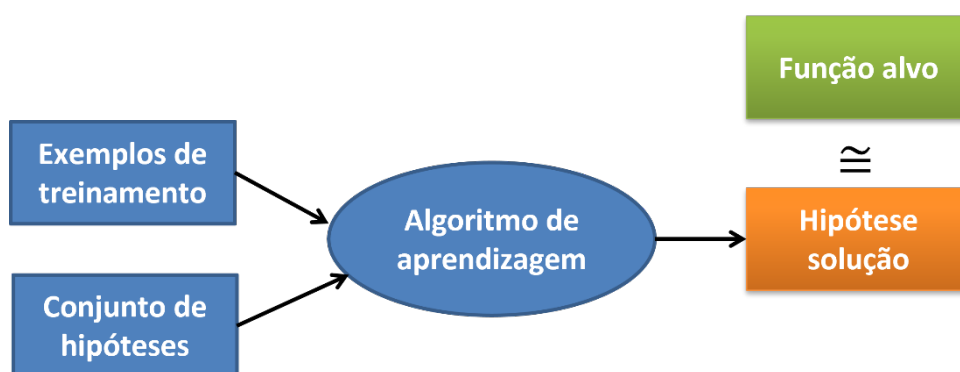
	<b>UNIVERSIDADE FEDERAL DA PARAÍBA</b> <b>CENTRO DE INFORMÁTICA</b>	
	<b>Disciplina</b>	Aprendizagem de Máquina
	<b>Professores</b>	Bruno Jefferson de Sousa Pessoa Gilberto Farias de Sousa Filho

## Lista de Exercícios

(Preparação para a prova)

1. O que o diagrama abaixo como um todo ilustra e o que cada componente representa?



2. Por que, em geral, apenas minimizar  $E_{in}$  não é suficiente? Em alguma situação pode ser suficiente?
3. A desigualdade de Hoeffding, no contexto de aprendizagem de máquina, com respeito a uma determinada hipótese é:

$$P(|E_{in}(h) - E_{out}(h)| > \epsilon) \leq 2e^{-2\epsilon^2 N}$$

Explique o que essa desigualdade significa.

4. Por que a desigualdade acima não é válida para a maioria dos modelos de aprendizagem de máquina? Como adaptá-la para que ela se torne válida? Explique.
5. Por que é importante provar que a função de crescimento para um conjunto de hipóteses é polinomial?
6. Relacione dicotomias com a dimensão VC.
7. É sabido que a dimensão VC relaciona-se com a expressividade do conjunto de hipóteses. Comente sobre isso.

8. Por que não podemos simplesmente trocar  $M$  em  $2Me^{-2\epsilon N}$  pela função de crescimento  $m_H(N)$ ?
9. Como podemos calcular o tamanho mínimo que uma amostra deve ter para que o erro de generalização seja no máximo  $\epsilon$ , com confiança de 95%?
10. Por que garantir apenas que  $|E_{in}(h) - E_{out}(h)| < \epsilon$  não é suficiente?
11. Suponha que você foi contratado para exercer o papel de engenheiro de *machine learning* em uma empresa e está trabalhando na construção de um modelo preditivo com poucos dados disponíveis. Como você usaria a teoria da dimensão VC para ajudá-lo nesse trabalho?
12. Considerando os três principais modelos lineares de aprendizagem de máquina, responda os itens abaixo:
  - Em linhas gerais, explique como os algoritmos de aprendizado funcionam.
  - Descreva a função erro  $E_{in}(w)$  que cada algoritmo procura minimizar.
  - Em quais problemas práticos cada modelo deve ser usado?
13. Para que serve dividir a base de dados em dados de treino e teste durante a fase de aprendizado?
14. O que é *overfitting* na técnica de aprendizado de máquina? Como identificar sua ocorrência?
15. Para quê serve a regularização durante a fase de aprendizado de máquina? Como é sua relação com a função erro medida sobre os dados? Como a validação pode ser usada para executar a regularização?
16. Qual a função da heurística *Weight Decay*? Como ela atua durante o processo de aprendizado de máquina?
17. O que é o processo de validação na técnica de aprendizado de máquina? Quais são suas duas principais funções? Descreva o uso dos dados utilizados no aprendizado dentro do processo de validação.
18. Qual a relação entre o tamanho da amostra de validação e a qualidade da validação realizada?