Submission code: 59RX7RLD9M

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

LightGBM and K-Means perfomance in identifying permo-porosity facies in the pre-salt layer of the Búzios oil field, Santos Basin, Brazil

Maria José Oliveira (UFMG), Caique Pinheiro de Carvalho (UFMG), Henrique Versieux (UFMG), Pedro Baptista (UFMG), Tiago Novo (UFMG), Humberto Reias (HR CONSULTING LTDA.)

LightGBM and K-Means perfomance in identifying permo-porosity facies in the pre-salt layer of the Búzios oil field, Santos Basin, Brazil

Please, do not insert author names in your submission PDF file.

Abstract Summary

Permo-porous facies are very important to characterize reservoirs, it is possible to use different well logs info and split the data according to its petrophysical features such as porosity and permeability. In this work, we use different types of machine learning techniques, supervised and unsupervised, to classify well log data in three different facies, the first facie is represented by a low porosity lithology, the second facie with high porosity and low permeability caused by magnesian-clay, and the third facie with high porosity and high permeability. The algorithms use as input the acoustic impedance and the vp/vs ratio of the medium. The supervised method achieved a better accuracy but the unsupervised was also able to correctly identify the third facie, which is related to the reservoir.

Introduction

The discovery of hydrocarbons in the Santos Basin dates back to the 1970s. Currently, the largest producers' fields are the Tupi and Búzios fields, which host large oil reserves in the pre-Salt carbonate reservoir. Only the Búzios field contributes with 25% of the current Brazilian pre-salt hydrocarbon production, operating since 2018 (ANP, 2024). Santos Basin (Figure 1) was formed during the Mesozoic opening of the South Atlantic Ocean, limited Northeast and Southeast by the Cabo Frio High and the Florianópolis Platform, respectively. The basin is filled with siliciclastics, carbonate rocks and evaporites formed during three main tectonic episodes: rift, post-rift and drift.

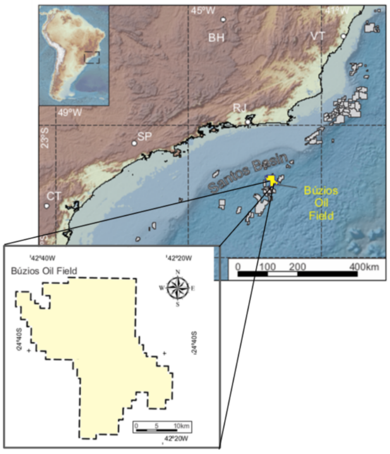

Figure 1: Map of the study area.

The work of De Castro and Lupinacci (2019) evaluated the Buzios field reservoirs within the Itapema and Barra Velha Formations. The authors highlighted the different types of carbonates to magnesian clay-rich reservoirs, whose quality may vary from poor to good porosity and permeability, which affects their quality. Due to the complexity and heterogeneity of the area, robust analysis and coherent interpretation combined with advanced techniques are necessary to understand the reservoirs. Although methods for classifying seismic facies using machine learning techniques have been extensively applied, we present an adaptation code for the supervised and unsupervised algorithms, therefore providing an estimation of geological accurate permo-porous facies.

Method and/or Theory

Methods for classifying seismic facies using machine learning are very resourceful when applied to reservoir characterization. Supervised learning consists of available training data that provides information on the pattern to be discovered. The model is trained by considering labeled data to learn how to predict a specific outcome. In the case of unsupervised learning, the goal is to find patterns or structures in the data for grouping, association, or dimensionality reduction, for example (Mohri, 2018).

Using the data from five wells, we classified three reservoir facies in carbonate strata of the pre-salt that were used as ground truth for applying a supervised and an unsupervised machine learning algorithms. The reservoir facies classification was based on porosity calculations using neutron and density logs and lab permeability measurements. For the first, we applied the formulae proposed by Dewan (1983), based on neutron porosity and density Using a cut of of 8% and 50mD for porosity and permeability, respectively, we classified three facies: i) facies 1, showing lower permo-porosities; ii) facies 2 exhibiting high porosity and low permeability; iii)facies 3, showing higher permo-porosities.

Afterward, we applied the LightGBM and K-means algorithms to predict the previously classified facies using acoustic impedance and the Vp/Vs ratio logs. The LightGBM is a supervised method developed to handle large-scale data. With a loss function that allows the model to take a smaller prediction step and prevent overfitting of the data, while K-Means is an unsupervised method based on clustering unlabeled data that share greater similarity than points from another group (MacQueen et al., 1967). The model was trained using four wells and validated on a single well.

Results

Here we present the 8-BUZ-27D well results. The clusters with the distribution of facies along the well were then generated first through the supervised learning algorithm, as shown in Figure 2 a. Facies 1 is green, facies 2 pink, and facies 3 blue in the image. For the current well, the calculated normalized accuracy is 0.7134. The K-means algorithm requires Vp, Vs, and Rho curves as input data for training. The result of the facies classification and the color pattern for facies distribution is the same followed by the supervised algorithm. Figure 2 b shows the result of the normalized accuracy, which for the K-means algorithm is 0.6158.

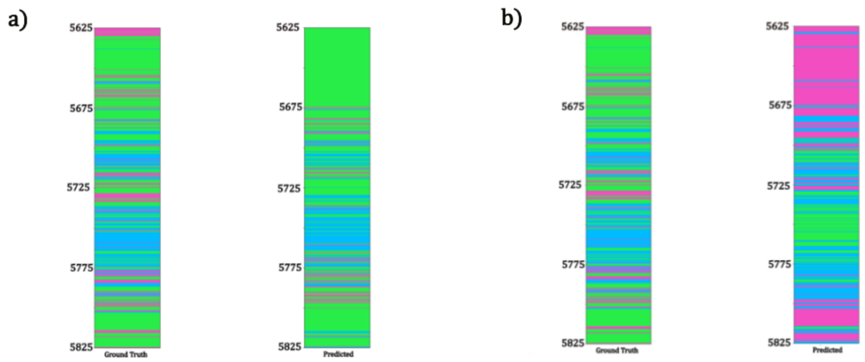

Figure 2: a) LightGBM classification algorithm. b) K-means classification algorithm

According to Figures 2 a and b for both algorithms, the lower portion of the results shows better permo-porosity properties and less presence of facies 2. In Figure 2, it was observed that the supervised algorithm LightGBM finds greater correlation with facies 3. The accuracy of the facies 1 correlation is the highest among them. Therefore, LightGBM accurately locates regions with low porosity and permeability, which could be classified as non-reservoir areas. On the other hand, the unsupervised algorithm (Figure 2 b) exhibits less consistent results. K-means was quite accurate in the porous carbonate facies (facies 3) compared to the ground truth. For facies 2 and 1, there was low correlation, often confusing one for the other, and failing to make a consistent classification.

Conclusions

After analyzing the classification images, some conclusions can be drawn about the effectiveness of the classification algorithms used, thereby providing a better understanding of the distribution of porosity and permeability in the pre-salt region studied. LightGBM is the one that presents the best accuracy when compared to the ground truth. K-means, in turn, loses accuracy but is also capable of identifying the zone with the best permo-porosity characteristics, such as the lower portion corresponding to the Itapema Formation. Also, the implication of facies 2 in the results highlights the occurrence of magnesian clays on the upper part which corresponds to the Barra Velha Formation.

Acknowledgments

We thank Petrobras for founding this research through the ' Inteligência artificial aplicada a exploração de petróleo na camada do pré-sal' project and technical support. We are also grateful to the Agência Ncional do Petroleo, Gás Natural e Biocombustíveis (ANP) for providing the seismic, well and aero-geophysical data.

References

ANP, 2024, Boletim da produção de petróleo e gás natural. número: Boletim ANP, **1**, 1–35.

De Castro, T., and W. Lupinacci, 2019, Evaluation of fine-grains in pre-salt reservoirs: , 1–6.

Dewan, J., 1983, Essentials of modern open-hole log interpretation: Pennwell Books,Tulsa, OK.

MacQueen, J., et al., 1967, Some methods for classification and analysis of multivariate observations: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Oakland, CA, USA, 281–297.

Mohri, M., 2018, Foundations of machine learning.