Final Year Project Proposal

Echo: A Multimodal Spatio-Temporal Framework for Forecasting Dyadic Conflict and Cooperation

Sharaf-Eddine Boukhezer
seb206@student.bham.ac.uk

Supervisor: Dr. Leandro Minku

University of Birmingham
MSci Computer Science

1 Problem Statement

Forecasting geopolitical dynamics remains a challenging task due to cooperation and conflict between polities being affected by multiple factors, including historical patterns, structural relations, and sudden events and shocks. While media narratives provide useful signals, current methods are mostly unimodal, relying on either pure time series, text-based sentiment models, or static risk indices, which restricts their ability to adapt across different time horizons and evolving geopolitical contexts.

2 Related Work

Prior work in the field spans different methodological strategies:

Paper	**Category**	**Model Input**	**Model Output**	**Target Level**	**Horizon**
Chen et al. (2020)	Time-Series Dyad Model	GDELT dyad event counts (4 QuadClass types) + top-k related dyads	Next week's material conflict count	Single dyad	1-step (weekly)
von der Maase (2025) – HydraNet	Global Spatio-Temporal Forecasting	Past fatalities (priogrid-month) as spatiotemporal tensor	Conflict intensity 1–36 months ahead	Global grid (not dyads)	1–36m
Croicu & von der Maase (2025)	Text-Based Escalation Prediction	News text embeddings (Factiva) + actor metadata	Escalation/de-escalation classification	Dyads/actors	Next period
Zakotianskyi (2025)	Feature-Rich Statistical / ML Models	100+ political/economic features from ViEWS & UCDP	Conflict onset probability (1, 3, 6m)	All dyads/country pairs	Fixed horizons
Liu & Shen (2025)	Graph-Based ML / Cyber Relations	Graph of cyber relations + node features + threat-report text	Binary cyberattack prediction	All dyads (graph-wide)	Next period

Table 1: Overview of related forecasting approaches.

Echo integrates elements from all of these, combining them into a single multimodal dyadic forecasting system.

3 Aim

Echo aims to build, evaluate, and analyse a multimodal forecasting system that predicts the future state of dyadic relations using four complementary signals:

1. Historical event patterns and trends.
2. News-derived signals.
3. Structural geopolitical relationships.
4. Country-level attributes (e.g., instability).

The system will produce unified forward-looking forecasts of both event intensity (QuadClass counts) and relational polarity (Goldstein score) for any country pair at time t + n, where n denotes the forecasting horizon.

- **Goldstein Scale (GS)**: A continuous measure of event polarity ranging from -10 (strong conflict) to $+10$ (strong cooperation), indicating the overall intensity and direction of interactions between two actors.
- **QuadClass Counts ($\mathbf{QC}_{1-4}$)**: Discrete event frequency counts across four categories derived from the CAMEO event ontology: *(1) Verbal Cooperation, (2) Material Cooperation, (3) Verbal Conflict, (4) Material Conflict.* These counts summarise the distribution of cooperative and conflictual actions for each dyad-week.

4 Scope & Boundaries

4.1 In scope

- Forecasting dyadic relations between countries in a global network.
- Single-horizon prediction setup ($t \to t + n$).
- Two targets: Goldstein score and QuadClass event distribution.
- Multimodal architecture combining time series, graph structure, text signals, and static features.
- Baseline comparison with shared data splits.
- Evaluation across increasing forecast horizons.
- Ablation study.

4.2 Out of scope

- Event-level prediction (Echo operates on aggregated time steps, not individual events).
- Forecasting specific event types or actor-level interactions beyond country dyads (focus limited to selected polities).

5 Methodology

5.1 Data & Preprocessing

Data Sources and Inputs

- **Time-Series History (TS)**: dyad-level event aggregates (Goldstein, QuadClass counts).
  Source: GDELT 1.0 Global Events Database (V2.0 if available).
- **News Signal (TX)**: BERT embeddings pooled per dyad-week from event descriptions/news text.
  Source: GDELT text fields + external news if needed.

- **Graph Structure (GR)**: dynamic country graph based on alliances, trade, co-event frequency, and borders.
  Sources: UCDP, ViEWS, COW, CEPII.

- **Static Dyad Features (ST)**: location, regime, instability, neighbour flags, etc.
  Sources: CEPII, WGI, COW, World Bank.

Train/Validation/Test Split

All splits are chronological to prevent temporal leakage:

$$\text{Train: } 2007-2010, \quad \text{Validation: } 2011, \quad \text{Test: } 2012-2013$$

Each forecast horizon $t + n$ is trained and evaluated separately (e.g., $n = 1, 4, 8, 12$).

5.2 Baseline Models

1. Uniform Random
2. Last Observation Carried Forward (LOCF)
3. Rolling Mean
4. Linear Regression
5. Vector Autoregression (VAR)
6. Random Forest
7. XGBoost
8. Vanilla LSTM
9. Temporal CNN (TCN)
10. BERT-only MLP (text-only)
11. Graph-only GNN
12. Different models may be proposed if relevant

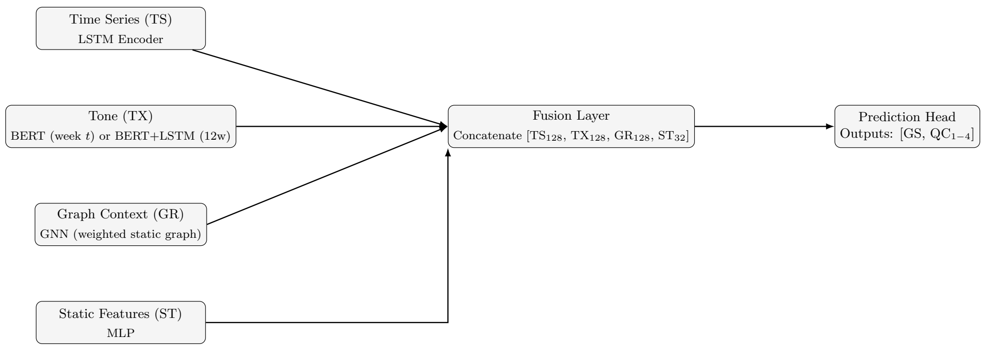
We also use the reported 70% accuracy from Chen *et al.* as a reference for comparison.

5.3 Proposed Model

Model Inputs (per dyad, per week)

- **Time-Series Window (TS)**: past k weeks of Goldstein and QuadClass counts, tone, and rolling statistics.
- **Textual Signal (TX)**: pooled BERT embedding for the current dyad-week (BERT frozen).
- **Graph Context (GR)**: node embeddings from a GNN (GraphSAGE/GAT), combined into dyad vectors.
- **Static Features (ST)**: distance, GDP ratio, regime difference, instability, etc.

Model Architecture



- Model predicts all 90 dyads in the global network jointly.
- Start with TS + TX only, then add GR (graph) and ST (static).
- Tone: compare snapshot (week t) vs. historical (12w) encoding.
- Graph: static weighted edges, later test temporal graph G_t.
- Test multi-head prediction for Goldstein (scalar) & QuadClass (4-way).

Note: different architectures will be tested during the ablation study in order to determine the best-performing one.

5.4 Output Definition

For each dyad (A, B) in the global country network at time t, the model produces a single-horizon forecast for time $t + n$ (in weeks).

The prediction for a dyad is a 5-dimensional output vector:

$$\hat{\mathbf{y}}_{A,B}(t+n) = \begin{bmatrix} \text{Goldstein}_{t+n} \\ \text{Quad}_1 \text{ (verbal cooperation)} \\ \text{Quad}_2 \text{ (material cooperation)} \\ \text{Quad}_3 \text{ (verbal conflict)} \\ \text{Quad}_4 \text{ (material conflict)} \end{bmatrix}$$

The model generates this vector *simultaneously for all dyads*, resulting in a prediction matrix of size:

$$\#\text{dyads} \times 5$$

Note: The analysis is limited to the 10 most event-active countries globally, selected by overall event frequency. This results in a dyadic set of 90 country pairs.

5.5 Evaluation Protocol

The model will be evaluated along four dimensions:

- Across models (Echo vs. all baselines)
- Across time horizons (short-, medium-, and long-term; objective up to t+36 months)
- Across targets (Goldstein polarity vs. QuadClass event distribution)
- Across dyad types (e.g. allies, rivals, neutral pairs, high- vs. low-activity dyads)

Evaluation Protocol

Up to three evaluation metrics will be selected to assess model performance from different aspects. A metric study will be conducted beforehand to determine the most appropriate choices.

- Goldstein → MAE, MSE, RMSE, R^2
- QuadClass → MSE, RMSE, Negative Binomial deviance, Distributional Calibration Error (proper scoring rule, distribution-level interpretation)

6 Work Plan

1. **Build full data pipeline:** Build ETL, clean data, generate weekly dyad table.
2. **Train baseline models:** Train all baseline models for t+1 horizon and for smaller data.
3. **Implement Echo v0 (minimal version).**
4. **Compare** Echo v0 vs. baselines on the same split and horizon.
5. **Micro ablation** (only if Echo v0 < best baseline).
6. **Full ablation study to find best Echo configuration:** Remove each modality, re-train, compare.
7. **Final evaluation:** Best Echo vs. top 2–3 baselines on full test set and all horizons.
8. **(Optional) Visualization Dashboard.**
9. **(Optional) LLM interpretability:** Use an LLM to interpret and explain the predictions from the input data.

7 Deliverables

- **Full Data Pipeline**: Reproducible ETL process and dataset builder (train/val/test splits).
- **Final Model (Echo)**.
- **Evaluation Report**: Baselines vs. Echo performance, assessment per horizon and per metric, with full reproducibility details.
- **Interactive Dashboard (Optional)**: Visualizes historical and predicted dyad trends.
- **Code & Documentation**: GitHub and GitLab repositories with technical documentation.
- **Final Written Report**.

8 Risks & Mitigation

- High computational cost: full-model training requires substantial GPU memory, especially with larger batch sizes.
- Risk of both overfitting (high-frequency dyads dominating the signal) and underfitting (sparse dyads lacking enough data to learn meaningful patterns).
- Data sparsity in low-activity dyads: many country pairs register near-zero events per month, destabilising training; smoothing or filtering to the most active dyads may be required.
- Noisy or weakly relevant text signals may dilute predictive power.
- Long-horizon performance degradation.
- Concept drift: geopolitical relations shift through coups, wars, alliances, and sanctions, making historical patterns partially obsolete; restricting the input window may help.
- Extreme and short-term shocks (e.g., invasions, regime collapses, economic crises) disrupt historical trends and produce outlier targets

9 References

9.1 Papers

[1] J. S. Goldstein, "A conflict–cooperation scale for weis events data," *Journal of Conflict Resolution*, vol. 36, no. 2, pp. 369–385, 1992. [Online]. Available: https://www.jstor.org/stable/174480

[2] D. Kinsella, "Goldstein scale for event data," Portland State University, retrieved October 30, 2025. [Online]. Available: https://web.pdx.edu/~kinsella/jgscale.html

[3] T. Stöehr, P. Jäger, and S. Schutte, "An ordinal latent variable model of conflict intensity," *arXiv preprint arXiv:2210.03971*, 2022. [Online]. Available: https://arxiv.org/pdf/2210.03971

[4] Y. Chen, A. Jatowt, and M. Yoshikawa, "Conflict or cooperation? predicting future tendency of international relations," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3341105.3373929

[5] e. a. Nguyen, "Next-generation conflict forecasting: Unleashing predictive patterns through spatiotemporal learning," *arXiv preprint arXiv:2506.14817*, 2025. [Online]. Available: https://arxiv.org/pdf/2506.14817

[6] M. Croicu and S. P. von der Maase, "From newswire to nexus: Using text-based actor embeddings and transformer networks to forecast conflict dynamics," *arXiv preprint arXiv:2501.03928*, 2025. [Online]. Available: https://arxiv.org/pdf/2501.03928

[7] O. Zakotianskyi, "Beyond closed doors: An open-source ai framework for forecasting armed conflict," 2025. [Online]. Available: https://euridice.eu/wp-content/uploads/2025/09/Beyond-Closed-Doors-An-Open-Source-AI-Framework-for-Forecasting-Armed-Conflict-1.pdf

[8] X. Liu and Y. Shen, "Inter-state cyberattack prediction with graph-text fusion," SSRN Preprint 5439839, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5439839

[9] B. J. Kinne, "Network dynamics and the evolution of international cooperation," *American Political Science Review*, vol. 107, no. 4, pp. 766–785, 2013. [Online]. Available: https://www.cambridge.org/core/journals/american-political-science-review/article/network-dynamics-and-the-evolution-of-international-cooperation/E686C749732F836F21371BD99061331B

[10] D. Shukla and S. Unger, "Sentiment analysis of international relations with artificial intelligence," *Athens Journal of Sciences*, vol. 9, no. 2, pp. 91–106, 2022. [Online]. Available: https://www.athensjournals.gr/sciences/2022-9-2-1-Shukla.pdf

9.2 Datasets

[1] J. S. Goldstein, "A conflict–cooperation scale for weis events data," *Journal of Conflict Resolution*, vol. 36, no. 2, pp. 369–385, 1992. [Online]. Available: https://www.jstor.org/stable/174480

[2] D. Kinsella, "Goldstein scale for event data," Portland State University, retrieved October 30, 2025. [Online]. Available: https://web.pdx.edu/~kinsella/jgscale.html

[3] T. Stöehr, P. Jäger, and S. Schutte, "An ordinal latent variable model of conflict intensity," *arXiv preprint arXiv:2210.03971*, 2022. [Online]. Available: https://arxiv.org/pdf/2210.03971

[4] Y. Chen, A. Jatowt, and M. Yoshikawa, "Conflict or cooperation? predicting future tendency of international relations," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3341105.3373929

[5] e. a. Nguyen, "Next-generation conflict forecasting: Unleashing predictive patterns through spatiotemporal learning," *arXiv preprint arXiv:2506.14817*, 2025. [Online]. Available: https://arxiv.org/pdf/2506.14817

[6] M. Croicu and S. P. von der Maase, "From newswire to nexus: Using text-based actor embeddings and transformer networks to forecast conflict dynamics," *arXiv preprint arXiv:2501.03928*, 2025. [Online]. Available: https://arxiv.org/pdf/2501.03928

[7] O. Zakotianskyi, "Beyond closed doors: An open-source ai framework for forecasting armed conflict," 2025. [Online]. Available: https://euridice.eu/wp-content/uploads/2025/09/Beyond-Closed-Doors-An-Open-Source-AI-Framework-for-Forecasting-Armed-Conflict-1.pdf

[8] X. Liu and Y. Shen, "Inter-state cyberattack prediction with graph-text fusion," SSRN Preprint 5439839, 2025. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5439839

[9] B. J. Kinne, "Network dynamics and the evolution of international cooperation," *American Political Science Review*, vol. 107, no. 4, pp. 766–785, 2013. [Online]. Available: https://www.cambridge.org/core/journals/american-political-science-review/article/network-dynamics-and-the-evolution-of-international-cooperation/E686C749732F836F21371BD99061331B

[10] D. Shukla and S. Unger, "Sentiment analysis of international relations with artificial intelligence," *Athens Journal of Sciences*, vol. 9, no. 2, pp. 91–106, 2022. [Online]. Available: https://www.athensjournals.gr/sciences/2022-9-2-1-Shukla.pdf