

**TUGAS 1 2**

**DATA SCIENCES USE CASE: TECHNOLOGY FOR OIL AND  
GAS INDUSTRY**

**DISUSUN OLEH**

Muhammad Luthfi Fahrezi

**DEPOK 2022**

Industri teknologi semakin meningkat setiap tahunnya. Teknologi bahkan merambat hingga berbagai bidang, salah satunya bidang eksplorasi energi. Teknologi *well logging* sangatlah penting pada bidang eksplorasi energi. *Well logging* adalah Teknik merekam karakteristik petrofisik pada area bawah permukaan. *Well logging* bertujuan untuk menentukan potensial hidrokarbon pada sebuah lapangan hidrokarbon.

Teknologi *well logging* juga berkembang dengan diterapkannya pendekatan *machine learning*. Pendekatan *machine learning* diperlukan mengingat untuk mendapatkan data sumur yang baik secara kualitas dan kuantitas akan memakan biaya yang sangat mahal. Sedangkan *machine learning* dapat memprediksi data sumur yang kurang. Sehingga jika hal tersebut dilakukan biaya eksplorasi hidrokarbon akan berkurang.

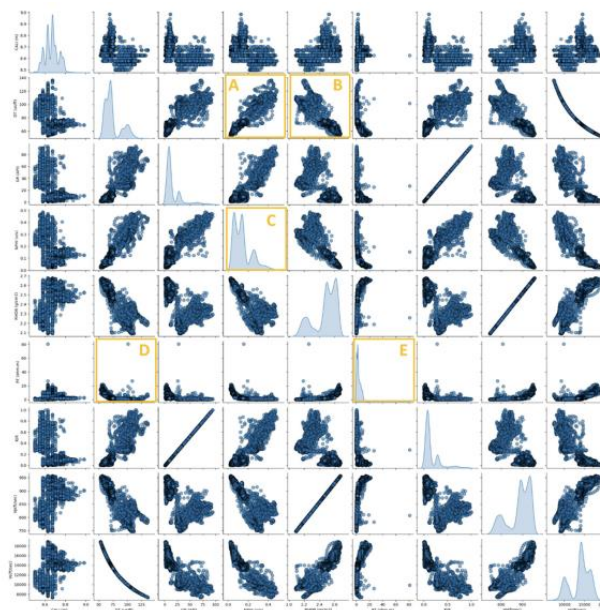
Pada jurnal Joshi et al. tujuan dari dilakukannya pendekatan *machine learning* ialah memprediksi data *log sonic* dan korelasi litologi. *log sonic* merupakan salah satu jenis data sumur yang didapatkan oleh teknologi *well logging*. Metode yang dilakukan untuk memprediksi data *log sonic* dan korelasi litologi ialah *machine learning*.

Data yang dibutuhkan untuk kegiatan ini ialah data lain yang didapatkan pada sumur tersebut berupa caliper, Gamma Ray, Resistivity, Densitas Bulk, dan NPHI (Neutron Porositas). Data yang digunakan pada jurnal ini bersumber dari *open source* di lapangan hidrokarbon Volve, Norwegia. Data yang digunakan berformat LAS dan disimpan dalam format excel seperti pada Gambar 1. Setelah data dimasukkan pada format excel dapat ditemukan data yang hilang, distribusi data, dan visualisasi data yang ada. Ditemukan bahwa data pada jurnal memiliki tipe distribusi skewed, korelasi positif, korelasi negative, outlier, dan distribusi spike yang diinterpretasikan dalam Gambar 2. Terlihat terdapat outlier pada data yang dapat menyebabkan noise pada data.

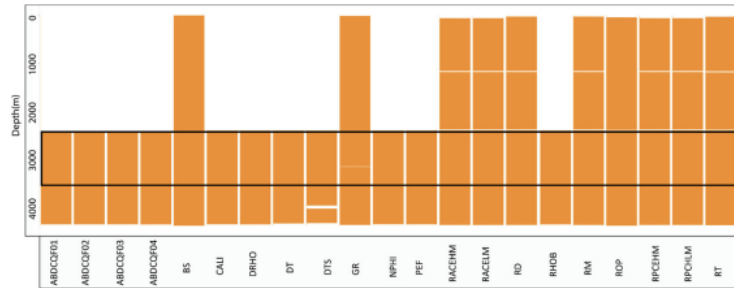
**Table 1** Showing excel data inventory prepared using different logs of well 159-F-11T2 for the required depth range (2577-3577m) used in current research work. The data obtained from the LAS 2.0 (logdata files) are converted into an excel file to prepare the excel data inventory for ML modelling. Different normalization procedures are applied to remove the abnormalities before ML-based modelling

Depth (m)	CALI (in)	DT (us/ft)	GR (API)	NPHI (v/v)	RACEHM (ohm.m)	RACELM (ohm.m)	RD (ohm.m)	RHOB (g/cm3)	RM (ohm.m)	RPCEHM (ohm.m)	RPCELM (ohm.m)	RT (ohm.m)
2577	8.766	109.385	74.7	0.403	0.868	0.823	0.75	2.347	0.751	0.751	0.75	0.751
2577.1	8.703	110.133	72.799	0.399	0.783	0.832	0.783	2.357	0.798	0.798	0.783	0.798
2577.2	8.703	110.843	71.631	0.39	0.818	0.831	0.776	2.36	0.75	0.75	0.776	0.75
2577.3	8.672	111.01	69.9	0.385	0.815	0.847	0.773	2.361	0.719	0.719	0.773	0.719
2577.4	8.672	110.797	69.04	0.382	0.824	0.858	0.757	2.363	0.693	0.693	0.757	0.693
2577.5	8.602	110.052	70.133	0.385	0.835	0.865	0.771	2.364	0.711	0.711	0.771	0.711
2577.6	8.578	108.809	72.008	0.372	0.85	0.882	0.775	2.363	0.699	0.699	0.775	0.699
2577.7	8.578	107.259	72.324	0.347	1	0.893	0.793	2.363	0.696	0.696	0.793	0.696
2577.8	8.602	106.379	65.954	0.317	0.947	0.894	0.871	2.366	0.839	0.839	0.871	0.839
2577.9	8.562	105.08	61.058	0.295	0.885	0.891	1	2.37	1.067	1.067	1	1.067
2578	8.547	103.452	56.124	0.28	0.911	0.885	1.069	2.375	1.238	1.238	1.069	1.238
2578.1	8.563	103.759	53.822	0.274	0.99	0.882	1.116	2.381	1.343	1.343	1.116	1.343
2578.2	8.602	103.968	54.241	0.279	1	0.877	1.174	2.386	1.561	1.561	1.174	1.561
2578.3	8.602	104.079	56.554	0.305	0.828	0.856	1.256	2.386	1.796	1.796	1.256	1.796
2578.4	8.578	105.944	59.192	0.344	0.808	0.86	1.15	2.376	1.418	1.418	1.15	1.418
2578.5	8.578	107.292	68.216	0.392	0.85	0.872	0.994	2.357	1.001	1.001	0.994	1.001
2578.6	8.578	107.911	74.277	0.436	0.877	0.873	0.852	2.337	0.794	0.794	0.852	0.794
2578.7	8.563	108.167	79.987	0.467	0.88	0.867	0.794	2.321	0.721	0.721	0.794	0.721
2578.8	8.578	108.126	82.846	0.475	0.852	0.857	0.75	2.309	0.655	0.655	0.75	0.655
2578.9	8.563	107.677	80.638	0.468	0.819	0.839	0.7	2.297	0.577	0.577	0.7	0.577
2579	8.578	107.583	78.152	0.451	0.8	0.835	0.679	2.289	0.536	0.536	0.679	0.536
2579.1	8.578	107.206	74.473	0.432	0.745	0.819	0.668	2.286	0.512	0.512	0.668	0.512
2579.2	8.578	106.195	71.828	0.411	0.705	0.818	0.667	2.287	0.532	0.532	0.667	0.532
2579.3	8.578	104.945	68.517	0.39	0.779	0.83	0.652	2.291	0.54	0.54	0.652	0.54
2579.4	8.547	103.515	63.701	0.373	0.738	0.812	0.67	2.295	0.607	0.607	0.67	0.607
2579.5	8.562	101.788	60.986	0.36	0.658	0.789	0.733	2.297	0.698	0.698	0.733	0.698
2579.6	8.547	100.624	59.558	0.353	0.711	0.805	0.724	2.298	0.708	0.708	0.724	0.708
2579.7	8.547	99.58	60.278	0.349	0.727	0.812	0.753	2.299	0.721	0.721	0.753	0.721
2579.8	8.578	98.404	61.809	0.347	0.742	0.822	0.766	2.298	0.731	0.731	0.766	0.731

Gambar 1. Data Sonic Log yang sudah berformat excel.

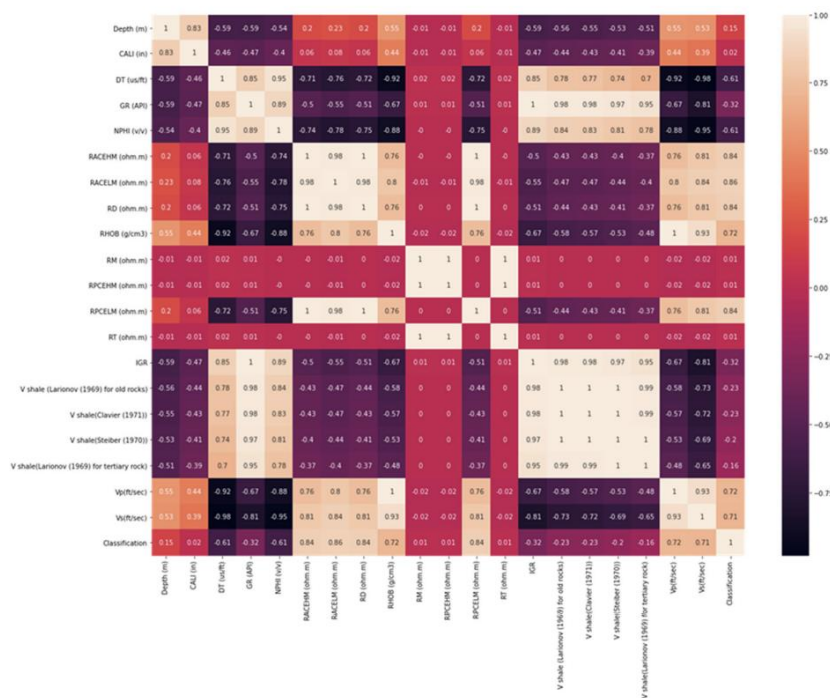


Gambar 2. *Pair-plot* yang menunjukkan korelasi parameter yang berbeda. A-Korelasi positif antara sonic log dan NPHI, B-Negatif korelasi antara sonic log dan RHOB, C-NPHI data termasuk positif skewness, D-outlier pada data, E-distribusi spike pada data resistivity.



Gambar 3. Box plot yang merepresentasikan data *well log* pada berbagai kedalaman.

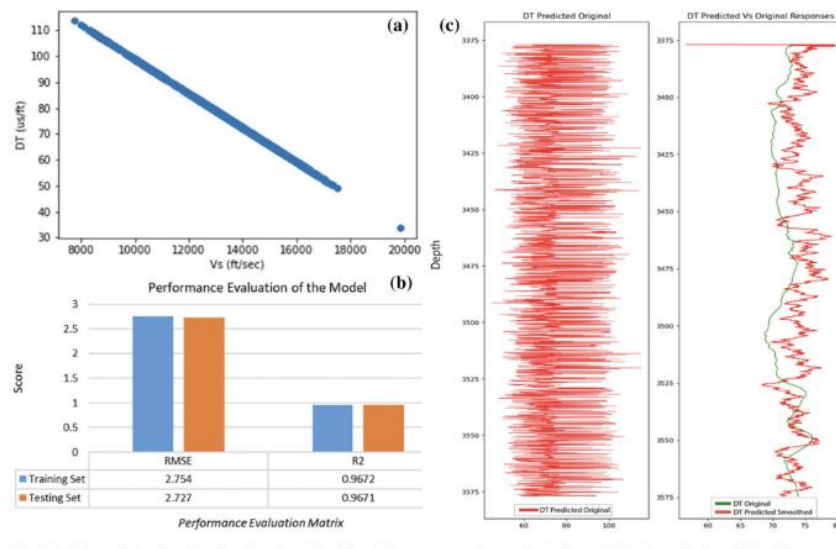
Selanjutnya dilakukan preparasi data yang bertujuan meningkatkan akurasi prediksi dari model yang buat. Data preparasi dilakukan dengan menormalisasi data dan melakukan koreksi data. Data yang sudah dilakukan preparasi terdiri dari nilai Gamma Ray dengan rentang 0.838 hingga 94.011 API, nilai Caliper dengan rentang 8.5 hingga 8.984 in, NPHI dengan rentang nilai 0.021 hingga 0.486, RHOB dengan rentang nilai 2.09 hingga 2.687 g/cm<sup>3</sup>. Setiap data diatas digunakan sebagai parameter untuk memprediksi sonic log. Namun, Vs (kecepatan gelombang sekunder) dijadikan parameter utama. Hal ini dikarenakan Vs memiliki korelasi tertinggi dengan sonic log. Nilai korelasi ini didapatkan berdasarkan matriks probabilitas pada Gambar 4. Nilai Vs didapatkan dari perhitungan petrofisika berdasarkan data *well logging*. Sehingga Nilai Vs sebagai variabel *independent* (bebas) dan sonic log sebagai variable *dependent* (terikat).



Gambar 4. Matriks korelasi dengan probabilitasnya dengan range -1 hingga 1 untuk menunjukkan tingkat korelasi dengan *sonic log*.

Tahap ketiga dilakukan *modelling* dengan tujuan membuat model prediktif, algoritma, model *machine learning*, dan lain-lain. Prediksi dari sonic log dilakukan dengan membuat model prediktif berupa linear regression. Pada jurnal ini 8,000 titik (80% dari keseluruhan data) sumur digunakan untuk membuat model prediksi, sedangkan 2,000 titik (20% dari keseluruhan data) digunakan untuk melakukan test pada hasil model prediksi. Pemodelan regressi yang dilakukan melakukan penempatan Vs sebagai sumbu x dan sonic log sebagai sumbu y. hubungan linear antara kedua parameter tersebut lalu dibuktikan oleh plot pada grafik. Tingkat akurasi prediksi dari sonic log ditentukan oleh nilai RMSE dan  $R^2$ . Pada data t

Tahap keempat yaitu melakukan evaluation data berupa interpretasi dan analisis. Berdasarkan data pada Gambar. 5a dan Gambar. 5b dapat dilihat nilai *training model* RMSE 2.622 dan  $R^2$  0.95. kedua nilai tersebut menunjukkan hasil yang akurat. Sedangkan nilai *testing point* RMSE 2.55 dan  $R^2$  0.96. kedua nilai tersebut menunjukkan hasil yang akurat. Pada Gambar. 5C dapat terlihat gambaran sonic log yang ada dengan prediksi dari sonic log. Prediksi dari sonic log menunjukkan tingkat akurasi 86.314%.



Gambar 5. (a) *regression plot*. (b) evaluasi model dari prediksis data log sonic dengan RMSE dan  $R^2$ . (c) log sonic original dan log sonic prediksi melalui machine learning.

Data sonic log sangatlah penting untuk eksplorasi dan produksi minyak dan gas bumi. Namun untuk mendapatkan data sonic log yang akurat dibutuhkan biaya yang mahal dikarenakan memerlukan kuantitas sumur eksplorasi yang banyak. Jurnal ini membuktikan bahwa

pendekatan *machine learning* dapat memprediksi data sonic log dengan cepat, akurat, dan murah daripada Teknik konvensional. Oleh karena itu, data sonic log pada jurnal ini digunakan untuk eksplorasi dan produksi dari lapangan minyak dan gas bumi.

Link GitHub: <https://github.com/luthfifhrz/Topic-1-2---Muhammad-Luthfi-Fahrezi>

## Daftar Pustaka

Joshi, D., Patidar, A., Mishra, A., Agarwajal, S. (2021). [Prediction of sonic log and correlation of lithology by comparing geophysical well log data using machine learning principles | SpringerLink](#). GeoJournal.