ChatGPT

Intraday Volatility Forecasting Using Hybrid GARCH-LSTM-VIX-EWMA Models: A Modular and Visual Framework

Abstract

This study presents a comprehensive and modular approach to forecasting intraday volatility of financial assets, using the SPY ETF 15-minute interval data. We develop and compare several statistical and deep learning models including GARCH, LSTM, hybrid GARCH-LSTM, GARCH-LSTM-VIX, and GARCH-LSTM-EWMA-VIX architectures. Each model is evaluated on the basis of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The study incorporates additional features like VIX and EWMA, and is structured around a production-ready, interactive Streamlit application which facilitates model comparison and visualization.

1. Introduction

Intraday volatility modeling is essential for financial risk management, high-frequency trading, and portfolio optimization. Traditional econometric models such as GARCH capture historical volatility well, while neural networks like LSTM are suited for sequence modeling and can extract nonlinear dependencies. This paper proposes a hybrid approach combining both statistical and machine learning techniques to capture intraday volatility more effectively.

2. Data Collection and Preprocessing

2.1 Datasets Used

- **SPY 15-minute intraday data** (Open, High, Low, Close, Volume)
- **VIX 15-minute intraday data**

2.2 Data Source

- Alpha Vantage API

2.3 Preprocessing Steps

- Time alignment of SPY and VIX datasets
- Calculation of log returns
- EWMA volatility estimation
- Lagged volatility features
- Train-test split and scaling for LSTM

3. Model Architectures

3.1 GARCH

Implemented using `arch` Python library to produce baseline volatility estimates.

3.2 LSTM

Neural network models created using TensorFlow/Keras with tuning of hyperparameters:

- Layers: 1, 2, or 3
- Activation functions: Tanh, ReLU
- Lookback windows: 5, 22, 66 intervals
- Loss functions: MAE, MSE

3.3 Hybrid GARCH-LSTM

Combines lagged volatility predictions from GARCH as additional input features for LSTM.

3.4 GARCH-LSTM-VIX

Adds VIX values (close) to LSTM input in combination with GARCH outputs.

3.5 GARCH-LSTM-EWMA-VIX

Further integrates EWMA volatility and rolling standard deviations for enhanced input diversity.

4. Evaluation Metrics

Each model was evaluated using:

- **Mean Absolute Error (MAE)**
- **Root Mean Squared Error (RMSE)**

5. Implementation Details

5.1 Modular Directory Structure

```
volatility_prediction_project/
|-- data/
|-- garch/
|-- lstm/
|-- lstm_garch/
|-- lstm_garch_vix/
|-- lstm_garch_ewma_vix/
|-- sensitivity/
|-- models_comparison/
|-- plots/
|-- app.py (Streamlit dashboard)
```

5.2 Key Scripts

- `garch_functions.py` : Training and prediction for GARCH
- `LSTM.py` : Base LSTM model creation
- `lstm_garch_intraday.py` : Hybrid model generation
- `ewma_functions.py` : EWMA volatility calculation
- `models_comparison.py` : Evaluation and plotting
- `app.py` : Interactive dashboard for model selection and plot viewing

6. Visualizations and Interpretation

Using interactive Plotly charts, users can view:

- Predictions vs Actuals for each model
- Sensitivity to lookback windows and activation functions
- EWMA vs Rolling Standard Deviation
- VIX vs EWMA overlays

7. Interactive Streamlit Dashboard

Features:

- Sidebar dropdowns for model selection
- Tabs for each model category (GARCH, LSTM, Hybrid, EWMA)
- Dynamic plot generation using Plotly
- Real-time display of MAE and RMSE

Deployment:

- Supports local and cloud deployment (Streamlit Cloud)
- Uses pre-trained model results stored in `data/`

8. Experimental Results

8.1 Key Metrics (Sample Output)

```
DataFrame: garch | MAE: 0.007087 | RMSE: 0.007192
DataFrame: lstm | MAE: 0.000072 | RMSE: 0.000091
DataFrame: lstm_garch | MAE: 0.000362 | RMSE: 0.000623
DataFrame: lstm_garch_vix | MAE: 0.000316 | RMSE: 0.000512
```

```
DataFrame: lstm_garch_vix_1_layer | MAE: 0.000254 | RMSE: 0.000489
DataFrame: ewma | MAE: 0.000257 | RMSE: 0.000412
```

8.2 Insights

- **LSTM-GARCH-VIX** outperforms plain LSTM in both metrics.
- **EWMA** closely tracks volatility but underperforms compared to deep learning.
- **Model tuning** (layers, lookbacks, loss) affects performance significantly.

9. Future Work

- Integrate real-time data ingestion
- Add models like HARCH, HAR, or transformers
- Improve interpretability using SHAP/attention mechanisms
- Compare performance across other ETFs and asset classes
- Extend to multi-asset volatility forecasting

10. Conclusion

This research builds a robust and extensible pipeline for intraday volatility forecasting. By blending GARCH's statistical strength with LSTM's sequence modeling capability and augmenting with market volatility indices (VIX) and technical indicators (EWMA), our hybrid models significantly improve volatility prediction performance. The accompanying dashboard offers an intuitive and flexible interface for real-time comparison and visualization.

Appendix

A. Tools and Libraries

- Python 3.11
- TensorFlow, Keras, pandas, numpy, scikit-learn, arch
- Streamlit, Plotly, Matplotlib

B. Repository

GitHub: https://github.com/himelku/volatility-prediction-project

---

Acknowledgements

Special thanks to the MQIM program at the University of New Brunswick, Alpha Vantage for API access, and Streamlit for enabling rapid development of data apps.