Let $\mathcal{C}_k$ be one of the target classes of the task, in this use case one of the possible states the patient can be while he/she is monitored, and let x_t be the observed sample at time instant t, then, for this use case, $P(\mathcal{C}_k \mid x_t)$ represents the *a posteriori* probability that the patient is in the state corresponding to target class $\mathcal{C}_k$ when the sample x_t has been observed.

Applying the Bayes' rule, considering each channel l as independent of the others (a Naive assumption) and expanding the joint probabilities to take into account the clustering, we get the following expression:

$$P(\mathcal{C}_k \mid x_t) = \frac{p(\mathcal{C}_k, x_t)}{\sum_{i=1}^{K} p(\mathcal{C}_i, x_t)} = \frac{\sum_{l=1}^{L} p(\mathcal{C}_k, x_t^l)}{\sum_{i=1}^{K} \sum_{l=1}^{L} p(\mathcal{C}_i, x_t^l)} = \frac{\sum_{l=1}^{L} \sum_{j=1}^{J} p(\mathcal{C}_k, w_j, x_t^l)}{\sum_{i=1}^{K} \sum_{l=1}^{L} \sum_{j=1}^{J} p(\mathcal{C}_i, w_j, x_t^l)}$$

where

K	is the number of target classes
J	is the number of clusters in the clustering, i.e., the number of mean vectors in the codebook
L	is the number of channels recorded in the EEG
x_t	is the sample at time t including all the channels in the EGG
x_t^l	is the sample at time t corresponding to the l-th channel of the EGG
$p(\mathcal{C}_k, x_t)$	is the joint probability of target class $\mathcal{C}_k$ and sample x_t
$p(\mathcal{C}_k, x_t^l)$	is the joint probability of target class $\mathcal{C}_k$ and sample x_t^l
$p(\mathcal{C}_k, w_j, x_t^l)$	is the joint probability of target class $\mathcal{C}_k$, cluster w_j and sample x_t^l

In more detail

$$p(\mathcal{C}_k, w_j, x_t^l) = p(x_t^l \mid w_j, \mathcal{C}_k) \cdot P(w_j \mid \mathcal{C}_k) \cdot P(\mathcal{C}_k)$$

where

$p(x_t^l \mid w_j, \mathcal{C}_k)$ is the conditional probability density of observing sample x_t^l when the patient is in the state corresponding to target class $\mathcal{C}_k$ and the generated samples in such state fall mainly in cluster w_j, this will be approached by $p(x_t^l \mid w_j)$,

$p(x_t^l \mid w_j)$ is the conditional probability density that observing sample x_t^l belongs to cluster w_j, that is computed simply by

$$p(x_t^l \mid w_j) = \begin{cases} 1 \texttt{ if } x_t^l \texttt{ falls in cluster } w_j \\ 0 \texttt{ otherwise} \end{cases}$$

$P(w_j \mid \mathcal{C}_k)$ is the conditional probability of observing samples falling in cluster w_j when the patient is in the state corresponding to target class $\mathcal{C}_k$, computed as

$$P(w_j \mid \mathcal{C}_k) = \frac{count(w_j, \mathcal{C}_k)}{\sum_{h=1}^{J} count(w_h, \mathcal{C}_k)}$$

$P(\mathcal{C}_k)$ is the *a priori* probablity of target class $\mathcal{C}_k$, computed as as

$$P(\mathcal{C}_k) = \frac{count(\mathcal{C}_k)}{\sum_{i=1}^{K} count(\mathcal{C}_k)}$$

Finally,

$$P(\mathcal{C}_k \mid x_t) \approx \frac{\sum_{l=1}^{L} \sum_{j=1}^{J} p(x_t^l \mid w_j) \cdot P(w_j \mid \mathcal{C}_k) \cdot P(\mathcal{C}_k)}{\sum_{i=1}^{K} \sum_{l=1}^{L} \sum_{j=1}^{J} p(x_t^l \mid w_j) \cdot P(w_j \mid \mathcal{C}_i) \cdot P(\mathcal{C}_i)}$$