

Um Arcabouço de Alto Desempenho para Análise de Fácies Sísmicas

Julio Cesar Faracco¹, Otávio Napoli¹, Edson Borin¹

¹Instituto de Computação – Universidade Estadual de Campinas (UNICAMP)
São Paulo – Brazil

jcfaracco@gmail.com, onapoli@lmcad.ic.unicamp.br, borin@unicamp.br

Abstract. *Seismic facies analysis is an important geological study to obtain relevant structural information from seismic data. These data are usually obtained by mapping a region using various capture methods and can be relatively large if the analyzed area is on the order of kilometers. Furthermore, with the increasing use of machine learning or deep learning techniques in this type of analysis, handling large data can require a greater amount of computational resources. In the context of seismic facies analysis, there are few tools and libraries dedicated to broadly addressing this area of HPC. As a result, the need for tools that support the most varied processing infrastructures is increasingly necessary. Therefore, the objective of this work is to present a standardized framework that can be easily used for seismic facies analysis, offering the maximum possible of joint acceleration techniques. Finally, we will present some results obtained through the use of this developed tool and how it can benefit geologists or geophysicists.*

Resumo. *A análise de fácies sísmicas é um importante estudo geológico para obtenção de informações estruturais relevantes de dados sísmicos. Esses dados são geralmente obtidos através de mapeamento de uma região utilizando vários métodos de captura e podem ser relativamente grandes se a área analisada for da ordem de quilômetros. Além disso, com o uso cada vez maior de técnicas de aprendizagem de máquina ou aprendizado profundo nesse tipo de análise, o manuseio de dados grandes pode exigir uma quantidade maior de recursos computacionais. No contexto de análise de fácies sísmicas, são poucas as ferramentas e bibliotecas que se dedicam a abordar amplamente essa área de HPC. Com isso, a necessidade de ferramentas que deem suporte as mais variadas infra-estruturas de processamento são cada vez mais necessárias. Sendo assim, o objetivo deste trabalho é apresentar um arcabouço padronizado que possa ser facilmente utilizado para análise de fácies sísmicas, oferecendo o máximo possível de técnicas de aceleração em conjunto. Por fim, apresentaremos alguns resultados obtidos através do uso desse ferramental desenvolvido e como ele pode beneficiar geólogos ou geofísicos.*

1. Introdução

Quando uma companhia de Óleo & Gás deseja perfurar um poço para extração de petróleo em águas profundas, ela possivelmente já conduziu uma série de estudos geológicos para justificar tal operação. É bastante provável que em algum ponto do estudo, ela precisou

realizar uma análise de fácies sísmicas para identificar quais rochas compõem aquela área, qual é o melhor caminho de perfuração até o poço, entre outros exemplos. Isso porque perfurar poços são tarefas caras e complexas. Portanto, uma análise bem executada é crucial para o sucesso da perfuração.

Ainda na análise de fácies sísmicas é comum que, durante o pré-processamento desses dados, sejam escolhidos atributos específicos para destacar características do subsolo e revelar informações úteis ao estudo. Normalmente, o processo é feito utilizando recursos computacionais e muitos especialistas utilizam métodos de aprendizado de máquina ou aprendizado profundo para melhorar ainda mais a qualidade da análise e até mesmo a interpretação dos dados. Existem exemplos na literatura utilizando K-Médias [Napoli et al. 2021, Troccoli et al. 2022], mapas auto-organizáveis (SOM) [Forest et al. 2019] e também utilizando aprendizado profundo [Waldeland et al. 2018].

Apesar da existência de diversos trabalhos na área, a experiência do nosso grupo indica que a reprodução dos experimentos é uma atividade geralmente complexa. Além do mais, um outro item de destaque é a falta de preocupação com desempenho em tempo de execução. Alguns trabalhos até exploram o tema, como em [Wu et al. 2019] que reduz o tempo de treinamento da rede neural utilizada de 2 horas para 3 minutos. Também, há menções nos trabalhos dos autores [Zhao 2018], justamente porque a biblioteca utilizada oferece suporte a GPUs.

2. Objetivos

Este trabalho visa desenvolver um arcabouço com a maioria das técnicas de análise de fácies sísmicas sob a ótica de computação de alto desempenho. Em uma área tão necessária, há uma escassez de ferramentas que possam prover tais técnicas para HPC. Logo, é fundamental disponibilizar ao especialista que o utilizará, a possibilidade de aceleração desses métodos de forma fácil, unificada e descomplicada. Sendo facilmente adaptável a computação paralela em aglomerados computacionais e/ou GPUs quando possível. Também é importante destacar que o arcabouço será primordialmente de código aberto já que as poucas ferramentas existentes são pagas.

3. Estudos de Casos

Esta seção apresenta dois casos de estudos utilizando o arcabouço desenvolvido. Ele consiste na classificação não supervisionada do conjunto de dados público *F3 Netherlands* com 1,3 GB utilizando 4 atributos sísmicos (RMS, Fase Instantânea, Frequência Instantânea e *Sweetness*) para realçar características Litológicas e dois diferentes métodos de aprendizado de máquina como comparação: K-Médias e SOM.

O segundo mostra a execução do trabalho de referência proposto por [Alaudah et al. 2019]. O trabalho propõe a interpretação estratigráfica do conjunto *F3* utilizando redes neurais. Nele, mostraremos ainda como é simples converter uma rede neural para o nosso modelo do arcabouço e, conseqüentemente, obter melhores desempenhos em diferentes arquiteturas.

3.1. Estudo de Caso 1

O arcabouço propõe uma fácil visualização do fluxo de dados. A Figura 1 ilustra o fluxo de execução da classificação onde cada bloco consiste em etapas do processamento e as

arestas os fluxos de dados. Cada bloco já faz parte da API disponibilizada pela ferramenta. Portanto, para que essa mesma estrutura seja executada em diferentes arquiteturas, basta que a mesma seja minimamente descrita no início da execução do processo. Isso inclui, números de GPUs, *threads* e *workers*.

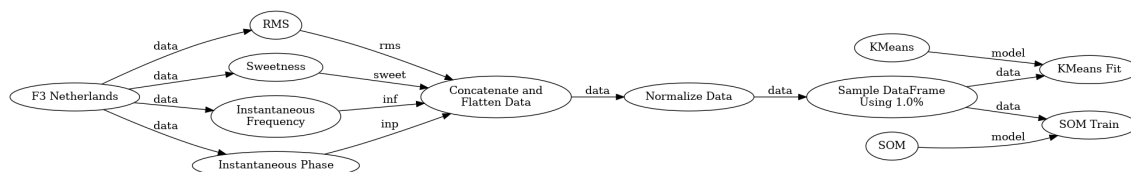


Figura 1. Exemplo de construção de um fluxo de análise de fásias sísmicas para a classificação de atributos sísmicos utilizando K-Médias e SOM.

Em nossa avaliação, executamos esse experimento em 3 diferentes cenários: utilizando somente CPU (Intel(R) Core(TM) i5-8500 com 16 GB de memória), utilizando uma GPU (NVIDIA(R) GTX 1080 com 8 GB de memória) e, por fim, utilizando dois nós com a mesma GPU. Os resultados das execuções são apresentados na Figura 2.

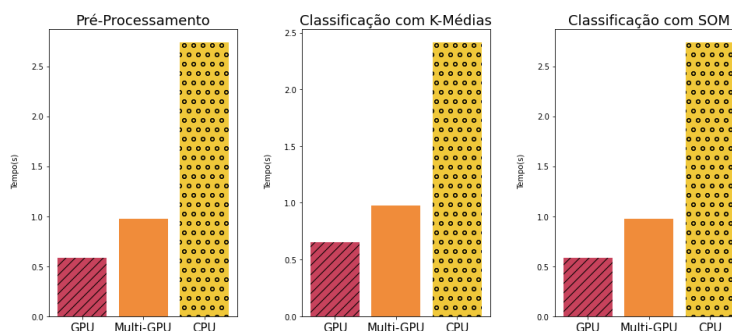


Figura 2. Tempo de execução do fluxo de aprendizado de máquina para 3 infra-estruturas distintas.

No caso da execução em Multi-GPUs, nota-se que o tempo foi maior que pro caso de uma GPU apenas. Existem diversos fatores que podem colaborar para o desempenho ruim do processo: sobrecarga na transferência de dados entre nós, sobrecarga na troca de mensagens com o escalonador ou outras causas.

3.2. Estudo de Caso 2

No segundo estudo de caso, reproduzimos a execução distribuída de uma rede neural. No arcabouço, somente desenvolvimento e inclusão do modelo proposto pelo trabalho é necessária. Toda configuração de execução na plataforma alvo é realizada pelo arcabouço. Dessa forma, utilizando dois dispositivos NVIDIA Tesla A100, testamos uma execução com uma GPU e com dois nós Dask com uma GPU cada. O resultado pode ser visto na Figura 3. Em ambos os casos, a acurácia da predição se manteve em média 80%. Valor compatível com os resultados do trabalho original para esse modelo de rede. Além disso, pode-se perceber que o treino é escalável, mas o tempo total nem tanto.

4. Considerações Finais

Nos dois estudos de caso, mostramos que nosso arcabouço já consegue fazer o uso de técnicas de aceleração para GPUs e múltiplas GPUs. Tanto para modelos mais simples

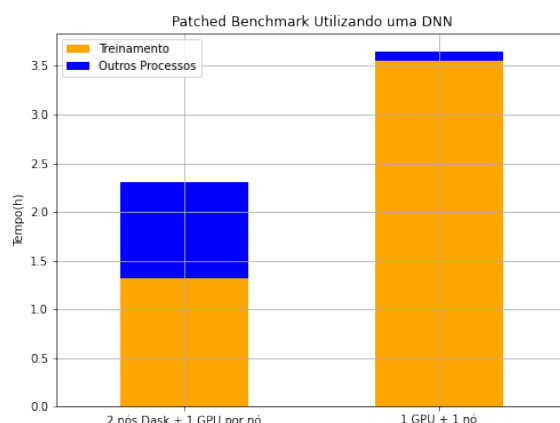


Figura 3. Tempo de execução do treinamento e teste da Rede Neural com duas plataformas diferentes.

como é o caso do K-Médias e o SOM, quanto para modelos mais complexos de aprendizado profundo. Nos próximos passos, iniciaremos estudos de melhorias dos próprios modelos implementados, principalmente para multi-GPUs, que obteve resultados não tão bons como o esperado. Também pretendemos adicionar mais técnicas de aprendizado de máquina e de cálculo de atributos sísmicos ao arcabouço. Por fim, agradecemos ao CNPq (314645/2020-9) e à Fapesp (13/08293-7) pelo apoio financeiro.

Referências

- Alaudah, Y., Michałowicz, P., Alfarraj, M., and AlRegib, G. (2019). A machine-learning benchmark for facies classification. *Interpretation*, 7(3):SE175–SE187.
- Forest, F., Lebbah, M., Azzag, H., and Lacaille, J. (2019). Deep architectures for joint clustering and visualization with self-organizing maps. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 105–116. Springer.
- Napoli, O., Do Rosario, V., De Oliveira, J., Silva, P. E., and Borin, E. (2021). Accelerating multi-attribute unsupervised seismic facies analysis with rapids. volume 2021, pages 1–5. European Association of Geoscientists & Engineers.
- Troccoli, E. B., Cerqueira, A. G., Lemos, J. B., and Holz, M. (2022). K-means clustering using principal component analysis to automate label organization in multi-attribute seismic facies analysis. *Journal of Applied Geophysics*, 198:104555.
- Waldeland, A. U., Jensen, A. C., Gelius, L.-J., and Solberg, A. H. S. (2018). Convolutional neural networks for automated seismic interpretation. *The Leading Edge*, 37(7):529–537.
- Wu, X., Liang, L., Shi, Y., and Fomel, S. (2019). Faultseg3d: Using synthetic data sets to train an end-to-end convolutional neural network for 3d seismic fault segmentation. *Geophysics*, 84(3):IM35–IM45.
- Zhao, T. (2018). Seismic facies classification using different deep convolutional neural networks. In *SEG Technical Program Expanded Abstracts 2018*, pages 2046–2050. Society of Exploration Geophysicists.