

CC0101: INTRODUÇÃO À MINERAÇÃO DE DADOS
ATIVIDADE 6
ALUNO: DANIEL DE AMARAL DA SILVA

1 Sobre o Conjunto de Dados e algumas observações

O Conjunto de Dados utilizado nessa atividade foi o MNIST, já que trabalhamos com o Iris diversas vezes no semestre.

Como o dataset é grande (70000x28x28), acabei por plotar também o tempo de processamento do algoritmo para fins de comparação. Por ser grande e homogêneo, não houve diferença considerável ao mudar os valores dos mini-lotes k , já que com esse dataset se consegue um bom ajuste mesmo com "poucas" (≈ 23000) observações.

Foi utilizado o Algoritmo 3-NN para o experimento, já que o processo com outros algoritmos de classificação seriam mais custosos.

2 Resultados

2.1 Tabela de Desempenho

k	Acurácia (%)
500	96,64
700	96,65
900	96,64
1000	96,64
2000	96,64
3000	96,65
5000	96,68
10000	96,64
14000	96,64
17000	96,65
20000	96,67

2.2 Visualização dos Resultados

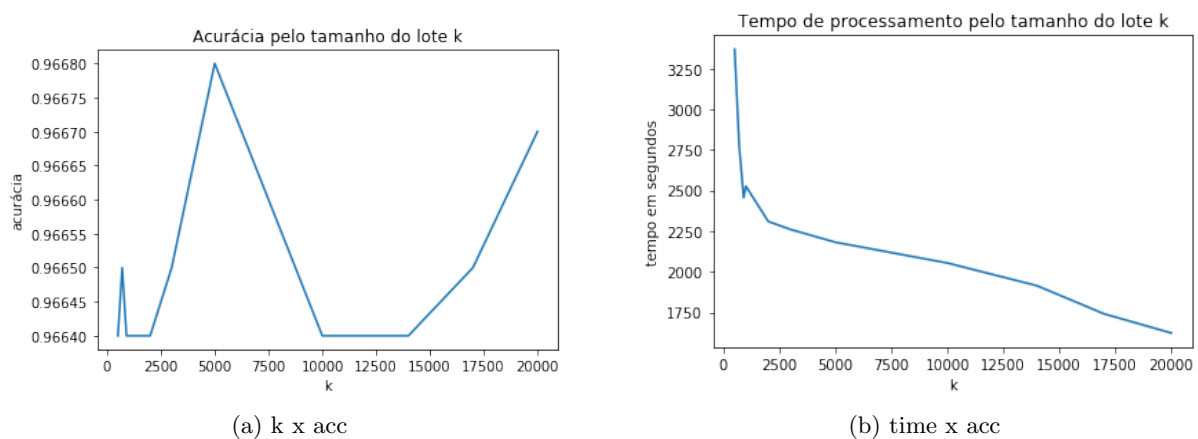


Figura 1: Visualização dos Resultados