In a regression task, the outputs are

a. Continuous.

b. Binary.

c. Discrete.

Latent features are

a. Correlated features.
b. Non observable features.
c. Non correlated features.

Let $\mathcal{H}$ be a set of hypotheses. Let $\mathcal{Z} = X \times Y$ be the joint space over the feature space X and label space Y. A loss function ℓ is defined as:

a. $\ell : \mathcal{H} \times Z \to \mathbb{R}$
b. $\ell : \mathcal{X} \times X \to \mathbb{R}^+$
c. $\ell : X \times Y \to \mathbb{R}^+$

Which risk is defined by $\mathbb{E}_{z \sim S} \ell(h, z)$?

a. The true risk.
b. The empirical risk.
c. The generalization risk.

Let $z = (x, y)$ be a labeled data and h a learned classifier. Let $\ell_{0,1}(h, z)$ be the $(0/1)$-loss.

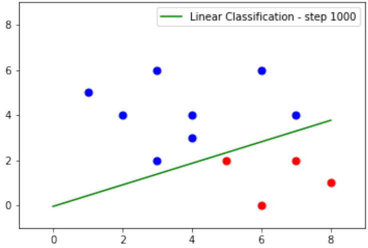
a. $\ell_{0,1}(h, z)$ is convex.
b. $\ell_{0,1}(h, z)$ is differentiable everywhere.
c. $\ell_{0,1}(h, z) = 1$ if $yh(x) \leq 0$ and 0 otherwise.

Let $\theta = (\theta_1, \theta_2)$ be a set of parameters. The gradient of the function $f_\theta(x) = 4\theta_1^2 x + 6y\theta_2\theta_1$ with respect to θ_1 and θ_2 is:

a. $8\theta_1 x + 6y\theta_2$.
b. $(8\theta_1 x + 6y\theta_2, 6y\theta_1)$.
c. $6y\theta_1$.

Which ℓ_p-norm is non convex?

a. $\ell_{0.5}$-norm
b. ℓ_1-norm
c. ℓ_2-norm



Let S be a set of training examples plotted in the figure above.

What can you conclude from the figure?

a. The true risk equals 0.
b. The empirical risk equals 0.
c. Both the true and empirical risks equal 0.

A differentiable function F:

a. decreases fastest if one goes in the direction of the negative gradient of F.
b. decreases fastest if one goes in the direction of the positive gradient of F.
c. increases fastest if one goes in the direction of the negative gradient of F.