Rock facies classification enhanced by an automatic bedding discriminator algorithm

Nayguel Costa[1], Pedro Diaz[2]
[1] Federal University of Bahia, Brazil
[2] Pontifical Catholic University of Rio de Janeiro – PUC-Rio

Abbreviated abstract: Interpretation of well log data to predict rock facies is fundamental in oil and gas exploration; however, it is time consuming and a hard to accomplish task. In this study, a novel rock facies classification methodology is proposed. The proposed methodology consists of applying an automatic bedding discriminator algorithm prior to the use of well-known machine learning models. The use of this pre-processing step in the processing pipeline has proved to be effective by different metrics (Accuracy, F-1 Score, and MCC).

Related publications:
– M.K. Dubois *et al*, Computers & Geosciences (33), 599-617 (2007)
– Potratz et al, Anuário do Instituto de Geociências (44), 35024 (2021)

nayguele@gmail.com

3rd Conference on Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021

Problem and Previous Works

- The best way to recognize lithological sequences in the subsurface is the description of well-cores. However, collecting cores is expensive.
- Geophysical profiles provide a close approximation of the rocks but need a specialist to determine which facies best represent the interval.
- Several authors have tested well-known machine learning algorithms (ex: Dubois et al., 2007) or more sophisticated techniques (Potratz et al., 2021), achieving good results.

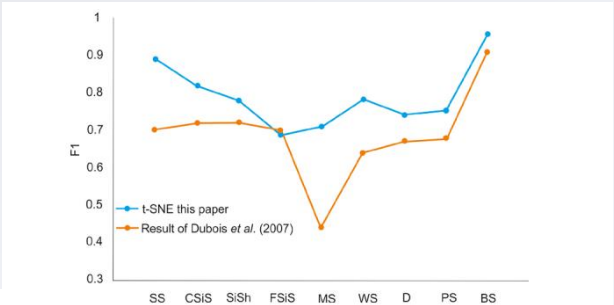
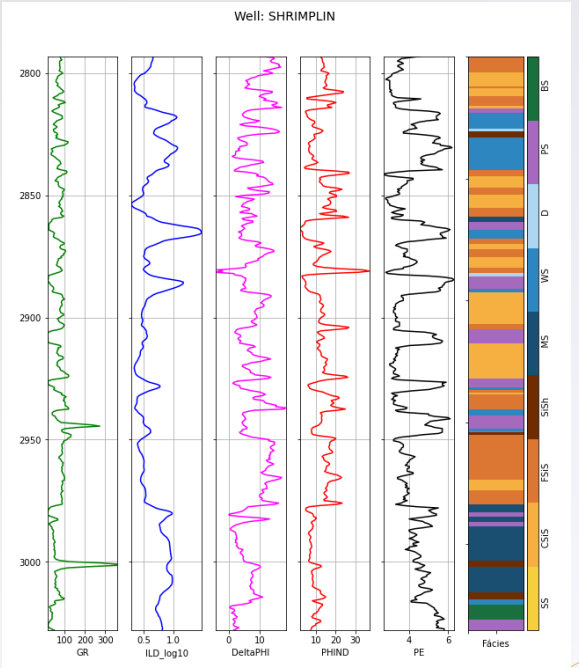


Figure 6 Evaluation metrics for each facies calculated for the classifications with t-SNE+K-NN (in blue) and based on the artificial neural network used by [6] (in orange).



Example of well from Hugoton Field Dataset. Geophysical logs and Facies annotation.

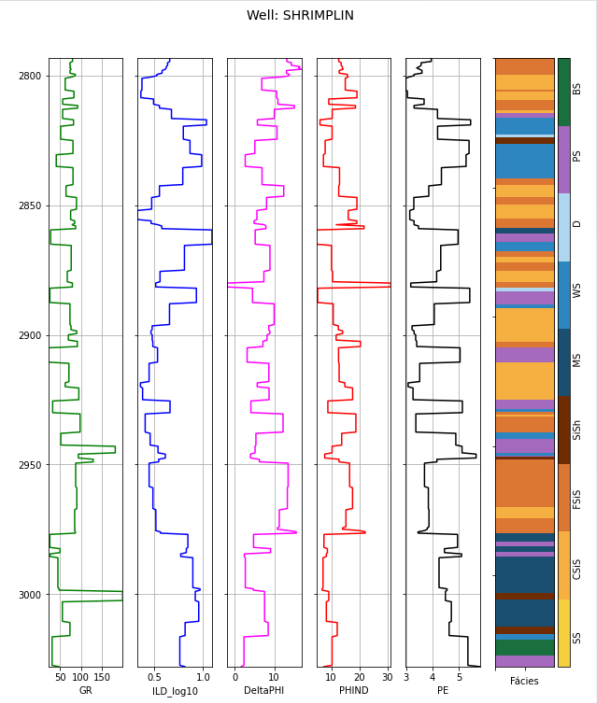
UFBA Universidade Federal da Bahia

PUC RIO

nayguele@gmail.com

3rd Conference on Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021

Technique



Automatic **B**edding **D**iscriminator* (ABD) is applied to profiles prior to the use of a machine learning model.
ABD first applies a centered moving mean to the logs, to deal with noise. Then, a second moving mean with a much larger window is applied, in order to catch the log's general trend. By comparing both filtered logs, it is possible to detect changes in layers based on the intersections. The average value of the log is assigned to each layer.

Once this is done, we can apply any machine learning model.
In this work , Support Vector Machine (SVM) and Random Forest (RF) are employed.

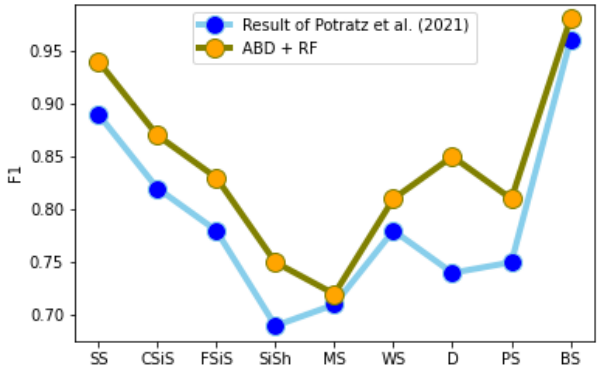
* This segmentation method is based on the work of Reid (Reid et al., 1989)

UFBA Universidade Federal da Bahia
PUC RIO
nayguele@gmail.com

3rd Conference on Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021

Results and Conclusions

Model	Precision	Recall	F1-score	Accuracy	Adjacent Facies Accuracy	MCC
SVM	0.63	0.63	0.61	0.63	0.92	0.56
ABD + SVM (substituted)	0.58	0.62	0.60	0.62	0.90	0.55
ABD + SVM (added)	0.66	0.66	0.64	0.66	0.93	0.59
RF	0.74	0.73	0.73	0.73	0.94	0.68
ABD + RF (substituted)	**0.84**	**0.84**	**0.84**	**0.84**	**0.95**	**0.81**
ABD + RF (added)	0.82	0.82	0.82	0.82	**0.95**	0.78
t-SNE + KNN (paper)	0.79	0.79	0.79	-	-	-



From the point of view of different metrics, it is worth including ABD in the processing pipeline. Both SVM and RF observed improvements with the application of the method.
Particularly, the **ABD + RF** combination produced outstanding results, surpassing the reference results obtained in works that used the same dataset and respected the same protocols as this one.
This preprocessing step (ABD) has a low computational cost, which further encourages its use.

UFBA Universidade Federal da Bahia
PUC RIO
nayguele@gmail.com

3rd Conference on Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021