Introduction to Pattern Recognition and Machine Learning
Homework - Week 2.

Minh Quoc Nguyen
290130 .

From the problem we have the information:

$$P(+ \mid cancer) = 0{,}98 \qquad (1)$$
$$P(- \mid \neg cancer) = 0{,}97 \qquad (2)$$
$$P(cancer) = 0.008 \qquad (3)$$

$(3) \Rightarrow P(\neg cancer) = 1 - P(cancer) = 0{,}992$.

~~Since a person has cancer and a person does not have cancer are mutually exclusive events with $P(cancer) + P(\neg cancer) = 1$.~~
~~Then~~

$$\cancel{P(+) = P(+ \mid cancer)P(cancer) + P(+ \mid \neg cancer)P(\neg cancer)}$$
$$\cancel{\Rightarrow P(+ \mid \neg cancer) = \frac{P(+) - P(+ \mid cancer)P(cancer)}{P(\neg cancer)}}$$
$$\cancel{=}$$

We have

$$P(+ \mid cancer) + P(- \mid cancer) = 1$$
$$\Rightarrow P(- \mid cancer) = 1 - P(+ \mid cancer) = 0{,}02.$$

$$P(+ \mid \neg cancer) + P(- \mid \neg cancer) = 1$$
$$\Rightarrow P(+ \mid \neg cancer) = 1 - P(- \mid \neg cancer) = 0{,}03$$

Maximum a posteri hypothesis h_{MAP} for cancer and not cancer.

$$h^{+}_{MAP} = \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(c \mid D)$$
$$= \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(c \mid +)$$
$$= \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(+ \mid c)P(c).$$
$$= \text{not cancer}$$

$$h^{-}_{MAP} = \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(c \mid -)$$
$$= \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(- \mid c)P(c).$$
$$= \text{not cancer}.$$

Maximum likelihood:

$$h^{+}_{ML} = \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(+ \mid c)$$
$$= \text{cancer}.$$

$$h^{-}_{ML} = \underset{c \in \{cancer, \neg cancer\}}{\arg\max} P(- \mid c)$$
$$= \text{not cancer}.$$

Compare the error if the doctor using the estimate above and all the population go to the doctor.

The doctor uses MAP:

$$P^{MAP}_{error} = P(h_{MAP} = cancer \wedge \neg cancer) + P(h_{MAP} = \neg cancer \wedge cancer)$$
$$= 0 + P(cancer)$$
$$= 0{,}008.$$

Another doctor uses ML:

$$P^{ML}_{error} = P(h_{ML} = cancer \wedge \neg cancer) + P(h_{ML} = \neg cancer \wedge cancer)$$
$$= P(+ \wedge \neg cancer) + P(- \wedge cancer)$$
$$= P(+ \mid \neg cancer)P(\neg cancer) + P(- \mid cancer)P(cancer)$$
$$= 0{,}02992.$$

Now $P^{MAP}_{error} < P^{ML}_{error}$.

but

$$P(h_{MAP} = \neg cancer \wedge cancer) = 0{,}08$$
$$> P(h_{ML} = \neg cancer \wedge cancer) = 0{,}00016.$$