From Sentiment to Forecast: Predicting Natural Gas Prices Using News Article Analysis

DATA SCIENCE AND ENERGY TRANSITION

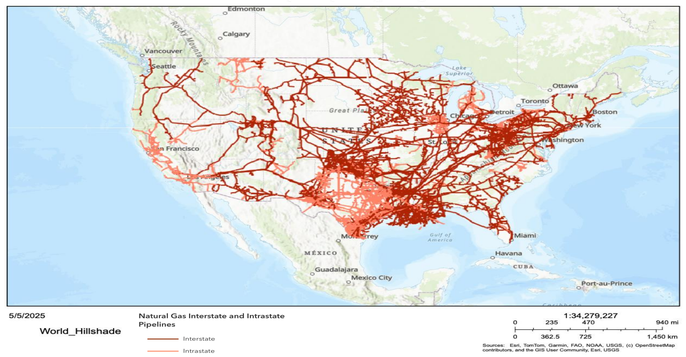
DR. HARDIK GOHEL
Associate Professor
& Director of Research

Divija Kalluri, Pooja Bejjanki, Junda Yin, Fatima Tanvir, Sadman Abdullah, Gael Cruz, Akanksha Samindla

DR. YUN WAN
Professor of Computer Information Systems

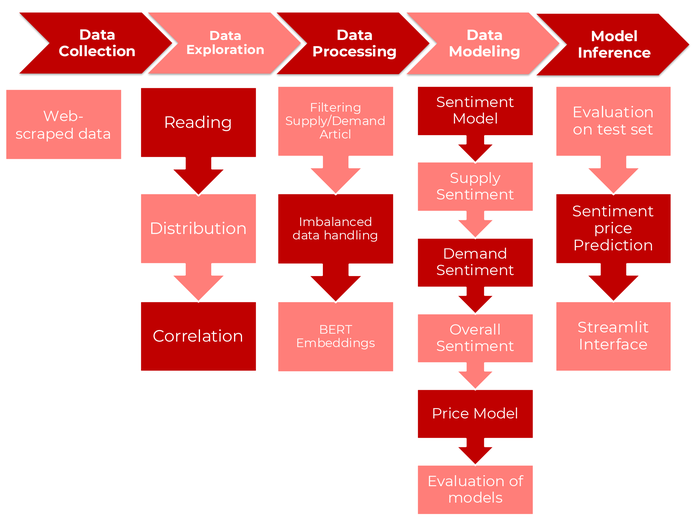
Introduction

The natural gas industry plays a vital role in the global energy market, with price fluctuations impacting both the economy and policy (McKinsey & Co., 2023). While traditional analysis relies on quantitative data, it often ignores the influence of media framing and public sentiment (Clarke et al., 2015).



This research develops a methodology that analyzes the public sentiment in natural gas news articles and predicts the corresponding market prices using machine learning and natural language processing techniques and evaluating the results.

Methodology



Data Acquisition

Natural gas news articles, natural gas prices and stock futures, storage, and consumption data were collected for analysis.

Data Distribution

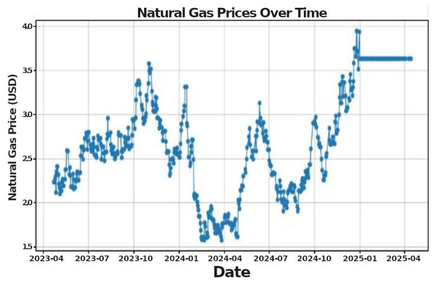


Figure1

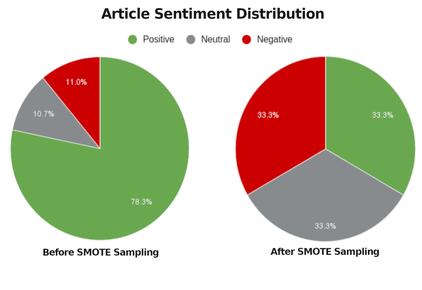


Figure2

Figure 1 reflects the Natural Gas price trends over the past 2 years. Figure 2 addresses the issue of data imbalance in the news article sentiment dataset by applying SMOTE to generate synthetic samples and balance the class distribution prior to model training.

Sentiment Model

Supply-Demand Heuristics (Model's Interpretation)

High Supply ↑ High Demand ↑	Storage ↑ Consumption ↑	Neutral sentiment - *Indicates market stability, leading to stable prices.*
Low Supply ↓ Low Demand ↓	Storage ↓ Consumption ↓	*Minimal market activity, resulting in stable prices and neutral sentiment.*
Low Supply ↓ High Demand ↑	Storage ↓ Consumption ↑	*Causes price increases, leading to positive sentiment.*
High Supply ↑ Low Demand ↓	Storage ↑ Consumption ↓	*Excess supply with low demand causes price declines, triggering negative sentiment.*

Table1. Sentiment Effect on Natural Gas Prices

Sentiment Model Performance Metric

Model	Precision	Recall	F1 Score	Overall Accuracy
Logistic Regression	**0.94**	**0.92**	**0.93**	**0.95**
Random Forest Clfr	**0.92**	**0.91**	**0.91**	**0.93**
Deep Learning (BERT)	**0.82**	**0.84**	**0.83**	**0.88**

Table2. Sentiment Model Performance Metrics

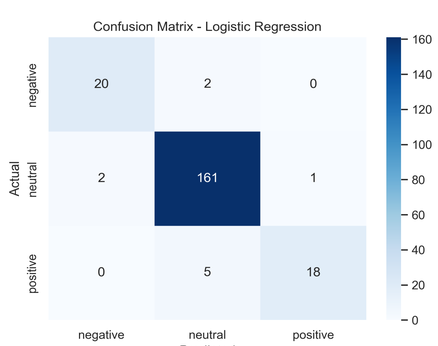


Figure3

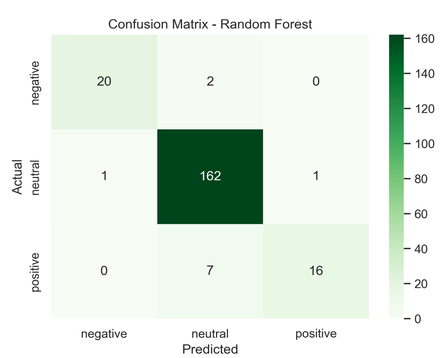


Figure4

Price Prediction Model

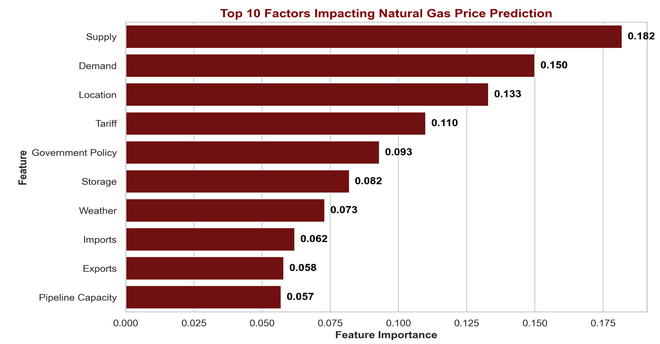
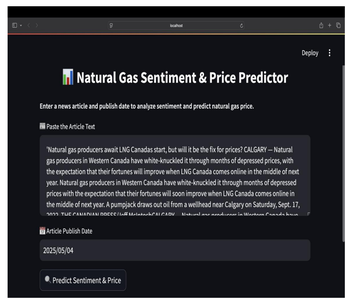
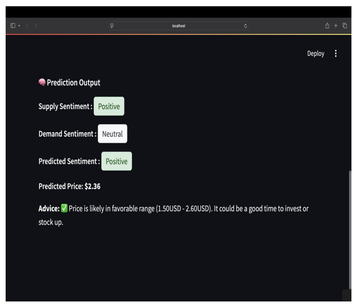


Figure5. Key Drivers Identified Through Feature Importance Scores for Natural Gas Price Prediction Using a Random Forest Regressor

Streamlit Inference





A Streamlit-based interface was developed to process natural gas-related news articles with predicted output corresponding to sentiment scores and predicted prices.

Inferences & Future Scope

This project demonstrates that news sentiment provides valuable insights into market trends through NLP and price prediction models.

Future work could extend this approach to other energy sectors, such as oil and solar, and incorporate factors like geopolitical events and weather patterns. Advanced models could further enhance prediction accuracy and scalability.

References

McKinsey & Company. (2023). Global energy perspective 2023: Natural gas outlook. https://www.mckinsey.com/industries/oil-and-gas/our-insights/global-energy-perspective-2023-natural-gas-outlook

Clarke, C. E., Hart, P. S., Schuldt, J. P., Evensen, D. T. N., Boudet, H. S., Jacquet, J. B., & Stedman, R. C. (2015). Public opinion on energy development: The interplay of issue framing, top-of-mind associations, and political ideology. *Energy Policy, 81*, 131-140. https://doi.org/10.1016/j.enpol.2015.02.019