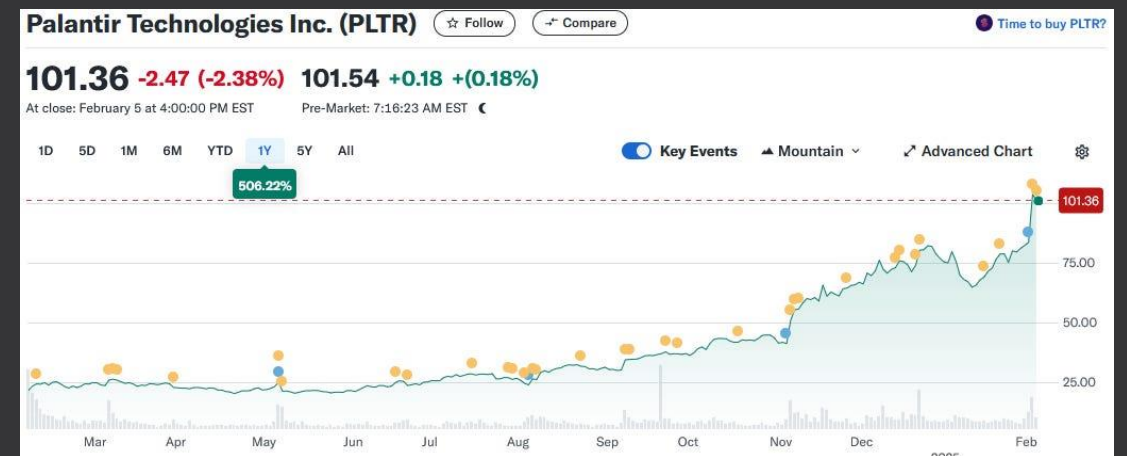# Forecasting 5-day Returns of Palantir (PLTR) Stocks

Emma Wunderly (leader), Andre Galistinos, and Ashley Nguyen

Project Overview

Context: PLTR has had ~1700% growth since going public in September 2020 currently sitting at ~$180 per share with a market cap of ~$430 billion

Motivation: Their stock has fluctuated heavily within the past 5 years, often reflecting developments in artificial intelligence and government contracts. These relationships raise ethical concerns about privacy, surveillance, and the implications of AI-driven decision-making, adding an additional layer of curiosity to our project.

Research Question: Can we create a model that accurately predicts Palantir's stock prices given past daily stock price and number of AI and war-related keywords were mentioned in the US news that day?

Hypothesis: An increase in the number of AI and war-related keywords in the US news will correlate to an increase in Palantir (PLTR) stock prices.

Modeling Approach: Hybrid model b/n an Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) model
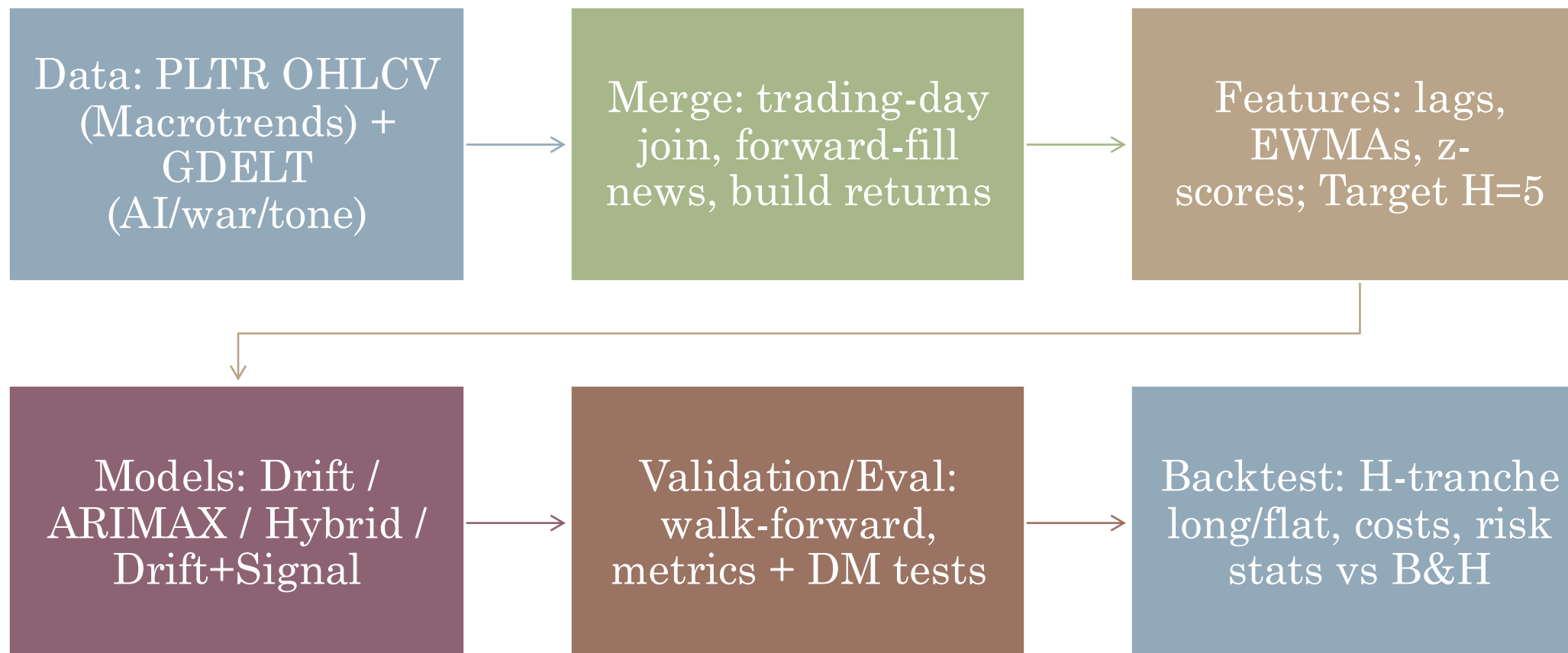
Goal: Forecast 5-day returns of PLTR stock

Data Acquisition

- Macrotrends → PLTR prices

- GDELT Project BigQuery → daily counts of AI- and war-themed words in world news

- Merge two datasets → pltr_with_gkg_daily-2.csv

Field	Type	Description	Notes / Uncertainty	Example
date	datetime	Trading day (NY market calendar).	Non-trading days omitted after merge.	2024-11-15
open, high, low, close	float	Daily OHLC prices (USD).	Likely unadjusted close (splits/dividends not applied).	open=21.34, high=22.10, low=20.95, close=21.87
volume	int/float	Daily shares traded.	Units = shares.	38,245,100
return	float	Simple daily return of close.	Undefined on first row.	0.012 (±1.2%)
log_return	float	log(close).diff().	Good for additivity/stationarity.	0.0119
cnt_ai	int	Daily count of GKG rows with AI/ML/LLM themes.	Keyword coverage may evolve; regex documented in query.	57
cnt_war	int	Daily count of GKG rows with war/conflict/military themes.	Much higher baseline than AI globally (expected).	512,304
avg_tone	float	Average overall V2Tone for that day.	Centered near 0; negative values common.	-1.62

Analysis Plan and Justification

Data: PLTR OHLCV (Macrotrends) + GDELT (AI/war/tone) → Merge: trading-day join, forward-fill news, build returns → Features: lags, EWMAs, z-scores; Target H=5

Models: Drift / ARIMAX / Hybrid / Drift+Signal → Validation/Eval: walk-forward, metrics + DM tests → Backtest: H-tranche long/flat, costs, risk stats vs B&H

Tricky Analysis Decision

- Our biggest decision was to go beyond the purely technical Palantir data set by merging Palantir stock data with three GDELT news features: "AI" theme count, "War" theme count and average tone.

- While predicting the Palantir stock prices was an interesting use of time series data, we wanted to bring more real-world implications to our predictions.

- By aiming for real world implications, we carried two major risks:
  - Increased model complexity: adding exogenous variables like "cnt_ai," "cnt_war," and "avg_tone," we were forced into the ARIMAX framework and making the hybrid ARIMAX-LSTM model, a significantly more complex and more prone to overfitting model compared to ARIMA alone.
  - Signal-to-noise ratio: GDELT data is noisy and heterogeneous, so we had to make a subjective judgment call on which keywords were relevant to PLTR (such as AI and WAR) and how to handle non-trading days.

- The impact of this decision was a tradeoff. Essentially, we sacrificed directional prediction power for a superior risk-managed trading strategy, providing the value of the news data as a volatility indicator rather than a pure price indicator.

Bias and Uncertainty

Data
- Keyword misclassification (AI out of context)
- Queried specified that related to AI- and war- but there are more words out there
- Timing mismatch – PLTR stock prices only exist on trading days (Mon-Fri) while news data is daily and continuous

Model
- AIRMAX assumes news is exogenous (external variables influence the stock price, but not vice versa) – not true b/c news and prices can move together where rising stock prices might to lead to more AI-related coverage

Results/Conclusion

- **Hypothesis:** GDELT AI/war/tone signals improve PLTR return forecasts beyond price drift.

- **Forecasting:** Drift (Train-Mean) DirAcc ≈ 0.59; ARIMAX/Hybrid ≈ 0.55. DM tests: no significant improvement → news adds only modest signal.

- **Backtest:** Debiased long/flat (H, 5 bps): Sharpe 3.3 > 2.0, vol 26.6% < 68.5%, MaxDD −24% < −37%, exposure ~89%

- **Conclusion:** Drift dominates; news adds small, non-significant edge. Long/flat reduces risk but can lag B&H in strong bull runs.

- **Implication:** To beat drift, add richer signals (event intensities/sentiment by source/region), regime detection, and broader assets/horizons.

```
=== Aggregate metrics (Drift+Signal) — mean across folds ===
                          RMSE      MAE       R2       DirAcc   BigMoveHit
model
Train-Mean (fold drift)  0.102695  0.073687  -0.060233  0.592239  0.622222
Zero                     0.103845  0.074909  -0.086618  0.002827  0.000000
Hybrid+Drift             0.120182  0.088839  -0.495452  0.562361  0.544444
ARIMAX+Drift             0.120225  0.088853  -0.496244  0.558116  0.538889
Naive(H)                 0.146867  0.111156  -1.203146  0.466077  0.411111
```

```
Debiased backtest (H-tranche long/flat, cost 5 bps per switch):
  Annualized mean − 83.296%, vol − 26.316%, Sharpe − 3.17
  Final wealth: 9.35×
```


Debiased Long/Flat — Cumulative Wealth vs Buy&Hold (H=5)
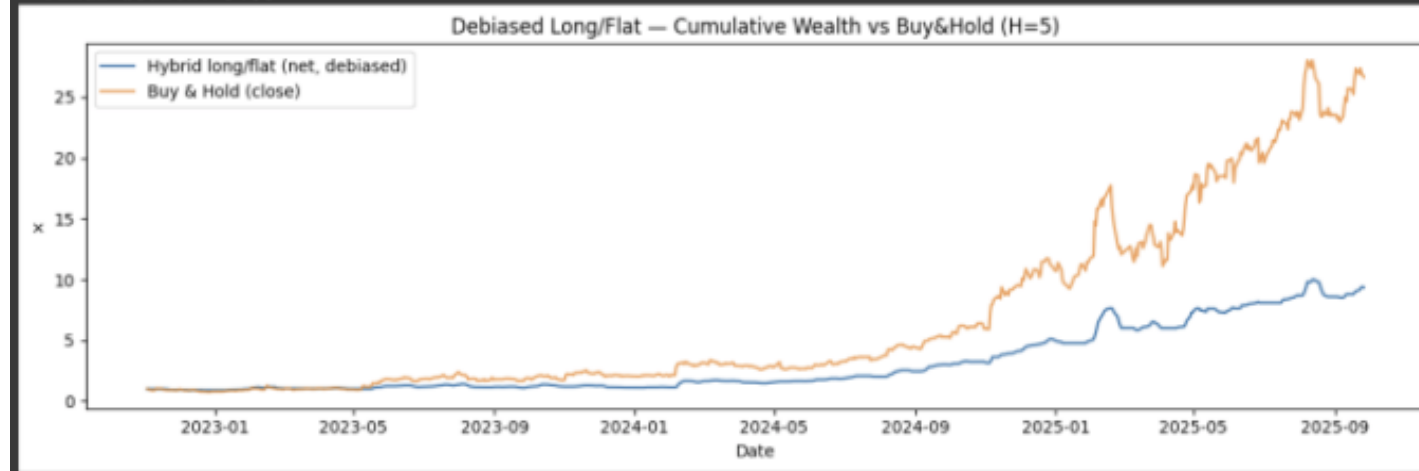
Next Steps

- Our new lines of exploration were the discovery of the model's performance in the back test. Despite finding poor directional accuracy (approx. 0.55), the long/flat strategy achieved a Sharpe ratio of 3.3 and greatly reduced volatility and Maximum Drawdown compared to Buy-and-Hold. Therefore, the hybrid model is not a strong directional predictor but an excellent risk/volatility filter.

- Improvements that could have been made are a switch to a dedicated sentiment score, as the news features, we used to be simple counts and lacked the nuance of actual sentiment. To improve, we could have used a large financial language model or a specialized library to analyze the headlines/snippets related to PLTR to yield a specific, and more continuous sentiment.

- New questions that we ask:
  - Does the effort required to gather, clean, and integrate the complex news data justify the statistically non-significant improvement in directional accuracy?
  - What is the dollar value threshold for a feature set to be deemed "useful" in real-world trading?
  - Would the ARIMAX/Hybrid approach perform better on a different, non-volatile asset or an asset where news sentiment is known to be a stronger driver (ex., Biotech stocks)?
  - Why is the simple Train-Mean Drift model so effective (DirAcc was approx. 0.59)? Is this due to statistical properties of PLTR's return series (mild positive skew) or simply a result of the specific 2020–2025 time window?

References

[1] Macrotrends LLC, "Palantir Technologies (PLTR) Historical Stock Prices," Macrotrends. [Online]. Available: https://www.macrotrends.net/stocks/charts/PLTR/palantir-technologies/stock-price-history. Accessed: Oct. 13, 2025.

[2] The GDELT Project, "GDELT 2.0 Global Knowledge Graph (GKG)," Google BigQuery Public Datasets (dataset: gdelt-bq.gdeltv2.gkg_partitioned). [Online]. Available: https://www.gdeltproject.org/data.html and https://cloud.google.com/bigquery/public-data/gdelt. Accessed: Oct. 13, 2025.

[3] A. B. Author, "A study on the application of ARIMA and MLA-LSTM hybrid model in time series forecasting," Conf. Proc. SPIE, vol. 13576, no. 1C, 2025.

[4] Fortune, "The AI bubble will pop. Intelligence won't.," Fortune. [Online]. Available: https://fortune.com/2025/10/06/the-ai-bubble-will-pop-intelligence-wont/.

[5] Y. Li, S. Liu, C. Wang, and T. Shi, "Stock trend forecasting using a hybrid model with news sentiment analysis and ARIMAX-LSTM," in Proc. Int. Conf. on Neural Information Processing (ICONIP), 2021.

[6] G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ: Wiley, 2015.

[7] S. H. Alqaisi, S. S. Shraida, and A. A. Aljawarneh, "A novel time series prediction model using deep learning and statistical approach: ARIMA-LSTM," in Proc. Int. Conf. on Computer Science and Information Technology (ICCSIT), 2021, pp. 1-6.

PLTR vs GDELT News Counts (AI & WAR)