

Disciplina: Aprendizagem de Máquina.
Professor: Tiago B. A. de Carvalho.

Exercícios da Semana 03 – Implementação do k-NN

1. (30 pontos) Construa sua própria implementação do classificador pelo vizinho mais próximo (1 – NN) utilizando distância euclidiana. Avalie este classificador utilizando metade dos exemplos de cada classe da base Iris (archive.ics.uci.edu/ml/datasets/iris) como conjunto de teste e o restante como conjunto de treinamento. *Atenção:* construa também a implementação da distância euclidiana.
2. (15 pontos) Refaça a questão anterior implementando a distância de Minkowski, descrita abaixo. Calcule os resultados para $p = 1$, $p = 2$ e $p = 4$.

$$d(\mathbf{x}_i, \mathbf{x}_j) = \left(\sum_{k=1}^d |x_{ik} - x_{jk}|^p \right)^{1/p}$$

3. (30 pontos) Construa sua própria implementação dos classificadores (a) 7-NN com peso e (b) 7-NN sem peso e avalie os classificadores utilizando metade dos exemplos de cada classe da base Iris como conjunto de teste e a outra metade como conjunto de treinamento.
4. (15 pontos) Utilizando sua implementação do k-NN e, divida a base Wine archive.ics.uci.edu/ml/datasets/Wine utilizando 50% da classe para treino e o restante para teste. Avalie vários valores de k e determine qual é aquele que gera a maior taxa de acerto.
5. (10 pontos) Refaça o experimento da questão anterior removendo a última coluna da base.
 - (a) Compare com os resultados da questão anterior.
 - (b) Calcule a média e o desvio padrão de cada característica e utilize estas informações para justificar a diferença de resultados em relação a questão anterior.