1.

a.

P(Sail=yes) = 7/10 = 0.7

P(Sail=no) = 3/10 = 0.3

b.

P(Sky|Sail) for all values of Sky ∈ {clear, cloudy, overcast} and Sail ∈ {yes, no}

P(clear|yes) = (2/10)/(7/10) = 2/7 = 0.2857

P(cloudy|yes) = (3/10)/(7/10) = 3/7 = 0.4286

P(overcast|yes) = (2/10)/(7/10) = 2/7 = 0.2857

P(clear|no) = (1/10)/(3/10) = 1/3 = 0.3333

P(cloudy|no) = (1/10)/(3/10) = 1/3 = 0.3333

P(overcast|no) = (1/10)/(3/10) = 1/3 = 0.3333

c.

P(Sea|Sail) for all values of Sea ∈ {blue, gray} and Sail ∈ {yes, no}

P(blue|yes) = (5/10)/(7/10) = 5/7 = 0.7143

P(gray|yes) = (2/10)/(7/10) = 2/7 = 0.2857

P(blue|no) = (0/10)/(3/10) = 0/3 => (0 + 1)/ (3 + 2) = 1/5 = 0.2

P(gray|no) = (3/10)/(3/10) = 3/3 = 1

d.

P(Sun|Sail) for all values of Sun ∈ {true, false} and Sail ∈ {yes, no}

P(true|yes) = (4/10)/(7/10) = 4/7 = 0.5714

P(false|yes) = (3/10)/(7/10) = 3/7 = 0.4286

P(true|no) = (0/10)/(3/10) = 0/3 => (0 + 1)/(3 + 2) = 1/5 = 0.2

P(false|no) = (3/10)/(3/10) = 3/3 = 1

e.

P(Sail=yes | Sky=overcast, Sea=gray, Sun=true) =

= P(Sail=yes) Π P(Sky=overcast | Sail=yes) P(Sea=gray | Sail=yes) P(Sun=true | Sail=yes)

= (7/10) \* (2/7) \* (2/7) \* (4/7) = α0.032653

P(Sail=no | Sky=overcast, Sea=gray, Sun=true) =

= P(Sail=no) Π P(Sky=overcast | Sail=no) P(Sea=gray | Sail=no) P(Sun=true | Sail=no)

= (3/10) \* (1/3) \* (3/3) \* (1/5) = α0.02

so:

P(Sail | Sky=overcast, Sea=gray, Sun=true) = α<0.032653, 0.02>

α = 1/(0.032653 + 0.02) = 18.9922

P(Sail | Sky=overcast, Sea=gray, Sun=true) = <0.62, 0.38>

therefore:

P(Sail=yes | Sky=overcast, Sea=gray, Sun=true) = 0.62

P(Sail=no | Sky=overcast, Sea=gray, Sun=true) = 0.38

f.

0.62 > 0.38

Naïve Bayes would choose Sail = yes

2.

a.

|  |  |  |  |
| --- | --- | --- | --- |
| Sky | Sea | Sun | Sail |
| 0 | 0 | 0 | Yes |
| 0 | 1 | 1 | Yes |
| 0 | 1 | 0 | No |
| 1 | 0 | 1 | Yes |
| 1 | 0 | 0 | Yes |
| 1 | 1 | 1 | Yes |
| 1 | 1 | 0 | No |
| 2 | 0 | 1 | Yes |
| 2 | 0 | 0 | Yes |
| 2 | 1 | 0 | No |

b.

New instance: <Sky=overcast, Sea=gray, Sun=true> => <Sky=2, Sea=1, Sun=1>

Euclidean distance between 3-dimensional points =

(x1, y1, z1) = (2, 1, 1) (the new instance values)

|  |  |  |  |
| --- | --- | --- | --- |
| Line | Calculation | Distance | Class |
| 1 |  | 2.45 | Yes |
| 2 |  | 2 | Yes |
| 3 | =2.24 | 2.24 | No |
| 4 | =1.41 | 1.41 | Yes |
| 5 | =1.73 | 1.73 | Yes |
| 6 |  | 1 | Yes |
| 7 | =1.41 | 1.41 | No |
| 8 |  | 1 | Yes |
| 9 | =1.41 | 1.41 | Yes |
| 10 |  | 1 | No |

c.

3 nearest neighbors, there are three points at distance 1, two are yes, one is no, majority of nearest neighbors are yes, so yes.

3NN would classify <Sky=overcast, Sea=gray, Sun=true> as Yes

d.

7 nearest neighbors: 1:yes, 1:yes, 1:no, 1.41:yes, 1.41:no, 1.41:yes, 1.73:yes

5 yes’es, 2 no’s

7NN would classify <Sky=overcast, Sea=gray, Sun=true> as Yes