



# Modeling Framework

## Predictive Analytics Pilot

Version 1.0 - February 2020  
United Nation Global Pulse

Unit of analysis	<p>1.1 What is the unit of interest? If people - what populations? Migrants, refugees, IDPs, ethnic groups, all? Will we model individual units or aggregates (groups of units)?</p> <p>1.2 At what geographic level (national, admin1, admin2)?</p> <p>1.3 At what frequency (daily, weekly, monthly, quarterly, yearly)?</p>
Target variable (what we're trying to predict)	<p>2.1 What is the target variable and what is its type? Predict continuous numbers or categories (classification)?</p> <p>2.2 Are we trying to predict the level of the indicator, changes in the indicator and/ or "flags/alerts" for potential changes?</p> <p>2.3 What is the forecast horizon? What is the time scale? How far into the future should we predict?</p>
Feature variables (what we're using to predict)	<p>3.1 Which features (environment, displacement, health, conflict, governance)? Are these features available in a standardized manner across all regions? Is the data longitudinal (over time) or cross-sectional (one point in time)?</p> <p>3.2 How much historical data will we incorporate? Will the data be regularly updated? How?</p> <p>3.3 How will we handle missing data?</p>
Model selection	<p>4.1 What modeling approach will we use? Machine learning, econometrics/statistical forecasting, agent-based models, scenario simulation?</p> <p>4.2 How will models be evaluated? What is "good" or "bad" performance? Which outcome metric will we use? (How to quantify "error"?) Do we need estimates of how confident our predictions are?</p> <p>4.3 How important is model interpretability/explainability?</p>
Applications	<p>5.1 What will be done with the predictions? What action will they trigger?</p> <p>5.2 Will we build the model once, or update it over time?</p> <p>5.3 Will the model incorporate feedback? If so, how and what type of feedback?</p>