TEAM DISASTER

Predicting Disaster-Prone Areas

AUTHORS

Stefan Bobick, Alberto Espinosa, Michael Sexton

FIGURE 1: HOUSES ARE PARTIALLY SUBMERGED FOLLOWING A DAM COLLAPSE DUE TO HEAVY RANINFALL IN MAIDUGURI, NIGERIA, SEPTEMBER 10, 2024

SOURCE: ALJAZEERA. (2024). RETRIEVED FROM HTTPS://WWW.ALJAZEERA.COM/NEWS/2024/9/11/NIGERIA-FLOODS-AFFECT-ONE-MILLION-PEOPLE-AFTER-DAM-COLLAPSE

PROBLEM SPACE

Natural disasters have a significant global impact: widespread community displacement, infrastructure damage, economic losses, injuries and deaths. More severe effects are posed to developing countries due to their particular vulnerability to extreme weather events, which are only further exacerbated by climate change. This project aims to utilize past historic disaster data to predict at-risk areas to amplify preparedness.

APPROACH

Data Preprocessing:

- Removed features with high frequency of missing data
- Encoded categorical variables
- Data imputation methods: median, mode, multivariate imputer that estimates each feature from all others

Methodology:

- Random Forest Ensemble Method
- Gradient Boosting Classifier
- Support Vector Machine

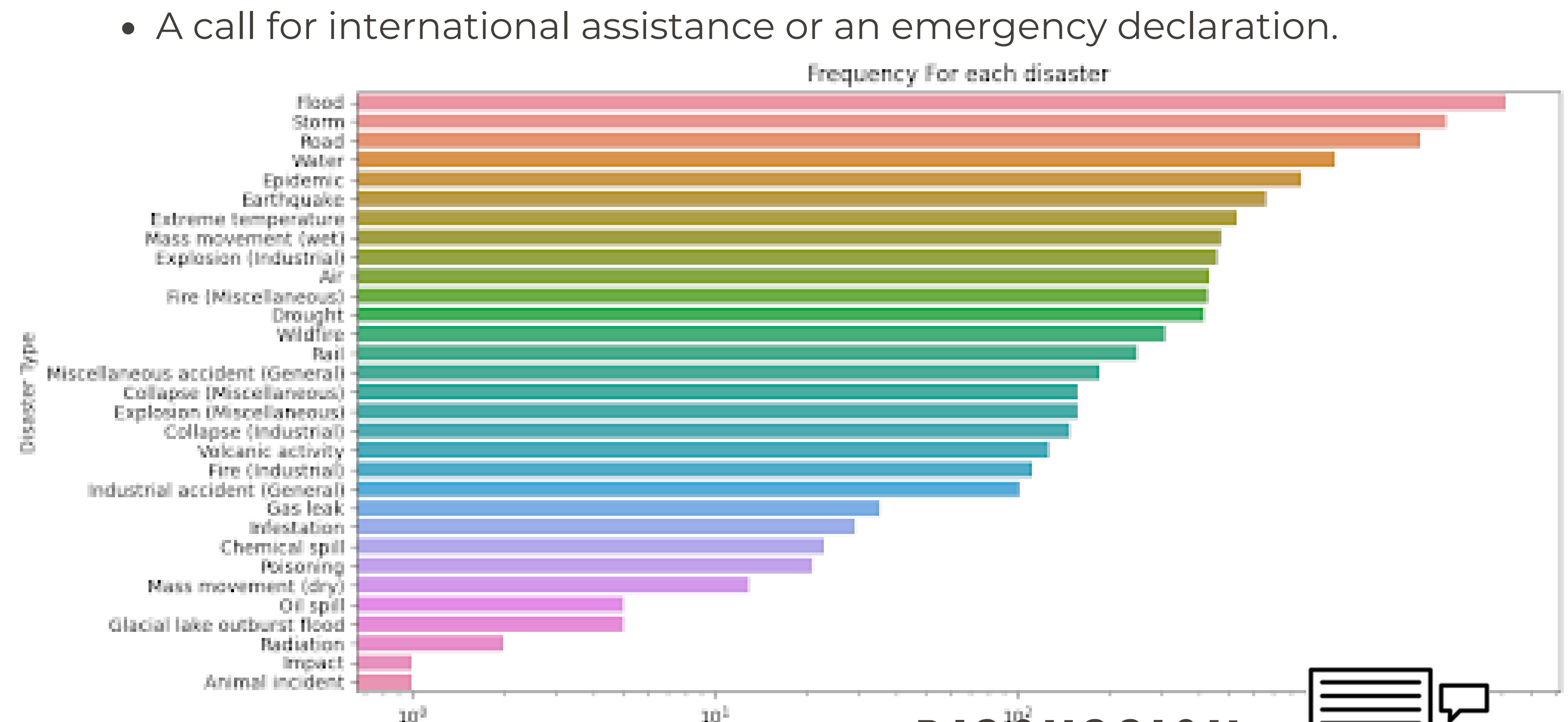
DATA

EM-DAT (Emergency Events Database) - 26,000 entries, 46 features

- **Classification** -> Disaster group, Subgroup, Type, Subtype
- **Location** -> Country, Region, Subregion
- **Financial impact** -> Gov't Response, Aid Contribution, Total damage, Insured Damage
- **Loss of Life** -> Total deaths, Number injured, Number homeless, Number affected
- **Temporal** -> Start Month, Start Day, Year, Duration.

The samples gathered meet one of the following criteria:

- At least **ten deaths** (including dead and missing).
- At least **100 affected** (people affected, injured, or homeless).
- A call for international assistance or an emergency declaration.

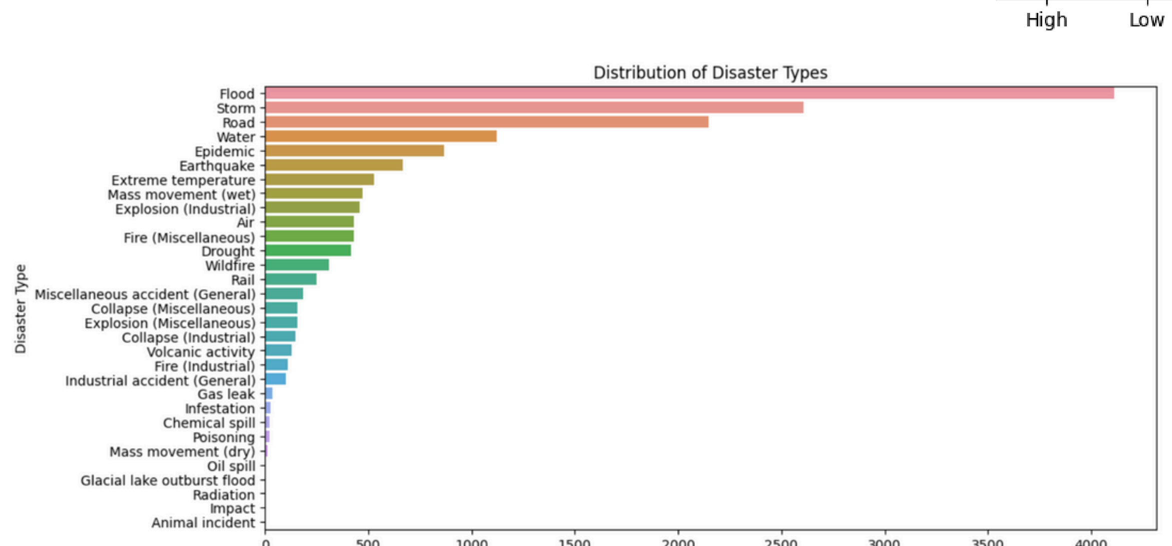


RESULTS

Random Forest Classifier:

Achieved accuracy: 57%

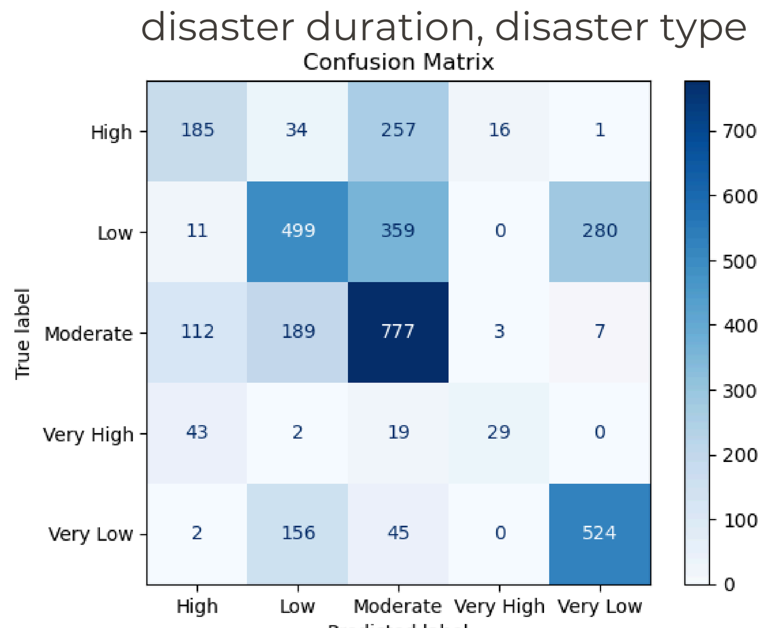
- Average performance in classifying the type of disaster based on total affected, total deaths region, subregion, country, start month, start year, event duration, start month
- This classifier should be able to determine which disaster types are the worst historically in terms of actual damage they caused. Then decide on where to reinforce infrastructure that would mitigate the potential damage.



Gradient Boosting Classifier:

Achieved Accuracy: ~60%

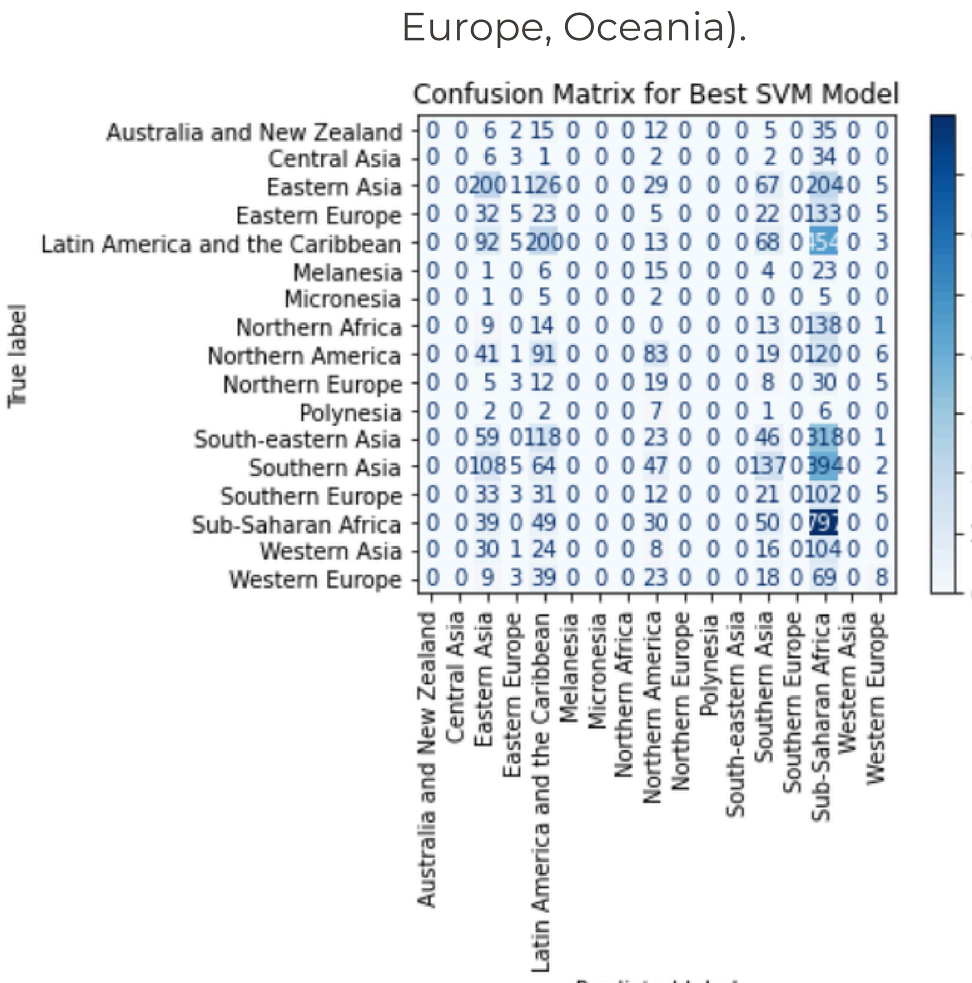
- Improved performance in classifying impact severity based on binning the normalized value of total affected to 'high', 'low', 'moderate', 'very high', and 'very low'
- Features trained on: region, sub region, country, start month, disaster duration, disaster type



Support Vector Classifier:

Achieved Accuracy: 28%

- Weak performance in classifying sub region based on features: disaster type, total deaths, total affected, no. injured, no. affected, CPI, start/end months, magnitude scale.
- Major class imbalance with Asia leading to heavy misclassification on minority classes (Americas, Africa, Europe, Oceania).



DISCUSSION

Outcome Goals

- Helps predict disaster type and disaster location
- Opens up conversation for optimal resource/financial allocation
- Amplifies readiness, planning, and infrastructure improvement

Directions to Investigate

- Integrate global financial response data
- Increased disaster frequency due to climate change

Potential Future Plans

- Integrate additional datasets with relevant features for more complex modeling
- Apply a neural network

RELATED LITERATURE

1. Biardeau, L. T., & Sahli, M. (2024). Investigating the non-linear impacts of seven types of natural disasters on inbound tourism: Insights from the EM-DAT database. Tourism Economics, 30(4), 900–923.
2. Jain, H., Dhupper, R., Shrivastava, A., Kumar, D., & Kumari, M. (2023). Leveraging machine learning algorithms for improved disaster preparedness and response through accurate weather pattern and natural disaster prediction. Frontiers in Environmental Science, 11.
3. Linardos, V., Drakaki, M., Tzionas, P., & Karnavas, Y.L. (2022). Machine Learning in Disaster Management: Recent Developments in Methods and Applications. *Mach. Learn. Knowl. Extr.*, *4*, 446-473.
4. Painter, W. L. (2020). The Disaster Relief Fund: Overview and issues. Specialist in Homeland Security and Appropriations. Updated April 16, 2020.