

**STUDI KASUS PREDIKSI DISTRIBUSI FASIES VERTIKAL DI
SUMUR RESERVOIR BATU PASIR MENGGUNAKAN
ALGORITMA RANDOM FOREST**

TUGAS AKHIR

Diajukan Guna Penyusunan Tugas Akhir Program Studi Teknik Perminyakan

Oleh

FAJAR RAMADHAN

163210672



**PROGRAM STUDI TEKNIK PERMINYAKAN
UNIVERSITAS ISLAM RIAU
PEKANBARU
2021**

**STUDI KASUS PREDIKSI DISTRIBUSI FASIES VERTIKAL DI
SUMUR RESERVOIR BATU PASIR MENGGUNAKAN
ALGORITMA RANDOM FOREST**

TUGAS AKHIR

Diajukan Guna Penyusunan Tugas Akhir Program Studi Teknik Perminyakan

Oleh

FAJAR RAMADHAN

163210672



PROGRAM STUDI TEKNIK PERMINYAKAN

UNIVERSITAS ISLAM RIAU

PEKANBARU

2021

HALAMAN PENGESAHAN

Tugas Akhir ini disusun oleh :

Nama : Fajar Ramadhan

NPM : 163210672

Program Studi : Teknik Perminyakan

Judul Skripsi : Studi Kasus Prediksi Distribusi Fasies Vertikal di Sumur Reservoir Batu Pasir Menggunakan Algoritma Random Forest.

Telah berhasil dipertahankan di depan Dewan Penguji dan diterima sebagai salah satu syarat guna memperoleh Gelar Sarjana Teknik pada Program Studi Teknik Perminyakan, Fakultas Teknik, Universitas Islam Riau

DEWAN PENGUJI

Pembimbing : Fiki Hidayat, ST., M.Eng (.....)

Penguji I : Novia Rita, S.T., M.T (.....)

Penguji II : Novrianti, ST., MT (.....)

Ditetapkan di : Pekanbaru

Tanggal : 25 Februari 2022

Disahkan Oleh :

**KETUA PROGRAM STUDI
TEKNIK PERMINYAKAN**

NOVIA RITA, ST., MT

PERNYATAAN KEASLIAN TUGAS AKHIR

Dengan ini saya menyatakan bahwa tugas akhir ini merupakan karya saya sendiri dan semua sumber yang tercantum didalamnya baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar sesuai ketentuan. Jika terdapat unsur penipuan atau pemalsuan data maka saya bersedia dicabut gelar yang telah saya peroleh.

Pekanbaru, 25 Februari 2022

Fajar Ramadhan
NPM.163210672



KATA PENGANTAR

Rasa syukur saya ucapkan kepada Allah SWT berkat rahmat dan karunia-Nya sehingga penulis mampu untuk menyelesaikan tugas akhir ini. Penulisan tugas akhir ini merupakan syarat terakhir untuk mendapatkan gelar Sarjana Teknik, Program Studi Perminyakan, Universitas Islam Riau. Banyak pihak yang telah membantu penulis dalam bentuk pengetahuan, materil dan moral dalam penyelesaian tugas akhir ini. Oleh sebab itu penulis ingin mengucapkan terimakasih sebesar-besarnya kepada:

1. Kedua orang tua beserta keluarga penulis yang telah memberikan dukungan moral, materil, dan do'a yang selalu diberikan.
2. Bapak Fiki Hidayat, ST., M.Engselaku Dosen Pembimbing Tugas Akhir, yang telah menyediakan waktu, tenaga dan pikiran untuk memberikan masukan dalam penyusunan tugas akhir ini.
3. Ibu Novia Rita, ST., MT selaku Ketua Prodi sertadosen-dosen yang telah banyak membantu terkait perkuliahan, ilmu pengetahuan dan hal lain yang tidak dapat saya sebutkan satu per satu.
4. Ketua dan Sekretaris Program Studi Teknik Perminyakan serta dosen-dosen lainnya yang belum bisa saya sebutkan satu persatu dalam tugas akhir ini, yang telah membantu dan mengajari penulis selama masa perkuliahan.
5. Sahabat dan yang terkasih Mhd. Tengku Sofyan Atsauri, Ade Novian Permana, Supandi Lahagu, Miranti Alivia, Hanifa, Perti Insani, senior dan teman-teman kampus yang tidak bisa deisebutkan satu persatu.

Harapan saya semoga AllahSWT memberikan balasan atas segala kebaikan kepada semua pihak yang telah membantu. Semoga skripsi ini membawa manfaat bagi pengembangan ilmu pengetahuan.

Pekanbaru, 25 Februari 2022

Fajar Ramadhan

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
PERNYATAAN KEASLIAN TUGAS AKHIR.....	iii
KATA PENGANTAR	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	vii
DAFTAR SINGKATAN	viii
DAFTAR SIMBOL.....	ix
ABSTRAK	x
ABSTRACT.....	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian.....	3
1.3 Manfaat Penelitian.....	3
1.4 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 State Of The Art	5
BAB III METODOLOGI PENELITIAN.....	16
3.1 Uraian Metode Penelitian.....	16
3.2 Lokasi Penelitian	17
3.3 Data Penelitian	18
3.4 Flow Chart	22
BAB IV HASIL DAN PEMBAHASAN	23
4.1 Penentuan Jumlah Kelas Fasies.....	23
4.2 Analisis Kelas Fasies	28
4.3 Pembentukan Model Prediksi.....	30
4.4 Hasil Prediksi Fasies Menggunakan Random Forest	31
BAB V PENUTUPAN.....	38
5.1 Kesimpulan.....	38
5.2 Saran	38
DAFTAR PUSTAKA	39

DAFTAR GAMBAR

Gambar 3. 1 Data sumur b5	18
Gambar 3. 2 Data sumur b10	19
Gambar 3. 3 Data sumur b30	19
Gambar 3. 4 Data sumur b36	20
Gambar 3. 5 Data sumur b38	20
Gambar 3. 6 Data sumur b52	21
Gambar 3. 7 Data sumur b58	21
Gambar 3. 1 Flow Chart.....	22
Gambar 4. 1 Grafik Metode Elbow pada sumur b5	23
Gambar 4. 2 Grafik Metode Elbow pada sumur b10	23
Gambar 4. 3 Grafik Metode Elbow pada sumur b30	24
Gambar 4. 4 Grafik Metode Elbow pada sumur b36	24
Gambar 4. 5 Grafik Metode Elbow pada sumur b38	24
Gambar 4. 6 Grafik Metode Elbow pada sumur b52	25
Gambar 4. 7 Grafik Metode Elbow pada sumur b58	25
Gambar 4. 8 Hasil Clustering Fasies pada sumur b 10	26
Gambar 4. 9 Sentroid K pada setiap kelas fasies pada sumur b10.....	26
Gambar 4. 10 Hasil klasifikasi sumur b10.....	27
Gambar 4. 10 Deskripsi running model pertama Randomized Search CV	30
Gambar 4. 11 Deskripsi running model kedua Randomized Search CV	31
Gambar 4. 13 Hasil Prediksi Fasies Sumur b5.....	32
Gambar 4. 14 Hasil Prediksi Fasies Sumur b30.....	33
Gambar 4. 15 Hasil Prediksi Fasies Sumur b36.....	34
Gambar 4. 16 Hasil Prediksi Fasies Sumur b38.....	35
Gambar 4. 17 Hasil Prediksi Fasies Sumur b52.....	36
Gambar 4. 18 Hasil Prediksi Fasies Sumur b58.....	37

DAFTAR TABEL

Tabel 3. 1 Jadwal Penelitian.....	21
-----------------------------------	----



DAFTAR SINGKATAN



ML	Machine Learning
RF	Random Forest
GR	Gamma Ray
SP	Spontaneous Potential
LLD	Laterolog Deep
LLS	Shallow Laterolog
MSFL	Micro Spherically Focused Log
API	American Petroleum Institute
mV	mili Volt
ft	feet
g/cc	gram/cubic centimeter
ohmm	ohm meter
v/v	volt/volt
Rw	Resistivity Water
Rmf	Resitivity mud filtrate
PU	Porosity Unit

DAFTAR SIMBOL

>
<
 Σ

Lebih besar

Lebih kecil

Sigma



Dokumen ini adalah Arsip Miik :

Perpustakaan Universitas Islam Riau

STUDI KASUS PREDIKSI DISTRIBUSI FASIES VERTIKAL DI SUMUR RESERVOIR BATU PASIR MENGGUNAKAN ALGORITMA RANDOM FOREST

FAJAR RAMADHAN

163210672

ABSTRAK

Pengetahuan tentang fasies merupakan hal yang penting dalam kegiatan karakterisasi *reservoir*. Informasi tentang fasies dapat diketahui melalui data inti dan pembacaan log sumur. Data inti dapat memberikan gambaran litofasies batuan dan pembacaan log menunjukkan elektrofases batuan di berbagai kedalaman *reservoir*. Telah banyak dilakukan penelitian dalam 20 tahun terakhir untuk melakukan penentuan distribusi fasies reservoir melalui metode klasifikasi elektrofases. Metode klasifikasi elektrofases semakin diminati saat mulai berkembangnya penggunaan ML dalam kegiatan klasifikasi data. Penggunaan ML dalam klasifikasi data memungkinkan untuk melakukan prediksi menggunakan parameter atau atribut dalam kegiatan *training* dan *testing*. Penelitian ini melakukan prediksi fasies pada reservoir batu pasir pada sebuah lapangan minyak menggunakan data berbagai jenis data log seperti GR, SP, LLD, LLS, Densitas dan Neutron yang diperoleh dari 7 sumur. Salah satu sumur akan dipilih untuk menjadi sumur acuan yang berfungsi sebagai pusat penentuan fasies dan pembentukan model prediksi. Pendekatan dua langkah diterapkan dalam penelitian ini. Langkah pertama melakukan penentuan jumlah fasies menggunakan algoritma K-Means. Langkah selanjutnya pembentukan model prediksi melalui klasifikasi elektrofases menggunakan algoritma RF. Pembentukan model pertama dilakukan menggunakan empat atribut log yaitu log GR, SP, LLS dan Densitas. Pembentukan model kedua dilakukan dengan empat atribut log yang berbeda yaitu log LLS, LLD, Densitas dan Neutron. Kedua langkah penelitian ini akan dilakukan menggunakan bahasa pemrograman Python. Hasil penelitian ini menunjukkan bahwa fasies *reservoir* dapat diprediksi melalui analisis clustering. Pendekatan K-means memprediksi klasifikasi fasies optimal menjadi tiga jenis fasies. Pembentukan model prediksi menggunakan algoritma RF pada kedua model memiliki tingkat akurasi yang tinggi yaitu 95% pada model pertama dan 97% pada model kedua, sehingga penerapan metode klasifikasi ini menjadi lebih efisien dibandingkan menggunakan metode konvensional dalam melakukan karakterisasi *reservoir*.

Kata Kunci: Prediksi, klasifikasi, Elektrofases, Fasies, Algoritma, RF, K-Means

STUDY OF VERTICAL FACIES DISTRIBUTION PREDICTION IN SANDSTONE RESERVOIR WELL USING RANDOM FOREST ALGORITHM

FAJAR RAMADHAN
16321072

ABSTRACT

Knowledge of facies is important in reservoir characterization activities. Information about facies can be obtained through core data and reading well logs. Core data can provide an overview of rock lithofacies and reading logs shows rock electrofacies at various reservoir depths. In the last 20 years, many studies have been carried out to distribute reservoir facies through the electrofacies classification method. Electrofacies classification methods are increasingly in demand when the use of ML in data classification activities begins to develop. The use of ML in data classification makes it possible to make predictions using parameters or attributes in training and testing activities. This study predicts facies in a sandstone reservoir in an oil field using various types of log data such as GR, SP, LLD, LLS, Density and Neutron obtained from 7 wells. One of the wells will be selected to be a reference well that will serve as a center for facies development and the formation of predictive models. A two-step approach was applied in this study. The first step is to do the number of facies using yahoo K-Means. The next step is the formation of a predictive model through electrofacies classification using the RF algorithm. The first model is formed using four log attributes, namely GR, SP, LLS and Densiras logs. The second model is formed with four different attributes, namely LLS log, LLD, Density and Neutron. The two steps of this research will be carried out using the Python programming language. The results of this study indicate that reservoir facies can be predicted through clustering analysis. The K-means approach predicts the optimal facies classification into three types of facies. The formation of the prediction model using the RF algorithm in both models has a high level of accuracy that is 95% in the first model and 97% in the second model so that the application of this classification becomes more efficient than using conventional methods to characterize reservoirs.

Keywords : Prediction, classification, Electrofacies, Facies, Algorithm, RF, K-Means

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis dan klasifikasi fasies telah menjadi prosedur penting diantara ahli geologi dan perminyakan untuk beberapa dekade terakhir. Fasies didefinisikan sebagai unit batuan dengan ciri-ciri litologi yang khas, termasuk komposisi, ukuran butir, karakteristik lapisan, dan struktur sedimen (Middleton, 1973)(Busch, Fortney, Berry, & others, 1987)(Rabaute, 1998)(Galloway, 1989). Catatan fasies sedimen yang dapat digunakan untuk menilai kualitas unit *reservoir* (Serra, Sulpice, & others, 1975) (Delfiner, Peyret, Serra, & others, 1987)(Van Wagoner, Mitchum, Campion, & Rahmanian, 1990)(Nichols, 1993)(Rider, 1990)(Posamentier & Allen, 1999)(Catuneanu, 2006).

Sejak diperkenalkan oleh Serra & Abbot pada awal 1980-an, penggunaan elektrofases dalam karakterisasi *reservoir* telah menarik minat, dengan teknik berbeda yang ditawarkan dalam analisis dan pemodelan fasies masukan dari log sumur (Serra & Abbott, 1982)(Serra, 1986)(Moline & Bahr, 1995)(Euzen, Delamaide, Feuchtwanger, Kingsmith, & others, 2010). Dalam aplikasi analisis log, elektrofases biasanya digunakan sebagai indikator litologi dan lingkungan pengendapan (Wolf, Pelissier-Combescure, & others, 1982)(Anxionnaz, Delfiner, & Delhomme, 1990). Menggabungkan elektrofases, berdasarkan karakteristik fisik dan sifat petrofisika, dengan pengukuran inti mengarah ke kelas yang dapat mewakili fasies geologi atau rentang sifat petrofisika esensial. Kemajuan yang cukup besar telah dibuat dalam mengembangkan berbagai metode klasifikasi otomatis untuk industri(Lim, Kang, Kim, & others, 1997)(Qi & Carr, 2006)(Skalinski, Gottlib-Zeh, Moss, & others, 2006)(Hong Tang, Meddaugh, Toomey, & others, 2011).

Memprediksi distribusi fasies yang tepat dalam sumur merupakan langkah penting untuk meningkatkan hubungan antara porositas dan permeabilitas, untuk secara akurat memperkirakan properti dalam interval non-inti, dan untuk mengidentifikasi zona permeabel tinggi dan rendah dan kemudian untuk melestarikan heterogenitas(Dorfman, Newey, & Coates, 1990)(Doveton &

Prensky, 1992)(Buchebe, Evans, & others, 1994)(Lim et al., 1997)(Moline & Bahr, 1995)(Lee, Datta-Gupta, & others, 1999)(Watanabe, Suzuki, Eto, Samizo, & others, 1999). Metode konvensional konsisten dengan menetapkan kesesuaian secara manual oleh penerjemah manusia dan merupakan proses yang sangat membosankan dan memakan waktu. Oleh karena itu, beberapa pendekatan alternatif untuk masalah klasifikasi fasies dari data sumur telah diusulkan. Karya pertama didasarkan pada metode statistik multivariat klasik(Wolf et al., 1982)(Busch et al., 1987) Kemudian, Wolf beserta rekannya di tahun 1982 mengusulkan penggunaan jaringan saraf untuk klasifikasi batuan(Baldwin, Bateman, Wheatley, & others, 1990)(Rogers, Fang, Karr, & Stanley, 1992).

Sementara itu, terutama dalam beberapa tahun terakhir, didorong oleh pertumbuhan *big data* dan dengan meningkatnya daya komputasi, telah terjadi peningkatan minat baru pada teknik pembelajaran mesin. Baru-baru ini metodologi ini semakin dieksplorasi, untuk aplikasi yang berbeda, juga oleh komunitas geofisika(Smith & Treitel, 2010)(Zhang, Frogner, Araya-Polo, & Hohl, 2014)(Zhao, Jayaram, Roy, & Marfurt, 2015)(Kobrunov & Priezzhev, 2016).

Machine learning (ML) telah menjadi salah satu alat paling ampuh selama dekade terakhir di antara banyak industri. Tidak mengherankan, pembelajaran mesin juga telah menjangkau industri minyak & gas dengan banyak aplikasi dari pengembangan geofisika seperti peningkatan deteksi kesalahan, identifikasi klasifikasi fasies, dan bahkan pemilihan atribut untuk pemahaman yang lebih baik tentang data seismik(Silva, 2020)(Xu, Misra, Srinivasan, Ma, & others, 2019).

Beberapa algoritma telah dikembangkan untuk mengotomatiskan klasifikasi fasies seismik. Secara khusus, teknik pembelajaran mesin memberikan alur kerja yang efektif dalam klasifikasi dan prediksi fasies. Teknik pembelajaran mesin sering diklasifikasikan tergantung pada apakah teknik tersebut tidak diawasi atau diawasi. Pembelajaran tanpa pengawasan menganalisis data masukan berdasarkan distribusi atau struktur yang mendasari data masukan. *K-means clustering*, *Principal Component Analysis* (PCA), *Self Organizing Maps* (SOM), dan *Linear Discriminant Analysis* (LDA) diperkenalkan untuk mengelompokkan dan memetakan data seismik dengan properti serupa dalam bentuk gelombang atau informasi atribut(Kim, Hardisty, Torres, & Marfurt, 2018).

Banyak algoritma klasifikasi yang diawasi telah diadopsi untuk pemodelan dan prediksi fasies menggunakan data *logging* sumur seperti *Logistic Boosting Regression* (LogitBoost), *Neural Network* (NN), *Nave Bayes Classifier* (NBC), *Multinomial Logistic Regression* (MLR), *Tree-Based Classification Model* (RPART), *Kernel Support Vector Machine* (KSVM), dan *Random Forest*. Akurasi algoritma ini bergantung pada pembentukan model dan jumlah serta jenis log sumur yang digunakan sebagai parameter masukan (Wong, Jian, & Taggart, 1995) (Lee et al., 1999) (Avseth, Mukerji, & others, 2002) (Hong Tang, White, Zeng, Gani, & Bhattacharya, 2004) (H Tang, 2008) (Adoghe, Aniekwe, Nwosu, & others, 2011) (Bagheri & Riahi, 2015) (Al-Mudhafer, 2014) (W. J. Al-Mudhafar & others, 2016) (W. J. Al-Mudhafar, 2017) (Puskarczyk, 2019) (Silva, 2020) (W. Al-Mudhafar, 2020).

Berdasarkan latar belakang diatas untuk memprediksi distribusi fasies vertikal sumur peneliti melakukan pendekatan dua langkah. Pertama peneliti melakukan pengelompokan data log sumur kedalam jenis elektrofases. Pengelompokan ini dilakukan menggunakan algoritma klasifikasi yang tidak diawasi yaitu K-Means. K-Means akan melakukan prediksi jumlah kelompok fasies berdasarkan kedalaman sumur. Kemudian pada langkah kedua dari data kelompok fasies akan digunakan sebagai model pada algoritma klasifikasi yang diawasi yaitu Random Forest untuk memprediksi sumur lain yang memiliki kehilangan data. Atribut yang akan digunakan berupa data *logging* diantaranya GR, SP, LLD, LLS, MSFL, Densitas dan Neutron. Penelitian ini dilakukan menggunakan aplikasi bahasa pemrograman Python.

1.2 Tujuan Penelitian

Tujuan dilakukannya penelitian ini adalah menentukan model statistik yang dapat diandalkan untuk memprediksi dan memodelkan fasies vertikal di sumur uji dan kemudian menggunakan model ini di sumur target untuk memprediksi fasies vertikal saat data inti tidak tersedia.

1.3 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah:

- a. Studi ini dapat dijadikan sebagai dasar untuk meningkatkan hubungan porositas dan permeabilitas.
- b. Studi ini dapat dijadikan sebagai dasar untuk memperkirakan properti *reservoir* non-inti.
- c. Studi ini dapat dijadikan sebagai dasar untuk mengidentifikasi zona permabel.
- d. Studi ini dapat dijadikan dasar untuk melestarikan heterogenitas *reservoir*.
- e. Studi ini juga dapat dijadikan publikasi berskala nasional dan internasional.

1.4 Batasan Masalah

Penelitian ini memiliki beberapa fokus utama untuk membatasi masalah dalam penelitian. Fokus utama peneliti untuk membatasi masalah yaitu:

- a. Penelitian ini hanya menggunakan bahasa pemrograman Python.
- b. Penelitian ini hanya dilakukan pada *reservoirsandstone* yang bersifat heterogen.
- c. Penelitian ini hanya menggunakan algoritma K-Means untuk pengelompokan data.
- d. Penelitian ini hanya menggunakan algoritma Random Forest untuk prediksi dan klasifikasi data.
- e. Penelitian ini hanya menggunakan data *logging* dan data *core*.
- f. Penelitian ini tidak membahas pertimbangan ekonomi.

BAB II

TINJAUAN PUSTAKA

2.1 State Of The Art

Mengetahui jenis fasies pada *reservoir* merupakan langkah penting untuk karakterisasi *reservoir*. Pengetahuan tentang jenis fasies memudahkan insinyur perminyakan untuk melakukan analisis tentang nilai porositas dan permeabilitas serta hubungan antara keduanya (Wong et al., 1995) (H Tang, 2008).

Banyak metode yang telah dilakukan pada beberapa dekade yang lalu untuk mengidentifikasi dan memprediksi jenis fasies yang ada di *reservoir*. Salah satu teknik yang sering digunakan adalah metode klasifikasi elektrofases. Data yang akan digunakan untuk klasifikasi elektrofases adalah data *welllogging* yang dikombinasikan dengan data *core* (Buche et al., 1994) (Serra & Abbott, 1982) (Wolf et al., 1982).

Defliner beserta rekannya (1987) melakukan penelitian penentuan litologi sumur secara otomatis. Penentuan litologi dilakukan dengan cara menggabungkan pengukuran *wireline modern* menggunakan basis data litofases untuk deskripsi litologi secara otomatis kemudian dibandingkan dengan deskripsi *core* dan log sumur. Litofases berbasis data dari pengetahuan petrografi diterjemahkan dalam bentuk respon log kemudian dengan fungsi diskriminan (aturan keputusan bayesian) dihasilkan kelompok fasies utama yaitu batu pasir, batu gamping, dolomit, serpih, batubara dan evaporit (Defliner et al., 1987).

Isimbabi dan lainnya (2018) melakukan Identifikasi elektrofases secara otomatis menggunakan metode pemodelan 3D dari desain aplikasi P-facies. Pemodelan ini pada dasarnya adalah membuat representasi atau replika dari sesuatu yang ada. Aplikasi berbasis excel yang dapat mengenali bentuk log dan dengan demikian, dapat mengidentifikasi hingga lima fasies yang berbeda: *Channel Sands*, *Upper Shoreface*, *Lower Shoreface*, *Heteroliths* and *Shale* di lingkungan laut dangkal. Data yang akan digunakan untuk identifikasi fasies adalah log sinar gama. Teknik ini memiliki keterbatasan yang melekat pada elektrofases, dan membutuhkan QAQC yang tepat oleh pengguna sebelum

membangun model fasies yang didasarkan pada hasilnya, terutama karena kemungkinan artefak interpretasi (Isimbabi & others, 2018).

Setahun kemudian Gkortsas dan Liang (2019) juga melakukan klasifikasi elektrofases secara otomatis menggunakan metode klasifikasi *unsupervised* model campuran Gaussian (GMM) untuk pengelompokan data otomatis. Pengembangan alur kerja klasifikasi elektrofases tanpa pengawasan yang menggunakan input berbagai log dan memberikan pengelompokan yang sesuai. Log tersebut adalah sinar *gamma* (GR), resistivitas (ILD_log10), efek fotolistrik (PE), perbedaan porositas kerapatan neutron (DeltaPHI) dan porositas kerapatan neutron rata-rata (PHIND). Variabel pembatas geologi adalah indikator *nonmarine / marine* (NM_M), yaitu bendera biner yang diatribusikan oleh para ahli untuk membedakan antara fasies *marine* dan *nonmarine* setelah pemeriksaan data) dan posisi relatifnya. Dari pengukuran 5 log kabel sebagai input jumlah *cluster* yang diperoleh adalah 9. Setelah diperoleh 9 kluster optimal, kemudian dilakukan pemetaan ulang pada 9 *cluster* menjadi 2 *cluster* untuk di uji pada indikator *nonmarine / marine*. Indikator *nonmarine / marine* akan berfungsi sebagai kriteria untuk menguji keakuratan hasil *clustering*. Hasil yang diperoleh adalah *clustering* remapping memiliki akurasi 85% (Gkortsas & Liang, 2019).

H Tang (2008) melakukan penelitian untuk peningkatan klasifikasi fasies *reservoir* menggunakan metode klasifikasi *Probabilistic Neural Network* (PNN) dan fungsi basis radial (RBF), yang merupakan fungsi tipe Gaussian simetris. Set pelatihan dipilih secara acak dari setiap zona dari Sumur M303 dan M307. Data inti digunakan untuk memvalidasi kemampuan prediksi. Data log yang digunakan untuk prediksi fasies meliputi tujuh atribut yaitu log GR, SP, DT, NPHI, RHOB, PE, dan RT. Prediksi fasies dan inti memiliki kecocokan yang baik pada sumur M303, hasil yang ditunjukkan oleh prediksi PNN memperoleh akurasi 50% pada set 1/8 jumlah sampel. Sedangkan pada pengujian sumur M307 menghasilkan akurasi 75% yang lebih baik dengan menggunakan 1/4 pada set sampel. Perbedaan PE dan RHOB-NPHI merupakan indikator yang baik untuk keberadaan batugamping dan dolomit di daerah penelitian. Fasies yang berbeda memiliki respons log yang berbeda. Jenis log dari berbagai fasies digunakan untuk mengidentifikasi fasies dan menghapus penciran dari set pelatihan (H Tang, 2008).

Peningkatan model elektrofases kemudian dilakukan lagi oleh Adoghe et al., (2011). Adoghe dkk menggunakan PCA klasifikasi *unsupervised* hirarki untuk menghasilkan variable eigen kemudian menggunakan *Indexation phase self oragnizing map* (IPSON) yang berdasarkan pada algoritma jaringan saraf untuk klasifikasi *supervised* otomatis. Untuk klasifikasi *unsupervised* empat log digunakan (*Gamma-ray*, *Density & Neutron porosity* dan *Compressional slowness*). Dengan mempertimbangkan sebagian besar peningkatan model elektrofases di zona yang sama dengan fasies inti, model elektrofases yang diturunkan dari variabel eigen secara jelas memberikan kecocokan yang sama atau lebih baik dengan fasies inti (meskipun dengan lebih detail) dari pada model elektrofases yang didasarkan langsung pada log sumur. Teknik penggunaan variabel eigen (komponen utama) yang berasal dari log sumur untuk pemodelan elektrofases telah terbukti menjadi metode yang kuat dan konsisten secara internal dalam menciptakan elektrofases, di mana fasies inti tidak tersedia (klasifikasi *unsupervised*), selain memberikan sebuah prediksi elektrofases yang lebih baik di mana fasies inti tersedia sebagai kontrol (klasifikasi *supervised*)(Adoghe et al., 2011).

Teh, Willhite, Doveton, dan rekan lainnya (2012) melakukan peningkatan karakterisasi reservoir menggunakan klasifikasi petrofisik dalam elektrofases. Metodologi untuk prediksi permeabilitas di lapangan Ogallah di Kansas barat, yang terdiri dari *reservoir* karbonat dan batupasir. Langkah pertama menggunakan karakterisasi elektrofases untuk mengelompokkan log sumur dan mengidentifikasi kluster dengan deskripsi litofases inti. Pengelompokan dilakukan menggunakan PCA, MCA, discriminant analysis, dan ACE berdasarkan data log yaitu sinar *gamma* (GR), log lateral (Rt) dan log sensitivitas mikro (Rxo). Langkah kedua kemudian mengidentifikasi pengklasifikasi petrofisika RFN pada *reservoir* karbonat dan FZI untuk batu pasir, yang paling mewakili karakteristik aliran batuan dan mengembangkan korelasi log sumur pengklasifikasi menggunakan analisis regresi non-parametrik. Permeabilitas dihitung dari pengklasifikasi petrofisika, menggunakan nilai porositas. Pengklasifikasi RFN dan FZI diprediksi dalam setiap kelompok elektrofases dari data log sumur untuk dolomit dan batupasir. Permeabilitas dihitung menggunakan

pengklasifikasi petrofisika. Pendekatan ini divalidasi dengan perbandingan permeabilitas yang diprediksi dengan permeabilitas inti untuk 18 sumur inti. Dalam kedua kasus tersebut, metodologi memprediksi lonjakan dan penurunan dengan cermat. Prediksi menunjukkan *goodness of fit* yang memuaskan dibandingkan dengan estimasi semi-log sederhana. Metode tersebut memberikan cara untuk memprediksi permeabilitas pada sumur yang tidak memiliki analisis inti (Teh, Willhite, Doveton, & others, 2012).

W. Al-Mudhafar (2014) melakukan penelitian tentang Regresi Logistik Multinomial untuk Estimasi Bayesian pada Pemodelan Fasies Vertikal pada *Reservoir* Batupasir Heterogen. Penelitian ini melakukan estimasi fasies menggunakan Multinomial Logistic Regresi (MLR) sehubungan dengan log sumur dan data inti pada sumur di West Africa *Sandstone Reservoir*. teorema Bayes telah digunakan untuk mendapatkan probabilitas yang diprediksi (posterior) dari fasies dengan mempertimbangkan distribusi beta fasies itu sendiri sebagai pengetahuan sebelumnya. Seluruh variabel independen adalah log kaliper (CCL), log induksi dalam, log induksi sedang, sinar *gamma*, porositas neutron, porositas inti, resistivitas dalam, resistivitas sedang, potensial spontan (SP), kepadatan & kepadatan terkoreksi, selain permeabilitas inti. Prediksi probabilitas untuk distribusi fasies vertikal telah diestimasi dari regresi logistik multinomial, pengetahuan sebelumnya diperoleh dari Distribusi Beta. Akibatnya distribusi probabilitas akhir (posterior) telah dihitung sesuai dengan teorema Bayes. Distribusi beta yang merepresentasikan kepadatan tiap fasies atau jumlah total tiap interval fasies dibagi jumlah total fasies untuk semua jenis: 0.1355, 0.0098, 0.0672, 0.1331, 0.0110, 0.2247, 0.2100, 0.2088. dengan kata lain dapat terlihat terdapat 8 jenis fasies berdasarkan densitas setiap fasies yang diidentifikasi (Al-Mudhafer, 2014).

Tiga tahun kemudian W. J. Al-Mudhafar (2017) mengintegrasikan mesin kernel support vector untuk klasifikasi fasies batuan. Distribusi probabilitas posterior diskrit dan bersyarat dari litofasies sumur dimodelkan dan diprediksi melalui *Kernel Support Vector Machines* (KSVM) sebagai fungsi interpretasi log sumur di dalam sumur di Anggota Batupasir Atas formasi Zubair di Ladang Minyak Rumaila Selatan, berlokasi di Irak. Data log meliputi porositas neutron,

saturasi air, dan volume serpih. Faktor respon multinomial adalah urutan litofasies vertikal terukur yang sebagian besar terdiri dari pasir, serpih, dan pasir serpih. Litofasies yang diprediksi divalidasi dengan menghitung persentase total yang benar dari matriks jumlah fasies yang diprediksi, yang diperkirakan dengan KSVM. Pemisahan nonlinier komponen yang ditangani oleh KSVM menghasilkan tingkat akurasi yang tinggi dari prediksi litofasies dan mencapai 99,55% dari total persen yang benar. Setelah menggambarkan distribusi vertikal pasir, serpih, dan pasir serpih, terlihat bahwa prediksi KSVM memiliki kesesuaian antara nilai posterior pasir dengan catatan porositas neutron yang tinggi dan interval volume serpih yang rendah (W. J. Al-Mudhafar, 2017).

Bestagini, Lipari, & Tubaro (2017) melakukan penelitian pendekatan *machine learning* untuk klasifikasi fasies menggunakan *well log*. Metode dalam klasifikasi elektroafisies pada penelitian ini menggunakan *supervised* mesin learning berupa random forest. Sebelum melakukan klasifikasi untuk melakukan pemisahan kelas agar lebih sederhana penelitian ini menggunakan pendekatan augmentasi fitur. Fitur yang digunakan terdiri dari 7 atribut, yaitu *Gamma-ray*, Resistivitas, *Photoelectric effect*, Pengukuran porositas kerapatan neutron, porositas kerapatan rata-rata, Indikator *nonmarine/marine*, dan Posisi relatif. Selanjutnya akan di karakterisasi kedalam 9 fasies yaitu Batupasir *Nonmarine* (SS), Batulanau Kasar *Nonmarine* (CSiS), Batuan lumpur halus *nonmarine* (FSiS), Batulanau dan serpih laut (SiSh), Batulumpur (MS), *Wackestone* (WS), Dolomit (D), *Packstone-grainstone* (PS) dan *Filoid-alga baf eston* (BS). Algoritme ini mencapai skor dengan rata-rata 0,61 selama 10 lipatan yang tersedia. Hal ini sejalan dengan skor sekitar 0,62 yang diperoleh pada sumur yang benar-benar tidak dikenal menurut Hall dan Hall (2017), dengan demikian kategori generalisasi kapabilitas yang baik dari strategi yang diusulkan. Selain itu, juga dilakukan perhitungan *f-score* yang dapat dicapai tanpa menggunakan fitur augmentasi. Nilai ini adalah 0,55 sehingga 6 persen lebih rendah yang menggunakan augmentasi fitur, yang sekali lagi untuk membuktikan yang terpenting (Bestagini, Lipari, & Tubaro, 2017).

Setahun setelahnya Kim et al (2018) melakukan klasifikasi fasies seismik menggunakan algoritma random forest dan menentukan pentingnya setiap atribut

dalam mengklasifikasikan fasies seismik (yaitu sensitivitas). Metode yang diterapkan adalah algoritma pembelajaran random forest untuk memprediksi litofasies seismik yang dibantu oleh *wireline* log dan interpretasi stratigrafi. Dengan beberapa atribut seismik yang digunakan yaitu atribut, I_p , I_s , V_p , V_s , $\mu\rho$, V_p / V_s . Dari hasil penggunaan metode ini diperoleh dua fasies batuan yaitu batu kapur dan batu serpih, dan untuk atribut yang paling memiliki pengaruh untuk klasifikasi adalah atribut I_p . Tingkat akurasi rata rata dari model adalah 0,91 (Kim et al., 2018).

Di tahun yang sama Shi et al (2019) melakukan peningkatan permeabilitas dan estimasi produktivitas dengan klasifikasi dan data inti dari berbagai lapangan minyak. Langkah yang dilakukan pada penelitian ini yaitu, pertama MRGC digunakan untuk mengklasifikasikan elektrofases dari kurva *logging* di bagian coring. Data *logging* yang akan digunakan adalah log *gamma* ray, densitas dan neutron. Kedua, algoritma KNN digunakan untuk mempelajari hasil klasifikasi elektrofases menjadi bagian-bagian tanpa *core*. Akhirnya, model permeabilitas berdasarkan kendala elektrofases ditetapkan. Parameter *Neighbor Index* (NI) dalam algoritma dapat dengan cepat mengklasifikasikan data sampel, dan parameter *Kernel Representative Index* (KRI) dapat memilih kelas yang optimal dari hasil beberapa klasifikasi. Berdasarkan standarisasi kurva *logging*, klasifikasi elektrofases dari beberapa ladang minyak dan kemudian mempelajari klasifikasi elektrofases untuk bagian sumur yang tidak dikoreksi. Menurut karakteristik respons yang berbeda dari kurva *logging*, kurva *logging* konvensional dapat dengan cepat diklasifikasikan berdasarkan algoritma MRGC, dan menggabungkan data cutting dan data analisis *core*, hubungan yang sesuai antara litofases dan elektrofases dapat ditentukan; Berdasarkan hubungan korespondensi litofases dan elektrofases, model evaluasi permeabilitas regional Huanghekou sag dibuat dengan masalah elektrofases (Shi et al., 2019).

Hossain, Wataada, Hermana, dan Aziz (2020) menggunakan mesin *supervised* untuk klasifikasi elektrofases menggunakan teori *Rough Set Theory* (RST). Ide utama dari penelitian ini adalah untuk memprediksi elektrofases secara akurat yang memiliki pengaruh yang sangat signifikan pada perhitungan parameter *reservoir* lainnya seperti permeabilitas, litologi, fasies pengendapan dan

lain-lain. Untuk mengklasifikasikan elektrofases digunakan metode Induksi Aturan berbasis RST untuk mengidentifikasi klaster dari deskripsi unik log sumur. tanggapan yang mencerminkan mineral dan litofases dari interval log. Dari 10 atribut log (DPHI, NPHI, PE, DCT, DRHO, DRHO, GR, HDIA, CT10, CALI, dan BHVT) dilakukan seleksi fitur menggunakan Extra Trees Classifier (ETC). Setelah pemilihan fitur data akan dibagi kedalam kelompok menggunakan PCA dan selanjutnya akan RST akan mengekstraksi pola elektrofases dalam bentuk aturan keputusan yang memungkinkan peneliti geosains untuk mengkorelasikan elektrofases dengan litofases dari himpunan kasar yang diekstraksi (RS). Hasil yang diperoleh menunjukkan bahwa klasifikasi EF berbasis RST akan membantu dalam melakukan prediksi kelas EF yang akurat dan aturannya dapat digunakan untuk analisis lebih lanjut untuk menentukan litologi studi lain. Dalam perluasan penelitian ini, akan membangun model untuk mendefinisikan litolgi dengan aturan RST yang diekstraksi (Hossain, Wataada, Hermana, & Aziz, 2020).

Para peneliti mulai melakukan studi perbandingan untuk mengetahui teknik yang lebih tepat untuk melakukan klasifikasi dan prediksi fasies *reservoir*. seperti yang dilakukan oleh Lee et al., (1999), mereka melakukan penelitian tentang karakterisasi elektrofases dan prediksi permeabilitas di *reservoir* karbonat menggunakan analisis multivariat dan regresi nonparametrik. Dalam penelitian ini dilakukan pendekatan dua langkah untuk prediksi permeabilitas yang menggunakan regresi non-parametrik dalam hubungannya dengan analisis statistik multivariat. Pertama, mengklasifikasikan data log sumur ke dalam jenis elektrofases. Klasifikasi ini tidak memerlukan subdivisi buatan dari populasi data tetapi mengikuti secara alami berdasarkan karakteristik unik dari pengukuran log sumur yang mencerminkan mineral dan litofases dalam interval log. Kombinasi analisis komponen utama, analisis klaster berbasis model dan analisis diskriminan digunakan untuk mengkarakterisasi dan mengidentifikasi jenis elektrofases. Kedua, menerapkan teknik regresi non-parametrik untuk memprediksi permeabilitas menggunakan log sumur dalam setiap elektrofases. Tiga pendekatan non-parametrik diperiksa yaitu ekspektasi bersyarat alternatif (ACE), model aditif umum (GAM) dan jaringan saraf (NNET) dan keuntungan dan kerugian relatif dieksplorasi. Di antara tiga metode non-parametrik, model ACE

tampaknya mengungguli model lain dalam hal kemampuan prediktifnya. Model NNET cenderung melebih-lebihkan atau meremehkan dalam beberapa interval formasi. Terutama kinerja prediksi yang buruk di sumur buta menunjukkan bahwa jaringan saraf terlatih tidak selalu memberikan prediksi yang akurat untuk data yang tidak digunakan selama pelatihan. Model GAM memberikan hasil yang memuaskan untuk sebagian besar interval di sumur buta.

Avseth et al., (2002) melakukan klasifikasi litofasies seismik dari log sumur menggunakan statistik fisika batuan. Terdapat tiga metode yang akan digunakan kemudian dibandingkan yaitu metode *Mahalanobis discriminant analysis* (MLDA), *probability density function* (PDF), dan klasifikasi *neural network* (NN). Untuk metode MLDA, Metode ini hanya menggunakan sarana dan kovarian data pelatihan. Sampel diklasifikasikan menurut jarak Mahalanobis minimum untuk setiap kluster dalam data pelatihan. Pada metode PDF perkiraan V_p versus GR yang mewakili berbagai fasies di sumur jenis. Pada metode NN klasifikasi menggunakan log porositas dan densitas. Mahalanobis (MLDA), klasifikasi fungsi kepadatan probabilitas (PDF), dan klasifikasi jaringan saraf (NN) semuanya memiliki tingkat keberhasilan sekitar 80% ketika v_p , dan log sinar γ digunakan secara bersamaan. NN sedikit lebih baik dari MLDA, yang lagi-lagi sedikit lebih baik dari PDF. Namun, NN membutuhkan lebih banyak upaya komputasi daripada MLDA, atau PDF. Keunggulan PDF dibandingkan MLDA adalah dapat dengan mudah mengungkapkan jenis litofasies lain yang tidak dipertimbangkan dalam data pelatihan dan / atau pengukuran log yang salah.

Perez, Datta-Gupta, Mishra, & others, (2003) melakukan penggunaan pohon klasifikasi sebagai analisis perbandingan untuk prediksi permeabilitas dengan unit elektrofaseies, litofasies dan aliran hidrolik. penelitian ini menggunakan pohon klasifikasi untuk klasifikasi fasies dan memprediksi permeabilitas pada sumur yang tidak dikoreksi. Partisi data dilakukan dengan tiga pendekatan yaitu elektrofaseies, litofasies, dan HFU kemudian ketiganya akan dibandingkan untuk akurasi prediksi permeabilitas. Untuk klasifikasi digunakan 7 jenis log yaitu (GR, LLD, MSFL, DT, NPHI, RHOB, dan PEF). Metodologi analisis pohon klasifikasi dapat menjadi alat yang ampuh dalam mengidentifikasi elektrofaseies, litofasies,

atau HFU dalam sumur yang tidak digali berdasarkan respons log sumur. Pendekatannya sederhana dan intuitif dan tidak memerlukan koefisien atau persamaan yang telah ditentukan sebelumnya untuk klasifikasi dan regresi. Dari tiga metode partisi data yang dipelajari di sini, pendekatan elektrofases menghasilkan pohon terkecil dan juga berkinerja terbaik dalam mode prediksi. Baik litofases maupun HFU menghasilkan pohon yang jauh lebih besar dan lebih kompleks dengan prediksi yang kurang akurat.

Bagheri & Riahi(2015)melaakuan analisa fasies seismik dari data *logging* menggunakan skema klasifikasi beberapa teknik *supervisedmachine learning*. Dalam analisis fasies seimik digunakan beberapa teknik *supervisedmachine learning* berupa SVC, MLP, Parzen, KNN dan Fisher. EFA dilakukan dengan menggunakan log neutron *porosity* (NPHI), *Bulk density* (RHOB), *sonic log* (DT), dan *gamma ray* (GR), sebagai data masukan utama, untuk mengontrol litologi *reservoir*. Berdasarkan hasil empiris, SVC adalah pengklasifikasi nonlinier yang paling efektif untuk memasukkan data seismik ke dalam estimasi litologi di mana fasies adalah data nonlinear yang dapat dipisahkan dan tidak dapat dibedakan menggunakan pengklasifikasi linier.

W. J. Al-Mudhafar, Al Lawe, Noshi, & others(2019) melakukan eksperimen peningkatan karakterisasi *reservoir* karbonat menggunakan analisis klaster. Untuk meningkatkan karakterisasi *reservoir* dilakukan analisis klaster dengan menggunakan alogritma K-Means dan Calinski-Harabasz menggunakan perangkat lunak R. Analisis klaster diadopsi dalam penelitian ini sebagai pendekatan statistik untuk memodelkan dan memprediksi distribusi spasial dari *reservoir* litofases yang diberikan data *well log* meliputi GR, SP, Densitas, Porositas Neutron, Porositas Total, Resistivitas, Induksi, Volume Shale, Saturasi Air, beserta nilai porositas dan permeabilitas dari analisis inti. Pada penggunaan algoritma K-Means diperoleh 5 jenis kelas fasies sedangkan pada algoritma Calinski-Harabasz diperoleh 3 jenis kelas fasies karena kurangnya kepekaan pada pendekatan ini. Dari hasil kecocokan dengan data inti K-Mean membagi klaster lebih tepat dibandingkan dengan pembagian klaster menggunakan algoritma Calinski-Harabasz.

Silva (2020) melakukan penelitian tentang klasifikasi elektrofases menggunakan algoritma *supervised*. Menggunakan machine learning *supervised* SVM dan KNN untuk memprediksi fasies pengendapan sumur target dalam lingkungan transisi untuk validasi digunakan matrix kebingungan. Data variabel log konvensional yang digunakan adalah *Gamma ray* (GR), *Bulk density* (RHOB), *Neutron Porosity* (NPHI), *Deep resistivity* (RD), *Travel time* (DT), dan *Photoelectric factor* (PEF). Selain data log variabel petrofisik juga digunakan seperti volume *sahle* (VSH), saturasi air (SW), permeabilitas (KLogH), porositas densitas neutron (PHIF), porositas densitas (PHID), dan perbedaan antara densitas dan porositas neutron (N-D). Baik KNN & SVM memberikan kinerja klasifikasi yang cukup baik. Dengan mempertimbangkan jumlah data latih (2401 sampel), rata-rata model F-score 32% untuk KNN dan 40% untuk SVM, mengingat kinerja yang lebih tinggi dari yang terakhir ini. SVM menunjukkan kinerja representatif yang lebih baik bahkan dengan mendefinisikan bagian atas dan bawah setiap fasies pengendapan. Ada kesalahan klasifikasi mengenai pola yang tumpang tindih karena kesamaan tanggapan sumur *logging*. Lingkungan pengendapan transisi adalah lingkungan pengendapan yang paling sulit untuk mendefinisikan fasies karena sebagian besar fasies tersebut dapat dipahami berdasarkan bentuk geometrisnya.

Pada penelitian terbaru yang dilakukan oleh W. Al-Mudhafar (2020) dua pendekatan pembelajaran mesin non-parameterik yang diawasi digunakan untuk klasifikasi elektrofases di *reservoir* karbonat diteliti menggunakan *K-Nearest Neighbours* (KNN) dan *Random Forests* (RF). Algoritma klasifikasi *Random Forests* (RF) jauh lebih akurat untuk memprediksi klasifikasi elektrofases di *reservoir* karbonat daripada di KNN. Tingkat akurasi tinggi yang diperoleh dalam penelitian ini jauh lebih banyak dari yang telah diindikasikan dalam berbagai penelitian tentang *reservoir* karbonat, dimana tingkat akurasi berkisar antara 40% -60%.

Dalam studi kasus prediksi distribusi vertikal pada sumur di *reservoir* batupasir ini, penulis akan menggunakan data lapangan yang memiliki kumpulan data *logging* pada 7 sumur. Data log yang akan digunakan adalah GR, SP, LLD, LLS, MSFL, Densitas dan Neutron. Penelitian ini melakukan pendekatan dua

langkah untuk melakukan prediksi pada sumur uji. Langkah pertama akan dilakukan pembagian klaster menggunakan algoritma K-Means dan pada langkah kedua akan dilakukan prediksi fasies menggunakan algoritma *Random forest* pada sumur uji dengan menggunakan data yang tidak lengkap.



BAB III

METODOLOGI PENELITIAN

3.1 Uraian Metode Penelitian

Dalam penelitian ini menggunakan data lapangan dari *reservoir sandstone* heterogen. Data tersebut terdiri dari berbagai jenis log seperti GR, SP, LLD, LLS, MSFL, Densitas dan Neutron. Penelitian ini akan menggunakan data sumur dan sumur lainnya akan dijadikan sumur uji untuk prediksi.

Pendekatan dua langkah akan diterapkan dalam penelitian ini yaitu, pertama melakukan pembagian klaster menggunakan K-Means yang merupakan algoritma *unsupervised learning* sederhana dan pada langkah kedua menggunakan algoritma *supervised learning* berupa Random Forest untuk melakukan prediksi fasies pada sumur yang kekurangan data atau sebagian data yang dihilangkan. Konsep ini diimplementasikan dengan menggunakan bahasa pemrograman Python.

K-Means merupakan pendekatan sederhana untuk mengklasifikasi data yang dimiliki kedalam sejumlah klaster atau kelompok tertentu. Pendekatan ini didasari oleh sentriod K yang mendefinisikan penempatan setiap klaster. Setelah penempatan sentroid selesai, langkah selanjutnya adalah menentukan titik-titik yang sesuai dengan kumpulan data tertentu dan menetapkannya ke sentroid terdekat. Kemudian, sentroid K baru yang akan dihitung agar dapat ditetapkan sebagai pusat yang memungkinkan untuk kluster dan titik kumpulan data harus dikaitkan ke sentroid baru. Pusat awal yang dipilih berdasarkan nilai sentriod K, akan dihitung dengan menggunakan rumus *Euclidean Distance* :

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3.1)$$

Dimana :

- x_i : objek x ke- i
- y_i : daya y ke- i
- n : banyaknya objek
- I : $1, 2, 3, \dots, n$

Fungsi rumus ini untuk mencari jarak terdekat antara titik sentroid dengan data. Proses ini diiterasi sampai lokasi sentroid yang konstan akan dicapai (Merliana,

Santoso, & others, 2015)(W. J. Al-Mudhafar et al., 2019).

Penentuan nilai K jumlah sentroid K dapat ditentukan menggunakan metode *Elbow* melalui persentase hasil perbandingan antara jumlah kluster yang akan membentuk siku pada suatu titik. Sudut siku dengan penerunan titik yang paling terbesar akan menjadi jumlah nilai kluster yang tepat. Persentase hasil perbandingan antara jumlah kluster diperoleh melalui perhitungan nilai *Sum of Square Error*(SSE) dari masing-masing nilai kluster. Semakin besar jumlah nilai kluster K, maka nilai SSE akan semakin kecil. Rumus SSE ditunjukkan oleh rumus 3.2 :

$$SSE = \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (3.2)$$

Dimana:

SSE : *Sum of Square Error*

N : jumlah kasus

X : kasus

C : sentroid

Nilai K yang diperoleh melalui metode *Elbow* akan menjadi jumlah dan jenis fasies pada setiap sumur (Muningsih & Kiswati, 2018).

Random Forest adalah algoritma klasifikasi *supervised learning* berbasis pohon keputusan. Algoritma ini dimulai dengan *Bootstrap* atau *Bagging* data dengan memilih kasus N sebagai perwakilan data, dari seluruh kumpulan data untuk menghasilkan subset rangkaian. Kemudian menghasilkan beberapa pohon pada sampel *Bootstrap* yang berbeda dari data subset pelatihan dan akhirnya mengurangi *varians* pada pohon dengan membuat rata-rata, yang mengarah pada peningkatan kinerja pohon keputusan pada data subset pengujian menjadi *Overfitting* (Hastie, Tibshirani, & Friedman, 2009).

3.2 Lokasi Penelitian

Penelitian ini berlokasi di Laboratorium Komputer Teknik Perminyakan Universitas Islam Riau pada 1 Oktober – 30 November 2021 dengan perincian kegiatan seperti yang terlihat di tabel dibawah ini.

Tabel 3. 1 Jadwal Penelitian

No	Jenis kegiatan	Waktu Pelaksanaan (Minggu)							
		Oktober				November			
		1	2	3	4	1	2	3	4
1	Study literature								
2	Penentuan klaster menggunakan K-Means								
3	Klasifikasi data dan prediksi fasies menggunakan Random Forest								
4	Analisis hasil								
5	Laporan akhir								

3.3 Data Penelitian

Penelitian ini menggunakan data lapangan. Adapun data log sumur yang akan digunakan adalah log GR, SP, LLD, LLS, MSFL, Densitas dan Neutron. Rincian data logging sumur dapat dilihat pada website di bawah ini :

<https://drive.google.com/file/d/1mC4oX25ccNGpdalAhNPkrt4GeRE6Kcof/view?usp=sharing>

	Depth_(ft)	GR_(API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	density_(g/cc)
0	117	272.814510	-17.995239	131.328009	14.134101	1.871583
1	119	282.131096	-14.863115	409.691012	211.964066	1.901766
2	120	259.617974	-11.541445	555.010702	194.759967	1.891909
3	121	286.202763	-6.963452	647.117875	205.511670	1.886228
4	123	263.942729	-0.893649	651.211146	319.481097	1.886780
5	124	284.132430	-0.775592	679.864040	310.879048	1.850777
6	126	247.011865	-1.437969	679.864040	429.147783	1.846466
7	127	264.632828	-1.416789	679.864040	431.297437	1.850142
8	128	213.487619	-1.161446	679.864040	459.253239	1.908699
9	130	254.940531	-0.946426	653.257781	502.260051	1.980545
10	131	252.548124	-1.356867	651.211146	556.018567	1.976378
11	132	265.522318	-2.002515	663.490957	603.326404	1.935058

Gambar 3. 1 Data sumur b5

	Depth_(ft)	GR_(API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	density_(g/cc)	neutron_(v/v)
0	130	88.889855	-7.883801	90.398701	21.644230	2.001557	0.465857
1	131	55.654038	-5.553851	128.866727	24.929539	2.049000	0.508517
2	132	38.887490	-11.440925	93.587997	40.247270	2.080241	0.498412
3	133	39.313108	-14.150928	122.465978	79.267054	2.080638	0.474837
4	134	39.593731	-12.812154	269.742679	235.190977	2.121180	0.483819
...
442	642	109.052023	-1.731528	747.980168	802.488703	2.173208	0.316548
443	645	105.358074	-0.654800	741.018497	804.007723	2.128763	0.356961
444	646	107.873357	-0.500977	753.933458	795.497546	2.182731	0.345734
445	647	106.867025	-0.418441	740.141006	781.425558	2.179557	0.370994
446	649	109.027299	-0.375620	738.194204	758.614151	2.157335	0.345734

447 rows × 7 columns

Gambar 3. 2 Data sumur b10

	Depth_(ft)	GR_(API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	MSFL_(ohmm)	density_(g/cc)	neutron_(v/v)
0	159	59.997620	-39.177121	301.560448	436.024021	346.580025	2.118345	0.471800
1	160	68.432037	-63.361765	549.517413	647.547176	690.593664	1.985891	0.550869
2	161	71.539696	-43.207877	687.741690	746.235054	719.952566	1.923172	0.504682
3	162	69.763891	11.956133	669.108526	750.849619	684.640285	1.889492	0.507770
4	164	92.848043	3.318768	729.066634	729.526748	708.266355	1.882270	0.567788
...
83	436	92.780619	0.209334	762.945499	792.193175	707.822420	2.142340	0.482369
84	438	85.939873	-0.136352	894.791934	907.314381	813.315369	2.190179	0.421782
85	441	74.515023	-0.136352	1032.507628	1125.745683	1010.710751	2.299377	0.274433
86	448	79.169616	-2.727475	1049.133881	935.966348	793.138108	1.988572	0.379071
87	449	106.567845	-3.245786	896.549594	911.769967	712.975870	2.044916	0.416915

88 rows × 8 columns

Gambar 3. 3 Data sumur b30

	Depth_(ft)	GR_(API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	MSFL_(ohmm)	density_(g/cc)	neutron_(v/v)
0	160	75.067550	-20.066044	166.103659	133.295523	81.229100	1.940982	0.052465
1	166	91.979898	-4.670697	336.358655	464.194303	638.534530	1.958659	0.172865
2	167	99.496227	5.376632	603.825919	697.208716	628.545804	1.785523	0.116508
3	173	117.357509	7.321329	724.558725	710.643781	732.796716	1.926034	0.572494
4	181	87.415806	2.135539	796.727355	804.152588	704.169531	2.052481	0.472586
...
58	544	112.283602	-1.142601	959.278803	924.400829	486.968958	1.900702	0.507650
59	545	122.639462	-1.142601	947.467922	940.817645	174.925665	1.848957	0.506793
60	546	99.856168	-0.920937	1000.586808	970.490555	720.385266	1.817113	0.502295
61	550	125.401024	-0.699070	958.444466	1031.755937	1044.715537	2.109678	0.469714
62	560	143.351584	2.626293	854.821648	865.117455	808.838610	2.127590	0.491802

63 rows × 8 columns

Gambar 3. 4 Data sumur b36

	Depth_(ft)	GR_(API)	SP_(mV)	LLS_(ohmm)	MSFL_(ohmm)	density_(g/cc)	neutron_(v/v)
0	171	48.232194	-57.484055	441.513358	633.829825	1.850033	0.528392
1	175	65.701947	2.912049	583.362978	557.912334	1.884584	0.521498
2	176	69.253116	-0.328090	590.354403	565.504112	1.884584	0.553834
3	177	72.633469	-0.391009	604.092386	570.143499	1.878004	0.568735
4	179	64.745842	-0.391009	624.250170	568.954579	1.881294	0.569117
...
202	421	96.773918	-4.995556	636.362052	665.464156	1.825352	0.497454
203	422	90.559412	-2.970699	646.990945	645.641501	1.815479	0.501510
204	425	100.120159	-3.662811	638.894279	642.585924	1.767764	0.469827
205	426	95.658437	-4.354923	638.894279	656.277289	1.732883	0.469046
206	427	91.037456	-4.292004	658.125476	646.695586	1.712005	0.508854

207 rows × 7 columns

Gambar 3. 5 Data sumur b38

	Depth_(ft)	GR (API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	density_(g/cc)	neutron_(v/v)
0	173	58.059600	10.154273	481.719233	481.719233	1.846913	0.091573
1	175	66.037695	12.816046	466.664731	466.664731	1.810149	0.025482
2	177	69.792062	13.481425	493.972898	493.972898	1.822122	0.061208
3	179	69.322908	14.405503	493.972898	493.972898	1.847441	0.134444
4	181	70.026742	9.061099	484.782649	484.782649	1.849902	0.032627
...
306	681	86.539853	15.034019	738.243456	738.243456	2.019061	0.439897
307	683	82.499774	17.251993	780.723961	780.723961	2.054205	0.437635
308	684	81.397934	18.014517	780.723961	780.723961	2.053151	0.445001
309	685	82.426318	18.028429	780.723961	780.723961	2.046359	0.429455
310	690	91.976112	17.141043	824.842257	824.842257	2.007993	0.464650

311 rows × 7 columns

Gambar 3. 6 Data sumur b52

	Depth_(ft)	LLS_(ohmm)	LLD_(ohmm)	MSFL_(ohmm)	density_(g/cc)	neutron_(v/v)
0	155	471.557049	463.197272	94.988684	1.831815	0.091870
1	156	470.180252	461.880784	72.040815	1.791430	0.068012
2	157	469.107829	460.650763	16.917798	1.759419	0.014521
3	158	468.434175	459.516457	315.538784	1.798135	0.033445
4	159	468.210107	458.487144	640.676884	1.842592	0.063542
...
408	565	806.214305	820.978660	773.163169	2.007587	0.468651
409	566	804.608620	836.560419	757.481292	2.019750	0.464571
410	567	803.653823	875.673561	783.925648	2.043363	0.480136
411	568	805.425635	912.309957	812.579599	2.042736	0.488140
412	569	809.582774	943.309707	786.197699	2.016646	0.477498

413 rows × 6 columns

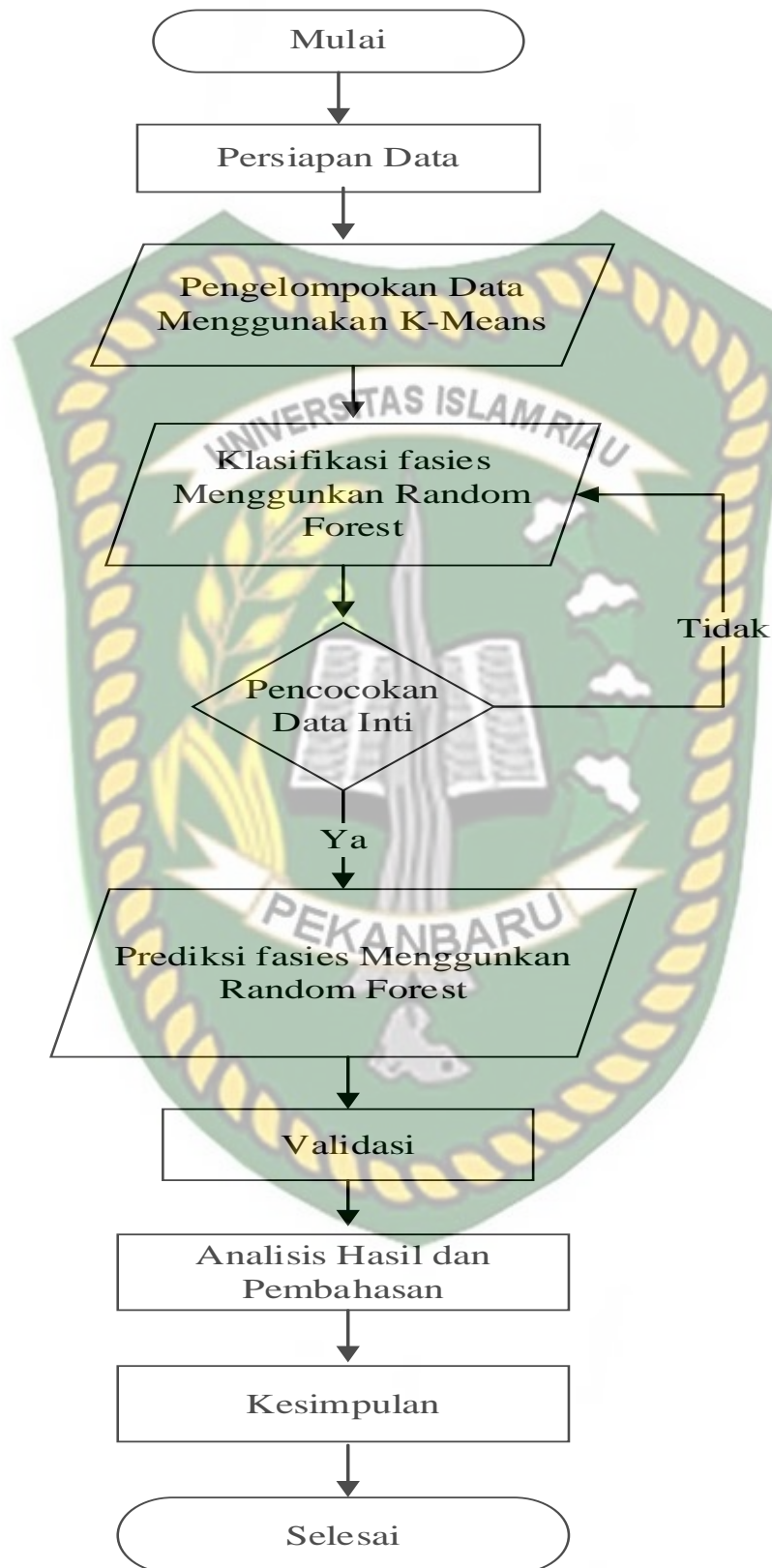
Gambar 3. 7 Data sumur b58

Ringkasan data yang akan digunakan dapat dilihat dari tabel dibawah ini:

Tabel 3. 2 Ringkasan Data Sumur

Jenis log	Nama Sumur						
	b5	b10	b36	b30	b38	b52	b58
GR	√	√	√	√	√	√	-
SP	√	√	√	√	√	√	-
LLS	√	√	√	√	√	√	√
LLD	√	√	√	√	-	√	√
Densitas	√	√	√	√	√	√	√
Neutron	-	√	√	√	√	√	√
MSFL	-	-	√	√	√	-	√

3.4 Flow Chart

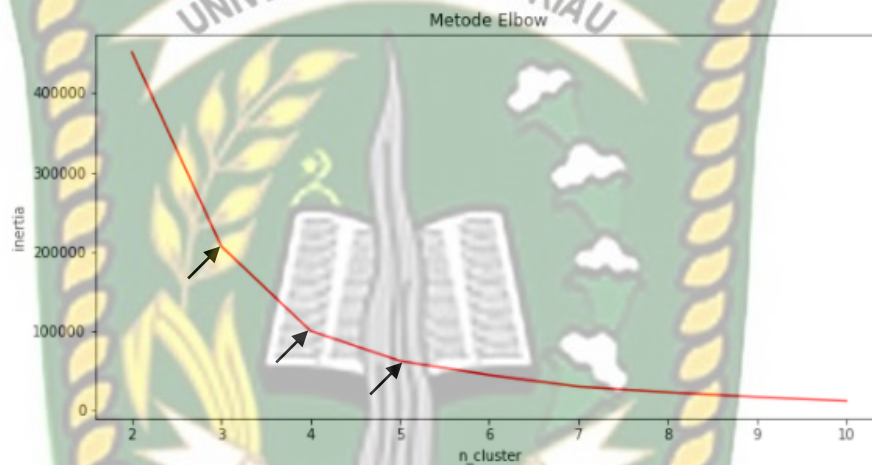


Gambar 3. 8 Flow Chart

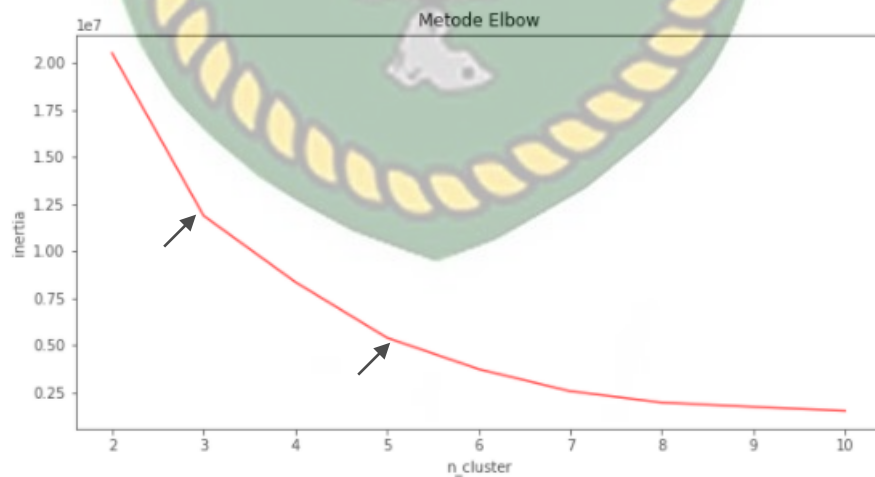
BAB IV HASIL DAN PEMBAHASAN

4.1 Penentuan Jumlah Kelas Fasies

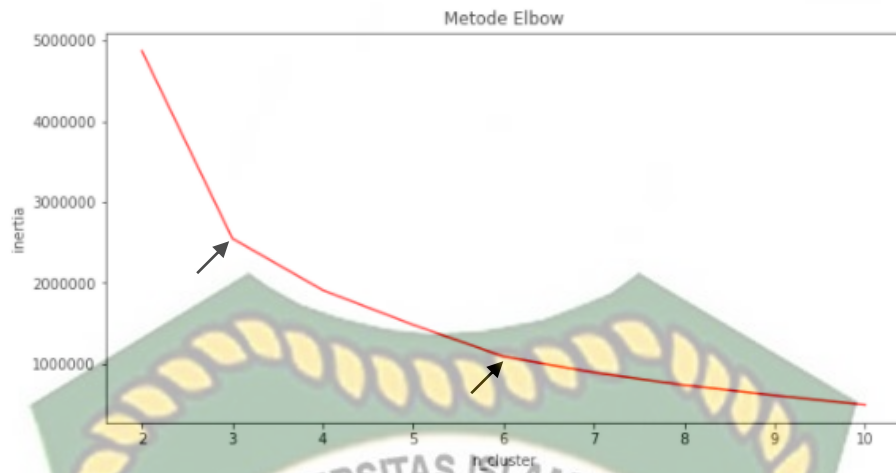
Penentuan klasifikasi fasies pada setiap sumur menggunakan algoritma K-Means dimulai dengan melakukan penentuan jumlah K. Jumlah K ditentukan berdasarkan perhitungan optimasi jumlah K menggunakan metode *Elbow*. Pada setiap sumur telah dilakukan perhitungan jumlah K seperti yang ditampilkan pada gambar dibawah ini.



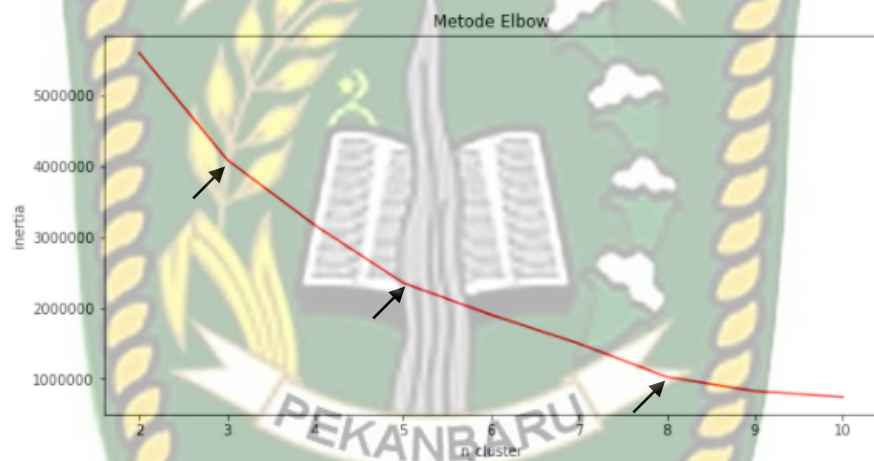
Gambar 4. 1 Grafik Metode *Elbow* pada sumur b5



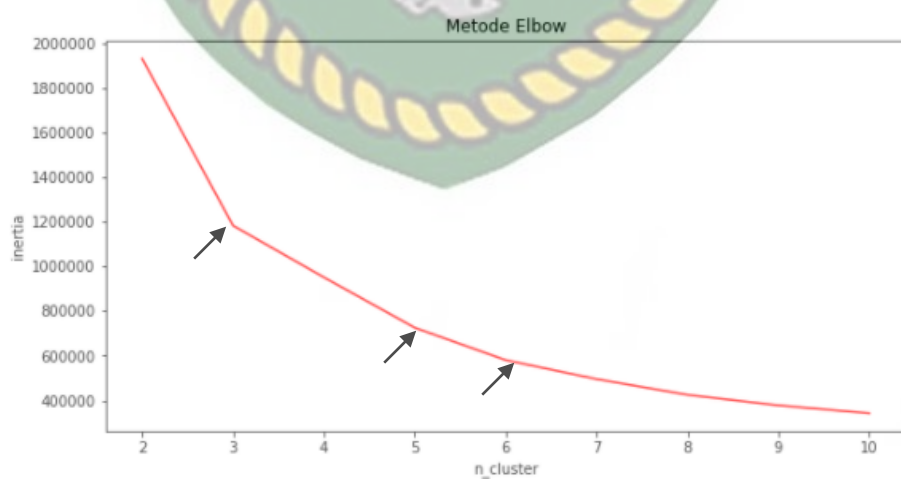
Gambar 4. 2 Grafik Metode *Elbow* pada sumur b10



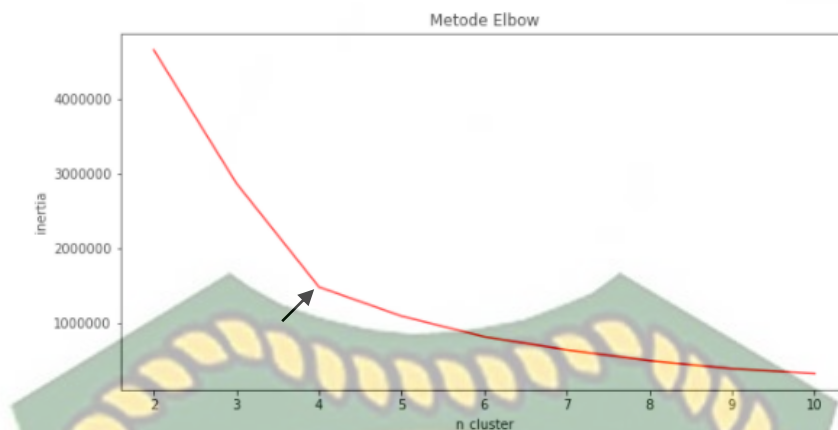
Gambar 4. 3 Grafik Metode *Elbow* pada sumur b30



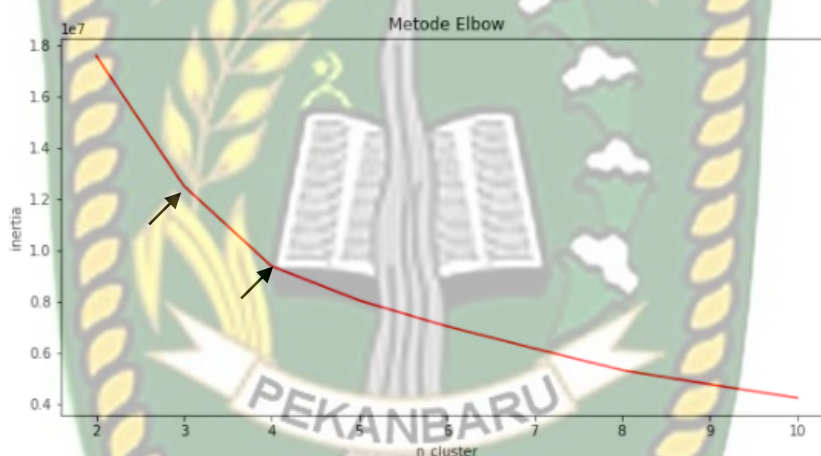
Gambar 4. 4 Grafik Metode *Elbow* pada sumur b36



Gambar 4. 5 Grafik Metode *Elbow* pada sumur b38



Gambar 4. 6 Grafik Metode *Elbow* pada sumur b52



Gambar 4. 7 Grafik Metode *Elbow* pada sumur b58

Pada gambar 4.1-4.7 menunjukkan hasil analisis metode *Elbow* di setiap sumur. Dari tampilan gambar untuk setiap sumur memiliki siku pada tiga n-klaster kecuali pada sumur w52, sehingga dari total tujuh sumur mayoritas berada pada tiga klaster maka untuk tahap selanjutnya *clustering* menggunakan algoritma K-Means akan menggunakan nilai tiga sebagai jumlah sentroid K.

Setelah dilakukan *clustering* pada ke tujuh sumur, peneliti memilih sumur b10 sebagai data acuan untuk melakukan *clustering* menggunakan algoritma K-Means karena sumur b10 memiliki data kedalaman sumur terbanyak dan memiliki enam atribut log yaitu GR, SP, LLS, LLD, Density, dan Neutron. Hasil dari *clustering* pada sumur b10 ditunjukkan oleh gambar dibawah ini

	GR_(API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	density_(g/cc)	neutron_(v/v)	Fasies
0	86.669855	-7.883801	90.398701	21.644230	2.001557	0.465857	2
1	55.654036	-5.553851	128.866727	24.929539	2.049000	0.508517	2
2	38.887490	-11.440925	93.587997	40.247270	2.060241	0.498412	2
3	39.313106	-14.150928	122.455978	79.267054	2.080636	0.474837	2
4	39.593731	-12.812154	269.742679	235.190977	2.121180	0.483819	2
...
442	109.052023	-1.731528	747.980168	802.488703	2.173208	0.316546	2
443	105.356074	-0.654600	741.016497	804.007723	2.128763	0.356961	2
444	107.873357	-0.500977	733.933458	795.497546	2.182731	0.345734	2
445	106.867025	-0.419441	740.141006	781.425556	2.179557	0.370994	2
446	109.027299	-0.375620	738.194204	758.614151	2.157335	0.345734	2

447 rows × 7 columns

Gambar 4. 8 Hasil *Clustering* Fasies pada sumur b 10

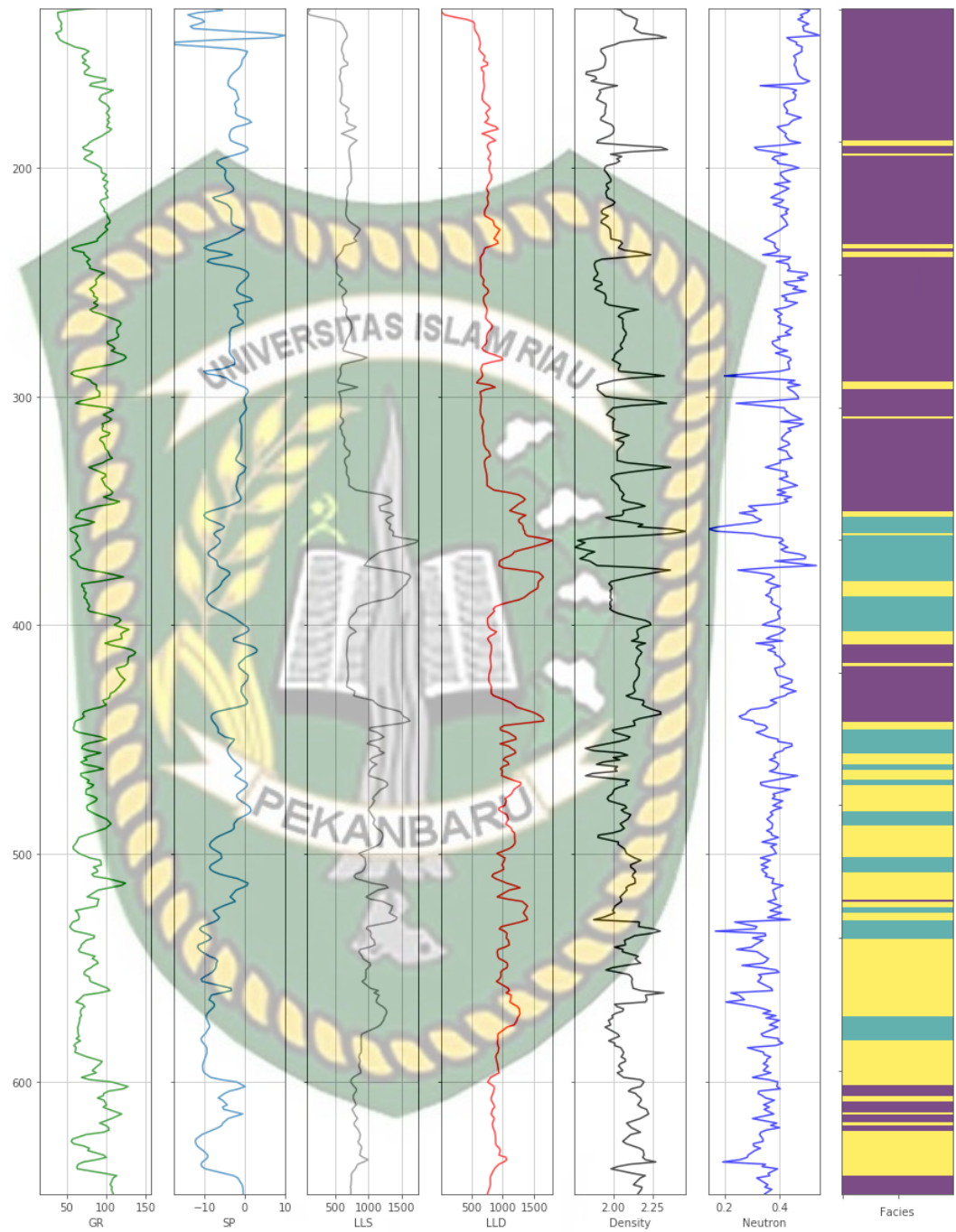
Fasies pada setiap kelas memiliki sentroid K sebagai pusat pengelompokan data. Setiap data dikelompokkan kedalam kelompok fasies 0, fasies 1, dan fasies 2 yang ditunjukkan pada gambar 4.9

	GR_(API)	SP_(mV)	LLS_(ohmm)	LLD_(ohmm)	density_(g/cc)	neutron_(v/v)	Fasies
0	92.035356	-2.788260	654.993190	720.820800	2.042408	0.416267	0
1	77.278104	-5.813320	1338.182683	1337.706095	2.062149	0.348064	1
2	81.281536	-6.331419	965.696639	984.772807	2.063489	0.360122	2

Gambar 4. 9 Sentroid K pada setiap kelas fasies pada sumur b10

Pada gambar 4.10 ditampilkan hasil klasifikasi fasies menggunakan algoritma K-Means. Setiap kelas fasies ditampilkan dengan warna yang berbeda yaitu kelas fasies 0 ditampilkan dengan warna ungu, kelas fasies ke 1 ditampilkan dengan warna kuning dan kelas fasies ke 2 ditampilkan dengan warna biru.

Well: Well10



Gambar 4. 10 Hasil klasifikasi sumur b10

4.2 Analisis Kelas Fasies

Analisis fasies pada sumur b10 dilakukan berdasarkan korelasi log GR, SP, LLS, LLD, Densitas dan Neutron. Setiap log memiliki fungsi yang saling terhubung sehingga dapat menentukan karakteristik fasies batuan.

Log GR merupakan sebuah log yang digunakan untuk mengukur kandungan unsur radio aktif dalam suatu formasi. Unsur radio aktif seperti Uranium, Thorium, dan Potasium umumnya ditemukan pada lapisan batuan serpih. Kandungan radio aktif sedikit sekali ditemukan di batuan *sandstone*, *dolomite* dan *limestone* sehingga log ini dapat menentukan daerah permeabel dan tidak permeabel. Semakin tinggi nilai log GR maka akan semakin tinggi kemungkinan zona formasi tersebut tidak permeabel (Ismahesa, 2016).

Log SP adalah suatu log yang digunakan untuk mengukur beda potensial pada elektroda yang ada diatas permukaan dan yang bergerak dalam lubang sumur. Beda potensial tersebut akan menghasilkan defleksi akibat perbedaan nilai salinitas lumpur pemboran (R_{mf}) dan fluida formasi (R_w). Jika nilai salinitas mengalami defleksi kekiri atau pun kanan maka zona tersebut di indikasikan sebagai zona permeabel dan apabila tidak terjadi defleksi maka zona tersebut diindikasikan sebagai zona *tight* atau impermeabel. Pada zona yang mengalami defleksi kekiri nilai $R_w \ll R_{mf}$ lapisan permeabel tersebut mengandung *saline water* dan pada zona yang memiliki defleksi kekanan nilai $R_w \gg R_{mf}$ lapisan tersebut mengandung *fresh water*. Hidrokarbon menyebabkan berkurangnya nilai SP pada zona permeabel (Schon, 2015).

Log LLS dan LLD merupakan log yang digunakan untuk menentukan resistivitas formasi. Kedua log ini berfungsi untuk menentukan lapisan reservoir dan non-reservoir, jenis fluida dan batas kontak fluida. Log LLS mengukur resistivitas disekitar zona transisi dan log LLD mengukur resistivitas di zona tidak terinvasi. Log LLS dan LLD biasanya ditampilkan dalam satu kolom, ketika defleksi kekanan oleh kedua log tersebut mengindikasikan zona reservoir. Zona hidrokarbon diindikasikan dengan defleksi kekanan serta nilai LLD lebih besar dibandingkan nilai LLS, jika nilai LLD lebih kecil dibandingkan dengan nilai LLS maka akan mengindikasikan zona air dan jika nilai kedua log sama maka akan mengindikasikan zona tidak permeable (Ellis & Singer, 2007).

Log densitas adalah suatu log yang berfungsi untuk menghitung densitas batuan dengan menggunakan bahan radio aktif berupa sinar gamma dan prinsip hamburan compton. Log densitas menghitung desitas batuan secara menerus pada formasi dalam satuan gr/cc. Interpretasi menggunakan log ini berdasarkan karakteristik grafik dengan melihat defleksi. Defleksi yang lebih rendah mengindisikan fluida gas dan defleksiyang tinggi mengindikasikan air maupun minyak (Goldberg, 2001).

Log neutron merupakan suatu log yang berfungsi sebagai penghitung porositas batuan dan pendeteksi zona gas. Persentase pori batuan diukur berdasarkan intensitas atom hidrogen yang diasumsikan sebagai hidrokarbon atau air. Log neutron mengukur intensitas hidrogen dalam satuan Porosity Unit (PU). Pada formasi yang mengndung minyak atau air memiliki kandungan hidrogen yang tinggi sehingga menyebabkan nilai Porosity Unit juga tinggi. Sedangkan pada formasi yang mengandung gas memiliki kandungan hidrogen yang rendah sehingga nilai Porosity Unit-nya rendah dan fenomena ini disebut juga dengan gas effect (Schon, 2015).

Hasil klasifikasi fasies menggunakan K-Means menghasilkan tiga kelas fasies yang memiliki sentriod pada setiap jenis log. Kelas fasies 0 memiliki nilai log GR tertinggi dibanding yang lainnya dan juga memiliki nilai SP yang juga tinggi ini mengindikasikan bahwa fasies batuan tersebut adalah batuan serpih atau lempung, selain itu pada nilai log LLS dan LLD juga memiliki nilai yang rendah dibandingkan kelas fasies yang lain sehingga semakin tinggi kemungkinan bataun tesebut adalah batuan non-reservoir. Nilai log densitas pada batuan tersebut juga memiliki nilai yang terendah tetapi nilai log Neutron pada zona ini memiliki nilai tertinggi sehingga ada kemungkinan air terdapat dalam batuan. Dari keseluruhan karakteristik yang ditunjukan oleh setiap log, batuan pada fasies 0 dapat diinterpretasikan sebagai batuan serpih.

Kelas fasies 1 memiliki nilai log GR terendah dibandingkan kelas faseis lainnya, selain itu pad nilai pada log SP juga menunjukan nilai yang terendah sehingga batuan pada zona ini diindikasikan sebagai batu pasir atau gamping. Pada log LLS dan LLD zona ini memiliki nilai tertinggi, ini mngindikasikan terdapat hidrokarbon pada batuan tersebut. Nilai log densitas dan Neutron pada

zona ini juga memiliki nilai yang tinggi menunjukkan adanya minyak pada batuan ini sehingga zona ini dapat diproduksi.

Kelas fasies 2 memiliki nilai log GR yang tinggi namun memiliki nilai log SP yang paling rendah sehingga ini mengindikasikan zona ini adalah batuan *shaly-sand*. Nilai log LLS dan LLD yang cukup tinggi mengindikasikan zona ini permeabel tetapi tidak sebaik zona pada kelas fasies 1. Pada log densitas dan neutron menginterpretasikan terdapat hidrokarbon di zona ini tetapi jika diproduksi tidak lebih menguntungkan dibandingkan kelas fasies 1.

4.3 Pembentukan Model Prediksi

Setelah diperoleh kelas atau fasies di setiap lapisan pada sumur b10, selanjutnya dibentuk model untuk melakukan prediksi fasies pada sumur lainya menggunakan algoritma Random Forest. Dari perolehan data setiap sumur terdapat beberapa perbedaan kondisi seperti perbedaan jumlah atribut jenis log sehingga model prediksi dibentuk berdasarkan jumlah atribut log pada setiap sumur.

Model pertama dibentuk berdasarkan kepada 4 jenis log yaitu GR, SP, LLS, dan Neutron dengan menggunakan *hyperparameter tuning* berupa *Randomized Search, Cross Validation* yang digunakan adalah 3, iterasi dilakukan sebanyak 50 kali, sehingga menghasilkan 150 *fitting* model dengan *score* R^2 terbaik pada model Random Forest Classification diperoleh yaitu 0.98 pada data *training* dan 0.95 pada data *test*, sehingga model pertama memiliki akurasi 95%.

Fitting 3 folds for each of 50 candidates, totalling 150 fits

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 4 concurrent workers.
[Parallel(n_jobs=-1)]: Done 42 tasks | elapsed: 16.5s
[Parallel(n_jobs=-1)]: Done 150 out of 150 | elapsed: 41.9s finished
```

```
{'algo__max_depth': 58, 'algo__max_features': 0.8168886881742097, 'algo__min_samples_leaf': 15, 'algo__n_estimators': 171}
0.9859943977591037 0.9831932773109244 0.9555555555555556
```

Gambar 4. 11 Deskripsi *running* model pertama *Randomized Search CV*

Hasil model prediksi Random Forest yang pertama ini digunakan untuk memprediksi sumur target. Sumur target pada model ini adalah sumur b5, b30, b36, b38, dan b52.

Model kedua dibentuk berdasarkan kepada 4 jenis log yaitu LLS, LLD, Densitas dan Neutron dengan menggunakan *hyperparameter tuning* berupa

Randomized Search, *Cross Validation* yang digunakan adalah 3, iterasi dilakukan sebanyak 50 kali, sehingga menghasilkan 150 *fitting* model dengan *score* R^2 terbaik pada model Random Forest Classification diperoleh yaitu 0.99 pada data *training* dan 0.97 pada data *test*, sehingga model kedua memiliki akurasi 97%.

Fitting 3 folds for each of 50 candidates, totalling 150 fits

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 4 concurrent workers.
[Parallel(n_jobs=-1)]: Done 42 tasks      | elapsed: 10.5s
[Parallel(n_jobs=-1)]: Done 150 out of 150 | elapsed: 37.4s finished

{'algo__max_depth': 58, 'algo__max_features': 0.8168886881742097, 'algo__min_samples_leaf': 15, 'algo__n_estimators': 171}
0.9971988795518207 0.9943977591036415 0.9777777777777777
```

Gambar 4. 12 Deskripsi *running* model kedua *Randomized Search* CV

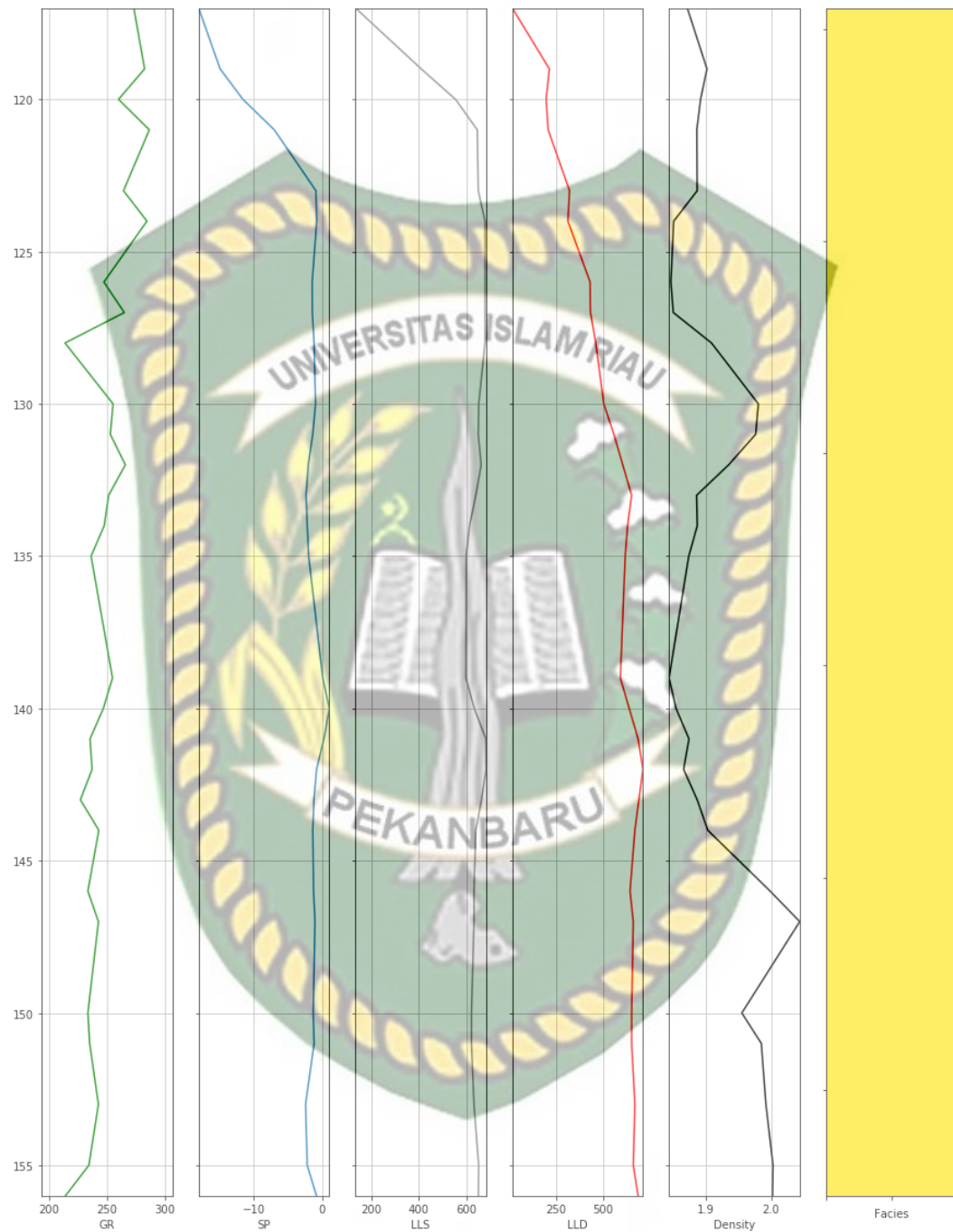
Hasil model prediski Random Forest yang kedua ini akan digunakan untuk prediksi sumur b58.

4.4 Hasil Prediksi Fasies Menggunakan Random Forest

Setelah dilakukan pembentukan dua model menggunakan algoritma Random Forest berdasarkan jenis data log yang tersedia pada setiap sumur dapat dilihat pada gambar 4.13-4.18. Pada model pertama prediksi fasies dibentuk berdasarkan log GR, SP, LLS, dan Neutron. Model ini memprediksi sumur b5, b5, b30, b36, b38, dan b52.

Pada sumur b5 prediksi dilakukan pada kedalaman 117-156ft, hasil prediksi menunjukan pada kedalaman tersebut fasies batuan dikategorikan kedalam fasies 2 dimana fasies ke 2 ini merupakan batuan *shlay-sand*. Karakteristik log menunjukan lapisan tersebut mengandung hidrokarbon seperti yang terlihat pada nilai log LLD dan LLS yang tinggi serta nilai log SP yang rendah, namun pada nilai log GR lapisan ini memiliki nilai yang sangat tinggi yaitu 213 hingga 286 API. Pada lapisan ini data logging yang dimiliki sedikit sehingga kedalaman yang diprediksi hanya sepanjang 39 ft.

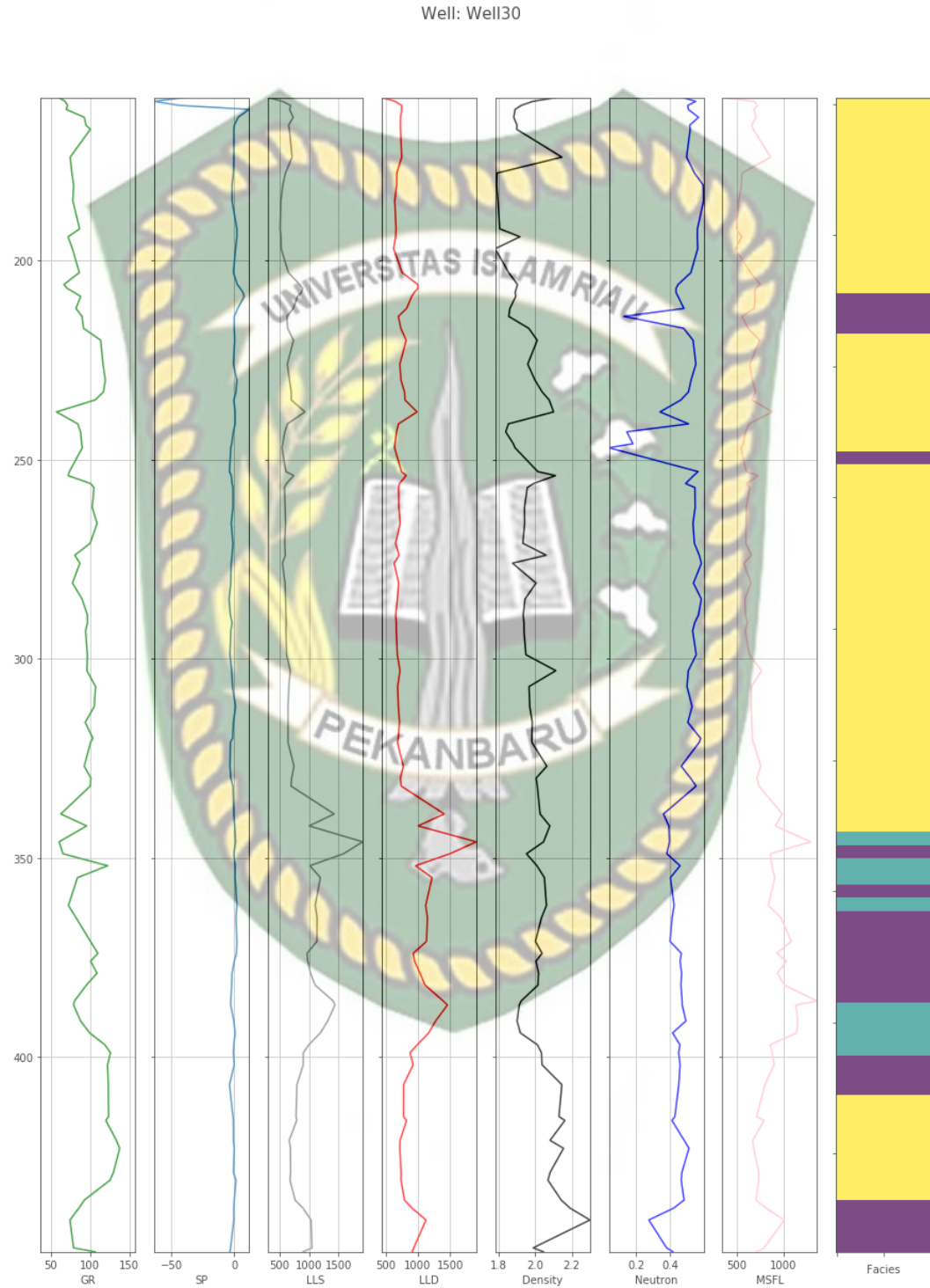
Well: Well5



Gambar 4. 13 Hasil Prediksi Fasies Sumur b5

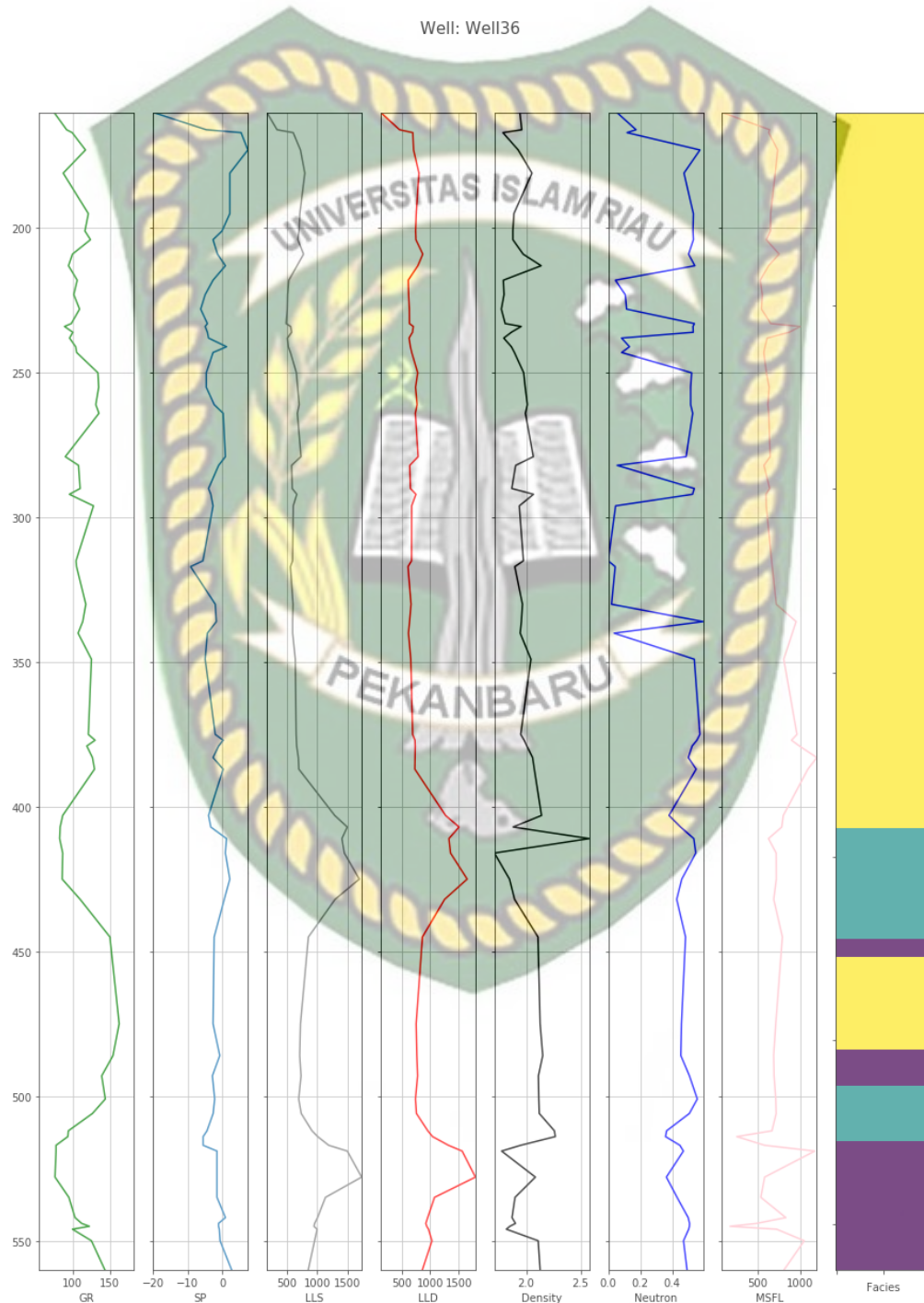
Pada sumur b30 prediksi dilakukan pada kedalaman 159-449ft, hasil prediksi mengidentifikasi 3 jenis fasies. Pada kedalaman 159-332ft fasies batuan diidentifikasi sebagai fasies 2 yang diselingi dengan fasies 0 pada kedalaman 206-208ft dan 238ft. Pada kedalaman 339-402ft terdapat perselingan antara fasies 1

dan fasies 0. Pada lapisan 407-463ft fasies yang diidentifikasi adalah fasies 2 dan pada lapisan 438-449ft fasies yang ditemukan adalah fasies 0.



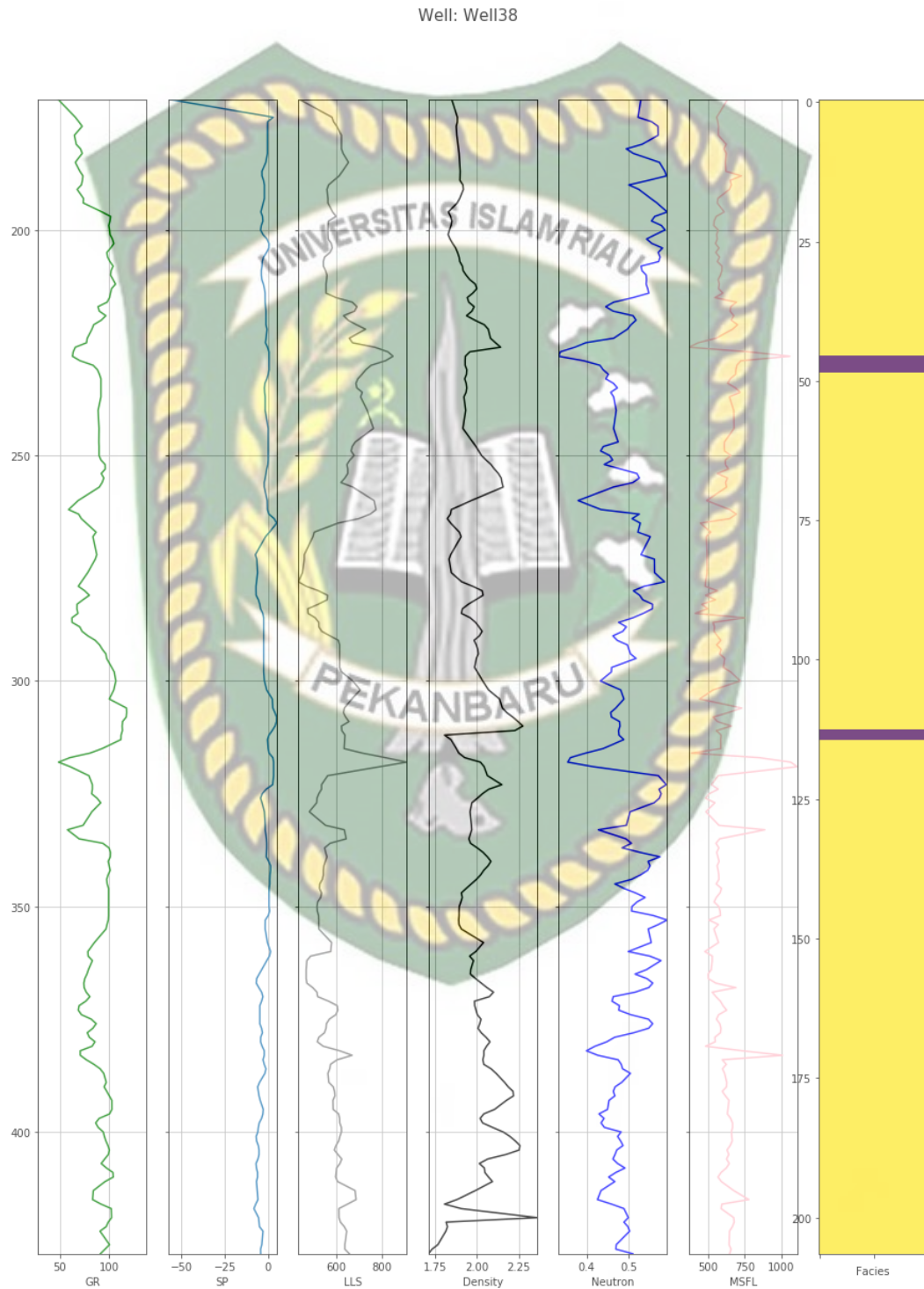
Gambar 4. 14 Hasil Prediksi Fasies Sumur b30

Pada sumur b36 prediksi dilakukan pada kedalaman 160-560ft, hasil prediksi mengidentifikasi 3 jenis fasies. Pada kedalaman 160-387ft fasies batuan diidentifikasi sebagai fasies 2. Pada kedalaman 403-432ft fasies batuan diidentifikasi sebagai fasies 1. Pada kedalaman 445-560ft fasies batuan diidentifikasi sebagai fasies 0 yang diselingi oleh fasies 2 dan fasies 1.



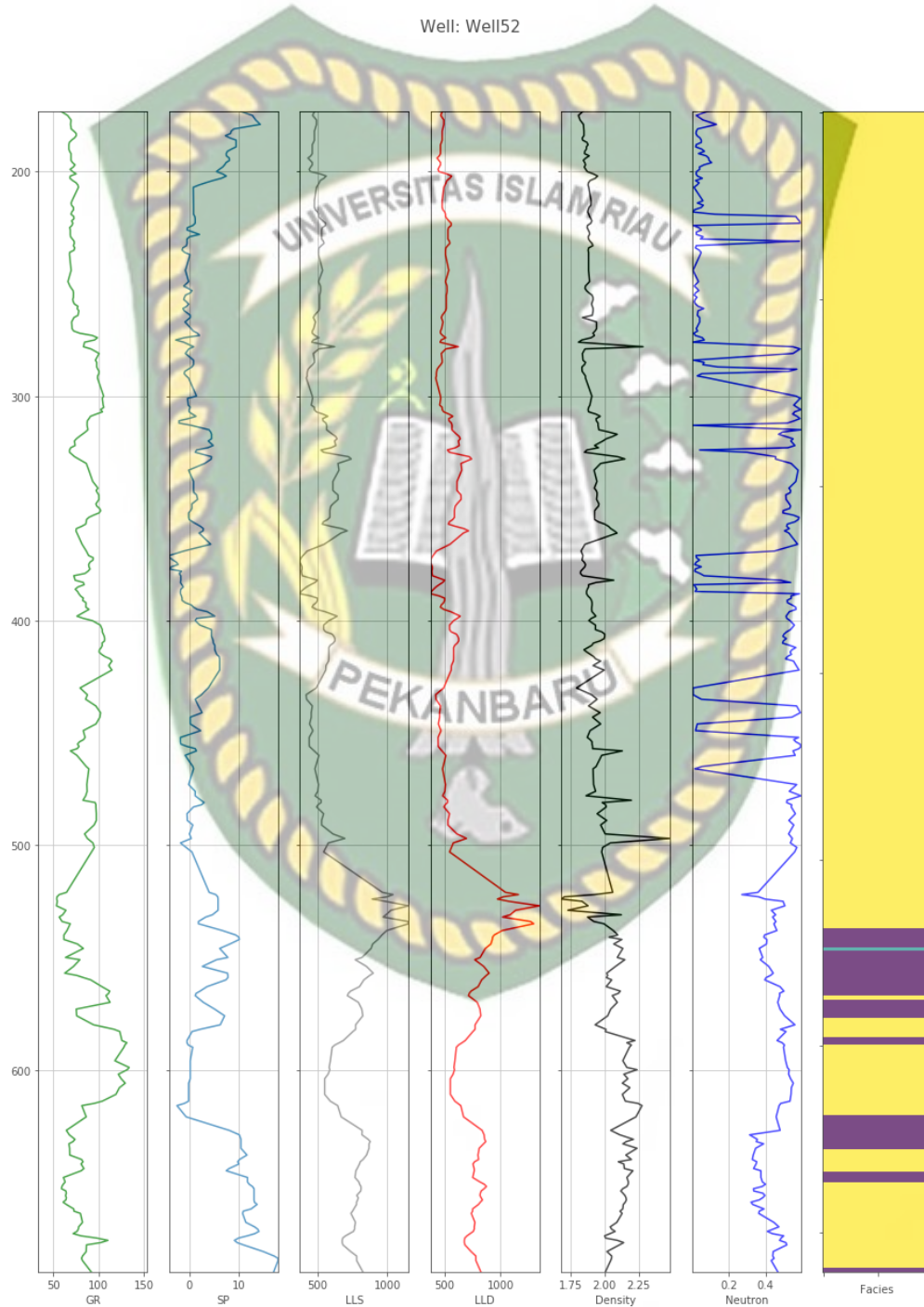
Gambar 4. 15 Hasil Prediksi Fasies Sumur b36

Pada sumur b38 prediksi dilakukan pada kedalaman 171-427ft, hasil prediksi mengidentifikasi 2 jenis fasies yaitu fasies 0 dan fasies 2. Pada sumur ini formasi batuan didominasi oleh fasies 2. Hanya pada kedalaman 227-229 ft dan 317-318ft fasies 0 teridentifikasi.



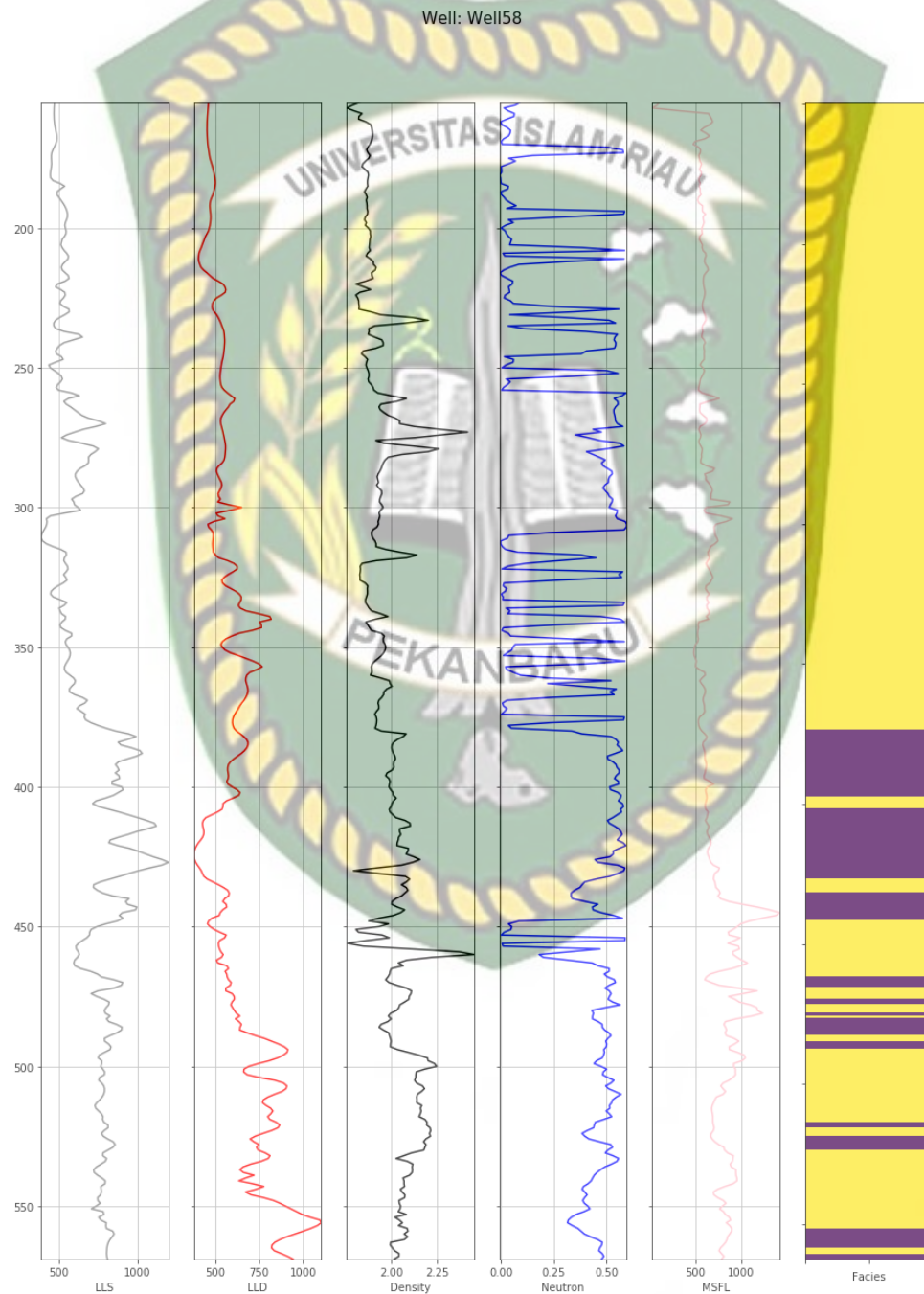
Gambar 4. 16 Hasil Prediksi Fasies Sumur b38

Pada sumur b52 prediksi dilakukan pada kedalaman 173-690ft, hasil prediksi mengidentifikasi 3 jenis fasies. Pada kedalaman 173-503ft fasies batuan diidentifikasi sebagai fasies 2. Pada kedalaman 521-576ft terdapat perselingan tiga fasies formasi batuan yang didominasi oleh fasies 1. Pada kedalaman 563-690ft terdapat perselingan dua fasies yaitu fasies 2 dan fasies 0.



Gambar 4. 17 Hasil Prediksi Fasies Sumur b52

Pada sumur b58 prediksi fasies dilakukan menggunakan model ke 2 karena jenis data log yang dimiliki berbeda dengan sumur lainya sehingga dibutuhkan model dengan jenis log yang berbeda untuk prediksi fasies. Prediksi dilakukan pada kedalaman 155-569ft. Hasil prediksi menunjukan terdapat dua jenis fasies diidentifikasi yaitu fasies 2 dan fasies 0. Formasi batuan pada sumur ini didominasi oleh fasies 2.



Gambar 4. 18 Hasil Prediksi Fasies Sumur b58

BAB V

PENUTUPAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan dari yang telah dituliskan sebelumnya, kesimpulan dari penelitian ini adalah klasifikasi yang dilakukan menggunakan algoritma K-Means berhasil menentukan tiga jenis fasies. Prediksi fasies menggunakan algoritma Random Forest dilakukan dengan menggunakan dua model berdasarkan jenis atribut log yang berbeda. Model pertama memiliki akurasi 95% digunakan kepada sumur b5, b30, b36, b38, dan b52 berdasarkan pada data log GR, SP, LLS, dan Neutron. Model kedua memiliki akurasi 97% digunakan pada sumur b58 berdasarkan pada data log LLS, LLD, Densitas dan Neutron.

5.2 Saran

Berdasarkan pembahasan yang telah dijabarkan dalam penelitian ini, diharapkan pada penelitian selanjutnya dapat menggunakan data core untuk mengetahui jenis fasies sehingga akurasi prediksi fasies semakin tepat.

DAFTAR PUSTAKA

- Adoghe, L. I., Aniekwe, O. S., Nwosu, C., & others. (2011). Improving electrofacies modeling using multivariate analysis techniques: a deepwater turbidite case study. *Nigeria Annual International Conference and Exhibition*.
- Al-Mudhafar, W. (2020). *Advanced Supervised Machine Learning Algorithms for Efficient Electrofacies Classification of a Carbonate Reservoir In a Giant Southern Iraqi Oil Field*. <https://doi.org/10.4043/30906-MS>
- Al-Mudhafar, W. J. (2017). Integrating kernel support vector machines for efficient rock facies classification in the main pay of Zubair formation in South Rumaila oil field, Iraq. *Modeling Earth Systems and Environment*, 3(1), 12.
- Al-Mudhafar, W. J., Al Lawe, E. M., Noshi, C. I., & others. (2019). Clustering analysis for improved characterization of carbonate reservoirs in a Southern Iraqi Oil Field. *Offshore Technology Conference*.
- Al-Mudhafar, W. J., & others. (2016). Applied Geostatistical Reservoir Characterization in R: Review and Implementation of Permeability Estimation Modeling and Prediction Algorithms-Part II. *Offshore Technology Conference*.
- Al-Mudhafer, W. J. M. (2014). Integrating Markov chains for Bayesian estimation of vertical facies sequences through linear discriminant analysis. *76th EAGE Conference and Exhibition 2014*, 2014(1), 1–3.
- Anxionnaz, H., Delfiner, P., & Delhomme, J. P. (1990). Computer-generated corelike descriptions from open-hole logs. *AAPG Bulletin*, 74(4), 375–393.
- Avseth, P., Mukerji, T., & others. (2002). Seismic lithofacies classification from well logs using statistical rock physics. *Petrophysics*, 43(02).
- Bagheri, M., & Riahi, M. A. (2015). Seismic facies analysis from well logs based on supervised classification scheme with different machine learning techniques. *Arabian Journal of Geosciences*, 8(9), 7153–7161.

- Baldwin, J. L., Bateman, R. M., Wheatley, C. L., & others. (1990). Application of a neural network to the problem of mineral identification from well logs. *The Log Analyst*, 31(05).
- Bestagini, P., Lipari, V., & Tubaro, S. (2017). A machine learning approach to facies classification using well logs. In *Seg technical program expanded abstracts 2017* (pp. 2137–2142). Society of Exploration Geophysicists.
- Bucheb, J. A., Evans, H. B., & others. (1994). Some applications of methods used in electrofacies identification. *The Log Analyst*, 35(01).
- Busch, J. M., Fortney, W. G., Berry, L. N., & others. (1987). Determination of lithology from well logs by statistical analysis. *SPE Formation Evaluation*, 2(04), 412–418.
- Catuneanu, O. (2006). *Principles of sequence stratigraphy*. Elsevier.
- Delfiner, P., Peyret, O., Serra, O., & others. (1987). Automatic determination of lithology from well logs. *SPE Formation Evaluation*, 2(03), 303–310.
- Dorfman, M. H., Newey, J.-J., & Coates, G. R. (1990). New techniques in lithofacies determination and permeability prediction in carbonates using well logs. *Geological Society, London, Special Publications*, 48(1), 113–120.
- Doveton, J. H., & Prensky, S. E. (1992). Geological applications of wireline logs: a synopsis of developments and trends. *The Log Analyst*, 33(3), 286–303.
- Ellis, D. V., & Singer, J. M. (2007). *Well logging for earth scientists* (Vol. 692). Springer.
- Euzen, T., Delamaide, E., Feuchtwanger, T., Kingsmith, K. D., & others. (2010). Well log cluster analysis: an innovative tool for unconventional exploration. *Canadian Unconventional Resources and International Petroleum Conference*.
- Galloway, W. E. (1989). Genetic stratigraphic sequences in basin analysis I: architecture and genesis of flooding-surface bounded depositional units. *AAPG Bulletin*, 73(2), 125–142.
- Gkortsas, V.-M., & Liang, L. (2019). Automatic electrofacies classification from

well logs. In *SEG Technical Program Expanded Abstracts 2019* (pp. 2629–2633). Society of Exploration Geophysicists.

Goldberg, D. (2001). *Well Logging for Physical Properties: A Handbook for Geophysicists, Geologists, and Engineers*. Wiley Online Library.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Hossain, T. M., Wataada, J., Hermana, M., & Aziz, I. A. (2020). Supervised machine learning in electrofacies classification: A rough set theory approach. *Journal of Physics: Conference Series*, 1529(5), 52048.

Isimbabi, O. A., & others. (2018). Automation of Electrofacies Identification-A Case for Digitalization. *SPE Nigeria Annual International Conference and Exhibition*.

Ismahesa, A. (2016). *Karakterisasi Reservoir Karbonat, Berdasarkan Data Log Sumur, Seismik dan Batuan di Blok X Formasi Baturaja, Cekungan Sumatera Selatan*.

Kim, Y., Hardisty, R., Torres, E., & Marfurt, K. J. (2018). Seismic facies classification using random forest algorithm. In *SEG Technical Program Expanded Abstracts 2018* (pp. 2161–2165). Society of Exploration Geophysicists.

Kobrunov, A., & Priezzhev, I. (2016). Hybrid combination genetic algorithm and controlled gradient method to train a neural network. *Geophysics*, 81(4), IM35--IM43.

Lee, S. H., Datta-Gupta, A., & others. (1999). Electrofacies characterization and permeability predictions in carbonate reservoirs: role of multivariate analysis and nonparametric regression. *SPE Annual Technical Conference and Exhibition*.

Lim, J.-S., Kang, J. M., Kim, J., & others. (1997). Multivariate statistical analysis for automatic electrofacies determination from well log measurements. *SPE Asia Pacific Oil and Gas Conference and Exhibition*.

- Merliana, N. P. E., Santoso, A. J., & others. (2015). *Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means Clustering*.
- Middleton, G. V. (1973). Johannes Walther's law of the correlation of facies. *Geological Society of America Bulletin*, 84(3), 979–988.
- Moline, G. R., & Bahr, J. M. (1995). Estimating spatial distributions of heterogeneous subsurface characteristics by regionalized classification of electrofacies. *Mathematical Geology*, 27(1), 3–22.
- Muningsih, E., & Kiswati, S. (2018). Sistem aplikasi berbasis optimasi metode elbow untuk penentuan clustering pelanggan. *Joutica*, 3(1), 117–124.
- Nichols, G. (1993). RG Walker & NP James (eds) 1992. Facies Models. Response to Sea Level Change. v+ 409 pp. Not a third edition of Facies Models; the book has been completely rewritten, hence new title. Order from: Geological Association of Canada, Publications, Department. *Geological Magazine*, 130(3), 405–406.
- Perez, H. H., Datta-Gupta, A., Mishra, S., & others. (2003). The role of electrofacies, lithofacies, and hydraulic flow units in permeability predictions from well logs: a comparative analysis using classification trees. *SPE Annual Technical Conference and Exhibition*.
- Posamentier, H. W., & Allen, G. P. (1999). *Siliciclastic sequence stratigraphy: concepts and applications* (Vol. 7). SEPM (Society for Sedimentary Geology) Tulsa, Oklahoma.
- Puskarczyk, E. (2019). Artificial neural networks as a tool for pattern recognition and electrofacies analysis in Polish palaeozoic shale gas formations. *Acta Geophysica*, 67(6), 1991–2003.
- Qi, L., & Carr, T. R. (2006). Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas. *Computers & Geosciences*, 32(7), 947–964.
- Rabaute, A. (1998). Inferring a continuous lithology and mineralogy from multivariate statistical analysis of well-logging data: Examples from some sedimentary structures associated with tectonic plates convergence zones

(Ocean drilling program leg 134, 156 and 160). *Mémoires Géosciences-Montpellier*.

Rider, M. H. (1990). Gamma-ray log shape used as a facies indicator: critical analysis of an oversimplified methodology. *Geological Society, London, Special Publications*, 48(1), 27–37.

Rogers, S. J., Fang, J. H., Karr, C. L., & Stanley, D. A. (1992). Determination of lithology from well logs using a neural network. *AAPG Bulletin*, 76(5), 731–739.

Schon, J. (2015). *Basic well logging and formation evaluation*. Bookboon.

Serra, O. (1986). Fundamentals of well-log interpretation of logging data. *Developments in Petroleum Science*, 15.

Serra, O., Sulpice, L., & others. (1975). Sedimentological analysis of shale-sand series from well logs. *SPWLA 16th Annual Logging Symposium*.

Serra, O. t, & Abbott, H. T. (1982). The contribution of logging data to sedimentology and stratigraphy. *Society of Petroleum Engineers Journal*, 22(01), 117–131.

Shi, X., Chen, H., Li, R., Yang, X., Liu, H., Li, T., & others. (2019). Improving permeability and productivity estimation with electrofacies classification and core data collected in multiple oilfields. *Offshore Technology Conference*.

Silva, J. C. M. (2020). Electrofacies classification using Supervised learning algorithms. *First EAGE Conference on Machine Learning in Americas*, 2020(1), 1–6.

Skalinski, M., Gottlib-Zeh, S., Moss, B., & others. (2006). Defining and predicting rock types in Carbonates-preliminary results from an integrated approach using core and log data from the Tengiz field. *Petrophysics*, 47(01).

Smith, T., & Treitel, S. (2010). Self-organizing artificial neural nets for automatic anomaly identification. In *SEG Technical Program Expanded Abstracts 2010* (pp. 1403–1407). Society of Exploration Geophysicists.

Tang, H. (2008). Improved carbonate reservoir facies classification using artificial

neural network method. *Canadian International Petroleum Conference*.

- Tang, Hong, Meddaugh, W. S., Toomey, N., & others. (2011). Using an artificial-neural-network method to predict carbonate well log facies successfully. *SPE Reservoir Evaluation & Engineering*, 14(01), 35–44.
- Tang, Hong, White, C., Zeng, X., Gani, M., & Bhattacharya, J. (2004). Comparison of multivariate statistical algorithms for wireline log facies classification. *AAPG Annual Meeting Abstract*, 88, 13.
- Teh, W., Willhite, G. P., Doveton, J. H., & others. (2012). Improved reservoir characterization in the Ogallah field using petrophysical classifiers within electrofacies. *SPE Improved Oil Recovery Symposium*.
- Van Wagoner, J. C., Mitchum, R. M., Campion, K. M., & Rahmanian, V. D. (1990). *Siliciclastic sequence stratigraphy in well logs, cores, and outcrops: concepts for high-resolution correlation of time and facies*.
- Watanabe, K., Suzuki, S., Eto, K., Samizo, N., & others. (1999). Reservoir Simulation for a Low Productive and Strongly Heterogeneous Carbonate Reservoir in the Arabian Gulf by Geostatistical Approach. *SPE Asia Pacific Oil and Gas Conference and Exhibition*.
- Wolf, M., Pelissier-Combescure, J., & others. (1982). FACIOLOG-automatic electrofacies determination. *SPWLA 23rd Annual Logging Symposium*.
- Wong, P. M., Jian, F. X., & Taggart, I. J. (1995). A critical comparison of neural networks and discriminant analysis in lithofacies, porosity and permeability predictions. *Journal of Petroleum Geology*, 18(2), 191–206.
- Xu, C., Misra, S., Srinivasan, P., Ma, S., & others. (2019). When petrophysics meets big data: What can machine do? *SPE Middle East Oil and Gas Show and Conference*.
- Zhang, C., Frogner, C., Araya-Polo, M., & Hohl, D. (2014). Machine-learning based automated fault detection in seismic traces. *76th EAGE Conference and Exhibition 2014*, 2014(1), 1–5.
- Zhao, T., Jayaram, V., Roy, A., & Marfurt, K. J. (2015). A comparison of

classification techniques for seismic facies recognition. *Interpretation*, 3(4), SAE29--SAE58.



Dokumen ini adalah Arsip Miik :
Perpustakaan Universitas Islam Riau