



Universidade Federal de Pernambuco
Graduação em Ciência da Computação
Centro de Informática



Agrupamento em Subespaço com Mapas de Kohonen em GPU

Proposta de Trabalho de Graduação

Aluno: Eduardo Peixoto Macedo
Orientador: Hansenclever Bassani

Recife, Setembro de 2016

Índice

Introdução	3
Objetivo.....	3
Cronograma	4
Referências.....	5
Possíveis Avaliadores.....	5
Assinaturas	5

Introdução

A Aprendizagem de Máquina evoluiu a partir de estudos nas áreas de reconhecimento de padrões e teoria de aprendizagem computacional. Ela é uma área da Inteligência Artificial que visa o desenvolvimento de técnicas computacionais de aprendizado e a construção de sistemas capazes de adquirir conhecimento automaticamente. Utilizando técnicas de aprendizagem de máquina é possível a construção de algoritmos que podem realizar previsões e tomadas de decisão a partir de um conhecimento obtido previamente.

Existem dois paradigmas principais de Aprendizado:

- Aprendizado Supervisionado:
 - Existe o auxílio de um professor que diz algo a respeito dos objetos observados. O algoritmo recebe uma entrada e é informado da saída desejada. O objetivo é aprender uma forma de mapear as entradas para as saídas. Este método se relaciona com problemas de classificação e de regressão.
- Aprendizado Não-Supervisionado:
 - A informação da saída desejada não é dada ao algoritmo. Mesmo assim ainda é possível identificar estrutura e padrões nos dados. Esta técnica é utilizada em problemas de agrupamento e geração de regras de associação.

Este trabalho dará foco aos métodos de Aprendizado Não-Supervisionado. Mais especificamente, às técnicas *DSSOM* (*Dimension Selective Self-Organizing Maps*) [1] e *LARFDSSOM* (*Local Adaptive Receptive Field Dimension Selective Self-Organizing Maps*) [2] que se baseiam em *SOM* (*Self-Organizing Maps*) [3] para agrupamento em subespaço.

Agrupamento em subespaço é uma extensão da técnica tradicional de *clustering*. Com os avanços tecnológicos, a obtenção de dados se tornou mais fácil, o que resulta em bases de dados maiores e mais complexas. Bases com muitas dimensões muitas vezes têm dados irrelevantes que podem esconder *clusters* e confundir os algoritmos tradicionais. Algoritmos de agrupamento em subespaços são capazes de descobrir *clusters* que existam em múltiplos subespaços.

Objetivo

Este documento apresenta uma proposta de estudo teórico e prático sobre aprendizado não-supervisionado. O objetivo geral do projeto é o estudo, entendimento e implementação das técnicas *DSSOM* e *LARFDSSOM*. Serão utilizadas duas abordagens para implementação das técnicas. A primeira utilizando apenas processamento em CPU e a segunda, utilizando GPU. Ademais, será realizada uma análise comparativa de desempenho envolvendo as duas implementações em ambos os algoritmos.

Cronograma

Atividades	Setembro				Outubro				Novembro				Dezembro			
Levantamento de material e estudo	X	X	X	X	X	X										
Concepção da proposta					X	X	X	X	X	X						
Testes e experimentos									X	X	X	X				
Elaboração do relatório									X	X	X	X				
Preparação da apresentação											X	X	X			

Referências

- [1] BASSANI, HANSENCLEVER F. ; ARAUJO, ALUIZIO F. R. . Dimension Selective Self-Organizing Maps for clustering high dimensional data. In: 2012 International Joint Conference on Neural Networks (IJCNN 2012 Brisbane), 2012, Brisbane. The 2012 International Joint Conference on Neural Networks (IJCNN). p. 1-8.
- [2] Bassani, H.F.; Araujo, A.F.R., "Dimension Selective Self-Organizing Maps With Time-Varying Structure for Subspace and Projected Clustering," Neural Networks and Learning Systems, IEEE Transactions on , vol.PP, no.99, pp.1,1
- [3] T. Kohonen, "The self-organizing map," Proceedings of the IEEE, vol. 78, no. 9, pp. 1464–1480, 1990.

Possíveis Avaliadores

- Tsang Ren
- George Darmington

Assinaturas

Eduardo Peixoto Macedo
(Aluno)

Hansenclever Bassani
(Orientador)