Let $\mathcal{L} = (x_1, ..., x_n)$ – training data with labels (y_1, y_n) and $\mathcal{U}$ – set of examples models doesn't know labels for. Let f – function, interpolating training dataset, with minimum norm (this condition is gotten from the experiments – such models usually have good interpolating properties. Let's define $f_t^u(x)$ - as minimum-norn function that interpolates training data combined with the point $u \in \mathcal{U}$ with label t. The label $t(u)$ we will choose by one of these ways:

$$t^{(1)}(u) = argmin_{t\in\{-1,1\}} \|f_t^u(x)\|,\ t^{(2)}(u) = \begin{cases} +1 \text{ if } f(u) \geq 0, \\ -1 \text{ if } f(u) < 0 \end{cases} \tag{1}$$

Defined $t(u)$, let $f^u(x) = f_{t(u)}^u(x)$. Introduce also *score*-functions:

$$score^{(1)}(u) = \|f^u(x)\|,\ score^{(2)}(u) = \|f^u(x) - f(x)\|, \tag{2}$$

In the first case the most *score* get the least smooth function, in the second case – the function, that differs the most from previous one. We expect that point with the largest *score* is the most imformative.
Then the next point for labeling is

$$u^* = argmax_{u\in\mathcal{U}} score(u).$$

In (1) и (2) we intentionally do not define specific norm for variety of *score*-functions. If these norms are the same and we chose the first *score*-function, then the new point u^* can be determined by

$$\|f^{u^*}(x)\| = max_{u\in\mathcal{U}} min_{t\in\{-1,1\}} \|f_t^u(x)\|. \tag{3}$$