Q1.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Color = Yello, Size = 2lb, Temp = Warm-blooded, Bath = Everyday | | | | | | |
|  | class | color | size | temp | bath |  |
| P(Cat|New pet) | 0.343 | 0.122 | 0.25 | 0.921 | 0.154 | 0.0014823141 |
| P(Dog|New pet) | 0.451 | 0.269 | 0.255 | 0.898 | 0.340 | 0.0094495238 |
| P(Small mammal|New pet) | 0.02 | 0.250 | 0.286 | 0.600 | 0.167 | 0.0001428571 |
| P(Reptile|New pet) | 0.078 | 0.214 | 0.077 | 0.182 | 0.167 | 0.0000389610 |
| P(Fish| New pet) | 0.049 | 0.091 | 0.100 | 0.250 | 0.222 | 0.0000247475 |
| P(Bird|New pet) | 0.029 | 0.111 | 0.250 | 0.667 | 0.286 | 0.0001534392 |
| P(Other|New pet) | 0.029 | 0.111 | 0.125 | 0.333 | 0.286 | 0.0000383598 |
| \* pick bath 'More frequently than once a month' cause it's closest to Everyday bath | | | | | | |

From the above table, I assume that the pet is a Dog

Q2.

|  |  |  |  |
| --- | --- | --- | --- |
| cats | 58.4 | 251.3 | 0.232391564 |
| dogs | 76.8 | 251.3 | 0.305610824 |
| small mammal | 6.2 | 251.3 | 0.024671707 |
| reptile | 6 | 251.3 | 0.023875846 |
| fish | 76.3 | 251.3 | 0.30362117 |
| birds | 22.9 | 251.3 | 0.091126144 |
| other | 4.7 | 251.3 | 0.018702746 |
|  | 251.3 |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | class | color | size | temp | bath |  |
| P(Cat|New pet) | 0.232391564 | 0.122 | 0.25 | 0.921 | 0.154 | 0.00100430696 |
| P(Dog|New pet) | 0.305610824 | 0.269 | 0.255 | 0.898 | 0.340 | 0.00640327440 |
| P(Small mammal|New pet) | 0.024671707 | 0.250 | 0.286 | 0.600 | 0.167 | 0.00017622648 |
| P(Reptile|New pet) | 0.023875846 | 0.214 | 0.077 | 0.182 | 0.167 | 0.00001192600 |
| P(Fish| New pet) | 0.30362117 | 0.091 | 0.100 | 0.250 | 0.222 | 0.00015334403 |
| P(Bird|New pet) | 0.091126144 | 0.111 | 0.250 | 0.667 | 0.286 | 0.00048214891 |
| P(Other|New pet) | 0.018702746 | 0.111 | 0.125 | 0.333 | 0.286 | 0.00002473908 |

Re-calculate the probability of each Class and plug in the value into the table again. It is still considered a dog.

Q3.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Color = Yello, Size = 2lb, Temp = Warm-blooded, Bath = Everyday | | | | | | |
|  | class | color | size | temp | bath |  |
| P(Cat|New pet) | 0.343 | 0.122 | 1 | 0.921 | 0.154 | 0.0059292563 |
| P(Dog|New pet) | 0.451 | 0.269 | 1 | 0.898 | 0.340 | 0.0370712088 |
| P(Small mammal|New pet) | 0.02 | 0.250 | 1 | 0.600 | 0.167 | 0.0005000000 |
| P(Reptile|New pet) | 0.078 | 0.214 | 1 | 0.182 | 0.167 | 0.0005064935 |
| P(Fish| New pet) | 0.049 | 0.091 | 1 | 0.250 | 0.222 | 0.0002474747 |
| P(Bird|New pet) | 0.029 | 0.111 | 1 | 0.667 | 0.286 | 0.0006137566 |
| P(Other|New pet) | 0.029 | 0.111 | 1 | 0.333 | 0.286 | 0.0003068783 |
| \* pick bath 'More frequently than once a month' cause it's closest to Everyday bath | | | | | | |

I think I can just ignore the size probability and treat it like just 1 not to affect other probabilities. The result still is a Dog

Q4.

Difference: The difference between the two methods lies in the possibility of data separation. If your data is clearly separable with lines, use a hard margin. If there are some data that make finding a linear classifier impossible, it is better to use a soft margin to allow some data points to be misclassified.

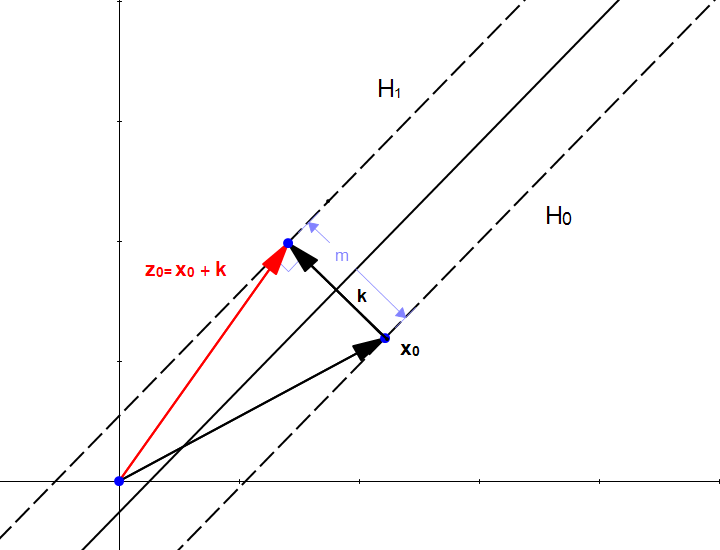
Hard margin advantages: No misclassification occurs.

Soft margin advantages: It is applicable even if a linear boundary is not feasible.

Which one and which situation: Use a hard margin when your data is linearly and reliably separable. If there are data points that get in the way of linear classification, use a soft margin.

Extra.

If we start from the point x0 and add k we find that the point z0=x0+k is in the hyperplane H1 as shown on figure below.



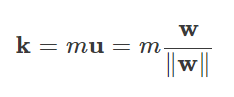
The fact that z0 is in H1 means that

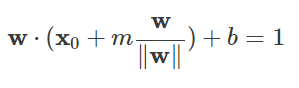


We can replace z0 by x0+k because that is sum of the vector as we can see above figure.

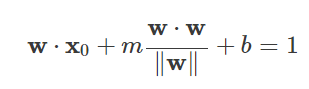


We can now replace k, cause k is

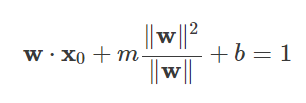


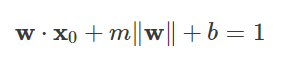


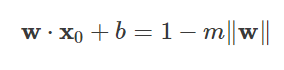
We now expand the equation



The dot product of a vector with itself is the square of its norm so :







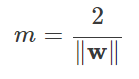
We can substitute left side with this



As a result, we can get







Finally, we get m in the figure above.

If we want to maximize m, we need to minimize w value.