1. Consider the learning problem of logistic regression: Let $X = \{ x \in \mathbb{R}^d : x \leq B \}$, for some scalar $B > 0$, let $Y = \{ \pm 1 \}$, and let the loss function $l$ be defined as $l(w, (x, y)) = \log(1 + \exp(-y \langle w, x \rangle))$. Show that the resulting learning problem is both convex-Lipschitz-bounded and convex-smooth-bounded. Specify the parameters of Lipschitzness and smoothness.

2. Let $q \in (1, 2)$ and consider the $l_q$-norm $\| w \|_q = \left( \sum_{i=1}^d |w_i|^q \right)^{1/q}$. It can be shown that the function $R(w) = \frac{1}{2(q-1)} \| w \|_q^2$ is 1-strongly convex with respect to $\| w \|_q$. Show that if $q = \frac{\log(d)}{\log(d) - 1}$ then $R(w)$ is $1/3 \log(d)$-strongly convex with respect to the $l_1$-norm over $\mathbb{R}^d$.

3. Let $S = ((x_1, y_1), \ldots, (x_m, y_m)) \in (\mathbb{R}^d \times \{ \pm 1 \})^m$. Assume that there exists $w \in \mathbb{R}^d$ such that for every $i \in [m]$ we have $y_i \langle w, x_i \rangle \geq 1$, and let $w^*$ be a vector that has the minimal norm among all vectors that satisfy the preceding requirement. Let $R = \max_i \| x_i \|$. Define a function $f(w) = \max_{i \in [m]} (1 - y_i \langle w, x_i \rangle)$. Show that $\min_{w : \| w \| \leq \| w^* \|} f(w) = 0$.

4. **Growth function of product:** For $i = 1, 2$, let $F_i$ be a set of functions from $X$ to $Y_i$. Define $H = F_1 \times F_2$ to be the Cartesian product class. That is, for every $f_1 \in F_1$ and $f_2 \in F_2$, there exists $h \in H$ such that $h(x) = (f_1(x), f_2(x))$. Prove that $\tau_H(m) \leq \tau_{F_1}(m) \tau_{F_2}(m)$.

5. Let $f : [-1, 1]^n \to [-1, 1]$ be a $\rho$-Lipschitz function. Fix some $\epsilon > 0$. Construct a neural network $N : [-1, 1]^n \to [-1, 1]$, with the sigmoid activation function, such that for every $x \in [-1, 1]^n$ it holds that $|f(x) - N(x)| \leq \epsilon$. 