

PENERAPAN ALGORITMA LINIER REGRESI PADA SISTEM PREDIKSI HARGA GAS ALAM

Moh.Budi Nurcahyono¹ Bambang Wisnu Widagdo²

¹Teknik Informatika, Universitas Pamulang, Jl. Raya Puspitek No.10 Serpong,
Tangerang Selatan, 15310

e-mail: mbudinurcahyono22@gmail.com

e-mail: ²dosen02092@unpam.ac.id

ABSTRACT

Fluctuating natural gas prices are one of the challenges in making predictions. Therefore, predicting natural gas prices is important for analyzing natural gas prices using the linear regression method. Natural gas prices are predicted based on several factors such as supply and demand factors, weather and also supply and stock. The problem usually faced in this research is that in predicting natural gas prices, it is difficult to read and analyze changes or movements in natural gas prices, and can result in errors in decision and policy making. Linear regression is a data analysis technique that predicts unknown data values using other related and known data values. Data collection was carried out by searching for historical natural gas data through website sources, and through mathematical relationships between variables, namely the independent variable (X) opening 14.41-97.80, highest 15.67-99.87, lowest 14.32-92.00, vol. 120K - 4592K, % change -16.66 – 21.89 and the last dependent variable (Y) is 14.82-96.47, each of which has a minimum value and a maximum value. This analysis is to determine whether the value of the variable has increased or decreased. Natural gas prices were collected from 2015 to 2024 using a test model and training model, the test model used 2374 lines of data and the training model used 112 lines of data. The Anaconda Navigator application is used to carry out the data analysis process. The results of this research show that machine learning can be used to predict and analyze natural gas prices with an MSE of 2823.5, MAE of 35.221, and R-Squared of 0.9850. Thus, these results show that the linear regression method has an accuracy level of 98.50% according to the R-Squared value.

Keywords: Natural Gas, Prediction, Linear Regression

ABSTRAK

Harga gas alam yang fluktuatif menjadi salah satu tantangan dalam melakukan prediksi. Oleh karena itu, prediksi harga gas alam menjadi penting untuk melakukan analisis harga gas alam dengan menggunakan metode regresi linier. Harga gas alam diprediksi berdasarkan beberapa faktor seperti, faktor penawaran dan permintaan, cuaca dan juga persediaan dan stok. Masalah yang biasa dihadapi dalam penelitian ini adalah dalam prediksi harga gas alam, sulit untuk membaca dan menganalisis perubahan atau pergerakan harga gas alam tersebut, serta dapat mengakibatkan kesalahan dalam pengambilan keputusan dan kebijakan. Regresi linier adalah teknik analisis data yang memprediksi nilai data yang tidak diketahui dengan menggunakan nilai data lain yang terkait dan diketahui. Pengumpulan data dilakukan dengan mencari data histori gas alam melalui sumber *website*, dan melalui hubungan matematis antara variabel, yaitu variabel *independent* (X) pembukaan 14,41-97,80, tertinggi 15,67-99,87, terendah 14,32-92,00, vol. 120K - 4592K, perubahan % -16.66 – 21.89 dan variabel *dependent* (Y) terakhir 14,82-96,47 yang masing-masing memiliki nilai minimal dan nilai maksimal. Analisis ini untuk mengetahui nilai dari variabel mengalami kenaikan atau penurunan. Harga gas alam dikumpulkan dari tahun 2015 hingga 2024 menggunakan model uji dan model *training*, model uji menggunakan 2374 baris data dan model *training* menggunakan 112 baris data. Aplikasi *anaconda navigator* digunakan untuk melakukan proses analisis data. Hasil penelitian ini menunjukkan bahwa *machine learning* dapat digunakan untuk memprediksi dan menganalisis harga gas

alam dengan MSE sebesar 2823.5, MAE sebesar 35.221, dan *R-Squared* sebesar 0.9850. Dengan demikian, hasil tersebut menunjukkan bahwa metode regresi linier memiliki tingkat akurasi sebesar 98.50% sesuai dengan nilai *R-Squared*.

Kata Kunci : Gas Alam, Prediksi, Regresi linier

1. PENDAHULUAN

Gas alam merupakan salah satu sumber energy didunia yang mana pemanfaatannya di masa depan diperkirakan semakin tinggi dan menjadi sumber energi yang penting [1]. Gas alam merupakan salah satu jenis sumber energi yang berasal dari fosil tanaman, hewan, mikroorganisme yang tersimpan dibawah tanah selama ribuan bahkan jutaan tahun [2]. Energi adalah salah satu elemen penting dalam kehidupan manusia, dan kebutuhan akan energi terus meningkat seiring dengan perkembangan dunia [3]. Untuk memenuhi kebutuhan energi tersebut kebanyakan dari manusia menggunakan energi fosil terutama minyak bumi, akan tetapi dalam ketersediaannya jumlah minyak bumi di dunia semakin menipis. Salah satu potensi sumber energi yang bisa kita gunakan selain dari minyak bumi adalah gas bumi atau gas alam. Gas alam memiliki banyak keunggulan dibandingkan dengan energi lain [4]. Ini disebabkan sumber dayanya yang melimpah, ramah lingkungan, serta sumber energi yang efisien terutama jika dilihat dari biaya modal yang lebih rendah dalam mengubah energi menjadi listrik.

Gas alam merupakan pendapatan utama Indonesia yakni sekitar 25% hingga 30%, dalam hal basis cadangan gas alam, Indonesia menempati posisi ke empat belas di dunia [5]. Selain menjadi sumber energi yang tidak terbarukan, gas alam memainkan peran penting dalam memenuhi kebutuhan energi primer di seluruh dunia. Untuk mewujudkan kemandirian energi dan ketahanan energi nasional, pemerintah menerbitkan PP No. 79 tahun 2014 tentang Ketahanan Energi Nasional (KEN). Dalam PP tersebut, KEN menargetkan bauran energi nasional pada tahun 2025 adalah batu bara 30%, minyak bumi 25%, gas bumi 22%, dan EBT 23%. Di tahun 2050, bauran energi batu bara dan minyak bumi diturunkan menjadi 25% batu bara dan 20% minyak bumi, sedangkan bauran energi gas bumi ditingkatkan menjadi 24% dan EBT menjadi 31%. Karena peran besarnya dalam energi dunia, gas alam dianggap sebagai industri utama dalam pasar energi, namun trend harga gas alam sangat fluktuatif dan sangat sulit untuk dianalisa dan diprediksi. Untuk itu diperlukan upaya prediksi terkait gas alam yang akurat.

Proses pengolahan dalam *machine learning* memerlukan data yang digunakan sebagai data pelatihan, pembelajaran atau *training* dan data yang digunakan untuk dipelajari sebagai data latih (*training set*). Proses *machine learning* akan dilakukan pembagian dataset yang akan diolah menjadi 2 bagian yaitu data yang menjadi data *training* sebanyak 80% dan data *testing* sebanyak 20%. Semua pengolahan diuji dengan menggunakan jupyter notebook dengan memanfaatkan *library* yang mendukung dalam *regresi linear* seperti *pandas*, *matplotlib*, *seaborn* dan *sklearn*.

Fungsi dari *library* yang digunakan antara lain *library pandas* digunakan untuk analisis data yang akan digunakan. Data yang digunakan dapat berupa file Microsoft Excel, CSV maupun basis data. *Pandas* dapat digunakan untuk membersihkan data asli ke dalam sebuah bentuk data yang sesuai yang diinginkan. Kegunaan lain diantaranya menyelaraskan data untuk perbandingan dan penggabungan set data, penanganan data yang hilang atau tidak sesuai, *library matplotlib* digunakan untuk visualisasi data, *library sklearn* digunakan dalam proses olah statistika terutama dalam regresi linier dan *library seaborn* digunakan visualisasi data statistik, seperti pembuatan grafik dan visualisasi yang dapat merangkum data serta proses menggambarkan distribusi.

Penerapan data mining menggunakan metode regresi linier, setiap data harga gas alam yang terdiri dari 7 atribut, data tersebut dikonversi menjadi data yang bertipe *object* ke dalam bentuk numerik, sehingga menjadi 6 feature dan 1 output yang terdiri dari tanggal, pembukaan, tertinggi, terendah, vol., perubahan %, dan terakhir, yang memungkinkan data dapat dianalisa lebih lanjut untuk pengambilan keputusan yang akurat. Diharapkan dengan adanya *machine learning* ini dapat memprediksi harga gas alam yang *fluktuatif* dengan tepat, cepat, dan akurat agar tidak terjadi kesalahan dalam pengambilan keputusan dan kebijakan dalam menentukan harga gas alam serta dapat memudahkan analisa dan prediksi perubahan trend harga gas alam.

Pada penelitian akan menerapkan metode regresi linier untuk memprediksi harga gas alam agar dapat memudahkan proses analisa dan prediksi harga gas alam, peneliti akan menggunakan metode regresi linier.

2. METODE

2.1 Metode Penelitian

Metodologi yang digunakan dalam penelitian ini melalui beberapa tahapan, yaitu:

a. Observasi.

Melakukan observasi dengan melihat secara langsung dengan objek penelitian.

b. Studi Pustaka

Tahap ini dilakukan dengan cara mengkaji dan mempelajari literature dan referensi berupa naskah ilmiah, buku konsep *machine learning*, serta cara kerja metode Regresi linier dan metode lainnya sehingga dapat menunjang metodologi yang akan diterapkan pada penelitian.

3. HASIL

3.1 Pengumpulan Data (Dataset)

Seperti disebutkan pada sub bab 3.2 bahwa data yang digunakan untuk melakukan penelitian adalah *dataset* data *historis* gas alam yang berasal dari website [investing.com](https://id.investing.com/commodities/natural-gas-historical-data) (<https://id.investing.com/commodities/natural-gas-historical-data>) yang memiliki atribut dengan keterangan sebagai berikut:

Tabel 3. 1 *Dataset* Data Historis Gas Alam

Nama Atribut	Keterangan
Tanggal	Waktu perdagangan gas
Terakhir	Harga penutupan perdagangan gas pada hari tersebut
Pembukaan	Harga pembukaan perdagangan gas pada hari tersebut
Tertinggi	Harga tertinggi perdagangan gas pada hari tersebut
Terendah	Harga terendah perdagangan gas pada hari tersebut
Vol.	Jumlah kuantitas gas yang diperdagangkan pada hari tersebut
Perubahan %	Persentase perubahan harga gas pada hari tersebut dibandingkan dengan hari sebelumnya

3.2 Pengolahan Data Awal

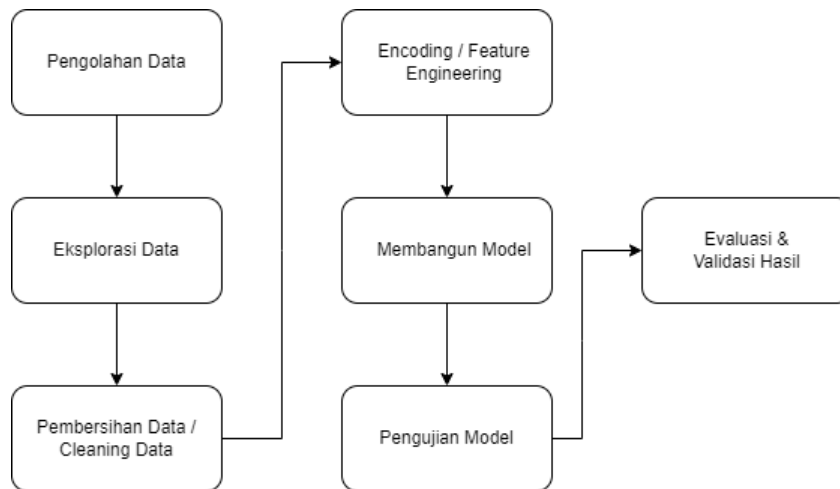
Pada pengolahan data awal ini dilakukan pemilihan atribut atau tahap preprocessing yang akan digunakan dalam penelitian ini yang sebelumnya sebanyak 2374 data dengan 7 atribut, dengan struktur awal data untuk setiap atribut dan *Class* seperti tabel di bawah ini:

Tabel 3. 2 Struktur Awal *Dataset*

Nama Atribut	Nilai	Tipe Data
Tanggal	Tanggal Awal: 2015-01-01 Tanggal Akhir: 2023-12-29	Object
Terakhir	14,82 - 96,47	Object
Pembukaan	14,41 - 97,80	Object
Tertinggi	15,67 - 99,87	Object
Terendah	14,32 - 92,00	Object
Vol.	120K - 4592K	Object
Perubahan %	-16.53 - 21.89	Object

3.3 Perancangan Penelitian

Dalam penelitian ini menggunakan metode Regresi linier untuk memprediksi harga gas alam, dimana data yang diambil merupakan *Dataset* DHGA yang diambil langsung dari website investing.com(<https://id.investing.com/commodities/natural-gas-historical-data>). Tahapan-tahapan yang dibutuhkan dalam penelitian ini adalah: Pengolahan data, *Eksplorasi data*, *Cleaning data*, *Encoding*, kemudian Membangun model, lalu Pengujian Model dan terakhir Evaluasi & Validasi hasil seperti terlihat pada gambar dibawah ini:



Gambar 3. 1 Rancangan Penelitian

a. Pengolahan Data

Tabel 3. 3 Tabel *Dataset* Latih (*Training*)

Inde x	Tanggal	Terakhir (Rp)	Pembukaan (Rp)	Tertinggi (Rp)	Terendah (Rp)	Vol. (Kubik)	Perubahan (%)
0	29/12/2023	2,514	2,539	2,561	2,489	89,60K	-1,68%
1	28/12/2023	2,557	2,455	2,576	2,412	116,06K	-2,37%
2	27/12/2023	2,619	2,580	2,722	2,562	3,93K	2,71%
3	26/12/2023	2,550	2,545	2,581	2,454	50,76K	5,46%
4	25/12/2023	2,418	2,425	2,442	2,415	NaN	-7,36%
...
2369	07/01/2015	2,871	2,949	3,012	2,825	134,83K	-2,28%
2370	06/01/2015	2,938	2,921	2,955	2,811	111,34K	1,94%
2371	05/01/2015	2,882	3,076	3,176	2,875	141,78K	-4,03%
2372	02/01/2015	3,003	2,845	3,096	2,805	109,12K	5,70%
2373	01/01/2015	2,841	2,916	2,916	2,825	NaN	-1,66%

Tabel 3. 4 Tabel *Dataset* Pengujian (*Testing*)

Inde x	Tanggal	Terakhir (Rp)	Pembukaan (Rp)	Tertinggi (Rp)	Terendah (Rp)	Vol. (Kubik)	Perubahan (%)
0	31/05/2024	2,587	2,570	2,624	2,518	139,12K	0,58%
1	30/05/2024	2,572	2,662	2,682	2,562	160,06K	3,17%
2	29/05/2024	2,493	2,594	2,669	2,417	2,66K	-3,75%
3	28/05/2024	2,590	2,503	2,621	2,475	86,03K	-6,83%
4	27/05/2024	2,780	2,747	2,853	2,742	NaN	1,16%
...
107	05/01/2024	2,893	2,845	2,906	2,681	187,50K	2,55%
108	04/01/2024	2,821	2,688	2,863	2,681	206,31K	5,73%
109	03/01/2024	2,668	2,575	2,719	2,537	166,47K	3,89%
110	02/01/2024	2,568	2,605	2,675	2,524	132,45K	-3,71%
111	01/01/2024	2,667	2,614	2,670	2,614	NaN	6,09%

- b. *Eksplorasi Data*
- c. *Pembersihan Data*
- d. *Encoding / Feature Engineering*
- e. *Membangun Model*
- f. *Pengujian Model*

Untuk menguji model, berbagai metrik dan metode digunakan, seperti:

1. Akurasi (*Accuracy*)
2. Presisi (*Precision*)
3. MSE (*Mean Squared Error*)
4. *R-Squared*
5. MAE (*Mean Absolute Error*)

Tujuan penelitian ini adalah untuk menerapkan metode algoritma Regresi linier untuk memprediksi harga gas alam. Penelitian ini dilakukan terdapat *Dataset Realtime* dengan *range*.

3.4 Persiapan Data

Dalam penelitian ini terdapat satu *dataset* yaitu *dataset realtime* DHGA dari situs resmi *investing.com* yang dapat diunduh disitus <https://id.investing.com/commodities/natural-gas-historical-data>. Seperti pada tabel dibawah ini:

Tabel 4. 1 Dataset Awal

No	Nama	Record	Feature	Class
1	DHGA – Data Historis Gas Alam	2374	7	1

Dataset ini merupakan *dataset* publik dengan record yang cukup banyak dengan berbagai tipe, tetapi dalam penelitian ini peneliti hanya menggunakan harga gas alam dengan 7 *Feature* dan 2374 *Record*.

4. PEMBAHASAN

4.1 Pengolahan Data

Pada tahap ini dilakukan pengolahan data dengan mengkonversi data yang bertipe *object* menjadi tipe data numerik, hal ini bertujuan untuk memudahkan proses eksplorasi dan analisa *dataset* yang akan dilakukan pada tahap berikutnya. Seperti pada tabel dibawah ini.

Tabel 4. 2 Tabel Pengolahan Data

Nama Atribut	Nilai	Type
Tanggal	2015/01/01 – 2023/12/29	<i>Datetime</i>
Terakhir	14,82-96,47	<i>Float</i>
Pembukaan	14,41-97,80	<i>Float</i>
Tertinggi	15,67-99,87	<i>Float</i>
Terendah	14,32-92,00	<i>Float</i>
Vol.	120K - 4592K	<i>Float</i>
Perubahan %	-16.53 – 21.89	<i>Float</i>

4.2 Eksplorasi Data

Pada tahap ini dilakukan proses analisa dari *Dataset* dengan tujuan untuk mengetahui *dataset* yang digunakan terkait dengan dimensi, tipe, serta pola data yang ada pada *dataset*.

4.3 Analisa Data

Dalam analisa data pada *dataset* dilakukan pemisahan terkait dengan *Class* atau *label* yang berfungsi sebagai target atau *output* dengan atribut atau *feature* yang berfungsi sebagai input.

Tabel 4. 3 Tabel Analisa Data

Nama Kolom	Type	Keterangan
Tanggal	<i>Datetime</i>	<i>Feature</i>
Terakhir	<i>Float</i>	<i>Output</i>
Pembukaan	<i>Float</i>	<i>Feature</i>
Tertinggi	<i>Float</i>	<i>Feature</i>
Terendah	<i>Float</i>	<i>Feature</i>
Vol.	<i>Float</i>	<i>Feature</i>
Perubahan %	<i>Float</i>	<i>Feature</i>

4.4 Pembersihan Data

a. *Class* atau *Label*

Class atau *label* pada *dataset* ini adalah kolom terakhir yang merupakan harga penutupan atau harga terakhir dari penjualan gas alam pada hari tersebut. Akan berfungsi sebagai *output* dari model *machine learning* yang dibuat, hasil analisa dapat dilihat pada tabel dibawah ini.

Tabel 4. 4 Tabel *Class*

Keterangan	Nilai
Nilai Minimal	14,82
Nilai Maksimal	96,47
Jumlah Baris	2374 Baris Data
Keunikan Setiap Data	2374 Baris Data

b. Atribut atau *feature*

1. Tanggal

Tabel 4. 5 Tabel Tanggal

Keterangan	Nilai
Jumlah Baris	2374 Baris Data
Keunikan Setiap Data	2374 Baris Data
Tanggal Awal	02/01/2015
Tanggal Akhir	29/12/2023

2. Pembukaan

Tabel 4. 6 Tabel Pembukaan

Keterangan	Nilai
Jumlah Baris	2374 Baris Data
Keunikan Setiap Data	1545 Baris Data
Data yang sering muncul	2,619
Frekuensi kemunculan	7

3. Tertinggi

Tabel 4. 7 Tabel Tertinggi

Keterangan	Nilai
Jumlah Baris	2374 Baris Data
Keunikan Setiap Data	1511 Baris Data
Data yang sering muncul	2,955
Frekuensi kemunculan	9

4. Terendah

Tabel 4. 8 Tabel Terendah

Keterangan	Nilai
Jumlah Baris	2374 Baris Data
Keunikan Setiap Data	1507 Baris Data
Data yang sering muncul	2,822
Frekuensi kemunculan	7

5. Vol.

Tabel 4. 9 Tabel Vol.

Keterangan	Nilai
Jumlah Baris	2257 Baris Data
Keunikan Setiap Data	2155 Baris Data
Data yang sering muncul	157,12K
Frekuensi kemunculan	3

6. Perubahan %

Tabel 4. 10 Tabel Perubahan %

Keterangan	Nilai
Jumlah Baris	2374 Baris Data
Keunikan Setiap Data	1065 Baris Data
Data yang sering muncul	0,00%
Frekuensi kemunculan	13

4.5 Statistik Data

Pada tahap ini dilakukan proses penggalian informasi terkait dengan nilai-nilai statistik pada *dataset*, khususnya pada kolom - kolom yang bertipe numerik, adapun informasi terkait statistik pada *dataset* ini, dapat dilihat dibawah ini.

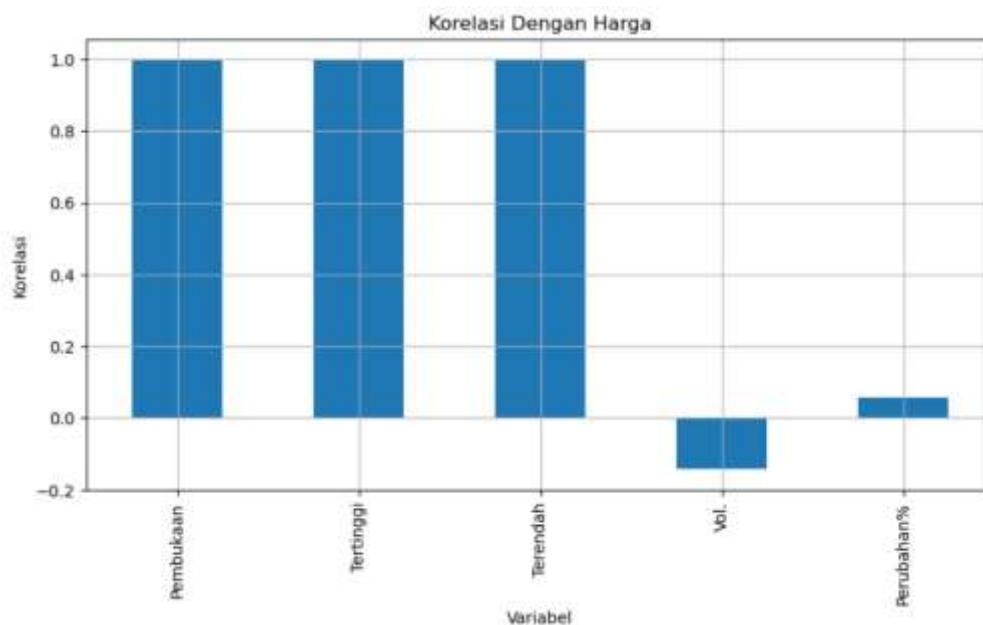
4.6 Korelasi Atribut (Input) terhadap Label (Ouput)

Korelasi *atribut* terhadap *label* mengacu pada hubungan statistik antara atribut atau fitur suatu *dataset* dengan *label* atau target yang ingin diprediksi.

Tabel 4. 11 Korelasi Atribut

Input	Korelasi	Output
Pembukaan	0.995366	Terakhir
Tertinggi	0.997881	Terakhir
Terendah	0.998361	Terakhir
Vol.	-0.143712	Terakhir
Perubahan %	0.055978	Terakhir

Berdasarkan tabel di atas dapat dijelaskan bahwa terdapat dua korelasi antara atribut terhadap *label*, yaitu korelasi positif pada kolom atau *atribut* pembukaan, tertinggi, terendah, dan perubahan sedangkan korelasi negatif terdapat pada kolom vol. sebesar -0.143712. Dari data pada tabel diatas maka dapat dilakukan visualisasi dalam bentuk grafik, seperti gambar di bawah ini.



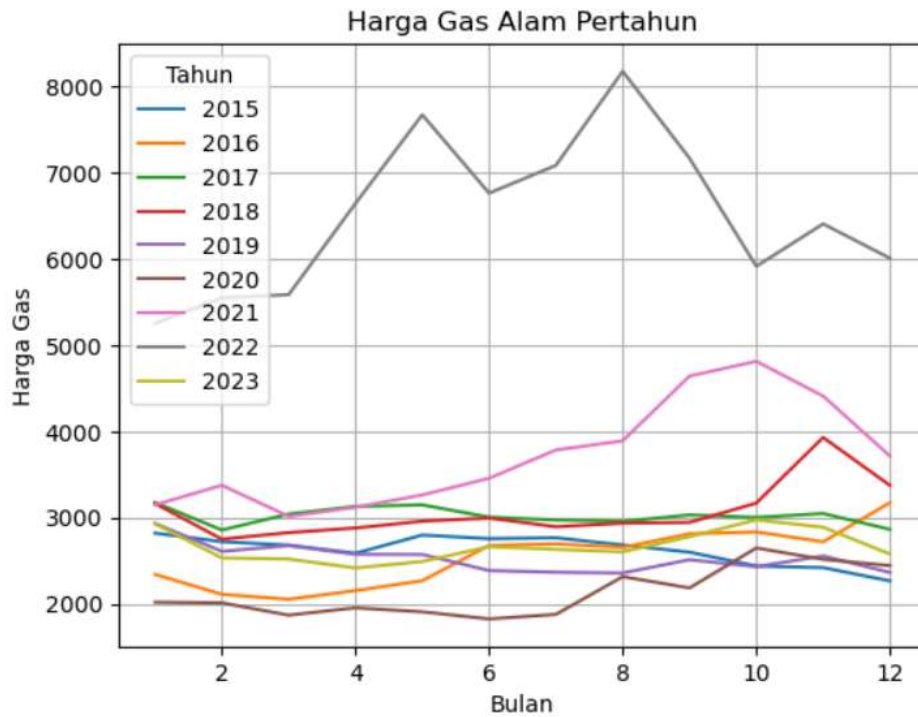
Gambar 4. 1 Korelasi Atribut

Ketika angka semakin mendekati angka 1, korelasi dinyatakan semakin baik dalam hubungan antar kolom. Artinya, semakin dekat nilai korelasi dengan 1, semakin kuat hubungan linier antara dua variabel. Jika nilai korelasi mendekati 1, maka perubahan dalam satu variabel cenderung disertai oleh perubahan yang sebanding dalam variabel lainnya. Dengan kata lain, terdapat ketergantungan yang kuat antara kedua variabel tersebut.

4.7 Visualisasi Data

Pada tahap ini akan dilakukan *visualisasi* data menggunakan grafik atau diagram untuk memberikan gambaran trend atau pola data.

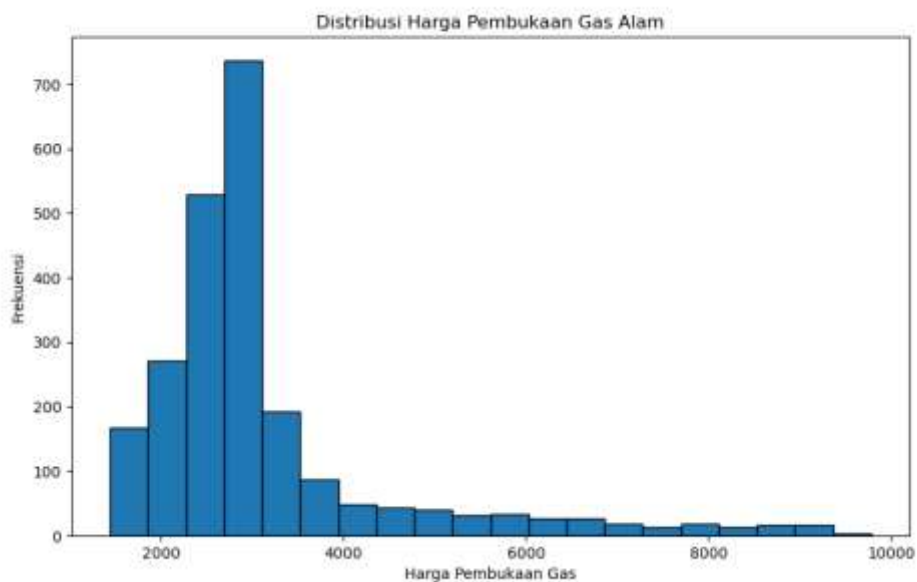
a. *Class* atau *Label* pada gambar dibawah ini merupakan visualisasi atau gambaran harga gas alam sepanjang tahun 2015 sampai dengan tahun 2023.



Gambar 4. 2 Harga Gas Alam Pertahun

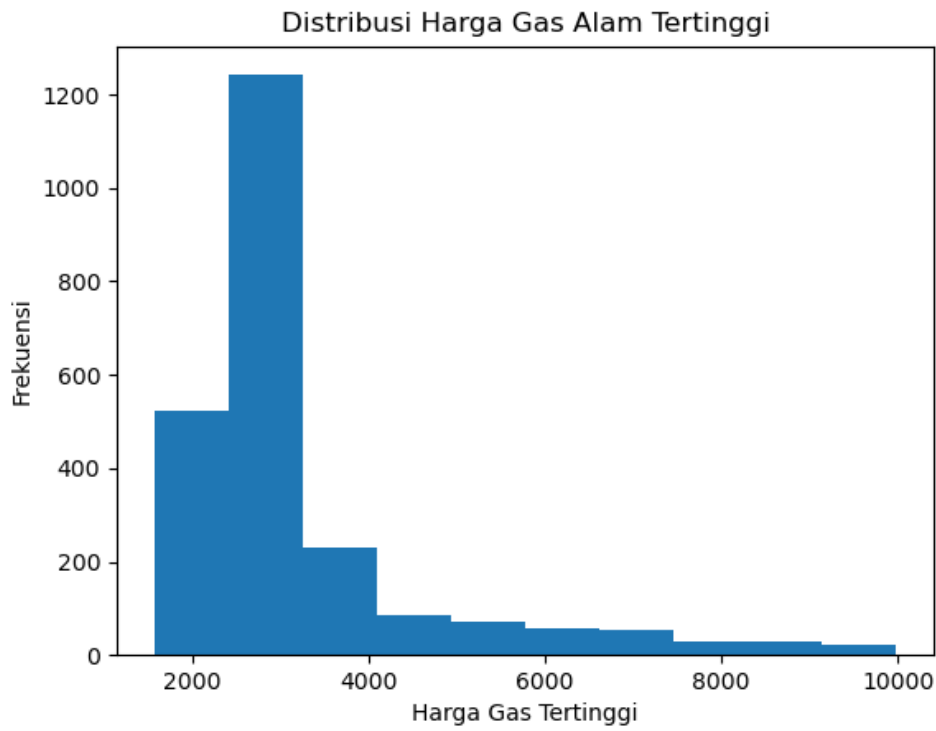
b. Atribut atau feature Pada model prediksi harga gas alam, atribut atau feature yang berfungsi sebagai input antara lain:

1. Distribusi Harga Pembukaan Gas Alam



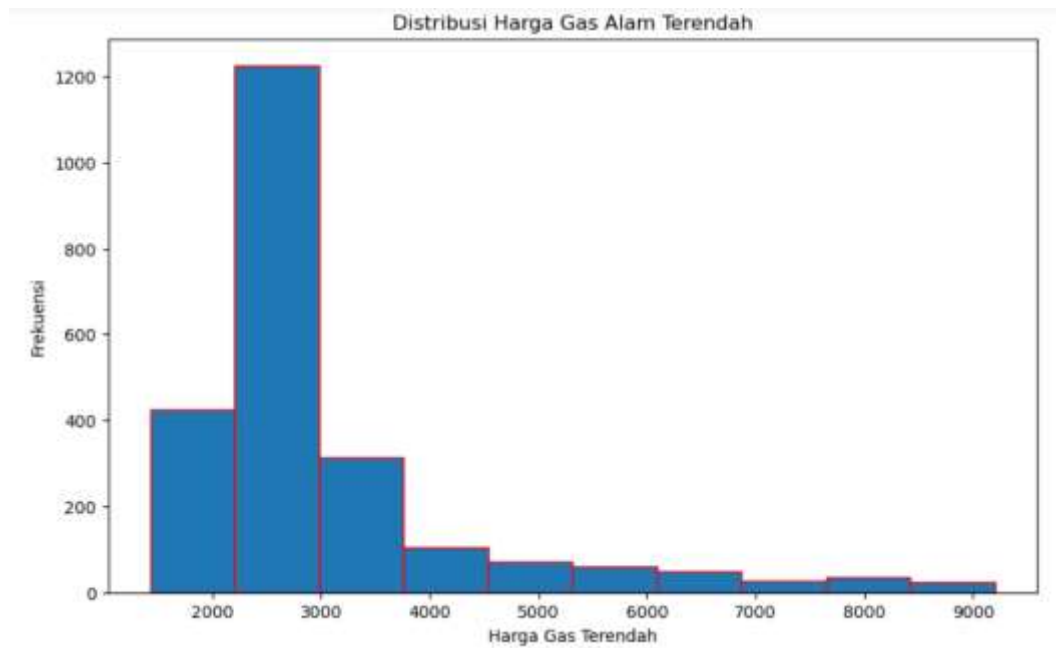
Gambar 4. 3 Grafik Pembukaan

2. Distribusi Harga Gas Alam Tertinggi



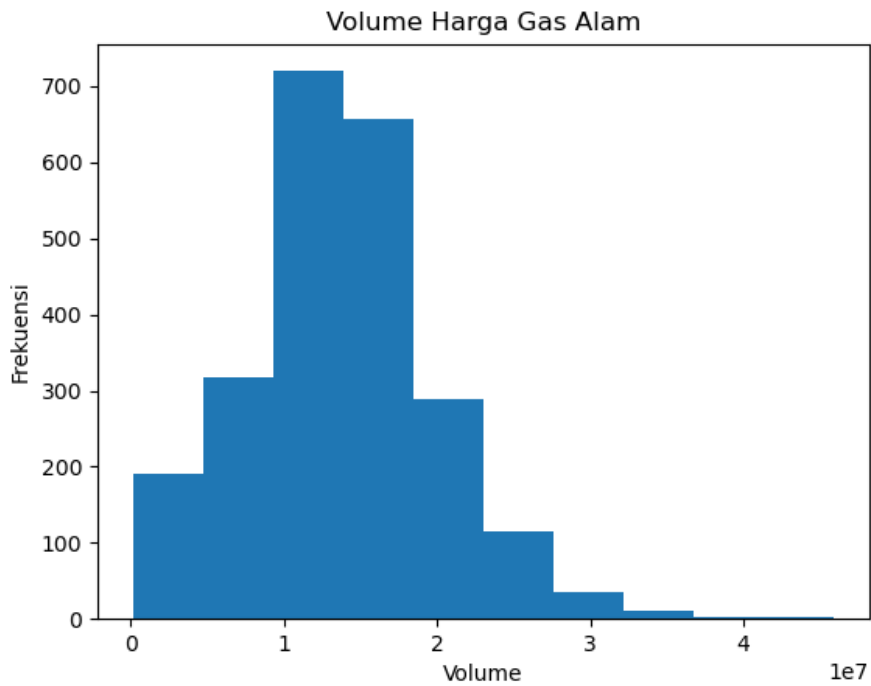
Gambar 4. 4 Grafik Tertinggi

3. Distribusi Harga Gas Alam Terendah



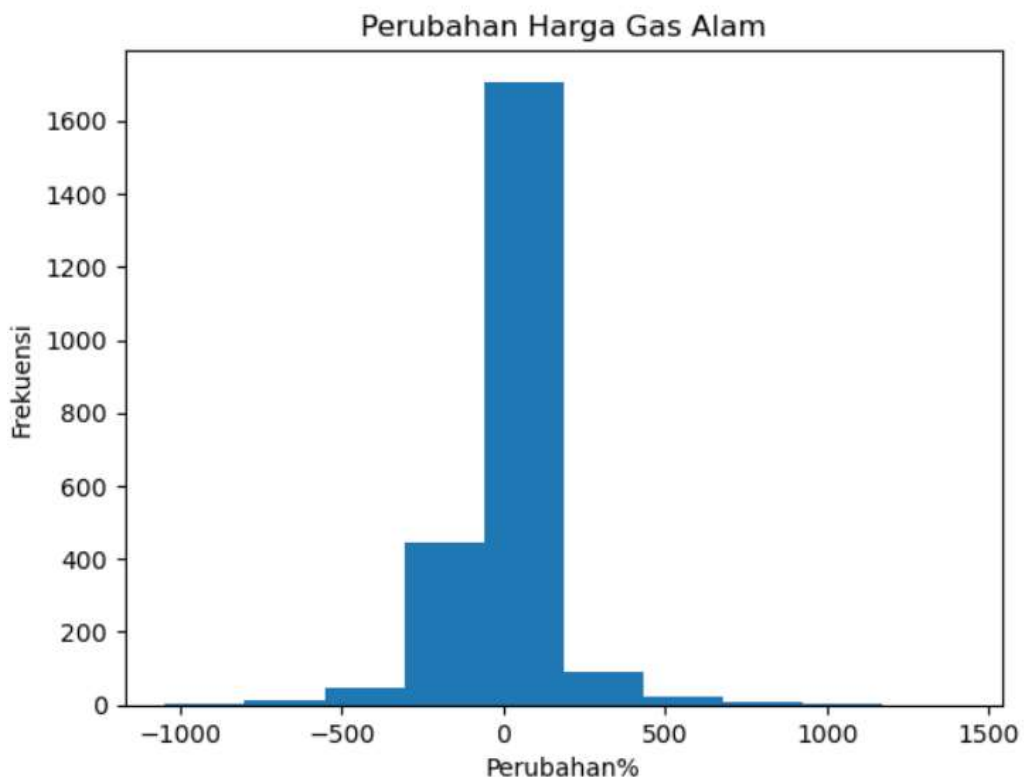
Gambar 4. 5 Grafik Harga Gas Alam Terendah

4. Grafik Volume Penjualan Gas Alam



Gambar 4. 6 Volume Penjualan Gas Alam

5. Grafik Perubahan Harga Gas Alam



Gambar 4. 7 Perubahan Harga Gas Alam

4.8 Encoding (Feature Engineering)

Encoding merupakan proses konversi data yang bertipe *object* ke dalam bentuk numerik, sehingga memungkinkan analisa lebih lanjut dan pengambilan keputusan yang akurat. Seperti pada tabel berikut:

Tabel 4. 12 Tabel Encoding (Feature Engineering)

No	Kolom	Tipe Data Sebelum Di Konversi	Kolom	Tipe Data Setelah Di Konversi
1.	Tanggal	<i>Object</i>	<i>Hari</i>	<i>Int</i>
			<i>Minggu</i>	<i>Int</i>
			<i>Bulan</i>	<i>Int</i>
			<i>Tahun</i>	<i>Int</i>
2.	Terakhir	<i>Object</i>	Terakhir	<i>Float</i>
3.	Pembukaan	<i>Object</i>	Pembukaan	<i>Float</i>
4.	Tertinggi	<i>Object</i>	Tertinggi	<i>Float</i>
5.	Terendah	<i>Object</i>	Terendah	<i>Float</i>
6.	Vol.	<i>Object</i>	Vol.	<i>Float</i>
7.	Perubahan%	<i>Object</i>	Perubahan%	<i>Float</i>

Pada *dataset* diatas setiap kolom merupakan tipe data objek yang harus dikonversikan kedalam bilangan numerik agar dapat diproses dengan sistem matematis.

4.9 Membangun Model

Membangun model merupakan langkah krusial dalam analisis data, dimana kita menciptakan suatu representasi matematis yang dapat menggambarkan hubungan antar variabel dalam *dataset*. Pada penelitian ini peneliti menggunakan metode regresi linier untuk membangun model *machine learning* ini.

Tabel 4. 13 Dataset Yang Digunakan

<i>Dataset Yang Digunakan</i>	<i>Dimensi</i>
Data Historis Gas Alam	2374 Baris Data & 7 Kolom

a. *Split Data*

Pada tahapan ini dilakukan pembagian data menjadi data latih dan data uji, dimana data *training* diambil dari *dataset* keseluruhan sebanyak 80%, sementara data *testing* diambil sebanyak 20% dari keseluruhan data.

Tabel 4. 14 Split Data

<i>Keterangan</i>	<i>Ukuran</i>	<i>Dimensi</i>
<i>Data Training</i>	80%	1805 Baris Data & 7 Kolom
<i>Data Testing</i>	20%	452 Baris Data & 7 Kolom

b. *Training Model*

Menggunakan model Regresi linier untuk menganalisa dan mengambil keputusan berdasarkan data yang ada. Model Regresi linier adalah teknik analisis data yang memprediksi nilai data yang tidak diketahui dengan menggunakan nilai data lain yang terkait dan diketahui.

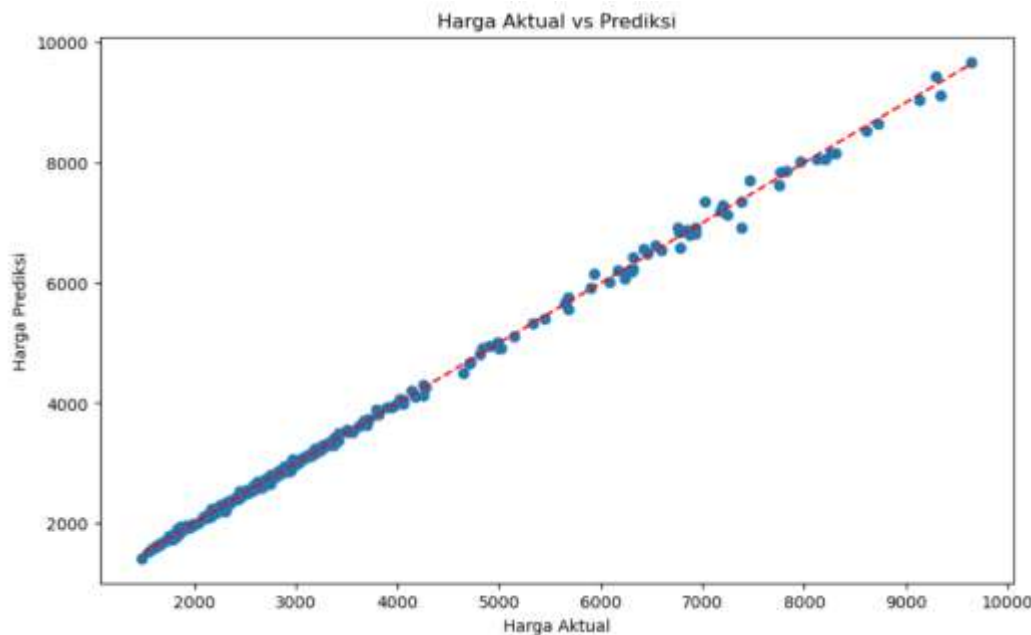
4.10 Pengujian Model

Pengujian ini dilakukan untuk memastikan bahwa model dapat memberikan prediksi yang akurat dan dapat diandalkan dalam mengklasifikasikan data baru. Dalam pengujian ini, beberapa metrik evaluasi seperti Mean Squared Error (MSE), *Mean Absolute Error* (MAE), dan *R-Squared* digunakan untuk mengukur sejauh mana model dapat mengklasifikasikan data dengan benar.

Tabel 4. 15 Parameter Model

<i>Model</i>	<i>Hasil</i>
<i>Mean Squared Error (MSE)</i>	2569.4
<i>Mean Absolute Error (MAE)</i>	27.959
<i>R-Squared</i>	0.9988

Berdasarkan evaluasi dan pengujian model yang telah dilakukan, maka memperoleh hasil pengujian dengan nilai akurasi *MSE* 2569.4, *MAE* 27.959, *R-Squared* sebesar 0.9988.



Gambar 4. 8 Grafik Harga Aktual

Terlihat pada grafik diatas dapat dijelaskan bahwa hasil aktual harga gas alam memiliki tingkat akurasi yang tinggi yaitu sebesar 99.88%, dimana harga asli aktual sesuai atau tepat dengan harga prediksi atau harga perkiraan.

4.11 Evaluasi Dan Validasi Hasil

Melakukan proses pengujian model *machine learning* yang telah dibuat dengan menggunakan pola *dataset* yang baru yang telah disiapkan sebelumnya, adapun *dataset* yang digunakan dapat dilihat pada tabel dibawah ini:

Tabel 4. 16 Dataset Uji

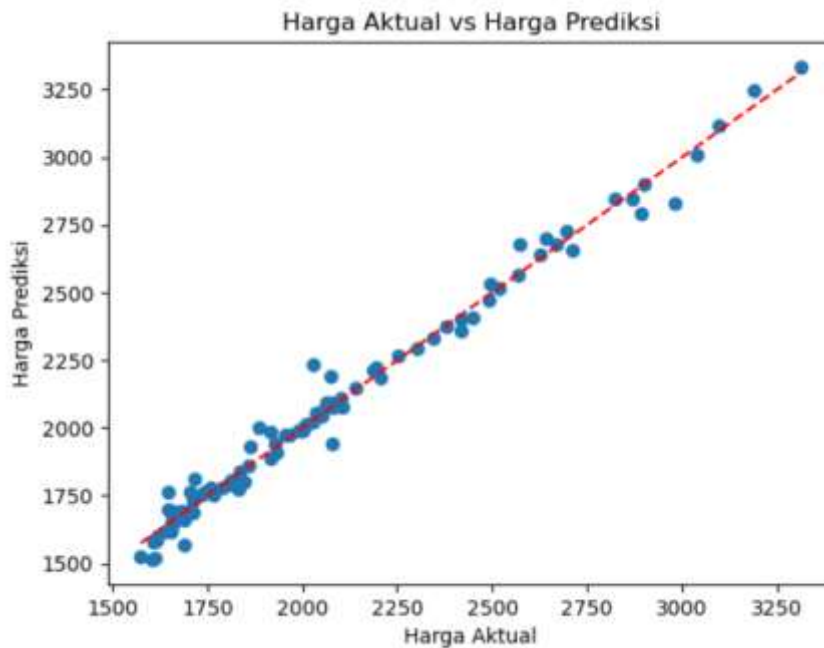
Nama	Dimensi
<i>Dataset</i> Historis Gas Alam	112 baris data & 7 kolom

Pada tabel diatas merupakan *dataset* yang akan digunakan untuk pengujian evaluasi dan validasi hasil, menggunakan data uji sebanyak 112 data dengan 7 kolom.

Tabel 4. 17 Parameter Model Uji

Model	Hasil
<i>Mean Squared Error (MSE)</i>	2823.5
<i>Mean Absolute Error (MAE)</i>	35.221
<i>R-Squared</i>	0.9850

Hasil dari model yang ditentukan terjadi penurunan tingkat akurasi yang cukup signifikan yaitu sebesar 98.50%, hal ini disebabkan karena data yang dipakai adalah data yang berbeda dari data sebelumnya atau data yang belum pernah diuji sebelumnya.



Gambar 4. 9 Grafik Harga Prediksi

Terlihat pada grafik diatas dapat dijelaskan bahwa hasil prediksi harga gas alam memiliki tingkat penurunan yang cukup signifikan yaitu sebesar 98.56%, dimana harga asli prediksi sesuai atau tepat dengan harga aktual atau harga yang sesungguhnya.

5. KESIMPULAN

Hasil dari implementasi dan pengujian yang telah dilakukan pada bab sebelumnya menunjukkan bahwa metode Regresi linier dapat diterapkan ke dalam model *machine learning* untuk menganalisa dan memprediksi harga gas alam dengan hasil nilai MSE sebesar 2823.5, MAE sebesar 35.221, dan *R-Squared* sebesar 0.9850. Dengan demikian, hasil tersebut menunjukkan bahwa metode Regresi linier memiliki tingkat akurasi sebesar 98.50% sesuai dengan nilai *R-Squared*.

6. SARAN

Penelitian ini menggunakan *dataset* dari investing.com, yang memiliki tautan atau link ke (<https://id.investing.com/commodities/natural-gas-historical-data>) data historis gas alam. Peneliti tidak membahas penggunaan *dataset* dari sumber lain. Peneliti dapat memberikan saran dengan membuat rekomendasi berikut:

- a. Penelitian ini dapat dilanjutkan dengan memasukkan model *machine learning* ke dalam aplikasi berbasis web atau lainnya.
- b. Penelitian ini dapat dilanjutkan dan dikembangkan dengan menggunakan data nyata/realset perusahaan.

7. UCAPAN TERIMAKASIH

- a. Terima kasih kepada Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya menyelesaikan penelitian ini.
- b. Terima kasih kepada Bapak Bambang Wisnu Widagdo, S.T., M.Sc.IT. selaku Dosen Pembimbing yang telah memberikan bimbingan dan motivasi kepada penulis.
- c. Terima kasih kepada teman seangkatan dan pihak yang terlibat atas doa, usaha dan motivasi yang telah dicurahkan.

- d. Terima kasih kepada keluarga penulis, khususnya kepada orang tua tercinta dan keluarga besar yang tidak dapat penulis sebutkan satu persatu yang telah memberi dukungan spiritual maupun materi serta doa kepada penulis.
- e. Terima kasih kepada kerabat-kerabat penulis yang tidak dapat penulis sebutkan satu persatu yang telah memberi dukungan serta doa kepada penulis.
- f. Tidak lupa juga mengucapkan terima kasih kepada diri sendiri atas ketekunan dan keteguhan hati selama studi ini.

8. DAFTAR PUSTAKA

- [1] A. P. Pratama and Y. B. W. Tama, "Prediksi Harga Komoditas Gas Alam Menggunakan Model Brownian Motion Dan Geometric Brownian Motion," *Proximal: Jurnal Penelitian Matematika dan Pendidikan Matematika*, pp. 73-81, 2023.
- [2] H. D. Bhakti and H. Abror, "Aplikasi Adaptive Neuro Fuzzy System(ANFIS) Untuk Mem-prediksi Kebutuhan Gas Bumi Indonesia," *Jurnal Teknologi Informasi dan Multimedia*, pp. 73-84, 2022.
- [3] R. R.Saputra, "PERBANDINGAN HASIL PREDIKSI HARGA GAS ALAM DUNIA MENGGUNAKAN FUZZY TIME SERIES RUEY CHYN TSAUR DAN FUZZY TIME SERIES LEE," 2023.
- [4] S. S, H. E and Jondri, "Model Autoregressive yang Dioptimasi Menggunakan Algoritma Genetika Untuk Memprediksi Ketersediaan Energi Sumber Daya Mineral di Indonesia," *Ind. Journal on Computing*, pp. 1-11, 2016.
- [5] R. C. Tsaur, "A Fuzzy Time Series Forecasting Model Based on Data Granulation," *Journal of the Chinese Institute of Engineers*, 2021.
- [6] G. N. Ayuni and D. Fitriyah, "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ," *Jurnal Telematika*, 2019.
- [7] A. A. Suryanto and A. Muqtadir, "PENERAPAN METODE MEAN ABSOLUTE ERROR (MEA) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI," *SAINTEKBU: Jurnal Sains dan Teknologi*, 2019.
- [8] H. e. a. Abror, "al. "Evaluasi Prediksi Konsumsi Gas Bumi Menggunakan Artificial Neural Network (Ann).",," *urnal Informatika Kaputama (JIK) 5.1* , pp. 108-115., (2021).
- [9] K. A. Revaldi, F. D. Cahyono, M. Hakimah and R. R. Muhima, "Penerapan Metode Regresi Linier Sederhana dalam Memprediksi Jumlah Kebutuhan Ekspor Migas dan Non-Migas di Indonesia," *Institut Teknologi Adhi Tama Surabaya*, pp. 268-275, 2023.
- [10] D. Kertayuga, E. Santoso and N. Hidayat, "Prediksi Nilai Ekspor Impor Migas Dan Non-Migas Indonesia Menggunakan Extreme Learning Machine (ELM)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, pp. 2792-2800, 2021.
- [11] W. A. Medyanti, M. Faisal and H. Nurhayati, "Optimasi Metode Single Exponential Smoothing Dengan Grid Search Pada Prediksi Nilai Ekspor Migas," *Sintech Journal*, pp. 59-69, 2024.

- [12] H. Abror, A. Triono, H. D. Bhakti and E. E. DwiSaputri, "Evaluasi Prediksi Konsumsi Gas Bumi Menggunakan Artificial Neural Network (Ann)," *Jurnal Informatika Kaputama (JIK)*, pp. 108-115, 2021.
- [13] P. F, V. G and D. É, "Scikit-learn: Machine learning in Python," *the Journal of machine Learning research*, pp. 2825-2830, 2011.
- [14] H. Colak, T. Partal and İ. Ceylan, "Modeling the Hydraulic Conductivity of Soil Using ANFIS and MLR Approaches," *Environmental Monitoring and Assessment*, 2011.
- [15] Edison and David, "Prediksi Volume Penjualan Gas PT PGN (Persero) Menggunakan Regresi Linear Berganda," *Journal of Digital Ecosystem for Natural Sustainability (JoDENS)*, pp. 14-18, 2021.
- [16] M. Y. Fathoni, Darmansah and A. T. Rosalinda, "Perbandingan Holt' Method Dan Winter's Method Pada Peramalan Ekspor Minyak Dan Gas Bumi Di Indonesia," *Jurnal Teknik Informatika dan Sistem Informasi*, pp. 2509-2520, 2022.