



**SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON  
BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE  
*RANDOM FOREST***

**SKRIPSI**

**ADITYA NUR'AHYA**

**1910512025**

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI S-1 SISTEM INFORMASI**

**2023**



**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA**

**SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON  
BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE  
*RANDOM FOREST***

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana  
komputer**

**ADITYA NUR'AHYA  
1910512025**

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI S-1 SISTEM INFORMASI**

**2023**

## LEMBAR PERSETUJUAN SIDANG

Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Aditya Nur'ahya

NIM : 1910512025

Program Studi : S1 Sistem Informasi

Judul : Sistem Prediksi Zona Potensial Hidrokarbon Berdasarkan Data  
*Well-Log* Menggunakan Metode *Random Forest*

Telah disetujui untuk diujikan oleh Tim Penguji pada ujian sidang Tugas Akhir sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S-1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.

Mengetahui,



**Helena Nurramdhani Irmanda, S.Pd., M.Kom.**

Ketua Program Studi

Menyetujui,



**Ati Zaidiah, S.Kom., MTI.**

Dosen Pembimbing 1

Ditetapkan di : Jakarta

Tanggal Persetujuan : 23 Desember 2022

## PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri dan sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Aditya Nur'ahya  
NIM : 1910512025  
Program Studi : S1 Sistem Informasi  
Judul : Sistem Prediksi Zona Potensial Hidrokarbon Berdasarkan Data  
*Well-Log Menggunakan Metode Random Forest*  
Tanggal : 23 Desember 2022

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 23 Desember 2022

Yang menyatakan,



(Aditya Nur'ahya)

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK  
KEPENTINGAN AKADEMIS**

---

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Aditya Nur’ahya  
NIM : 1910512025  
Fakultas : Ilmu Komputer  
Program Studi : S1 Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (*Non - exclusive Royalty Free Right*) atas skripsi saya yang berjudul:

**SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON  
BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE  
*RANDOM FOREST***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media/memformatkan, mengelola dalam bentuk pangkalan data (basis data), merawat dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 23 Desember 2022

Yang menyatakan,



(Aditya Nur’ahya)

## LEMBAR PENGESAHAN

Dengan ini menyatakan bahwa Tugas Akhir berikut:

Nama : Aditya Nur'ahya

NIM : 1910512025

Program Studi : S1 Sistem Informasi

Judul Skripsi : Sistem Prediksi Zona Potensial Hidrokarbon Berdasarkan Data  
*Well-Log* Menggunakan Metode *Random Forest*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai persyaratan untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.



**Dr. Widya Cholil, S.Kom., M.I.T.**

Penguji I



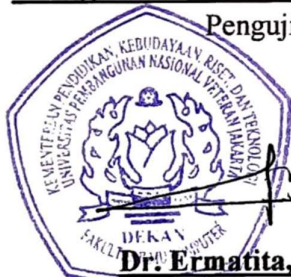
**Kraugusteeliana, SKom., MKom., MM.**

Penguji II



**Ati Zaidiah, S.Kom., M.TI.**

Pembimbing



**Dr. Ermatita, M.Kom.**

Dekan Fakultas Ilmu Komputer



**Helena Nurramdhani Irmanda, S.Pd., M.Kom.**

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 6 Januari 2023



# SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE *RANDOM FOREST*

Aditya Nur'ahya

## ABSTRAK

Kegiatan interpretasi data geofisika merupakan aktivitas analisis yang dilakukan oleh ahli geofisika dalam menginterpretasikan data hasil akuisisi dengan data – data yang berkaitan. Sehingga, mendapatkan pendugaan awal untuk menjelaskan keadaan bawah permukaan. Penelitian ini dilakukan untuk membuat prediksi yang dapat membantu mempercepat dugaan awal kandungan hidrokarbon yang dilakukan oleh seorang ahli geofisika sebagai kegiatan interpretasi. Kegiatan yang interpretasi membutuhkan kecepatan dan ketelitian yang tinggi sehingga dapat menghasilkan hasil analisis yang tepat. Penerapan konsep *data mining* digunakan untuk mengklasifikasikan kandungan hidrokarbon pada lapisan tanah yang di prediksi menggunakan *data well-log*. Salah satu algoritma klasifikasi data mining yang digunakan pada penelitian ini yaitu algoritma *Random Forest*. Algoritma ini digunakan untuk mengklasifikasikan kandungan hidrokarbon. Pada penelitian ini digunakan *dataset* yang didapatkan dari situs resmi *National Offshore Petroleum Information Management System* (NOPIMS) dengan *dataset* awal berjumlah 28.378 data dan setelah dibersihkan data yang digunakan berjumlah 4.556 data, kemudian data dilakukan pembagian menjadi data latih dan data uji untuk pelatihan model. Pada penelitian ini dilakukan evaluasi model yang mencakup pengukuran akurasi, presisi, nilai *recall*, nilai *f1-measure*, dan nilai AUC score yang dilakukan dengan mengukur setiap perbandingan rasio pembagian data. Dengan perbandingan rasio 80- 20 nilai akurasi yang didapatkan sebesar 0,9923 atau 99,23%, nilai presisi 0,97 atau 97%, nilai *recall* 0,98 atau 98%, nilai *f1-measure* 0,97 atau 97%, dan nilai AUC Score 0,9919. Diharapkan implementasi model pada sistem dapat membantu ahli geofisika dalam menginterpretasikan *data well-log* dalam menentukan zona potensial hidrokarbon.

**Kata kunci:** Prediksi, Hidrokarbon, *Random Forest*

**SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON  
BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE  
*RANDOM FOREST***

**Aditya Nur'ahya**

***ABSTRACT***

*The activity of interpreting geophysical data is an analytical activity carried out by geophysicists in interpreting the acquired data with related data. Thus, obtaining initial estimates to explain the subsurface conditions. This research was conducted to make predictions that could help speed up the initial prediction of the hydrocarbon content made by a geophysicist as an interpretation activity. Interpretation activities require high speed and accuracy so as to produce the right analysis results. The application of the concept of data mining is used to classify the hydrocarbon content in the soil layer which is predicted using well-log data. One of the data mining classification algorithms used in this study is the Random Forest algorithm. This algorithm is used to classify the hydrocarbon content. In this study, datasets obtained from the official website of the National Offshore Petroleum Information Management System (NOPIMS) were used with an initial dataset of 28,378 data and after cleaning the data used amounted to 4,556 data, then the data was divided into training data and test data for model training. In this study, an evaluation of the model was carried out which included measurements of accuracy, precision, recall value, f1-measure value, and AUC score value which was carried out by measuring each ratio of data sharing ratios. With a ratio of 80-20, the accuracy value obtained is 0.9923 or 99.23%, the precision value is 0.97 or 97%, the recall value is 0.98 or 98%, the f1-measure value is 0.97 or 97%, and AUC Score value of 0.9919. It is hoped that the implementation of the model in the system can assist geophysicists in interpreting well-log data in determining potential hydrocarbon zones.*

***Keyword: Prediction, Hydrocarbon, Random Forests***



## KATA PENGANTAR

Alhamdulillahirabbil'alamin, segala puji bagi Allah *azza wajalla* yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan tugas akhir ini dengan judul “Sistem Prediksi Zona Potensial Hidrokarbon Berdasarkan Data *Well-Log* Menggunakan Metode *Random Forest*” tepat pada waktunya.

Penyelesaian tugas akhir ini pula tidak lepas dari bimbingan dan dorongan dari berbagai pihak. Dalam kesempatan ini, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Ayah dan Ibu tercinta, selaku orang tua yang selalu mendidik, mengajar, membimbing, merawat, dan membesarkan penulis dengan penuh kasih serta tanpa henti mendoakan keberhasilan penulis di dunia dan akhirat.
2. Ibu Dr. Ermatita, M. Kom., selaku dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta.
3. Ibu Helena Nurramdhani Irmanda, S.Pd., M.Kom. selaku Kepala Program Studi S-1 Sistem Informasi.
4. Ibu Ati Zaidiah, S. Kom., MTI. selaku dosen pembimbing tugas akhir.
5. Ibu Kraugusteeliana, S.Kom., M.Kom., MM. selaku dosen pembimbing akademik.
6. Seluruh Bapak/Ibu dosen Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta atas seluruh ilmu yang bermanfaat yang telah diberikan
7. Semua pihak yang telah memberikan bantuan dan dukungan serta kerja sama yang baik sehingga laporan tugas akhir ini dapat diselesaikan dengan baik.
8. Kepada istriku tercinta yang akan membaca tulisan ini suatu hari nanti, karena telah menanti saya dengan sabar dan menerima saya apa adanya.
9. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting.*

Jakarta, 25 Desember 2022

Penulis

## DAFTAR ISI

LEMBAR JUDUL .....	ii
LEMBAR PERSETUJUAN SIDANG .....	iii
PERNYATAAN ORISINALITAS .....	iv
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS .....	v
LEMBAR PENGESAHAN .....	vi
ABSTRAK .....	vii
<i>ABSTRACT</i> .....	viii
KATA PENGANTAR .....	ix
DAFTAR ISI .....	x
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xiv
DAFTAR RUMUS .....	xv
DAFTAR LAMPIRAN .....	xvi
BAB 1 PENDAHULUAN .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Batasan Penelitian .....	3
1.4. Tujuan Penelitian .....	3
1.5. Manfaat Penelitian .....	4
1.6. Luaran yang diharapkan .....	4
1.7. Sistematika Penulisan .....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1. Prediksi .....	6
2.2. Hidrokarbon .....	6
2.3. <i>Well-Logging</i> .....	7
2.4. <i>Data Mining</i> .....	8
2.5. <i>Algoritma Random Forest</i> .....	8
2.6. Evaluasi Model .....	11
2.7. <i>Python</i> .....	14
2.8. <i>Unified Modeling Language</i> .....	14
2.9. <i>React.js</i> .....	15
2.10. <i>Application Programming Interface (API)</i> .....	15

2.11. <i>Flask</i> .....	15
2.12. <i>Blackbox Testing</i> .....	16
2.13. Penelitian yang Relevan.....	16
BAB III METODOLOGI PENELITIAN.....	21
3.1. Tahapan Penelitian.....	21
3.2. Identifikasi Masalah.....	21
3.3. Studi Pustaka.....	22
3.4. Persiapan Data .....	22
3.5. Pemahaman Data .....	22
3.6. <i>Pre-Processing</i> .....	23
3.6.1. <i>Data Cleansing</i> .....	23
3.6.2. <i>Data Selection</i> .....	23
3.6.3. <i>Data Transformation</i> .....	23
3.7. Pembagian Data .....	24
3.8. Implementasi Model Algoritma <i>Random Forest</i> .....	24
3.9. Evaluasi Model .....	24
3.10. Implementasi Sistem.....	24
3.11. Pengujian Sistem.....	25
3.12. Alat Bantu Penelitian .....	25
3.13. Jadwal Rencana Penelitian.....	26
BAB IV PEMBAHASAN.....	27
4.1. <i>Dataset</i> .....	27
4.2. <i>Data Pre-Processing</i> .....	29
4.2.1. <i>Data Cleansing</i> .....	30
4.2.2. <i>Data Selection</i> .....	32
4.2.3. <i>Data Transformation</i> .....	34
4.2.4. Pembagian Data .....	34
4.3. Implementasi Algoritma <i>Random Forest</i> .....	35
4.4. Evaluasi Model .....	38
4.4.1. Evaluasi Model dengan Rasio 70% - 30%.....	39
4.4.2. Evaluasi Model dengan Rasio 80% - 20%.....	42
4.4.3. Evaluasi Model dengan Rasio 90% - 10%.....	45
4.5. Rancangan Sistem Prediksi.....	48
4.6. Implementasi Sistem Prediksi.....	50

4.6.1.	Halaman <i>Login</i> atau <i>Sign In</i> .....	51
4.6.2.	Halaman Registrasi .....	52
4.6.3.	Halaman Beranda ( <i>Home</i> ).....	53
4.6.4.	Halaman Input Data .....	54
4.6.5.	Halaman Hasil Prediksi.....	55
4.6.6.	<i>Blackbox Testing</i> .....	55
BAB V PENUTUP.....		57
5.1.	Kesimpulan .....	57
5.2.	Saran .....	58
DAFTAR PUSTAKA .....		59
RIWAYAT HIDUP.....		63
LAMPIRAN.....		65

## DAFTAR TABEL

Tabel 1 <i>Confusion Matrix</i> .....	11
Tabel 2 Penelitian yang relevan .....	17
Tabel 3 Jadwal Rencana Penelitian.....	26
Tabel 4 Daftar Variabel pada Dataset .....	27
Tabel 5 Perbandingan Rasio Pembagian Data Latih dan Data Uji .....	35
Tabel 6 Parameter Pelatihan Model Algoritma.....	35
Tabel 7 <i>Confusion Matrix</i> Percobaan Pertama .....	39
Tabel 8 Presisi Percobaan Pertama .....	39
Tabel 9 <i>Recall</i> Percobaan Pertama.....	40
Tabel 10 <i>Confusion Matrix</i> Percobaan Kedua .....	42
Tabel 11 Presisi Percobaan Kedua.....	42
Tabel 12 <i>Recall</i> Percobaan Kedua .....	43
Tabel 13 <i>Confusion Matrix</i> Percobaan Ketiga .....	45
Tabel 14 Presisi Percobaan Ketiga.....	45
Tabel 15 <i>Recall</i> Percobaan Ketiga .....	46
Tabel 16 Evaluasi Model .....	48
Tabel 17 <i>Blackbox Testing</i> .....	55

## DAFTAR GAMBAR

Gambar 1 Konsep <i>Random Forest</i> .....	9
Gambar 2 Alur Penelitian.....	21
Gambar 3 <i>Dataset</i> Awal.....	30
Gambar 4 <i>Dataset</i> Berisi Nilai Kosong bernilai -999.25.....	31
Gambar 5 <i>Dataset</i> yang sudah diisi dengan nilai NaN .....	31
Gambar 6 <i>Dataset</i> Bersih.....	32
Gambar 7 Pemilihan Variabel Hipotesis Awal menggunakan <i>feature_importances</i> .....	33
Gambar 8 <i>Dataset</i> dengan kolom yang diseleksi.....	34
Gambar 9 Contoh Pohon Model <i>Decision Tree</i> .....	36
Gambar 10 Sampel <i>Bootstrapping</i> pada <i>Dataset</i> .....	37
Gambar 11 <i>Dataset Features Importance</i> .....	38
Gambar 12 Grafik <i>ROC Curve</i> Percobaan Pertama.....	41
Gambar 13 Grafik <i>ROC Curve</i> Percobaan Kedua .....	44
Gambar 14 Grafik <i>ROC Curve</i> Percobaan Ketiga .....	47
Gambar 15 Rancangan Use Case Sistem Prediksi Zona Potensial Hidrokarbon..	49
Gambar 16 Rancangan <i>Component Diagram REST API</i> untuk <i>Frontend</i> dan <i>Backend</i> .....	49
Gambar 17 Menjalankan modul <i>Frontend Website</i> Sistem Prediksi dengan <i>React.js</i> .....	50
Gambar 18 Menjalankan modul <i>Backend API</i> Sistem Prediksi dengan <i>Framework</i> <i>Flask</i> menggunakan <i>Python</i> .....	51
Gambar 19 Halaman <i>Login</i> atau <i>Sign In</i> .....	51
Gambar 20 Halaman <i>Register</i> atau Registrasi .....	52
Gambar 21 Halaman <i>Home</i> atau Beranda .....	53
Gambar 22 Bagian <i>About Us</i> .....	53
Gambar 23 Halaman <i>Form Input Data</i> .....	54
Gambar 24 Halaman <i>Form Input Data</i> dengan Hasil Prediksi .....	55

## DAFTAR RUMUS

Rumus 1 Rumus <i>Information Gain</i> .....	10
Rumus 2 <i>Information Gain</i> Kelas Target.....	10
Rumus 3 <i>Information Gain</i> Partisi Data.....	10
Rumus 4 Perumusan Nilai Accuracy .....	12
Rumus 5 Perumusan Nilai <i>Precision</i> .....	13
Rumus 6 Perumusan Nilai <i>Recall</i> .....	13
Rumus 7 Perumusan Nilai <i>F-1 Score Measure</i> .....	13

## DAFTAR LAMPIRAN

Lampiran 1 Dataset Awal.....	66
Lampiran 2 Dataset setelah Pre-processing .....	91
Lampiran 3 Summary Plot sebagai dasar pelabelan data.....	117
Lampiran 4 <i>Python Code</i> Implementasi Model setelah Data <i>Pre-Processing</i> ....	121
Lampiran 5 <i>Python Code File Rest API Routes</i> (routes.py).....	122
Lampiran 6 Wawancara Validasi Data dengan Ahli.....	123
Lampiran 7 CV Ilham Diaz Rahmat Nugroho .....	125
Lampiran 8 Hasil Turnitin.....	128