Let $h^* \in \mathcal{H}$ be the optimal hypothesis.

a. $h^* = arg \min_{h_i \in \mathcal{H}} \hat{\mathcal{R}}(h_i)$
b. $h^* = min_{h_i \in \mathcal{H}} \mathcal{R}(h_i)$
c. $h^* = arg \min_{h_i \in \mathcal{H}} \mathcal{R}(h_i)$

What does the term PAC mean in statistical learning theory?

a. Probably Almost Correct.
b. Probably Approximately Correct.
c. Potentially Approximately Correct.

In the following generalization bound

$$\mathcal{R}(h) \leq \hat{\mathcal{R}}(h) + \sqrt{\frac{1}{2m} log \frac{2|\mathcal{H}|}{\delta}},$$

$\hat{\mathcal{R}}(h)$ represents:

a. A measure of complexity of h.
b. The bias.
c. The variance.

The VC-dimension of a hyperplane in $\mathbb{R}^3$ is equal to:

a. 3
b. 4
c. 5

Which statement is true?

a. $\hat{\mathcal{R}}(h^*) \leq \hat{\mathcal{R}}(h)$
b. $\mathcal{R}(h^*) \geq \mathcal{R}(h)$
c. $\mathcal{R}(h^*) \leq \mathcal{R}(h)$

The Bayesian error is supposed to be $\epsilon_B = 3\%$. A classifier h has a training error=5% and a validation error=10%.

a. The avoidable error is 2% and the variance is 5%.
b. The avoidable error is 5% and the variance is 2%.
c. The avoidable error is 3% and the variance is 7%.

In the following optimization problem:

$$arg \min_{h_\theta \in \mathcal{H}} \hat{\mathcal{R}}^{\ell}(h_\theta) + \lambda ||\theta||_p^p$$

a. θ, ℓ, λ and p are all parameters.
b. θ, ℓ and p are parameters and λ is a hyperparameter.
c. θ is a parameter and p, ℓ and λ are hyperparameters.