Predictive Analytics for Weather-Induced Infrastructure Failures: Societal Impact Assessment with Multi-Modal Deep Learning

Swathi Sri Lasya Mayukha Ramachandruni [1] **Sneha Nalla** [1]

Abstract

Extreme weather events increasingly disrupt infrastructure and communities. This project proposes a multi-modal predictive analytics framework to forecast power outage probabilities and durations and to assess associated societal impacts. We integrate the US Power Outage Database (2000–2020) (Mukherjee et al., 2018), NOAA Storm Events (NCEI, 2025), EIA grid data (EIC, 2025), and CDC Social Vulnerability Index (CDC, 2024), harmonizing spatial and temporal features. The technical approach uses temporal-spatial feature engineering, deep learning (LSTM), and ensemble methods (Random Forest, XGBoost) to generate accurate predictions and risk scores. Key challenges include merging heterogeneous datasets, imputation of missing data, and aligning multi-resolution records. Expected outcomes include outage prediction models with 75–80% accuracy for severe events, risk maps for emergency planning, duration estimation with sub-4-hour MAE, and actionable policy guidance for resilience improvement. This research advances spatiotemporal machine learning for disaster management and AI for social good.

1. Motivation

Climate-driven extreme weather disrupts critical infrastructure and disproportionately affects vulnerable groups. Current tools lack the capacity to capture multidimensional risk. Our project aims to identify high-risk regions and estimate societal impacts to improve emergency preparedness.

2. Technical Approach

We will:

- Harmonize four national datasets by spatial and temporal indexing
- Engineer features from meteorological, infrastructural, and social vulnerability data
- Train deep learning models (LSTM) and ensemble predictors (Random Forest, XGBoost) for outage/duration forecasting
- Validate with temporal/spatial cross-validation and performance metrics (F1, MAE)

3. Expected Outcomes

- Outage occurrence and duration models (75–80% accuracy; MAE < 4 hours)
- Vulnerability risk maps for targeted resource allocation
- Open-source modeling framework
- Policy recommendations for resilient infrastructure and equitable emergency response

4. Broader Impact

Our project will enable utility and emergency agencies to respond proactively, especially in underserved regions. It will also promote fair resource allocation and strengthen national resilience capacity through open access to models and data.

References

CDC. Social vulnerability index, July 2024. URL https://www.atsdr.cdc.gov/place-health/php/svi/?CDC_AAref_Val.

Ciszyński, M. and Chromiński, K. Applying machine learning techniques in weather forecast modeling. *Procedia Computer Science*, 246:4133–4141, 2024. ISSN 1877-0509. doi: https://doi.org/10.1016/j.procs.2024.09.252. URL https://www.sciencedirect.com/science/article/pii/S1877050924022713. 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES 2024).

EIC. Electricity data browser, June 2025. URL https://www.eia.gov/electricity/data/browser/.

Mukherjee, S., Nateghi, R., and Hastak, M. Data on major power outage events in the continental u.s. *Data in Brief*, 19:2079–2083, 2018. ISSN 2352-3409. doi: https://doi.org/10.1016/j.dib.2018.06.067. URL https://www.sciencedirect.com/science/article/pii/S2352340918307182.

NCEI. Storm events database, May 2025. URL https://www.ncei.noaa.gov/stormevents/.

[1]Department of Computer Science, San Jose State University, CA, USA. Correspondence to: Swathi Sri Lasya Mayukha Ramachandruni <swathisrilasya-mayukha.ramachandruni@sjsu.edu>, Sneha Nalla <sneha.nalla@sjsu.edu>.

Proceedings of the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.