

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

Aditya Nur'ahya<sup>1</sup>, Ati Zaidiah, S.Kom., MTI.<sup>2</sup>  
S1 Sistem Informasi / Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional Veteran Jakarta<sup>1,2</sup>  
adityan@upnvj.ac.id<sup>1</sup>, atizaidiah@upnvj.ac.id<sup>2</sup>

**Abstrak.** Kegiatan interpretasi data geofisika merupakan aktivitas analisis yang dilakukan oleh ahli geofisika dalam menginterpretasikan data hasil akuisisi. Pendugaan awal dilakukan 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. Penerapan konsep *Data Mining* digunakan untuk mengklasifikasikan kandungan hidrokarbon pada lapisan tanah dilakukan dengan menerapkan Algoritma *Random Forest* yang digunakan untuk mengklasifikasikan kandungan hidrokarbon. Penerapan *Data Mining* dilakukan dengan konsep CRISP-DM dalam pengolahan dataset hingga pengembangan sistem prediksi yang dapat digunakan. Pada penelitian ini digunakan dataset yang didapatkan dari situs resmi NOPIMS dengan dataset awal berjumlah 28.378 data dan data yang digunakan berjumlah 4.556 data. Dengan perbandingan rasio 80-20 nilai akurasi yang didapatkan sebesar 0,9923 atau 99,23%, nilai presisi 0,97, nilai *recall* 0,98, nilai *f1-measure* 0,97, dan nilai *AUC Score* 0,9919.

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

## 1 Pendahuluan

Hidrokarbon memiliki peranan yang sangat penting dalam pemenuhan kebutuhan energi saat ini. Terlebih kebutuhan manusia untuk energi semakin meningkat seiring perkembangan zaman. Kebutuhan akan energi yang semakin tinggi ini memerlukan langkah yang efektif guna meningkatkan produktivitas minyak dan gas bumi. Peningkatan produksi dapat dilakukan melalui kegiatan eksplorasi.

Metode *Well-Logging* merupakan metode geofisika yang bekerja dengan cara merekam respon dari alat *log*. *Logging* bekerja dengan cara memasukan alat log ke dalam sumur sehingga mendapat respon berupa sifat fisik dari batuan dan fluida. Hasil yang didapat dari pengukuran logging berupa kurva yang mengandung informasi keadaan bawah permukaan. Dalam melakukan pekerjaan logging kurva tersebut perlu dilakukan interpretasi berdasarkan data-data dan kurva yang telah diakuisisi. Hal ini dilakukan guna melakukan analisis kandungan bawah permukaan, menentukan lapisan prospek hidrokarbon, serta sebaran zona hidrokarbon. Sementara itu, potensi zona hidrokarbon diukur dengan memperkirakan saturasi airnya (*Sw*), saturasi minyak (*So*) dan parameter petrofisika lainnya. Nilai log yang merepresentasikan kandungan hidrokarbon Gas dalam lapisan batuan yaitu *Water Saturation* (*Sw*) dimana nilai *Sw* untuk hidrokarbon gas berkisar antar 0 – 25% [1].

Metode yang digunakan dalam interpretasi data *well-log* semakin beragam, sejalan dengan perkembangan teknologi. Analisa *well-log* membutuhkan kecepatan dan ketelitian yang tinggi agar eksplorasi hidrokarbon dapat berjalan dengan optimal. Penggunaan metode *data mining* diharapkan dapat membantu peran geofisikawan dalam melakukan interpretasi data *well-log*. Metode *data mining* memungkinkan pengolahan data dengan membaca pola dan kombinasi pada dataset yang berukuran besar dalam waktu singkat dengan ketepatan yang tinggi.

Metode *data mining* digunakan untuk membantu menganalisa data dalam jumlah yang banyak sehingga dapat mengefisienkan waktu yang digunakan geofisikawan untuk menginterpretasikan data *well-log*. *Data mining* adalah proses menemukan informasi baru dengan cara mencari pola atau aturan yang ada dalam kumpulan data yang besar. Salah satu metode data mining yang dapat digunakan yaitu klasifikasi yang digunakan untuk melakukan prediksi [5]. *Random Forest* merupakan salah satu algoritma yang dapat digunakan untuk membantu mengklasifikasikan data dalam proses interpretasi data *well-log*. Interpretasi data *well-log* yang dihasilkan dapat dilakukan prediksi dengan menggunakan metode *data mining*. Menggabungkan pengetahuan domain berbasis petrofisika dengan pembelajaran mesin memungkinkan algoritma dapat diskalakan ke dalam kumpulan data yang lebih besar, meningkatkan efisiensi, dan membantu interpretasi [4].

Hal ini sangat membantu penerapan interpretasi data log dalam menentukan zona potensial hidrokarbon seperti penelitian yang dilakukan oleh Iskandar, dkk. [6].

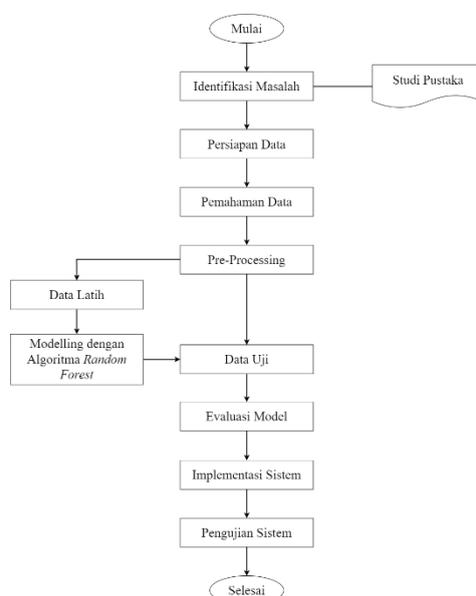
Pada penelitian yang dilakukan oleh Ilham Diaz Rahmat Nugroho dilakukan pengukuran estimasi parameter petrofisika yang menghasilkan luaran bahwa pengukuran data petrofisika dengan penerapan metode algoritma *data mining* efektif dalam mengestimasi parameter petrofisika dari data log[8].

Pada penelitian yang dilakukan oleh Agustyaningrum, dkk [2] didapatkan bahwa algoritma *Random Forest* memiliki nilai akurasi sebesar 89,9%, algoritma *Support Vector Machine (SVM)* memiliki nilai akurasi 88,02%, dan algoritma *Naive Bayes* sebesar 80,9%. Pada penelitian tersebut digunakan data numerik yang relevan dengan data *well-log* yang digunakan pada penelitian ini. Hal ini menunjukkan perbandingan kuat dan keunggulan dari algoritma *Random Forest* dalam memprediksi nilai dari data numerik.

Berdasarkan latar belakang yang peneliti uraikan di atas. Penelitian ini akan memprediksi zona potensial hidrokarbon dengan memanfaatkan data *well-log* dan menerapkan algoritma *Random Forest* yang diterapkan pada sistem prediksi sederhana yang dapat dipergunakan untuk memprediksi[11]. Penerapan sistem prediksi menggunakan konsep *REST API* untuk sarana komunikasi *frontend* dan *backend website*, dimana *frontend* dikembangkan dengan *React.js* dan *backend* dikembangkan dengan *framework flask* yang menyediakan beberapa fungsi beserta kumpulan kode yang dapat digunakan kembali untuk mendukung pembuatan *website*. [3]

## 2 Metode Penelitian

Penelitian ini dibuat dengan mengadaptasi tahapan *Data Mining* pada metode CRISP-DM atau *Cross-Industry Standard Process for Data Mining*. Secara bertahap alur kerangka kerja pada metode penelitian ini dijelaskan pada diagram berikut.



Gambar 1 Metode Penelitian

### 2.1 Persiapan Data

Data pada penelitian ini bersumber dari dataset yang dipublikasikan secara *open-source* melalui situs *website* <https://nopims.dmp.wa.gov.au/>. Dataset yang telah didapatkan kemudian dilakukan validasi dengan melakukan wawancara kepada pakar atau Ahli Geofisika. Wawancara yang dilakukan dengan Ilham Diaz Rahmat Nugroho yang bekerja sebagai Ahli Geofisika Muda yang bekerja di Pertamina Hulu Rokan 2.

### 2.2 Pemahaman Data

Variabel yang dipublikasikan diantaranya *depth* (pengukuran kedalaman), *badhole* (kategori lubang bor semakin besar karena runtuh), *coal (Coal Volume)*, *cali (Caliper)*, *gr (Gamma Ray)*, *dct (Compressional Sonic)*, *dts (Shear Sonic)*, *neut (Neutron Porosity)*, *pef (Photo-Electric Effect)*, *dens (Density)*, *rt (Resistivity)*, *rxo (Invaded Zone Resistivity)*, *perm (Permeability)*, *phie (Effective Porosity)*, *phit (Total Porosity)*, *swe (Effective Water Saturation)*, *swt (Total Water Saturation)*, dan *vsh (Volume of Wet Shale)*. Serta variabel kategori yaitu *Label* sebagai target prediksi.

## 2.3 Data Pre-Processing

Pada tahapan ini, data akan dibersihkan dengan memeriksa informasi (fitur) yang hilang dan mereduksi duplikasi data yang ada sehingga menghasilkan data yang layak untuk dilakukan pemodelan[10]. Beberapa metode pengolahan data yang dilakukan diantaranya *Data Cleansing*, *Data Selection*, dan *Data Transformation*.

## 2.4 Pembagian Data

Pemrosesan data yang akan digunakan dalam pemodelan yaitu dengan membagi dataset menjadi data latih dan data uji sebagai input data untuk dapat melatih model algoritma *Random Forest*[11]. Penelitian ini akan membagi dataset menjadi beberapa persentase yang berbeda untuk dilakukan beberapa percobaan yang ditujukan untuk mendapatkan akurasi yang tinggi.

## 2.5 Implementasi Algoritma *Random Forest*

Data yang telah melalui tahapan *pre-processing* dan pembagian data akan dilakukan pengimplementasian model dengan algoritma *Random Forest* sebagai proses melatih model yang akan dijadikan model utama prediksi pada sistem.

## 2.6 Evaluasi Model

Evaluasi model algoritma *Random Forest* dilakukan dengan mengevaluasi kemampuan prediksi menggunakan data pelatihan untuk menghasilkan nilai akurasi dari metode yang telah diterapkan dilakukan untuk mengetahui performa model.[7] Data yang telah dibagi baik menjadi data latih maupun data uji akan dianalisis secara mendalam. Keluaran dari proses ini adalah persentase akurasi dari model yang dihasilkan, nilai akurasi, nilai dari precision, recall, dan *f1-score*, nilai AUC, dan *ROC Curve*.

**Tabel 1** *Confusion Matrix* Evaluasi Model

Nilai Aktual	Nilai Prediksi	
	Label 1	Label 2
Label 1	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
Label 2	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

## 2.7 Implementasi Sistem

Implementasi sistem dibangun dengan menggunakan *framework Flask*. *Flask* menggunakan bahasa pemrograman *Python* sehingga mendukung untuk mengimpor model prediksi yang telah dibuat dengan pelatihan model menggunakan data latih. Model yang telah dilatih dikonversi menjadi file berekstensi *.pkl* atau *pickle* yang kemudian akan diimpor ke dalam sistem. Sistem yang dibangun menerapkan konsep *backend website* dengan basis API sebagai sarana komunikasi dengan *frontend website*. Sedangkan untuk *frontend website*, sistem dibangun menggunakan *library Javascript* yaitu *React.js* untuk tampilan antarmuka.

## 2.8 Pengujian Sistem

Pengujian sistem dilakukan tahapan uji menggunakan *blackbox testing*. Pengujian sistem ini berfokus pada informasi yang ditampilkan oleh sistem. Pengujian dilakukan untuk melihat respon sistem dalam skenario uji input pada sistem prediksi. Pengujian ini akan dapat menghasilkan hasil uji skenario yang dapat dievaluasi apakah sistem berjalan sesuai skenario dan dapat digunakan.

# 3 Hasil dan Pembahasan

## 3.1 Dataset

*Dataset* yang digunakan merupakan hasil interpretasi dari proses *well-logging* pada sumur lepas pantai (*offshore*) di ladang minyak Gorgon, lepas pantai negara bagian Barat Australia yang dioperasikan oleh Chevron Australia Pty Ltd. Data well-log ini diinterpretasikan sebagai tahapan pengembangan bor di cekungan Northern Carnarvon, Barat Laut Australia Barat. Dataset ini di rilis pada 7 Januari 2019 pada situs resmi *National Offshore Petroleum Information Management System*. Dataset yang diperoleh berupa file berekstensi *.las*. Peneliti mendapatkan data sejumlah 28.378 data dari hasil interpretasi *well-logging*. Dataset ini memiliki 19 (sembilan belas) variabel yang terdiri 18 (delapan belas) variabel dependen dan 1 (satu) variabel independen yang berupa kelas kategori.

### 3.2 Data Pre-processing

Dataset yang didapatkan merupakan file berjenis .las yang tidak langsung digunakan dalam tahapan pemodelan. Untuk itu perlu dilakukan beberapa tahapan *Data Preprocessing*. Pada tahapan ini data yang telah didapatkan terlebih dahulu diubah format ekstensi file menjadi .xlsx dengan menggunakan aplikasi *Microsoft Excel*. Maka didapatkan hasil dataset setelah konversi ekstensi file terdiri dari 28.378 record.

**Tabel 2** Dataset Awal

DEPTH	BAD HOLE	CALI	COAL	DENS	DTC	DTS	GR	NEUT	PEF	PERM	PHIE	PHIT	RT	RXO	SWE	SWT	VSH	LABEL
616.0008	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
616.1532	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
616.3056	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
616.458	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
616.6104	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4940.046	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
4940.198	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
4940.351	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
4940.503	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected
4940.656	0	-999.25	0	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	No HC Detected

Data yang kosong kemudian harus dihilangkan. Dikarenakan data yang kosong masih diinterpretasikan dengan nilai -999.25 maka harus diganti dengan nilai "null" supaya memudahkan untuk menyeleksi data yang kosong.

**Tabel 3** Dataset setelah nilai -999.25 di transformasi menjadi NaN

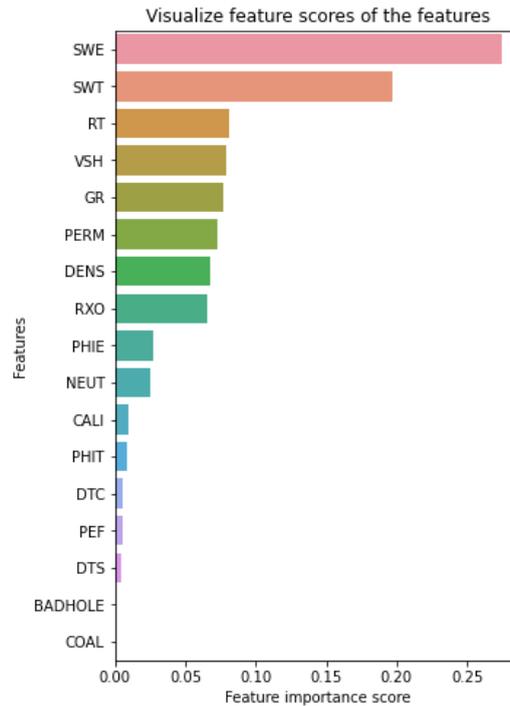
DEPTH	BAD HOLE	CALI	COAL	DENS	DTC	DTS	GR	NEUT	PEF	PERM	PHIE	PHIT	RT	RXO	SWE	SWT	VSH	LABEL
616.0008	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
616.1532	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
616.3056	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
616.458	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
616.6104	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4940.046	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
4940.198	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
4940.351	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
4940.503	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected
4940.656	0	NaN	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	No HC Detected

Kemudian data yang telah diganti kemudian dilakukan pembersihan dengan menghilangkan seluruh baris data yang terdapat nilai kosong atau NaN.

**Tabel 4** Dataset yang telah dibersihkan

DEPTH	BAD HOLE	CALI	COAL	DENS	DTC	DTS	GR	NEUT	PEF	PERM	PHIE	PHIT	RT	RXO	SWE	SWT	VSH	LABEL
4138.117	0.000	9.200	0.000	2.529	81.683	148.503	164.954	0.252	4.415	0.044	0.079	0.102	3.929	3.929	1.000	1.000	0.805	No HC Detected
4138.270	0.000	9.203	0.000	2.538	81.616	148.424	160.507	0.247	4.214	0.032	0.074	0.096	4.493	4.493	1.000	1.000	0.781	No HC Detected
4138.422	0.000	9.205	0.000	2.561	81.149	148.757	159.180	0.254	4.042	0.018	0.063	0.085	5.006	5.006	1.000	1.000	0.773	No HC Detected
4138.574	0.000	9.185	0.000	2.572	80.459	150.031	162.500	0.262	4.379	0.014	0.057	0.080	4.990	4.990	1.000	1.000	0.792	No HC Detected
4138.727	0.000	9.188	0.000	2.571	79.716	151.241	163.587	0.274	4.602	0.015	0.058	0.081	4.416	4.416	1.000	1.000	0.798	No HC Detected
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4898.898	0.000	9.097	0.000	2.616	82.585	156.997	151.947	0.186	3.362	0.005	0.027	0.050	9.060	9.060	1.000	1.000	0.797	No HC Detected
4899.050	0.000	9.090	0.000	2.627	81.606	156.804	166.220	0.171	3.434	0.003	0.016	0.042	12.528	12.528	1.000	1.000	0.885	No HC Detected
4899.203	0.000	9.078	0.000	2.633	79.545	156.171	171.873	0.167	3.656	0.002	0.011	0.038	18.549	18.549	0.608	0.883	0.920	Gross
4899.355	0.000	9.058	0.000	2.633	78.768	152.936	164.605	0.154	3.902	0.002	0.011	0.036	21.138	21.138	0.537	0.862	0.875	Gross
4899.508	0.000	9.046	0.000	2.634	71.119	148.237	140.251	0.130	3.997	0.001	0.011	0.032	22.658	22.658	0.986	0.995	0.726	Gross

Variabel yang akan digunakan untuk pelatihan model algoritma pada tahapan lanjutan yaitu berjumlah 14 (empat belas) variabel, diantaranya cali (*Caliper*), coal (*Coal Volume*), dens (*Density*), dtc (*Compressional Sonic*), dts (*Shear Sonic*), gr (*Gamma Ray*), neut (*Neutron Porosity*), pef (*Photo-Electric Effect*), perm (*Permeability*), phie (*Effective Porosity*), phit (*Total Porosity*), rt (*Resistivity*), rxo (*Invaded Zone Resistivity*), swe (*Effective Water Saturation*), swt (*Total Water Saturation*), dan vsh (*Volume of Wet Shale*).



**Gambar 2** Hipotesis awal menggunakan *feature\_importances* untuk seleksi variabel

Data yang telah dibersihkan untuk pelatihan model pada tahapan sebelumnya diperoleh sebanyak 4556 baris data. Kemudian dataset dibagi ke dalam 2 (dua) bagian untuk pelatihan model, yaitu dibagi menjadi data latih dan data uji. Rincian pembagian data latih dan data uji sebagai berikut.

**Tabel 5** Pembagian Data Latih dan Data Uji berdasarkan rasio data

Persentase Data Latih	Persentase Data Uji	Jumlah Data Latih	Jumlah Data Uji	Total
70%	30%	3189	1367	4556
80%	20%	3644	912	4556
90%	10%	4100	456	4556

### 3.3 Implementasi Algoritma *Random Forest*

Tahapan implementasi model, data latih dan data uji yang telah dilakukan pembagian rasionya kemudian digunakan dalam pelatihan model algoritma *Random Forest*. Algoritma *Random Forest* yang digunakan dilakukan pelatihan model dengan beberapa parameter yang digunakan.

**Tabel 6** Parameter Pelatihan pada implementasi model Algoritma *Random Forest*

Parameter	Nilai
<i>n_estimators</i>	100
<i>bootstrap</i>	<i>True</i>
<i>class_weight</i>	<i>balanced</i>

### 3.4 Evaluasi Sistem

Setiap nilai akurasi, presisi, *recall*, *f1-measure*, dan *AUC Score* yang dihasilkan dari pengujian perbandingan rasio pembagian data disajikan dengan tabel dibawah ini sebagai hasil penerapan algoritma *Random Forest*.

**Tabel 7** Hasil Evaluasi Model dengan percobaan berdasarkan rasio pembagian data

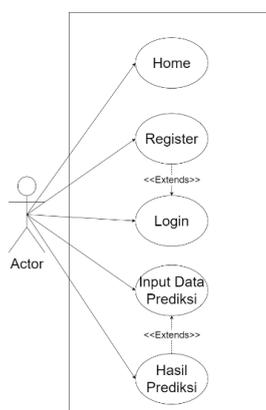
Percobaan Ke-	Rasio Pembagian Data	Akurasi	Presisi	<i>Recall</i>	<i>F1-measure</i>	<i>AUC Score</i>
1	70 – 30	0.9897	0.97	0.97	0.97	0.9874
2	80 – 20	0.9923	0.97	0.98	0.97	0.9919

Percobaan Ke-	Rasio Pembagian Data	Akurasi	Presisi	Recall	F1-measure	AUC Score
3	90 – 10	0.9868	0.96	0.98	0.97	0.9897

Berdasarkan evaluasi model yang disajikan dengan tabel di atas didapatkan bahwa model yang lebih baik dari keseluruhan evaluasi model yaitu model yang dilatih dan divalidasi dengan data latih dan data uji rasio pembagian 80 – 20. Dengan 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 0,99. Maka model dengan rasio pembagian data dan evaluasi sebagaimana dijelaskan di atas yang akan diimpor untuk diterapkan dalam pembangunan sistem prediksi sederhana untuk memprediksi zona potensial hidrokarbon.

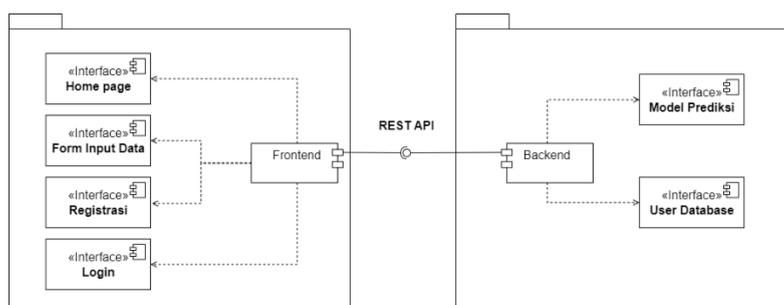
### 3.5 Rancangan Sistem Prediksi

Sistem prediksi dirancang untuk dapat menerima input dari pengguna menggunakan tampilan *frontend website React.js* yang kemudian akan data yang akan di prediksi diteruskan ke sistem backend *Flask*.



**Gambar 3** Use Case Diagram Sistem Prediksi

Komunikasi *frontend* dan *backend website* ini diterapkan penggunaan *Application Programming Interface (API)* yang lebih memudahkan dan efisien untuk mengirim data input pengguna dan data hasil prediksi ke dalam *frontend website*. Perancangan sistem prediksi digambarkan melalui *Component Diagram* dari komponen *frontend website* dan *backend website*.



**Gambar 4** Component Diagram REST API untuk Frontend dan Backend

### 3.6 Implementasi Sistem Prediksi

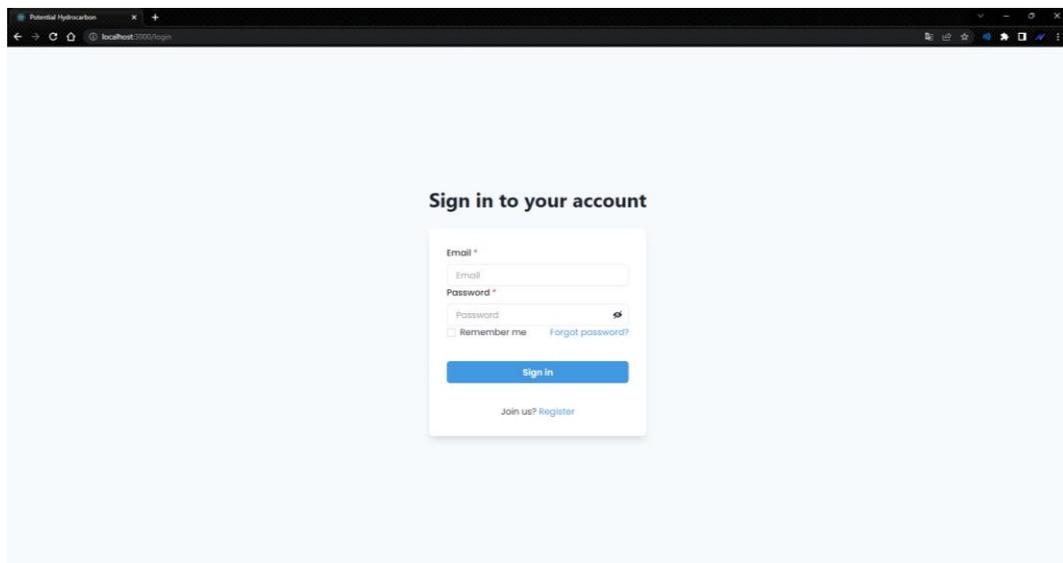
Sistem yang telah dirancang akan diimplementasikan dengan menggunakan *React.js* dan *Framework Flask*. Dalam tahapan ini dilakukan pengembangan sistem prediksi dengan 2 (dua) modul *website* yang saling terhubung dengan API yang disematkan pada *backend website*. *Frontend website* dikembangkan dengan *library javascript* yaitu *React.js* yang memudahkan peneliti untuk dapat langsung menerapkan sistem antarmuka yang efisien dibandingkan dengan *library* lainnya.

```
Compiled successfully!  
  
You can now view skripsi in the browser.  
  
Local:           http://localhost:3000  
On Your Network: http://192.168.18.100:3000  
  
Note that the development build is not optimized.  
To create a production build, use npm run build.  
  
webpack compiled successfully
```

**Gambar 5** Modul *Frontend* menggunakan *React.js*

1. Halaman *Login*

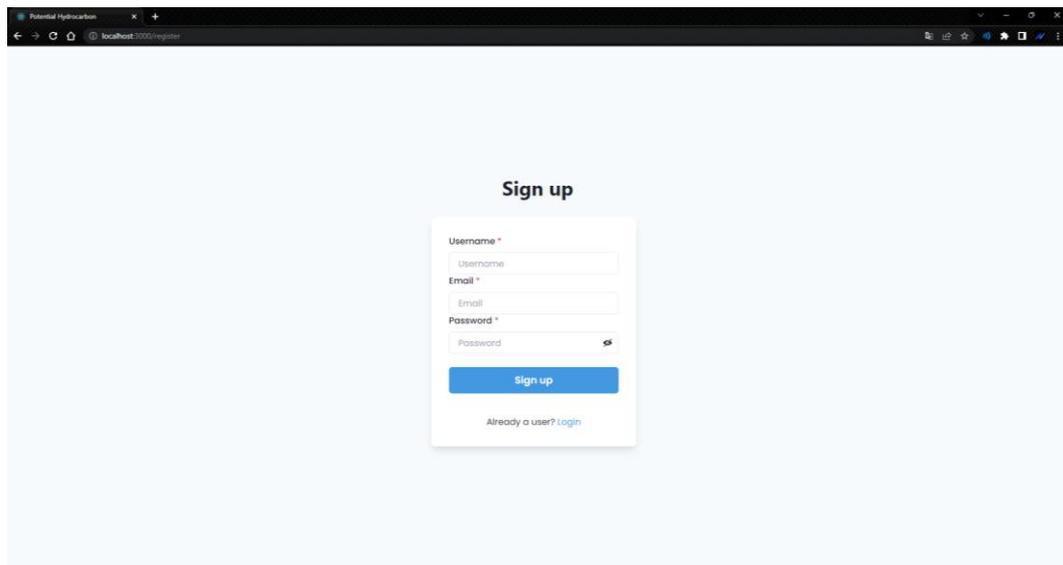
Pada Gambar 6 di bawah ini menampilkan halaman dari antarmuka pengguna untuk melakukan *input* data pengguna untuk mengakses sistem prediksi zona potensial hidrokarbon dengan akun yang telah diregistrasi.



**Gambar 6** Halaman *Login*

## 2. Halaman Registrasi

Pada Gambar 7 di bawah ini menampilkan halaman dari antarmuka pengguna untuk melakukan input data pengguna yang ditujukan untuk membuat akun pada sistem prediksi zona potensial hidrokarbon.



The screenshot shows a web browser window with the URL 'localhost:3000/register'. The page has a light blue background and a central white registration form. The form is titled 'Sign up' and contains the following fields: 'Username \*' with a placeholder 'Username', 'Email \*' with a placeholder 'Email', and 'Password \*' with a placeholder 'Password' and a toggle icon. Below the fields is a blue 'Sign up' button. At the bottom of the form, there is a link that says 'Already a user? Login'.

**Gambar 7** Halaman Registrasi

## 3. Halaman Beranda

Pada Gambar 8 di bawah ini menampilkan halaman utama dari antarmuka pengguna pada sistem prediksi zona potensial hidrokarbon.



**Gambar 8** Halaman Beranda

#### 4. Halaman Input Data

Pada Gambar 9 di bawah ini menampilkan halaman dari antarmuka pengguna untuk melakukan input data yang akan dilakukan prediksi dengan menggunakan sistem prediksi zona potensial hidrokarbon.

**Potential Hydrocarbon Prediction**

**Form Prediction**

Depth 3872.063	Caliper (CALI) 8.799	Density (DENS) 2.505
Compressional Sonic (DTC) 88.634	Shear Sonic (DTS) 151.319	Gamma Ray (GR) 139.895
Neutron Porosity (NEUT) 0.342	Permeability (PERM) 0.124	Effective Porosity (PHE) 0.097
Total Porosity (PHIT) 0.122	Resistivity (RT) 8.007	Invaded Zone Resistivity (RXO) 8.718
Effective Water Saturation (SWE) 0.478	Total Water Saturation (SWT) 0.583	Volume of Wet Shell (VSH) 0.858

**Variable Explanation**

- Depth
- Caliper (CALI)

**Gambar 9** Halaman Prediksi

#### 5. Tampilan Hasil Prediksi

Pada Gambar 10 di bawah ini menampilkan halaman dari antarmuka pengguna untuk yang menampilkan hasil prediksi beserta nilai probabilitas potensi hidrokarbon dan penjelasan mengenai variabel yang digunakan dalam memprediksi zona potensial hidrokarbon.

**Potential Hydrocarbon Prediction**

**Results**

Potential Hydrocarbon is **gas**  
Probability of Potential Hydrocarbon 0.56

**Variable Explanation**

Depth	Kedalaman reservoir yang terukur.
Caliper (CALI)	
Density (DENS)	
Compressional Sonic (DTC)	
Shear Sonic (DTS)	
Gamma Ray (GR)	
Neutron Porosity (NEUT)	
Permeability (PERM)	
Effective Porosity (PHE)	
Total Porosity (PHIT)	
Resistivity (RT)	
Invaded Zone Resistivity (RXO)	
Effective Water Saturation (SWE)	
Total Water Saturation (SWT)	
Shale Volume (VSH)	

**Gambar 10** Halaman Menampilkan Hasil Prediksi

### 3.7 Blackbox Testing

Metode *blackbox testing* digunakan untuk menguji sistem prediksi yang ditujukan untuk melihat respon dari sistem ketika dilakukan input dari pengguna. Skenario uji dilakukan dengan melihat kesesuaian antara input yang diberikan pengguna dengan output yang dihasilkan berdasarkan aturan dari model yang telah ditetapkan di awal pembuatan sistem, yang akan memberikan feedback yang dihasilkan berupa informasi kesimpulan dimana hasil pengujian sudah sesuai atau masih perlu dilakukan perbaikan.[9]

Tabel 8 Hasil *Blackbox Testing* pada sistem prediksi

Modul	Input	Proses	Output	Status
Halaman Beranda ( <i>Home</i> )	Mengakses laman URL	Menampilkan tampilan halaman awal	Halaman Awal ditampilkan	Berhasil
Halaman <i>Input Data</i> Prediksi	Klik Tombol <i>Get Prediction</i>	Mengarahkan ke halaman <i>input data</i>	Halaman <i>Input Data</i> ditampilkan	Berhasil
	Klik Tombol <i>Submit</i>	Mengirim data untuk diprediksi	Hasil Prediksi ditampilkan	Berhasil
Halaman <i>Login</i>	Klik Tombol <i>Reset</i>	Mengatur ulang nilai <i>input</i>	Data input dikosongkan	Berhasil
	Klik Tombol <i>Sign In</i>	Mengirim data <i>user</i> untuk <i>login</i> aplikasi	Halaman utama ditampilkan	Berhasil
Halaman Registrasi	Klik Tombol <i>Sign Up</i>	Mengirim data <i>user</i> untuk registrasi aplikasi	Halaman login ditampilkan	Berhasil

Dengan dilakukan pengujian yang disajikan pada tabel di atas, semua modul dan fungsi sistem menghasilkan *output* yang diinginkan dari perancangan.

#### 4 Simpulan dan Saran

Berdasarkan penelitian yang telah dilakukan, terdapat beberapa kesimpulan yang didapatkan dari penelitian ini, diantaranya:

- Algoritma *Random Forest* dapat secara optimal diterapkan dalam prediksi zona potensial hidrokarbon berdasarkan data hasil *well-log* yang telah diinterpretasikan. Algoritma *Random Forest* secara optimal memprediksi menggunakan 14 variabel yang merupakan parameter *well-log* yang meliputi cali (*Caliper*), coal (*Coal Volume*), dens (*Density*), dtc (*Compressional Sonic*), dts (*Shear Sonic*), gr (*Gamma Ray*), neut (*Neutron Porosity*), pef (*Photo-Electric Effect*), perm (*Permeability*), phie (*Effective Porosity*), phit (*Total Porosity*), rt (*Resistivity*), rxo (*Invaded Zone Resistivity*), swe (*Effective Water Saturation*), swt (*Total Water Saturation*), dan vsh (*Volume of Wet Shale*).
- Evaluasi model algoritma yang dilakukan terhadap percobaan pelatihan model algoritma yang dibangun lebih optimal dengan perbandingan rasio pembagian data latih dan data uji yang terdiri dari 80% data latih dan 20% data uji. Evaluasi meliputi pengukuran nilai akurasi sebesar 99,23% dari nilai maksimum 100%, nilai presisi sebesar 0,97 dari nilai maksimum 1,00, nilai *recall* sebesar 0,98 dari nilai maksimum 1,00, nilai *f1-measure* sebesar 0,97 dari nilai maksimum 1,00, dan *AUC Score* yang merupakan ukuran dalam keberhasilan atau sensitivitas model terhadap data yang akan diprediksi sebesar 0,99 dari nilai maksimum 1,00.
- Sistem prediksi zona potensial hidrokarbon dibangun menggunakan *React.js* untuk *frontend website* dan *framework flask* untuk *backend website*. Kedua sistem saling berkomunikasi dengan API yang dibangun menggunakan *framework Flask* yang menggunakan bahasa pemrograman *python*. Modul API ini digunakan untuk menyimpan model algoritma *Random Forest* yang telah diuji validasi dengan hasil evaluasi terbaik. Sehingga sistem prediksi zona potensial hidrokarbon yang dibangun dapat menampilkan hasil prediksi yang ditampilkan pada *frontend website*.

Beberapa saran yang dapat menjadi pengembangan dari penelitian ini diantaranya.

- Dapat melakukan perbandingan lain dengan penerapan algoritma pembelajaran mesin lainnya yang serupa menerapkan konsep *ensemble* seperti *AdaBoost* dan *GradientBoosting* untuk mengetahui perbedaan hasil dan pola pada dataset yang sama.
- Dapat mengembangkan fitur sistem untuk dapat menerima inputan berupa file *dataset* lengkap tentang data *well-log*.
- Dapat mengembangkan fitur visualisasi pada data hasil prediksi dengan inputan file *dataset well-log* yang lengkap.

#### Referensi

- Abduh, M. et al. 2020. "Analisis Petrofisika Untuk Penentuan Zona prospek Reservoir Hidrokarbon di Lapangan 'X' cekungan banggai," KONSTAN - JURNAL FISIKA DAN PENDIDIKAN FISIKA, 5(1), hlm. 1–9.
- Agustyaningrum, C., Gata, W., Nurfalih, R., Radiyah, U. dan Maulidah, M., 2020. Komparasi Algoritma Naive Bayes, Random Forest dan SVM Untuk Memprediksi Niat Pembelanja Online. Jurnal Informatika, 20(2), hlm. 164-173.

- [3] Amalia, A., Zaidiah, A. dan Isnainiyah, I.N, 2022. “Prediksi Kualitas udara menggunakan algoritma K-nearest neighbor,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 7(2), hlm. 496–507.
- [4] Dwihusna, Nadima, 2020. *Seismic And Well Log Based Machine Learning Facies Classification In The Panoma-Hugoton Field, Kansas And Raudhatain Field, North Kuwait*. 2020 - Mines Theses & Dissertations.
- [5] Etriyanti, E., Syamsuar, D., dan Kunang, Y.N., 2020. Implementasi Data Mining Menggunakan Algoritme Naïve Bayes Classifier dan C4.5 untuk Memprediksi Kelulusan Mahasiswa. *Telematika*. 13 (1). hlm. 56-67.
- [6] Iskandar, I.N.E.N., Rohmana, R.C. and Atmoko, W. 2019. “Interpretasi Seismik Dalam Menentukan Zona Potensial Hidrokarbon Di Formasi Jatibarang Dan Talangakar, Sub-Cekungan Jatibarang, C Ekungan Jawa Barat Utara,” *Prosiding TAU SNAR-TEK Seminar Nasional Rekayasa dan Teknologi*, 1(1), hlm. 173–177.
- [7] Larassati D, Zaidiah A, dan Afrizal S, 2022. Sistem Prediksi Penyakit Jantung Koroner Menggunakan Metode Naïve Bayes. “*JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*”. 7(2). hlm. 533-546.
- [8] Nugroho, I. D. R., 2021. Estimasi Parameter Petrofisika Pada Log Sumur Menggunakan Algoritma Random Forest, XGBoost, dan Support Vector Regression (SVR) di Lapangan Kenali Asam (Kas) Cekungan Sub-Jambi, Jambi. Skripsi. Universitas Gadjah Mada. Yogyakarta. <http://etd.repository.ugm.ac.id/penelitian/detail/203861>.
- [9] Widiastiwi Y, Zaidiah A, Indriana I, 2020. Pengujian Model Aplikasi User Interface E-Anjal Dengan Menggunakan Metode Black Box. *Jurnal Informatik*. 16(2) hlm. 106-115.
- [10] Depari, D. H., Widiastiwi, Y., & Santoni, M. M. (2022). Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *Informatik: Jurnal Ilmu Komputer*, 18(3), 239-248.
- [11] Rahmawati, A., & Hananto, B. (2018). PREDIKSI PENENTUAN PROGRAM STUDI MENGGUNAKAN ALGORITMA K-NN PADA PELAJAR SMAN 6 KOTA DEPOK JURUSAN IPA. *PROSIDING SEINASI-KESI*, 1(1), 192-197.