How well can machine learning tools for humanitarian forecasting be used for predicting the consequences of forced displacement?

Author: Lia Petrova, l.p.petrova-1@student.tudelft.nl
Supervisors: Cynthia Liem, Marijn Roelvink

1. Introduction

- Displacement: widespread, uneven, unfair
- Forecasting: plan funding, mitigate crises
- ML: efficient, comprehensive, enabling
- Data: sparse, biased, unavailable
- Tools: numerous, diverse, disaggregated
- This review: summarises, compares, makes recommendations

2. Research Questions

What predictive ML systems are most effective in terms of countries assisted, accuracy of predictions, and ease of implementation, and under what circumstances they are useful?

- What ML tools for predicting displacement exist/are currently in use?
- How do they work, what can they achieve?
- How well can they be used in mitigating the consequences of displacement?

3. Methodology

- Systematized literature review/SALSA
- Engines: IEEExplore, Scopus, Web of Science, Humdata database, the UK Humanitarian Innovation Hub, and inspiration from Google Scholar and Research Gate.
- Search queries: "machine learning", "displacement", and "forecast" + synonyms. Papers using non-ML methods, predicting or assessing non-forced or non-human migration, or not in English were not considered.
- Comparison metrics: use of data (categories, sources, variables, missing data), algorithms and accuracy (+ benchmarks), operational usability
- Results: After removing irrelevant results, filtering by abstract and title, and deduplication, the final set of papers corresponding to the search was 113, of which 19 were included in the review.

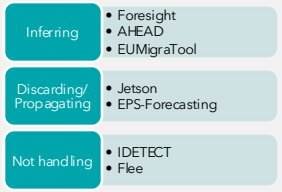
4. Results

A.

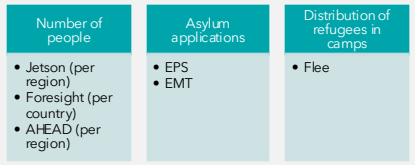
Data Category	Jetson	Foresight	EPS	IDETECT	AHEAD	EMT	Flee
Internet Search Keywords			✓	✓		✓	
Diseases and Healthcare	✓				✓	✓	
Socio-demographics		✓				✓	✓
Food Security		✓			✓	✓	
Political Events/Governance		✓	✓			✓	
Violence/Conflict	✓	✓	✓	✓	✓	✓	✓
Economy	✓	✓	✓			✓	
Travelling Distances	✓						✓
Natural Disasters/Climate	✓	✓		✓	✓	✓	
Social Unrest			✓				

B.

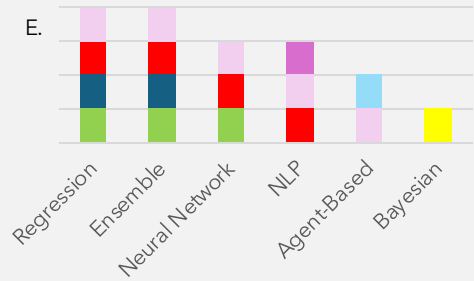
C.



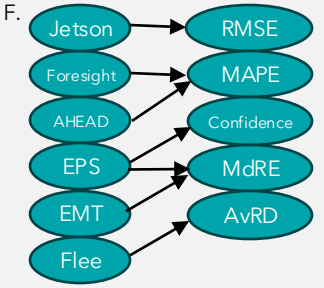
D.



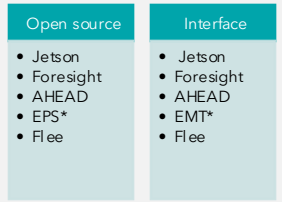
E.



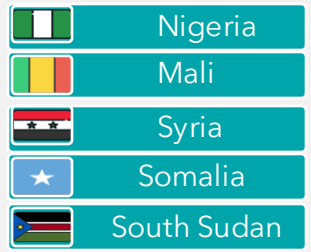
F.



G.

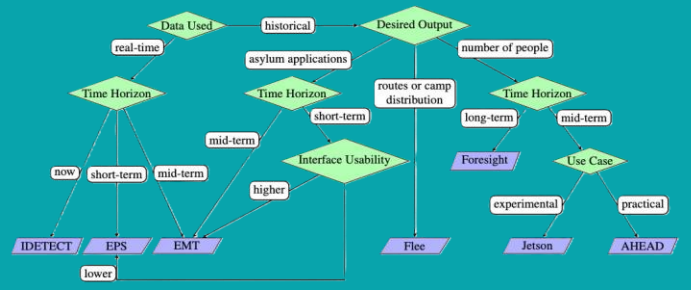


H.



5. Discussion

- Short-/mid-term forecasting: data that captures sudden triggers and is updated often;
- Long-term predictions: data of structural slow changes that need longer time to monitor and update;



6. Limitations and Future Work

- Bigger span and broader scope of the review
- Experimental testing of accuracy and reproducibility
- Ethical considerations and viability of recommendations

A. Data categories per model
B. Time horizon of predictions
C. Handling missing data
D. Dependent variable
E. ML algorithms used
F. Error rates
G. Accessibility
H. Most common countries of analysis

TU Delft