First we show that whenever $w_{t-1}^T x_t y_t < 0$ (i.e., an update is performed),

$$\|w_t - w_*\|^2 \le \|w_{t-1} - w_*\|^2 - 1$$

To see this, notice that

$$\|w_t - w_*\|^2 = \left\|w_{t-1} - w_* + y_t \frac{x_t}{\|x_t\|}\right\|^2$$

$$= \|w_{t-1} - w_*\|^2 + \underbrace{2\,\frac{w_{t-1}^T x_t \cdot y_t}{\|x_t\|}}_{\le 0} - 2\,\underbrace{\frac{w_*^T x_t y_t}{\|x_t\|}}_{\ge 1} + \underbrace{\left\|\frac{x_t}{\|x_t\|}\right\|^2}_{=1}$$

$$\le \|w_{t-1} - w_*\|^2 - 1$$

But this means that the algorithm can perform at most $\|w_0 - w_*\|^2$ updates because otherwise $\|w_t - w_*\|^2$ would become negative, which is impossible.