Classificador de dados sísmicos para medidas em fundo oceânico

Lucas Lobato, Nicolas Giovanardi, Josafat Ribeiro e Stephan Paul.
Universidade Federal de Santa Catarina. Laboratório de Vibrações e Acústica. Florianópolis, Brasil.

Copyright 2024, SBGf - Sociedade Brasileira de Geofísica

Este texto foi preparado para a apresentação no IX Simpósio Brasileiro de Geofísica, Salvador, 08 a 10 de outubro de 2024. Seu conteúdo foi revisado pelo Comitê Técnico do X SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.

Resumo

Ocean Bottom Nodes são equipamentos utilizados para aquisição de dados de pressão sonora e velocidade no fundo oceânico para prospecção de óleo e gás. Por se tratarem de sensores autônomos, são equipamentos que medem dados ativos, passivos de microssismos. Neste trabalho são avaliados modelos para classificar de traços sísmicos medidos no fundo oceânico em ativos, passivos e microssismo. Os modelos foram desenvolvidos com base em algoritmos de aprendizado de máquina com diferentes *features*. Os resultados são explorados em termos de matriz de confusão e mostram que os modelos utilizando com um maior número de features apresentam os melhores resultados.

Introdução

Na indústria de hidrocarbonetos, a aquisição de dados sísmicos é o método geofísico mais amplamente utilizado e conhecido. técnica (Kearey et al., 2013). De acordo com Kearey et al. (2013), o método de reflexão sísmica, que geralmente consiste em emitir um pulso acústico e medir o som e a vibração refletidos, pode revelar detalhes sobre as estruturas geológicas desde dezenas de metros até toda a litosfera. Em cenários subaquáticos, o reflexo da onda sonora pode ser medido por sistemas de hidrofones, chamados streamers, na superfície do mar. Sistemas de sensores baseados no fundo oceânico, como Ocean Bottom Cables (OBC), sistemas de monitoramento permanente de reservatórios (PRM) e Ocean Bottom Nodes (OBN) para aquisição de dados sísmicos apresentam diversas vantagens quando comparados aos sistemas streamer convencionais, como medição de azimute mais amplo, possibilidade de caracterização de ondas P e S, por meio de pressão sonora e aceleração ou velocidade, respectivamente, ser menos sensível ao ruído produzido por embarcações adjacentes e melhor repetibilidade (Zhang et al., 2021; Hays et al., 2008).

Os dados sísmicos contínuos no tempo medidos são normalmente segmentados, sendo cada segmento denominado *traço*. Os traços podem, a princípio, representar três diferentes tipos: ativo, passivo e microssismo. O primeiro relativo ao uso de alguma fonte sonora, como air-gun. Os traços passivos contêm apenas ruído de fundo, enquanto os traços de microssismos contêm ruído de fundo e a resposta da coluna d'água e fundo do mar devido a excitações de microssismos (Artman, 2006).

O objetivo deste trabalho é construir um classificador de dados sísmicos medidos por sistemas de aquisição sísmica de fundo oceânico. Entretanto, por se tratar de um trabalho inicial, apenas traços de pressão sonora serão considerados. Para isso, dois algoritmos consolidados de aprendizado de máquina são usados e diferentes subconjuntos de *features* são testados.

Metodologia

Os dados sísmicos do campo de Jubarte (Thedy et al., 2014) foram utilizados neste trabalho, separados em conjunto A (dado ativo, estação 510), B (dado passivo, L0048C) e C (dado de microssismo, L0017). Todos os conjuntos foram gentilmente cedidos pela Petrobras S.A. em formato SEG-Y. Neste trabalho, apenas os dados de pressão sonora foram considerados. Dois diferentes algoritmos de classificação foram testados: um baseado em árvore de decisão, denominado aqui DT, usando o modelo *ExtraTreeClassifier* implementado com o pacote Scikit-learn 1.3.2 e outro baseado em redes neurais, denominado aqui NN, usando o modelo Sequential, implementado com o pacote Keras 2.15.0. Foi usada uma rede neural *feed-forward*, foi construída com quatro camadas, sendo o número de neurônios e função de ativação de cada camada apresentados na Tabela 1. A dimensão de entrada é definida pelo número de *features* utilizadas. O otimizador Adam, a função de perda *categorical_crossentropy* e 80 épocas foram usados no processo de treinamento.

Tabela 1: Descrição da rede neural utilizada.

Camada	Neurônios	Função
1	500	Relu
2	100	Relu
3	50	Relu
4	3	Softmax

Como *features*, foram utilizados dois subconjuntos de features, chamados s1 e s2. O subconjunto s1 é composto apenas pela *feature* de valor rms (*root mean square*), enquanto o subconjunto s2 é composto pelo valor rms, percentil 80%, desvio padrão, curtose e

skewness. Todas as *features* foram calculadas sobre a forma de onda bruta, sem pré-processamento. Foram utilizados 1500 traços de pressão sonora (500 ativos, 500 passivos e 500 de microssismos), sendo 50% usados para treinamento e 50% para teste, com a separação feita aleatoriamente usando a função *train_test_split* do pacote Scikit-learn 1.3.2.

Resultados

A Figura 1 apresenta os resultados dos classificadores em termos de matrizes de confusão para classificador DT com features s1 e s2 e classificador NN com features s1 e s2. Os resultados mostram que o desempenho do modelo de classificação dos traços ativos não foi afetado pelo conjunto de características testadas neste estudo. Para os traços ativos, o *recall* e *precision* (Grandini et al., 2020) foram de 100% para ambos os modelos DT e NN. Porém, para a classificação de traços passivos e de microssismo, o conjunto de características utilizadas apresentou um impacto considerável.

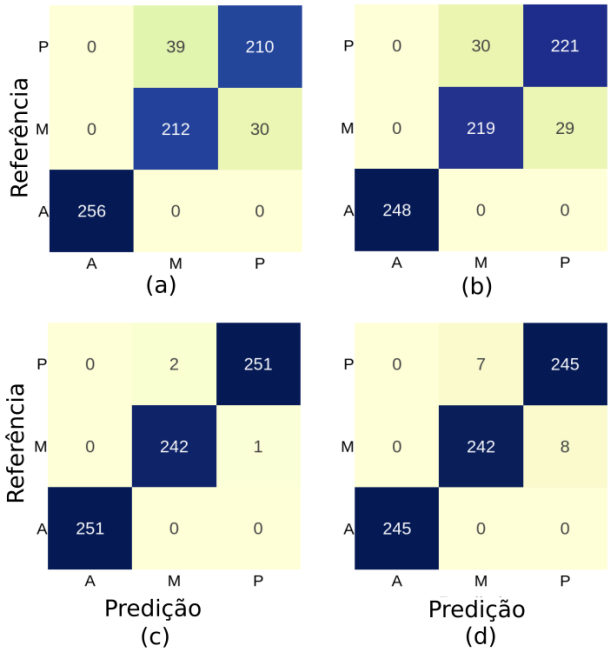

Figura 1: Matrizes de confusão dos resultados de teste para *features* s1 com classificador DT (a) e NN (b) e com *features* s2 com classificador DT (c) e NN (d).

Ao usar o subconjunto de *features* s1, os modelos DT e NN apresentaram acurácias de 90,8% e 92,1%, respectivamente. O modelo NN, para o subconjunto s1, apresentou resultados ligeiramente melhores na classificação de dados passivos e microssismos, com *recall* e *precision* de 88,3% e 87,9%, respectivamente, para os dados de microssismos e 88,0% e 88,4%, respectivamente, para os dados passivos. Ao ser utilizado o subconjunto de *features* s2, ou seja, usando todas as cinco métricas descritas anteriormente, a performance dos modelos aumentaram consideravelmente.

Com o subconjunto de *features* s2, os modelos DT e NN apresentaram acurácias de 99,6% e 98,0%, respectivamente. Para este caso, o modelo DT apresentou resultados ligeiramente melhores. Para os dados passivos o *recall* e *precision* foram de 99,2% 99,6%, enquanto para o conjunto de dados de microssismos o *recall* e *precision* foram de 99,5% e 99,2%.

Discussão e Conclusões

Neste estudo, quatro modelos, utilizando algoritmos de aprendizado de máquina de árvore de decisão e redes neurais com dois conjuntos diferentes de*features* diferentes, foram desenvolvidos e comparados. Embora análises prévias realizadas (e não apresentadas aqui por brevidade) tenha mostrado pouca distinção entre traços passivos e de microssismo, os resultados alcançados usando modelos avaliados indicaram que os modelos que utilizam o subconjunto de *features* s2 (compreendendo cinco *features*) superaram aqueles com o subconjunto s1 (contendo apenas uma *feature*). Os resultados mostraram que os modelos são perfeitos na distinção entre traços ativos ou não, o que é esperado dada a natureza e diferença dos dados ativos para os demais. Além disso, os modelos também apresentaram uma boa capacidade de distinção entre os dados passivos e passivos com microssismos, especialmente utilizando o subconjunto de *features* s2.

Embora o classificador descrito neste trabalho não se destine a uma aplicação específica, tais classificadores de traços sísmicos podem ser usados em diversas aplicações. Por exemplo, um sistema de gestão de energia para OBNs pode ser criado para ativar o registo apenas quando forem identificados eventos ativos ou de microssismos. Além disso, um sistema de comunicação autónomo pode ser ativado para transmissão de um tipo específico de traços. Em todos esses cenários, um classificador de traços é essencial e os modelos discutidos neste estudo podem ser benéficos.

Apesar de bons resultados terem sido alcançados, alguns aspectos ainda merecem estudo e atenção. Nas próximas etapas, será investigado como essas configurações de algoritmo de classificação e conjunto de *features* são capazes de classificar traços ativos, passivos e microssismos em termos de aceleração e velocidade. Além disso, o algoritmo de rede neural permite inúmeras arquiteturas por meio de número de camadas, número de neurônios, otimizador, épocas utilizadas no processo de treinamento, funções de ativação e assim por diante. Portanto, uma otimização relacionada aos hiperparâmetros do classificador de traços baseado em redes neurais também será considerada em um futuro próximo, a depender da performance deste para classificação de traços em termos de aceleração e velocidade.

Agradecimentos

Os autores gostariam de agradecer à Petrobras S.A. pelo apoio financeiro.

Referências

Artman, B., 2006, Imaging passive seismic data: Geophysics, 71, SI177–SI187, doi: 10.1190/1.2209748.

Grandini, M., E. Bagli, and G. Visan, 2020, Metrics for multi-class classification: an overview: Presented at the . doi: 10.48550/arXiv.2008.05756.

Hays, D., K. Craft, P. Docherty, and F. Smit, 2008, in An Ocean Bottom Seismic node repeatability study: 55–59.

Kearey, P., M. Brooks, and I. Hill, 2013, An introduction to geophysical exploration: Wiley.

Thedy, E. A., R. Filho, W. L. Johann, S. N. Seth, S. C. Souza, and P. E. Murray, 2014, in Jubarte Permanent Reservoir Monitoring – Installation and First Results. 13th International Congress of the Brazilian Geophysical Society. 918–920.

Zhang, D., C. Tsingas, A. A. Ghamdi, M. Huang, W. Jeong, K. K. Sliz, S. M. Aldeghaither, and S. A. Zahrani, 2021, A review of OBN processing: challenges and solutions: Journal of Geophysics and Engineering, 18, 492–502, doi: 10.1093/jge/gxab024. Braz. J. Geophys.