The misclassification rate is the probability of a data point being classified into the wrong class.

Minimizing the misclassification rate involves assigning each point to the class for which the posterior probability is highest.

For the two-class case, the misclassification rate is given by:

where R_1 is the region of input space assigned to class C_1 and R_2 is the region assigned to class C_2.

To minimize this, we should assign each value of x to the class for which $p(x,C_k)$ is smaller. Since $p(x,C_k)=p(C_k|x)p(x)$, and the factor $p(x)$ is common to both terms, we can minimize the misclassification rate by assigning each value of x to the class for which the posterior probability $p(C_k|x)$ is largest.