Table

Description automatically generatedתז

תז

Q1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data | Z=i | Z=j | Z=k | Total |
| X=a | 10 | 18 | 15 | 43 |
| X=b | 2 | 10 | 20 | 32 |
| X=c | 12 | 8 | 3 | 23 |
| X=d | 12 | 17 | 9 | 38 |
| Total | 36 | 53 | 47 | 136 |

א)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| H(X) | X=a | X=b | X=c | X=d | Total |
| P(x) | 43/136=0.32 | 32/136=0.24 | 23/136=0.17 | 38/136=0.28 |  |
| - (x) | 1.64 | 2.06 | 2.56 | 1.84 |  |
| -p(x)\* (x) | 0.52 | 0.49 | 0.44 | 0.51 | 1.96 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| H(Z) | Z=i | Z=j | Z=k | Total |
| P(z) | 36/136=0.26 | 53/136=0.39 | 47/136=0.35 |  |
| - (z) | 1.94 | 1.36 | 1.51 |  |
| -p(x)\* (z) | 0.5 | 0.53 | 0.53 | 1.56 |

ב)

H(X)=4\*(-0.25\*=2

H(Z)=3\*(-0.33\*=1.58

ג)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| P(X,Z) | Z=i | Z=j | Z=k | Total |
| X=a | 10/136=0.07 | 18/136=0.13 | 15/136=0.11 | 0.31 |
| X=b | 2/136=0.01 | 10/136=0.07 | 20/136=0.15 | 0.23 |
| X=c | 12/136=0.09 | 8/136=0.06 | 3/136=0.02 | 0.17 |
| X=d | 0.09 | 0.13 | 0.07 | 0.29 |
| Total | 0.26 | 0.39 | 0.35 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| P(X|Z) | Z=i | Z=j | Z=k |
| X=a | 10/36=0.28 | 18/53=0.34 | 15/47=0.32 |
| X=b | 2/36=0.06 | 10/53=0.19 | 20/47=0.43 |
| X=c | 12/36=0.33 | 8/53=0.15 | 3/47=0.06 |
| X=d | 12/36=0.33 | 17/53=0.32 | 9/47=0.19 |
| Total | 1 | 1 | 1 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| P(Z|X) | Z=i | Z=j | Z=k | Total |
| X=a | 10/43=0.23 | 18/43=0.42 | 15/43=0.35 | 1 |
| X=b | 2/32=0.06 | 10/32=0.31 | 20/32=0.63 | 1 |
| X=c | 12/23=0.52 | 8/23=0.35 | 3/23=0.13 | 1 |
| X=d | 12/38=0.31 | 17/38=0.45 | 9/38=0.24 | 1 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Z=i | Z=j | Z=k | Total |
| X=a | 0.07\*=0.15 | 0.13\*0.16 | 0.11\*=0.17 | 0.48 |
| X=b | 0.01\*=0.04 | 0.07\*=0.12 | 0.15\*=0.1 | 0.26 |
| X=c | 0.09\*=0.08 | 0.06\*=0.09 | 0.02\*=0.06 | 0.23 |
| X=d | 0.09\*=0.15 | 0.13\*=0.15 | 0.07\*=0.14 | 0.44 |
| Total | 0.42 | 0.52 | 0.47 | 1.41 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Z=i | Z=j | Z=k | Total |
| X=a | 0.07\*=0.13 | 0.13\*=0.2 | 0.11\*=0.18 | 0.51 |
| X=b | 0.01\*=0.04 | 0.07\*=0.17 | 0.15\*=0.18 | 0.39 |
| X=c | 0.09\*=0.14 | 0.06\*=0.16 | 0.02\*=0.08 | 0.38 |
| X=d | 0.09\*=0.14 | 0.13\*=0.21 | 0.07\*0.17 | 0.52 |
| Total | 0.45 | 0.74 | 0.61 | 1.8 |

ד) I

I

|  |  |  |  |
| --- | --- | --- | --- |
| P(Z) | Z=i | Z=j | Z=k |
|  | 36/136=0.26 | 53/136=0.39 | 47/136=0.35 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| /p(z) | Z=i | Z=j | Z=k | Total |
| X=a | 0.07\*=0.01 | 0.13\*0.01 | 0.11\*=0 | 0.02 |
| X=b | 0.01\*=0.02 | 0.07\*=0.02 | 0.15\*=0.13 | 0.17 |
| X=c | 0.09\*=0.09 | 0.06\*=0.01 | 0.02\*=0.03 | 0.13 |
| X=d | 0.09\*=0.02 | 0.13\*=0.03 | 0.07\*=0.04 | 0.09 |
| Total | 0.14 | 0.07 | 0.2 | 0.41 |

I

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| P(x) | X=a | X=b | X=c | X=d |
|  | 43/136=/0.32 | 32/136=/0.24 | 23/136=/0.17 | 38/136=/0.28 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Z=i | Z=j | Z=k | Total |
| X=a | 0.07\*=0.01 | 0.13\*=0.01 | 0.11\*=0 | 0.02 |
| X=b | 0.01\*=0.02 | 0.07\*=0.02 | 0.15\*=0.13 | 0.17 |
| X=c | 0.09\*=0.09 | 0.06\*=0.01 | 0.02\*=0.03 | 0.13 |
| X=d | 0.09\*=0.02 | 0.13\*=0.03 | 0.07\*0.04 | 0.09 |
| Total | 0.14 | 0.07 | 0.2 | 0.41 |

I = I =0.41

Table

Description automatically generatedQ2

2) *= -0.5log0.5 – 0.5log05=1*

=0.031

=)+ ))

)= = =0.99

)= = =0.92

=+ =0.969

=0.128

)=

= =0.92

)=

= =0.8

)= =0.92

=+ =0.872

=0.515

)=

)=

)= =0.97

==0.485

=0.448

)= =0.92

)=

)= =0.92

)=

==0.552

=0.128

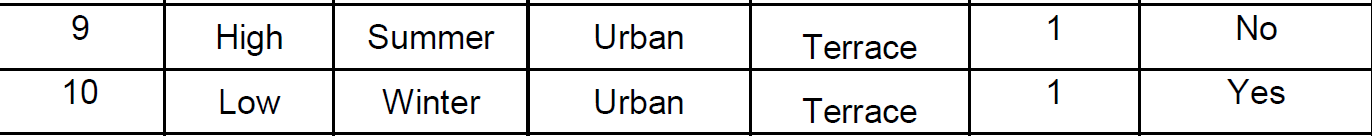
)= =0.92

)= =0.8

==0.872

|  |  |
| --- | --- |
| Attribute | Information Gain |
| Num of floors | 0.031 |
| House type | 0.128 |
| District | 0.515 |
| Season | 0.448 |
| Income | 0.128 |

root



Urban

Yes=3

No=2

Rural

Yes=2

No=0

Suburban

Yes=0

No=3

Table

Description automatically generated

*= -0.4log0.4 – 0.6log06=0.97*

=0.33

)= =0.8

)=

==0.64

=0.018

)= =0.92

)= =1

==0.0.952

=0.97

)=0

)=0

)=0

)==

= 0

=0.018

)=

)= =0.92

==0.872=0.952

|  |  |
| --- | --- |
| Attribute | Information Gain |
| Num of floors | 0.33 |
| House type | 0.018 |
| Season | 0.97 |
| Income | 0.018 |

root

district

Urban

Yes=3

No=2

Rural

Yes=2

No=0

Suburban

Yes=0

No=3

season

Summer

No=1

Yes=0

Autumn

Yes=0

No=1

Spring

Yes=1

No=0

Winter

Yes=2

No=0

3)Majority Accuracy =5/10=50%

4)Test Accuracy=2/6=33% Train Accuracy=1=100%

Normal Approximation to Binomial Distribution: 1>0.33+1.96

1>0.7

5)train accuracy>test accuracy => overfiting

6)

|  |  |
| --- | --- |
| ID | Bought |
| 16 | No |
| 17 | Yes |
| 18 | Yes |

7) q(v)=

q(t)==0.35