

DOPP Exercise 3 - Group 8

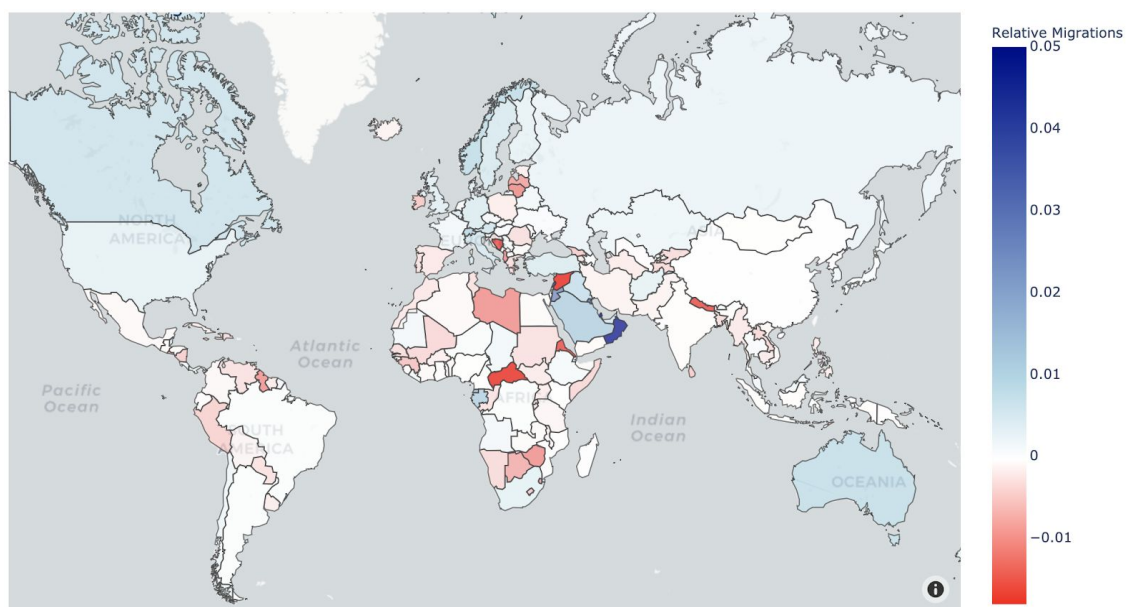
Are there typical characteristics of refugee origin and destination countries, and can countries that will produce large numbers of refugees be predicted?

To predict migration, and especially to prepare for large migration of a country, one first needs to understand which indicators could be a sign for forthcoming migration. Our group investigated different indicators of countries to prevent an unprepared migration, or even intervene early.

We investigated the following datasets:

1. A dataset from the United Nations containing total population numbers (male, female, total) per country and per year.
2. A dataset from the United Nations with the population per age group and gender.
3. A dataset, which contains multiple indicators of a country's population, e.g. the growth rate of a country or the relative migration.
4. One dataset containing the fragile states index, in which the total stability rank of a state and other factors, like the security apparatus of a country or public services are state.

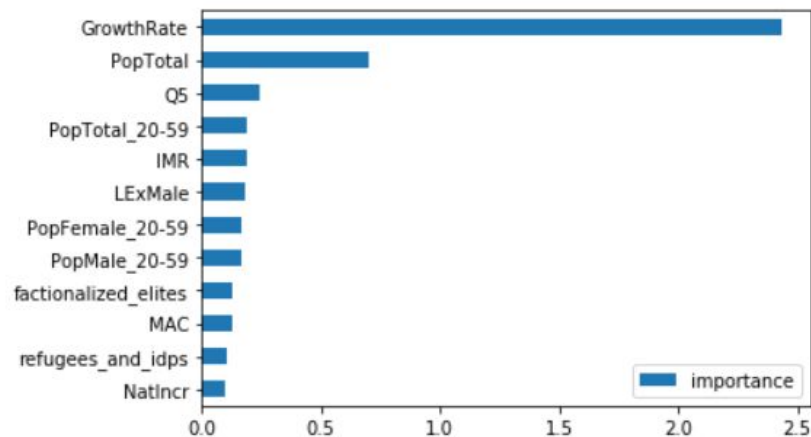
The figure below shows relative migration-rate for countries worldwide, with blue indicating immigration and red indicating migration. This reflects the timespan of 2010-2015.



Points of Interest:

We found that there are many more countries with low migration-rates (lower than 0.5%) than countries with high migration-rate (higher than 0.5%).

By looking into the data, we get on to the most relevant factors of a country to predict the migration-rate. We gathered these insights by training a Machine-Learning Model on the cleaned and transformed datasets.



The figure above shows the most significant factors for predicting the migration-rate of a country. We see in our results, that the important indicators for migration predictions are quite similar to those indicating a developing country. Two of our top ten parameters are factors for child mortality (Q5;IMR). The feature “fractionalized_elites” indicates an unfair distribution of power, which would imply the assumption of dictatorial leadership.

Results

The classification rate of our RidgeClassifier:

	Accuracy
Classification rate for high migration countries	81.3 %
Classification rate for low migration countries	72.3 %

We are able to predict low migration countries with an accuracy of 72.3%, the accuracy is even higher when we predict countries with high migration, with 81.3%.