Deep Learning Approaches to Commodity Price Forecasting

Author: Sulaiman Mahmood
Degree: MSc Data Science

Overview

This work formed my Master's dissertation and investigated the use of deep learning models for forecasting commodity prices, with a particular focus on whether more compact recurrent architectures could improve performance and training stability compared to standard approaches.

While Long Short-Term Memory (LSTM) networks are widely used in time-series forecasting, this project deliberately extended beyond LSTM by implementing and evaluating Gated Recurrent Unit (GRU) networks, which offer a simpler gating mechanism and fewer parameters.

Research Motivation

Commodity price time series are highly volatile and influenced by external shocks such as geopolitical events and supply chain disruptions. These characteristics introduce non-linear patterns and long-term dependencies that are difficult to model using classical statistical methods.

LSTM networks are commonly applied to such problems; however, they can be computationally heavy and sensitive to training instability when data is limited or noisy. This research explored whether GRU models could provide comparable or improved performance while reducing model complexity, effectively taking the modelling approach one step further.

Methodology

The modelling pipeline included:

1. Data cleaning, normalization, and exploratory analysis
2. Implementation of baseline models:
    - Linear regression
    - ARIMA
3. Deep learning models:
    - LSTM as a benchmark recurrent architecture
    - **GRU as an extension beyond LSTM**, with fewer gates and reduced parameter count
4. Evaluation using RMSE and MSE to enable direct comparison across models

All models were implemented in MATLAB using the Deep Learning Toolbox.

Results and Interpretation

Model	RMSE
Linear Regression	0.0569
ARIMA	0.0631
LSTM	0.0493
GRU	**0.0078**

The GRU model consistently achieved the lowest error. In addition to improved accuracy, GRU showed more stable convergence behaviour during training, supporting the hypothesis that a simpler recurrent structure can be advantageous for this class of time-series data.

Key Takeaway

A key outcome of this research was demonstrating that **moving beyond standard LSTM architectures to GRU-based models can lead to both improved predictive performance and more reliable training**, particularly in volatile financial time series with limited historical data.

This finding highlights the importance of architectural choice rather than assuming more complex models are always preferable.

Contribution

This dissertation demonstrates:

- independent model selection and evaluation
- critical comparison of deep learning architectures
- applied reasoning beyond textbook implementations
- research maturity in time-series forecasting

Availability

Code and supporting documentation are publicly available on GitHub as part of an academic and professional portfolio.