

Prediksi Tekanan Pori Berdasarkan Data *Logging* Sumur Menggunakan *Deep Neural Network*

Meredita Susanty¹; Raka Sudira Wardana²

¹ Program Studi Ilmu Komputer, Fakultas Sains dan Komputer, Universitas Pertamina

² Program Studi Teknik Perminyakan, Fakultas Teknik Eksplorasi dan Perminyakan, Universitas Pertamina

¹ meredita.susanty@universitaspertamina.ac.id

ABSTRACT

Formation pore pressures are essential data used in designing parameters in drilling operations. Errors in determining the pore pressure of the formation can cause problems in drilling, resulting in increased drilling costs and the loss of workers' lives. Pore pressure can be obtained through direct measurements or using empirical methods. However, direct measurements only perform at a certain depth due to its high costs. The empirical method requires accuracy in the Trendline Normal Compaction analysis also bounds for use in certain formations. This research explores the potential of artificial neural networks to predict pore pressure based on well logging data. Using an artificial neural network architecture that uses three hidden layers, the model is built with five activation functions and the Mean Absolute Error evaluation metric. All models perform well in predicting pore pressure, as indicated by a loss below 0.5. Among the five activation functions, the Exponential Linear Unit produces the lowest loss of 0.07 compared to models with other activation functions.

Keywords: Deep Neural Network, Artificial Intelligence, Activation Function, Pore Pressure, Well Log

ABSTRAK

Tekanan pori formasi merupakan data penting yang digunakan dalam mendesain parameter dalam operasi pengeboran. Kesalahan dalam menentukan tekanan pori formasi dapat menimbulkan permasalahan dalam pengeboran yang dapat menyebabkan bertambahnya biaya pengeboran hingga kehilangan nyawa pekerja. Tekanan pori bisa didapatkan melalui pengukuran langsung maupun menggunakan metode empiris. Pengukuran langsung tidak dilakukan di setiap kedalaman karena biayanya yang tinggi. Metode empiris memiliki keterbatasan dalam menentukan tekanan pori formasi dikarenakan metode ini memerlukan adanya keakuratan dalam analisis Normal Compaction Trendline dan batasan untuk digunakan pada formasi tertentu. Penelitian ini menggali potensi jaringan saraf tiruan untuk memprediksi tekanan pori berdasarkan data logging sumur. Dengan arsitektur jaringan saraf tiruan yang menggunakan tiga hidden layer, model dibangun dengan lima fungsi aktivasi dan metrik evaluasi Mean Absolute Error. Semua mampu memprediksi tekanan pori dengan baik yang ditunjukkan dengan nilai loss dibawah 0.5. Diantara kelima fungsi aktivasi, Exponential Linear Unit menghasilkan loss paling baik sebesar 0.07 dibanding model dengan fungsi aktivasi lainnya.

Kata kunci: Deep Neural Network, Kecerdasan Buatan, Fungsi aktivasi, Tekanan Pori, Logging Sumur

1. PENDAHULUAN

Tekanan pori formasi merupakan tekanan fluida yang berada di dalam pori formasi [1]. Pada operasi pengeboran minyak dan gas bumi, data tekanan pori formasi digunakan untuk mendesain parameter seperti lumpur pengeboran, casing dan semen [2]. Kesalahan dalam penentuan tekanan pori formasi bisa menyebabkan permasalahan pada saat pengeboran dilakukan, seperti terjadinya *kick* dan *blow out*.

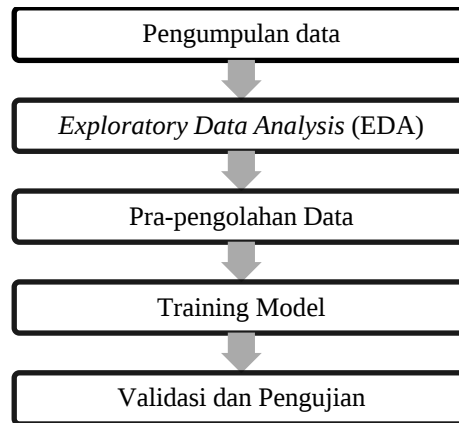
Tekanan pori formasi bisa diukur secara langsung menggunakan peralatan seperti *Drill Stem Test* (DST) ataupun *Repeat Formation Tester* (RFT) [3]. Namun pengukuran tekanan formasi secara langsung hanya dilakukan di beberapa titik kedalaman saja, terutama di zona reservoir. Hal ini dikarenakan biaya operasional RFT dan DST yang cukup tinggi. Padahal dalam mendesain parameter pengeboran, dibutuhkan data tekanan pori formasi di setiap kedalaman.

Untuk menentukan tekanan pori formasi di setiap kedalaman umumnya dilakukan dengan pendekatan menggunakan metode empiris seperti metode Hotmann and Johnson [4], Eaton [5], Bowers [6] dan Zhang [7]. Penentuan tekanan pori formasi menggunakan metode tersebut dibuat berdasarkan teori bahwa porositas akan berkurang seiring bertambahnya kedalaman dengan *trendline* tertentu yang disebut juga sebagai *Normal Compaction Trendline*. Adanya penyimpangan terhadap *Normal Compaction Trendline* akan menandakan adanya tekanan pori formasi abnormal. Oleh karena itu, akurasi tekanan pori formasi menggunakan metode tersebut akan dipengaruhi terhadap ketepatan dalam menentukan *Normal Compaction Trendline*. Penentuan *Normal Compaction Trendline* memerlukan analisa terhadap litologi formasi. Selain itu, ada beberapa limitasi dalam penentuan tekanan formasi menggunakan metode yang berbasis porositas [8] mengasumsikan formasi yang ditinjau memiliki litologi yang sama. Limitasi lainnya yaitu berkurangnya porositas juga tidak selalu disebabkan oleh *mechanical compaction*. Oleh karena itu penentuan tekanan pori formasi menggunakan persamaan berbasis perubahan porositas tidak selalu memberikan hasil yang akurat yang dikarenakan kesalahan asumsi atau kurangnya data [8].

Pendekatan *machine learning* memiliki potensi untuk mengatasi keterbatasan tersebut. Pendekatan ini telah dilakukan dalam beberapa penelitian sebelumnya dengan berbagai teknik seperti jaringan saraf tiruan, *support vector machine*, *radial basis function*, *fuzzy logic*, dan *functional networks* [9]–[11]. Penelitian-penelitian sebelumnya melakukan perbandingan beberapa metode, perbandingan jumlah *layer* dan *nodes* dalam jaringan syaraf tiruan. Selain jumlah *layer* dan *nodes* dalam jaringan syaraf tiruan, fungsi aktivasi yang digunakan dalam model juga berpengaruh terhadap performa model [12]. Penelitian ini bertujuan membandingkan beberapa fungsi aktivasi dan mencari fungsi aktivasi yang paling baik untuk digunakan dalam jaringan syaraf tiruan dalam memprediksi nilai tekanan pori.

2. METODE/PERANCANGAN PENELITIAN

Penelitian ini mengikuti tahapan yang dijelaskan pada gambar Gambar 1 yang terdiri dari pengumpulan data, analisis data, pra-pengolahan data, *training* model yang dibangun dengan menggunakan teknik *deep neural network*, serta validasi dan pengujian terhadap model yang dibuat.



Gambar 1. Metodologi Penelitian

Dataset dalam penelitian ini terdiri dari 5 atribut input yakni kedalaman sumur, *gamma ray*, *sonic density*, dan *accoustic slowness*, untuk memprediksi 1 atribut output tekanan pori seperti dijelaskan pada Tabel 1. Pada tabel 1, input variabel atau fitur diberi nama x dan output variabel diberi nama y. Dataset yang didapat dibagi menjadi 3 data *training*, data validasi dan data pengujian dengan perbandingan 3:1:1.

Selanjutnya dilakukan analisis terhadap data yang ada. Exploratory Data Analysis (EDA) merupakan investigasi awal terhadap data untuk menemukan pola, anomali menggunakan pendekatan statistika dan representasi grafis. Pada tahapan ini dilakukan pemeriksaan tipe data, rentang nilai, dan kelengkapan data pada setiap fitur. Selain itu juga dilakukan pemeriksaan untuk melihat korelasi setiap fitur terhadap output serta tingkat korelasi setiap fitur terhadap output. Jika korelasi antara fitur dan output linear, maka tingkat korelasi diukur menggunakan *Pearson's Rank Correlation*[13]. Sedangkan jika korelasinya tidak linear, maka digunakan *Spearman Rank Correlation*[14].

Tabel 1. Input dan Output Variabel

Atribut	Keterangan
X1	Kedalaman sumur
X2	<i>Gamma Ray</i>
X3	<i>Sonic Density</i>
X4	<i>Accoustic Slowness</i>
Y	<i>Tekanan Pori</i>

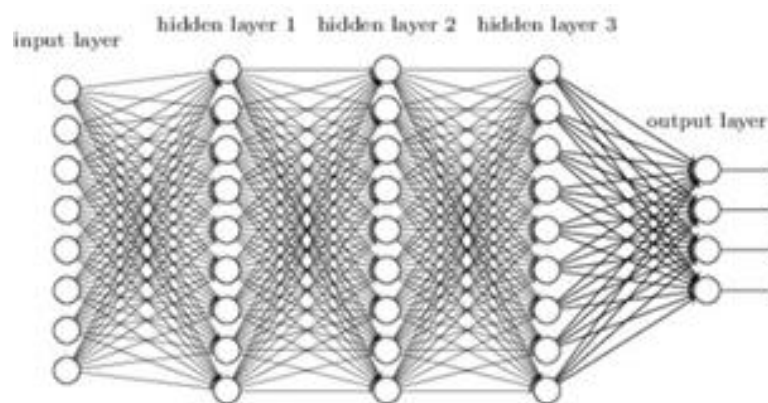
Pada tahapan pra-pengolahan data dilakukan normalisasi menggunakan *min-max scaling* jika hasil exploratory data analysis menunjukkan perbedaan rentang nilai untuk tiap fitur. *Min-max scaling* melakukan perubahan linear terhadap data asli [15]. Setiap nilai akan dipetakan menjadi nilai baru dalam range nilai yang sudah ditentukan menggunakan persamaan (1). Metode ini dipilih karena tetap mempertahankan *relationship* diantara data asli.

$$v' = \frac{v - \min_A}{\max_A - \min_A} (n\max_B - n\min_B) + n\min_B \quad (1)$$

$$\sigma(z) = 1 / (1 + \exp(-z)) \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - x_i| \quad (3)$$

Model untuk memprediksi tekanan pori dibangun menggunakan teknik *deep neural network* dimana model akan memiliki beberapa *hidden layer* seperti ditunjukkan pada Gambar 2. Penelitian ini menggunakan beberapa fungsi aktivasi – Rectified Linear Unit (ReLU), Scaled Exponential Linear Unit (SELU), Exponential Linear Unit (ELU), softsign, tanh – dan memilih salah satu fungsi aktivasi yang paling baik performanya. Tujuan utama dari penggunaan fungsi aktivasi ini adalah untuk membentuk non-linearity (ketidaklinieran) pada *neural network*. Tanpa adanya fungsi aktivasi, *neural network* hanya akan melakukan transformasi linier dari input ke output.



Gambar 2. Arsitektur JST [16]

Training merupakan proses mencari nilai yang memiliki error paling rendah [12]. Karena model digunakan untuk memprediksi nilai yang kontinu, performa model diukur berdasarkan nilai *loss* yang dihitung menggunakan *loss function* tertentu [17], semakin kecil nilai *loss* dari suatu model artinya nilai prediksi yang dihasilkan model semakin dekat dengan nilai yang sebenarnya. Untuk mengevaluasi performa model dalam penelitian ini digunakan *loss function Mean Absolute Error* (MAE) pada persamaan 3 [18]. *Training* dilakukan dengan menjalankan model terhadap data *training*.

Proses validasi bertujuan menilai performa model dan menyempurnakan parameter model seperti *weight* dan *bias* pada tiap neuron dan memastikan model yang dihasilkan tidak mengalami *overfitting* terhadap data tertentu maupun *underfitting*. Proses validasi dilakukan dengan menjalankan model yang sudah *ditraining* menggunakan data validasi. Model yang dipilih adalah model yang memiliki nilai *loss* paling rendah dan selisih nilai *loss* paling kecil antara hasil *training* dan validasi.

Pengujian dilakukan terhadap model yang sudah *ditraining* menggunakan data pengujian. Pengujian bertujuan menilai performa model yang sudah dibuat terhadap data baru.

3. HASIL DAN PEMBAHASAN

Dataset yang digunakan berasal dari 3 sumur dengan kedalaman antara sembilan ratus hingga tiga belas ribu kaki. Dataset yang digunakan berjumlah 1426 data yang dibagi menjadi 3 kelompok secara acak dengan proporsi 3:1:1 menjadi 912 data *training*, 228 data *validation* dan 286 data tes.

Analisis data dilakukan terhadap keseluruhan dataset. Analisis pertama yang dilakukan adalah memeriksa tipe data, rentang nilai, dan kelengkapan data pada setiap fitur. Semua data menggunakan

tipe data float dan tidak ada data yang kosong. Hasil pemeriksaan secara statistik pada Tabel 2 menunjukkan bahwa ada perbedaan rentang nilai untuk tiap fitur.

Tabel 2. Analisis Statistik

	X1	X2	X3	X4	Y
count	1426	1426	1426	1426	1426
mean	6369.59	42.52	2.38	104.10	9.94
std	2683.02	16.30	0.17	24.11	1.41
min	943	4.48	1.72	48.75	8.05
25%	4505.5	30.02	2.30	89.87	8.8
50%	6564	44.42	2.4	101.93	9.66
75%	8232.2	54.8	2.5	115.5	10.89
max	13250	89.07	2.74	177.5	15.98

Karena rentang nilai yang beragam, dilakukan normalisasi terhadap data menggunakan teknik *min-max scaling*. Hasil normalisasi ditunjukkan pada Tabel 3.

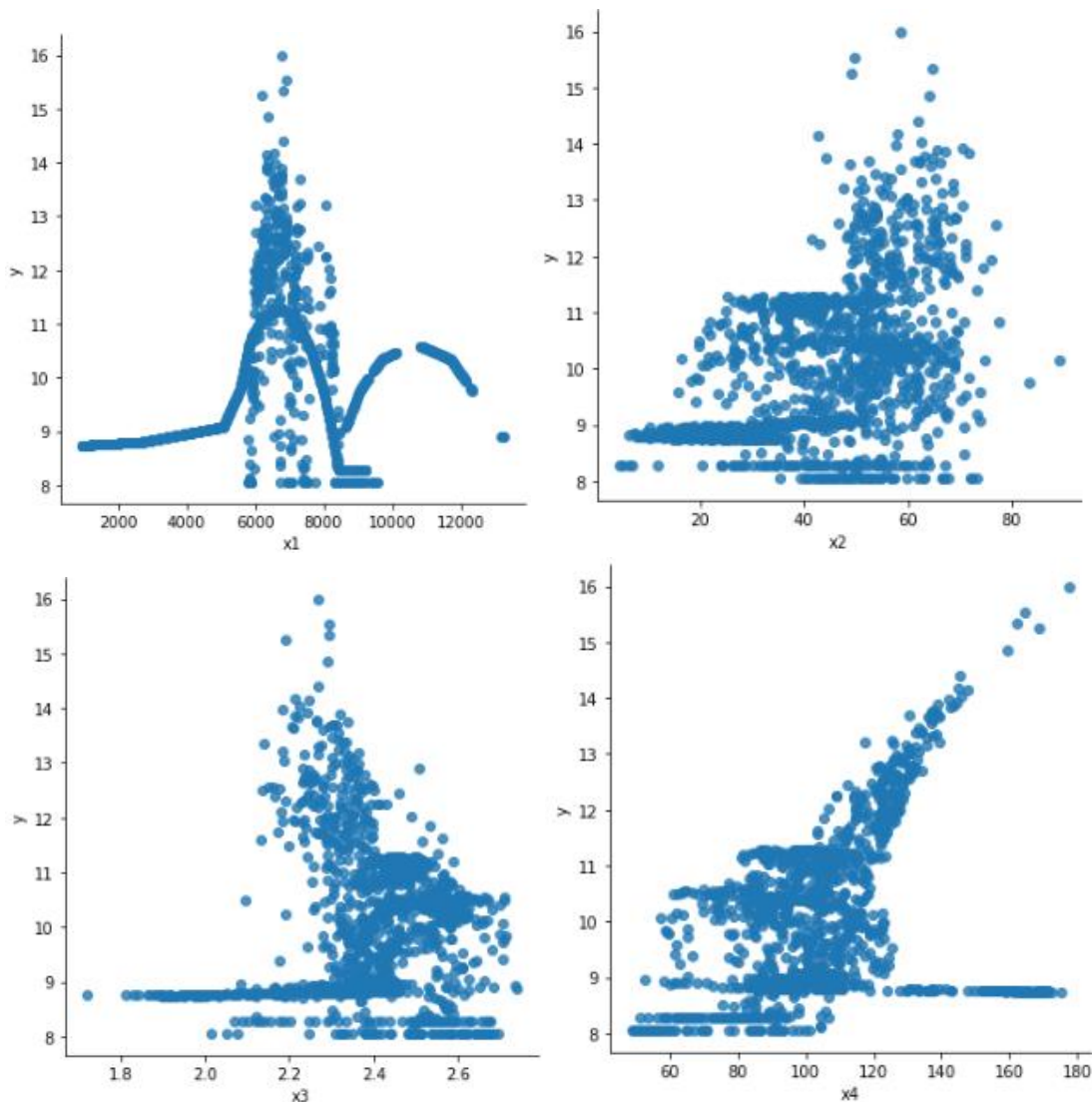
Tabel 3. Hasil Normalisasi

	X1	X2	X3	X4	Y
count	1426	1426	1426	1426	1426
mean	-1.195	-2.390	7.175	-8.968	-2.591
std	1.000	1.000	1.000	1.000	1.000
min	-2.023	-2.335	-3.885	-2.297	-1.336
25%	-0.695	-0.767	-0.452	-0.590	-0.810
50%	0.072	0.117	0.083	-0.090	-0.196
75%	0.694	0.754	0.686	0.473	0.674
max	2.565	2.857	2.109	3.046	4.281

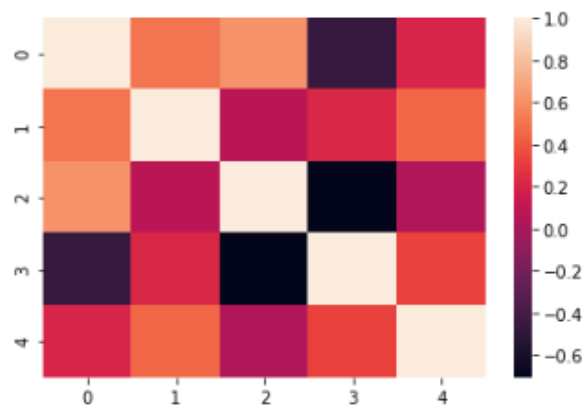
Berdasarkan data yang sudah dinormalisasi diperiksa korelasi antara setiap fitur terhadap output dengan menggunakan *scatter plot* pada Gambar 3. Dari visualisasi data yang ditampilkan di Gambar 3 bisa dilihat bahwa korelasi antara setiap fitur atau input terhadap output tidak linear. Selanjutnya dilakukan pemeriksaan seberapa besar korelasi masing-masing fitur terhadap output. Karena hubungan antara setiap fitur tidak linear terhadap output, untuk memeriksa tingkat korelasi digunakan *Spearman rank correlation coefficient*. Gambar 4 menunjukkan tingkat korelasi dengan gradasi warna, warna yang terang menunjukkan korelasi positif sedangkan warna gelap menunjukkan korelasi negatif. Dari Gambar 4 bisa dilihat bahwa semua fitur memiliki korelasi positif terhadap output dengan fitur *gamma ray* dan *acoustic slowness* memiliki korelasi yang lebih kuat dibandingkan fitur kedalaman dan *sonic density*.

Selanjutnya, model *deep neural network* dibangun dengan 1 input layer yang terdiri dari 180 nodes, 3 hidden layer yang masing-masing terdiri dari 270 nodes, dan 1 output layer berisi 2 nodes yang mewakili nilai tekanan pori. Arsitektur yang sama akan dilatih menggunakan fungsi aktivasi yang berbeda; ReLU, ELU, SELU, tanH dan softsign. *Training* model dilakukan terhadap data *training* dan data validasi sebanyak 1000 kali (*epoch*). Selama model berjalan, algoritma optimasi Adam digunakan untuk memperbarui *weight* berdasarkan data *training*. Untuk menghitung *loss* digunakan *loss function* MAE. Hasil *training* dan validasi untuk setiap fungsi aktivasi ditunjukkan

pada Gambar 5. Semua model menunjukkan penurunan nilai loss yang menandakan model belajar dengan baik. Nilai loss pada Gambar 5 yang semakin lama semakin mendekati 0 (turun) menunjukkan bahwa prediksi yang dilakukan oleh sistem semakin baik. Sistem belajar dan melakukan perbaikan nilai *weight* yang digunakan untuk *fitting network* dengan baik. Dari kelima model pada Gambar 5, terlihat bahwa nilai *loss* mulai melandai pada iterasi ke 50, semakin melandai di iterasi ke 200 dan relatif stabil pada iterasi ke 600. Jika ingin mempercepat jalannya simulasi pada model, parameter epoch pada model bisa diturunkan menjadi 600. Dari kelima model, fungsi aktivasi ELU menunjukkan performa paling baik karena nilai *loss* yang paling rendah.

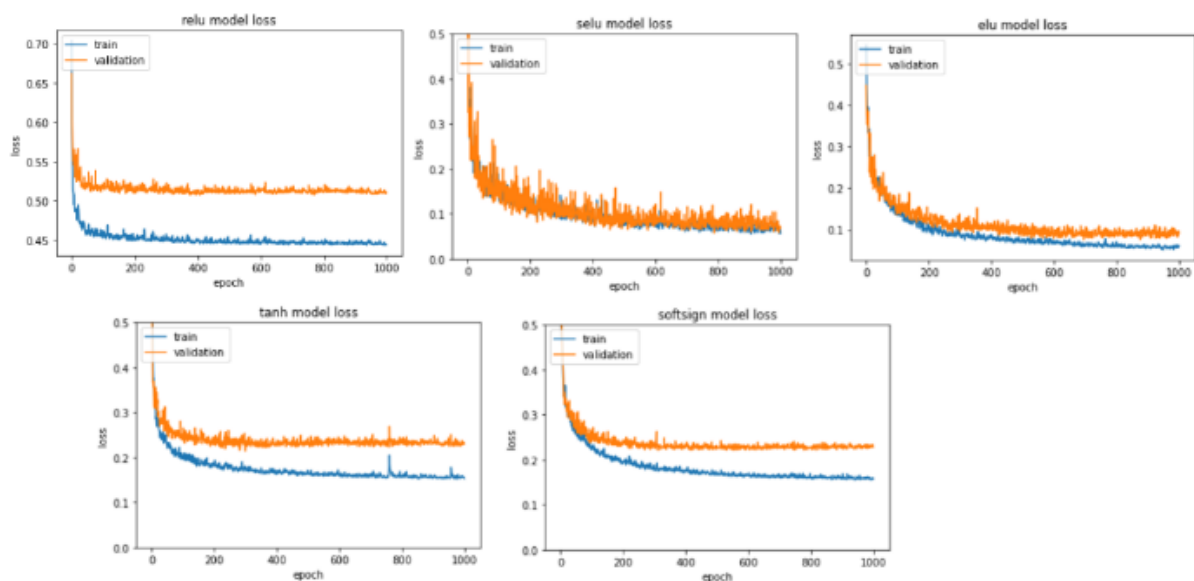


Gambar 3. Korelasi Setiap Fitur terhadap Output



Gambar 4. Tingkat Korelasi Setiap Fitur terhadap Output

Validasi digunakan menggunakan data validasi sebanyak 228 data yang sudah dinormalisasi. Pada Gambar 5 hasil validasi menunjukkan nilai loss yang sedikit lebih tinggi. Gambar 5 menunjukkan kemiripan nilai loss antara model train dan validation. Hal ini menunjukkan bahwa model tidak mengalami *overfitting*. Fungsi aktivasi SELU memiliki hasil validasi yang paling baik karena kesamaan nilai yang dihasilkan.



Gambar 5. Loss Pada Model

Hasil pengujian terhadap 286 data test ditunjukkan pada Tabel 4. Hasil pengujian menunjukkan nilai *loss* yang sedikit lebih tinggi dibanding hasil *training* namun mendekati nilai *loss* pada hasil validasi. Fungsi aktivasi ELU menunjukkan performa yang paling baik dibanding fungsi aktivasi lainnya.

Tabel 4. Hasil Pengujian

Fungsi aktivasi	Loss
ReLU	0.4735
ELU	0.0730
SELU	0.0870
tanH	0.1936
softsign	0.2431

Berdasarkan hasil *training*, validasi dan pengujian, model yang menggunakan fungsi aktivasi ELU menunjukkan performa paling baik. Berdasarkan metric MAE, model dengan fungsi aktivasi ELU memiliki nilai error yang paling rendah baik saat *training*, validasi dan testing.

4. KESIMPULAN DAN SARAN

Metode kecerdasan buatan menggunakan *deep neural network* memberikan solusi alternatif dan mengatasi keterbatasan dalam memprediksi tekanan pori menggunakan metode empiris yang saat ini digunakan. Untuk melakukan prediksi terhadap data yang kontinu, model dibangun menggunakan teknik regresi dan dievaluasi menggunakan metrics MAE memanfaatkan 5 parameter yang didapat dari data sumur; kedalaman, *Gamma Ray*, *Sonic Density*, dan *Acoustic Slowness*. Dalam penelitian ini, 5 fungsi aktivasi untuk data yang tidak linear dibandingkan dan diukur performanya menggunakan metrik tersebut. Hasil evaluasi menunjukkan bahwa fungsi aktivasi ELU memiliki performa yang paling baik dengan nilai loss sebesar 0.07. Selain menggunakan MAE, permasalahan regresi bisa dievaluasi menggunakan metrik lain seperti *Mean Squared Error*. Penelitian selanjutnya dapat membandingkan dampak fungsi error terhadap performa model.

DAFTAR PUSTAKA

- [1] N. J. Adams and T. Charrier, *Drilling Engineering A Complete Well Planning Approach*. Tulsa, Oklahoma, United States: Pennwell Books, 1985.
- [2] M. D. Zoback, *Reservoir Geomechanics*. Cambridge, United Kingdom: CAMBRIDGE UNIVERSITY PRESS, 2007.
- [3] M. E. C. and F. S. Y. J. A.T. Bourgoyne Jr, K.K. Millheim, *Applied Drilling Engineering*. Society of Petroleum Engineers, 1991.
- [4] C. E. Hottmann and R. K. Johnson, "Estimation of Formation Pressures from Log-Derived Shale Properties," *J. Pet. Technol.*, vol. 17, no. 06, pp. 717–722, Jun. 1965.
- [5] B. A. Eaton, "The Equation for Geopressure Prediction from Well Logs," in *Fall Meeting of the Society of Petroleum Engineers of AIME*, 1975.
- [6] G. L. Bowers, "Pore Pressure Estimation From Velocity Data: Accounting for Overpressure Mechanisms Besides Undercompaction," *SPE Drill. Complet.*, vol. 10, no. 02, pp. 89–95, Jun. 1995.
- [7] J. Zhang, "Pore pressure prediction from well logs: Methods, modifications, and new approaches," *Earth-Science Rev.*, vol. 108, no. 1–2, pp. 50–63, Sep. 2011.
- [8] R. E. Swarbrick, "Pore-Pressure Prediction: Pitfalls in Using Porosity," in *Offshore Technology Conference*, 2001.
- [9] A. Ahmed, S. Elkatatny, A. Ali, and A. Abdulraheem, "Comparative analysis of artificial intelligence techniques for formation pressure prediction while drilling," *Arab. J. Geosci.*, vol. 12, no. 18, 2019.

- [10] E. Z. Naeini, S. Green, I. Russell-Hughes, and M. Rauch-Davies, “An integrated deep learning solution for petrophysics, pore pressure, and geomechanics property prediction,” *Lead. Edge*, vol. 38, no. 1, pp. 53–59, 2019.
- [11] L. Hu et al., “A new pore pressure prediction method-back propagation artificial neural network,” *Electron. J. Geotech. Eng.*, vol. 18 S, pp. 4093–4107, 2013.
- [12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [13] M. Kamber, *Data Mining : Concepts and Techniques*, Second. Elsevier, 2016.
- [14] C. Xiao, J. Ye, R. M. Esteves, and C. Rong, “Using Spearman’s correlation coefficients for exploratory data analysis on big dataset,” *Concurr. Comput. Pract. Exp.*, vol. 28, no. 14, pp. 3866–3878, Sep. 2016.
- [15] J. Han, M. Kamber, and J. Pei, *Data Mining : Concepts and Techniques*, Third. Elsevier, 2012.
- [16] L. Anastasia Kyrykovich, “Deep Neural Networks.” [Online]. Available: <https://www.kdnuggets.com/2020/02/deep-neural-networks.html>. [Accessed: 05-Jun-2020].
- [17] F. Chollet, *Deep Learning with Python*. Manning Publications, 2017.
- [18] C. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Clim. Res.*, vol. 30, pp. 79–82, Dec. 2005.