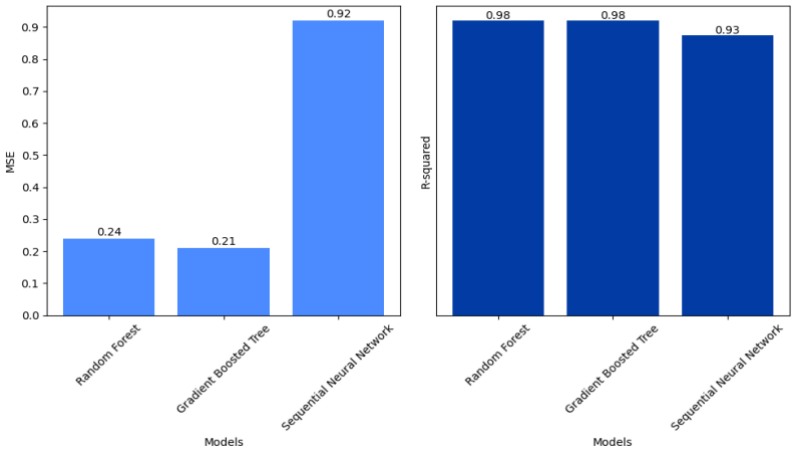
Executive Summary – G10 (United Nations Conflict Prediction)

Working with the United Nations Service for Geospatial, Information and Telecommunications Technologies (SGITT), group 10 was tasked with crafting a solution to develop an early-warning conflict system that could help the organization to both develop a deep understanding of the underlying factors at the most granular level as possible that lead to conflicts in the United Nations countries of operations and a quick-action lever for prevention of conflicts, fatalities and all the comprehensive economic losses encompassed.

For doing so, after a deep literature review in which the team clarified the variables to take into account, the models at their disposal and the approach; the main objective was set: to develop a predictive model to estimate the number of fatalities resulting from different types of conflicts on a sub-national scale, specifically in countries where the UN has a mission, with a prediction timeframe of 1-12 months ahead. The target variable was the total number of fatalities, which enabled better grasping a wider range of conflicts, their complexities and nuances by means of machine learning models (Schon & Kon, 2022). In terms of target countries, in line with talks with the United Nations, we selected 6 in Africa: Congo Democratic Republic, Lybia, Mali, Somalia, Central African Republic and South Sudan; for computing optimization.

The underlying datasets were a breakthrough in comparison with last year's project. Drawing from AfroGrid studies and their dataset curated against different time-windows, we were able to obtain data from different countries and administrative levels clustered in raster data by 60x60 squared kilometers regions. This, paired up with the fact that AfroGrid already contains variables that are of key importance for multiple conflict prediction (the dataset comes from academic studies), enabled a head start, to which we added the Human Development Index indicators and PrioGrid variables that could help the explainability and prediction objective.

With the data needed at our hands, a case study was developed for Central African Republic. Curating the dataset for avoiding target leakage, and prioritizing data maintenance by a restrictive approach using multicollinearity of 75% and feature importance in a first low-fidelity model, we cleaned our dataset and crafted a random forest regressor, a gradient boosted tree and a Keras Sequential Neural Network. The results for central African Republic enable to examine different feature importance perspectives backed up by -squared metrics above 90%.



From the gradient boosted tree, monthly predictions were extracted for Central African Republic. Using QGIS, we visualized the predictions on a grid of 60x60 squared kilometers, providing a visual representation of the conclusions. This finalized the project, offering a data-driven approach to conflict understanding and prediction for the United Nations.

The last step was a documentation of the project and its main assets, not only for the United Nations, but for next year project, so that we can enable them to go further and deeper with a new approach if decided.