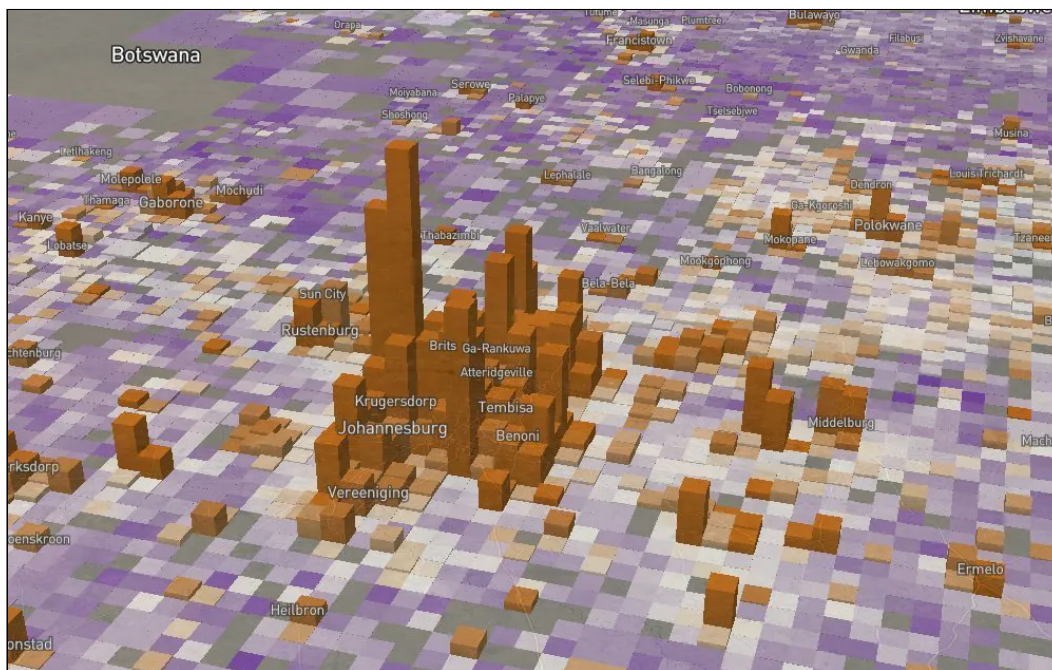


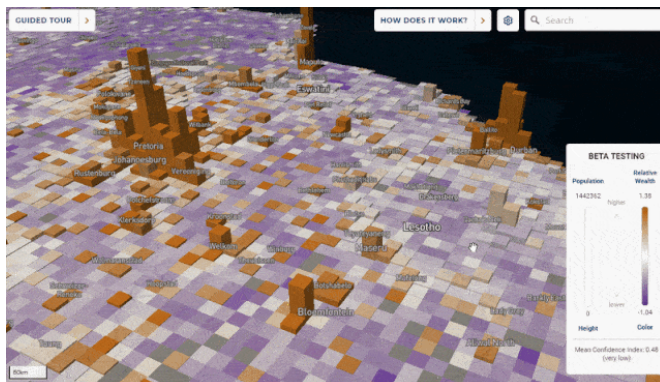
Estimating the True State of Global Poverty With Machine Learning

Originally published at <https://www.unite.ai/estimating-the-true-state-of-global-poverty-with-machine-learning/>

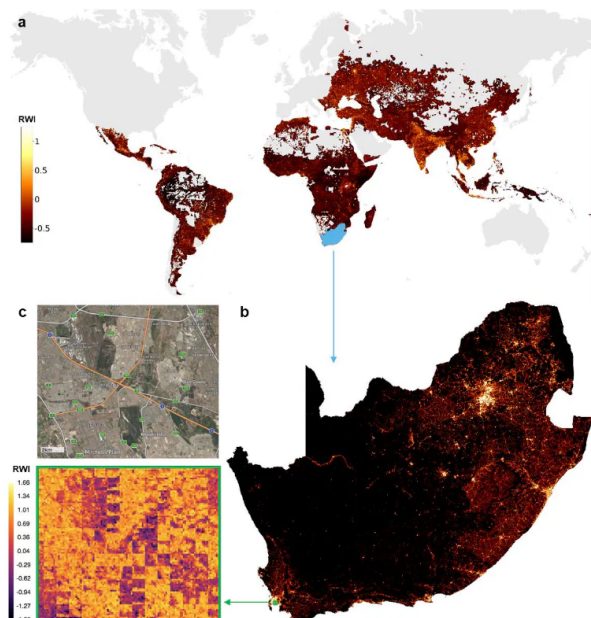


A collaboration from UoC Berkeley, Stanford University and Facebook offers a deeper and more granular picture of the actual state of poverty in and across nations, through the use of machine learning.

The [research](#), entitled *Micro-Estimates of Wealth for all Low-and Middle-Income Countries*, is accompanied by a [beta website](#) that allows users to interactively explore the absolute and relative economic state of fine-grained areas and pockets of poverty in low and middle-income countries.



The framework incorporates data from satellite imagery, topographic maps, mobile phone networks and aggregated anonymized data from Facebook, and is verified against extensive face-to-face surveys, for purposes of reporting relative wealth disparity in a region, rather than absolute estimates of income.



A map of global poverty, weighted towards the most affected areas. Lower, enlargements of South Africa and Lesotho (b); a 12sqkm area around the Khayelitsha township near Cape Town. Source: <https://arxiv.org/ftp/arxiv/papers/2104/2104.07761.pdf>

The system has been [adopted](#) by the government of Nigeria as a basis for administering social protection programs, and runs in tandem with the existing framework from the World Bank, the National Social Safety nets Project ([NASSP](#)). In February the first recipients under the scheme were paid a cash transfer of 5000 Nigerian naira, a benefit payable up to six months, until a threshold of one million naira is reached.

The paper asserts that [data poverty](#) makes a notable contribution to incorrect distribution of aid in countries with minimal data-gathering resources or limited infrastructure, and that politically motivated misreporting (a problem [not limited](#) to low-income countries) is also a factor in this regard.

Registering The ‘Unreported Poor’

The researchers’ simulations on the data demonstrated that, under existing regulations for allocation of aid resources, distribution based on this system increases payment to those most in need and decreases payment for existing recipients in higher-income brackets. The paper also notes the difficulty that social protection program administrators faced in allocating aid resources at the outset of the COVID-19 crisis, due to the lack of comprehensive or detailed data. In Nigeria, for example, the most recent survey data covers households in a mere 13.8% of all Nigerian districts, compared to the 100% coverage that the new scheme affords.

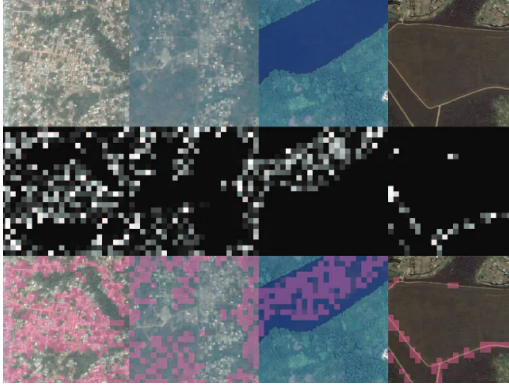
Previous work in AI-assisted poverty estimation has concentrated largely on satellite-obtained data (see below), but the researchers assert that data from mobile connectivity obtains a more accurate and granular insight into wealth disparity across regions, and this information stream provides half of all the contributing data for the project.

From the standpoint of generalization in machine learning data, the researchers observe that models trained in one country can be a beneficial and accurate template for models covering adjacent countries. They also note that the new framework is able not only to distinguish between urban and rural locales, but is capable of providing disparity maps within urbanized areas, which exceeds the scope of many recent research initiatives in this sector.

Satellite Imagery In Poverty Analysis

The principal behind satellite-based poverty analysis is the presumption that poor people have little money to run electric lights in the hour of darkness, or may have no electric light facility at all. Where the absence of pinpoint lights can be correlated to the presence of people, as ascertained by other means (such as mobile connectivity data), an index of deprivation can be generated.

This technique was proposed in 2016 in an [earlier Stanford paper](#) from another research group. The method detailed in that paper pioneered the use of evening-time satellite coverage provided by the United States Air Force Defense Meteorological Satellite Program (DMSP) via the National Oceanic and Atmospheric Administration (NOAA-NGDC).



Four convolutional filters identify, left to right, features that relate to urban zones, rural zones, water and roads. The top row shows source images from Google Maps, the middle the filter activation maps from machine learning analysis, and the bottom row an overlay of activation maps on the original map imagery.

Source: <https://science.sciencemag.org/content/sci/353/6301/790.full.pdf>

The Stanford project correlated the filtered evidence of night-lights in the satellite surveillance against its own database of DHS surveys for the year that both the surveys and the aggregated satellite results occurred. It was necessary to [establish](#) averages of the sum of night-light values as proxies for certain economic indicators.

Ground Truth For Global Poverty Statistics

For the new Stanford project, the researchers decided to derive the data framework from the existing Demographic and Health Survey ([DHS](#)) Program, even though, as they concede, this effectively replicates the DHS schema in the dataset. The researchers observe: ‘We elected to train our model exclusively on DHS data because it is the most comprehensive single source of publicly available, internationally standardized wealth data that provides household-level wealth estimates with sub-regional geo-markers.’

However, the project operates at a far higher resolution than DHS, and using the existing framework as ground truth provides two benefits: firstly, the DHS data does not rely on formal reporting of income, which is an unreliable indicator in the countries most affected by poverty, where black market economies are rife; and secondly, the data is collected in a standardized manner and to an international template that allows the researchers’ framework to encompass other countries that are subject to this method of measurement, rather than establishing equivalencies across competing frameworks.

Mobile Connectivity As An Economic Index

For people living in economically challenged areas, mobile connectivity has become a technological lifeline over the last two decades, since mobile phones are the minimum available technological platform that can be relied upon in such conditions. Mobile phones have also become *de facto* payment platforms for aid recipients who lack a bank account, or other conventional means of receiving money.

However, as has been [observed before](#), using mobile network indicators as an economic index for machine learning systems has some potential drawbacks: there are people in the affected regions that are so poor as to not even possess a mobile phone – the very people that the system is most designed to help; the system could potentially be gamed by users with multiple mobile phones in circumstances where a phone has become a proxy for unique ID hashes of citizens; and there are privacy implications to creating this kind of identification system, in cases where the local or national government retains some oversight of the project.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG’s Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



[Discover More](#)