

1. The conditional probabilities  $P(X_i = x_i | Y = y)$  for each feature  $X_i$  (e.g., age), its possible value  $x_i$  (e.g., 10-19), and each class label  $Y = y$  ( $y$  can be no-recurrence-events or recurrence-events).

$P(\text{age} = 10-19 | Y = \text{no-recurrence-events}) = 0.005050505050505051$   
 $P(\text{age} = 10-19 | Y = \text{recurrence-events}) = 0.011494252873563218$   
 $P(\text{age} = 20-29 | Y = \text{no-recurrence-events}) = 0.010101010101010102$   
 $P(\text{age} = 20-29 | Y = \text{recurrence-events}) = 0.011494252873563218$   
 $P(\text{age} = 30-39 | Y = \text{no-recurrence-events}) = 0.11111111111111111$   
 $P(\text{age} = 30-39 | Y = \text{recurrence-events}) = 0.1839080459770115$   
 $P(\text{age} = 40-49 | Y = \text{no-recurrence-events}) = 0.31313131313131315$   
 $P(\text{age} = 40-49 | Y = \text{recurrence-events}) = 0.3103448275862069$   
 $P(\text{age} = 50-59 | Y = \text{no-recurrence-events}) = 0.3282828282828283$   
 $P(\text{age} = 50-59 | Y = \text{recurrence-events}) = 0.25287356321839083$   
 $P(\text{age} = 60-69 | Y = \text{no-recurrence-events}) = 0.1919191919191919$   
 $P(\text{age} = 60-69 | Y = \text{recurrence-events}) = 0.19540229885057472$   
 $P(\text{age} = 70-79 | Y = \text{no-recurrence-events}) = 0.030303030303030304$   
 $P(\text{age} = 70-79 | Y = \text{recurrence-events}) = 0.011494252873563218$   
 $P(\text{age} = 80-89 | Y = \text{no-recurrence-events}) = 0.005050505050505051$   
 $P(\text{age} = 80-89 | Y = \text{recurrence-events}) = 0.011494252873563218$   
 $P(\text{age} = 90-99 | Y = \text{no-recurrence-events}) = 0.005050505050505051$   
 $P(\text{age} = 90-99 | Y = \text{recurrence-events}) = 0.011494252873563218$   
 $P(\text{menopause} = \text{lt40} | Y = \text{no-recurrence-events}) = 0.03125$   
 $P(\text{menopause} = \text{lt40} | Y = \text{recurrence-events}) = 0.012345679012345678$   
 $P(\text{menopause} = \text{ge40} | Y = \text{no-recurrence-events}) = 0.45833333333333333$   
 $P(\text{menopause} = \text{ge40} | Y = \text{recurrence-events}) = 0.38271604938271603$   
 $P(\text{menopause} = \text{premeno} | Y = \text{no-recurrence-events}) = 0.51041666666666666$   
 $P(\text{menopause} = \text{premeno} | Y = \text{recurrence-events}) = 0.6049382716049383$   
 $P(\text{tumor-size} = 0-4 | Y = \text{no-recurrence-events}) = 0.03980099502487562$   
 $P(\text{tumor-size} = 0-4 | Y = \text{recurrence-events}) = 0.022222222222222223$   
 $P(\text{tumor-size} = 5-9 | Y = \text{no-recurrence-events}) = 0.024875621890547265$   
 $P(\text{tumor-size} = 5-9 | Y = \text{recurrence-events}) = 0.011111111111111112$   
 $P(\text{tumor-size} = 10-14 | Y = \text{no-recurrence-events}) = 0.12935323383084577$   
 $P(\text{tumor-size} = 10-14 | Y = \text{recurrence-events}) = 0.022222222222222223$   
 $P(\text{tumor-size} = 15-19 | Y = \text{no-recurrence-events}) = 0.11442786069651742$   
 $P(\text{tumor-size} = 15-19 | Y = \text{recurrence-events}) = 0.07777777777777778$   
 $P(\text{tumor-size} = 20-24 | Y = \text{no-recurrence-events}) = 0.17412935323383086$   
 $P(\text{tumor-size} = 20-24 | Y = \text{recurrence-events}) = 0.15555555555555556$   
 $P(\text{tumor-size} = 25-29 | Y = \text{no-recurrence-events}) = 0.15920398009950248$   
 $P(\text{tumor-size} = 25-29 | Y = \text{recurrence-events}) = 0.21111111111111111$   
 $P(\text{tumor-size} = 30-34 | Y = \text{no-recurrence-events}) = 0.1691542288557214$   
 $P(\text{tumor-size} = 30-34 | Y = \text{recurrence-events}) = 0.25555555555555554$   
 $P(\text{tumor-size} = 35-39 | Y = \text{no-recurrence-events}) = 0.05970149253731343$   
 $P(\text{tumor-size} = 35-39 | Y = \text{recurrence-events}) = 0.08888888888888889$   
 $P(\text{tumor-size} = 40-44 | Y = \text{no-recurrence-events}) = 0.0845771144278607$   
 $P(\text{tumor-size} = 40-44 | Y = \text{recurrence-events}) = 0.07777777777777778$   
 $P(\text{tumor-size} = 45-49 | Y = \text{no-recurrence-events}) = 0.014925373134328358$   
 $P(\text{tumor-size} = 45-49 | Y = \text{recurrence-events}) = 0.022222222222222223$

$P(\text{tumor-size} = 50-54 \mid Y = \text{no-recurrence-events}) = 0.024875621890547265$   
 $P(\text{tumor-size} = 50-54 \mid Y = \text{recurrence-events}) = 0.044444444444444446$   
 $P(\text{tumor-size} = 55-59 \mid Y = \text{no-recurrence-events}) = 0.004975124378109453$   
 $P(\text{tumor-size} = 55-59 \mid Y = \text{recurrence-events}) = 0.011111111111111112$   
 $P(\text{inv-nodes} = 0-2 \mid Y = \text{no-recurrence-events}) = 0.7970297029702971$   
 $P(\text{inv-nodes} = 0-2 \mid Y = \text{recurrence-events}) = 0.4725274725274725$   
 $P(\text{inv-nodes} = 3-5 \mid Y = \text{no-recurrence-events}) = 0.08415841584158416$   
 $P(\text{inv-nodes} = 3-5 \mid Y = \text{recurrence-events}) = 0.17582417582417584$   
 $P(\text{inv-nodes} = 6-8 \mid Y = \text{no-recurrence-events}) = 0.039603960396039604$   
 $P(\text{inv-nodes} = 6-8 \mid Y = \text{recurrence-events}) = 0.12087912087912088$   
 $P(\text{inv-nodes} = 9-11 \mid Y = \text{no-recurrence-events}) = 0.01485148514851485$   
 $P(\text{inv-nodes} = 9-11 \mid Y = \text{recurrence-events}) = 0.06593406593406594$   
 $P(\text{inv-nodes} = 12-14 \mid Y = \text{no-recurrence-events}) = 0.009900990099009901$   
 $P(\text{inv-nodes} = 12-14 \mid Y = \text{recurrence-events}) = 0.03296703296703297$   
 $P(\text{inv-nodes} = 15-17 \mid Y = \text{no-recurrence-events}) = 0.019801980198019802$   
 $P(\text{inv-nodes} = 15-17 \mid Y = \text{recurrence-events}) = 0.04395604395604396$   
 $P(\text{inv-nodes} = 18-20 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 18-20 \mid Y = \text{recurrence-events}) = 0.01098901098901099$   
 $P(\text{inv-nodes} = 21-23 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 21-23 \mid Y = \text{recurrence-events}) = 0.01098901098901099$   
 $P(\text{inv-nodes} = 24-26 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 24-26 \mid Y = \text{recurrence-events}) = 0.02197802197802198$   
 $P(\text{inv-nodes} = 27-29 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 27-29 \mid Y = \text{recurrence-events}) = 0.01098901098901099$   
 $P(\text{inv-nodes} = 30-32 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 30-32 \mid Y = \text{recurrence-events}) = 0.01098901098901099$   
 $P(\text{inv-nodes} = 33-35 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 33-35 \mid Y = \text{recurrence-events}) = 0.01098901098901099$   
 $P(\text{inv-nodes} = 36-39 \mid Y = \text{no-recurrence-events}) = 0.0049504950495049506$   
 $P(\text{inv-nodes} = 36-39 \mid Y = \text{recurrence-events}) = 0.01098901098901099$   
 $P(\text{node-caps} = \text{yes} \mid Y = \text{no-recurrence-events}) = 0.1256544502617801$   
 $P(\text{node-caps} = \text{yes} \mid Y = \text{recurrence-events}) = 0.4$   
 $P(\text{node-caps} = \text{no} \mid Y = \text{no-recurrence-events}) = 0.8743455497382199$   
 $P(\text{node-caps} = \text{no} \mid Y = \text{recurrence-events}) = 0.6$   
 $P(\text{deg-malig} = 1 \mid Y = \text{no-recurrence-events}) = 0.2916666666666667$   
 $P(\text{deg-malig} = 1 \mid Y = \text{recurrence-events}) = 0.11111111111111111$   
 $P(\text{deg-malig} = 2 \mid Y = \text{no-recurrence-events}) = 0.5104166666666666$   
 $P(\text{deg-malig} = 2 \mid Y = \text{recurrence-events}) = 0.35802469135802467$   
 $P(\text{deg-malig} = 3 \mid Y = \text{no-recurrence-events}) = 0.19791666666666666$   
 $P(\text{deg-malig} = 3 \mid Y = \text{recurrence-events}) = 0.5308641975308642$   
 $P(\text{breast} = \text{left} \mid Y = \text{no-recurrence-events}) = 0.5078534031413613$   
 $P(\text{breast} = \text{left} \mid Y = \text{recurrence-events}) = 0.55$   
 $P(\text{breast} = \text{right} \mid Y = \text{no-recurrence-events}) = 0.49214659685863876$   
 $P(\text{breast} = \text{right} \mid Y = \text{recurrence-events}) = 0.45$   
 $P(\text{breast-quad} = \text{left\_up} \mid Y = \text{no-recurrence-events}) = 0.34536082474226804$   
 $P(\text{breast-quad} = \text{left\_up} \mid Y = \text{recurrence-events}) = 0.30120481927710846$   
 $P(\text{breast-quad} = \text{left\_low} \mid Y = \text{no-recurrence-events}) = 0.36597938144329895$   
 $P(\text{breast-quad} = \text{left\_low} \mid Y = \text{recurrence-events}) = 0.3855421686746988$   
 $P(\text{breast-quad} = \text{right\_up} \mid Y = \text{no-recurrence-events}) = 0.10824742268041238$   
 $P(\text{breast-quad} = \text{right\_up} \mid Y = \text{recurrence-events}) = 0.1686746987951807$

$P(\text{breast-quad} = \text{right\_low} \mid Y = \text{no-recurrence-events}) = 0.09278350515463918$   
 $P(\text{breast-quad} = \text{right\_low} \mid Y = \text{recurrence-events}) = 0.08433734939759036$   
 $P(\text{breast-quad} = \text{central} \mid Y = \text{no-recurrence-events}) = 0.08762886597938144$   
 $P(\text{breast-quad} = \text{central} \mid Y = \text{recurrence-events}) = 0.060240963855421686$   
 $P(\text{irradiat} = \text{yes} \mid Y = \text{no-recurrence-events}) = 0.15706806282722513$   
 $P(\text{irradiat} = \text{yes} \mid Y = \text{recurrence-events}) = 0.3875$   
 $P(\text{irradiat} = \text{no} \mid Y = \text{no-recurrence-events}) = 0.8429319371727748$   
 $P(\text{irradiat} = \text{no} \mid Y = \text{recurrence-events}) = 0.6125$

## 2. The class probabilities $P(Y = y)$ for each class label $Y = y$

$P(Y = \text{no-recurrence-events}) = 0.7063197026022305$   
 $P(Y = \text{recurrence-events}) = 0.2936802973977695$

## 3. For each test instance, given the input vector $X = [X1 = x1, \dots, X9 = x9]$ , give the calculated:

- score( $Y = \text{no-recurrence-events}, X$ )
- score( $Y = \text{recurrence-events}, X$ )
- predicted class of the input vector

Accuracy = 0.8

score( $Y = \text{no-recurrence-events}, ['50-59' \text{'premeno'} '50-54' '0-2' \text{'yes'} 2 \text{'right'} \text{'left\_up'} \text{'yes'}]$ ) = 4.017731924138001e-06

score( $Y = \text{recurrence-events}, ['50-59' \text{'premeno'} '50-54' '0-2' \text{'yes'} 2 \text{'right'} \text{'left\_up'} \text{'yes'}]$ ) = 7.096642912782161e-06

Predicted class: recurrence-events , Actual class: no-recurrence-events

score( $Y = \text{no-recurrence-events}, ['50-59' \text{'ge40'} '35-39' '0-2' \text{'no'} 2 \text{'left'} \text{'left\_up'} \text{'no'}]$ ) = 0.00033365875538186835

score( $Y = \text{recurrence-events}, ['50-59' \text{'ge40'} '35-39' '0-2' \text{'no'} 2 \text{'left'} \text{'left\_up'} \text{'no'}]$ ) = 2.6021024013534593e-05

Predicted class: no-recurrence-events , Actual class: no-recurrence-events

score( $Y = \text{no-recurrence-events}, ['50-59' \text{'premeno'} '10-14' '3-5' \text{'no'} 1 \text{'right'} \text{'left\_up'} \text{'no'}]$ ) = 4.707378797592419e-05

score( $Y = \text{recurrence-events}, ['50-59' \text{'premeno'} '10-14' '3-5' \text{'no'} 1 \text{'right'} \text{'left\_up'} \text{'no'}]$ ) = 9.715041077797865e-07

Predicted class: no-recurrence-events , Actual class: no-recurrence-events

score( $Y = \text{no-recurrence-events}, ['40-49' \text{'premeno'} '10-14' '0-2' \text{'no'} 2 \text{'left'} \text{'left\_low'} \text{'yes'}]$ ) = 0.00015163354504655905

score( $Y = \text{recurrence-events}, ['40-49' \text{'premeno'} '10-14' '0-2' \text{'no'} 2 \text{'left'} \text{'left\_low'} \text{'yes'}]$ ) = 1.0219165794406317e-05

Predicted class: no-recurrence-events , Actual class: no-recurrence-events

score( $Y = \text{no-recurrence-events}, ['50-59' \text{'ge40'} '15-19' '0-2' \text{'yes'} 2 \text{'left'} \text{'central'} \text{'yes'}]$ ) = 4.34523614378287e-06

score(Y = recurrence-events , ['50-59' 'ge40' '15-19' '0-2' 'yes' 2 'left' 'central' 'yes'] ) =  
1.9205993914751727e-06

Predicted class: no-recurrence-events , Actual class: no-recurrence-events

score(Y = no-recurrence-events , ['50-59' 'premeno' '25-29' '0-2' 'no' 1 'left' 'left\_low' 'no'] ) =  
0.0006000123072810078

score(Y = recurrence-events , ['50-59' 'premeno' '25-29' '0-2' 'no' 1 'left' 'left\_low' 'no'] ) =  
3.880403296274196e-05

Predicted class: no-recurrence-events , Actual class: no-recurrence-events

score(Y = no-recurrence-events , ['60-69' 'ge40' '25-29' '0-2' 'no' 3 'right' 'left\_low' 'no'] ) =  
0.0002071279390811688

score(Y = recurrence-events , ['60-69' 'ge40' '25-29' '0-2' 'no' 3 'right' 'left\_low' 'no'] ) =  
7.415569159796725e-05

Predicted class: no-recurrence-events , Actual class: no-recurrence-events

score(Y = no-recurrence-events , ['60-69' 'ge40' '20-24' '0-2' 'no' 1 'right' 'left\_up' 'no'] ) =  
0.00031504865826699285

score(Y = recurrence-events , ['60-69' 'ge40' '20-24' '0-2' 'no' 1 'right' 'left\_up' 'no'] ) =  
8.934762413310772e-06

Predicted class: no-recurrence-events , Actual class: recurrence-events

score(Y = no-recurrence-events , ['40-49' 'ge40' '30-34' '3-5' 'no' 3 'left' 'left\_low' 'no'] ) =  
3.9123945169671666e-05

score(Y = recurrence-events , ['40-49' 'ge40' '30-34' '3-5' 'no' 3 'left' 'left\_low' 'no'] ) =  
6.483884504037111e-05

Predicted class: recurrence-events , Actual class: recurrence-events

score(Y = no-recurrence-events , ['50-59' 'ge40' '30-34' '3-5' 'no' 3 'left' 'left\_low' 'no'] ) =  
4.1017039290784815e-05

score(Y = recurrence-events , ['50-59' 'ge40' '30-34' '3-5' 'no' 3 'left' 'left\_low' 'no'] ) =  
5.283165151437645e-05

Predicted class: recurrence-events , Actual class: recurrence-events