

Questão 1 :

Generalização: Considere a desigualdade de Hoeffding

$$P(|\nu - \mu| > \epsilon) \leq 2e^{-2\epsilon^2 n}.$$

- (a) Na primeira aula do módulo falámos de como a desigualdade ajuda a prever o erro de generalização de um modelo. Explique como.
- (b) Explique o que consegue garantir relativamente ao erro de generalização se tiver ϵ fixo e n a crescer.

Questão 2 :

Redução de dimensionalidade: O esboço de uma matriz (Matrix Sketch).

- (a) Identifique os factores que diferenciam a aprendizagem em streams da aprendizagem em batch.
- (b) Explique o que é o esboço de uma matriz e para que é usado no contexto de aprendizagem em streams.
- (c) Um dos algoritmos que conhecemos para redução de dimensionalidade é o PCA. Este algoritmo aplica-se diretamente à aprendizagem de dados em stream?