Thesis

Forecasting Natural Gas Prices in the West European Market
A Comparative Analysis of Forecasting Models

Author
Richard Constandine Allan

National and Kapodistrian University of Athens

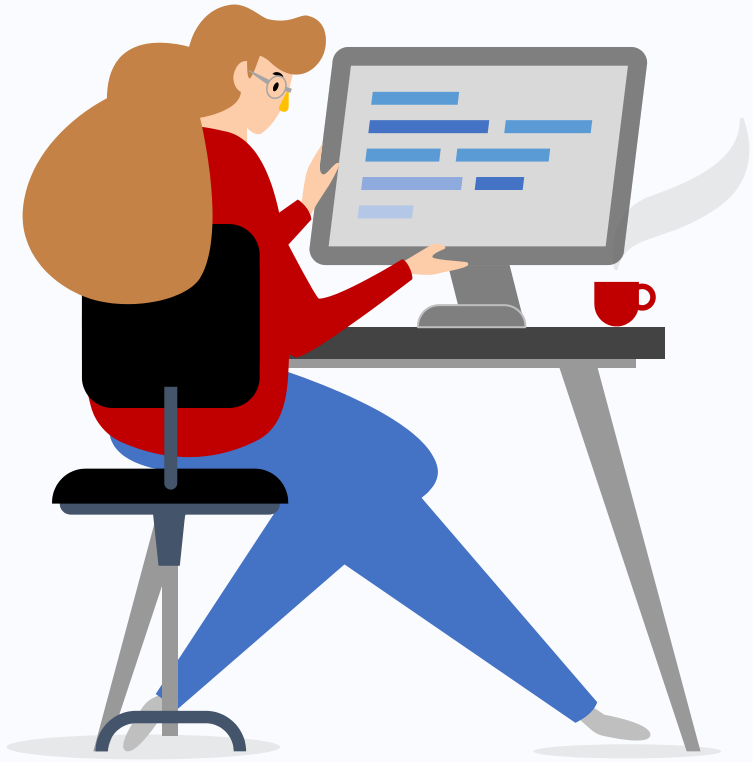
Department of Economics

MSc. Business Analytics, Administration and Information Systems

Supervisor
Dr. Kyriakos Drivas

Athens, June, 2023

Table of Contents



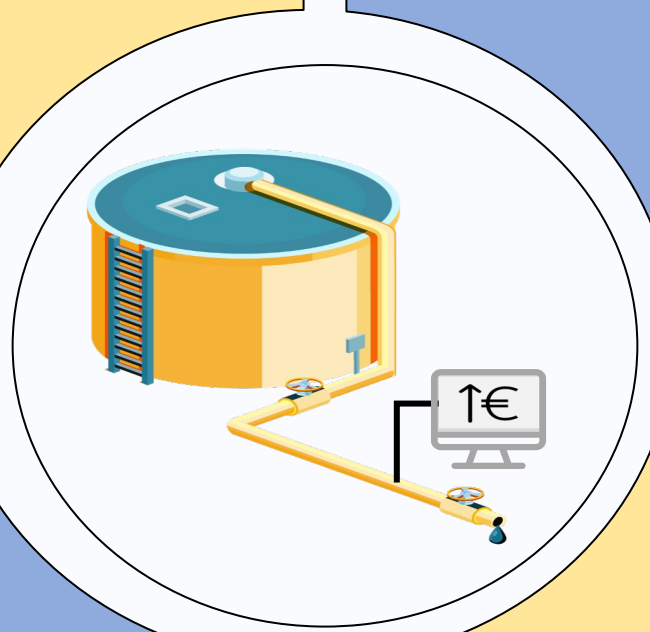
1. Introduction

Key Points

- Second Largest Energy Source
- West Europe's high demand for natural gas and its reliance on imports
- Challenges for forecasting models due to the complexity of the European energy market

Literature Review

- Wide use of Time Series models for natural gas price forecasting
- Neural Networks and Vector Machines outperform statistical models
- Combination of methods



Models Used for Forecasting Natural Gas Prices

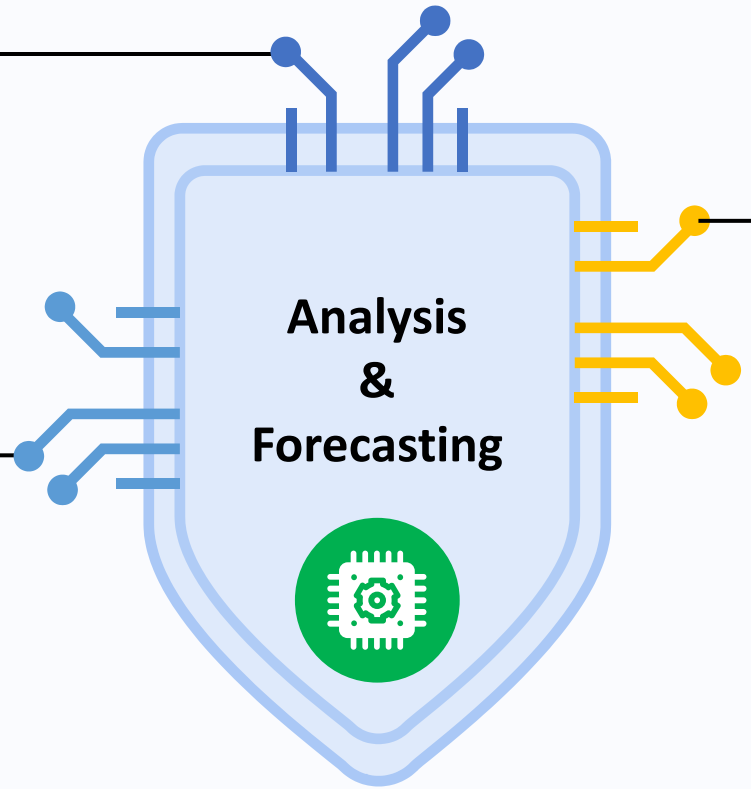
- Statistical Models such as ARIMA, ARIMAX, GARCH and VAR
- Machine Learning Models such as SVM, SVR, ANN, GPM and GPR

Variables Used for Forecasting Natural Gas Prices

- Prices of Energy Commodities
- Temperature & Heating Degree Days
- Natural Gas Storage Campacity
- Price of Carbon Credits

Aminu, 2019; Azadeh et al., 2012; Brown & Yucel, 2008; Erdogdu, 2010; Hosseinipoor et al., 2016; Hulshof et al., 2016; Li et al., 2017; Mu, 2007; Nick & Thoenes, 2014; Lv & Shan, 2013; Salehnia et al., 2013; Seo, 2021; Su et al., 2019

2. Aim of Study



01

Identification of the most accurate model for forecasting natural gas prices.

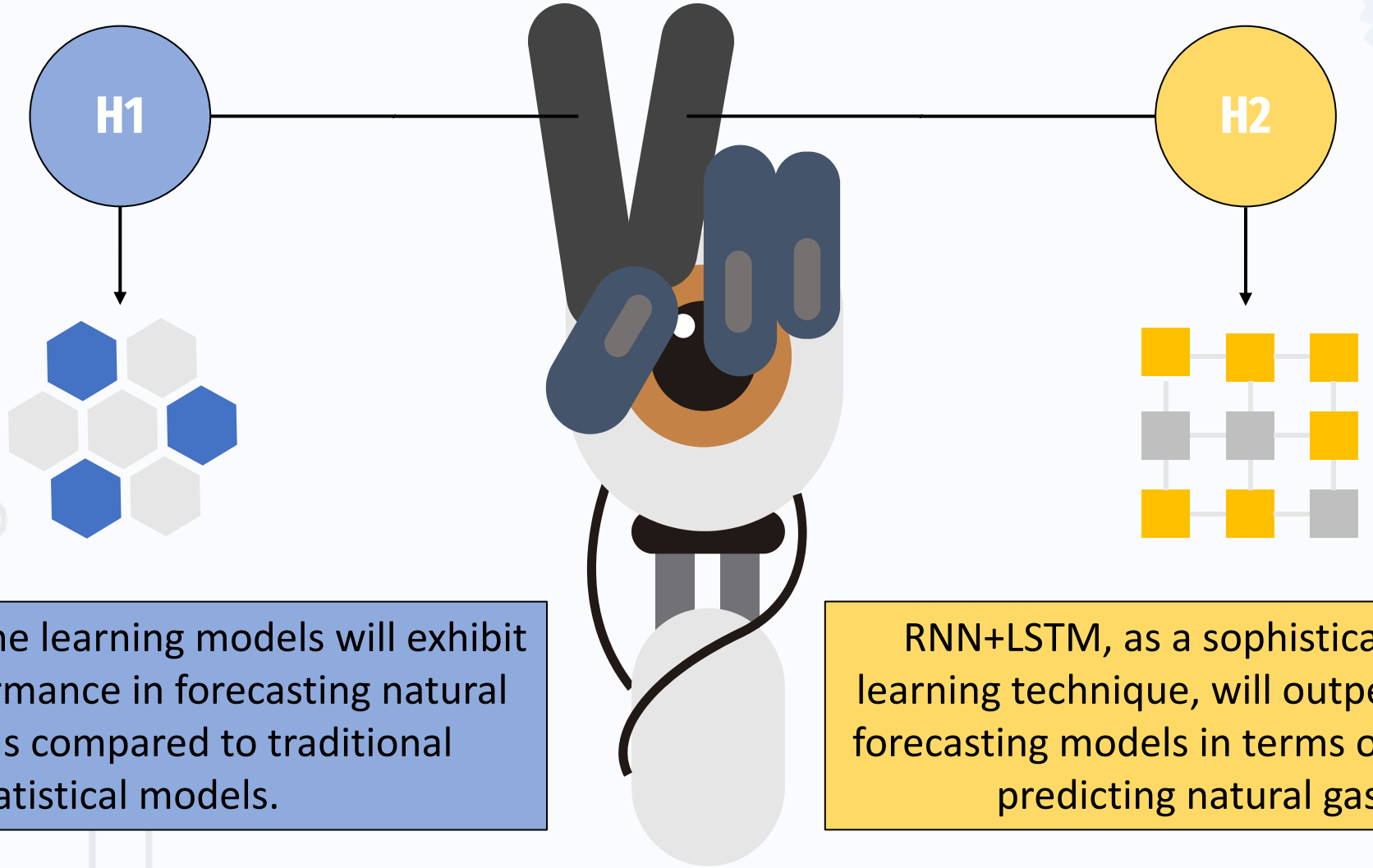
02

Evaluation of the effectiveness of complex machine learning models in comparison to traditional statistical models.

03

Improving the accuracy of forecasting techniques for natural gas prices.

3. Research Hypotheses

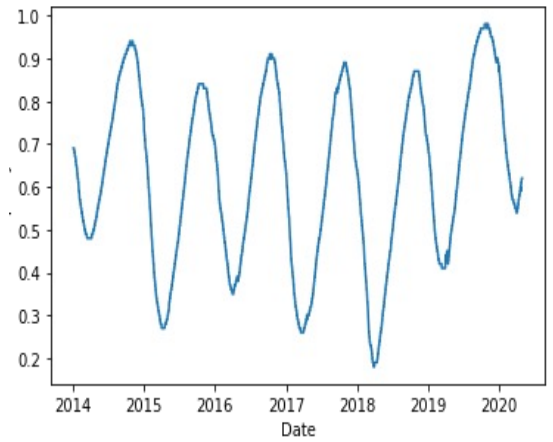


Complex machine learning models will exhibit superior performance in forecasting natural gas prices, as compared to traditional statistical models.

RNN+LSTM, as a sophisticated machine learning technique, will outperform all other forecasting models in terms of its accuracy in predicting natural gas prices.

4.1 Data – Time Series

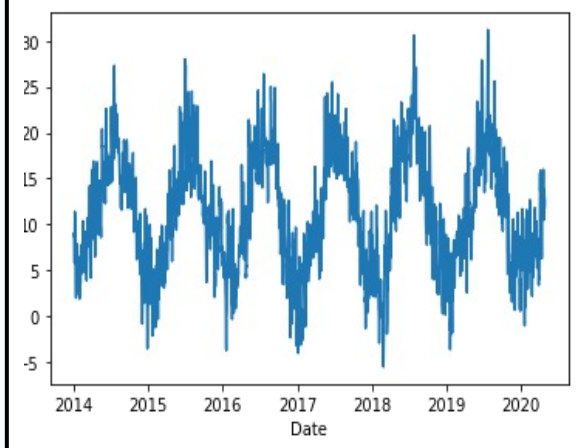
Capacity



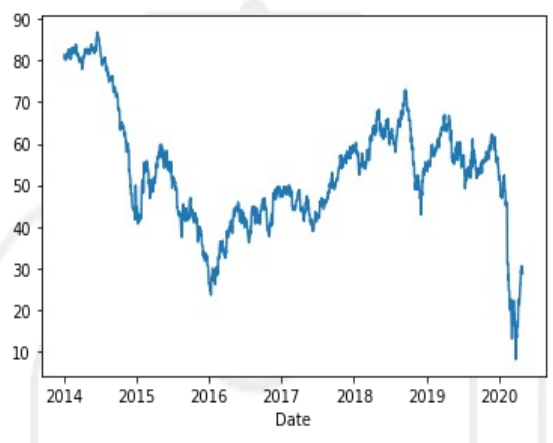
Time Series Matrix

	TTF Price	Temperature	Capacity	Coal Price	Brent Price
Date					
02-01-2014	26.30	8.8	0.69	60.88	81.38
03-01-2014	26.03	9.1	0.69	59.94	80.35
04-01-2014	26.75	8.1	0.69	59.94	80.35

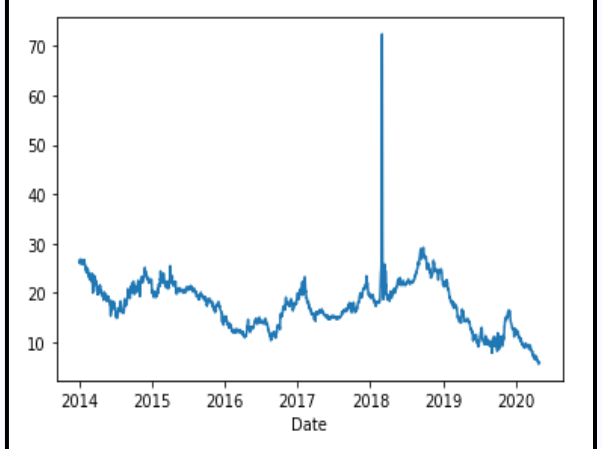
Temperature



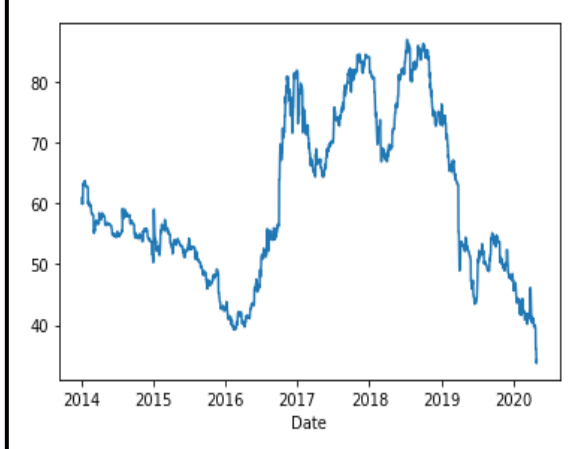
Brent Price



TTF Price



Coal Price



4.2 Data – Variable Description

Variable	Description	Source	Unit of Measure
TTF Price	TTF Spot Last Price	European Energy Excange (EEX)	EUR/MWh
Brent Price	Europe Brent Oil Spot Last Price FOB	Energy Information Administration (EIA)	EUR/Barrel
Coal Price	Coal (API2) CIF ARA Continuous Contract Last Price	MarketWatch	EUR/Tonne
Capacity	Storage Capacity of European Natural Gas	Gas Infrastructure Europe (GIE)	Percentage (%)
Temperature	Average Temperature between 13 West European Countries	European Climate Assessment & Dataset	°C

5.1 Methodology – Preparation

Tool

Python Programming Language & Jupyter Lab

Libraries

- Pandas
- NumPy
- SciPy
- Seaborn
- Marplotlib
- Statsmodels
- TensorFlow
- Sklearn
- Datetime
- Pmdarima

Models

- MLR
- SARIMAX
- RNN + LSTM
- SVR

Timeframe

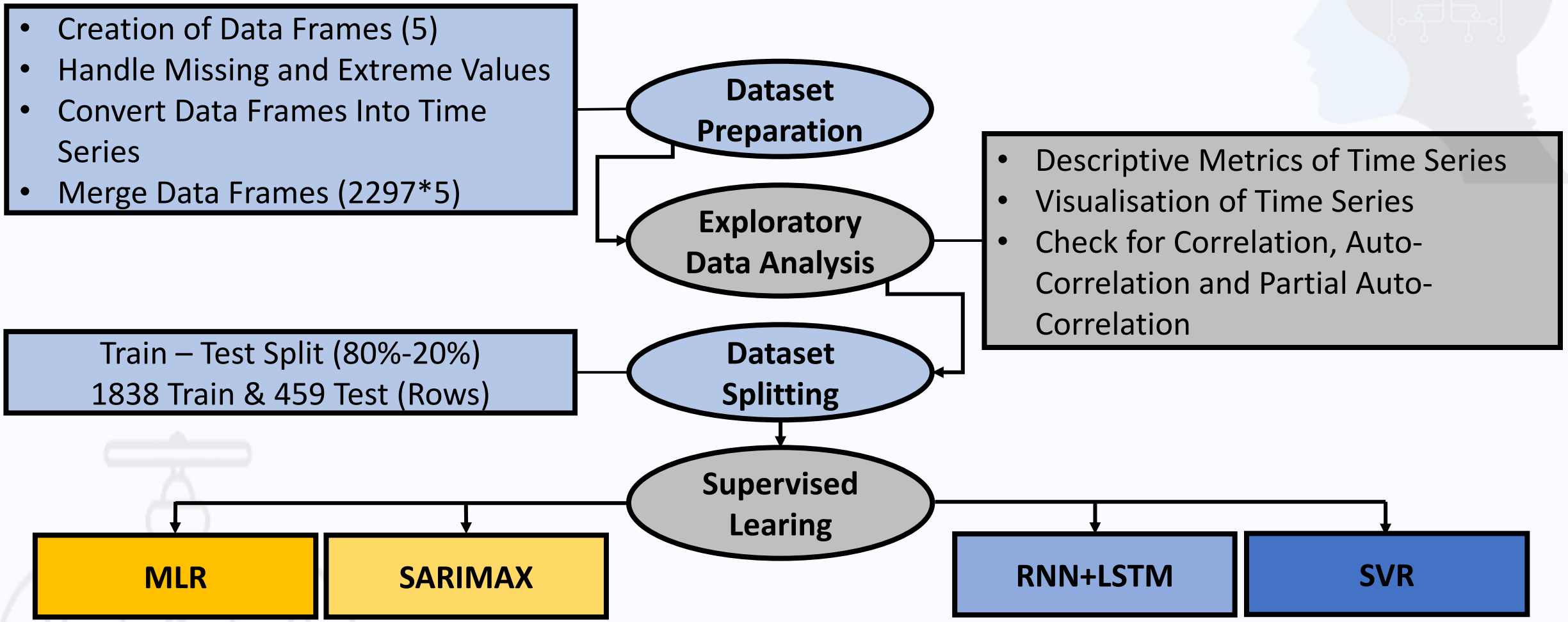
02/01/2014 – 29/04/2020

File Type

.csv

1. Review of Literature
2. Choice of Models and Variables
3. Request and Research for Data
4. Collection of Data
5. Choice of Timeframe of the Analysis
6. Preparation of Files for Analysis

5.2 Methodology – Process



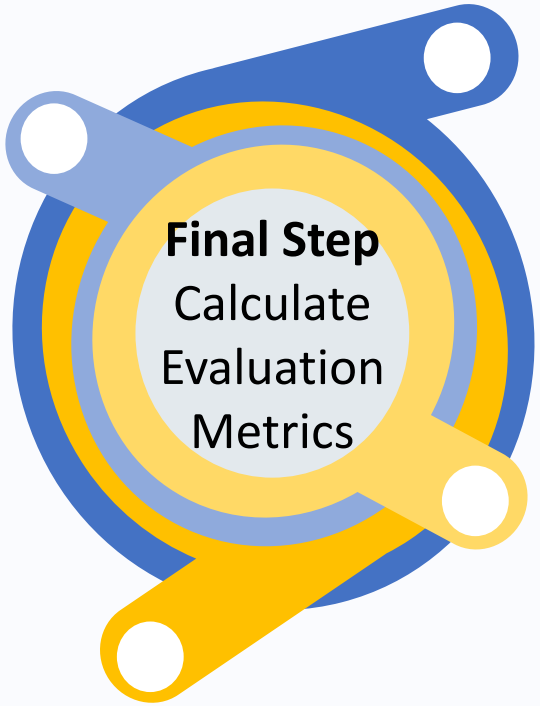
5.3 Methodology – Training & Testing Models

RNN+LSTM

1. Set Date Column as Index
2. Scale Datasets with MinMaxSCaler
3. Transform Time Series into Suitable Format for LSTM Model
4. Set Time Steps to Five
5. Create LSTM Model
6. Minimise Difference by Applying Adam Optimization Algorithm
7. Train Model for 200 Epochs on Training Set
8. Generate Predictions on Test Set
9. Repeat Process Nine Times & Calculate Mean of Evaluation Metrics

MLR

1. Set Training Dependent Variable & Regressors
2. Set Test Dependent Variable & Regressors
3. Performed Linear Regression on Training Set
4. Generate Predictions on Test Set



SVR

1. Set Date Column as Index
2. Standardize Regressors
3. Optimize Hyperparameters
4. Train the Model on Train Set
5. Generate Predictions on Test Set
6. Perform First Differencing on Data
7. Repeat Training Process
8. Generate Predictions on Test Set
9. Repeat Process Nine Times & Calculate Mean of Evaluation Metrics

SARIMAX

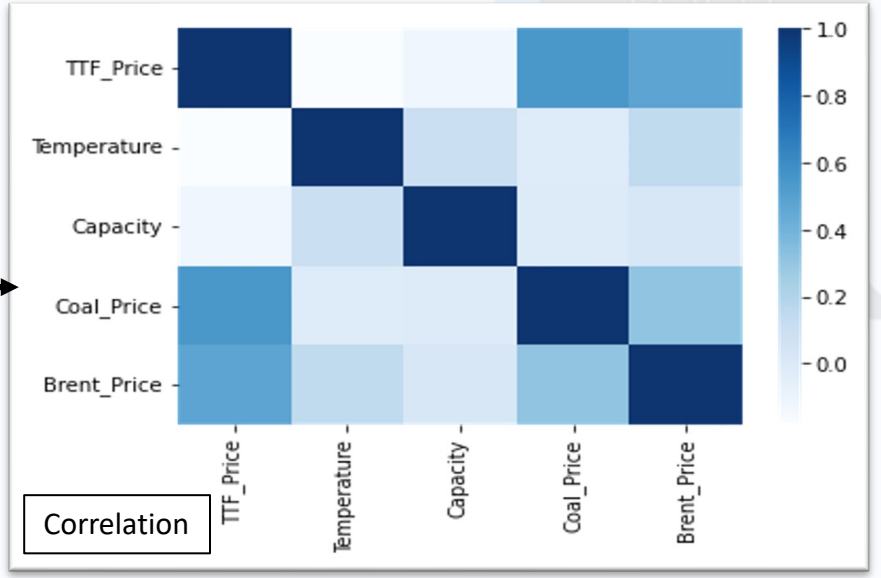
1. Test Dataset's Stationarity (ADF Test)
2. Set Date Column as Index
3. Create Model (2,1,1) & Apply Training Set
4. Set Model to Use Seasonal Patterns
5. Set Model to Use Stepwise Approach
6. Generate Results of Model
7. Perform Ljung-Box Test
8. Generate Predictions on Test Set

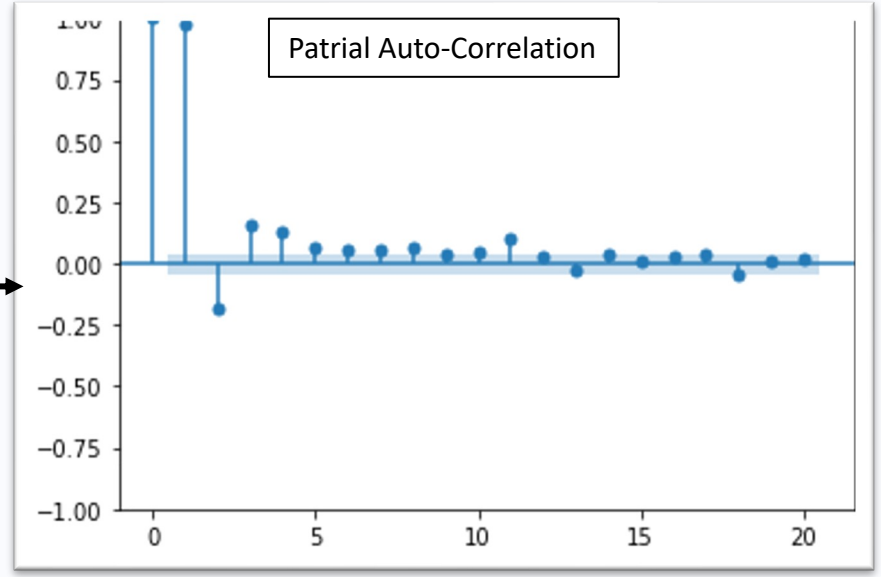
6.1 Results – Descriptive Statistics

Descriptive Statistics

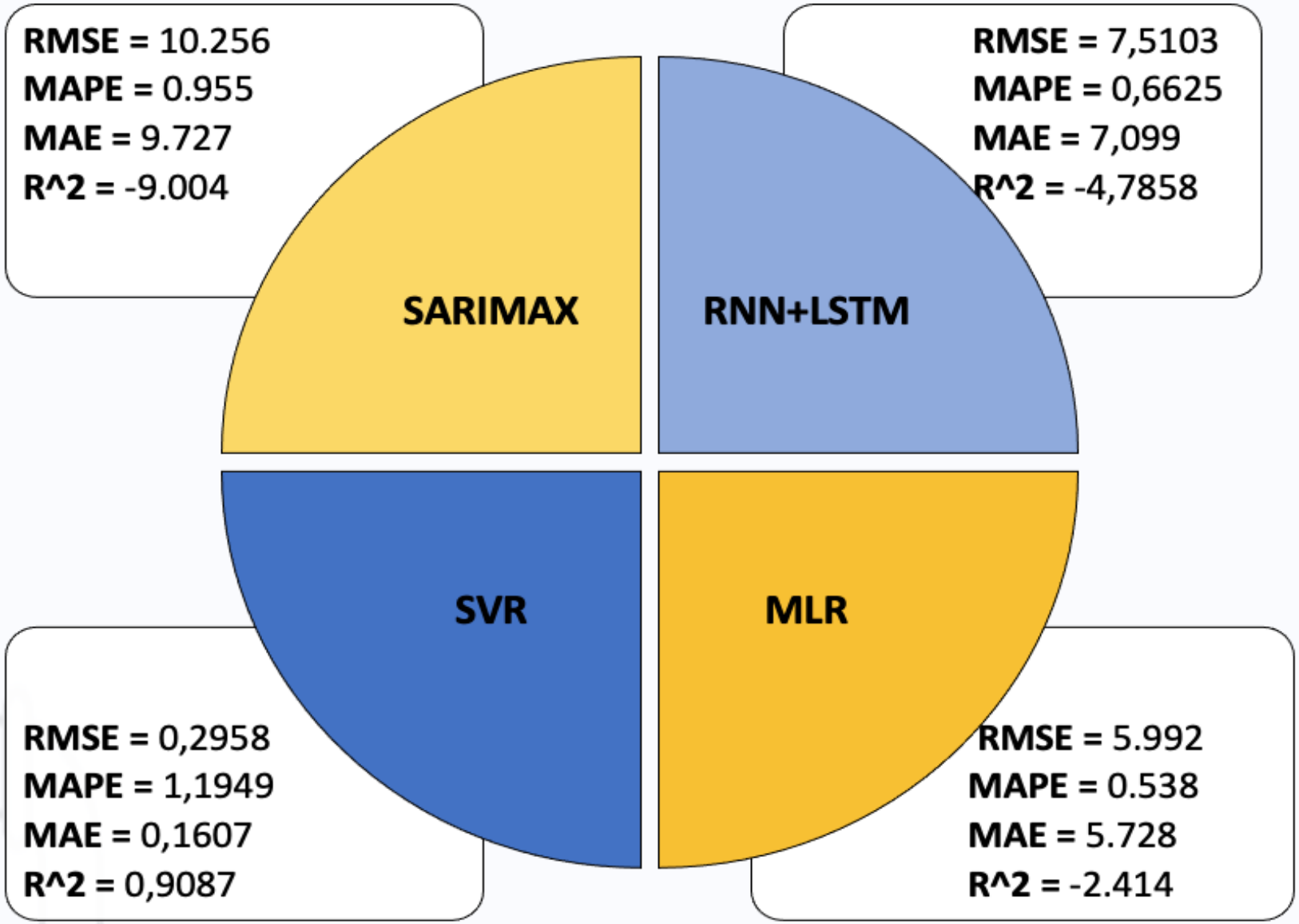
	TTF Price	Temperature	Capacity	Coal Price	Brent Price
Count	2297.00	2297.00	2297.00	2297.00	2297.00
Mean	17.59	11.20	0.62	60.52	53.22
Min	5.67	-5.50	0.18	33.72	8.00
25%	14.25	6.20	0.47	50.74	43.95
50%	17.82	11.00	0.63	55.95	52.88
75%	21.02	16.10	0.82	72.94	60.31
Max	72.37	31.20	0.98	86.97	86.85

	TTF Price	Temperature	Capacity	Coal Price	Brent Price
Std	4.92	6.33	0.20	13.55	14.21
Var	24.24	40.05	0.04	183.70	202.11
Skew	7.20	-0.60	-1.04	-1.10	0.22
Kurt	0.69	0.09	-0.18	-0.18	0.16





6.2 Results – Model Evaluation



7. Conclusions

Lowest MAE	**SVR**	Highest MAPE
Lowest RMSE		Highest R-Squared

01

High MAE	**MLR**	Lowest MAPE
High RMSE		Negative R-Squared

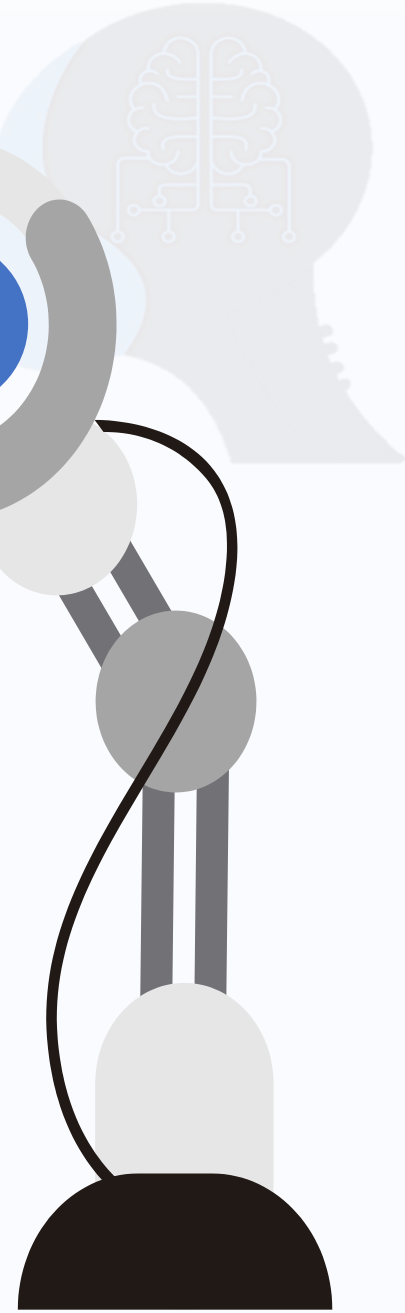
02

High MAE	**RNN+LSTM**	Low MAPE
High RMSE		Negative R-Squared

03

Highest MAE	**SARIMAX**	High MAPE
Highest RMSE		Negative R-Squared

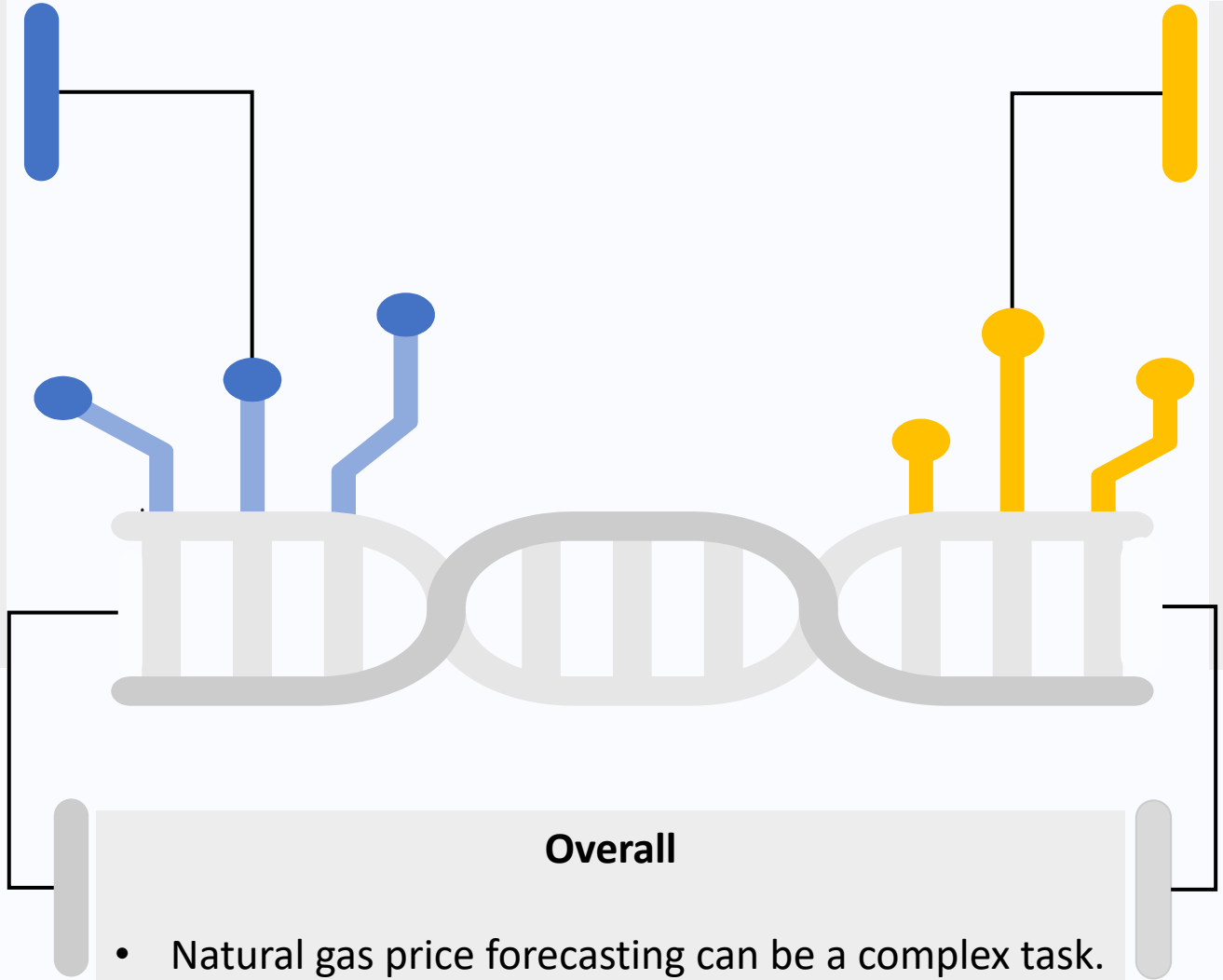
04



8. Discussion

1st Hypothesis

- Complex machine learning models did not necessarily outperform traditional statistical and time series models.
- Traditional statistical models such as MLR can also provide accurate forecasts of natural gas prices.



2nd Hypothesis

- RNN+LSTM did not outperform the rest of the models
- The RNN model can be sensitive to the size and quality of the dataset used for training

Overall

- Natural gas price forecasting can be a complex task.
- The appropriate forecasting model may depend on the specific characteristics of the data.

9. Limitations

- Uneventfull Timeframe 01
- Small Number of Models 02
- Small Number of Regressors 03
- Time Constraints 04
- Missing Values 05
- Group Testing of Variables 06

10. Suggestions

01 **Research with More Variables**

Geopolitical Events, Economic Indicators, and Technological Advancements

02 **Longer Timeframe of Available Data**

Should Include Periods of Recession (2020-2021)

03 **More Models**

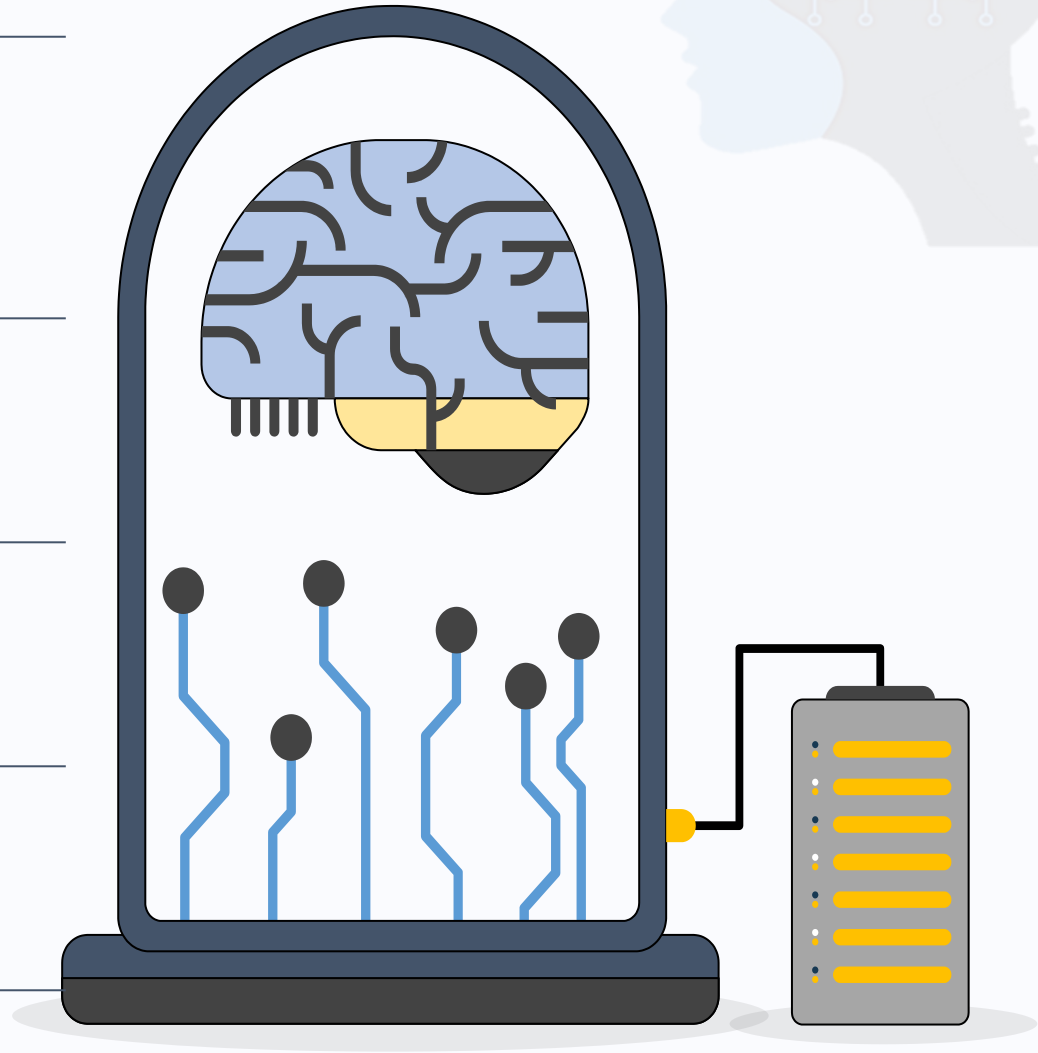
Such as Advanced Machine Learning and Ensemble Methods

04 **Testing Models with Each Variable Individually**

Examination of Each Variable’s Impact

05 **Examination of Other Gas Markets**

Data from the Asian and American Markets



References

- Aminu, N. (2019). Energy prices volatility and the United Kingdom: Evidence from a dynamic stochastic general equilibrium model. Energy, 172, 487-497.
- Azadeh, A., Sheikhalishahi, M., & Shahmiri, S. (2012). A hybrid neuro-fuzzy approach for improvement of natural gas price forecasting in vague and noisy environments: domestic and industrial sectors. In *Proceedings of the International Conference on Trends in Industrial and Mechanical Engineering* (ICTIME'2012). Dubai, United Arab Emirates 24-25. 2012.
- Brown, S. P. A., & Yucel, M. K. (2008). What drives natural gas prices?. The Energy Journal, 29, 45–60.
- Erdogdu, E. (2010). Natural gas demand in Turkey. *Applied Energy*, *87*(1), 211-219.
- Hosseinipoor, S., Hajirezaie, S., & Nejati, J. (2016). Application of ARIMA and GARCH Models in Forecasting the Natural Gas Prices. The *University of Oklahoma*.
- Hulshof, D., Van Der Maat, J. P., & Mulder, M. (2016). Market fundamentals, competition and natural-gas prices. Energy policy, 94, 480-491.
- Mu, X. (2007). Weather, storage, and natural gas price dynamics: Fundamentals and volatility. Energy Economics, 29(1), 46-63.
- Nick, S., & Thoenes, S. (2014). What drives natural gas prices?—A structural VAR approach. *Energy Economics*, *45*, 517-527.
- Li, H., Chen, L., Wang, D., & Zhang, H. (2017). Analysis of the price correlation between the international natural gas and coal. Energy Procedia, 142, 3141-3146.
- Lv, X., & Shan, X. (2013). Modeling natural gas market volatility using GARCH with different distributions. Physica A: Statistical Mechanics and its Applications, 392(22), 5685-5699.
- Salehnia, N., Falahi, M. A., Seifi, A., & Adeli, M. H. M. (2013). Forecasting natural gas spot prices with nonlinear modeling using Gamma test analysis. *Journal of Natural Gas Science and Engineering*, *14*, 238-249.
- Seo, S. H. (2021). Forecasting Korean LNG import price using ARIMAX, VECM, LSTM and hybrid models. *Ulsan National Institute of Science & Technology*.
- Su, M., Zhang, Z., Zhu, Y., Zha, D., & Wen, W. (2019). Data driven natural gas spot price prediction models using machine learning methods. *Energies*, 12(9), 1680.

THANKS!

Do you have any questions?