

Aplicação de Técnicas de Aprendizado de Máquina para Predição de Propriedades Petrofísicas em reservatórios do Pré-sal

Palavras-Chave: Machine Learning, Geofísica, Perfilagem de poços

Autores/as:

Fernando Antonio Novais Vieira – IG/UNICAMP

Prof. Dr. Gelvam André Hartmann (orientador) – IG/UNICAMP

INTRODUÇÃO:

O Pré-Sal é um importante reservatório de petróleo de interesse econômico. Porém, ele é de grande desafio para exploração por conta da heterogeneidade de formações geológicas carbonáticas e a espessa camada de sal que precede as rochas-reservatório (RICCOMINI; SANT'ANNA; TASSINARI, 2012). A grande heterogeneidade de características texturais e petrofísicas encontradas nas rochas carbonáticas impõem certos desafios, como a correlação de dados de porosidade e permeabilidade para que seja possível se identificar os melhores intervalos de reservatórios carbonáticos (AHR, 2008). Estes desafios implicam na necessidade de se obter inúmeras amostras a cada perfuração, para a análise laboratorial minuciosa de características petrofísicas e petrográficas, o que pode requerer tempo e investimentos. Assim como outros reservatórios mundiais, o pré-sal conta com ferramentas de perfilagem de poços, como o *Log While Drilling* (LWD) (DA ROCHA et al., 2019; NEVES et al., 2019; OLIVEIRA et al., 2019). A relação, por meio digital, dos diferentes perfis coletados ao longo da perfuração facilita o trabalho de avaliação das propriedades petrofísicas em intervalos de interesse, porém com o extenso tempo que as interpretações dos dados requerem, bem como as incertezas inerentes a elas.

Desta forma, novas tecnologias vêm sendo utilizadas como forma de automatizar o processo de identificação de intervalos promissores e preenchimento de lacunas de informação em certos poços. Uma dessas ferramentas é a utilização de técnicas de aprendizado de máquina (Machine Learning – ML). Por meio da aplicação de algoritmos de ML aos dados de perfilagem, é possível se construir associações mais eficazes que as modelagens tradicionais e adquirir relações mais

rápidas para os intervalos geológicos de maior interesse, em menor tempo do que aqueles determinados manualmente (QI; CARR, 2006; SILVA et al., 2015, 2020). Estas aplicações, associadas a ferramentas computacionais de análise de robustez (LIU; ÖZSU, 2009), permitem uma maior confiança na predição de características inerentes aos reservatórios carbonáticos, sejam elas qualitativas ou quantitativas.

Neste trabalho, pretende-se empregar métodos de ML aos dados extraídos na Bacia de Santos e fornecidos publicamente pela Agência Nacional de Petróleo, Gás e Biocombustíveis (ANP) com o objetivo de contribuir na predição e compreensão de propriedades petrofísicas em alvos do pré-sal brasileiro na Bacia de Santos. A presente pesquisa é desenvolvida por meio do Programa de Formação de Recursos Humanos em Exploração Petrolífera e Geologia de Reservatórios (PRH-ANP 19.1), desenvolvida no Instituto de Geociências da UNICAMP e conta com financiamento da Agência Nacional do Petróleo (ANP) e da Financiadora de Estudos e Pesquisas (FINEP).

ALVO DE ESTUDO E DADOS

A Bacia de Santos se apresenta como uma importante porção do Pré-Sal (Figura 1), com uma extensão de 352.000 km² e lâmina d'água com profundidades de até 3 km (GOMES et al., 2020), apresentando condições favoráveis para a exploração de petróleo na Formação Barra Velha. Estas características tornam a utilização das ferramentas computacionais de ML atraentes para o contexto da Bacia de Santos, uma vez que ela se apresenta como uma importante fonte de hidrocarbonetos, com grande interesse econômico. Dessa forma, os campos de Búzios e Mero foram escolhidos como alvos para a obtenção dos dados utilizados na pesquisa, pois apresentam uma base de dados robusta e substancial para aplicação de técnicas de ML.

Para o desenvolvimento dos testes iniciais foram selecionados dois poços inseridos no campo de Búzios. O poço A, está localizado na região central do campo enquanto o poço B localizado a cerca de 3,7 km a noroeste do poço A. Os dois poços apresentam conjuntos de dados semelhantes, os quais foram importados para o software utilizado para tratar este tipo de dado geofísico, no caso o software Techlog, da Schlumberger. Assim, foi possível estabelecer os intervalos para análise e predição das propriedades petrofísicas. Uma vez que a maior parte das rochas-reservatório da Bacia de Santos está localizada na Formação Barra Velha, optou-se por restringir o intervalo de profundidades à extensão vertical desta formação.

Até o presente momento, as curvas de raios gama (GR), porosidade de nêutrons (NPHI), fator fotoelétrico (PEF), densidade volumétrica (RHOZ), resistividade da formação (RT), micro resistividade (RXOZ) e saturação de água (WS) foram escolhidas para o treinamento dos

algoritmos. Este conjunto de curvas foi interpolado com 30 cm de resolução espacial, devido às classificações de fácies estarem descontínuas ao longo do intervalo.

METODOLOGIA

O fluxo de trabalho consiste: (1) seleção de poços de interesse, que depende do conjunto de dados disponíveis (2) tratamento e preparação dos dados para a sua posterior utilização nos algoritmos; (3) emprego das ferramentas de ML; (4) análises estatísticas de robustez dos algoritmos.

Os procedimentos de tratamento e preparação dos dados consistem na filtragem de ruídos nas curvas petrofísicas, na substituição ou remoção de trechos do conjunto de dados que apresentem a curva de calibre de poço (*caliper*) com arrombamento e remoção de dados nulos. Além destes procedimentos, foram realizadas associações de fácies com os dados de perfilagem, por meio da observação de perfis de acompanhamento geológico e utilização de descrições de lâminas, ambos fornecidos nos dados da ANP. Após a fase de tratamento, os dados são compilados em um arquivo de texto, contendo todas as informações das curvas e fácies.

Os scripts de ML estão sendo desenvolvidos em linguagem Python. A linguagem Python conta com bibliotecas de algoritmos, que podem ser facilmente instaladas de acordo com a necessidade do usuário. Uma dessas bibliotecas, o SciKit Learn, é a biblioteca utilizada em Python para trabalhos envolvendo inteligência artificial, e a biblioteca a ser utilizada no desenvolvimento deste projeto. Os algoritmos que devem ser utilizados neste projeto para determinação das predições são: *Decision Trees* (DT), *Random Forest* (RF) e *K-Nearest Neighbors* (KNN).

O modo como cada algoritmo funciona se baseia em diferentes tratamentos estatísticos para o estabelecimento de relações entre atributos variados e um atributo alvo. Por meio de observações contínuas de atributos, que no caso são as curvas de características petrofísicas, os algoritmos podem ser treinados para a construção de relações entre as características e as classes ou variáveis alvo, que no caso são as fácies. Dessa forma, é possível identificar ou realizar a predição das classes ou variáveis em um novo conjunto de dados.

RESULTADOS PRELIMINARES

Pelo algoritmo DT, alcançou-se uma taxa de precisão de cerca de 80% nas predições intra-poço, utilizando-se os dados do poço A. A métrica de precisão se baseia na razão entre o número de positivos verdadeiros e a soma de positivos verdadeiros e falso-positivos. A taxa de acerto no treinamento ficou em torno de 84%. Para os testes, foi estabelecido que 30% das amostras seriam

utilizadas para treinamento. A figura 1 representa as fácies encontradas no poço e as fácies previstas pelo algoritmo. Pela codificação por números inteiros para a utilização dentro da linguagem Python, 12 classes de fácies foram totalizadas para um dos poços. Optou-se, no primeiro teste, por realizar-se um ensaio intra-poço. Neste ensaio, foi utilizado o método *train-test split*, que randomiza as amostras escolhidas para o treinamento a cada execução do código.

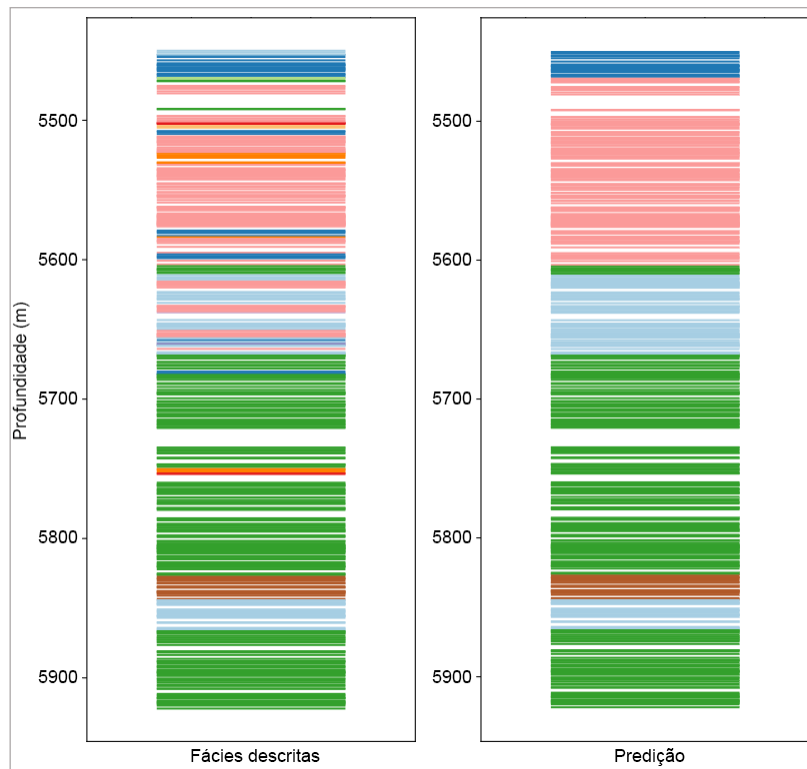


Figura 1 - Resultado Gráfico de teste preliminar, utilizando-se algoritmo DT, com 30% das amostras reservadas para teste.

DISCUSSÃO E CONCLUSÕES

Os resultados preliminares indicam que a predição de fácies por meio de algoritmos de ML em dados obtidos na Bacia de Santos é promissora. Apesar de randomizar a escolha das amostras para treinamento, foi utilizado apenas 30% delas, obtendo-se uma taxa de acerto em torno de 80%. Isso mostra uma importante relação entre o número de amostras exigido para que se obtenha taxas de acerto satisfatórias. Observou-se, também, que a maior parte dos erros de predição cometidos pelo algoritmo está localizada em regiões do poço que apresentam classes de fácies pouco presentes ao longo do intervalo, o que sugere uma necessidade de balanceamento do número de amostras por classe de fácies no conjunto de dados. Esta melhoria na qualidade do conjunto de dados fornecido para o treinamento do algoritmo deve ser abordada nas próximas tentativas de predição.

Como possibilidade de aumento na taxa de acerto, também é possível se utilizar a validação cruzada. Este tipo de validação terá a sua implementação possibilitada pela padronização do número de classes nos poços, bem como pelo balanceamento do número de amostras por classe.

BIBLIOGRAFIA

AHR, W. M. **Geology of Carbonate Reservoirs**. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2008.

DA ROCHA, H. O. et al. Petrophysical characterization using well log resistivity and rock grain specific surface area in a fractured carbonate pre-salt reservoir in the Santos Basin, Brazil. **Journal of Petroleum Science and Engineering**, v. 183, p. 106372, dez. 2019.

GOMES, J. P. et al. Facies classification and patterns of lacustrine carbonate deposition of the Barra Velha Formation, Santos Basin, Brazilian Pre-salt. **Marine and Petroleum Geology**, v. 113, p. 104176, mar. 2020.

LIU, L.; ÖZSU, M. T. (EDS.). **Encyclopedia of database systems**. New York: Springer, 2009.

NEVES, I. DE A. et al. Presalt reservoirs of the Santos Basin: Cyclicity, electrofacies, and tectonic-sedimentary evolution. **Interpretation**, v. 7, n. 4, p. SH33–SH43, 1 nov. 2019.

OLIVEIRA, V. C. B. et al. Electrofacies Characterization in Lacustrine Coquinas and Hybrid Deposits from Rift Phase: Pre-Salt, lower Cretaceous, Campos Basin, Brazil. **Anuário do Instituto de Geociências - UFRJ**, v. 42, n. 3, p. 178–191, 30 set. 2019.

QI, L.; CARR, T. R. Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas. **Computers & Geosciences**, v. 32, n. 7, p. 947–964, ago. 2006.

RICCOMINI, C.; SANT'ANNA, L. G.; TASSINARI, C. C. G. Pré-sal: geologia e exploração. **Revista USP**, v. 0, n. 95, p. 33, 30 nov. 2012.

SILVA, A. A. et al. Artificial neural networks to support petrographic classification of carbonate-siliciclastic rocks using well logs and textural information. **Journal of Applied Geophysics**, v. 117, p. 118–125, jun. 2015.

SILVA, A. A. et al. Petrofacies classification using machine learning algorithms. **GEOPHYSICS**, v. 85, n. 4, p. WA101–WA113, 1 jul. 2020.