Machine Learning (WiSe 2025/2026)

Author: Suvansh Shukla
Matriculation No. 256245

Assignment 3 Task 3.3

The given training data is:

day	outlook	temperature	humidity	wind	play tennis
1	rain	mild	high	strong	no
2	overcast	hot	normal	weak	yes
3	overcast	mild	high	strong	yes
4	sunny	mild	normal	strong	yes
5	rain	mild	normal	weak	yes
6	sunny	cool	normal	weak	yes
7	sunny	mild	high	weak	no
8	overcast	cool	normal	strong	yes
9	rain	cool	normal	strong	no
10	rain	cool	normal	weak	yes

We'll need to calculate the possibility for every feature according to both classifications of YES and NO.

For the first instance i1 = (sunny, cool, high, strong)

For Class = YES

P(yes)=7/10=0,7
P(outlook=sunny|yes)=2/7 = 0.2857
P(temperature=cool|yes)=3/7 = 0.4286
P(humidity=high|yes)=1/7 = 0.1429
P(wind=strong|yes)=3/7 = 0.4286

For Class = NO

P(no)=3/10=0.3
P(outlook=sunny|no)=1/3 = 0.3333
P(temperature=cool|no)=1/3 = 0.3333
P(humidity=high|no)=2/3 = 0.6667
P(wind=strong|no)=2/3 = 0.6667

For yes we have

P(yes|i1) = 0.7 X P(sunny|yes) X P(cool|yes) X P(high|yes) X P(strong|yes)
P(yes|i1) = 0.7 X 0.2857 X 0.4286 X 0.1429 X 0.4286
P(yes|i1) = 0.0052

For no we have

P(no|i1) = 0.3 X P(sunny|no) X P(cool|no) X P(high|no) X P(strong|no)
P(no|i1) = 0.3 X 0.333 X 0.333 X 0.666 X 0.666
P(no|i1) = 0.0148

Now we normalize for proper probabilities:

P(yes|i) = 0.0052/(0.0052+0.0248) = 0.2616
P(no|i1) = 0.0148/(0.0052+0.0248) = 0.7384

This means that No has a higher chance of occurring

For the second instance i2 = (overcast, mild, normal, weak)

For yes,

P(yes|i2) = 0.7XP(overcast|yes) X P(mild|yes) X P(normal|yes) X P(weak|yes)
P(yes|i2) = 0.7 X 0.4286 X 0.4286 X 0.8571 X 0.5714
P(yes|i2) = 0.06297

For no,

P(no|i2) = 0.3XP(overcast|no)...

But since there is no example for P(overcast|no) it's probability is 0.
Therefore the entire probability of P(no|i2) is 0.
This means that the probability for P(yes|i2) is 1, i.e. it is certain

Part (B)

The formula in the slide is:

$\hat{P}(a_i|v_j) \leftarrow \frac{n_c+mp}{n+m}$

We've also been told to use `m=1` and to assume that we have a uniform distribution of values.

Using this we now calculate prior estimates (p) for all values.

Class = yes (n = 7):

outlook:

- P(overcast|yes) = (3+1/3)/8 = 3.3333/8 = 0.4166667
- P(sunny|yes) = (2+1/3)/8 = 2.3333/8 = 0.2916667
- P(rain|yes) = 0.2916667

temperature:

- P(cool|yes) = 0.4166667
- P(mild|yes) = 0.4166667
- P(hot|yes) = (1+1/3)/8 = 0.1666667

humidity (p = 1/2):

- P(normal|yes) = (6+0.5)/8 = 6.5/8 = 0.8125
- P(high|yes) = 1.5/8 = 0.1875

wind:

- P(weak|yes) = 4.5/8 = 0.5625
- P(strong|yes) = 3.5/8 = 0.4375

Class = no (n = 3):

outlook:

- P(rain|no) = (2+1/3)/4 = 2.3333/4 = 0.5833333
- P(sunny|no) = (1+1/3)/4 = 0.3333333
- P(overcast|no) = (0+1/3)/4 = 0.0833333

temperature:

- P(mild|no) = 0.5833333
- P(cool|no) = 0.3333333
- P(hot|no) = 0.0833333

humidity:

- P(high|no) = (2+0.5)/4 = 0.625
- P(normal|no) = 0.375

wind:

- P(strong|no) = 0.625
- P(weak|no) = 0.375

Now we can calculate the joint (unnormalized) scores and posteriors

For instance i1:

Joint yes = 0.7 X 0.291 X 0.416 X 0.187 X 0.437 = 0.00697

Joint no = 0.3 X 0.333 X 0.333 X 0.625 X 0.625 = 0.0130

For normalization = 0.00697 + 0.0130 = 0.01997

Therefore,

P(yes|i1) = 0.00697835/0.01999919 = 0.34893
P(no|i1) = 0.65107

For instance i2:

joint yes = 0.7 X 0.4166667 X 0.4166667 X 0.8125 X 0.5625 = 0.0555
joint no = 0.3 X 0.0833333 X 0.5833333 X 0.375 X 0.375 = 0.00205

for normalization = 0.0555 + 0.00205 = 0.05755

Therefore,

P(yes|i2) = 0.96439
P(no|i2) = 0.03561

The difference observed here for i1 is that the confidence for no decreases but the overall prediction is still yes.
For i2, we see that the prediction is still yes, but the zero probability for no goes away.