Data Science/Machine Learning Advisor Technical Assessment

Jeff Wagg
February 22, 2024

Task 2 – Critical Evaluation

We consider the impact of a predictive model able to quantify the proliferation of yellow fever in Brazil. Although promising, caution should be exercised before blindly applying this model in other countries, or regions. Yellow fever is typically carried by monkeys, and the vector which transmits the virus is the Aedes mosquito whose bite will infect humans travelling outside of urban environments (see ECDC facts). The Aedes mosquito is also known to carry the dengue virus, and generally prefers warmer climates (20°C to 30°C, Liu et al. 2023, Front Cell Infect Microbiol., 13:1242173). As such, the model is unlikely to be applicable in cooler environments, or regions where the prevalence of monkeys differs greatly from that of Brazil.

We recommend obtaining a sample of model feature variable data and proliferation rates from the new environment going back at least one year to ensure that all seasonal variations are covered. It is also important to include both urban and non-urban data in the model. If the model applied to these test data exhibits a high accuracy and precision, then it could be used in the new region. Otherwise, the model fit parameters should be recalculated using sample data from this area.

Task 3 – Strategic Vision

The last decade has witnessed explosive growth in the use and effectiveness of artificial intelligence (AI) tools. Advances in both the hardware and machine learning (ML) algorithms have led to the widespread application of AI in nearly every aspect of life. The potential exists to dramatically improve the way in which aid and education are delivered to children. A great example of this is Mtabe[1], a Tanzanian ed-tech start-up that is using AI to personalize the learning journey for individual students.

Along with the positive benefits of using AI for children, one must also be aware of the risks[2]. Great care needs to be taken when using humanitarian datasets in order to ensure the sovereignty of the data and preserve the anonymity of vulnerable populations like children (see RD4C report[3]). Navigating these risks requires clear guidelines and principles for the use of AI models applied to humanitarian data, which have been developed by the European Commission[4] and also by UNICEF[5] to protect the safety of children.

In support of the UNICEF 2022-2025 strategic goals[6] of ensuring that every child acquires the skills required for the future (Goal 2), and that every child has an equitable chance in life (Goal 5), we recommend the development and implementation of new data science and AI tools. The process should leverage existing partnerships with organisations like Google Inc. and Meta, as

[1] https://mtabeapp.com/
[2] Spencer, S. W., 2021, *Humanitarian AI: The Hype, the hope and the future,* Network Paper 85
[3] https://rd4c.org/principles/
[4] https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines.1.html
[5] https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children
[6] https://www.unicef.org/media/115646/file/Strategic%20Plan%202022-2025%20publication%20English.pdf

well as academia, who have the long term resources available to develop and test new solutions. We also see tremendous benefit in engaging with, and funding AI start-ups from the regions where UNICEF operates, especially those in the Global South which has traditionally been under-represented in AI policy discussions.