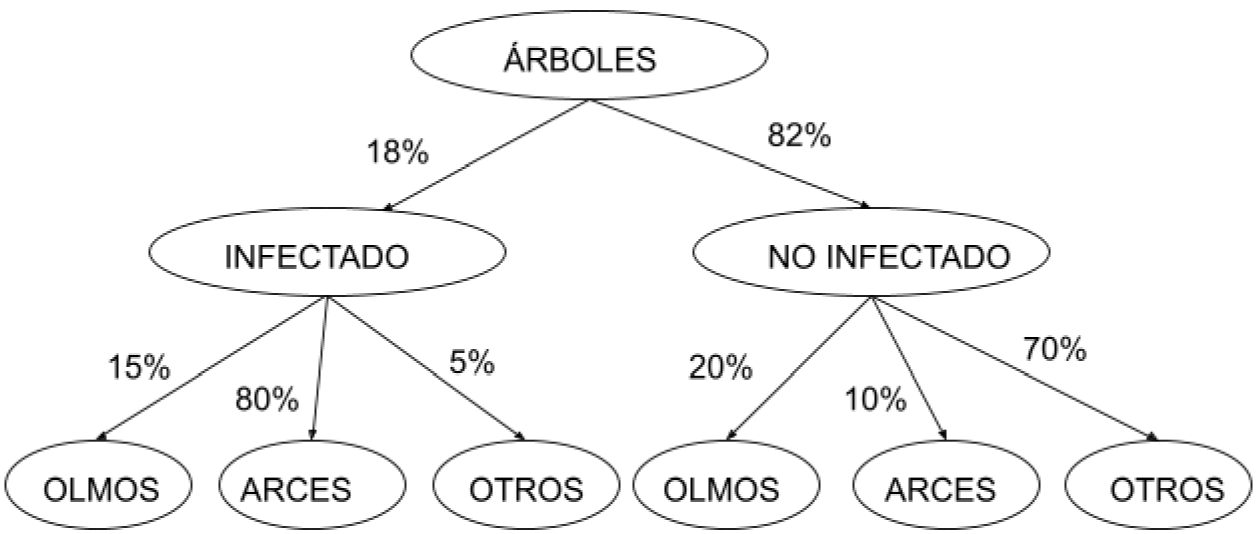
TAREA 2
INFERENCIA II

Exercise 2.5 (Sick trees) A local arboretum contains a variety of tree species, including elms, maples, and others. Unfortunately, 18% of all trees in the arboretum are infected with mold. Among the infected trees, 15% are elms, 80% are maples, and 5% are other species. Among the uninfected trees, 20% are elms, 10% are maples, and 70% are other species. In monitoring the spread of mold, an arboretum employee randomly selects a tree to test.

a. What's the prior probability that the selected tree has mold?
b. The tree happens to be a maple. What's the probability that the employee would have selected a maple?
c. What's the posterior probability that the selected maple tree has mold?
d. Compare the prior and posterior probability of the tree having mold. How did your understanding change in light of the fact that the tree is a maple?



RESPUESTAS:

a. La probabilidad a priori de que un árbol tenga moho es 18%. Al igual que en el ejemplo de las fake news, el problema se reduce a un problema booleano, donde no importa qué tipo de árbol es. Solo se tiene en cuenta la probabilidad previa de que un árbol cualquiera de todo el arbolario esté infectado, qué por letra es 18%.
b. La probabilidad de seleccionar un arce la podemos calcular como la probabilidad de seleccionar un arce infectado más la probabilidad de seleccionar un arce no infectado. Esto es, tal como se ve en el libro P(A/B) + P(A/ B complemento).

Entonces, P(Arce) = P(Arce / Infectado) + P(Arce / No infectado) = 18% * 80% + 10% * 82% = 0,144 + 0,082 = 0,226 = 22,6%.

c. Se pide calcular P(Infectado / Arce).
   Para despejarla utilizaremos el Teorema de Bayes, el cual plantea que:
   P(Infectado / Arce)= (P(Arce / Infectado) *P (Infectado)) / P (Arce) = (18% * 80%) / 22.6% = 63,7168.
   Es decir, dado que el árbol seleccionado fue un arce, la probabilidad que el mismo esté infectado es del 63, 7168%. Es una proporción sumamente alta de arces infectados.
d. El escenario cambia considerablemente. El priori de que un árbol esté infectado era de 18%, pero con la información de que es un arce, la probabilidad aumenta a 63,7% (posteriori).
   Claramente no es lo mismo decir que un árbol está infectado (caso más general) a decir que un árbol está infectado **dado que es un arce**. Si sabemos que es un arce, la probabilidad de que el árbol esté infectado se vuelve considerablemente mayor.
   En este ejemplo, la especie tiene un peso sumamente considerable, dado que la distribución de especies es **muy distinta entre infectados y sanos**.

Exercise 2.9 (Good mood, bad mood) Your roommate has two moods, good or bad. In general, they're in a good mood 40% of the time. Yet you've noticed that their moods are related to how many text messages they receive the day before. If they're in a good mood today, there's a 5% chance they had 0 texts, an 84% chance they had between 1 and 45 texts, and an 11% chance they had more than 45 texts yesterday. If they're in a bad mood today, there's a 13% chance they had 0 texts, an 86% chance they had between 1 and 45 texts, and a 1% chance they had more than 45 texts yesterday.

	Good mood	Bad mood	Total
0 texts	2%	7.8%	9.8%
1 - 45 texts	33.6%	51.6%	85.2%
46+ texts	4.4%	0.6%	5%
Total	40%	60%	100%

a. Use the provided information to fill in the table above.
b. Today's a new day. Without knowing anything about the previous day's text messages, what's the probability that your roommate is in a good mood? What part of the Bayes' Rule equation is this: the prior, likelihood, normalizing constant, or posterior?
c. You surreptitiously took a peek at your roommate's phone (we are attempting to withhold judgment of this dastardly maneuver) and see that your roommate received 50 text messages yesterday. How likely are they to have received this many texts if they're in a good mood today? What part of the Bayes' Rule equation is this?
d. What is the posterior probability that your roommate is in a good mood given that they received 50 text messages yesterday?

RESPUESTAS:

a. Para rellenar la tabla se calcularon las probabilidades de cada suceso.
b. La probabilidad de que el compañero esté en "Good Mood" es 40% = P(A). Este es el priori y es el total de probabilidad de que esté de buen humor, independientemente de la cantidad de mensajes que haya recibido.
c. P(46+ / Good Mood) = P(Good Mood ∩ 46+)/ P(Good Mood)=
(4.4%)/ 40%= 11%. Lo que estamos viendo de la ecuación es la verosimilitud. La

probabilidad de 46 + / Good Mood es igual a la verosimilitud de Good Mood / 46+ y es la Regla de Bayes para eventos.

d. P(Good Mood / 46+) = P(Good Mood $\cap$ 46+) / P(46+) = (4.4% / 5%) = 88%
Estamos calculando la probabilidad posterior. Es decir, dado que recibió 46 mensajes, cual es la probabilidad de que esté de buen humor.
Nuevamente la evidencia (46+ mensajes) cambia radicalmente el panorama donde estamos parados.
Mientras que el prior de que esté de buen humor era 40%, el posterior dado que recibió 46+ mensajes aumenta a 88%.