K - nearest neighbor algorithm for classification

1) Given a training set $(X_1, Y_1), \ldots, (X_l, Y_l)$, $Y_i = f(X_i), Y_i \in Y$ (a finite discrete set of values) consider a new point of query X_q
2) Calculate the distances between X_q and all the training data:

$$d(X_q, X_1), \ldots, d(X_q, X_l)$$

3) Take training points $X_{n_1}, X_{n_2}, \ldots, X_{n_k}$ with k smallest distances
4) Assign to $\hat{f}(X_q)$ the majority vote among $Y_{n_1}, Y_{n_2}, \ldots, Y_{n_k}$:

$$\hat{f}(X_q) = \text{argmax}_{Y_i \in Y} \sum_{j=1}^{k} \delta(Y_{n_j}, Y_i)$$

$$\text{where } \delta(a, b) = \begin{cases} 1, & if\ a = b \\ 0, & otherwise \end{cases}$$