

Conflict Model Predictor

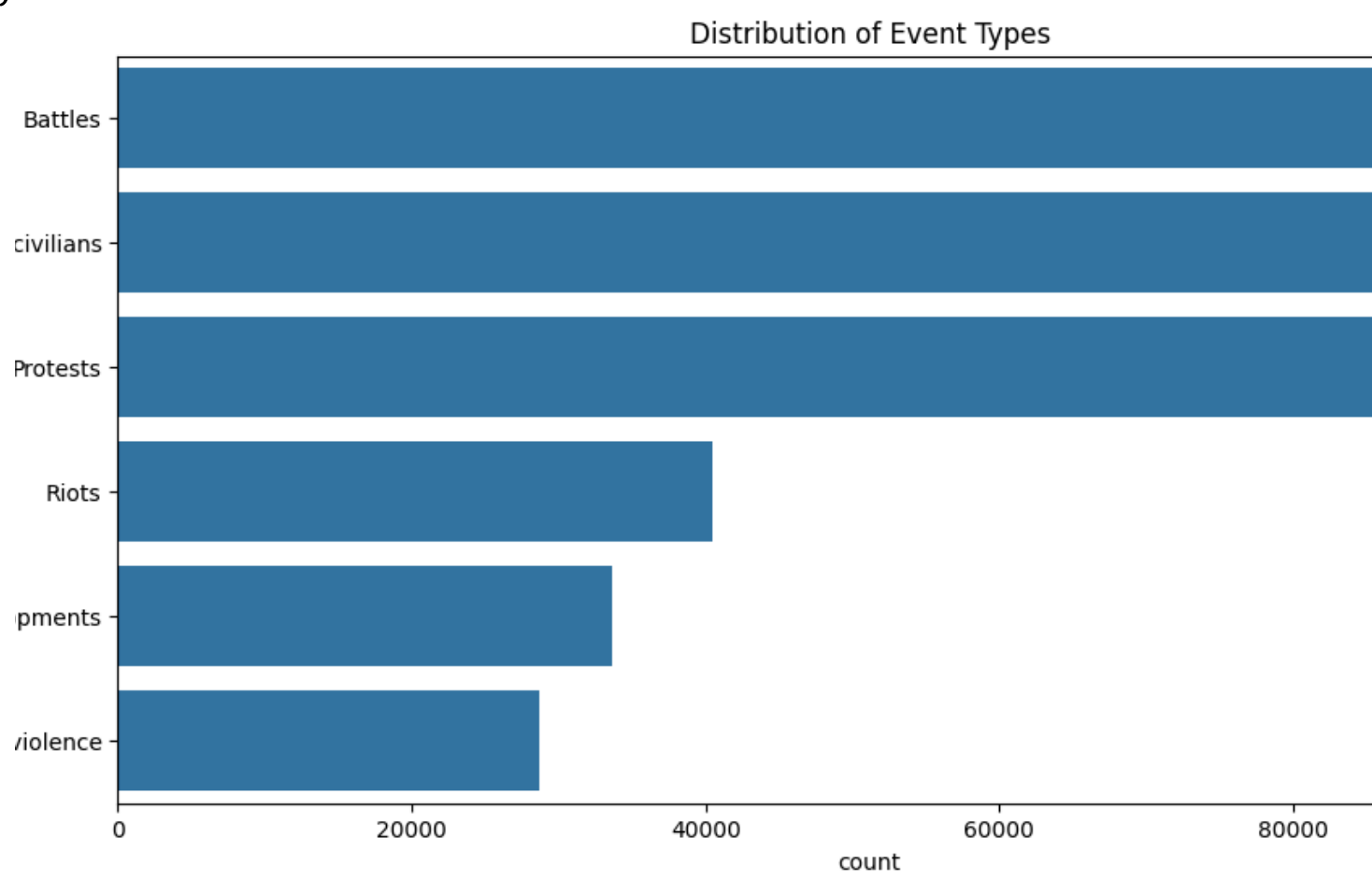
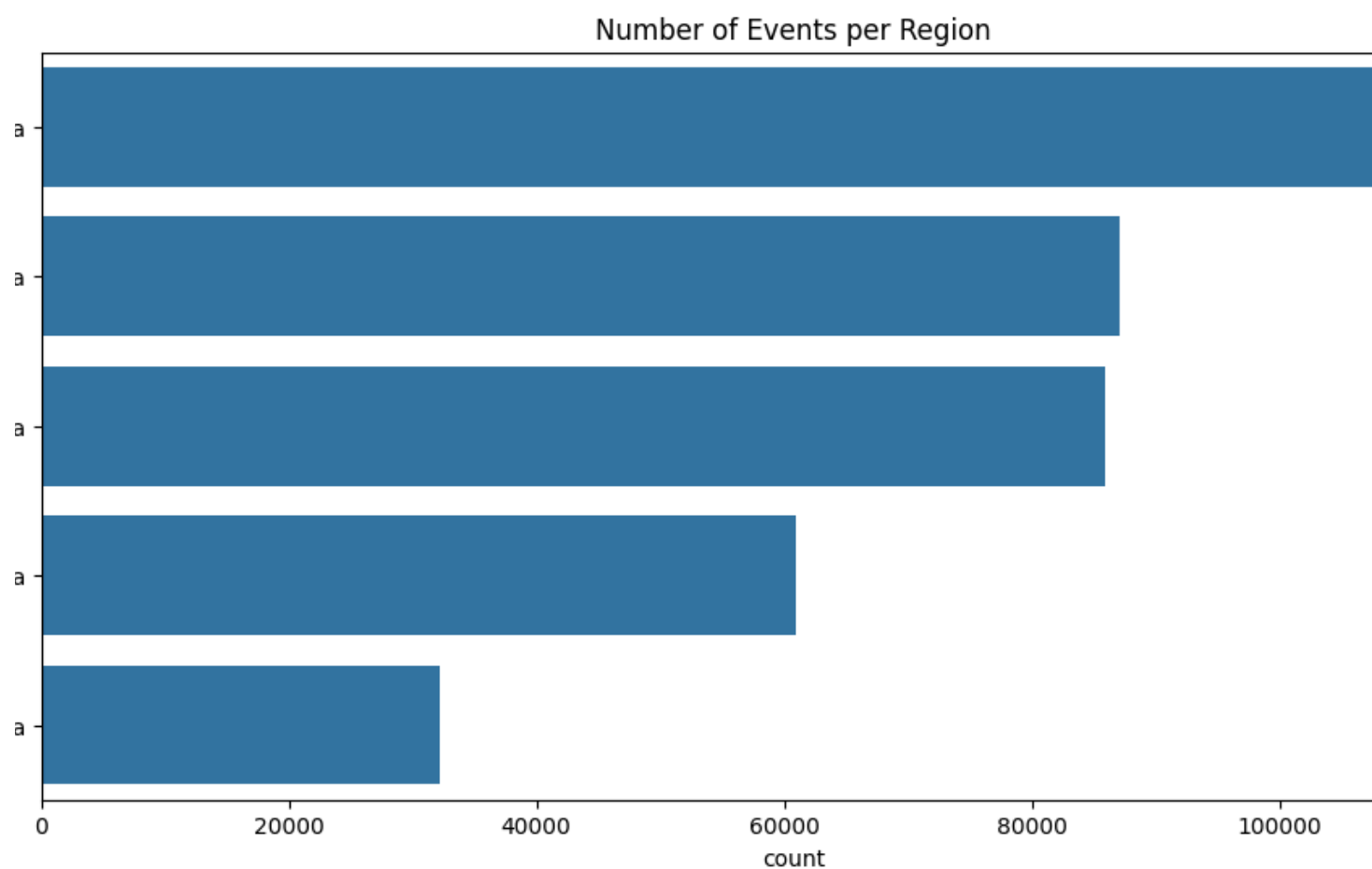
Analysis of Africa Conflicts
1990s –2023

Background and Objective



- Since the 1990s, conflicts in Africa have evolved from civil wars, such as in Rwanda and Liberia, to insurgencies driven by groups like Boko Haram and Al-Shabaab. Communal violence, political unrest, and terrorism have further destabilized regions, with ethnic tensions and resource competition fueling conflicts in places like South Sudan, Ethiopia, and the Sahel.
- These conflicts have shifted towards asymmetric warfare, targeting civilians, and expanding across borders. A conflict predictor using ACLED data would help governments, security agencies, and humanitarian organizations anticipate violence, allocate resources effectively, and design informed peacebuilding strategies, enabling proactive intervention and crisis prevention.

- The bar charts show regional disparities in the number of events across Africa. Eastern Africa has the highest count, nearing 120,000, indicating the most activity, followed by Western and Northern Africa with around 80,000 events each. Middle Africa shows moderate activity with just over 60,000 events, while Southern Africa has the lowest at approximately 40,000. This suggests Eastern Africa requires more focus, while Southern Africa might be under-represented or less active.



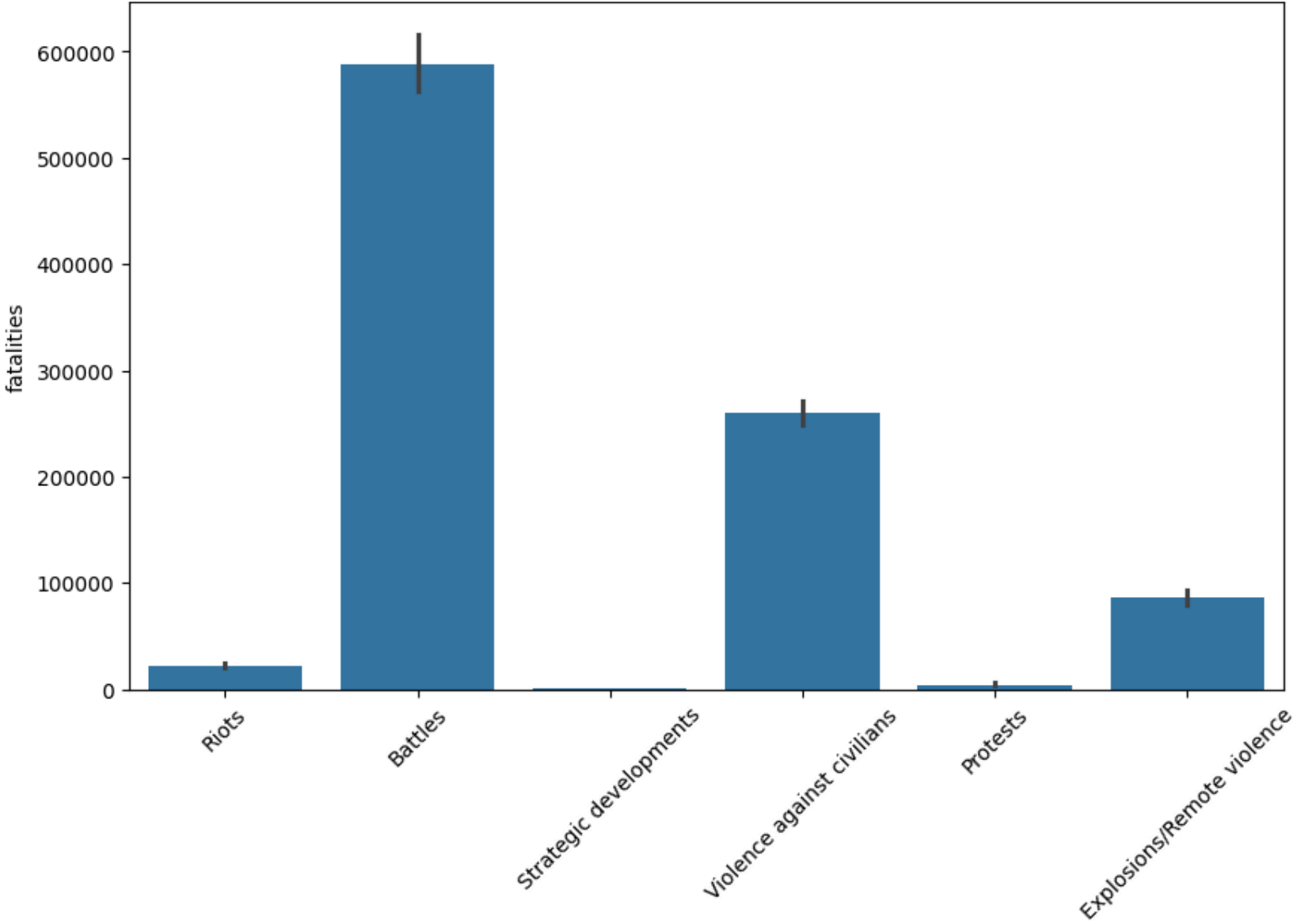
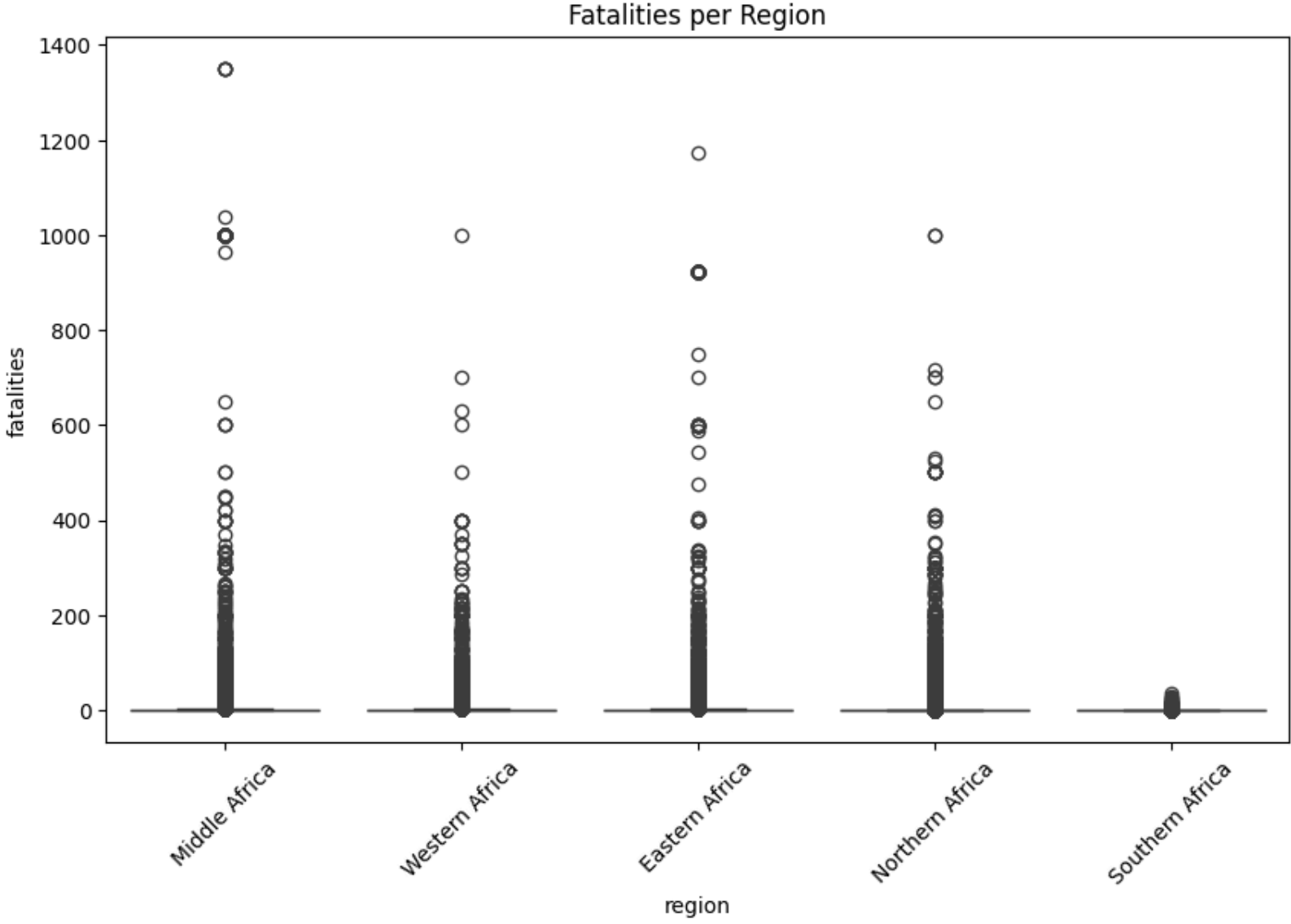
Conflicts Across Africa



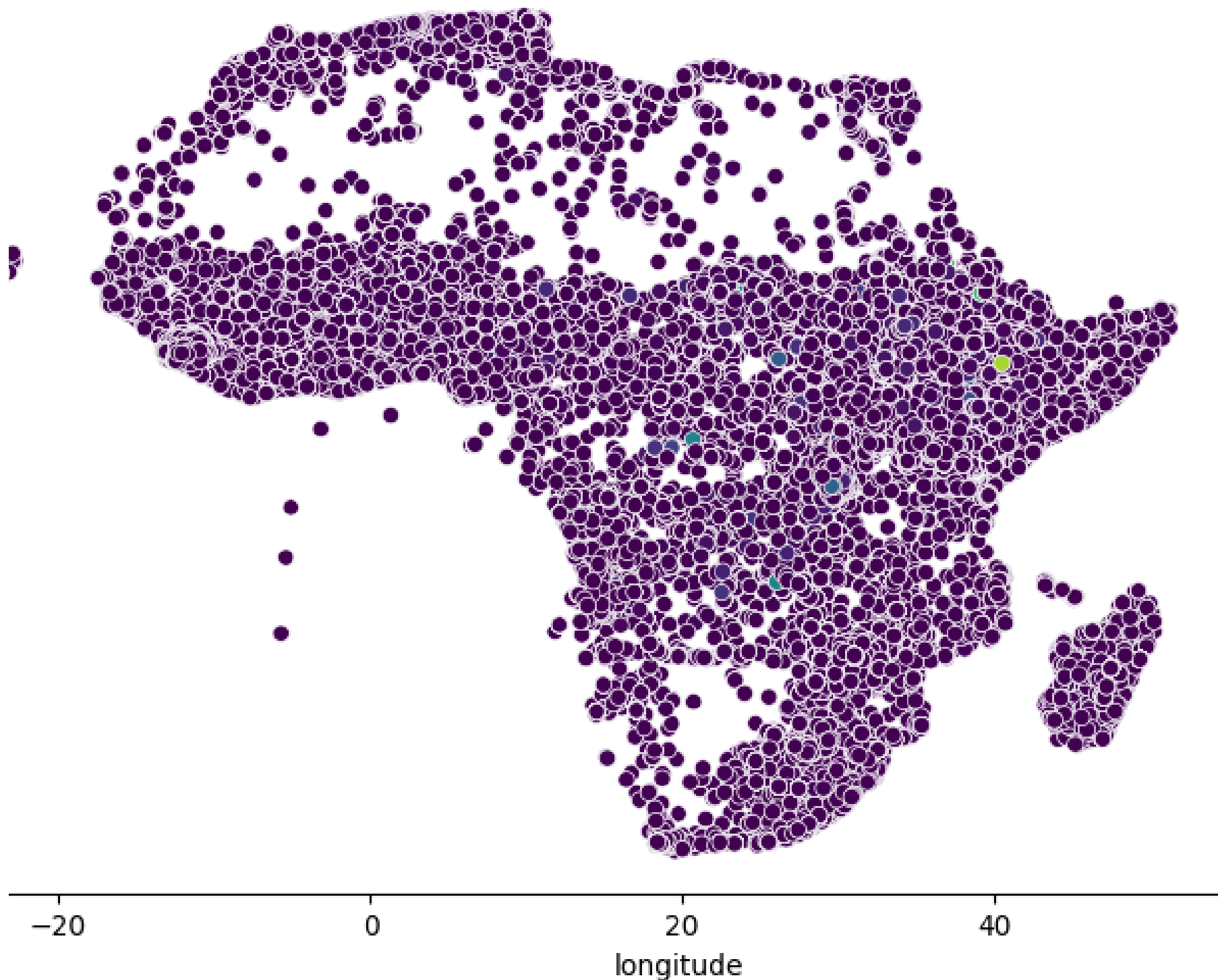
Variable Relationship

The leading cause of fatalities in conflict events, contributing to over 600,000 deaths, followed by violence against civilians and explosions/remote violence, which also have significant fatality counts

In contrast, riots, protests, and strategic developments lead to minimal deaths



Geographical Distribution of Events with Fatalities



Geographic representation of conflicts

Geographically, conflict events are widespread across Africa, with dense concentrations in regions with higher fatality counts.

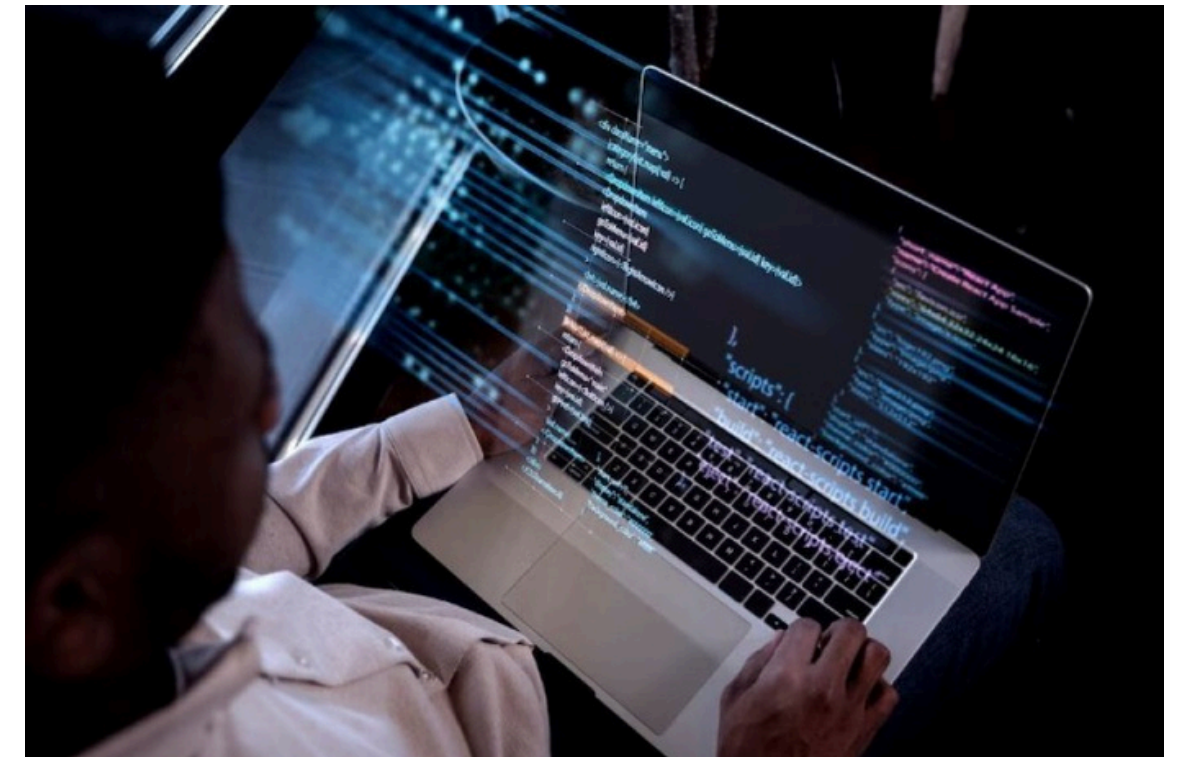
Model Selection & Performance



Logistic regression and Random Forest models proved to be the best machine learning models to train.



Firstly, the logistic regression provided an accuracy level of 0.71, after utilizing a pipeline method. In effort to improve the performance using the polynomial feature engineering. This tuning lowered the score to 0.69.



The Random Forest model performed the best with a 0.95 accuracy score, suggesting it may be the suitable model to predict the conflict classification problem.

Conclusion



The Random Forest model performs exceptionally well in classifying political violence events, especially for "Strategic Developments," "Violence against Civilians," and "Protests" (all with perfect scores).



While the model is highly accurate, it shows slightly lower performance in identifying "Battles," which could benefit from further fine-tuning

Recommendation

To enhance the prediction performance of the Random Forest model, the following top three recommendations should be considered:

1. Address Imbalanced Data: The significant class imbalance, particularly affecting the "Battles" class, impacts model accuracy. Mitigate this issue by oversampling the minority class, undersampling the majority class, or adjusting class weights to ensure fairness and accuracy across all classes.

2. Improve Class-specific Performance: The model shows poor precision and recall for the "Battles" class. To address this, collect more data for the "Battles" class or apply data augmentation techniques to enhance the model's ability to distinguish this class from others.

3. Optimize Hyperparameters: Although the model's overall accuracy is high, further improvements may be possible. Perform hyperparameter tuning using grid search or random search to find the optimal hyperparameters for the classification task.