

Theoretical part

The weighted vote assigns a weight to each point, the weight can be calculated as the inverse of each points distance from some point x . That is $\frac{1}{\text{Dist}(x, x_i)}$

Considering two target classes c_1, c_2 , without weights x would be assigned to the class with k nearest points. Therefore as k increases the points that are checked increases. If we assume that point x belongs to c_1 but c_2 has more neighbors then the algorithm decides that x belongs to c_2 .

Also, if the cardinality of $c_2 > c_1$ and $k = |c_2|$ then x would always belong to c_2 .

Giving weights to each vote allows the distance to those points to decide where x belongs, a cluster of points some distance d from x will hold a higher weight than another cluster distance $2d$ even if it contains more points.

As the number of k increases the calculation of weight stays the same, therefore if the cardinality of $c_2 > c_1$ and $k = |c_2|$ then the weights of the furthest points will be negligible.