2/9

P (Temperature=Hot | PlayTennis=No)   =   2/5
P (Temperature=Mild | PlayTennis=Yes)   =   4/9
P (Temperature=Mild | PlayTennis=No)   =   2/5
P (Temperature=Cool | PlayTennis=Yes)   =   3/9
P (Temperature=Cool | PlayTennis=No)   =   1/5

HUMIDITY:
P (Humidity=High | PlayTennis=Yes)   =   3/9
P (Humidity=High | PlayTennis=No)   =   4/5
P (Humidity=Normal | PlayTennis=Yes)   =   6/9
P (Humidity=Normal | PlayTennis=No)   =   1/5

WIND:
P (Wind=Weak | PlayTennis=Yes)   =   6/9
P (Wind=Weak | PlayTennis=No)   =   2/5
P (Wind=Strong | PlayTennis=Yes)   =   3/9
P (Wind=Strong | PlayTennis=No)   =   3/5

Considering an instance
x' = {Outlook=Sunny, Temperature=Cool, Humidity=High, Wind=Strong}

Therefore, **P(Yes | x')** = P (PlayTennis = Yes) * P (Outlook=Sunny | PlayTennis=Yes) * P (Temperature=Cool | PlayTennis=Yes) * P (Humidity=High | PlayTennis=Yes) * P (Wind=Strong | PlayTennis=Yes)

$$= \quad 9/14 * 2/9 * 3/9 * 3/9 * 3/9$$
$$= \quad 1/189$$
$$= \quad 0.00529100529 \quad \text{-------------------------------- (I)}$$

And, **P(No | x')** = P (PlayTennis = No) * P (Outlook=Sunny | PlayTennis=No) * P (Temperature=Cool | PlayTennis=No) * P (Humidity=High | PlayTennis=No) * P (Wind=Strong | PlayTennis=No)

$$= \quad 5/14 * 3/5 * 1/5 * 4/5 * 3/5$$
$$= \quad 18/875$$
$$= \quad 0.02057142857 \quad \text{-------------------------------- (II)}$$

From, (I) and (II), we see that, P(Yes | x') < P(No | x')
Hence, for the instance **x'** = {Outlook=Sunny, Temperature=Cool, Humidity=High, Wind=Strong} we conclude that not playing tennis is the maximum likelihood hypothesis.

Similarly, considering another instance
y' = {Outlook=Overcast, Temperature=Cool, Humidity=High, Wind=Strong}

Therefore, **P(Yes | y')** = P (PlayTennis = Yes) * P (Outlook= Overcast | PlayTennis=Yes) * P (Temperature=Cool | PlayTennis=Yes) * P (Humidity=High | PlayTennis=Yes) * P (Wind=Strong | PlayTennis=Yes)

      =   9/14 * 4/9 * 3/9 * 3/9 * 3/9

      =   2/189

      =   0.01058201058   -------------------------------- (III)

And, **P(No | y')** = P (PlayTennis = No) * P (Outlook= Overcast | PlayTennis=No) * P (Temperature=Cool | PlayTennis=No) * P (Humidity=High | PlayTennis=No) * P (Wind=Strong | PlayTennis=No)

      =   5/14 * 0/5 * 1/5 * 4/5 * 3/5

      =   0   -------------------------------- (IV)

From, (III) and (IV), we see that, P(Yes | y') > P(No | y')

Hence, for the instance y' = {Outlook=Overcast, Temperature=Cool, Humidity=High, Wind=Strong} we conclude that playing tennis is the maximum likelihood hypothesis.