In Perceptron Algorithm, Why use $y-\hat{y}$ Instead of $\hat{y}-y$?

In perceptron algorithm, error = $y - \hat{y}$. so if $\hat{y}$ is bigger than y, error must be negative value, and weight value will be decreasing.
The opposite case, in same logic, the height value will be increasing.

but if we denote error = $\hat{y} - y$, the weight value will be adjusted in the wrong direction.
In this algorithm, if $\hat{y} > y$, $\hat{y}$ must be decreased.
else if $\hat{y} < y$, $\hat{y}$ must be increased. so we use error = $y - \hat{y}$.