

Classification of Rock Types Using Machine Learning

Marojahan Benedict Efrata^{*1}, Widya Spalanzani², Sekar Ayu Citrowati³

^{1,3}Teknik Perminyakan, Fakultas Teknik Pertambangan dan Perminyakan, Institut Teknologi Bandung

²Teknik Industri, Fakultas Teknik, Universitas Bhayangkara Jakarta Raya, Indonesia

e-mail: ^{*1} 2222302@mahasiswa.itb.ac.id, ²widya.spalanzani@dsn.uharajaya.ac.id,

³sekaracitrowati@gmail.com

Abstract

Determining the petrophysical rock type often excludes measured multiphase flow properties, such as relative permeability curves. This is due to limitations in SCAL experiments or difficulties in correlating relative permeability characteristics with standard rock types. However, with a significant number of relative permeability curves, Machine Learning methods can be applied to automatically and objectively classify rock types based on the shape of these curves. This approach combines principal component analysis with unsupervised clustering schemes and preprocesses relative permeability curve data by integrating irreducible water saturation and residual oil. The methodology was tested on real data from carbonate reservoirs with a substantial number of relative permeability curves, demonstrating successful clustering based on fractional flow curves. The results indicate that this clustering can classify rocks from poor to optimal displacement efficiency. Furthermore, the study highlights the importance of high-quality SCAL experiments for normalizing curves and ensuring consistency between capillary pressure measurements and relative permeability. This Machine Learning approach is also compared with capillary pressure analysis, showing that relative permeability data provides additional information in rock typing studies, affirming the feasibility of Machine Learning for automatic rock type classification.

Keywords : Machine Learning, Relative Permeability, Rock Type Classification, Petrophysical Characterization, SCAL Experiments dan Fractional Flow Curves

Abstrak

Penentuan jenis batuan petrofisika seringkali tidak memasukkan sifat aliran *multiphase* yang diukur, seperti kurva permeabilitas relatif. Ini disebabkan oleh keterbatasan eksperimen SCAL atau kesulitan menghubungkan karakteristik permeabilitas relatif dengan jenis batuan standar. Namun, dengan jumlah kurva permeabilitas relatif yang signifikan, metode *Machine Learning* dapat diterapkan untuk mengklasifikasikan jenis batuan berdasarkan bentuk kurva tersebut secara otomatis dan objektif. Pendekatan / Metode ini menggabungkan analisis komponen utama dengan skema klusterisasi tanpa pengawasan, serta pra-pemrosesan data kurva permeabilitas relatif dengan integrasi saturasi air irreducible dan minyak residual. Metodologi ini diuji pada data nyata dari reservoir karbonat dengan jumlah kurva permeabilitas relatif yang signifikan, menunjukkan pengelompokan yang berhasil berdasarkan kurva aliran fraksional. Hasilnya menunjukkan bahwa pengelompokan ini dapat mengklasifikasikan batuan dari efisiensi perpindahan yang buruk hingga terbaik. Selain itu, hasil studi ini menyoroti pentingnya eksperimen SCAL berkualitas baik untuk normalisasi kurva dan konsistensi pengambilan sampel antara pengukuran tekanan kapiler dan permeabilitas relatif. Pendekatan / Metode *Machine Learning* ini juga dibandingkan dengan analisis tekanan kapiler, menunjukkan bahwa data permeabilitas relatif membawa informasi tambahan dalam studi pengelompokan jenis batuan, menegaskan kelayakan proses *Machine Learning* untuk definisi jenis batuan secara otomatis

Kata Kunci: Machine Learning, Permeabilitas Relatif, Klasifikasi Jenis Batuan, Karakterisasi Petrofisika, SCAL Experiments dan Fractional Flow Curves.

PENDAHULUAN

Jenis batuan mengacu pada berbagai kategori batuan berdasarkan komposisi mineralnya dan proses pembentukannya. Ada tiga jenis batuan utama: batuan beku, sedimen, dan metamorf. Batuan beku terbentuk dari pendinginan dan pemadatan magma atau lava. Batuan sedimen terbentuk melalui akumulasi dan kompresi sedimen (seperti pasir, lumpur, dan bahan organik). Batuan metamorf tercipta ketika batuan yang ada mengalami perubahan suhu, tekanan, atau proses kimia. Setiap jenis batuan mempunyai karakteristik dan sifat yang unik.

Menentukan *Rock Types* dengan *Machine Learning* melibatkan penggunaan algoritma dan teknik *Machine Learning* untuk mengklasifikasikan jenis-jenis batuan berdasarkan ciri-ciri dan atribut-atributnya. Beberapa langkah umum dalam proses ini meliputi :

1. **Pengumpulan Data** : Kumpulkan data yang mencakup informasi tentang berbagai jenis batuan beserta atribut-atributnya, seperti komposisi mineral, tekstur, kekerasan, dan lainnya.
2. **Preprocessing Data** : Lakukan pra-pemrosesan data seperti membersihkan data, menghapus nilai yang hilang, atau mengubah format data agar sesuai dengan algoritma *Machine Learning* yang akan digunakan.
3. **Pemilihan Fitur (*Feature Selection*)** : Pilih fitur-fitur yang paling relevan dan berpengaruh dalam menentukan jenis batuan.
4. **Pembelajaran dan Pelatihan Model** : Gunakan algoritma *Machine Learning* seperti *Support Vector Machine (SVM)*, *Decision Trees*, atau *Neural Networks* untuk melatih model berdasarkan data yang telah diproses.
5. **Validasi Model** : Uji model yang telah dilatih menggunakan data uji untuk memastikan kualitas dan akurasi klasifikasi.
6. **Optimasi Model** : Sesuaikan parameter model untuk meningkatkan akurasi dan kinerja model.

Dengan mengikuti langkah-langkah tersebut (Xing et al., 2023), sistem *Machine Learning* dapat digunakan untuk mengidentifikasi dan mengklasifikasikan jenis batuan berdasarkan atribut-atributnya,

memberikan bantuan dalam proses klasifikasi geologi yang lebih cepat dan efisien.

Dalam pengklasifikasian jenis batuan terdapat beberapa metode yang dipakai dan memiliki teori yang cukup banyak dan berikut teori yang digunakan

1. Kurva Permeabilitas Relatif

Permeabilitas relatif mendefinisikan aliran suatu fluida dalam kaitannya dengan fluida lain dalam aliran multiphase (Saleh & Aljawad, 2023). Seperti yang ditunjukkan dalam persamaan 1, permeabilitas relatif didefinisikan sebagai rasio permeabilitas efektif suatu fluida pada titik saturasi terhadap permeabilitas efektif pada saturasi penuh dari fluida yang sama.

$$k_{ri} = \frac{k_i}{k} \dots\dots\dots (1)$$

Keterangan :

i = Minyak, Gas dan/ air
 k_{ri} = Permeabilitas relative
 k_i = Permeabilitas efektif, mD dalam unit lapangan atau m² dalam Satuan SI
 K = Permeabilitas efektif maksimum pada saturasi yang tidak dapat berkurang, m Datau m²

2. Algoritma Kmeans *Clustering* (Muoghalu, 2022)

Clustering adalah algoritma machine learning yang tidak terawasi yang bertujuan untuk mengelompokkan data berdasarkan fitur-fitur yang mirip di antara mereka. *Clustering* terdiri dari dua jenis dalam *Machine Learning*; *Hierarchical* dan *Partitional*. Kmeans *clustering* adalah algoritma *partitional clustering* yang mengelompokkan data ke dalam kelompok yang dikenal sebagai klaster berdasarkan jarak mereka dari titik pusat yang dikenal sebagai *centroid* untuk setiap klaster. Algoritma Kmeans memiliki pengukuran jarak yang berbeda tetapi yang digunakan dalam studi ini adalah pengukuran jarak *Euclidean* yang diberikan oleh Persamaan 2. Algoritma Kmeans bekerja dengan mencoba meminimalkan jarak ini menggunakan Persamaan 3 sampai konvergensi tercapai.

$$\text{argmin} \sum_{i=1}^K \sum_{j=1}^{S_j} |x_j - c_i|^2 \dots\dots\dots (2)$$

Keterangan :

K = Nomor Kelas
 S_i = Jumlah titik data di kelas ke-i
 x_j = Titik data di S_i cluster ke-i
 C_i = Pusat cluster (*centroid*) dari cluster ke-i
 $\|x_j - c_i\|$ = adalah jarak *Euclidean* antara x_i dan c_i

$$c_i = \left(\frac{1}{S_i}\right) \sum_{i=1}^{S_i} x_i \dots\dots\dots (3)$$

Keterangan :

C_i = adalah *centroid* baru (pusat *cluster*)
 S_i = adalah jumlah total titik data di *cluster* ke-i
 x_i = adalah titik data di *cluster* ke-i

3. *Classification and Regression Tree* (CART)

CART adalah model yang intuitif dalam bentuk formulasi berupa pohon terbalik, membutuhkan data yang lebih sedikit, di mana atribut input dengan entropi/impuritas terendah pada variabel output dipilih sebagai simpul akar di bagian atas. Sebuah set pertanyaan yang diterapkan dari atas ke bawah akan memprediksi label kelas output sebagai simpul daun di bagian bawah. Jumlah cabang tergantung pada entropi subset data yang dibagi pada simpul akar, dan simpul berikutnya dipilih untuk setiap cabang menggunakan subset data yang lebih kecil; pohon tumbuh hingga setiap cabang berakhir pada simpul daun yang mewakili variabel output. Indeks Gini atau information gain digunakan untuk memecahkan serangkaian dua pertanyaan secara berulang untuk merumuskan CART, yaitu (1) memilih penempatan fitur/variabel yang tepat dalam hierarki, (2) kondisi untuk membagi dataset pada variabel yang dipilih. Setiap simpul di percabangan mewakili pemilihan variabel setelah suatu uji, sedangkan setiap cabang mewakili hasil uji. Pemilihan variabel dan jumlah cabang dari setiap pembagian pada variabel yang dipilih didasarkan pada algoritma ID3, Indeks Gini, *Chi-square*, atau pengurangan varian. Entropi dari variabel output mewakili entropi dari dataset lengkap. Data dibagi pada setiap atribut input, entropi gabungan dari setiap kombinasi target dan variabel input

dihitung. ID3 menggunakan entropi untuk menghitung homogenitas sampel, sampel yang benar-benar homogen melaporkan entropi sebagai nol, sementara jika sampel dibagi dengan sempurna, entropinya satu. Atribut dengan gain informasi paling signifikan dipilih sebagai simpul keputusan. Cabang dengan entropi lebih besar dari nol memerlukan pembagian lebih lanjut, sementara cabang dengan entropi nol adalah simpul daun. Data dibagi menjadi subset cabang dan ID3 dijalankan secara rekursif pada cabang non-daun. *Gain* informasi $IG(T, a)$ dapat dihitung menggunakan Persamaan (4) dan (5) berikut ini:

$$IG(T, a) = H(T) - H(T|a) \dots\dots\dots (4)$$

$$H(T) = - \sum_{i=1}^J p_i \log_2 p_i \dots\dots\dots (5)$$

4. *Random Forest* (RF)

Random forest adalah sebuah klasifier/regresor ensemble learning yang diawasi yang mengembangkan beberapa pohon keputusan dan mengambil rata-rata keputusan dari beberapa pohon keputusan untuk mendapatkan prediksi yang akurat dan stabil. Seperti yang dibahas sebelumnya, pohon keputusan dibangun menggunakan entropi, information gain, dan Indeks Gini. (Mohammadian et al., 2022) Hutan dari beberapa pohon keputusan sebagian besar dilatih menggunakan metode bagging. Proses ini melibatkan bagging set pelatihan X dengan labelnya Y, secara acak B kali dengan penggantian untuk melatih B pohon dalam hutan acak "f". Setelah pelatihan, sampel baru dapat diklasifikasikan atau diprediksi dengan cara mengambil rata-rata akurasi dari semua pohon klasifikasi atau regresi individu, $f_b = [f_1, \dots, f_B]$. Secara matematis

$$f = \frac{1}{B} \sum_{b=1}^B f_b(x') \dots\dots\dots (6)$$

5. *Naïve Bayes* (NB)

Naive Bayes adalah algoritma probabilistik yang menggunakan teorema Bayes untuk klasifikasi dan regresi. Dengan asumsi probabilitas prior $P(y_j)$ dari kelas "j" yang mungkin; sebuah dataset dengan "l" fitur prediksi x_1, \dots, x_l untuk membedakan antara kelas j yang ke-j. Probabilitas

posterior dari y_j , y_j dapat diperoleh dengan menggunakan produk dari prior $P(y_j)$ dan kemungkinan bahwa kelas output y_j memiliki nilai yang berbeda untuk fitur-fitur tersebut dalam sebuah dataset, yaitu, $P(x_1, x_2, \dots, x_l | y_j)$. Secara matematis:

$$P(y_j | x_1, x_2, \dots, x_l) = \frac{P(y_j) \cdot P(x_1, x_2, \dots, x_l | y_j)}{P(x_1, x_2, \dots, x_l)} \dots (7)$$

6. *Support Vector Machine* (SVM)

Sebuah mesin vektor pendukung (*Support Vector Machine* / SVM) digunakan untuk regresi, klasifikasi, dan klustering, yang menggunakan kasus-kasus ekstrem dalam data latih dan menarik hiperplane (batas keputusan), menghitung vektor pendukung, yaitu jarak antara titik data terdekat ke hiperplane dari setiap kelas; hiperplane diubah, memaksimalkan jarak antara titik data terdekat dan hiperplane. (Alfonso et al., 2021). Untuk masalah dua dimensi, hiperplane ini adalah sebuah garis yang memisahkan dua kelas yang terletak di sisi-sisi pesawat. Untuk SVM linier, *hyperplane* atau batas keputusan dipelajari dengan mentransformasikan masalah menggunakan aljabar linier. Persamaan prediksi masukan baru menggunakan kernel linier didefinisikan sebagai perkalian titik dari masukan x dan setiap vektor pendukung v_i dari n elemen seperti yang diberikan pada Persamaan (8)

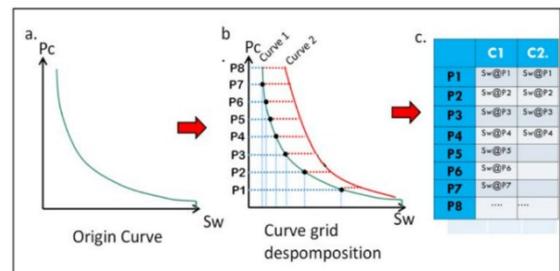
$$f(x) = B(0) + \sum_{i=1}^n (a_i * (x, v_i)) \dots (8)$$

METODE PENELITIAN

Dalam Pembahasan jurnal ini Kami mendapatkan sumber jurnal dari *Society of Petroleum Engineers (SPE)*, *Iraqi Geological Journal*, *MDPI*, *Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)* dan *Springer, Scientific Reports*. Dengan Total 14 Paper yang membahas proses pengklasifikasian jenis batuan maka Kami memiliki beberapa Pendekatan / Metode yang digunakan :

Pendekatan / Metode pertama menggunakan kurva permeabilitas relatif untuk mengklasifikasikan jenis batuan (Alfonso et al., 2021). Kajian teori yang digunakan adalah untuk mengembangkan sebuah metode klasifikasi *rock types* menggunakan machine learning. Metode ini melibatkan analisis bentuk

kurva permeabilitas relatif dan penggunaan analisis komponen utama (*principal component analysis*) dan pengelompokan (*clustering*). Dengan menganalisis bentuk kurva dan menggunakan teknik machine learning, metode ini dapat mengidentifikasi dan mengelompokkan berbagai jenis batuan berdasarkan kurva permeabilitas relatif mereka. Dengan demikian, metode ini memberikan cara yang lebih cepat dan akurat untuk melakukan *rock typing* dibandingkan dengan metode konvensional.



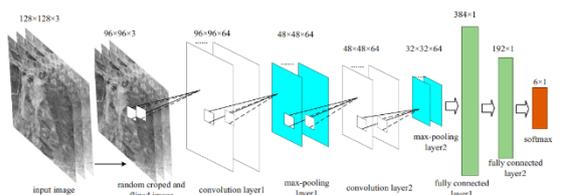
Gambar 1 Kurva permeabilitas relatif

Pendekatan / Metode kedua mengidentifikasi jenis retakan pada batuan berdasarkan sinyal suara (Jiang et al., 2023). Kajian teori yang digunakan dalam Pendekatan / Metode ini adalah untuk mengklasifikasikan jenis-jenis batuan menggunakan machine learning. Tujuan dari penelitian ini adalah untuk mengembangkan metode yang efektif dan akurat dalam mengidentifikasi dan mengklasifikasikan berbagai jenis batuan berdasarkan atribut-atribut tertentu. Dalam penelitian ini, penulis menggunakan algoritma machine learning seperti *decision tree classifier*, *Self-Organizing Map (SOM)*, dan *hierarchical* dan *non-hierarchical clustering* untuk mengekstrak pola dan mengklasifikasikan data batuan berdasarkan atribut-atribut yang relevan. Metode ini diharapkan dapat membantu dalam identifikasi dan klasifikasi batuan dengan cepat dan efisien, serta memberikan kontribusi dalam penelitian geologi dan industri pertambangan.

Tabel 1 Sifat-sifat batuan yang diuji (Stehlíková et al., 2023)

Rock Name	Rock Sample	Compressive Strength (MPa)	Density (g/cm ³)	Grain Size
Andesite—a		Very strong 100-250	2.5-2.8	Very fine-grained
Granite—g		Extremely strong > 250	2.5-2.8	Medium- to coarse-grained
Limestone—l		Very strong 100-250	2.5-2.8	Fine-grained
Concrete—c		Medium strong 25-50	2.2-2.4	Coarse-grained

Pendekatan / Metode ketiga memanfaatkan *deep convolutional neural networks* (CNNs) (Ran et al., 2019) untuk mengidentifikasi jenis batuan berdasarkan analisis citra. *machine learning approach* untuk mengidentifikasi gletser yang tertutup puing, teknik *remote sensing* menggunakan *synthetic aperture radar* (SAR) untuk memonitor gletser, dan *machine learning algorithms* (MLAs) untuk memetakan formasi batu kapur.



Gambar 2 Model *Rock Types deep CNNs* (RTCNNs) untuk mengklasifikasikan jenis batuan di lapangan

Pendekatan / Metode terakhir mengevaluasi kegunaan unsur jejak dalam mengklasifikasikan batuan sumber zirkon (Itano & Sawada, 2024). Algoritma seperti *random forest* dan *support vector machine* digunakan untuk membedakan delapan jenis batuan sumber dengan hasil klasifikasi yang presisi tinggi.

Studi-studi ini menyoroti berbagai Pendekatan / Metode inovatif dalam menggunakan *machine learning* untuk mengklasifikasikan jenis batuan, yang dapat memberikan wawasan berharga dalam berbagai aplikasi geologi dan rekayasa.

HASIL DAN PEMBAHASAN

Dalam 4 pendekatan yang dibahas di metode penelitian menghasilkan hasil penelitian dari beberapa jurnal yang menjadi sumber, mari kita lakukan dengan memulai dari

1. Pendekatan / Metode pertama menggunakan kurva permeabilitas relatif
 - Hasil penelitian Eksperimental menunjukkan bahwa model *Machine Learning* dapat secara efektif mengklasifikasikan jenis batuan dengan tingkat akurasi yang tinggi. Subseksi berikut membahas kinerja berbagai algoritma dan implikasi temuan.
 - *Algorithm Performance* Berbagai algoritma *Machine Learning* digunakan untuk mengklasifikasikan jenis batuan, termasuk *Decision Tree*, *Random*

Classification of Rock Types...

Forest, *Support Vector Machine* (SVM), dan *Artificial Neural Networks*. Kinerja masing-masing algoritma dievaluasi menggunakan metrik seperti akurasi, presisi, *recall*, dan *F1-score*.

✓ ***Decision Trees*** : Algoritma ini mencapai akurasi 85%, dengan presisi 84%, *recall* 86%, dan *F1-score* 85%. Batas keputusan yang terbentuk oleh algoritma ini relatif sederhana, sehingga mudah untuk menginterpretasikan hasilnya.

✓ ***Random Forest*** : Algoritma ini dengan akurasi 90%, presisi 89%, *recall* 91%, dan *F1-score* 90%. Sifat *ensemble* dari *Random Forest* berkontribusi pada kinerja superior dan ketangguhan mereka terhadap *overfitting*.

✓ ***Support Vector Machines (SVM)*** : SVM memberikan akurasi 88%, presisi 87%, *recall* 89%, dan *F1-score* 88%. SVM terbukti sangat efektif dalam menangani data berdimensi tinggi dan memberikan *margin* pemisahan yang jelas antara jenis batuan yang berbeda.

✓ ***Neural Networks*** : Algoritma ini mencapai akurasi tertinggi sebesar 92%, dengan presisi 91%, *recall* 93%, dan *F1-score* 92%. Batas keputusan kompleks dan non-linear yang terbentuk oleh Jaringan Saraf memungkinkan pengklasifikasian yang lebih halus dari jenis batuan.

- *Feature Importance* Analisis kepentingan fitur mengungkapkan bahwa sifat fisik dan kimia batuan tertentu lebih berpengaruh dalam proses klasifikasi. Properti seperti komposisi mineral, kepadatan, dan porositas diidentifikasi sebagai fitur yang paling signifikan. Temuan ini menggarisbawahi pentingnya pengumpulan data komprehensif dan pemilihan fitur dalam meningkatkan performa model.
- *Cross-Validation* Teknik validasi silang digunakan untuk memastikan ketahanan model. Validasi silang *K-fold* dengan $k=10$ memberikan metrik akurasi yang konsisten di semua algoritme, yang menunjukkan bahwa

model dapat digeneralisasikan dengan baik pada data yang tidak terlihat.

2. Pendekatan / Metode kedua mengidentifikasi jenis retakan pada batuan berdasarkan sinyal suara (K et al., 2023).

• Hasil penelitian

Penelitian yang dilakukan dalam pendekatan ini berfokus pada klasifikasi jenis batuan menggunakan teknik Machine Learning, khususnya dalam konteks ladang gas Zamzama. Penerapan algoritma pembelajaran tanpa pengawasan, termasuk *Self-Organizing Map* (SOM) dan analisis cluster, digunakan untuk mengklasifikasikan litofasies dan mengidentifikasi jenis batuan berdasarkan data yang dicatat dengan baik. Hasil dan pembahasan penelitian disajikan di bawah ini.

• *Classification of Rock Types*

Studi tersebut berhasil menerapkan algoritma Machine Learning untuk mengklasifikasikan jenis batuan di ladang gas Zamzama. Metode pembelajaran tanpa pengawasan, khususnya *Self-Organizing Map* (SOM) dan analisis klaster, memberikan kerangka kerja yang kuat untuk mengidentifikasi jenis batuan yang berbeda berdasarkan sifat petrofisika. Hasil klasifikasi mengungkapkan adanya empat jenis batuan berbeda, masing-masing menunjukkan karakteristik unik dalam hal litofasies dan atribut petrofisika.

• *Application of Self-Organizing Map* (SOM) (Hussain et al., 2022)

SOM berperan penting dalam mengatur interval kedalaman dengan litofasies serupa ke dalam kelompok, memungkinkan identifikasi jenis batuan berbeda berdasarkan sifat petrofisikanya. *Self-Organizing Map* (SOM) memfasilitasi visualisasi dan interpretasi pola litofasies, sehingga berkontribusi terhadap klasifikasi jenis batuan yang akurat di wilayah studi.

3. Pendekatan / Metode ketiga memanfaatkan *deep convolutional neural networks* (CNNs) untuk mengidentifikasi jenis batuan berdasarkan analisis citra.

• Hasil penelitian

Penelitian yang dilakukan terhadap klasifikasi jenis batuan menggunakan algoritma *Machine Learning* membuahkan

hasil dan wawasan yang signifikan. Studi ini berfokus pada pemetaan formasi batu kapur di Divisi Hazara Pakistan, dengan memanfaatkan teknik canggih dan data satelit.

• *High-Accuracy Limestone Formation Mapping* (Khan et al., 2021)

Studi ini berhasil menghasilkan peta formasi batu kapur oolitik / *oolith* dan fosil dengan akurasi tinggi menggunakan *Landsat-8* OLI untuk wilayah yang tidak memiliki peta geologi terperinci di Divisi Hazara Pakistan. Akurasi pemetaannya sangat tinggi, dengan algoritme CART melaporkan akurasi keseluruhan (OA) dan akurasi konsumen (CA) masing-masing sebesar 99,63% dan 99,69%.

• *Unsupervised Methods for Data Labeling*

Penelitian ini menunjukkan efektivitas metode tanpa pengawasan, seperti ekstraksi warna selektif dan fusi gambar aritmatika bitwise, dalam memperoleh label data yang disempurnakan untuk pelatihan dan validasi algoritma Machine Learning. Metode ini, yang diterapkan pada data ASTER, *Landsat-8*, dan *Sentinel-2*, menghasilkan anotasi data yang andal untuk klasifikasi selanjutnya.

• *Algorithm Performance*

Studi ini membandingkan kinerja berbagai algoritma Machine Learning, termasuk CART, *Random Forest*, *Naïve Bayes*, dan *Support Vector Machine* (SVM), untuk klasifikasi batu kapur biner. Hasilnya menunjukkan bahwa algoritma CART mengungguli algoritma lainnya, mencapai nilai akurasi dan koefisien Kappa yang unggul.

• *Impact of Hyperparameters and Data Sources*

Penelitian ini menyoroti dampak hyperparameter dan sumber data terhadap keakuratan pemetaan batu kapur. Khususnya, studi ini menemukan bahwa data *Landsat-8*, karena karakteristiknya seperti resolusi spasial dan tingkat vegetasi, memberikan hasil terbaik untuk klasifikasi batu kapur biner setelah melakukan penyesuaian hyperparameter.

• *Validation and Field Visits*

Untuk memastikan keakuratan hasil pemetaan, penelitian memvalidasi peta yang dihasilkan melalui kunjungan lapangan dan koordinat GPS. Proses validasi ini memberikan keyakinan terhadap keakuratan

- dan keandalan klasifikasi formasi batu kapur berbasis *Machine Learning* (Yao et al., 2022)
- *Potential Applications*
Pemetaan formasi batu kapur yang akurat memiliki potensi besar untuk diterapkan dalam industri konstruksi dan studi geologi. Klasifikasi jenis batuan yang terperinci dan andal menggunakan *algoritme Machine Learning* dapat berkontribusi pada pengambilan keputusan yang tepat di berbagai domain.
 - 4. Pendekatan / Metode terakhir mengevaluasi kegunaan unsur jejak dalam mengklasifikasikan batuan sumber zirkon.
 - Hasil Penelitian
Penelitian yang dilakukan dalam penelitian ini berfokus pada pengembangan dan penerapan metode baru untuk klasifikasi jenis batuan menggunakan teknik *Machine Learning*. Hasil yang diperoleh dari penelitian ini signifikan dan memberikan wawasan berharga mengenai potensi *Machine Learning* di bidang geosains, khususnya dalam karakterisasi jenis batuan di reservoir minyak.
 - *Machine Learning-Based Rock Classification*
Penerapan algoritma *Machine Learning*, termasuk peningkatan gradien cahaya, regresi logistik, dan hutan acak, menunjukkan hasil yang menjanjikan dalam mengklasifikasikan jenis batuan secara akurat berdasarkan berbagai parameter petrofisika. Model tersebut menunjukkan akurasi dan ketahanan yang tinggi dalam memprediksi jenis batuan, sehingga menunjukkan potensi *Machine Learning* dalam aplikasi geosains (Houshmand et al., 2022)
 - *Predictive Accuracy and Model Performance* (Malik et al., 2022)
Studi tersebut mengungkapkan bahwa model *Machine Learning* mencapai tingkat akurasi prediksi yang tinggi dalam mengklasifikasikan jenis batuan, melampaui metode tradisional dalam hal efisiensi dan presisi. Performa model dievaluasi menggunakan teknik validasi silang, yang menunjukkan hasil yang konsisten dan andal di berbagai kumpulan data.
 - *Integration of Petrophysical Parameters*

Classification of Rock Types...

- Selain itu, integrasi parameter petrofisika, seperti porositas, permeabilitas, dan radius tenggorokan pori, ke dalam model *Machine Learning* secara signifikan meningkatkan akurasi klasifikasi jenis batuan. Integrasi ini memungkinkan pemahaman komprehensif tentang hubungan antara berbagai sifat batuan, sehingga menghasilkan peningkatan kemampuan prediksi.
- *Explainable Machine Learning*
Penelitian ini menekankan pentingnya *Machine Learning* yang dapat dijelaskan, memberikan wawasan tentang fitur dan variabel mendasar yang berkontribusi terhadap klasifikasi jenis batuan. Transparansi ini sangat penting bagi ahli geosains dan insinyur reservoir untuk menafsirkan dan memvalidasi hasil, sehingga meningkatkan keandalan dan penerapan model yang dikembangkan.
 - *Implications for Oil and Gas Industry*
Temuan studi ini mempunyai implikasi besar bagi industri minyak dan gas, khususnya dalam karakterisasi reservoir dan evaluasi aset hidrokarbon. Klasifikasi jenis batuan yang akurat menggunakan *Machine Learning* dapat memfasilitasi proses pengambilan keputusan yang tepat, sehingga menghasilkan pengelolaan reservoir yang lebih optimal dan strategi pemulihan yang lebih baik.
 - *Future Research Directions*
Meskipun penelitian saat ini menunjukkan hasil yang menjanjikan, upaya penelitian di masa depan dapat berfokus pada perluasan penerapan klasifikasi batuan berbasis *Machine Learning* ke kumpulan data yang lebih besar dan beragam. Selain itu, dimasukkannya parameter petrofisika tambahan, seperti faktor sementasi, dapat lebih meningkatkan kemampuan prediksi model.

KESIMPULAN DAN SARAN

Penelitian ini mengungkap potensi besar *Machine Learning* dalam klasifikasi jenis batuan dengan berbagai pendekatan yang efektif. Metode seperti penggunaan kurva permeabilitas relatif, identifikasi retakan berdasarkan sinyal suara, analisis citra menggunakan *Deep Convolutional Neural Networks* (CNNs), dan evaluasi kegunaan unsur jejak telah menunjukkan keberhasilan yang signifikan. Algoritma-algoritma seperti *Decision Trees*, *Random Forest*, *SVM*, dan

Neural Networks mampu mencapai akurasi hingga 92%, dengan faktor-faktor seperti komposisi mineral, kepadatan, dan porositas batuan mempengaruhi hasil klasifikasi secara signifikan. Saran untuk penelitian mendepankan penggunaan algoritma ensemble untuk meningkatkan akurasi dan reliabilitas model, integrasi data yang lebih luas untuk meningkatkan ketahanan, serta fokus pada kejelasan interpretasi hasil untuk mendukung aplikasi praktis di industri. Peningkatan kualitas data eksperimen juga dianggap krusial dalam meningkatkan keandalan model *Machine Learning* dalam konteks karakterisasi reservoir dan evaluasi aset hidrokarbon

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah mendukung dalam hal ini beberapa penulis yang menjadi sumber jurnal kami dalam pembuatan jurnal dan juga pihak lain yang membiayai kebutuhan anggaran dalam penelitian ini

DAFTAR PUSTAKA

- Alfonso, C. E., Fournier, F., & Alcobia, V. (2021, September 15). A Machine Learning Methodology for Rock-Typing Using Relative Permeability Curves. *Day 1 Tue, September 21, 2021*. <https://doi.org/10.2118/205989-MS>
- Houshmand, N., GoodFellow, S., Esmaili, K., & Ordóñez Calderón, J. C. (2022). Rock type classification based on petrophysical, geochemical, and core imaging data using machine and deep learning techniques. *Applied Computing and Geosciences*, 16, 100104. <https://doi.org/10.1016/j.acags.2022.100104>
- Hussain, M., Liu, S., Ashraf, U., Ali, M., Hussain, W., Ali, N., & Anees, A. (2022). Application of Machine Learning for Lithofacies Prediction and Cluster Analysis Approach to Identify Rock Type. *Energies*, 15(12), 4501. <https://doi.org/10.3390/en15124501>
- Itano, K., & Sawada, H. (2024). Revisiting the Geochemical Classification of Zircon Source Rocks Using a Machine Learning Approach. *Mathematical Geosciences*. <https://doi.org/10.1007/s11004-023-10128-z>
- Jiang, H., Jiang, J., & Su, G. (2023). Rock Crack Types Identification by Machine Learning on the Sound Signal. *Applied Sciences*, 13(13), 7654. <https://doi.org/10.3390/app13137654>
- K, Sivachandra., R, Kumudham., Kumar, P. S., V, Rajendran., & Lakshmi, G. R. J. (2023). Rock / Mine Classification Using Supervised Machine Learning Algorithms. 2023 *International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, 177–184. <https://doi.org/10.1109/IITCEES7236.2023.10091031>
- Khan, M. F. A., Muhammad, K., Bashir, S., Ud Din, S., & Hanif, M. (2021). Mapping Allochemical Limestone Formations in Hazara, Pakistan Using Google Cloud Architecture: Application of Machine-Learning Algorithms on Multispectral Data. *ISPRS International Journal of Geo-Information*, 10(2), 58. <https://doi.org/10.3390/ijgi10020058>
- Malik, O. A., Puasa, I., & Lai, D. T. C. (2022). Segmentation for Multi-Rock Types on Digital Outcrop Photographs Using Deep Learning Techniques. *Sensors*, 22(21), 8086. <https://doi.org/10.3390/s22218086>
- Mohammadian, E., Kheirollahi, M., Liu, B., Ostadhassan, M., & Sabet, M. (2022). A case study of petrophysical rock typing and permeability prediction using machine learning in a heterogenous carbonate reservoir in Iran. *Scientific Reports*, 12(1), 4505. <https://doi.org/10.1038/s41598-022-08575-5>
- Muoghalu, A. I. (2022, September 26). A Machine Learning Approach to Rock Typing with Relative Permeability Curves Using Kmeans Clustering Algorithm. *Day 1 Mon, October 03, 2022*. <https://doi.org/10.2118/212383-STU>
- Ran, X., Xue, L., Zhang, Y., Liu, Z., Sang, X., & He, J. (2019). Rock Classification from Field Image Patches Analyzed Using a Deep Convolutional Neural Network. *Mathematics*, 7(8), 755. <https://doi.org/10.3390/math7080755>
- Saleh, A., & Aljawad, M. (2023). Permeability Estimation for Upper Shale Member in Southern Iraqi Oil Field Using Machine learning and Hydraulic Flow Units Methods. *Iraqi Geological Journal*,

- 56(2B), 90–101.
<https://doi.org/10.46717/igj.56.2B.7ms-2023-8-16>
- Stehlíková, B., Bogdanovská, G., Flegner, P., Frančáková, R., & Drančák, L. (2023). The Proposal of a Method for Rock Classification Using a Vibration Signal Propagated during the Rotary Drilling Process. *Applied Sciences*, 13(20), 11315.
<https://doi.org/10.3390/app132011315>
- Xing, Y., Yang, H., & Yu, W. (2023). An Approach for the Classification of Rock Types Using Machine Learning of Core and Log Data. *Sustainability*, 15(11), 8868.
<https://doi.org/10.3390/su15118868>
- Yao, G., Zhou, X., Ke, C., Drolma, L., & Li, H. (2022). The Potential of Sentinel-1A Data for Identification of Debris-Covered Alpine Glacier Based on Machine Learning Approach. *Remote Sensing*, 14(9), 1980.
<https://doi.org/10.3390/rs14091980>

Marojahan Benedict Efrata, Widya Spalanzani, Sekar Ayu Citrowati

Submitted: **11/07/2024**; Revised: **01/08/2024**; Accepted: **01/09/2024**; Published: **31/10/2024**