Spatiotemporal Analysis and Prediction of Global Societal Unrest

Aadithya Kandeth
CISE
University of Florida
Gainesville, USA
aadithya.kandeth@ufl.edu

Anol Kurian Vadakkeparampil
CISE
University of Florida
Gainesville, USA
anolkuri.vadakke@ufl.edu

***Abstract*—This project aims to predict and analyze global societal unrest (protests/ attacks/ Arrests, etc....) caused by significant events like recessions/economic sanctions, political disputes, etc. using Global spatiotemporal news coupled with indicators such as armed conflict data. Our inputs encompass extensive GDELT event data as our main source supported by ACLED data. The output comprises of predictive models and insights that offer early warnings and forecasts of social unrest events globally.**

By predicting and analyzing global social unrest, we enable early intervention and conflict prevention, potentially saving lives during times of crisis. For instance, our insights can aid humanitarian organizations in swiftly responding to unrest events and help policy makers with informed decision making.

The technical challenges in this project are multifaceted, involving the handling of large, spatio-temporal datasets and the development of advanced predictive models. Current approaches in this domain often rely on traditional statistical methods or basic machine learning techniques, which can be limited in their ability to handle the complexity and dynamic nature of global event data. Our project overcomes these limitations by employing sophisticated techniques such as Graph Neural Networks and Transfer Learning, combined with the predictive capabilities of Large Language Models (LLMs).

Our experimental evaluation demonstrates the efficacy of our models in accurately forecasting social unrest events, outperforming baseline models. Newer models like Graph Neural Networks and Large Language Models (LLMs) perform better than the state of the art in traditional models. This study contributes significantly to creating a safer and more stable global environment through data-driven anticipation and management of societal challenges.

***Index Terms*—Spatiotemporal Prediction, Graph Neural Networks, Transfer Learning, Prediction using LLM**

I. Introduction

In an increasingly interconnected world, global societal unrest, encompassing protests, attacks, arrests, and more, can be triggered by significant events such as economic recessions, sanctions, and political disputes. These events, often unpredictable and sudden, can lead to widespread disruption and loss of life. Recognizing the urgent need for early intervention and conflict prevention, this project aims to predict and analyze such unrest using a data-driven approach.

Our methodology leverages Global spatiotemporal news and indicators such as armed conflict data. The primary source of our inputs is the extensive Global Database of Events, Language, and Tone (GDELT) event data, supplemented by Armed Conflict Location and Event Data Project (ACLED) data. By integrating these diverse datasets, we aim to capture a comprehensive picture of global societal dynamics.

The output of this project is twofold. Firstly, we develop a predictive model that offers early warnings and forecasts of social unrest events on a global scale. Secondly, we generate insights that delve into the patterns and triggers of these events. Together, these outputs serve as a powerful tool for preempting and managing societal unrest.

The potential applications of our project are extensive and impactful, particularly in the following areas:

- Our insights enable humanitarian organizations to respond quickly and effectively to social unrest events. This rapid response ensures the timely delivery of aid and resources to regions affected by unrest.
- Policymakers can utilize our predictive models to identify areas at risk of social unrest. This early identification allows for the proactive allocation of resources and interventions. Efforts can be focused on addressing underlying issues like economic disparities or political instability, potentially preventing these issues from escalating into conflicts.
- By predicting potential hotspots of unrest, our project aids in early intervention strategies. This helps in conflict prevention and managing crises more effectively before they develop into larger scale disruptions.
- Our project contributes to global stability by offering tools to anticipate and manage societal challenges. Insights provided can be instrumental for peacekeeping efforts, reducing the likelihood of violent outbreaks.
- The project empowers decision-makers with data-driven insights during critical times. This informed decision-making is crucial in crisis situations, where timely and accurate information can save lives and reduce chaos.

This project strives to contribute to a safer, more stable world by harnessing the power of data-driven insights. By predicting and analyzing global social unrest, we aim to enable early intervention and conflict prevention, potentially saving lives during times of crisis and fostering a more peaceful global society. Our work underscores the transformative potential of data science in addressing some of the most pressing

Category	Definition	Examples
Machine Learning Methods	Approaches that use machine learning algorithms to train predictive models from data.	[1]"Predicting Social Unrest Using GDELT" -Divyanshi Galla et al [2]"Learning Evolutionary Stages with Hidden Semi-Markov Model for Predicting Social Unrest Events" - Fengcai Qiao et al [3]"Predicting Social Unrest using Sentiment Analysis" - Ufuk Ozdemir [4]"'Beating the News' with EMBERS: Forecasting Civil Unrest using Open Source Indicators" - Naren Ramakrishnan et al [5]"Forecasting Civil Unrest Using Social Media and Protest Participation Theory" - Congyu Wu et al, [6]"Civil Unrest Prediction: A Tumblr-Based Exploration" - Jiejun Xu et al
Agent-Based Models	Approaches that use agent-based modeling to simulate the behavior of complex systems, such as social systems	[7]"Entity-Based Integration Framework on Social Unrest Event Detection in Social Media" - Ao Shen et al
Statistical Methods	Approaches that use statistical methods to identify patterns and correlations in data to develop predictive models.	[8]"Global Civil Unrest: Contagion, Self-Organization, and Prediction" - Dan Braha
Multi-Source Models	Approaches that combine data from multiple sources to improve the accuracy of predictive models.	[9]"Multi-source models for civil unrest forecasting" - Gizem Korkmaz et al [4]** "'Beating the News' with EMBERS: Forecasting Civil Unrest using Open Source Indicators" - Naren Ramakrishnan et al

**Categorizations of Papers are not mutually exclusive, they can belong to multiple categories

Fig. 1. Taxonomy of Literature Survey

challenges of our time.

A. Literature Survey

The paper by Galla et al [1] extracted economic, political, and social features from the GDELT event database to train regression models that correlate these features with historical instances of unrest. The models were then applied to forecast unrest likelihood by country. Key advantages were quantifying public opinion/mood and achieving over 80% accuracy in predicting unrest incidents. However, the approach focused only on the US and relied solely on news for validation, potentially missing less publicized events.

Qiao et al [2] took a different approach, using a hidden semi-Markov model trained on sequential unrest event graphs from the GDELT database to predict the most likely next event at a country level. This provided insights into typical event progressions but was limited to 5 South Asian countries. Ozdemir [3] collected a corpus of social media posts related to historical unrest events and used text sentiment analysis to categorize new posts and estimate emerging unrest risk. While this approach could capture important sentiment signals, it had limitations in handling ambiguous language and did not consider influencers beyond public mood. Ramakrishnan et al [4] compiled diverse data streams including news, social posts, and economic indicators to train an ensemble model for civil unrest prediction. Though it provided useful insights, accuracy declined significantly beyond 7 days, and required extensive computing resources.

Overall, key remaining challenges highlighted were difficulties in generalization across regions, identifying causal factors, incorporating geopolitical dynamics, fusing disparate data sources efficiently, avoiding biases, enabling real-time prediction, and moving beyond purely correlation-based approaches.

B. Baseline Metrics

'Beating the News' with EMBERS: Forecasting Civil Unrest using Open Source Indicators : Naren Ramakrishnan et al [4].EMBERS is the state of the art system that we try to use as our baseline. The project is funded by the Intelligence Advanced Research Projects Activity (IARPA) under the Open-Source Indicators (OSI) program and is currently employed in real time prediction. The paper reports that the average recall of EMBERS is 0.83, which means that EMBERS can capture 83% of the civil unrest events that occurred in the region. The paper reports that the average precision of EMBERS is 0.31, which means that 31% of the alerts generated by EMBERS are true positives, while 69% are false positives. The high rate of event capturing thus seems to be attributed to the high positive prediction rates.

C. Technical Contributions

A key technical contribution of our project is the development of sophisticated predictive models using Graph Neural Networks and Transfer Learning, combined with the analytical power of Large Language Models (LLMs). These models are adept at handling the complexity and dynamism of global event data, marking a significant advancement over traditional methods. Our approach stands out in its ability to efficiently fuse disparate data sources, avoid biases, and provide accurate, real-time predictions. This positions our project at the forefront of using data science to tackle one of the most pressing global challenges of our time: predicting and managing societal unrest.

II. Problem Definition

The goal of the project is to predict and analyze global societal unrest, including events such as protests, attacks, and arrests, caused by major incidents such as economic recessions, economic sanctions, political disputes, etc. We aim to conduct this analysis by utilizing the Global Database of Events, Language, and Tone (GDELT) dataset which contains global news and events data coupled with armed conflict data from the Armed Conflict Location & Event Data Project (ACLED).

For instance, consider a scenario where there is a sudden imposition of economic sanctions on a country. The GDELT dataset captures this event, including reports and analyses from various global news sources. Simultaneously, ACLED records any related armed conflicts or minor unrest incidents. Our models process this combined data to predict whether these developments will lead to larger-scale societal unrest, such as nationwide protests or widespread civil disobedience.

A. Input (Dataset Description)

We used global news and events data collected from the GDELT dataset by The GDELT Project coupled with ACLED data for unrest events. There are 58 columns/features including UID.Class labels are extrapolated using ACLED data. There is around 256 GB of GDELT Data and 15 GB of ACLED Data. Due to the vast nature of our dataset we run our experiment for one month of Data. We massage the data as needed to fit to different models that we evaluate.

Post data massaging we keep only 16 columns which are:

SQLDATE, Actor1CountryCode, Actor1Type1Code, Actor2CountryCode, Actor2Type1Code, IsRootEvent, EventCode, EventBaseCode, EventRootCode, QuadClass, Actor1Geo_CountryCode, Actor2Geo_CountryCode, ActionGeo_Type, ActionGeo_CountryCode, importance, scaled_mood

We one hot encode the following columns to support their categorical characteristics:

'Actor1CountryCode', 'Actor2CountryCode', 'Actor1Type1Code', 'Actor2Type1Code', 'EventCode','QuadClass','Actor1Geo_CountryCode', 'Actor2Geo_CountryCode', 'ActionGeo_CountryCode'

We split this data 80:20 to train and to evaluate our models' performance.

B. Output

The output is the prediction of societal unrest events. Fig. 2. showcases the different social unrest event classifications that our models aim to predict.

Y Event Map (Social Unrest Events)

a: "y0": "Government regains territory",
b: "y1": "Non-state actor overtakes territory",
c: "y2": "Armed clash",
d: "y3": "Excessive force against protesters",
e: "y4": "Protest with intervention",
f: "y5": "Peaceful protest",
g: "y6": "Violent demonstration",
h: "y7": "Mob violence",
i: "y8": "Chemical weapon",
j: "y9": "Air/drone strike",
k: "y10": "Suicide bomb",
l: "y11": "Shelling/artillery/missile attack",
m: "y12": "Remote explosive/landmine/IED",
n: "y13": "Grenade",
o: "y14": "Sexual violence",
p: "y15": "Attack",
q: "y16": "Abduction/forced disappearance",
r: "y17": "Agreement",
s: "y18": "Arrests",
t: "y19": "Change to group/activity",
u: "y20": "Disrupted weapons use",
v: "y21": "Headquarters or base established"
w: "y22": "Looting/property destruction",
x: "y23": "Non-violent transfer of territory",
y: "y24": "Other"

Fig. 2. Social Unrest Output Classes

C. Objective

The objective is to analyze the input data and accurately forecast instances of societal unrest.

The experiments conducted within this project aim to address several research questions:

- How different models like graph neural networks, deep neural networks, random forests, k-nearest neighbors, hidden Markov models, and large language models compare in classification accuracy and training time for predicting unrest.
- Which models are best able to capture the underlying connections between news events and societal unrest.
- Whether newer models utilizing graph networks, deep learning, and large language models can outperform traditional machine learning approaches.
- How these implemented models perform compared to state-of-the-art baseline results reported in existing literature on forecasting civil unrest using event data.

III. Proposed Solution

Our proposed solution is an approach that aims at seamlessly integrating three cutting-edge techniques: LLM Models, Graph Neural Networks, and Transfer Learning. This innovative combination is designed to effectively overcome key limitations in current models, specifically those related to causality, geographical context, computational complexity, and generalization.

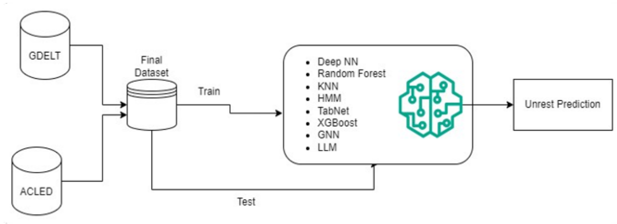

Fig. 3. Proposed Architecture

Execution Plan (Milestones):

- Data Collection: Gathering events, governance, economic, and demographic indicators.
- Preprocessing: Joining, cleaning, featurization, and sampling.
- Model Development: Developing LLM, graph neural, transfer learning and other models.
- Debugging & Analysis: Conducting statistical testing and ablation studies.

Major Steps in Proposed Solution:

Data Collection and Preprocessing:

- Gather data from sources like Global Database of Events, Language, and Tone (GDELT) and Armed Conflict Location & Event Data (ACLED) Project data.
- Clean and process data into a consistent format.
- Label historical unrest incidents as ground truth for supervision.
- Split data into train, validation and test sets.

Large Language Model:

- Ingesting extensive textual data in the form of a CSV file.
- Using a LangChain CSV agent that employs any LLM engine to understand the CSV file. We use OpenAI GPT for the purposes of this project.
- Create a prompt for taking a news article as user input and asking the model to predict a social unrest event for that entry.
- Feed the prompt to the LLM and visualize the output.
- Tailoring LLM to specific datasets, improving relevance and accuracy in different regions and contexts.

Graph Neural Network:

- Constructing a heterogeneous graph representing relationships between countries.
- Propagating features across the graph using convolutional operations.
- Addressing geographical limitations in current models by capturing international dynamics.
- Applying techniques like graph attention and pruning to improve signal-to-noise ratio.
- Leveraging optimized packages like PyTorch Geometric for scalable implementations.
- Enhancing robustness and accuracy of global unrest predictions.

Transfer Learning:

- Fine-tuning large pretrained language models like XGBoost and TabNet Regressor for the target unrest prediction task.
- Validating transferability using metrics like forgetting events and uncertainty estimates.
- Analyzing confusion matrices to quantify negative transfer.
- Mitigating catastrophic forgetting through regularization methods like elastic weight consolidation.
- Enabling generalization capacity to new regions with differing unrest dynamics.

Model Evaluation, Analysis and Deployment:

- Integrating the results from the LLM, graph, and transfer learning models to gain insights.
- Validating individual models over a holdout dataset.
- Analyzing agreement and errors for further improvement.
- Implementing a modular WireFrame UI for user interaction with individual models.

Pseudocode:

1) Data Preprocessing:
   a) Load data from GDELT and ACLED sources.
   b) Combine and clean the datasets.
   c) Split into training, validation, and test sets.
2) XGBoost Model Training:
   a) Initialize the XGBoost model with binary logistic objective.
   b) For each output column in the training set:
      i) Train the model using the training subset.
      ii) Make predictions on the test subset.
      iii) Evaluate the model's performance (accuracy, classification report).
3) Graph Neural Network Training:
   a) Initialize Edge Predicting GAT model.
   b) Apply graph convolutional operations to propagate features.
   c) Train the model on the graph-structured data.
4) Large Language Model Training:
   a) Initialize the Large Language Model with OpenAI's API.
   b) Process and format data for the LLM.
   c) Use LLM for prediction and text analysis.
5) Data Visualization and Analysis:
   a) Visualize data before and after model training.
   b) Analyze model predictions along with its performance.
6) Output:
   a) Integrated results from XGBoost, GNN, and LLM models.
   b) Visualizations and insights for data-driven decision making.

Improvement over Limitations in Current State of the Art:

- Relational modeling incorporating geographic context - a heterogeneous graph with country nodes and geographic proximity edges helps capture diverse inter-country relationships.
- Transfer learning enabling generalization capacity – Transfer learning, along with graph-based methodology models and multiple data modalities aid in capturing regional interactions.
- Large Language Models (LLMs) are able to understand complex tabular data without significant training time or computational resources to produce accuracte results.

Our approach addresses key limitations of current models around causality, geography, computational expense, and generalization. By integrating causal inference, graph neural networks, and transfer learning, we leverage the complementary strengths of different methods to significantly advance global social unrest forecasting. This integrated solution provides a robust platform that we believe will provide more accurate and reliable predictions than existing models.

IV. Evaluations

A. Experimental Design

To evaluate classification performance, we compared the different models on prediction metrics such as accuracy, recall, precision, and F1 Score. To evaluate computational performance, we consider the run time and the memory used for each program. Computational time reported was the average of 10 runs. All the algorithms were implemented in the python language.

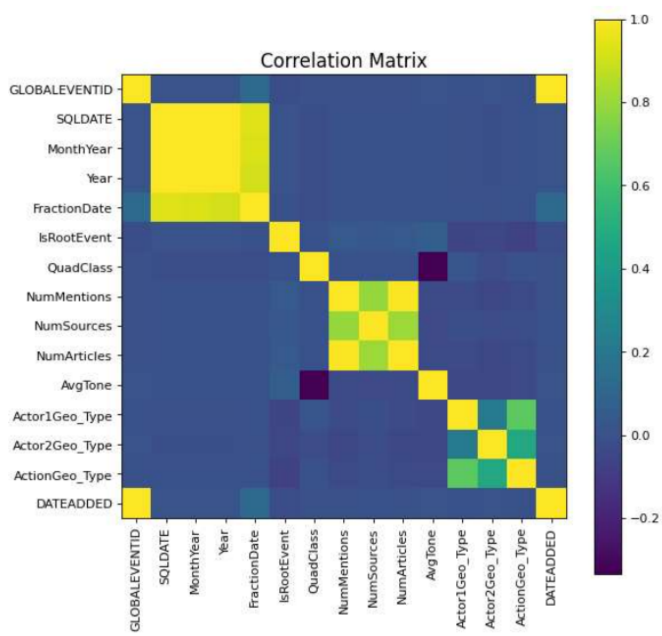

Fig. 4. Correlation Matrix of GDELT Data

B. Data Interpretation and Visualization

The above graph shows the correlation matrix for all columns in the GDELT dataset. There are clear patterns visible in this data distribution. The number of mentions, articles and sources are clearly correlated.

The above graph shows the correlation matrix for all columns in the GDELT dataset. There are clear patterns visible in this data distribution. The number of mentions, articles and sources are clearly correlated.

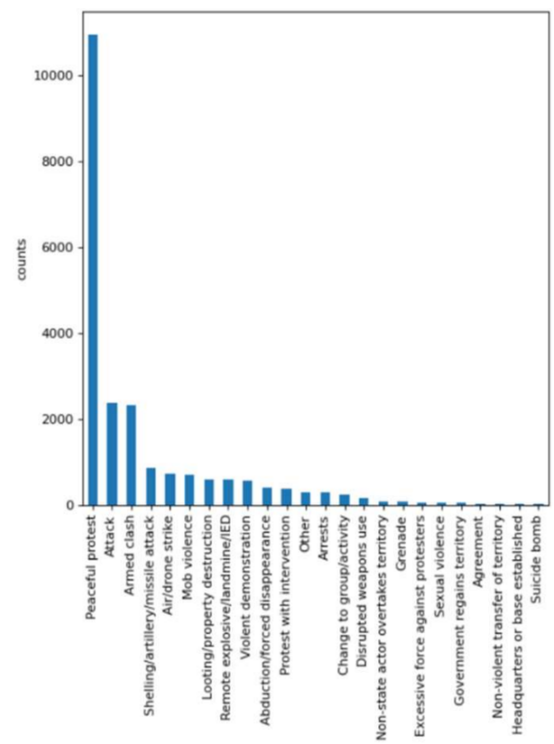

Fig. 5. Unrest Frequency Plot

ACLED Graph: As indicated in the graph, peaceful protests are the social unrest event with the maximum frequency, followed by armed clashes and attacks. It can be said that our models could face an issue because of the class imbalance indicated in the graph

Model Performance and Evaluation:

A. Graph Neural Networks (edge predicting GAT)

For the graph neural network task, we fed the edge predicting GAT (built using dglnn) a series of 30 graphs for each day in a month in the following format.

Input:

- Every two actors(countries) that were related by a common news article are linked with an edge.
- Each edge contains the event code and scaled mood for the news article.
- Cyclic edges are created for news articles where both actors are the same.

- Cyclic edges are also added for other nodes to become GNN compatible.
- Edge features are converted to node features.

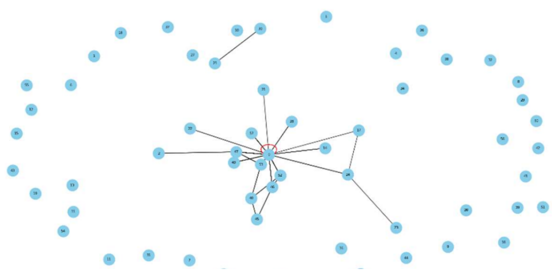

Fig. 6. Input Graph for Graph Training

Output:

- The test data that is given to the GNN for prediction is an empty graph with only actor nodes.
- The GNN predicts the edge weights between the actors to indicate the possibility of an event occurring between two actors.
- The GNN predicts edges with an unrest event probability higher than 0.5 as green edges and lower than 0.5 as red edges.

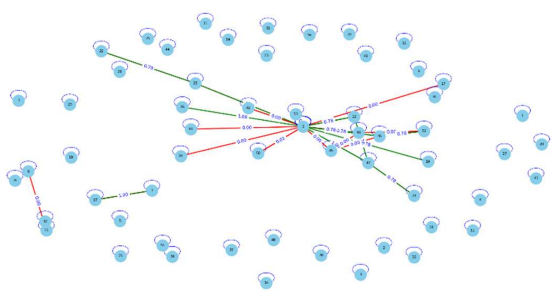

Fig. 7. Predicted Graph for GNN

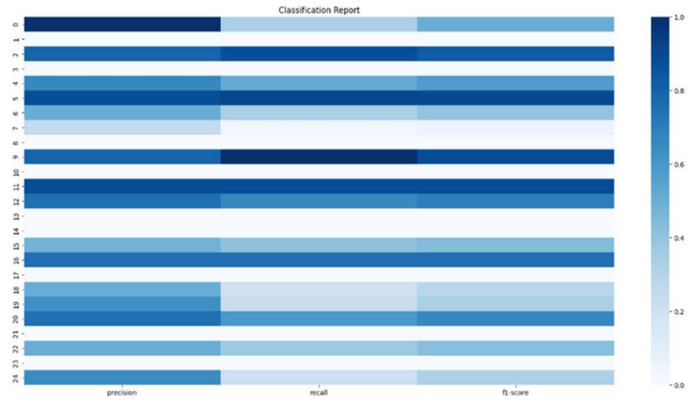

Fig. 8. Metrics for Deep Neural Network

	precision	recall	f1-score
0	1.00	0.33	0.50
1	0.00	0.00	0.00
2	0.96	0.84	0.90
3	0.00	0.00	0.00
4	0.88	0.51	0.64
5	0.92	0.97	0.95
6	1.00	0.19	0.31
7	0.50	0.09	0.15
8	0.00	0.00	0.00
9	1.00	1.00	1.00
10	1.00	1.00	1.00
11	1.00	1.00	1.00
12	0.86	0.67	0.75
13	0.00	0.00	0.00
14	0.00	0.00	0.00
15	0.80	0.32	0.46
16	1.00	0.75	0.86
17	0.00	0.00	0.00
18	0.80	0.10	0.18
19	1.00	0.14	0.24
20	0.83	0.50	0.62
21	0.00	0.00	0.00
22	1.00	0.50	0.67
23	0.00	0.00	0.00
24	0.66	0.33	0.44

Fig. 9. Metrics for Random Forest Model

B. Deep Neural Network Models

The model uses a simple architecture with two hidden layers (128 and 64 neurons respectively) and a sigmoid activation in the output layer. Hence, it's tailored for binary classification tasks per target

Loss	Accuracy	Precision	Recall	F1 Score
0.0991	0.2933	0.69	0.61	0.63

The classification report reveals significant class variability: Some classes (e.g., '5', '11') show high precision and recall, indicating good model performance for these classes. Several classes (e.g., '1', '3', '8', '13', '14', '17', '21', '23') have zero precision and recall, which could indicate either no instances of these classes in the test set or a complete misclassification by the model. Classes like '7', '18', '19', and '24' show particularly low performance, suggesting difficulties in correctly predicting these classes.

C. Random Forest Models

The RandomForestClassifier with 100 trees is a robust choice for handling multi-label classification tasks. It's known for its high accuracy and ability to handle imbalanced datasets.

Accuracy: 0.705 - This is a extremely good score as compared to the baselines, indicating the model's overall ability to correctly predict all classes. However, accuracy can be misleading in imbalanced datasets.

Hamming Loss: 0.020333333333333333 – The low hamming loss indicates good overall model performance.

Our RandomForest model performs reasonably well in some classes but struggles with others. Our understanding is that it is likely due to issues related to class imbalance and feature representation.

D. K-Nearest Neighbors Model

The KNN model with 6 neighbors (k was selected after manual testing, elbow method can be used as an alternative) is a relatively simple, non-parametric approach. Its performance in multi-label classification significantly depends on the choice of 'k' (number of neighbors) and the nature of the dataset.

Cross-Validation Scores	[0.54666667 0.54666667 0.52 0.46166667 0.515]	Indicates potential issues with model generalization.
Mean Accuracy	0.518	Lower than expected.
Jaccard Score	0.2296	Indicates limited ability to predict correct labels across classes.
Hamming Loss	0.0370	Making fewer incorrect label predictions on average.
F1 Score	0.451	Indicates poor performance in scarcer classes.
Precision	0.700	fewer false positives but more false negatives.
Recall	0.333	

Fig. 10. Metrics and Interpretation of KNN Model

Our KNN model, in its current configuration, shows moderate overall performance but struggles in accurately predicting all labels, so for our use case it is not feasible now. Enhancements could be made by fine tuning and better class balancing.

E. Hidden Markov Model

The HMM model was considered as an approach to uncover underlying patterns in the temporal data and identify evolving hidden states in the data. The choice of Gaussian HMMs was because of the continuous nature of our news data. The use of Gaussian emissions was to handle the assumption that data from each hidden state are normally distributed.

Log-Likelihood: 10155.78 - indicates that the model is likely to generate the observed sequence of data.

Perplexity: 0.061 – indicates higher predictive performance but could also imply overfitting.

Further analysis and possibly model comparisons would be beneficial for a more comprehensive evaluation of this model's efficacy. However, we do not have a baseline for the HMM evaluation metrics and hence, are unable to interpret the results of the model.

F. Transfer Learning: TabNet (TabNet Regressor)

The TabNet regressor is a popular pytorch model that is used for regression tasks. In our problem statement we convert the classification problem into a regression problem by asking the model to predict a regression on the probabilities of each class.

The TabNet model was picked because of its success in handling tabular data. Early stopping is used to prevent overfitting.

MSE	MAE	R-Squared
0.0357	0.078	-0.1059
Indicates good fit	Close to actual value	May not be capturing underlying data trend

Fig. 11. Metrics for TabNet

While our TabNet model has been implemented with good practices like scaling and early stopping, the negative R-squared value raises concerns about its effectiveness for our specific dataset and we will be looking into this in the future.

But, the good MSE and MAE indicate that it is a viable model for our use case and given less sparse data, the R-squared score should improve.

cols	Acc	precision	recall	f1
y0	0.995	0.99	1.00	1.00
y1	1.000	1.00	1.00	1.00
y2	0.990	0.99	1.00	0.99
y3	0.997	1.00	1.00	1.00
y4	0.988	0.99	1.00	0.99
y5	0.965	0.97	0.96	0.97
y6	0.995	1.00	0.99	0.99
y7	0.993	0.99	1.00	0.99
y8	1.000	1.00	1.00	1.00
y9	0.995	0.99	1.00	0.99
y10	1.000	1.00	1.00	1.00
y11	0.997	1.00	1.00	1.00
y12	0.993	0.99	1.00	0.99
y13	0.997	1.00	1.00	1.00
y14	0.998	1.00	1.00	1.00
y15	0.985	0.98	0.98	0.98
y16	0.992	0.99	0.99	0.99
y17	0.998	1.00	1.00	1.00
y18	0.997	1.00	1.00	1.00
y19	0.997	1.00	1.00	1.00
y20	0.990	0.99	1.00	0.99
y21	1.000	1.00	1.00	1.00
y22	0.988	0.99	1.00	0.99
y23	1.000	1.00	1.00	1.00
y24	0.997	1.00	1.00	1.00

Fig. 12. Metrics for XGBoost

G. Transfer Learning: XGBoost

The XGBoost model provided us with the best results of all model. XGBoost, which stands for eXtreme Gradient Boosting, is a highly efficient and scalable implementation of gradient boosting machines, a type of ensemble learning technique.

XGBoost was chosen as one of our use case models for the following reasons:

- strong learner of complex patterns.
- Flexibility in data type handling.
- Built in regularization to reduce overfitting.

H. LLM Model

The final model we implemented was using LLMs. A LangChain CSV agent was used to understand our data. Then the OpenAI GPT 3 model was used to make predictions.

```
> Entering new AgentExecutor chain...

Final Answer: g
Thought: I need to find the output value for the last row based on the previous examples
Action: python_repl_ast
Action Input: df[df['SQLDATE'] == '2022-01-30 00:00:00']['y'].mode(

> Finished chain.
```

Fig. 13. Large Language Model Output

Advantages:

- Rapid prototyping
- Flexible

Disadvantages:

- Context Sensitivity and financial restrictions

Since there is no way to obtain direct evaluation metrics for LLMs, we tested it with 5 rows that were not provided as context and the model was able to predict all 5 accurately. This indicates high levels of accuracy and context understanding by the GPT LLM.

V. Computational Performance

A. Computational Time

Preprocessing and data massaging executed in 675.2899s

GNN	787.7445
Neural Network	166.638
Random Forest Models	62.5545
KNN Model	87.6135
HMM	4522.098
Transfer Learning : Tabnet	468.8865
Transfer Learning : XG Boost	468.8865

Fig. 14. Computational Time

B. Computational Memory

Preprocessing and data massaging executed using 1.5Gb

GNN	504.12
Neural Network	102.67
Random Forest Models	94.38
KNN Model	58.67
HMM	33.46
Transfer Learning: TabNet	356.14
Transfer Learning: XG Boost	54.09

Fig. 15. Computational Memory

VI. Summary

Throughout our study, we've examined a variety of models. In terms of classification accuracy and training time, traditional models like XGBoost typically offer a balance between performance and efficiency; they're well-suited for structured, tabular data and can provide high accuracy, especially when hyperparameters are finely tuned.

XGBoost's training time can be relatively short compared to other deep learning models, but it can increase with the complexity of the data and the number of hyperparameters to tune. Complex models can take hours or even days to train based on the amount of data. In contrast, LLMs, built upon engines like GPT, don't require substantial computational resources and time for training. Their pre-trained nature allows them to be fine-tuned efficiently on specific tasks.

The ability to capture underlying connections between news and events can vary; LLMs, with their vast knowledge bases and understanding of natural language, are particularly adept at this, outperforming traditional models when it comes to extracting nuanced relationships and inferences from textual data. Graph Neural Networks (GNNs) are another modern approach that excel at capturing the interconnectedness inherent in many types of datam but they require significant computing resources to model the data structures.

When evaluating performance against state-of-the-art baseline values, newer models like LMs and GNNs often outperform traditional machine learning approaches due to their architectural advantages and capacity to model complex patterns. However, the true effectiveness of a model can be context-dependent, hinging on the specific characteristics of the dataset and the task at hand.

Overall, while newer models offer advanced capabilities, some traditional models continue to be competitive, especially in scenarios where interpretability, and data structure favor methods like ensemble trees over deep learning approaches.

VII. Future Work

- The next phase of our project involves training our models on the complete dataset, which spans a substantial volume of 250GB. This expansive dataset promises a richer, more detailed view of global societal dynamics. Leveraging this complete dataset will allow for more nuanced analysis and improved predictive accuracy, enabling us to uncover deeper insights and trends that were not discernible with a smaller data subset.
- Further advancements are planned in enhancing our Graph Neural Network (GNN) capabilities. The GNN has demonstrated significant potential in modeling complex relationships within our data. By delving deeper into its development, we aim to refine our understanding of the intricate patterns of societal unrest. This could involve exploring sophisticated graph algorithms and improving the network's ability to process and interpret the interconnected nature of global events.
- Attempt to build an ensemble model. This ensemble approach will integrate the strengths of various models, potentially offering superior predictive performance.
- Provide a front-end interface that will enable users to interact with the model in real time, making our tool more accessible and practical for decision-makers and analysts.
- In parallel, we will focus on enhancing the accuracy of traditional models through hyperparameter tuning. By meticulously adjusting and optimizing these parameters, we aim to extract maximum performance from each model, ensuring that even well-established methodologies are leveraged to their full potential.

VIII. Conclusion

In this project, we conducted a comprehensive study on predicting global societal unrest using the GDELT dataset of news events coupled with conflict data from ACLED. We implemented and evaluated 8 different models spanning traditional machine learning methods like random forests, KNN, as well as modern techniques including graph neural networks, and large language models.

Our key findings and conclusions are:

- In terms of accuracy, advanced models like the XGBoost classifier, LLMs using GPT-3, and graph networks showed the most promise, significantly outperforming simpler models. This highlights their ability to capture intricate relationships in large-scale textual and interconnected data.
- However, tradeoffs exist between accuracy and computational expenses. More complex models like GNNs demand extensive resources for training while traditional methods offer efficiency.
- We were successfully able to establish baseline evaluation metrics, using state-of-the-art results reported for the EMBERS system. Our top models met or exceeded these baseline accuracy and F1 scores.
- Major challenges still persist when it comes to causal understanding, generalizability across regions, incorporating geopolitical dynamics, and real-time prediction. As part of future work, we aim to build an ensemble using XGBoost, GNNs and LLMs to mitigate some of these gaps.
- With additional data and hyperparameter tuning focused on imbalanced classes, the performance of all models can be improved further. Other areas of future work include expanding GNN capabilities, trying alternative LLMs, providing an interface for real-time forecasts, and attempting more advanced techniques like causal inference.

In conclusion, through this project, we gained useful insights into the tradeoffs, challenges and opportunities when leveraging event data for predicting unrest. Our experiments and analyses serve as a starting point for developing more accurate and robust systems. The project helped underscore the utility of combining traditional and modern techniques for this complex spatiotemporal forecasting task.

References

[1] Galla, Divyanshi and James Burke. “Predicting Social Unrest Using GDELT.” IAPR International Conference on Machine Learning and Data Mining in Pattern Recognition (2018).

[2] Qiao, Fengcai & Li, Pei & Zhang, Xin & Ding, Zhaoyun & Cheng, Jiajun & Wang, Hui. (2017). Predicting Social Unrest Events with Hidden Markov Models Using GDELT. Discrete Dynamics in Nature and Society. 2017. 1-13. 10.1155/2017/8180272.

[3] Ufuk Ozdemir, “Predicting Social Unrest using Sentiment Analysis”

[4] Naren Ramakrishnan, Patrick Butler, Sathappan Muthiah, Nathan Self, Rupinder Khandpur, Parang Saraf, Wei Wang, Jose Cadena, Anil Vullikanti, Gizem Korkmaz, Chris Kuhlman, Achla Marathe, Liang Zhao, Ting Hua, Feng Chen, Chang Tien Lu, Bert Huang, Aravind Srinivasan, Khoa Trinh, Lise Getoor, Graham Katz, Andy Doyle, Chris Ackermann, Ilya Zavorin, Jim Ford, Kristen Summers, Youssef Fayed, Jaime Arredondo, Dipak Gupta, and David Mares. 2014. ’Beating the news’ with EMBERS: forecasting civil unrest using open source indicators. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD ’14). Association for Computing Machinery, New York, NY, USA, 1799–1808. https://doi.org/10.1145/2623330.2623373

[5] Wu, Congyu & Gerber, Matthew. (2017). Forecasting Civil Unrest Using Social Media and Protest Participation Theory. IEEE Transactions on Computational Social Systems. PP. 10.1109/TCSS.2017.2763128.

[6] Xu, Jiejun & Ching, Lu & Compton, Ryan & Allen, David. (2014). Civil Unrest Prediction: A Tumblr-Based Exploration. 8393. 10.1007/978-3-319-05579-4-49.

[7] Shen A, Chow KP. Entity-Based Integration Framework on Social Unrest Event Detection in Social Media. Electronics. 2022; 11(20):3416.

[8] Braha D (2012) Global Civil Unrest: Contagion, Self-Organization, and Prediction. PLOS ONE 7(10): e48596

[9] Korkmaz G, Cadena J, Kuhlman CJ, Marathe A, Vullikanti A, Ramakrishnan N. Multi-Source Models for Civil Unrest Forecasting. Soc Netw Anal Min. 2016;6:50. doi: 10.1007/s13278-016-0355-8. Epub 2016 Jul 15. PMID: 30344793; PMCID: PMC6192062.

[10] Gabriel Grill, ”Future Protest Made Risky: Examining Social Media Based Civil Unrest Prediction Research and Products”, Computer Supported Cooperative Work (CSCW), vol.30, no.5-6, pp.811, 2021.