

## ARBOLES DE DECISIÓN (Algoritmo ID3)

Genere el correspondiente árbol de decisión a los siguientes datos:

Escolaridad	Hacinamiento	Vivienda	o Servicios	Dependencia	Condición
No Asisten	Hay	Adecuada	Con serv	Alta depend	Pobre
No Asisten	No hay	Adecuada	Sin serv	Alta depend	Pobre
Asisten	No hay	Inadecuada	Sin serv	Sin depend	Pobre
No Asisten	No hay	Adecuada	Sin serv	Sin depend	Pobre
Asisten	No hay	Adecuada	Con serv	Sin depend	No Pobre
Asisten	No hay	Adecuada	Con serv	Sin depend	No Pobre
Asisten	No hay	Adecuada	Con serv	Sin depend	No Pobre
Asisten	No hay	Adecuada	Con serv	Sin depend	No Pobre
Asisten	No hay	Inadecuada	Sin serv	Sin depend	Pobre
No Asisten	No hay	Adecuada	Con serv	Alta depend	Pobre
No Asisten	Hay	Adecuada	Con serv	Sin depend	Pobre
No Asisten	Hay	Adecuada	Con serv	Sin depend	Pobre
No Asisten	Hay	Inadecuada	Con serv	Alta depend	Pobre
No Asisten	Hay	Adecuada	Con serv	Sin depend	Pobre
Asisten	Hay	Adecuada	Con serv	Alta depend	Pobre
Asisten	No hay	Adecuada	Con serv	Sin depend	No Pobre
Asisten	Hay	Adecuada	Con serv	Alta depend	Pobre
Asisten	No hay	Inadecuada	Sin serv	Sin depend	Pobre

Por definicion: Pobre = (+), No pobre = (-)

$$Entropia(S) = Entropia(13,5) = -\frac{13}{18} \log_2 \left( \frac{13}{18} \right) - \frac{5}{18} \log_2 \left( \frac{5}{18} \right) = 0.8524$$

$$Entropia(Escolaridad\_NoAsisten) = -\frac{8}{8} \log_2 \left( \frac{8}{8} \right) - \frac{0}{8} \log_2 \left( \frac{0}{8} \right) = 0$$

$$Entropia(Escolaridad\_Asisten) = -\frac{5}{10} \log_2 \left( \frac{5}{10} \right) - \frac{5}{10} \log_2 \left( \frac{5}{10} \right) = 1$$

$$Ganancia(S, \quad Escolaridad) = 0.852 - \frac{8}{18}(0) - \frac{10}{18}(1) = 0.2968$$

$$Entropia(Hacinamiento\_Hay) = -\frac{7}{7} \log_2 \left( \frac{7}{7} \right) - \frac{0}{7} \log_2 \left( \frac{0}{7} \right) = 0$$

$$Entropia(Hacinamiento\_NoHay) = -\frac{6}{11} \log_2 \left( \frac{6}{11} \right) - \frac{5}{11} \log_2 \left( \frac{5}{11} \right) = 0.994$$

$$Ganancia(S, \quad Hacinamiento) = 0.852 - \frac{7}{18}(0) - \frac{11}{18}(0.994) = 0.2445$$

## ARBOLES DE DECISIÓN (Algoritmo ID3)

$$Entropia(Vivienda\_Adecuada) = -\frac{9}{14} \log_2 \left( \frac{9}{14} \right) - \frac{5}{14} \log_2 \left( \frac{5}{14} \right) = 0.9402$$

$$Entropia(Vivienda\_Inadecuada) = -\frac{4}{4} \log_2 \left( \frac{4}{4} \right) - \frac{0}{4} \log_2 \left( \frac{0}{4} \right) = 0$$

$$Ganancia(S, Vivienda) = 0.852 - \frac{14}{18} (0.9402) - \frac{4}{18} (0) = 0.1207$$

$$Entropia(Servicios\_ConServ) = -\frac{8}{13} \log_2 \left( \frac{8}{13} \right) - \frac{5}{13} \log_2 \left( \frac{5}{13} \right) = 0.961$$

$$Entropia(Servicios\_SinServ) = -\frac{5}{5} \log_2 \left( \frac{5}{5} \right) - \frac{0}{5} \log_2 \left( \frac{0}{5} \right) = 0$$

$$Ganancia(S, Servicios) = 0.852 - \frac{13}{18} (0.961) - \frac{5}{18} (0) = 0.1579$$

$$Entropia(Dependencias\_AltaDep) = -\frac{6}{6} \log_2 \left( \frac{6}{6} \right) - \frac{0}{6} \log_2 \left( \frac{0}{6} \right) = 0$$

$$Entropia(Dependencias\_SinDep) = -\frac{7}{12} \log_2 \left( \frac{7}{12} \right) - \frac{5}{12} \log_2 \left( \frac{5}{12} \right) = 0.9798$$

$$Ganancia(S, Dependencias) = 0.852 - \frac{6}{18} (0) - \frac{12}{18} (0.979) = 0.1988$$

Se selecciona *Escolaridad* por ser la que otorga mayor ganancia con respecto a S

$$Entropia(Escolaridad\_Asisten) = 1$$

$$Entropia(Asisten\_Hay) = -\frac{2}{2} \log_2 \left( \frac{2}{2} \right) - \frac{0}{2} \log_2 \left( \frac{0}{2} \right) = 0$$

$$Entropia(Asisten\_NoHay) = -\frac{3}{8} \log_2 \left( \frac{3}{8} \right) - \frac{5}{8} \log_2 \left( \frac{5}{8} \right) = 0.9544$$

$$Ganancia(Escolaridad\_Asisten, Hacinamiento) = 1 - \frac{2}{10} (0) - \frac{8}{10} (0.9544) = 0.2364$$

$$Entropia(Asisten\_Adecuada) = -\frac{2}{7} \log_2 \left( \frac{2}{7} \right) - \frac{5}{7} \log_2 \left( \frac{5}{7} \right) = 0.8631$$

$$Entropia(Asisten\_Inadecuada) = -\frac{3}{3} \log_2 \left( \frac{3}{3} \right) - \frac{0}{3} \log_2 \left( \frac{0}{3} \right) = 0$$

$$Ganancia(Escolaridad\_Asisten, Vivienda) = 1 - \frac{7}{10} (0.8631) - \frac{3}{10} (0) = 0.3958$$

## ARBOLES DE DECISIÓN (Algoritmo ID3)

$$Entropia(Asisten\_ConServ) = -\frac{2}{7}\log_2\left(\frac{2}{7}\right) - \frac{5}{7}\log_2\left(\frac{5}{7}\right) = 0.8631$$

$$Entropia(Asisten\_SinServ) = -\frac{3}{3}\log_2\left(\frac{3}{3}\right) - \frac{0}{3}\log_2\left(\frac{0}{3}\right) = 0$$

$$Ganancia(Escolaridad\_Asisten, Servicios) = 1 - \frac{7}{10}(0.8631) - \frac{3}{10}(0) = 0.3958$$

$$Entropia(Asisten\_AltaDep) = -\frac{2}{2}\log_2\left(\frac{2}{2}\right) - \frac{0}{2}\log_2\left(\frac{0}{2}\right) = 0$$

$$Entropia(Asisten\_SinDep) = -\frac{3}{8}\log_2\left(\frac{3}{8}\right) - \frac{5}{8}\log_2\left(\frac{5}{8}\right) = 0.9544$$

$$Ganancia(Escolaridad\_Asisten, Dependencias) = 1 - \frac{3}{10}(0) - \frac{8}{10}(0.9544) = 0.2364$$

Como la ganancia de Vivienda y Servicios con respecto a Escolaridad, son iguales, se selecciona cualquiera de las dos, por orden seleccionaremos Vivienda

$$Entropia(Asisten\_Adecuada) = 0.8631$$

$$Entropia(Adecuada\_Hay) = -\frac{6}{6}\log_2\left(\frac{6}{6}\right) - \frac{0}{6}\log_2\left(\frac{0}{6}\right) = 0$$

$$Entropia(Adecuada\_NoHay) = -\frac{3}{8}\log_2\left(\frac{3}{8}\right) - \frac{5}{8}\log_2\left(\frac{5}{8}\right) = 0.9544$$

$$Ganancia(Asisten\_Adecuada, Hacinamiento) = 0.8631 - \frac{6}{14}(0) - \frac{8}{14}(0.9544) = 0.3177$$

$$Entropia(Adecuada\_ConServ) = -\frac{7}{12}\log_2\left(\frac{7}{12}\right) - \frac{5}{12}\log_2\left(\frac{5}{12}\right) = 0.9798$$

$$Entropia(Adecuada\_SinServ) = -\frac{2}{2}\log_2\left(\frac{2}{2}\right) - \frac{0}{2}\log_2\left(\frac{0}{2}\right) = 0$$

$$Ganancia(Asisten\_Adecuada, Servicios) = 0.8631 - \frac{12}{14}(0.9798) - \frac{2}{14}(0) = 0.0232$$

$$Entropia(Adecuada\_AltaDep) = -\frac{5}{5}\log_2\left(\frac{5}{5}\right) - \frac{0}{5}\log_2\left(\frac{0}{5}\right) = 0$$

$$Entropia(Adecuada\_SinDep) = -\frac{4}{9}\log_2\left(\frac{4}{9}\right) - \frac{5}{9}\log_2\left(\frac{5}{9}\right) = 0.991$$

$$Ganancia(Asisten\_Adecuada, Dependencias) = 0.8631 - \frac{5}{14}(0) - \frac{9}{14}(0.991) = 0.226$$

# ARBOLES DE DECISIÓN (Algoritmo ID3)

Se selecciona *Hacinamiento* por ser la que otorga mayor ganancia con respecto a *Vivienda*

$$Entropia(Adecuada\_NoHay) = 0.9544$$

$$Entropia(NoHay\_ConServ) = -\frac{1}{6} \log_2 \left( \frac{1}{6} \right) - \frac{5}{6} \log_2 \left( \frac{5}{6} \right) = 0.65$$

$$Entropia(NoHay\_SinServ) = -\frac{5}{5} \log_2 \left( \frac{5}{5} \right) - \frac{0}{5} \log_2 \left( \frac{0}{5} \right) = 0$$

$$Ganancia(Adecuada\_NoHay, Servicios) = 0.9544 - \frac{6}{11} (0.65) - \frac{5}{11} (0) = 0.5998$$

$$Entropia(NoHay\_AltaDep) = -\frac{2}{2} \log_2 \left( \frac{2}{2} \right) - \frac{0}{2} \log_2 \left( \frac{0}{2} \right) = 0$$

$$Entropia(NoHay\_SinDep) = -\frac{4}{9} \log_2 \left( \frac{4}{9} \right) - \frac{5}{9} \log_2 \left( \frac{5}{9} \right) = 0.991$$

$$Ganancia(Adecuada\_NoHay, Dependencias) = 0.9544 - \frac{2}{11} (0) - \frac{9}{11} (0.991) = 0.1435$$

Se selecciona *Servicios* por ser la que otorga mayor ganancia con respecto a *Hacinamiento*

## Árbol de decisión

