Lithology identification based on interpretability integration learning

Xiaochun Lin
National Research Center for Geoanalysis
Shitao Yin (840467121@qq.com)
China University of Geosciences (Beijing)

Research Article

Keywords: Logging data, Lithology Identification, Ensemble Learning Stacking, Interpretability,

Posted Date: March 27th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2716684/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Earth Science Informatics on May 29th, 2023. See the published version at https://doi.org/10.1007/s12145-023-01024-5.

Abstract

A lithology intelligent identification interpretability model is proposed based on Ensemble Learning Stacking, Permutation Importance (PI) and Local Interpretable Model-agnostic Explanations (LIME). The method aiming to provide more accurate geological information and more scientific theoretical support for oil and gas resource exploration. Two logging datasets from the public domain were used as experiments, and support vector machine (SVM), random forest (RF) and naive bayes (NB) were used as primary learners, and SVM as secondary learners, to classify lithology through stacking algorithm. Then, the evaluation indexes such as Area Under Curve (AUC), precision, recall and F1-score were used to verify its accuracy, and PI and LIME were used to explain the lithology identification model. The study shows that the results of the stacking algorithm have the best indexes and the highest prediction accuracy. In terms of overall interpretation, PHIND, GR and RT have the most influence on lithology identification of a natural gas protection area in the United States; DEN, CAL and PEF have the most influence on lithology identification in Daqing Oilfield in China. Interpreted from the perspective of a single sample, the LIME algorithm is able to give a quantitative prediction probability and the degree of influence of the characteristic variables.

1. Introduction

Oil and gas are kinds of non-renewable resources. With the exploitation and consumption of human beings, the speed and accuracy of oil and gas resource exploration are becoming higher and higher. In the exploration of oil and gas resources, lithology identification is the premise of accurately determining rock porosity and oil saturation, and also the basis of studying geological reservoir characteristics, calculating reserves and geological modeling. Therefore, rapid and accurate lithology identification using machine learning methods has become a research topic.

Data-driven machine learning method can effectively mine the complex nonlinear relationship between high-dimensional features. Machine learning has developed rapidly in recent years and has been widely used in many fields, including geosciences (Saporetti et al. 2018, Sun et al. 2019, Saporetti et al. 2019, Asante-Okyere et al. 2020). In lithologic identification, there are also many related studies, such as Adaboost (Han et al. 2021), random forest (RF) (Ao et al. 2019), support vector machine (SVM) (Bressan et al. 2020), and artificial neural networks (Asante-Okyere et al. 2020).

Stacking integration is one of ensemble learning algorithms that uses a parallel learning approach and an untyped algorithm (called "primary learner") to obtain the initial prediction values and a meta-learner to further optimize the initial prediction values to obtain the final prediction results. In the literature (Liu et al., 2020), a load prediction method based on a multi-model fusion Stacking ensemble learning approach is proposed, using long short-term memory (LSTM), gradient decision tree, RF, and SVM as primary learners, and then the results of the primary learners are further optimized by a meta-learner. The method makes full use of the advantages of each model and has good prediction results for conventional loads.

At present, the research of intelligent lithologic identification focuses on improving the accuracy of the model,, while the predict results of the models lack sufficient interpretation. More accurate machine learning algorithms are less interpretable (Ibrahim et al., 2019), which limits the progress of the identification model in lithology identification. In order to improve the interpretability of machine learning algorithms and increase the reliability of identification model, some interpretive algorithms have emerged in recent years. For example, SHAP based on coronary heart disease mortality prediction (Wang et al., 2021), applied interpretable machine learning to estimate crop yields (Mateo-Sanchis et al., 2021), and the LIME based on traffic safety interpretability study (Das et al., 2021).

Based on the previous research results, the ensemble learning model represented by random forest, support vector machine and stacking algorithm is widely adopted in lithology identification. In this paper, in order to further verify the generalization ability of the ensemble learning model in lithology recognition and prediction evaluation and improve the interpretability of the model, In the Council Grove gas reserve located in Kansas, USA (Bohling and Dubois, 2003; Dubois et al, 2007), and Daqing Oilfield, in China, two public logging data sets as an example, With SVM, RF, and naive bayes (NB) as primary learners, SVM as a secondary learner, Classification prediction of lithology is made by Stacking. Precision, recall, F1-score, Area Under Curve (AUC) were verified, and the interpretability of the identification model was studied by two explanatory algorithms: PI and LIME.

2. Materials

2.1 Data description and pre-processing

Two datasets from the public domain were used in the study. Datasets 1 (Data 1) is Facies logs from nine wells from the Council Grove gas reserve located in Kansas, USA, and datasets 2 (Data 2) has 12 wells comes from Daqing Oilfield, China, with same logs and lithologies (Cao, 2018).

Data 1 has 9 lithologies include Nonmarine sandstone, Nonmarine coarse siltstone, Nonmarine fine siltstone, Marine siltstone and shale, Mudstone (limestone), Wackestone (limestone), Dolomite, Packstone-grainstone (limestone), Phylloid-algal and bafflestone (limestone) which are studied from core samples in every half foot and matched with logging data in well location. The dataset contains 3165 samples in total. Feature variables include five from wireline log measurements. See Table 1 and Fig. 1.

Table 1
Data 1 description

Feature variables	Interpretation	Lithological categories	Interpretation
GR	Gamma ray log	Litho1	Nonmarine sandstone
RT	Resistivity measurement	Litho2	Nonmarine coarse siltstone
PE	Photoelectric effect log	Litho3	Nonmarine fine siltstone
POR	Porosity index	Litho4	Marine siltstone and shale
PHIND	Average of neutron and density log	Litho5	limestone
		Litho6	Dolomite

Data 2 is the actual logging data of a tight sandstone working area in Daqing Oilfield, which has been repeatedly tested by experts. This area is a key area for tight sandstone oil exploration mainly contains five different lithology categories, i.e., mudstone, siltstone, argillaceous siltstone, silt mudstone, and oil shale. The dataset contains 5978 samples in total. Feature variables include eight conventional logging curves common to each well. See Table 2 and Fig. 2.

Table 2
Data 2 description

Feature variables	Interpretation	Lithological categories	Interpretation
GR	Gamma ray log	Litho1	Mudstone
CAL	Caliper log	Litho2	Siltstone
SP	Spontaneous potential log	Litho3	Argillaceous siltstone
AC	Acoustic log	Litho4	Silt mudstone
DEN	Density log	Litho5	Oil shale
LLD	Deep resistivity log		
LLS	Shallow resistivity log		
PE	Photoelectric effect log		

To ensure the reliability of the data, data preprocessing is performed on the original logging data, including the following: remove outliers, filter, and the normalization method of the maximum and minimum values is used for the logging data. The calculation method is as follows:

$$X_{nor} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

1

where X_{nor} is the normalized logging data, and X_{max} and X_{min} are the maximum value and the minimum value of the original logging data of the single attribute, respectively.

2.2 Correlation analysis of feature variables

There may be complex and unknown relationships between feature variables, so the degree of correlation should be found and quantified. This can help to better prepare the data to meet the expectations of machine learning algorithms, such as linear regression, whose performance decreases as these correlations emerge. In this paper, Spearman coefficient is used to measure the correlation of the selected feature variables, that is, whether the correlation coefficient of the characteristic variables meets $|R| \leq 0.75$ (Chen et al., 2022a). Figure 3. (a) shows that the eight feature variables used by Data 1 all meet the requirement of being independent of each other. As can be seen from Fig. 3. (b), among the feature variables used by Data 2, DEN and AC have high correlation, so they choose to remove AC, and other feature variables meet the requirements of mutual independence.

3. Methods

3.1.1 SVM

SVM constructs the widest classification boundary of the data accurately and solves the non-linear classification problem (Wang et al., 2017). The algorithm calculates a hyperplane that maximizes the distance between different categories of samples (Shankar et al., 2020), which can be applied to linearly separable and linearly nonseparable data. SVM uses the hinge loss function to calculate the empirical risk and adds a regularization term to the solution system to optimize the structural risk. It is a sparse and robust classification method. SVM can perform nonlinear classification through kernel method, which is one of the common kernel learning methods (Hsieh et al., 2009).

3.1.2 RF

The Bagging algorithm is a very widely used ensemble learning algorithm. In the Bagging algorithm, the training set sample of each learner is obtained from the original training set with a randomly selected sample of the original training set, and the size of the new training set is equal to the original training set. Random Forest algorithm is an ensemble learning method based on Bagging algorithm with high classification accuracy and good tolerance for outliers and noise.RF can handle high dimensional data and does not have to do feature selection. (Breiman 2001, Genuer et al. 2017,)

3.1.3 NB

NB is a classification method based on Bayes theorem and independent assumptions of feature conditions, and is one of the most widely used models. The NBC model has stable classification efficiency while requires few estimated parameters and is less sensitive to missing data. The naive Bayesian algorithm assumes that the properties of the data set are independent of each other, so the logic of the algorithm is very simple, and it will not show too much difference for different types of data sets. Theoretically, it has a minimal error rate compared with other classification methods.

3.1.4 Stacking

Ensemble learning is a basic method of data science, which depends on the results of multiple models, that is, the results of multiple weak learners are organized, often better than a single strong model. Stacking is a method in ensemble learning.

The framework of Stacking is shown in Fig. 4, which contains two layers of prediction model, the first layer of prediction model is called primary learner, and second layer prediction model is called meta-learner. The Stacking prediction method first inputs the raw data into each primary learner to obtain the prediction results of the primary learner. Then, the prediction results of the primary learners are used as the input of the meta-learner to get the final prediction results.

The Stacking prediction method combines the advantages of different learners through the integration of multiple primary learners to make the prediction model with strong generalization ability; further, the meta-learner is used to optimize the output results of primary learners to improve the overall prediction accuracy (Xu, 2020).

3.2 Evaluation identification results

To evaluation identification results of the four models, precision, recall, F1-score and AUC were selected as indicators. All of the evaluation metrics are the results obtained on the validation set.

Precision is the probability of actually being positive among all the predicted positive samples. Recall r is the probability of being predicted as a positive sample in the actual positive sample. F1-score is the harmonic average of precision and recall.

Receiver Operating Characteristic curve (ROC) and AUC are measures used to evaluate the performance of classification model. The abscissa of the ROC curve is the false positive rate (The probability of determining a positive case but not a real case), and the ordinate is the true positive rate (The probability that a positive case is also a real case). AUC is the area under the curve. When comparing different classification models, the ROC curve of each model can be drawn. The area under the curve can be used as an indicator of the advantages and disadvantages of the model. The higher the AUC value, the higher the accuracy of the classifier (Swets, 1988).

3.3 Interpretive model for lithology identification

The unexplained lithology recognition model cannot accurately analyze the geological information, because the machine learning model is similar to the black box, which cannot control the operation within the model, and can only be tried between different parameters. Therefore, we need to introduce the interpretability algorithm PI and LIME.

3.3.1 Permutation Importance

PI is an effective method to measure the importance of features unrelated to the model. It proves the importance of a feature value by changing the size of the feature value. The greater the change, the greater the importance (H. M. et al., 2022). The specific steps are as follows: 1) select a feature after the model training; 2) place random numbers for the feature and calculate the new prediction results; 3) the influence of the feature on the model prediction can be obtained by comparing the old and new prediction results.

3.3.2 Local Interpretable Model-agnostic Explanations

LIME is a model-independent local interpretability algorithm that was proposed by Marco (Ribeiro et al., 2016) in 2016. LIME can truly reflect the behavior of the classifier when predicting samples. It targets a single sample and assumes that the local model is a simple linear model to explain the local data points. LIME makes the input values with small perturbations around the local points, observes the prediction behavior of the model, and assigns weights according to the distance between the disturbance points and the original data, so as to obtain an interpretable model and prediction results. The specific formula is as follows:

$$\mathrm{M}_{\text{explanation}}\ (\mathrm{x}) = \underset{\mathrm{g}\in\mathrm{G}}{\operatorname{argmin}} \mathrm{L}\left(\mathrm{f}, \mathrm{g}, \pi_{\mathrm{x}}\right) + \Omega_{(\mathrm{g})}$$

1

For the explanatory model g of example x, the approximation of model g and the original model f is compared by minimizing the loss function. Formula (6): $\Omega_{(\mathrm{g})}$ represents the model complexity of the explanatory model g, G represents all possible explanatory models, π_{x} defines the neighborhood of x, and makes the model interpretable by minimizing L.

4. Results And Discussion

4.1 Results of models

4.1.1 Experimental Results in Data 1

Four models, SVM, RF, NB, and Stacking, were used to identify the lithology in Dataset 1. The results of precision, recall and F1-score indicators are shown in Table 3. The results show that the four models constructed in the training set have good performance in lithology intelligent recognition, and the

identification result of Stacking is the best. The three evaluation indicators of Stacking have the highest mean. The precision of Stacking was 86.40%, recall was 92.23% which is more than 6% higher than the three primary models, and F1-score was 89.12%, showing the effectiveness of ensemble learning. At the same time, it can be seen from the standard deviation analysis that Stacking also has certain advantages in stability and can overcome the imbalance of category number to a certain extent.

Table 3
Lithology identification result in Data 1

Models	Precision	Standard Deviation	Recall	Standard Deviation	F1 value	Standard Deviation
SVM	80.00%	11.25%	70.65%	15.18%	76.13%	15.48%
RF	84.24%	8.83%	79.32%	10.39%	82.90%	9.34%
NB	81.63%	7.72%	86.39%	14.84%	84.38%	7.53%
Stacking	86.40%	4.58%	92.23%	7.38%	89.12%	2.98%

In this study, the ROC curves were analyzed for each of the four models (Fig. 4). The micromean AUC of SVM was 0.82 and macro mean AUC value was 0.72, with Class5 AUC value of 0.85, which was the best discrimination, and Class4 AUC value of 0.59, the worst discrimination. The micro-mean AUC values for both the RF and NB were 0.90, and the macro-mean AUC values were 0.82 and 0.87, respectively. In RF prediction results, Class4 and Class6 identified poorly, and the AUC values did not exceed 0.7. In the identification results of NB, the identification effect of Class2 was relatively poor, but the AUC value reached 0.8, indicating the overall identification is good; the AUC of Stacking was 0.95, the macro average was 0.92, the lowest AUC value was 0.86, and the highest AUC value was 0.96, Class1 and Class5, indicating that the model identification results are very accurate, whether in a single category or on the whole.

4.1.2 Experimental Results in Data 2

Four models, SVM, RF, NB, and Stacking, were used to identify the lithology in Dataset 2. The results of precision, recall and F1-score indicators are shown in Table 4. The lithology recognition result of SVM is poor, precision was 50.36%, recall was 68.28% and F1-score was 62.06%; The lithologic identification results of RF and NB are not particularly excellent, and Stacking has the best lithologic identification results. In Stacking, precision was 84.48%, recall was 89.43%, F1-score was 86.09%, more than 7% higher than the other three models showing the effectiveness of ensemble learning in complex situations. Meanwhile, it can be seen from the standard deviation analysis that Stacking also has certain advantages in stability.

Table 4
Lithology identification result in Data 2

Models	Accuracy	Standard Deviation	Recall	Standard Deviation	F1 value	Standard Deviation
SVM	50.36%	21.67%	68.28%	19.35%	62.06%	15.98%
RF	70.34%	9.41%	73.48%	13.84%	71.90%	9.56%
NB	74.49%	8.46%	82.73%	15.48%	78.74	8.67%
Stacking	84.48%	5.93%	89.43%	6.72%	86.09%	3.85%

In this study, ROC curves were analyzed for the four models (Fig. 5). The micromean AUC value of SVM was 0.77, and the best identified macro-average AUC value was 0.53. The AUC value of Class5 was 0.35, indicating the worst discrimination effect. RF and NB have the micromean AUC values of the 0.86 and the macro-mean AUC values of 0.70 and 0.82, respectively. In RF prediction results, Class2 and Class5 were less well identified, and the AUC values did not exceed 0.6. In the identification results of NB, the recognition effect of Class2 is relatively poor, and the AUC value does not reach 0.7, indicating that the overall identification is poor; the AUC of Stacking is 0.91 and the macro average is 0.86, and the overall performance is better than the other three models. The AUC value of Class2 was the lowest, 0.72, a large improvement over other models; Class5 had the highest AUC value of 0.96, the highest identification accuracy in all categories of all models.

4.2 Interpretability of the model

4.2.1. The importance of factors in Data 1

Different evaluation factors affect the accuracy of lithology identification differently, so this study determines the influence size of the evaluation factors from a global perspective. That is, the importance of each evaluation factor is calculated, and the higher the value of importance, the greater the effect of this factor on the evaluation results. The importance order from large to small is Average of neutron and density log, Gammara ylog, Resistivity measurement, Porosity index and Photoelectric effect log (Fig. 6).

4.2.2. Lithology recognition model local interpretation in Data 1

The LIME algorithm selects samples for categories 5 and 3 (Fig. 7 and Fig. 8). The results show the contribution of each feature to the prediction result of this sample. The positive number in the figure represents that the feature has a positive effect on the result, and the negative number is the opposite. From Fig. 7, the model has an 82% probability for sample Data 1–27 being identified as a Class5. The reason for the model classifying this sample was based on such factors as PHIND, RE and PRO, with PHIND having the highest contribution of 0.494, consistent with the results explained by the global model. In addition, the contribution of RT and GR is negative, which becomes the interference that predicts the sample to be Class5.

The model with 66% probability for sample 43 being identified as Class3. The rationale for the model to classify this sample is based on POP, GR, PE, PHIND, and RT, indicating that all features are supporting the result of this prediction, but the effect size is different.

4.2.3 The importance of factors in Data 2

The importance order from large to small is Density log Caliper log Photoelectric effect log Gamma ray log Spontaneous potential log Shallow resistivity log Deep resistivity log(Fig. 9).

4.2.4. Lithology recognition model local interpretation in Data 2

LIME identifies categories 3 and 1 of category 1 (Fig. 10 and Fig. 11) and interpreted them. The results show that the model has a 64% probability for sample Data 2-111 being identified as Class3. The rationale for the model to classify this sample was based on such factors as DEN, CAL and GR, of which DEN had the highest contribution of 0.195. The results are consistent with the global interpretation. In addition, SP, LLD and LLS are the interference to predict Class3, but less interference. The model has an 89% probability for sample Data 2-243 to be identified as a Class1. The reason for the classification of this sample is based on such factors as PEF, CAL, DEN, SP, LLS, and LLD. In addition, GR becomes the interference that predicts the sample to be Class1.

4.3 Discussion

Previous machine learning-based lithology identification studies generally reflect the importance of features, but it is related to the machine learning algorithm itself, so the method is not universal. However, according to the different influence degree of different characteristics on the lithology identification

results, this study adopts the PI algorithm unrelated to the model algorithm itself from a global perspective, which can more objectively explain the importance of identifying different input features of the model.

In addition, in view of the problem of previous lithology recognition models, this study introduced LIME algorithm to locally explain the model and reveal the specificity of different samples for different categories. The results show that different input characteristics will have different effects on the classification of each sample, including support and interference, and the size of the effect is also different.

Based on the results of the two interpretation algorithms, we know that in Data 1, PHIND, GR and RT are the most important feature variables of the model to identify lithology, and PHIND usually has a positive effect on the identification results. In Data 2, DEN is the most important variable of the model to identify lithology and has a high positive effect; SP, LLD and LLS contribute less, but usually have a negative impact on the identification results. The results of this interpretation have some significance in reality, indicating that the geological background of the two regions is obviously different, and there are obvious differences in the lithology of the rocks storing natural gas and oil, which are the different influencing factors leading to lithology identification, indicating that we can trust our prediction model to a certain extent.

5. Conclusion

Taking a natural gas protection area in the United States and Daqing Oilfield, China, GR, R, PE, POR, GR, PHIND, CAL, SP, AC, DEN, LLD, LLS and PE are selected as the characteristic variables for identification, and the ensemble learning model is combined with PI algorithm and LIME algorithm to identify lithology and obtain high identification accuracy. The main conclusions are:

1. The Ensemble Learning Stacking algorithm is most suitable for lithology identification of the two data sets, with Data 1 identification accuracy 86.40%, recall 92.23%, F1 score 89.12%, Data 2 identification accuracy 84.48%, recall 89.43%, F1 score 86.09%, and each index are better than the other three machine learning models. According to the standard deviation analysis results, Stacking has certain advantages for lithology identification in stability, and can effectively overcome the imbalance of category number.
2. From the perspective of global interpretation, the three characteristic variables PHIND, GR and RT have the greatest influence on the lithology identification of a natural gas protection area in the United States; DEN, CAL and PEF have the greatest influence on the lithology identification of Daqing Oilfield in China. From the perspective of single sample interpretation, the LIME algorithm is able to give a quantitative prediction probability and the degree of influence of characteristic variables.

Declarations

Author’s Contribution

All authors contributed to the study conception and design. Data collection were collected by Xiaochun Lin. The experimental models were constructed by Xiaochun Lin and Shitao Yin. The development and the testing of the presented methods were performed by Xiaochun Lin and Shitao Yin. The manuscript was written by Xiaochun Lin. All authors attend to comment the manuscript, and all authors read and approved the final manuscript.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Ao Y, Li H, Zhu L, Ali S, Yang Z (2019) Identifying channel sand-body from multiple seismic attributes with an improved random forest algorithm. *JOURNAL OF PETROLEUM SCIENCE AND ENGINEERING*, 173, 781–792.Ao, Y., H. Li, L. Zhu, S. Ali & Z. Yang (2019) Identifying channel sand-body from multiple seismic attributes with an improved random forest algorithm. JOURNAL OF PETROLEUM SCIENCE AND ENGINEERING, 173, 781–792
2. Asante-Okyere S, Shen C, Ziggah YY, Rulegeya MM, Zhu X (2020) A Novel Hybrid Technique of Integrating Gradient-Boosted Machine and Clustering Algorithms for Lithology Classification, vol 29. NATURAL RESOURCES RESEARCH, pp 2257–2273
3. Breiman L (2001) Random forests. Mach Learn 45:5–32
4. Bressan TS, de Souza MK, Girelli TJ & F. Chemale Junior (2020) Evaluation of machine learning methods for lithology classification using geophysical data.COMPUTERS & GEOSCIENCES,139
5. Chen Z, Chang R, Guo H, Pei X, Zhao W, Yu Z, Zou L (2022a) Prediction of Potential Geothermal Disaster Areas along the Yunnan–Tibet Railway Project. Remote Sens 14:3036
6. Das S, Datta S, Zubaidi HA, Obaid IA (2021) Applying interpretable machine learning to classify tree and utility pole related crash injury types. IATSS Res 45:310–316
7. Freidman JH (2008) Greedy Function Approximation: A Gradient Boosting Machine. Institue Math Stat 29:1189–1232
8. Genuer R, Poggi J-M, Tuleau-Malot C (2017) Random Forests for Big Data, vol 9. BIG DATA RESEARCH, pp 28–46. & N. Villa-Vialaneix
9. Han R, Wang Z, Wang W, Xu F, Qi X, Cui Y (2021) Lithology identification of igneous rocks based on XGboost and conventional logging curves, a case study of the eastern depression of Liaohe Basin,

vol 195. JOURNAL OF APPLIED GEOPHYSICS, p 104480

10. He M, Wang Q (2022) Excavation compensation method and key technology for surrounding rock control. Eng Geol 307:106784
11. Hsieh WW (2009) Machine Learning Methods in the Environmental Sciences: Contents. Chapter 7, pp.157–169
12. Ibrahim M, Modarres C, Louie M, Paisley J (2019) Global explanations of neural network: Mapping the landscape of predictions, in: AIES 2019 - Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society
13. Liu Y, Yu Z, Chen C, Han Y, Yu B (2020) Prediction of protein crotonylation sites through LightGBM classifier based on SMOTE and elastic net. Anal Biochem 609:113903
14. Mohammadi NM, Hezarkhani A (2018) Application of support vector machine for the separation of mineralised zones in the Takht-e-Gonbad porphyry deposit, SE Iran, vol 143. JOURNAL OF AFRICAN EARTH SCIENCES, pp 301–308
15. Mateo-Sanchis A, Piles M, Amorós-López J, Muñoz-Marí J, Adsuara JE, Moreno-Martínez Á, Camps-Valls G (2021) Learning main drivers of crop progress and failure in Europe with interpretable machine learning. Int J Appl Earth Obs Geoinf 104:102574
16. Raeesi M, Moradzadeh A, Ardejani FD, Rahimi M (2012) Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks. JOURNAL OF PETROLEUM SCIENCE AND ENGINEERING, pp 82–83
17. Ribeiro M, Singh S, Guestrin C (2016) "Why Should I Trust You?": Explaining the Predictions of Any Classifier, in: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 97–101
18. Saporetti CM, da Fonseca LG, Pereira E, de Oliveira LC (2018) Machine learning approaches for petrographic classification of carbonate-siliciclastic rocks using well logs and textural information. J Appl Geophys v 155:217–225
19. Saporetti CM, da Fonseca LG, Pereira E (2019) A Lithology Identification Approach Based on Machine Learning With Evolutionary Parameter Tuning. IEEE Geosci Remote Sens Lett 16:1819–1823
20. Sun F, Yao Y, Chen M, Li X, Zhao L, Meng Y, Sun Z, Zhang T, Feng D (2017) Performance analysis of superheated steam injection for heavy oil recovery and modeling of wellbore heat efficiency. ENERGY 125:795–804
21. Sun J, Li Q, Chen M, Ren L, Huang G, Li C, Zhang Z (2019) Optimization of models for a rapid identification of lithology while drilling-A win-win strategy based on machine learning, vol 176. JOURNAL OF PETROLEUM SCIENCE AND ENGINEERING, pp 321–341
22. Shankar K, Lakshmanaprabu SK, Gupta D, Maseleno A, de Albuquerque VHC (2020) Optimal feature-based multi-kernel SVM approach for thyroid disease classification.J. Supercomput.76

23. Sebtosheikh MA, Salehi A (2015) Lithology prediction by support vector classifiers using inverted seismic attributes data and petrophysical logs as a new approach and investigation of training data set size effect on its performance in a heterogeneous carbonate reservoir, vol 134. JOURNAL OF PETROLEUM SCIENCE AND ENGINEERING, pp 143–149
24. Swets JA (1988) Measuring the Accuracy of Diagnostic Systems. Science (80-.). 240, 1285–1293
25. Shepherd C, Clegg C, Stride C (2009) Opening the Black Box: A Multi-Method Analysis of An Enterprise Resource Planning Implementation. J Inf Technol 24:81–102
26. Thi Ngo PT, Panahi M, Khosravi K, Ghorbanzadeh O, Kariminejad N, Cerda A, Lee S (2021) Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geosci Front 12:505–519
27. Wang H, Xiong J, Yao Z, Lin M, Ren J (2017) Research survey on support vector machine, in: International Conference on Mobile Multimedia Communications (MobiMedia)
28. Wang Y, Fang Z, Hong H (2019) Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. Sci Total Environ 666:975–993
29. Wang K, Tian J, Zheng C, Yang H, Ren J, Liu Y, Han Q, Zhang Y (2021) Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med 137:104813
30. Zhanghua XU, Xuying HUANG, Lu LIN et al BP neural networks and random forest models to detect damage by Dendrolimus punctatus Walker[J].Journal of Forestry Research,2020, 31(1):107–121
31. Yi-hua Z, Rong LI (2009) Application of Principal Component Analysis and Least Square Support Vector Machine to Lithology Identification. Well Logging Technology 33:425–429
32. Zhang X, Ding S, Xue Y (2017) An improved multiple birth support vector machine for pattern classification. NEUROCOMPUTING 225:119–128

Figures

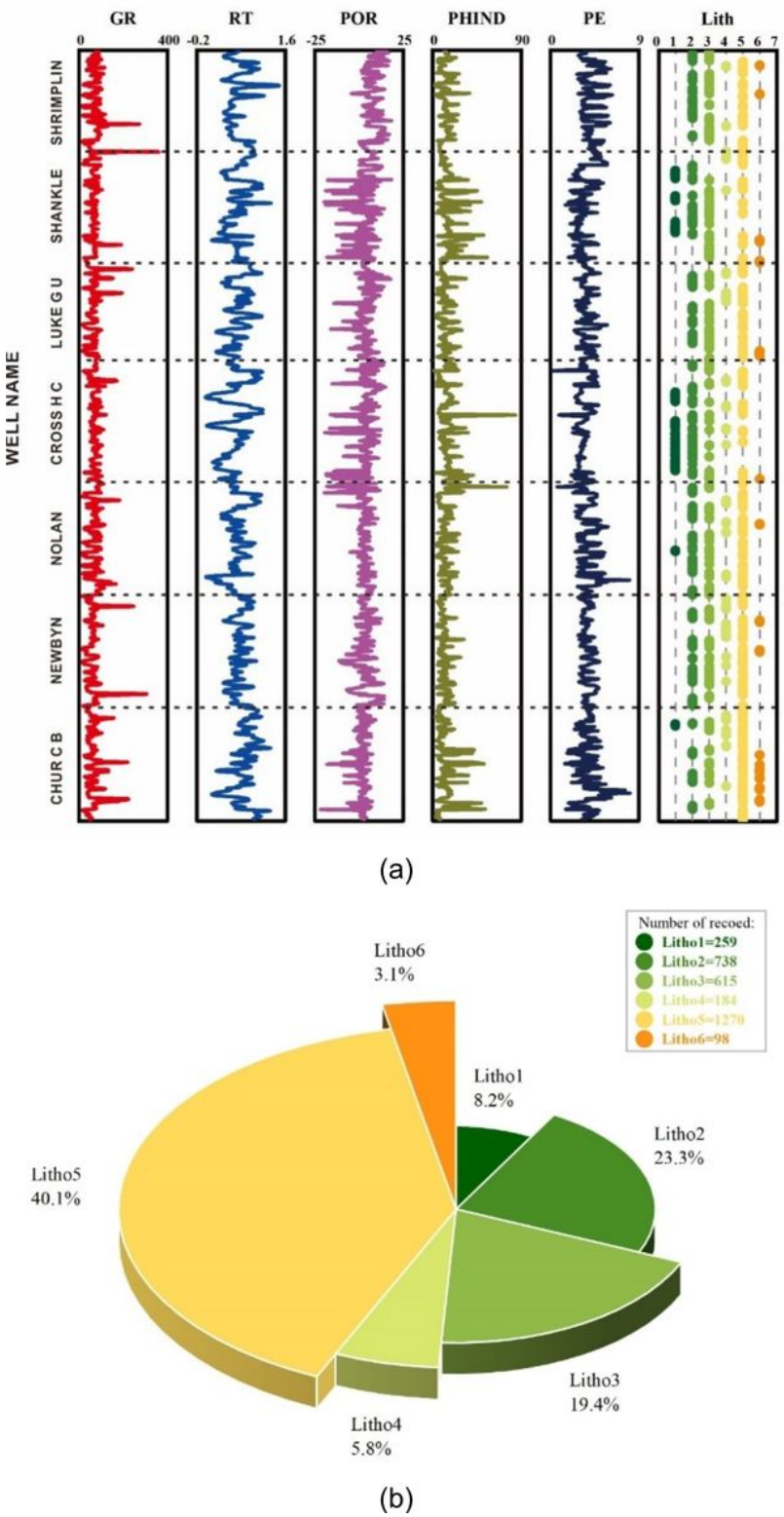

Figure 1

Data 1 (a) Well log data (b) Distribution of lithologies

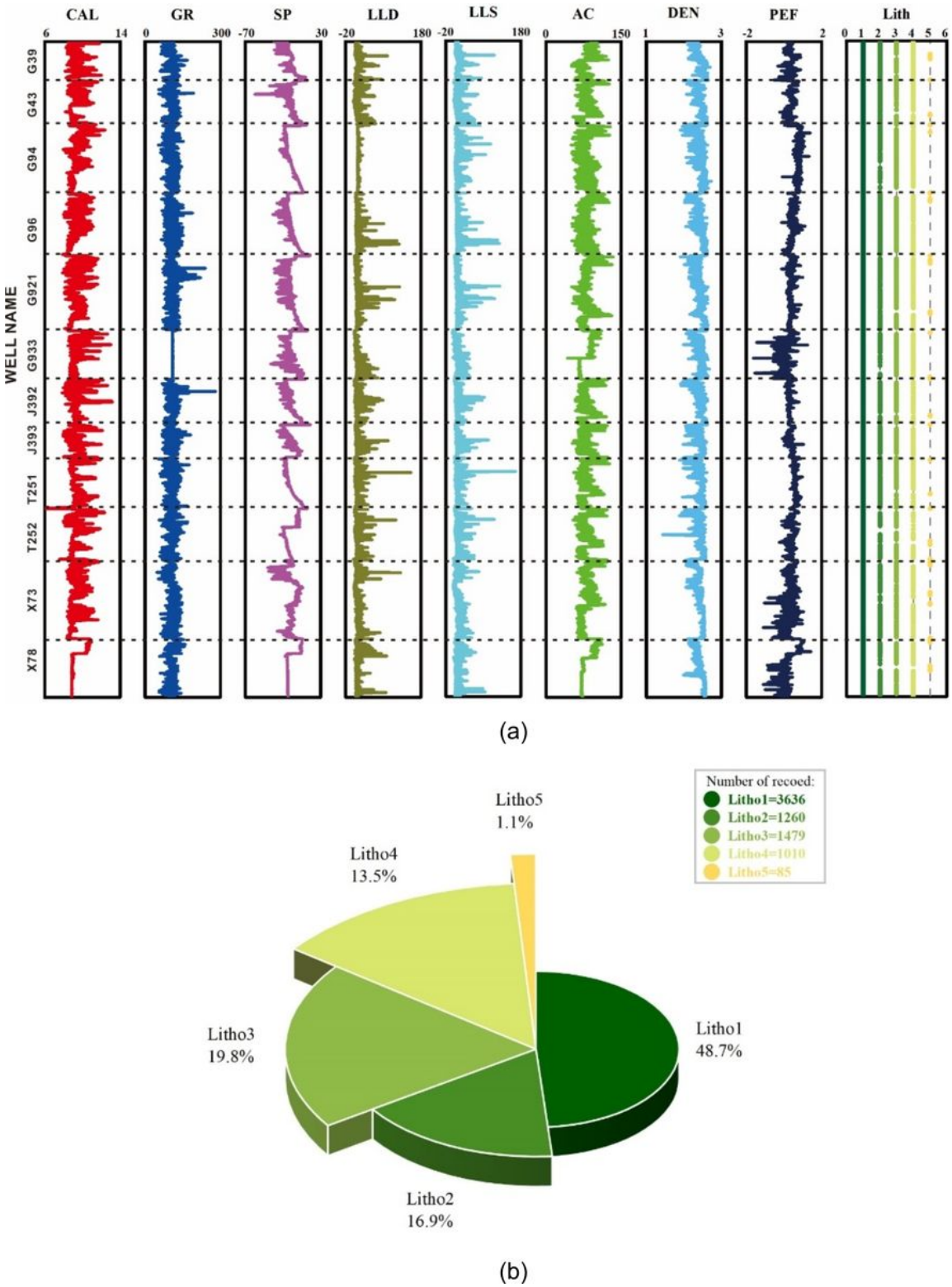

Figure 2

Data 2 (a) Well log data (b) Distribution of lithologies

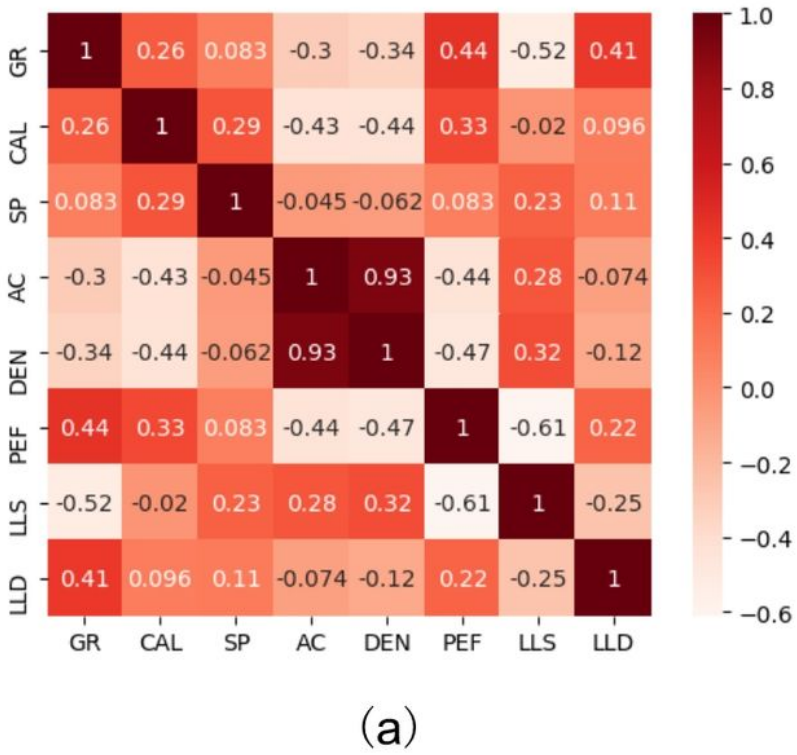

Figure 3

Characteristic variables correlation heat map of Data 1 (a) and Data 2 (b)

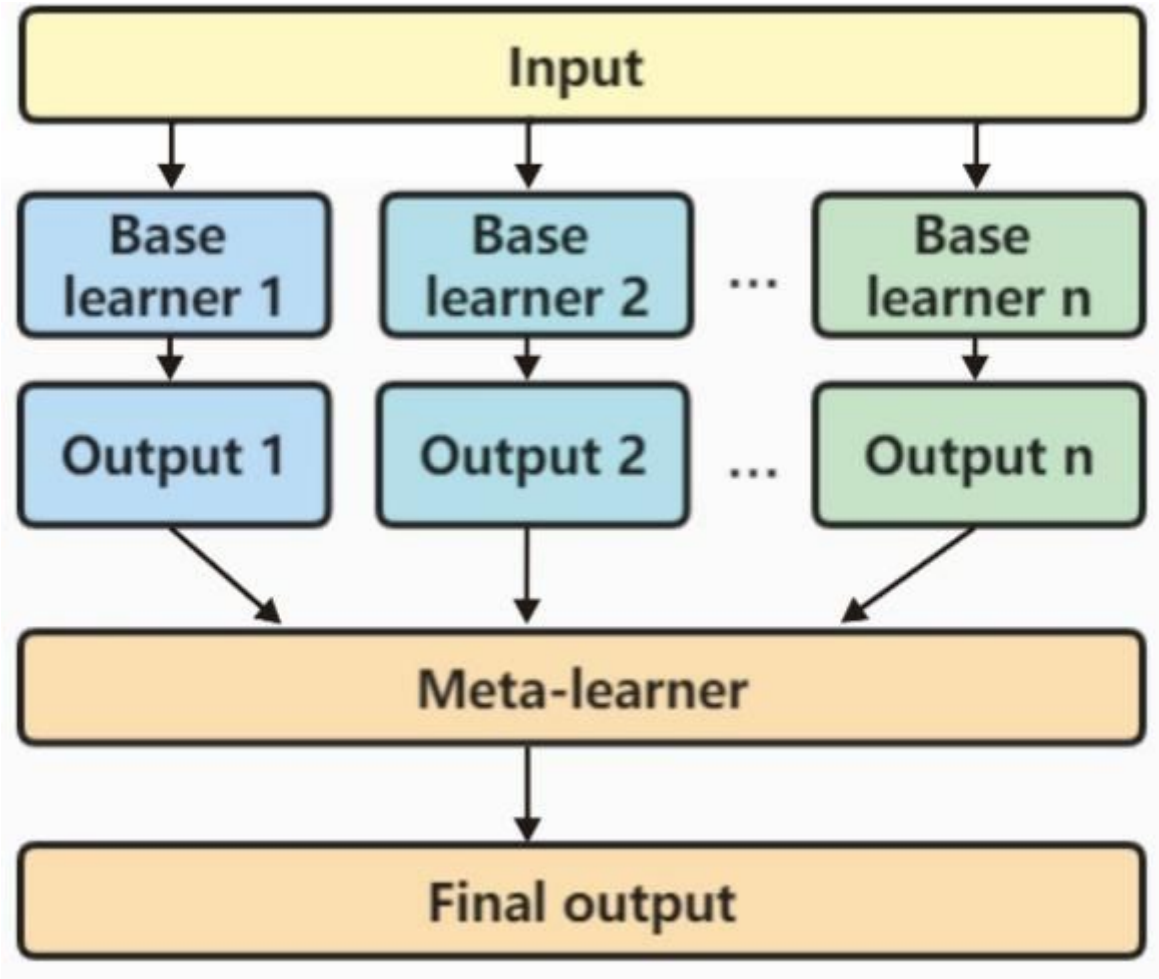

Figure 4

The framework for Stacking

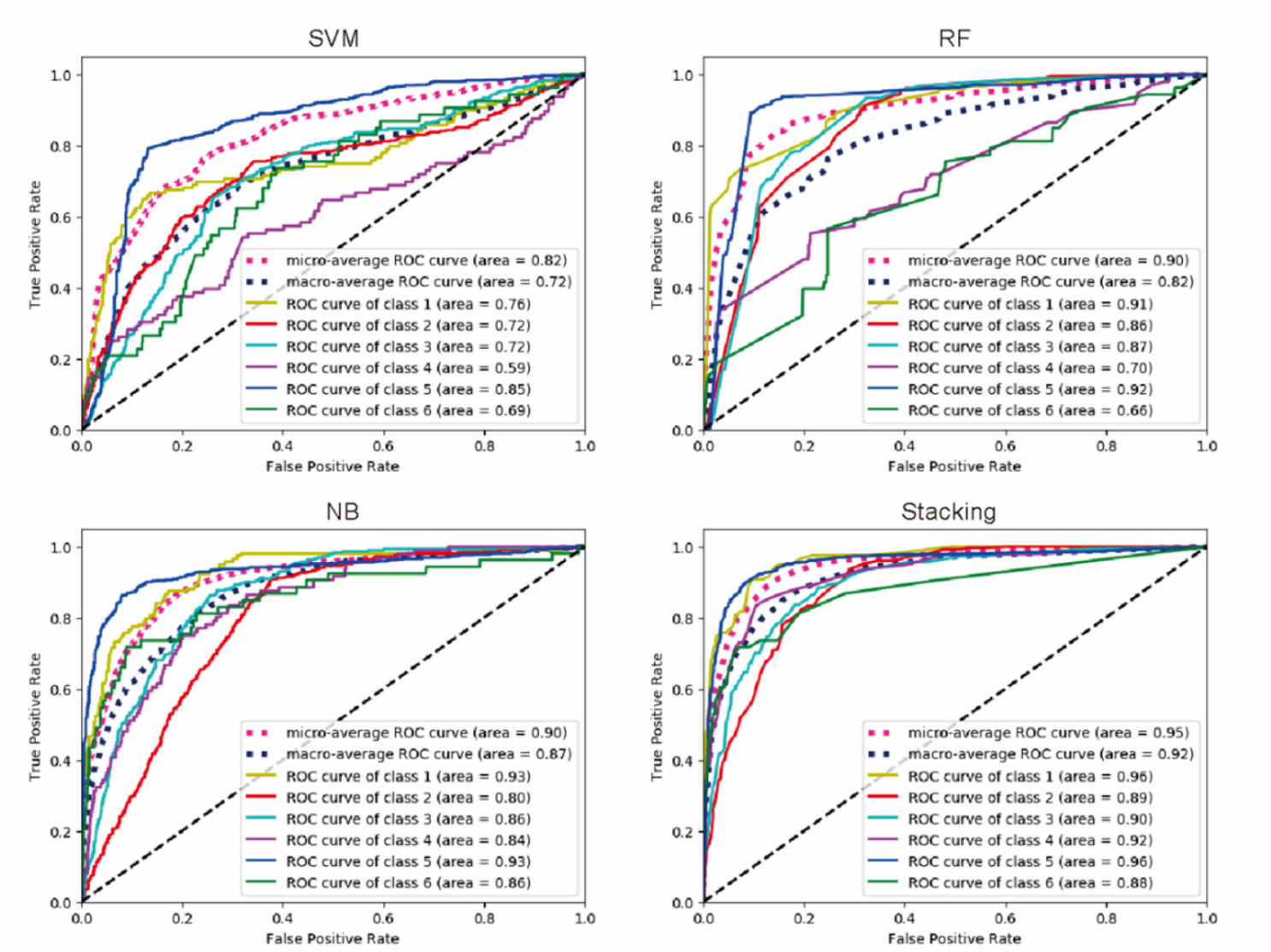

Figure 5

Fig. 4 ROC curve analysis plots of the four models in Data 1

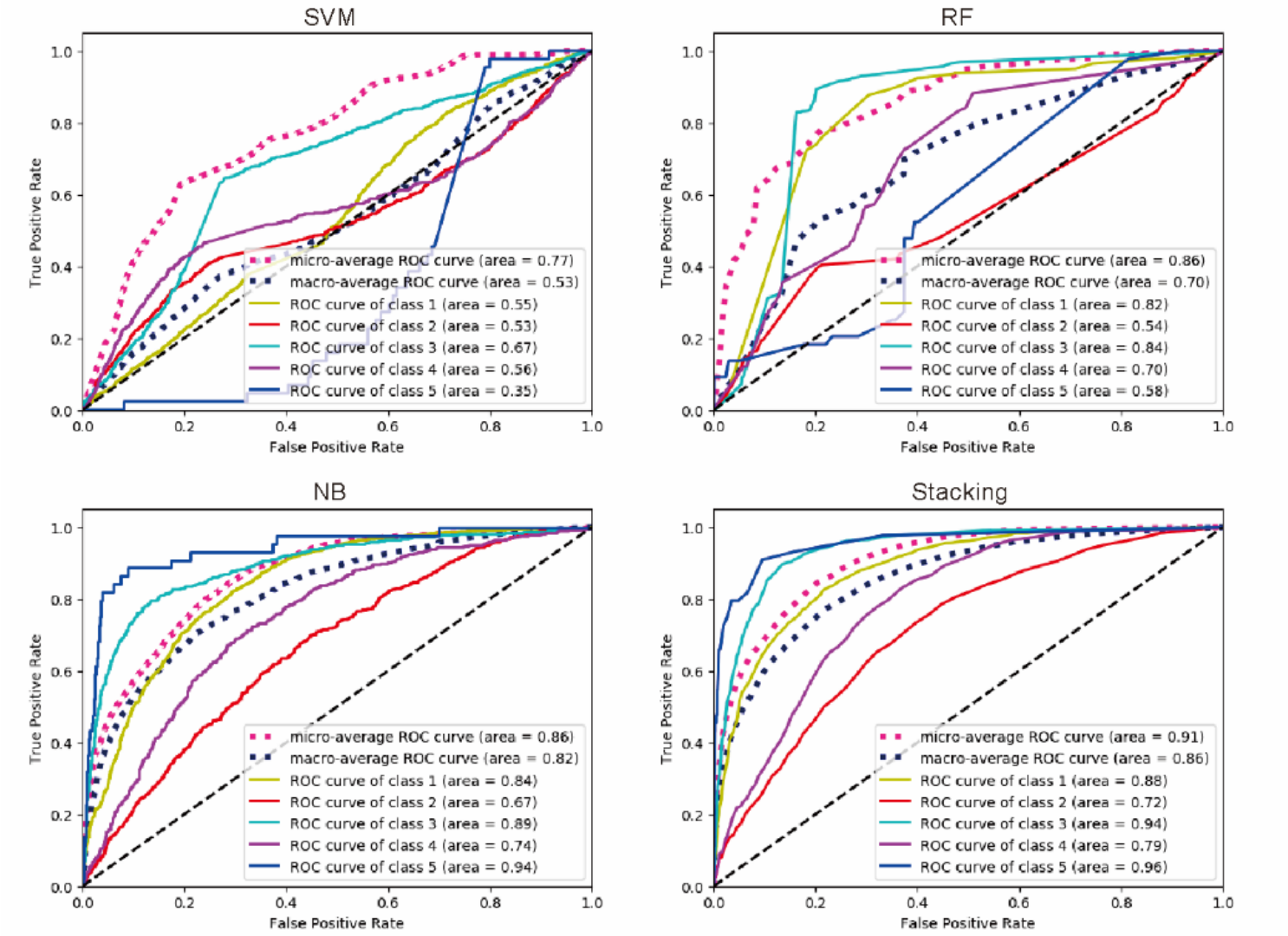

Figure 6

Fig. 5. ROC curve analysis plots of the four models in Data 2

Weight	Feature
0.1005 ± 0.0097	PHIND
0.0951 ± 0.0135	GR
0.0759 ± 0.0135	RT
0.0258 ± 0.0065	POR
0.0254 ± 0.0078	PE

Figure 7

Fig. 6. Importance of the feature variable in Data 1

y=1 (probability 0.000)		y=2 (probability 0.090)		y=3 (probability 0.020)		y=4 (probability 0.070)		y=5 (probability 0.820)		y=6 (probability 0.000)	
Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature
0.013	GR	0.094	GR	0.055	GR	0.057	GR	0.494	PHIND	-0.001	POR
0.008	PE	0.053	RT	0.028	RT	0.02	POR	0.158	PE	-0.002	RT
0.002	RT	0.009	POR	-0.062	PE	0.011	PHIND	0.044	POR	-0.003	PHIND
-0.006	POR	-0.032	PE	-0.065	POR	-0.007	RT	-0.073	RT	-0.005	PE
-0.105	PHIND	-0.263	PHIND	-0.135	PHIND	-0.067	PE	-0.203	GR	-0.016	GR

Figure 8

Fig. 7. LIME interpretation of sample 27 in Data 1

y=1 (probability 0.010)		y=2 (probability 0.210)		y=3 (probability 0.660)		y=4 (probability 0.020)		y=5 (probability 0.040)		y=6 (probability 0.060)	
Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature
0.022	PHIND	0.087	PHIND	0.321	POR	0.018	GR	-0.024	PE	0.037	GR
0.015	RT	0.061	RT	0.058	GR	0.003	PE	-0.071	RT	0.006	PHIND
-0.005	PE	0.001	GR	0.05	PE	-0.012	RT	-0.071	POR	-0.001	POR
-0.038	GR	-0.021	PE	0.018	PHIND	-0.013	PHIND	-0.075	GR	-0.003	PE

Figure 9

Fig. 8. LIME interpretation of sample 43 in Data 1

Weight	Feature
0.0635 ± 0.0082	DEN
0.0178 ± 0.0012	CAL
0.0063 ± 0.0037	PEF
0.0015 ± 0.0034	GR
-0.0005 ± 0.0026	SP
-0.0023 ± 0.0025	LLS
-0.0029 ± 0.0017	LLD

Figure 10

Fig. 9. Importance of the feature variable in Data 2

y=1 (probability 0.090)		y=2 (probability 0.210)		y=3 (probability 0.640)		y=4 (probability 0.060)		y=5 (probability 0.000)	
Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature
0.008	LLS	0.04	LLS	0.195	DEN	0.017	LLS	0.004	DEN
0.005	SP	0.037	LLD	0.098	CAL	0.012	CAL	0	SP
0.003	LLD	0.032	SP	0.05	GR	0.001	PEF	0	GR
-0.016	PEF	0.014	PEF	0.008	PEF	-0.006	LLD	-0.001	LLD
-0.025	GR	-0.004	CAL	-0.028	SP	-0.01	SP	-0.003	CAL
-0.103	CAL	-0.009	GR	-0.032	LLD	-0.016	GR	-0.005	LLS
-0.149	DEN	-0.031	DEN	-0.06	LLS	-0.019	DEN	-0.008	PEF

Figure 11

Fig. 10. LIME interpretation of sample 111 in Data 2

y=1 (probability 0.890)		y=2 (probability 0.100)		y=3 (probability 0.000)		y=4 (probability 0.000)		y=5 (probability 0.010)	
Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature	Contribution	Feature
0.157	PEF	0.019	GR	0.006	GR	0.005	GR	0.009	GR
0.105	CAL	0.015	DEN	-0.002	LLD	-0.005	LLD	0.002	PEF
0.046	DEN	0.004	LLD	-0.003	LLS	-0.006	LLS	0	DEN
0.036	SP	-0.014	SP	-0.01	SP	-0.011	DEN	-0.001	SP
0.035	LLS	-0.022	CAL	-0.031	PEF	-0.011	SP	-0.001	LLS
0.005	LLD	-0.025	LLS	-0.045	CAL	-0.034	CAL	-0.002	LLD
-0.04	GR	-0.078	PEF	-0.05	DEN	-0.05	PEF	-0.004	CAL

Figure 12

Fig. 11. LIME interpretation of sample 243 in Data 2