Problema 3 PEP2 2022-2

Atributos:

- Tipo de Ovejas = {>26 (A1), 27 ≤ x ≤ 32 (A2), <37 (A3)}
- Progenitor = {Si (P), No ($\bar{P}$)}

Atributo clase:

- Germen = {Si, No}

Probabilidades:

- P(G) = 8/12 = 0,6̄ = 66,6̄ %
- P($\bar{G}$) = 4/12 = 0,3̄ = 33,3̄ %
- P(A1) = 2/12 = 0,16̄ = 16,6̄ %
- P(A2) = 3/12 = 0,25 = 25 %
- P(A3) = 7/12 = 0,58 = 58,3̄ %
- P(P) = 6/12 = 0,5 = 50 %
- P($\bar{P}$) = 6/12 = 0,5 = 50 %

Probabilidades condicionadas:

- P(G|A1) = 0 = 0 %
- P(G|A2) = 1/3 = 0,3̄ = 33,3̄
- P(G|A3) = 1 = 100 %
- P(G|P) = 3/6 = 0,5 = 50 %
- P(G|$\bar{P}$) = 5/6 = 0,83 = 83,3̄ %
- P($\bar{G}$|A1) = 1 = 100 %
- P($\bar{G}$|A2) = 1 - 1/3 = 0,6̄ = 66,6̄ %
- P($\bar{G}$|A3) = 0 = 0 %
- P($\bar{G}$|P) = 0,5 = 50 %
- P($\bar{G}$|$\bar{P}$) = 0,16̄ = 16,6̄ %

Cálculo de Información

- $Inf(Germina) = -(P(G)\cdot lg(P(G)) + P(\bar{G})\cdot lg(P(\bar{G})))$

$= -(8/12 \cdot lg(8/12) + 4/12 \cdot lg(4/12))$

$= 0,918$

- $Inf(Áreas) = -P(A1)\cdot(P(G|A1)\,lg(P(G|A1)) + P(\bar{G}|A1)\cdot lg(P(\bar{G}|A1))) -$
$P(A2)\cdot(P(G|A2)\cdot lg(P(G|A2)) + P(\bar{G}|A2)\cdot lg(P(\bar{G}|A2))) -$
$P(A3)\cdot(P(G|A3)\cdot lg(P(G|A3)) + P(\bar{G}|A3)\cdot lg(P(\bar{G}|A3)))$

$= -(2/12)\cdot(0) - (3/12)\cdot(1/3\cdot lg(1/3) + 2/3\cdot lg(1/3)) -$
$(7/12)\cdot(0)$

$= 0,918$

- $Inf(Progenitor) = -P(P)\cdot(P(G|P)\cdot lg(P(G|P)) + P(\bar{G}|P)\,lg(P(\bar{G}|P)))$
$-P(\bar{P})\cdot(P(G|\bar{P})\cdot lg(P(G|\bar{P})) + P(\bar{G}|\bar{P})\cdot lg(P(\bar{G}|\bar{P}))$

$= -(6/12)\cdot(3/6\cdot lg(3/6) + 3/6\cdot lg(3/6)) - (6/12)\cdot$
$(5/6\cdot lg(5/6) + 1/6\cdot lg(1/6))$

$= 0,825$

$Ganancia(Área) = Inf(Germina) - Inf(G|Área)$

$= 0,918 - 0,918 = 0$

$Ganancia(Progenitor) = Inf(Germina) - Inf(G|Progenitor)$

$= 0,918 - 0,825 = 0,093$

∴ Progenitor es más eficiente para clasificar