

# INTERPRETAÇÃO DE PERFIS LITOLÓGICOS UTILIZANDO INTELIGÊNCIA ARTIFICIAL

GUSTAVO GARCIA<sup>1</sup>; DANIEL PALOMINO<sup>2</sup>

<sup>1</sup>Universidade Federal de Pelotas – [gustavogarcianh@hotmail.com](mailto:gustavogarcianh@hotmail.com)

<sup>2</sup>Universidade Federal de Pelotas – [dpalomino@inf.ufpel.com.br](mailto:dpalomino@inf.ufpel.com.br)

## 1. INTRODUÇÃO

Em pesquisa publicada no ano de 2019 pela Deloitte destaca-se que a indústria do petróleo se encontra em último lugar no que se refere à maturidade digital. No índice definido pela pesquisa, a indústria do petróleo apresenta um índice de maturidade de 1,3, valor bem abaixo quando comparado com outros setores como o automotivo (2,8) e o bancário (2,9). Dentre as principais barreiras para a adoção de iniciativas digitais são destacadas a falta de habilidades digitais na força de trabalho (5º) e de investimento (9º). Sendo assim, são necessários esforços para aumentar a integração de tecnologias digitais aos processos da indústria do petróleo aumentando a automação e reduzindo custos.

Dentre os vários temas que podem ser endereçados com a integração de tecnologias digitais, destaca-se o relacionado a caracterização geológica de reservatórios, mais especificamente em relação à área de petrofísica. Os dados petrofísicos apresentam características que possibilitam o emprego de tecnologias digitais, como volume, variabilidade e visualização (XU et al., 2019). Adicionalmente, um dos temas de caracterização de reservatórios que será abordado neste projeto através da integração de aprendizado de máquina é a interpretação de perfis.

Como descrito por ANTARIKSA et al. (2022), as interpretações dos dados de perfilagem desempenham um papel crítico na avaliação da litologia do reservatório, impactando na exploração do mesmo (SHI et al., 2022). Os dados de perfilagem consistem no registro da magnitude de uma propriedade física de uma formação rochosa, como a resistividade, por exemplo.

Ressalta-se que a implementação do aprendizado de máquina é vantajosa por acelerar o processo de caracterização de reservatórios, que usualmente é dependente da avaliação de profissionais especializados, os quais necessitam de anos de experiência para adquirir o conhecimento necessário para interpretação e consequente caracterização. Adicionalmente, destaca-se possibilidade de identificação de padrões e oportunidades de aprimoramento a interpretações previamente realizadas por especialistas.

Portanto, o presente trabalho irá focar na utilização de aprendizado de máquina como ferramenta para auxiliar o processo de interpretação de perfis para classificação litológica, objetivando identificar os melhores métodos, com o desempenho mais adequado ao banco de dados proposto.

## 2. METODOLOGIA

Para iniciar o trabalho, a primeira etapa consistiu em obter acesso ao banco de dados, o qual foi localizado dentro do repositório de informações da Agência Nacional do Petróleo (ANP). Em seguida, procedemos à verificação da presença das interpretações litológicas neste banco de dados, de modo a prosseguir com a pesquisa. Além disso, conduzimos uma análise estatística abrangente desse

conjunto de informações. Foi constatado que o banco de dados continha informações cruciais, incluindo a profundidade de cada leitura realizada, os perfis litológicos, como resistividade, densidade e raio gama, a impedância acústica e impedância acústica real, a porosidade e, por fim, as *facies*. É importante destacar que o tamanho original deste banco de dados é de 4.907 registros.

Foi necessário realizar primeiramente um pré-processamento dos dados, substituindo a separação decimal de “vírgula” por “ponto”, possibilitando a correta leitura pelo software. Além disso, foi identificado dados discrepantes, informações detectadas como ruídos, além de equalizar os valores, para permitir que o banco de dados tivesse a mesma quantidade de informações em cada atributo, não tendenciando a tomada de decisão do algoritmo. Esse procedimento foi realizado diretamente no software Weka.

O software empregado consiste em uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. O Weka oferece uma gama de ferramentas que permite a rápida tomada de decisão na interpretação de perfis litológicos, por exemplo. Existem algoritmos de classificação, regressão, agrupamento, regras de associação, bem como a visualização, estatística prévia e pré-processamento dos dados antes da execução dos classificadores. Ressalta-se, que não usamos os mesmos dados de treino para teste, foi utilizado o método *K-fold cross validation*, com 10 *folds*.

Por fim, foi analisado o modelo *Random Tree*. A escolha do método ideal desempenha um papel crucial na construção de modelos de classificação robustos e precisos. Por isso, foram avaliadas principais métricas: *TP Rate*, *FP Rate*, *Precision* e *Recall*, além da acurácia de cada modelo.

### 3. RESULTADOS E DISCUSSÃO

O primeiro resultado obtido através do software é um sumário estatístico do banco de dados que pode ser visualizado na tabela 1.

Tabela 1: Sumário estatístico do banco de dados.

Atributos	Mínimo	Máximo	Média	Desvio Padrão
<b>Profundidade</b>	48,15	1499,7	777,482	392,174
<b>Resistividade</b>	1350,445	2346,629	1823,754	327,109
<b>Densidade</b>	368,115	667,921	522,102	65,661
<b>Raio Gama</b>	1,479	129,177	49,999	30,83
<b>Impedância Acústica</b>	2035266,625	5914055	3635572,5	982623,16
<b>Impedância Acústica Real</b>	-617319,75	629867	-48230,982	256061,118
<b>Porosidade</b>	0,204	0,705	0,323	0,06

A análise dos atributos revela informações importantes. Na profundidade, observamos uma variação que vai de 48,15 a 1499,7, com uma média de 777,482 e uma dispersão moderada em relação à média. Já na resistividade, a faixa se estende de 1350,445 a 2346,629, com uma média de 1823,754 e uma dispersão relativamente baixa. Quanto à densidade, os valores variam entre 368,115 e 667,921, com uma média de 522,102 e uma dispersão pequena. O raio gama, por sua vez, apresenta uma amplitude de 1,479 a 129,177, com média de 49,999 e uma dispersão considerável. No que diz respeito à impedância acústica, a variação é notável, abrangendo de 2.035.266,625 a 5.914.054,5, com média de

3.635.572,493 e uma alta dispersão. Já a impedância acústica real apresenta uma amplitude de menos 617.319,75 a 629.867, com média de -48.230,982 e uma dispersão significativa. Por último, a porosidade possui valores que variam de 0,204 a 0,705, com média de 0,323 e uma dispersão pequena em relação à média. Esses dados estatísticos oferecem insights valiosos sobre a amplitude e a variabilidade dos atributos, sendo fundamentais para análises estatísticas e tomadas de decisão relacionadas a esses dados.

Agora aplicando a técnica de classificação e verificando a acurácia do algoritmo de 99,4% e analisando suas matrizes de confusão, tabela 2, verifica se que o modelo possui uma grande precisão, pois os dados são classificados perfeitamente e quase sem falsos positivos ou negativos.

Tabela 2: Matriz de confusão.

a	b	c	d	< Classificado como
2395	1	24	0	a = F
0	2420	0	0	b = S
21	0	2395	4	c = C
0	0	0	2420	d = A

Na tabela acima, temos que “a” é representado por “F” que seria uma rocha argilosa, o folhelho, “b” no qual é uma rocha sedimentar clástica, siltito, “c” equivale também a uma rocha sedimentar, o calcário e “d” é o arenito. Observando a matriz de confusão é possível dizer que o arenito e o siltito possuem grande relevância no modelo, pois não houveram nenhum valor falso positivo ou negativo, entretanto se levarmos em consideração o folhelho e o calcário, existem alguns falsos positivos ou negativos.

Salienta-se ainda, outro fator importante na interpretação do modelo foi a avaliação das métricas, tabela 3, que nos mostra uma visão de acurácia de cada classe do banco de dados.

Tabela 3: Métricas do modelo.

TP rate	FP rate	Precision	Recall	Classe
0,990	0,030	0,991	0,990	F
1,000	0,000	1,000	1,000	S
0,990	0,030	0,990	0,990	C
1,000	0,010	0,998	1,000	A

A tabela 2 mostra que temos uma ótima sensibilidade (*Recall*), isso demonstra que nosso modelo é realmente confiável, outro parâmetro analisado foi a *Precision*, no qual exibiu-se altas precisões, que indica existências de poucos falsos positivos, como já comentado anteriormente.

#### 4. CONCLUSÕES

É possível classificar perfis litológicos com grande acurácia e precisão por meio de inteligência artificial e sua técnica de aprendizado de máquina. A inteligência artificial permitirá automatizar processos na empresa de sementes e aumentar a velocidade de conclusão de tarefas.

## 5. AGRADECIMENTOS

Os autores agradecem ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pela bolsa de estudos e auxílio financeiro que possibilitou a dedicação integral ao projeto e a realização do estudo.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ANTARIKSA, Gian; MUAMMAR, Radhi; LEE, Jihwan. Performance evaluation of machine learning-based classification with rock-physics analysis of geological lithofacies in Tarakan Basin, Indonesia. **Journal of Petroleum Science and Engineering**, v. 208, p. 109250, 2022.

SHI, Mingjiang et al. Logging curve prediction method based on CNN-LSTM-attention. **Earth Science Informatics**, v. 15, n. 4, p. 2119-2131, 2022.

XU, Chicheng et al. When petrophysics meets big data: What can machine do? In: **SPE Middle East Oil and Gas Show and Conference**. OnePetro, 2019.