

	Universidade Federal do Ceará	
	Disciplina: TÓPICOS AVANÇADOS EM APRENDIZAGEM DE MÁQUINA / TÓPICOS ESPECIAIS EM LÓGICA II	Código: CK0255/CKP8377
	Professor(a): João Paulo P. Gomes / César Lincoln C. Mattos	
	Semestre: 2022.2	
	Discente:	Matrícula:
	Curso: Ciência da Computação	

Trabalho 4

Leia as Instruções:

- O trabalho é individual.
- As simulações poderão ser realizadas em quaisquer linguagens de programação.
- Para a avaliação do trabalho deverá ser submetido um arquivo pdf com texto e figuras referentes aos resultados das simulações.
- Para a avaliação do trabalho deverão ser enviados os códigos fonte.

1. Implemente um modelo de **Probabilistic Principal Components Analysis (PPCA)** para os dígitos “5” do **dataset MNIST**, conforme disponibilizado no arquivo **mnist_5.csv**. O procedimento de inferência deve ser feito via **algoritmo Expectation-Maximization (EM)**.
 - O arquivo possui 6313 dígitos com 784 atributos (imagens com 28×28 pixels).
 - Avalie o seu modelo nas tarefas de **reconstrução** e de **geração** de novas amostras.
 - Experimente com diferentes valores para a dimensão do espaço latente (e.g. 2, 10, 20).
2. Considere os dados unidimensionais disponibilizados no arquivo **gaussian_1d.csv**. Implemente e execute um procedimento de **inferência variacional** para aproximar a posteriori dos parâmetros μ, τ de uma Gaussiana univariada $\mathcal{N}(x|\mu, \tau^{-1})$ a partir desses dados.
 - O arquivo possui 500 valores. Faça experimentos considerando somente os primeiros 10, 50, 100 pontos e, finalmente, todos os pontos.
 - Indique o valor esperado da posteriori aproximada para os parâmetros μ, τ .
 - Ilustre a aproximação obtida plotando as distribuições variacionais para μ, τ .