

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Optimizing Predictive Models for LNG Production and Temperature Forecasting

Frederick Etumnu^{*a*}, Prof. Ipeghan. J. Otaraku^{*b*}, Dr. Matthew Ehikhamenle^{*c*}

^a PhD Student, Center for Information and Telecommuication Engineering (CITE), University of Port Harcourt, Rivers State, Nigeria ^b Former Director, NLNG Centre for Gas, Refining & Petrochemicals, University of Port Harcourt, Rivers State, Nigeria

^c Assistant Director, Center for Information and Telecommuication Engineering (CITE), University of Port Harcourt, Rivers State, Nigeria.

ABSTRACT

Predicting LNG (Liquefied Natural Gas) production and associated rundown temperatures is crucial for efficient operations in the LNG industry. Accurate predictions enable better resource allocation, cost management, and maintenance planning. In this study, we compared the performance of various machine learning algorithms for their ability to forecast LNG production and rundown temperatures. We evaluated the predictions using error metrics, specifically the R2score and Root Mean Square Error (RMSE). The study revealed that the K-neighbors algorithm outperformed other models, exhibiting the highest R2 score and the lowest RMSE, making it the optimal choice for accurate predictions. Additionally, the feature selection process played a critical role in model accuracy, with certain features proving to be more suitable for training the models. Detailed error analysis and data visualization further demonstrated the effectiveness of the chosen algorithm. This study highlights the significance of selecting the right predictive model and features for LNG production and temperature forecasting, providing valuable insights for optimizing operations in the LNG industry.

Keywords: LNG Production Forecasting, Rundown Temperature Prediction, Machine Learning Algorithms, R2 Score Analysis, Root Mean Square Error (RMSE) Evaluation and KNeighbour

1. INTRODUCTION

The world is moving to natural gas from fossil fuels. Natural gas has supplanted coal as the world's cleanest and most ecologically friendly energy source due to its lower carbon emissions and affordability (Mofid & Fetanat, 2019; Salehi, 2018; Wang, 2017). Natural gas use is expected to rise 40% between 2014 and 2040. (BP, 2017). Jackson, Eiksund, and Brodal found that natural gas-powered plants provided 37% of fossil energy in 2030, up from 30% in 2013. (2017).

LNG is rapidly replacing fossil fuels as a key energy source worldwide because to its reduced greenhouse gas emissions and cleaner burning. These LNG benefits have attracted attention recently due to the energy crisis (Sang et al 2020). Natural gas is mostly transported via liquefaction and pipelines. Energy companies have routinely employed liquefying natural gas to transport it across long distances instead of pipes. The pipeline pressure difference limits pipeline gas supply, the transit route is rigid, and long-term contracts are needed (Lee et al 2020).



Figure 1. Global trend in energy consumption (U.S. Energy Information Administration, 2022)



Figure 2. History and Projection of U.S. Energy Consumption (U.S. Energy Information Administration, 2022)

Liquefaction is the process by which gases are transformed into liquids. Liquefied natural gas (LNG) is natural gas that has been cooled to -162 degrees Celsius while maintaining an atmospheric pressure. Natural gas is more convenient to transport in this state since only one-six hundredth as much is utilized as when it is in its gaseous state (Khalilpour & Karima, 2009).

When temperatures drop below the critical temperature of natural gas, the gas changes phase and becomes a liquid. Refrigerants must be employed to reach such low temperatures, and the heating curves of these refrigerants must be as similar to the cooling curve of natural gas as is practically possible. A refrigerant is the fluid used to transport heat in air conditioning and refrigeration systems. Maintaining a cold temperature for the LNG product in LNG facilities is crucial for its storage and transport.

As a consequence, production will suffer if this section is inefficient. There is a need for environmentally friendly, fast-acting refrigerants that also save energy. Over the years, a number of distinct LNG production methods and refrigerant types have arisen, each with its own unique process configuration. Some examples are the turbo-expander process, the cascade process, and single/dual mixed refrigerant (SMR/DMR) technology. The primary distinctions between these methods are capital and operational costs. A company's capital expenditures (CAPEX) and operating expenditures (OPEX) are affected by a number of factors, including its production capacity, the quantity of equipment, and the cost of labor. However, the use of mixed-refrigerant (MR) processes greatly increases the complexity of process design and operation due to the increased number of thermodynamic interactions, making process management and optimization exceptionally challenging (Shukri, 2004). Factors like as desired temperature range, ease of access, cost, and past experience all factor into the decision of which refrigerant to use. Ethane and propane, for instance, may be on hand in a natural gas processing facility, whereas ethylene and propylene are on hand in an olefins factory. There is a substantial danger of contamination if propane or propylene were used in an ammonia facility, but ammonia might serve the same function. Due to their inability to catch fire, halocarbons have seen widespread use. The Propane Precooled Mixed Refrigerant (C3MR) system is one of the most popular refrigeration methods utilized today. Before entering the mixed refrigeration system consisting of methane, ethane, propane, and nitrogen, the liquefied natural gas is precooled to -35Oc in a propane refrigeration system (Bahadori et al., 2014).

2.1 Methodology

2.1.1 Material

The material used in this research include: An artificial intelligent (Python programming language) software, PI Process book software 2015 version 3.6.2.271, PI datalink, Visual studio (VS) code editor and the Microsoft Excel 365. Python is an interpreted, high-level programming language that may be used for a wide variety of projects. The principle behind its design prioritizes the readability of the code by heavily indenting it. The PI process book and PI datalink add-in were basically used for data collection from the plant site. While the VS code editor is mainly an Integrated Development Environment (IDE) source code editor used to debug, highlight syntax and for coding of the GUI script. It is a user-friendly coding environment.



Figure 3. Research design framework for the optimization of LNG production using AI.

2.1.3 Process Optimization Description

Figure 4 shows the sequential order or steps used to achieve the aim and objectives of this research. It depicts the schematic breakdown of the optimization process using the artificial intelligence data driven approach.



Figure 4. Process description of Artificial Intelligence optimisation

2.1.4 Data Collection

The data collection for this work was done using PI Processbook 2015 software version 3.6.2.271 R2 and PI datalink. PI processbook is an OSIsoft vendor software that enable users to retrieve real-time data from the PI system which is linked to a live process plant. The software application can create dynamical graphical display, trends from historical and real time data. To retrieve the data used for the work, the PI datalink was connected to the PI server and then to the liquefaction plant via several process control schemes as shown in figure 3.3. The PI datalink is a Microsoft Excel add-in feature

linked to the PI software. The PI datalink's sample data multiple value function was used to retrieve about 10 years liquefaction unit data set at an hourly interval.



Figure 5. PI System data Collection scheme

2.1.5 Data Pre-processing and Cleaning

In this phase, three important steps were followed which are the data importation, cleaning of data, and exporting of data for next stage.

• Importation of library and loading of data

The data used for this project contains about 70,000 rows and 43 columns of about 10 years sample data. Microsoft excel was used to get a preview of how the data looks. The images below show the preview of the data in MS excel. The data seen above, as expected, contains quite a lot of invalid data and errors. Jupyter notebook version 6.4.5 environment was used, which comes as part of a software called Anaconda. Anaconda comes with a lot of libraries that will be so useful for me in this project. The first library that I used is the Pandas. Pandas with the help of python programming language was used to load the data into the Jupyter notebook environment. Below is the image that shows the data already imported into Jupyter notebook using Panda's python library. Figure 3.2 consist of data collected from the industry. Figure 3.3 displays the data loaded for the optimization process (first five rows of the data).

• Cleaning the data.

3. As mentioned earlier, the data loaded above contains many invalid inputs from the data collection source. Neglecting these errors will lead to making an inaccurate model. Hence, there is need for me to handle and get rid of them. These involve creating the data in a floating type. Data cleansing involved the following processes.

- i. Eliminated erratic values in LNG flow column.
- ii. Found out the error values in "MCHE cold bundle dp" column.
- iii. Found out the errors in "volumetric MR flow to MR component discharge" column.
- iv. Found out errors in "percentage Nitrogen in MR" column.
- v. Removed unnamed columns.
- vi. Eliminated all the null values.

		1	and the second					16			1 14	
				CONTROL W	VRIATELE IN MICHEL	CACE		1		1 1		
0.Fire	i Miji Rundowe Teorgentekere	MRI Converting	pag Operating Margin	MRT Compressor Rection Pressure	MEDIE Warm Sygnass Textgeneditors	Mill sellens Tanganoster	MRI Expander Laber Terrgeroture	Temp Top Temp Tep	A MEL OF Votion Votion Openning	MORE Cost	Mill Flow In: Mil somp. distriction	-
12896.043	I. (100 Apo)	2.8154	18,3890	1349	-41,001	-36.9446	437-4629	4 1000	00.0184	169.7031	Ale Dana	_
10803-041	9. (88.8358	2.9219	18.0275	2.3219	-#1.3048	-37.0312	-127-4072	A 1890	02.4368	100.0208	No Date:	
12110.108	8. 193 8358	3.1083	. 15.0399	. 2.3297	42.403	-37.6793	127.4081	4,1913	02.0184	100.3028	NetDete	
12121.300	7 (188,8252	2.9685	15 (0350	2.347	-42.108	10.110	427,294	A 1967	04.0094	154.3478	NetDate	
10391-246	7 160 8404	1.010	18.6330	1.344	-43.000	07 0600	127.264	4.2996	05 7162	161.3129	Per Date	
12220 209	4 160 6104	3.0814	28.4161		-42-9164	-07.1172	-47.204	4,8941	86.0043	164,2913	No Cara	
12952-518	1 105 8308	3.1366	29.7883	2 3328	-42.241	-17 1367	-127.5338	4,1216	85.9445	188,7578	No Own	
. 12M7.016	4 /199.8356	3.121	21.190	- 7.3059	-01.9482	17.1623	477.3438	4.909	04.0722	100.0104	An Den	
12103.007	3	2.9889	28.7847	1.3271	-42,9829	07.1841	477.2528	4.0278	89.6799	142,0198	Ne Date	
11997.072	C	1.000	18.7990	2.3178			117,200	4.1992	82.4918	102,0072	No Date	
12110.694	A	2.917	17.6780	2 3068		38.9239	127.4962	4.6771	302,7999	160.8281	No Date	
12801.218	8 (100.7764	2 7190	18.8789	3.2044	-62.0184	34,6310		4858	73.0003	163,8008	No Data	
11757/043	8 183,8908	2.184	7,087	2,2034	-41.9034	-74.6804	- 526 3067	4.8810	09.8133	144.8028	TALEMA	
11090.007	9	2.000	1.4486	2.2199	-41.9081		120.5404	4.1300	08.0068	142,7094	SkiDen	
11853.052	3 -153.8298	2,9199	3,8850	2.2734	-41,9981	38.1250	-120.1488	4.1853	01.6786	143.9078	No Date	
11412,007	2	0.000	2.5492	0.2779	-41 910	38,8472	-120,7908	4.1899	10.0498	142.4018	No Date	
11897,000	6	0.000	4.0840	0.2963	AU-0181	-34.9690	120.1041	4.8829	100,7143	121.0404	Nether	
11557.668	4 (168,802)	3.1455	4.0890	2.3029	-42.2583	-36.0677	129,6968	4.3309	. 70.5679	143.4375	PAGEDatte	
11682.707	101.0203	3.0071	6.0787	1.1394	-41-8771	-26.0469	426.3054	4.260	72,7282	152.0078	No Deta	
11380.201	2 188.8202	3.4636	8.3824	1.04	-42.5867	06.0VT6	100.704	4.3799	71.6248	141.1953	All Della	
11002.345	2 192 8 954	2.1100	0.0910	2.2794	-41.001	-38.9463	428.2012	4.801	19.0094	100.1028	No Deta	
11573-648	2	- 5.2884	7.3139	2.7/84	-42.2383	-78.2852	120.1781	4.4571	77.7147	147.9944	No Dela	
71820.418	111.890	3.0871	1,7389	2.1711	-41,998		127 8182	4 8580	92.0947	101.6112	TRY DWG	
	100.0014	1.004	8.1881	3.3943	-42.3034	(4.440)	107,0004	4 1414	00.045.0	466.3128	No Date	
11756-934	8. 169,7528	3.0172	6.78%	2.3404	46.898	36.0890	120.0618	8.8638	11.3671	153.8068	THE DWA	
11772.574	7. 199 MOR	3.385	7.6690	7.840	-47.6045	-38.1890	120.126	4.3252	77.8949	155 8 391	NO CHER	
11007.328	1 -100.7910	3.0368	1.0999	0.3102	-61,9032	25.5601	420.2011	4.1309	78.307	149.0008	Ne Data	
11001.349	e 155.8908		8,1980	1.3408	-42,038	-38.0795	- 120.0782	4,3904	70.7638	110.6813	The Elimin	

Figure 5. Data library from Microsoft Excel

t	OAD D	ATA												
1.1.1	ilepath = " f = pd_read f.head()	namonta je excel (file	alos" path, skiprov	p=1]										
	Unsamed: Ø	Unnamed: 1	UNG Flow	LNG Rundown Temperature	MR Operating Margin	PR Operating Margin	MR Corepressor Suction Pressure	MCHE Warm bypass Temperature	MR return Temperature	MR Expander Inlet Temperature		Unnamed: 33	Low Pressure Propare kettle level	Medium Pressure Propane/Natural Gas kettle level
0	C Ne	NW	120804075	150420312	2,659555	18.209555	2317578	418358	-16.9973	127,463(8	-	No Dida	57,310078	64,28125
1	1.100	NeV	12093041016	+150834902	2.921342	19.027308	2.327148	-41,934805	-37,029156	-127.407227	-	No Data	56773438	64,299062
2	in the second	NaN	12115.566018	+150.834961	3.106348	15.838821	2.329688	-42,452344	-37.0708.12	-127.458055	1	No Data	57,652031	64,648638
3	en time	348	12121388672	150,825195	2,939482	19.634064	234707	42.186477	-37,117188	-127.275391	-	No Deta	57213672	54,261719
	(c) 19200	1000	12070 336254	INVERSE	1,01808	12,521/08	1141211	42,000/10	-17.002001	127.250477		No Dete	56,042,011	441477

Figure 6 Previewed data for optimization process.

Rang	geIndex: 70129 entries, 0 to 70128		
Data	a columns (total 43 columns):		
#	Column	Non-Null Count	Dtype
0	Unnamed: 0	0 non-null	float6
1	Unnamed: 1	0 non-null	float64
2	LNG Flow	70129 non-null	object
3	LNG Rundown Temperature	70129 non-null	object
4	MR Operating Margin	70129 non-null	object
5	PR Operating Margin	70129 non-null	object
6	MR Compressor Suction Pressure	70129 non-null	object
7	MCHE Warm bypass Temperature	70129 non-null	object
8	MR return Temperature	70129 non-null	object
9	MR Expander inlet Temperature	70129 non-null	object
10	LNG Shell Temp Top Bundle	70129 non-null	object
11	LMR JT Valve Valve Opening	70129 non-null	object
12	MCHE Cold bundle dp	70129 non-null	object
13	Volumetric MR Flow to MR comp. discharge	70129 non-null	object
14	HMR Flow	70129 non-null	object
15	HMR/MR Ratio	70129 non-null	object
16	LNG Flow Ex MCHE	70129 non-null	object
17	MCHE Warm Bypass Temperature	70129 non-null	object
18	MR IGV Position	70129 non-null	object
19	MR Vessel Level	70129 non-null	object
20	MR Vessel Pressure	70129 non-null	object
21	Percentage Nitrogen in MR	70129 non-null	object
22	Percentage Methane in MR	70129 non-null	object
23	Percentage Ethane in MR	70129 non-null	object
24	Percentage Propane in MR	70129 non-null	object
25	Percentage Butane in MR	70129 non-null	object
26	Nitrogen Make-up flow	70129 non-null	object
27	Methane Make-up flow	70129 non-null	object
28	Ethane Make-up flow	70129 non-null	obiect

Figure 7. LNG data before cleaning

0	LNG Flow	32024 non-null	float64
1	LNG Rundown Temperature	32024 non-null	float64
2	MR Operating Margin	32024 non-null	float64
3	PR Operating Margin	32024 non-null	float64
4	MR Compressor Suction Pressure	32024 non-null	float64
5	MCHE Warm bypass Temperature	32024 non-null	float64
6	MR return Temperature	32024 non-null	float64
7	MR Expander inlet Temperature	32024 non-null	float64
8	LNG Shell Temp Top Bundle	32024 non-null	float64
9	LMR JT Valve Valve Opening	32024 non-null	float64
10	MCHE Cold bundle dp	32024 non-null	float64
11	Volumetric MR Flow to MR comp. discharge	32024 non-null	float64
12	HMR Flow	32024 non-null	float64
13	HMR/MR Ratio	32024 non-null	float64
14	LNG Flow Ex MCHE	32024 non-null	float64
15	MCHE Warm Bypass Temperature	32024 non-null	float64
16	MR IGV Position	32024 non-null	float64
17	MR Vessel Level	32024 non-null	float64
18	MR Vessel Pressure	32024 non-null	float64
19	Percentage Nitrogen in MR	32024 non-null	float64
20	Percentage Methane in MR	32024 non-null	float64
21	Percentage Ethane in MR	32024 non-null	float64
22	Percentage Propane in MR	32024 non-null	float64
23	Percentage Butane in MR	32024 non-null	float64
24	Nitrogen Make-up flow	32024 non-null	float64
25	Methane Make-up flow	32024 non-null	float64
26	Ethane Make-up flow	32024 non-null	float64
27	Propane Make-up flow	32024 non-null	float64
28	Ambient Condition	32024 non-null	float64
29	Scrub Column inlet Temperature	32024 non-null	float64
30	Natural Gas temp at MCHE warm bundle	32024 non-null	float64
31	Low Pressure Propane kettle level	32024 non-null	float64
32	Medium Pressure Propane/Natural Gas kettle level	32024 non-null	float64
0 \$	1 🤠 Python 3 (ipykernel) Idle		

Figure 8 LNG data after cleansing showing floating columns.

2.1.6 Model Selection Process

Several algorithms were tested on the two data set generated to determine the best LNG production and LNG temperature. In this process, LNG production (Uni-output) and LNG production and LNG rundown temperature (multi-output) were predicted. The data set up were loaded alongside the pandas, numpy and sklearn libraries. List of libraries and function loaded are listed below viz:

- Functions for calculating absolute, mean and standard deviation from the Numpy library
- Algorithm program from the sklearn library
- Programs chaining the algorithm available from the sklearn library
- Programs to process the data to make them suitable for the algorithm to work with
- Programs for metric calculations





The following algorithms were imported:

- Linear Regression
- KNeighbors Regressor
- Decision Tree Regressor
- Random Forest Regressor

These are algorithms are considered fit for the nature of task to be solve. This task is a numerical input, continuous target problem. It is numerical input because all our input features are of numeric datatype, and it is continuous (or numerical) target because the variables that I intend to predict are also numeric in nature.

• Linear Regression

The multiple linear regression model was used in the process due to several independent variables. The underlying equation for the multiple linear regression.

$$\beta_o + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \epsilon$$
 Equation 1

- y_i is the dependent or predicted variable
- β_0 is the y-intercept.
- β₁ are the regression coefficients representing the change in y relative to a one-unit change in xi₁, xi₂,... xip respectively.
- β_p is the slope coefficient for each independent variable
- ϵ is the model's random error (residual) term.

The cost function provides the best possible values for $\beta 0$, $\beta 1$, $\beta 3$... βp to make the best fit line for the data points. The algorithm converts this problem into a minimization problem to get the best values for β . The error is minimized in this problem between the actual value and the predicted value.

Equation 3

minimize $\frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2$	Equation 2
--	------------

 $J = \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2$

• KNeighbors Regressor

The straightforward algorithm K nearest neighbors predicts the numerical target based on a similarity metric while storing all of the available cases (e.g., distance functions).

Distance function

Euclidean	$\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$	Equation 4
Manhattan	$\sum_{i=1}^{k} x_i - y_i $	Equation 5

Equation 6

2.1.6 Model Development

The algorithm with more acceptable metrics from the model evaluation was selected.

This works best for the prediction to be performed. The following steps were used to prepare the machine learning model and load library dataset and the useful libraries model.



Figure 10. Flow chart for Model Development

The KNN algorithm was used to train the model.

Model Testing and Input Optimisation Using Optimisation Algorithms

The model was combined with several optimization algorithm on new set of inputs from another LNG plant data. The optimisation algorithm uses the model to find the best control variables or set points that will give an optimised LNG flow is shown below. Then compared with the new result.

The diagram below shows the model testing algorithm.





3.0 Result and Discussion

3.1 Model performance (Kneighbor Regressor)

Table 4. 1 Error analysis of algorithms

S/No	Algorithm	R2 - Score	RMSE
1.	Linear Regressor	0.9411	628.9384
2.	K-neighbors	0.9853	313.8738
3.	Decision Tree	0.9748	411.4508
4.	Random Forest	0.8981	826.8076

Table 2. Error	Analysis on th	he Algorithm	performance	for the actual	LNG flow a	and actual 1	LNG Rundown	temperature

Data Row No	Actual LNG Flow	Actual LNG	Rundown	
		Temperature		
18778 13575.379883		-145.473633		
10519	11672.666016	-145.039062		
31757	13300.155273	-145.786133		
6400 12687.996094		-145.610352		
17354 11994.082031		-145.502930		

Table 3 Error Analysis on Algorithm performance for the Predicted LNG flow and predicted LNG Rundown temperature

Data Row No	Predicted LNG Flow	Predicted LN	G Rundown	
		Temperature		
0	13363.826953	-146.179688		
1	12060.130078	-145.521484		
2	13391.834180	-145.701172		
3	12811.619531	-145.620117		
4	11980.771289	-145.903320		

Table 4 Comparison of predicted and actual LNG flow using the K-Neighbor regressor Model.

Actual LNG Flow	Predicted LNG Flow	Error
13575.3799	13363.8269	211.553
11672.66	12060.1301	387.4641
13388.1553	13391.8341	3.6788
12687.9961	12811.6195	123.6234
11994.082	11980.7713	13.3107
Actual Rundown Temperature	Predicted Rundown Temperature	Error
(°C)	(°C)	
-145.4736	-146.1797	0.7061
-145.0391	-145.5215	0.4824
-145.7861	-1455.7012	0.0849
-145.6104	-145.6201	0.0097
-145.5029	-145.9033	0.4004



Figure 12. LNG Rundown Temperature Prediction



Figure 13. Final Visualization of Algorithm performance (actual LNG flow)



Figure 14. Final Visualization of Algorithm performance (LNG rundown temperature)

Table 5.. Improved prediction after hyper parameter tuning.

Data Row No	Predicted LNG Flow	Predicted LNG Rundown
		Temperature
0	13509.923222 -145.550721	
1	12170.430456	-145.727184
2	13404.951410	-145.707795
3	12735.611069	-145.785677
4	11952.358026	-146.056697



Figure 15. GUI run test 1

11672.66601	UNG Rundov 5625 -145.0390625	en T - MR IGV Positio 52.58515930175	2m HMR Flow 578 17607.126953125	HR/MR Ratio 6.21599340438842	% Nitrogen in MR 2.85791015625	- % Methane in MR 42.20449447631836	
	% Ethane in MR 43.97090148925781	% Propane in MR 10.588183403015133	% Butane in MR 0.2811718583106994	Ambient Condition 27.05110549926757	LP Propane/MR Ket 60.626953125	tle Level	
load file Ci/Users/owner/Downloads/Ing/Final work/inputs.csv					Next	Next Input	
stput:							
mber of	iteration: 254						
mber of stimized 2.3192777	iteration: 254 Input: 2004543					Optimize	
mber of timized 1.3192777 0684.3070 0.0926905	iteration: 254 Input: 2804543 49236737 2988536					Optimoz Clear	
mber of ctimized 2.3192777 0684.3070 5.0926905 20033314	iteration: 254 Input: 2004543 49236737 2988536 56656656				j	Optomoz Clear	
mber of timized 1.3192777 0684.3070 0.0926905 20033314 24246373 0.0985733	iteration: 254 Input: 2804543 49236737 2988536 56456456 1431627 1238709				J	Optimize Clear	
mber of cimized 2.3192777 0654.3070 0.0926905 20033314 24246373 1.0985733 7234975	iteration: 254 Input: 2804543 49236737 2988536 56456456 1431627 1238709 797636					Optimae Clear	
mber of timized .3192777, 664.3070 .0926905 20033314 24246373 .0985733 .7234975 15709267 2424416	iteration: 254 Input: 2804543 49236737 2988536 56456456 1431627 1238709 787636 04045453 522381					Optimae Clear	
mber of ctimized 1.3192777, 664.3070 0.0926905 20033314 24246373 1.0985733 1.7234975 1.5709267 1.3664416 1.7547947	Iteration: 254 Input: 2004543 49236737 2988536 56456456 1431627 1238709 797636 94045453 5222351 947948					Optimae Clear	

Figure 16. GUI run test 2



Figure 17. GUI run test 3



Figure 18. relationship between predicted and actual LNG flow

Table 6. Improved prediction for hyperparameter tuning.

Actual LNG Flow	Predicted LNG Flow	Error	
13575.3799	13509.9223	65.458	
11672.66	12170.4305	497.7645	
13388.1553	13404.9514	16.7961	
12687.9961	12735.6111	47.615	
11994.082	11952.358	41.724	
Actual Rundown Temperature	Predicted Rundown Temperature	Error	
(°C)	(°C)		
-145.4736	-145.5507	0.07061	
-145.0391	-145.7272	0.6881	
-145.7861	-145.7077	0.0784	
-145.6104	-145.7856	0.1752	
-145.5029	-146.0567	0.5538	



Figure 18. Visualization of Algorithm performance (based on LNG flow)



Figure 15. Visualization of Algorithm performance (based MR return and predicted temperature)

3.2 Discussion

To interpret the algorithm results, we compare the errors (R2 error and Root Mean Square Error (RMSE)). A good prediction should have a high R2_score (range between 0 to 1) and a low RMSE. The RMSE has to be as low as possible. The lower the RMSE, the more accurate the prediction. The goal here is to determine which of the algorithms will give the best R2_score and RMSE. The algorithms were trained with the train dataset command that was selected by intuition to learn how to predict LNG production (uni-output regressor). It was deduced that all the different algorithms' predictions had very good R2 scores. The K-neighbours algorithm gave the best regressor, with an R2 score of 0.9853, followed by the decision tree, whose R2 score was 0.9748, and the linear regressor, whose score was 0.9411. Although still a very good score, Random Forest had the lowest R2 score of 0.8981 amongst the other algorithms. Observing the RMSE metrics, it can also be seen that the K-neighbours had the lowest error in their prediction compared to the others. The RMSE of K-neighbours was 313.8738, and the Decision Tree came closest with an error of 414.4508.

Comparing Figure 12 and Figure 14, we see that the features chosen by intuition serve as the best to train and build the model for predicting LNG production and LNG rundown temperature. From the figures, we chose K-neighbours because it gave the best R2-score and the lowest RMSE. Figure 15 shows the actual part of the output in the test dataset, while Figure 16 shows the predicted output in the test dataset from the K-neighbour regressor model.

Using K-Neighbours, the actual LNG flow and rundown temperatures were compared to what was predicted, and the errors were also calculated. The lowest error (3.6788) was computed when the flow of the LNG was 13388.1553 and the predicted flow was 13391.8341; next to the actual flow of 11994.082 and a predicted flow of 11980.7713; the error computed by the K-neighbours was 13.3107. The corresponding temperatures are (-145.7861 and -145.7012; -145.5029 and -145.9033) while the errors are 0.0849 and 0.4004 See Table 1 above. Figure 17 and Figure 18 show the scatter plot representing the relationship between the actual output and the predicted output of LNG production and LNG rundown temperatures, respectively. We

observe a sharper, high-positive correlation in Figure 18, which explains that the difference between the actual LNG production and predicted LNG production is close and within an acceptable range.

Acknowledgements

Special appreciation goes to Prof. Ipeghan J. Otaraku and Dr. Matthew Ehikhamenle for providing deep guidance and supervision throughout the period of this research work. Profound thanks also goes out to Prof B. O. Omijeh, director, Centre for Information and Telecommunication Engineering (CITE) and all the staff of the centre for the support all through the course of this project work.

References

Mofid, N. & Fetanat, A. (2019). Liquefied Natural Gas (LNG) as a Promising Fuel for Road Transportation in Iran. Energy Sources, Part B: Economics, Planning, and Policy, 14(5), 331-337.

Salehi, A. (2018). Liquefied natural gas: A review on development and environmental impacts. Renewable and Sustainable Energy Reviews, 81, 2088-2109.

Wang, X. (2017). The prospects for natural gas as a transport fuel in China. Energy Policy, 107, 640-647.

BP. (2017). BP Energy Outlook 2017. Retrieved from https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/energy-outlook/bp-energy-outlook-2017.pdf

Jackson, E., Eiksund, G., & Brodal, H. (2017). Liquefied natural gas in Norway: A study of the potential for developing a new export industry. International Journal of Energy Economics and Policy, 7(6), 117-127.

Sang, Y., Jia, Y., Ma, X., Zhang, X., Cai, Z., & Li, Y. (2020). Optimization of a liquefied natural gas supply chain with carbon capture and storage using mixed-integer linear programming. Journal of Cleaner Production, 252, 119855.

Lee, J., Kim, Y., Lee, J., & Lee, C. (2020). A study on the economic feasibility of small-scale liquefied natural gas plants in Korea. Energy Policy, 138, 111234.

Khalilpour, R., & Karima, B. (2009). Natural gas liquefaction technologies: A brief review. Journal of Natural Gas Chemistry, 18(4), 443-452.

Shukri, M. N. (2004). The use of mixed-refrigerant processes for natural gas liquefaction: A review. Applied Thermal Engineering, 24(18), 2757-2775.

Bahadori, A., Vuthaluru, H., & Ghalami, M. (2014). LNG production using propane pre-cooled mixed refrigerant process. Journal of Natural Gas Science and Engineering, 21, 1130-1143.