



Studi Prediksi Porositas Dengan Menggunakan Metode Deterministik dan Machine Learning Pada Lapangan “X”

Babas Samudera Hafwandi ^{1*)}

¹⁾ Teknik Perminyakan, Universitas Jember

* corresponding email: babashafwandi@unej.ac.id

ABSTRACT

Porosity is one of the most critical parameters in reservoir characterization, as it directly influences hydrocarbon storage capacity. Accurate porosity prediction becomes even more essential in fields with limited core data, such as Field “X”, located in the South Sumatra Basin. This study compares two different porosity prediction approaches: a deterministic method based on well log interpretation using NPHI and RHOB logs, and various Machine Learning (ML) algorithms, including Random Forest (RF), K-Nearest Neighbor (KNN), Gradient Boosting (GBR), AdaBoost (ADA), Support Vector Machine (SVM), and Decision Tree (DT). Data preprocessing involved feature selection using Pearson, Spearman, and Kendall correlation coefficients to identify the most influential log parameters. The dataset was then divided into training (70%) and testing (30%) subsets. Model performance was evaluated using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The deterministic method yielded an MAE of 0.0658 and RMSE of 0.0906, while the best ML model, Random Forest, achieved an MAE of 0.0329 and RMSE of 0.0434 on the testing dataset. In conclusion, Machine Learning, especially the Random Forest model, proves to be a more reliable and accurate tool for porosity prediction in geologically complex fields, offering significant potential for enhancing reservoir modeling and field development planning.

Keywords: Porosity, Machine Learning, Deterministic Method

I. PENDAHULUAN

Porositas adalah parameter penting dalam proses karakterisasi *reservoir* karena menunjukkan seberapa besar kemampuan batuan untuk menyimpan fluida hidrokarbon. Akurasi dalam estimasi nilai porositas sangat berperan dalam menentukan cadangan *reservoir* serta mendukung pengambilan keputusan dalam pengembangan lapangan [1], [2]. Oleh sebab itu, berbagai pendekatan untuk memprediksi porositas secara akurat terus menjadi perhatian utama dalam penelitian petrofisika dan karakterisasi *reservoir*.

Metode deterministik dengan menggunakan pendekatan empiris berbasis *log* sumur masih banyak digunakan karena kemudahan interpretasi yang dilakukan serta kemudahan perhitungan yang dilakukan [3], [4]. Meskipun demikian pendekatan ini memiliki keterbatasan dalam menghadapi heterogenitas *reservoir* yang kompleks [5].

Perkembangan teknologi komputasi yang semakin pesat dan meningkatnya ketersediaan data, metode pembelajaran mesin (*Machine Learning*) mulai banyak diadopsi dalam studi petrofisika khususnya dalam memprediksi porositas dalam mengevaluasi karakteristik *reservoir* [6], [7], [8]. Algoritma regresi yang umum digunakan seperti *Random Forest* dan *K-Nearest Neighbor* dapat mengatasi kompleksitas data *reservoir* dan dapat meningkatkan akurasi prediksi nilai porositas dalam interval kedalaman yang tidak memiliki data *core* [5], [8], [9].

Lapangan “X” merupakan lapangan yang memiliki karakteristik geologi yang kompleks dengan ketersediaan data *log* sumur yang cukup lengkap, namun dengan data *core* yang terbatas. Kondisi tersebut menjadi tantangan tersendiri dalam melakukan permodelan petrofisika, khususnya dalam melakukan prediksi nilai porositas. Studi ini dilakukan untuk memprediksi nilai porositas dengan menggunakan metode deterministik dengan memanfaatkan data *log* sumur khususnya data *log* NPHI dan RHOB dan menggunakan metode *Machine Learning*. Hasil prediksi kedua metode tersebut kemudian divalidasi dengan data *core* yang tersedia, dengan menggunakan matriks evaluasi MAE (*Mean Absolute Error*) dan RMSE (*Root Mean Square Error*), dari hasil evaluasi tersebut dapat menunjukkan metode terbaik untuk memprediksi nilai porositas pada Lapangan “X” dan dapat berkontribusi dalam pemodelan *reservoir* yang lebih presisi.

II. METODOLOGI PENELITIAN

Adapun *flowchart* dari studi prediksi porositas dapat dilihat pada Gambar 1. Pertama, melakukan input data berupa data *log* sumur dan data *core* dan melakukan *feature selection* yaitu mencari fitur data dan memilih data yang berpengaruh dalam prediksi porositas dengan menggunakan 3 korelasi yaitu *spearman*, *pearson* dan *kendall*, digunakan 3 korelasi untuk menguji konsistensi fitur data yang berpengaruh dalam prediksi porositas, setelah dilakukan *feature selection*,



kemudian dilakukan pemisahan antara data *training* dan data *testing*, dalam studi ini 70% dari total keseluruhan data digunakan sebagai data *training* dan 30 % sebagai data *testing*. Dalam studi penentuan porositas yang dilakukan, digunakan 6 model *machine learning* yang bisa terlihat pada Tabel 1 dibawah dan selanjutnya dilakukan proses *testing* dan *training* untuk menentukan model *machine learning* manakah yang terbaik dalam memprediksi nilai porositas pada Lapangan “X”, hasil dari prediksi porositas tersebut kemudian dibandingkan dengan data *core* dan hasil prediksi porositas dengan metode deterministik.

Dalam prediksi porositas yang dilakukan secara deterministik, dapat dilakukan dengan tiga cara, dengan menggunakan *log* densitas, dan kombinasi *log* densitas dan *log neutron* [10], [11] dan *sonic log* [10], [11], dalam studi ini perhitungan dilakukan dengan memanfaatkan kombinasi *Log NPHI* dan *RHOB*, yang dimana dapat memberikan perhitungan porositas yang akurat pada lithologi formasi yang kompleks [3] secara matematis dapat dirumuskan sebagai:

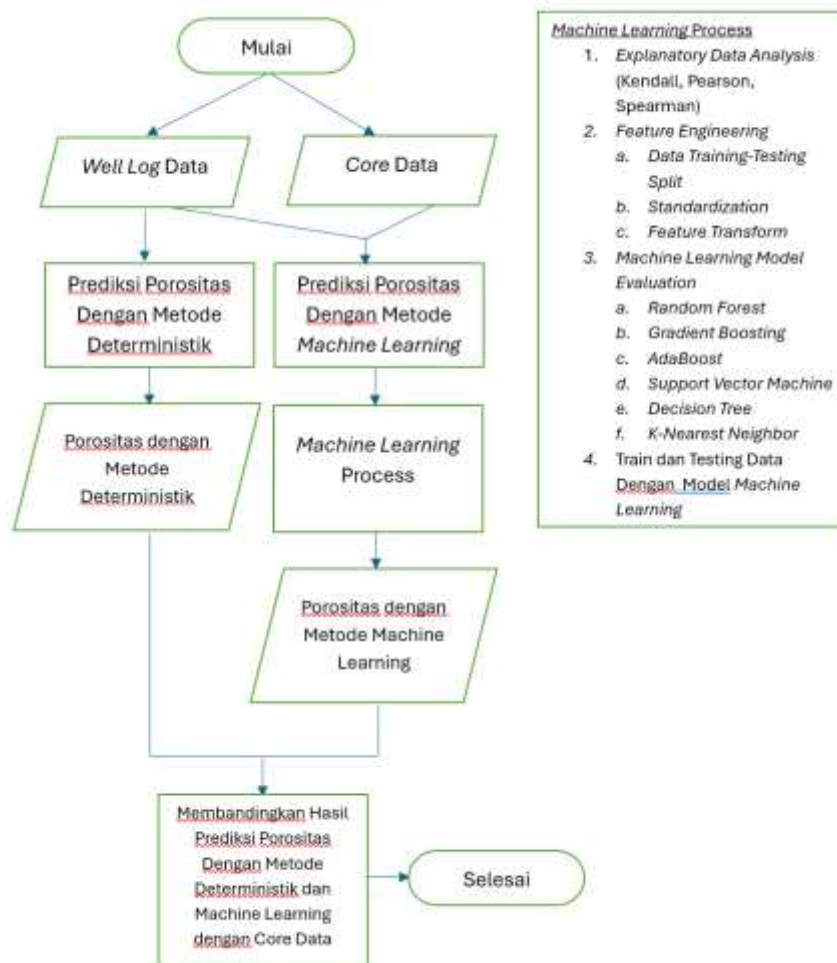
$$\phi_{N-D} = \left(\frac{\phi_N^2 - \phi_D^2}{2} \right)^{\frac{1}{2}} \quad (1)$$

dimana :

ϕ_{N-D} = Porositas *Neutron – Density* (fraksi)

ϕ_N = Porositas *Neutron* (fraksi)

ϕ_D = Porositas *Density* (fraksi)



Gambar 1. Bagan Alir (Flowchart) Studi Penentuan Porositas



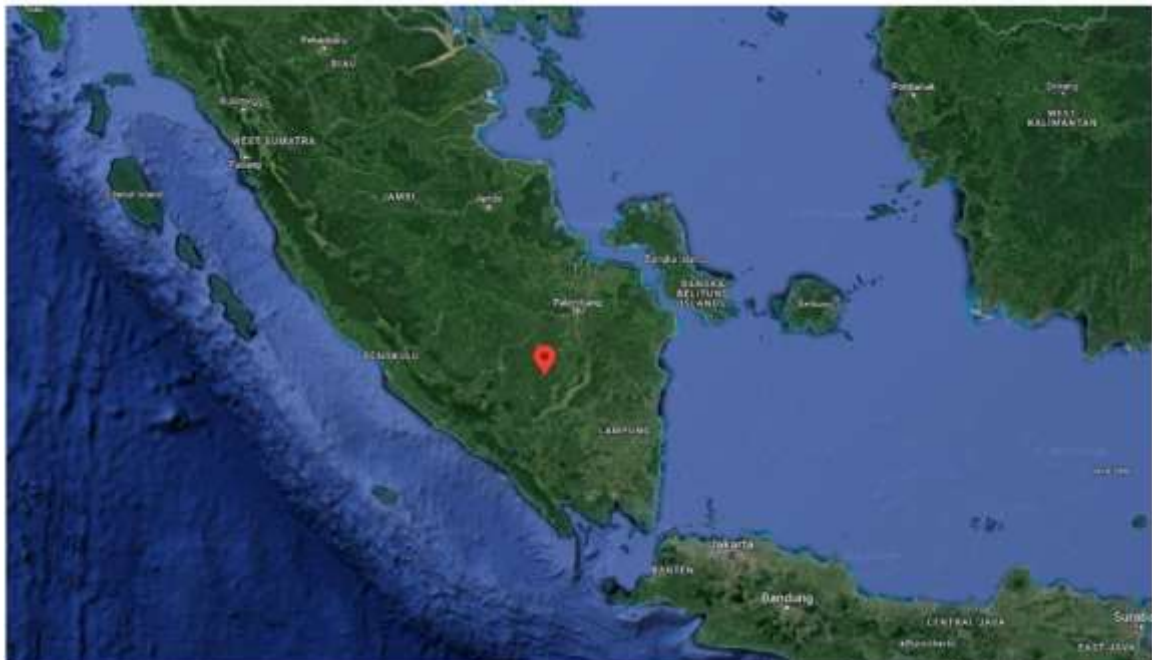
Tabel 1. Daftar Model *Machine Learning* Yang Digunakan

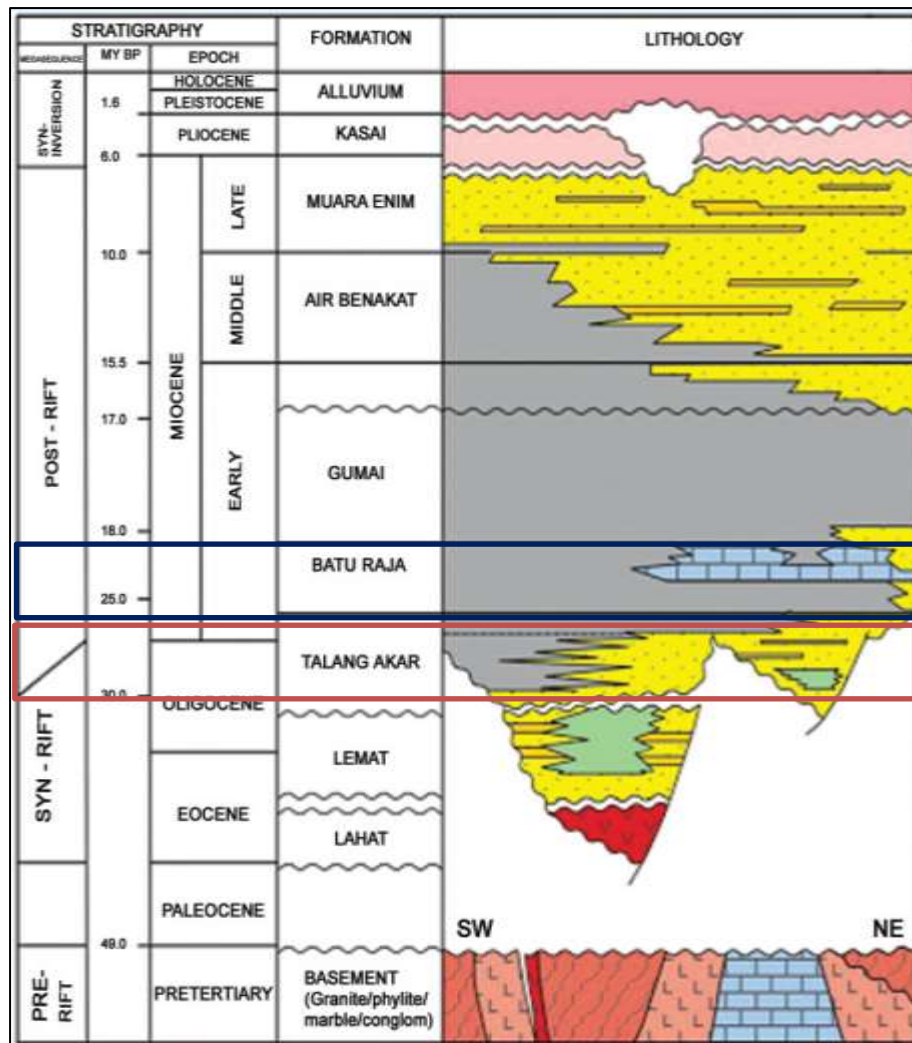
Model	Singkatan
DT	<i>Decision Tree</i>
GBR	<i>Gradient Boosting</i>
ADA	<i>AdaBoost</i>
RF	<i>Random Forest</i>
SVM	<i>Support Vector Machine</i>
KNN	<i>K-Nearest Neighbor</i>

III. HASIL DAN PEMBAHASAN

Lapangan “X” berada pada 105 km dari Kota Palembang, Provinsi Sumatera Selatan, dan berada pada cekungan Sumatera Selatan. Cekungan Sumatera Selatan berada disebelah timur pegunungan bukit barisan dan meluas ke daerah *offshore* dan paparan sunda disebelah timur laut [12]. Kolom stratigrafi Cekungan Sumatera Selatan dapat dilihat pada Gambar 3. Formasi dari berusia tua ke muda adalah *Basement*, Formasi Lahat, Formasi Talang Akar, Formasi Baturaja, Formasi Gumai, Formasi Air Benakat, Formasi Muara Enim dan Formasi Kasai. Fokus dari studi yang dilakukan adalah pada Formasi Baturaja dan Talang Akar. Formasi Baturaja merupakan formasi yang didominasi oleh batuan karbonat dan formasi talang akar didominasi oleh batu pasir yangh diselingi oleh *shale* [12]. Untuk persebaran data *well log* dan data *core* yang meliputi Formasi Baturaja dan Talang Akar dapat dilihat pada Tabel 2 dan Tabel 3.

Gambar 2. Lokasi Lapangan “X”





Gambar 3. Cekungan Sumatera Bagian Selatan Yang Menggambarkan Formasi Baturaja dan Formasi Talang Akar

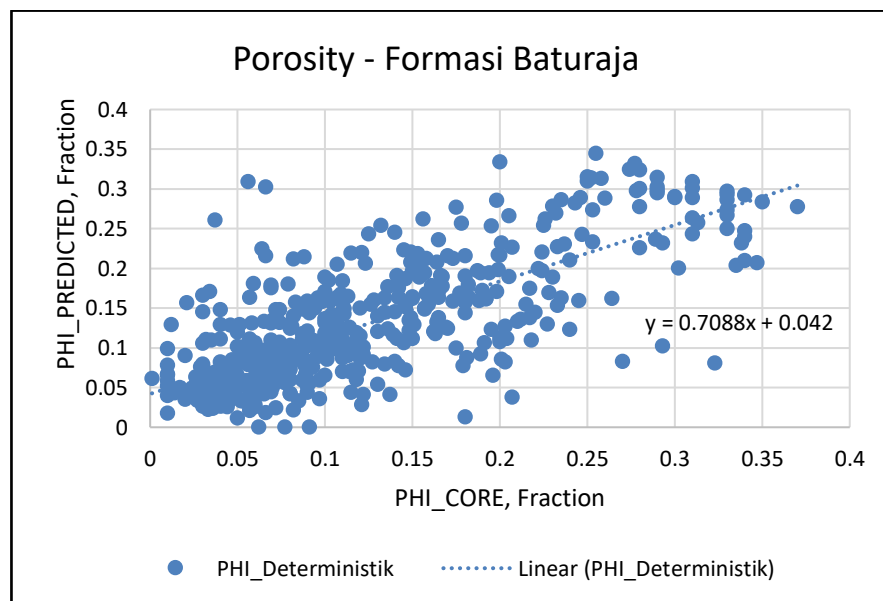
Tabel 2. Statistik Data *Well Log* dan Data *Core* pada Formasi Baturaja

<i>Well Log Data/ Well Core Data</i>	<i>count</i>	<i>mean</i>	<i>Standar Deviasi</i>	<i>min</i>	<i>25%</i>	<i>50%</i>	<i>75%</i>	<i>max</i>
CALI	4935	9.1197	0.6412	8.0285	8.5780	9.0030	9.5736	11.3110
CXO	5823	0.0849	0.0821	0.0006	0.0315	0.0598	0.1139	0.7307
SP	6042	-0.6135	81.3050	182.7625	38.0962	18.0330	66.2098	109.8735
DT	6647	69.6346	8.3572	53.2940	63.7595	67.8360	74.1279	132.4170
GR	7411	61.0114	22.8109	18.4378	44.8208	55.9033	72.7646	194.0810
LLD	7411	19.5558	33.4880	0.6197	7.1046	13.3003	22.3224	697.7159
LLS	7411	18.1422	17.8344	0.5932	7.5650	13.8520	22.4624	372.7881
MLL	7411	29.8704	142.7289	0.5112	8.4892	16.7382	31.6113	11048.0830
NPHI	6971	0.0980	0.0616	-0.0189	0.0532	0.0880	0.1362	0.4390
PEF	6451	5.0563	1.0692	2.3811	4.4449	4.9232	5.5835	22.7193
RHOB	6451	2.5031	0.1161	1.9548	2.4423	2.5315	2.5931	2.7318
PHI_CORE	491	0.1216	0.0833	0.0010	0.0570	0.0990	0.1695	0.3700

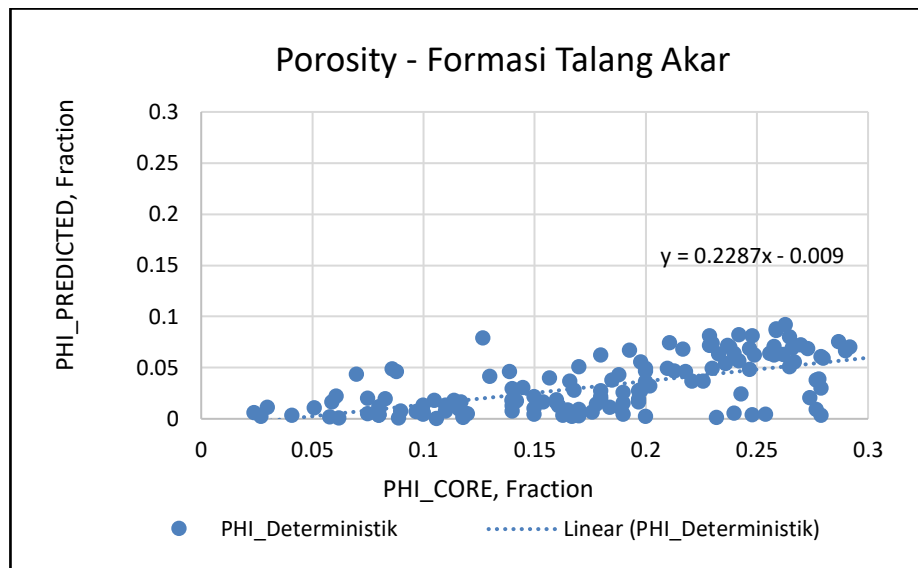
Tabel 3. Statistik Data *Well Log* dan Data *Core* pada Formasi Talang Akar

<i>Well Log Data/ Well Core Data</i>	<i>count</i>	<i>mean</i>	<i>Standar Deviasi</i>	<i>min</i>	<i>25%</i>	<i>50%</i>	<i>75%</i>	<i>max</i>
CALI	1328	8.9540	0.7299	7.8940	8.3935	8.6598	9.5320	13.4199
CXO	1932	0.2124	0.1774	0.0114	0.1140	0.1751	0.2398	1.2554
SP	2058	2.2552	67.2599	154.7500	-3.1488	15.0500	46.9973	83.6010
DT	2414	86.1628	9.5144	57.9121	80.2096	85.7508	91.1688	135.3620
GR	2423	102.5172	39.7233	23.1296	69.2700	102.0193	135.0227	258.5710
LLD	2432	5.5108	4.2587	1.3536	3.6043	4.6236	6.2783	64.2568
LLS	2432	5.8213	4.0441	1.7050	3.8897	5.0564	6.5553	56.5256
MLL	2432	7.4174	7.8839	0.8758	3.8748	5.7507	8.3361	132.3905
NPHI	2432	0.2505	0.0927	0.0192	0.1958	0.2407	0.2960	0.9743
PEF	2432	3.0199	0.8322	1.2607	2.3830	2.9203	3.4803	6.9902
RHOB	2432	2.4036	0.1691	1.3213	2.3154	2.4454	2.5124	2.9115
PHI CORE	144	0.1784	0.0711	0.0240	0.1180	0.1820	0.2400	0.3070

Dalam prediksi porositas dengan menggunakan metode deterministik dengan menggunakan kombinasi *log* NPHI dan RHOB dihasilkan Nilai MAE sebesar 0.0658 dan RMSE sebesar 0.0906 untuk Lapangan “X”. Selanjutnya, hasil dari prediksi porositas dengan metode deterministik kemudian dilakukan *plotting* terhadap data *core* pada Formasi Baturaja dan Talang Akar seperti yang terlihat pada Gambar 3 dan Gambar 4 di bawah.

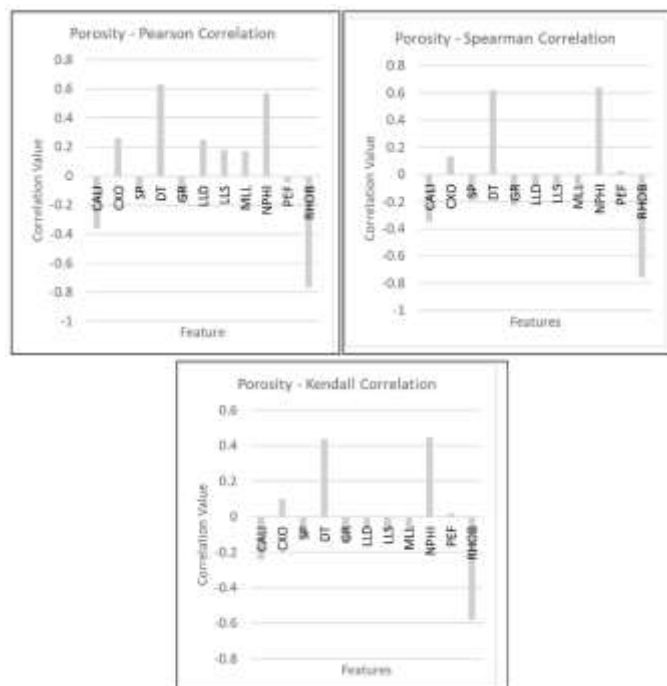


Gambar 3. Plot Validasi Porositas *Core* Terhadap Hasil Prediksi Porositas Dengan Metode Deterministik Pada Formasi Baturaja



Gambar 4. Plot Validasi Porositas *Core* Terhadap Hasil Prediksi Porositas Dengan Metode Deterministik Pada Formasi Talang Akar

Setelah dilakukan prediksi porositas dengan metode deterministik, selanjutnya dilakukan prediksi porositas dengan metode *machine learning*, sebelum dilakukan proses *testing* dan *training*, dilakukan *feature selection* dengan menggunakan 3 korelasi yaitu korelasi *Pearson*, *Spearman* dan *Kendall* yang ditampilkan dalam bentuk grafik untuk konsistensi dari pengaruh data log terhadap porositas. Berdasarkan tiga korelasi tersebut seperti yang terlihat pada Gambar 5 di bawah, data *log* yang sangat berpengaruh terhadap penentuan porositas berdasarkan tiga korelasi tersebut adalah CALI, NPHI, RHOB, DT dan GR. Dan selanjutnya data *log* tersebut digunakan sebagai data input dalam prediksi porositas dengan menggunakan *machine learning*. Hasil *training*, *testing* dan *cross-validation* dari prediksi masing-masing model *machine learning* dapat dilihat pada Tabel 4.



Gambar 5. Grafik hubungan korelasi *log* dengan Porositas dengan *Pearson Correlation*, *Spearman Correlation* dan *Kendall Correlation*



Tabel 4. Skor Model *Machine Learning* Untuk Prediksi Porositas

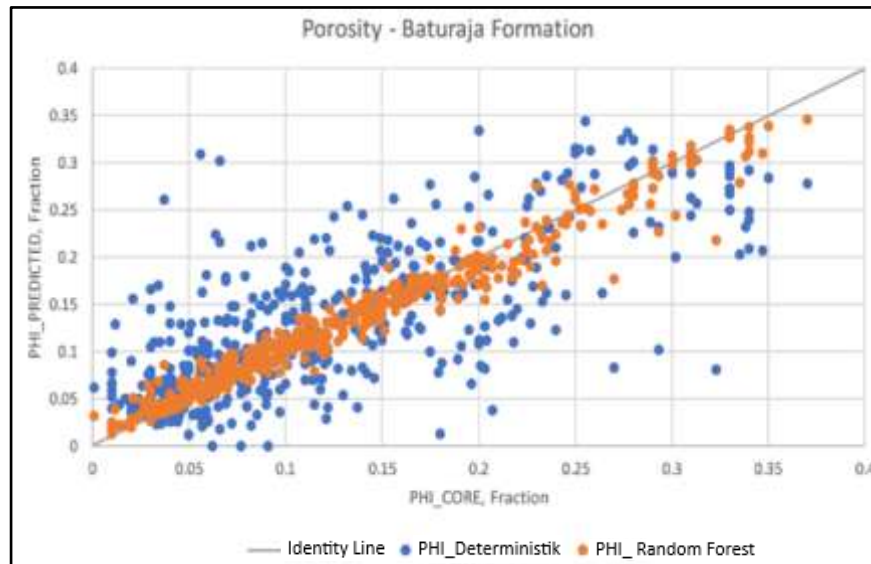
Training			
	Model	MAE	RMSE
RF	<i>Random Forest</i>	0.0134	0.0185
KNN	<i>K-Nearest Neighbor</i>	0.0302	0.0413
GBR	<i>Gradient Boosting</i>	0.0196	0.0255
ADA	<i>AdaBoost</i>	0.0366	0.0436
SVM	<i>Support Vector Machine</i>	0.0540	0.0614
DT	<i>Decision Tree</i>	0.0000	0.0000
Testing			
	Model	MAE	RMSE
RF	<i>Random Forest</i>	0.0329	0.0434
KNN	<i>K-Nearest Neighbor</i>	0.0346	0.0450
GBR	<i>Gradient Boosting</i>	0.0356	0.0463
ADA	<i>AdaBoost</i>	0.0425	0.0515
SVM	<i>Support Vector Machine</i>	0.0426	0.0586
DT	<i>Decision Tree</i>	0.0599	0.0674
Cross-Validation			
	Model	MAE	RMSE
RF	<i>Random Forest</i>	0.0357	0.0492
KNN	<i>K-Nearest Neighbor</i>	0.0367	0.0497
GBR	<i>Gradient Boosting</i>	0.0377	0.0509
ADA	<i>AdaBoost</i>	0.0433	0.0544
SVM	<i>Support Vector Machine</i>	0.0562	0.0653
DT	<i>Decision Tree</i>	0.0479	0.0648

Setelah dilakukan *training* dan *testing*, terlihat pada Tabel 4 di atas, pada data *training*, terlihat model terbaik dalam prediksi porositas adalah *Decision Tree*, namun performa model dalam memprediksi porositas pada data *testing* dan *cross validation* menjadi rendah, hal ini mengindikasikan model tersebut mengalami *overfitting*, pada data *testing* dan *cross validation* terlihat model *Random Forest* merupakan model yang terbaik dalam memprediksi porositas dengan MAE sebesar 0.0329 dan RMSE 0.0434 pada data *testing* dan 0.0357 dan 0.0492 pada *cross validation*, sehingga model *Random Forest* merupakan model terbaik dalam memprediksi porositas pada Lapangan “X”. Selanjutnya hasil dari prediksi porositas dengan menggunakan *machine learning* tersebut dibandingkan dengan metode deterministik dan hasilnya dapat dilihat pada Tabel 5, dan plot antara hasil prediksi dan data *core* untuk masing-masing metode pada masing-masing formasi dapat dilihat pada Gambar 6 dan Gambar 7.

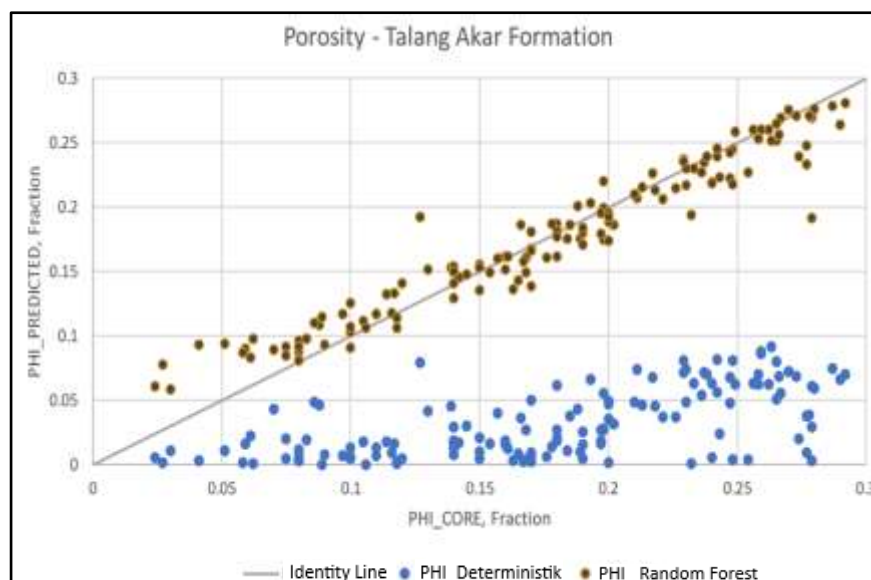
Seperti terlihat pada Tabel 5, Gambar 6 dan Gambar 7, metode *Machine Learning* dapat memprediksikan nilai porositas lebih baik dibandingkan dengan metode deterministik. Adapun prediksi porositas dapat dipengaruhi oleh adanya porositas sekunder yang terjadi secara alami pada Formasi Baturaja dan adanya perselingan batuan *shale* pada Formasi Talang akar yang dapat mempengaruhi pembacaan *log* khususnya NPHI dan RHOB yang menjadi parameter utama dalam prediksi porositas dengan metode deterministik. Selain itu hasil dari prediksi porositas dengan metode deterministik lebih tersebar (*scatter*) pada Formasi Baturaja dan terjadi *underfit* pada Formasi Talang Akar jika dibandingkan dengan metode *machine learning*. Hal ini terjadi diakibatkan oleh perbedaan kondisi pengujian dimana data *log* melakukan perekaman properti *reservoir* pada kondisi di *reservoir* jika dibandingkan dengan data *core* yang diperoleh dengan proses *coring* dan diuji di laboratorium. Selain adanya perbedaan pengujian, adanya deformasi pada *core* yang diuji akan berdampak signifikan terhadap *Specific Surface Area* yang berkontribusi pada perubahan nilai porositas jika dibandingkan dengan data *log*[13]. Adapun visualisasi hasil prediksi per sumur dapat dilihat pada Gambar 8 di bawah.

Tabel 5. Hasil Prediksi Porositas Untuk Setiap Metode

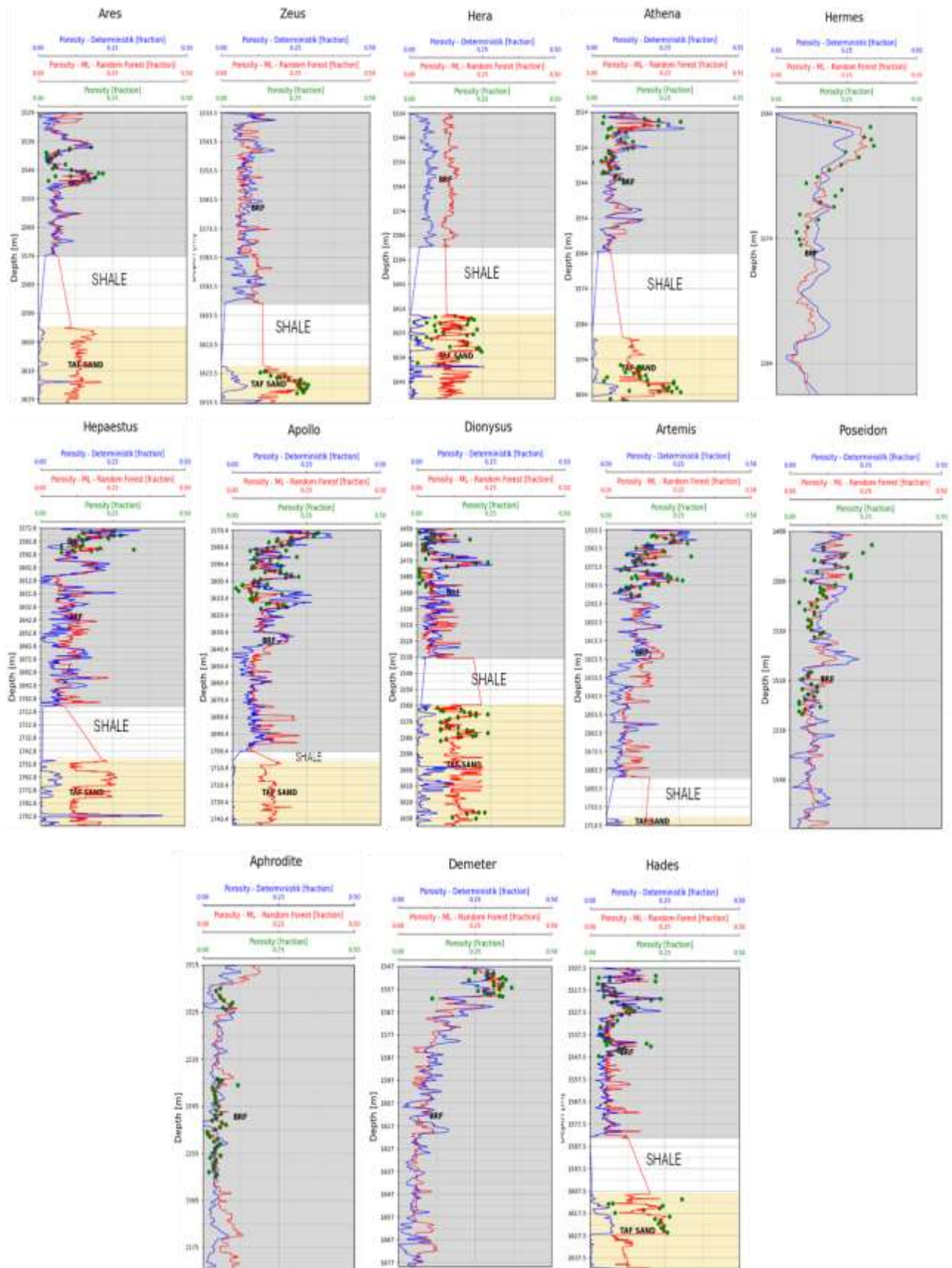
Metode Prediksi Porositas	MAE	RMSE
<i>Machine Learning</i>	0.0329	0.0434
Deterministik	0.0658	0.0906



Gambar 6. Plot Prediksi Porositas vs Porositas *Core* Pada Formasi Baturaja



Gambar 7. Plot Prediksi Porositas vs Porositas *Core* Pada Formasi Talang Akar



Gambar 8. Plot Hasil Porositas Dengan Metode Deterministik dan Metode *Machine Learning* dan Data *Core* Pada Masing-Masing Sumur Pada Lapangan “X”



IV. KESIMPULAN

Berdasarkan Hasil dan Pembahasan di atas, dapat disimpulkan :

1. Prediksi Porositas dengan dua metode yaitu metode deterministik dan metode *Machine Learning*, dan menghasilkan nilai MAE dan RMSE sebesar 0.0658 dan 0.0906 untuk metode deterministik dan 0.0329 dan 0.0434 untuk metode *Machine Learning*, dari nilai MAE dan RMSE tersebut, metode *Machine Learning* dapat digunakan sebagai alat untuk memprediksi porositas.
2. Hasil dari prediksi porositas tersebut dapat dipengaruhi oleh adanya porositas sekunder pada Formasi Baturaja dan adanya perselingan batuan *shale* pada Formasi Talang Akar selain itu adanya hasil prediksi porositas yang *scatter* pada Formasi Baturaja dan *underfit* pada Formasi Talang Akar pada metode deterministik, hal ini terjadi karena adanya perbedaan kondisi pengujian antara data *core* dan data *log*.

DAFTAR PUSTAKA

- [1] D. Tiab dan E. C. Donaldson, *Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties*. dalam EngineeringPro collection. Elsevier Science, 2012. [Daring]. Tersedia pada: <https://books.google.co.id/books?id=F2wxxHmpWjgC>
- [2] J. Wu, R. Luo, L. Luo, C. Lei, dan X. Chen, "Advanced machine learning for low-data porosity and permeability prediction in tight sandstones," *Geophysics*, vol. 90, no. 3, hlm. M31–M44, Mar 2025, doi: 10.1190/geo2024-0340.1.
- [3] G. Asquith dan D. Krygowski, *Basic well log analysis*. American Association of Petroleum Geologists, 2004. doi: 10.1306/Mth16823.
- [4] C. Zou, L. Zhao, F. Hong, Y. Wang, Y. Chen, dan J. Geng, "A comparison of machine learning methods to predict porosity in carbonate reservoirs from seismic-derived elastic properties," *Geophysics*, vol. 88, no. 2, hlm. B101–B120, Mar 2023, doi: 10.1190/geo2021-0342.1.
- [5] Y. Qi dkk., "Prediction of Carbonate Reservoir Porosity Based on CNN-BiLSTM-Transformer," *Applied Sciences (Switzerland)*, vol. 15, no. 7, Apr 2025, doi: 10.3390/app15073443.
- [6] Aliyuda Ali, Abdulwahab Muhammed Bello, dan Jerry Raymond, "Machine Learning Algorithms for Predicting Reservoir Porosity using Stratigraphic-dependent Parameters," *Global Journal of Computer Science and Technology*, vol. 22, no. G1, hlm. 15–25, Mei 2022, [Daring]. Tersedia pada: <https://gjst.com/index.php/gjst/article/view/2094>
- [7] M. Hussain, S. Liu, W. Hussain, Q. Liu, H. Hussain, dan U. Ashraf, "Application of Deep Learning for Reservoir Porosity Prediction and Self Organizing Map for Lithofacies Prediction," *J Appl Geophy*, vol. 230, hlm. 105502, 2024, doi: <https://doi.org/10.1016/j.jappgeo.2024.105502>.
- [8] A. Isah, Z. Tariq, A. Mustafa, M. Mahmoud, dan E. R. Okoroafor, "A Review of Data-Driven Machine Learning Applications in Reservoir Petrophysics," *Arab J Sci Eng*, 2025, doi: 10.1007/s13369-025-10329-0.
- [9] R. C. Rohmana, D. Triwanti, dan P. R. Setiyaningrum, "Penerapan Machine Learning dalam Penentuan Porositas Batuan: Studi Kasus Menggunakan Regresi Linier Berganda dan Regresi KNN pada Data Log Sumur Application of Machine Learning in Rock Porosity Determination: Case Study Using Multiple Linear Regression and KNN Regression on Well Log Data," *Jurnal Teknik*, vol. 13, hlm. 42–50, 2024, [Daring]. Tersedia pada: <http://jurnal.umt.ac.id/index.php/jt/index>
- [10] O. Serra, *Fundamentals of Well-log Interpretation: The interpretation of logging data*. dalam Developments in petroleum science. Elsevier, 1984. [Daring]. Tersedia pada: <https://books.google.co.id/books?id=OFYZAQAAIAAJ>
- [11] Schlumberger, *Log Interpretation Principles/Application*. USA: Schlumberger Educational Services, 1989.



- [12] D. Ginger, K. Fielding, dan I. P. Association, "IPA05-G-039 The Petroleum Systems and Future Potential of the South Sumatra Basin," dalam *Annual convention; 30th, Indonesian Petroleum Association*, Jakarta: IPA;, hlm. 67–90. [Daring]. Tersedia pada: <https://www.tib.eu/de/suchen/id/BLCP%3ACN059573091>
- [13] H. B. Helle, A. Bhatt, dan B. Ursin, "Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study," *Geophys Prospect*, vol. 49, no. 4, hlm. 431–444, Jul 2001, doi: <https://doi.org/10.1046/j.1365-2478.2001.00271.x>.