Shedding Light on Philippine Poverty

Mapping Poverty using Machine Learning, Satellite Images, and Open Geospatial Data

Isabelle Tingzon | February 17, 2019

ABOUT US

We use **data** to solve real-world problems

Manila | Singapore

Founded 2015

PUBLIC SECTOR

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

UNDP

Office of the President of the Philippines
National Anti-Poverty Commission

ADB ASIAN DEVELOPMENT BANK

Save the Children

MMDA

Teach For All

PH Department of Science and Technology

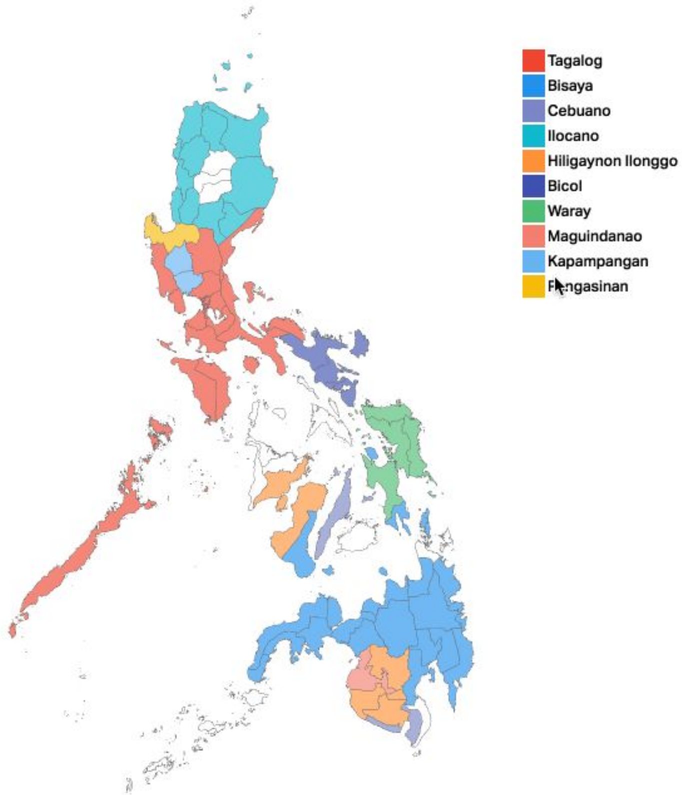
University of the Philippines

OUR WORK

Geospatial Analysis and Visualization

90% of Filipinos speak one of 10 languages

While there are nearly 200 unique languages and dialects spoken by the Philippines' nearly 100 million residents, over 90 percent of Filipino households speak one of just 10 languages. This map shows the provinces where these top 10 languages are the most widely-spoken.



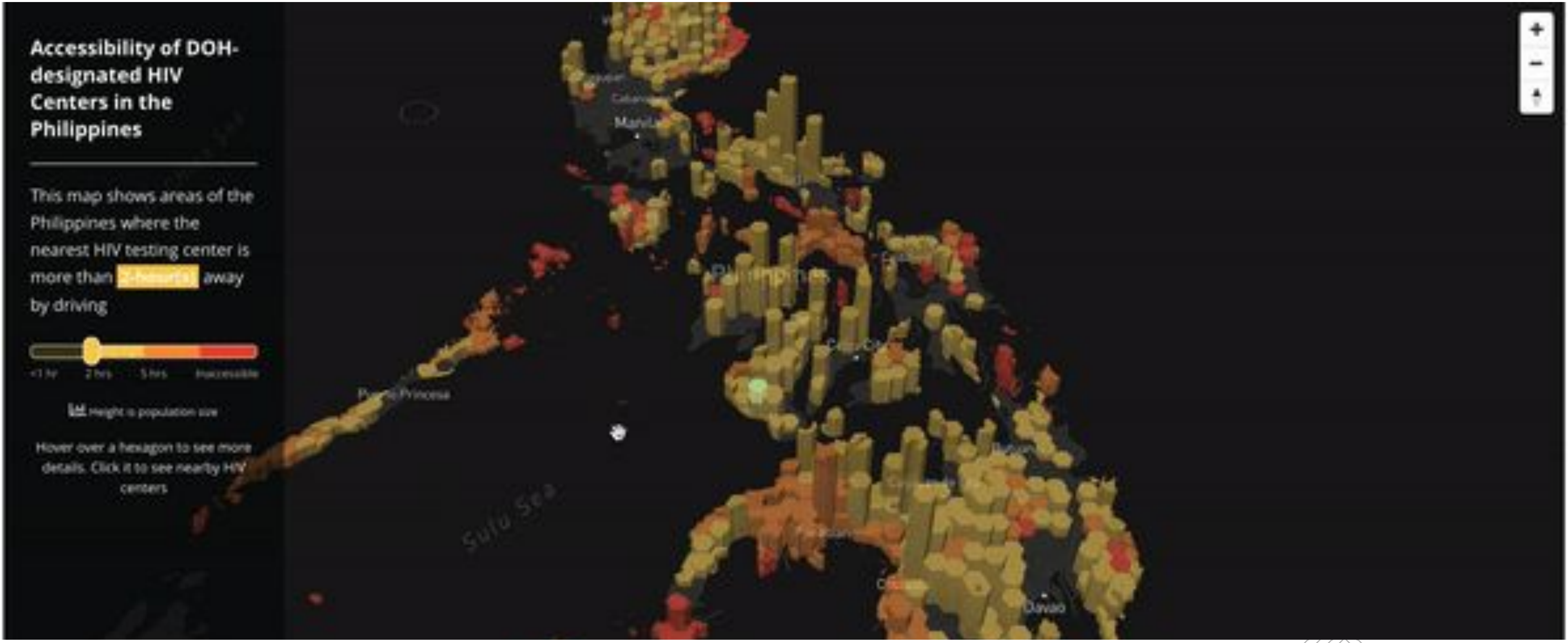
Language diversity across the country

But just looking at the most widely-spoken language in a place can obscure its linguistic richness.

Example: The Language Landscape of the Philippines in 4 Maps

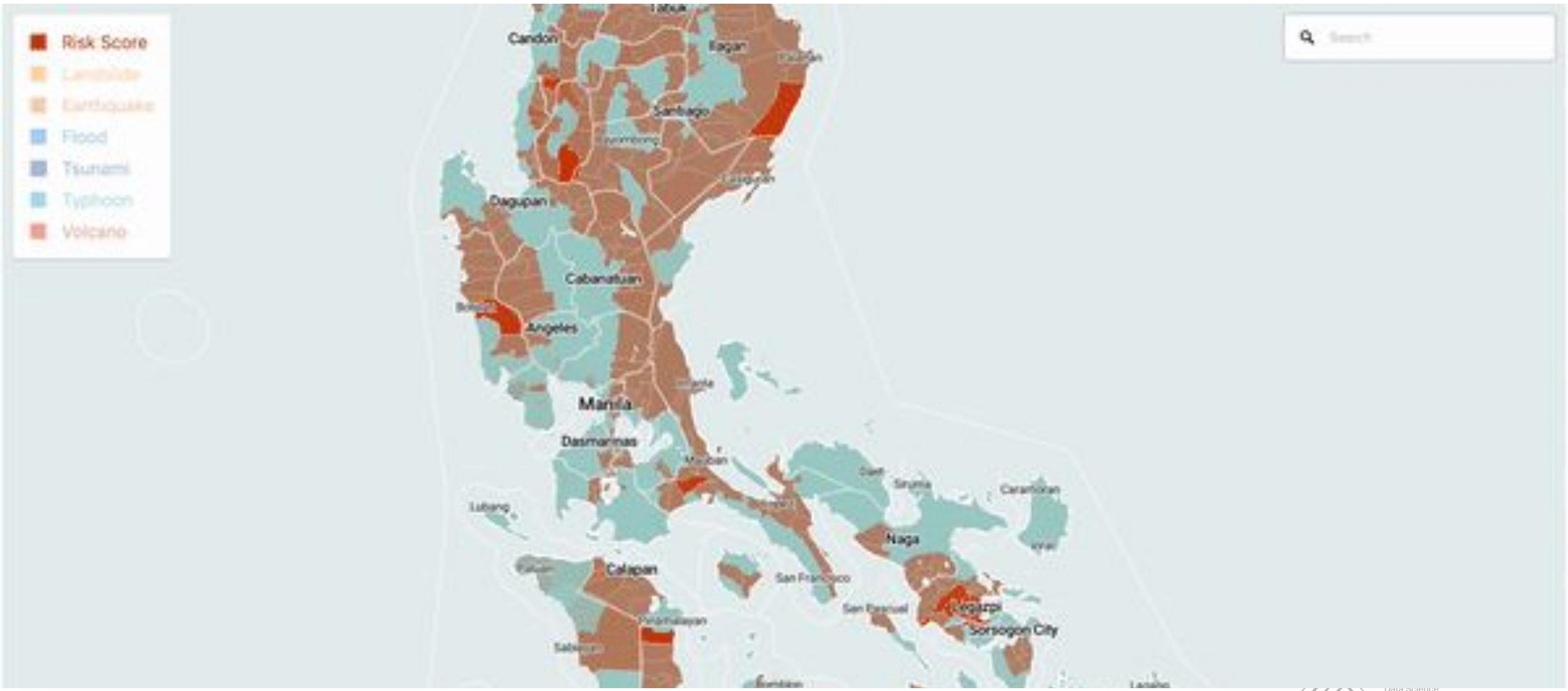
Thinking Machines
Data Science

Map of HIV Center Accessibility across the Philippines



OUR WORK

Hazard Exposure Map of the Philippines



OUR MISSION

Our social impact mission is to empower evidence-based policy and action by:

- Filling critical data gaps
- Making data open and useful
- Innovating with purpose

unicef INNOVATION

The UNICEF Innovation Fund provides financial and technological support for companies that are using technology in innovative ways to improve children's lives.

$ 4.1 M INVESTED | 58 PROJECTS | 38 COUNTRIES

Thinking Machines Data Science

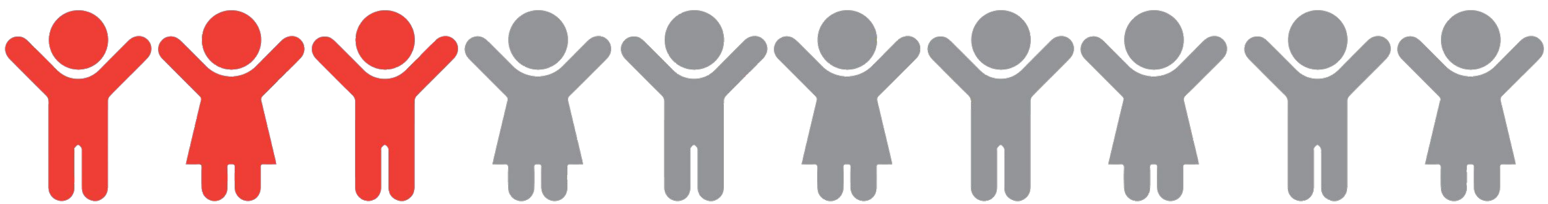
MOTIVATION

22M or 21.6% of Filipinos live below the national poverty line.

***Sources**: Philippine Statistical Authority 2015 Survey, National Economic Development Authority*

MOTIVATION

Child Poverty in the Philippines



3 in every **10 children** belong to poor families

Thinking Machines
Data Science

Source: Official Poverty Statistics among the Basic Sectors 2015, PSA

MOTIVATION

Conducting Household Surveys

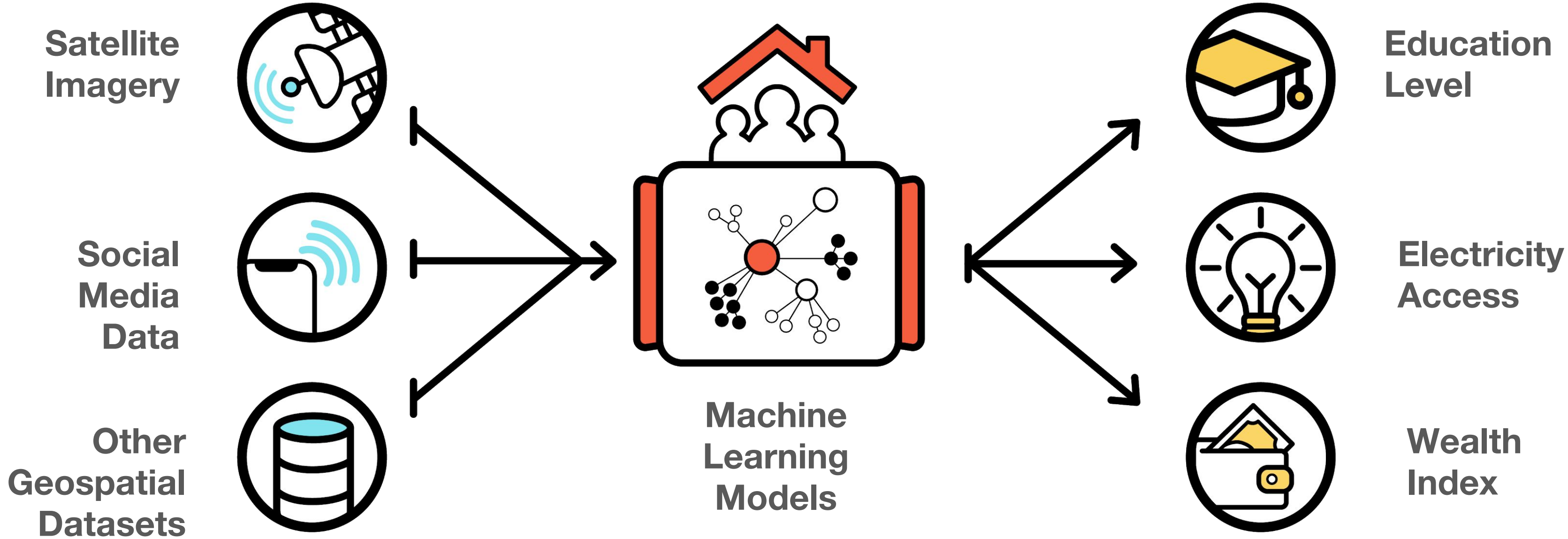
Can cost
millions of dollars

Conducted only
once every 3-5 years

Granularity is on a
regional/provincial level

SOLUTIONS

Use unconventional, readily-available data sources to predict socioeconomic indicators



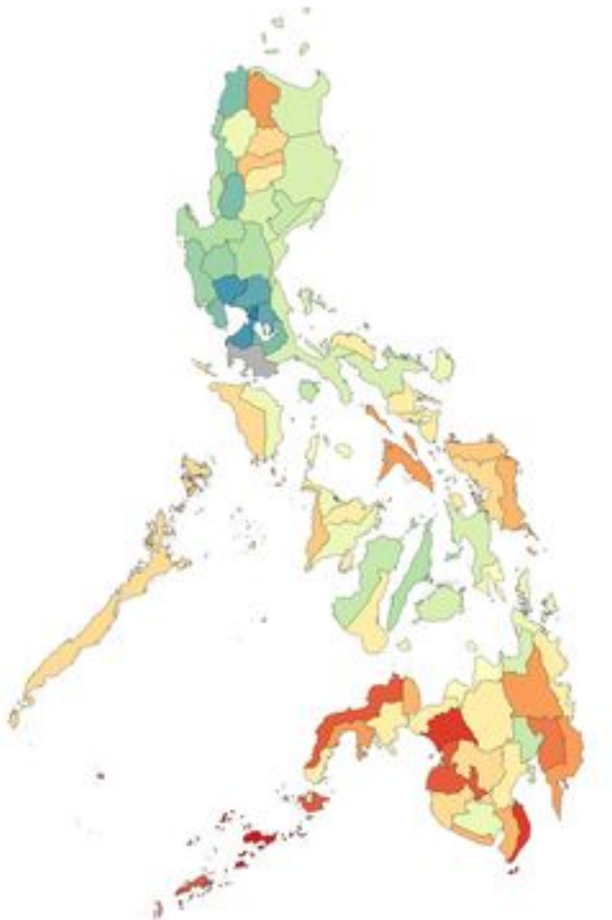
Thinking Machines
Data Science

SOLUTION

Machine Learning for Wealth & Poverty Prediction

- **Goal:** ***Faster***, ***cheaper***, and ***more granular*** reconstruction of poverty measures in the Philippines
- Replicated a study by **Jean et al.*** from the Stanford Sustainability and AI Lab
  - Estimated asset-based wealth for five sub-Saharan African countries
- Crowdsourced geospatial information for poverty prediction

Provincial-level Wealth Index
(Ground truth)



*Jean, Neal, et al. "Combining satellite imagery and machine learning to predict poverty." *Science* 353.6301 (2016): 790-794.

APPROACH

Using satellite images to predict wealth

Problem: Data Scarcity

Need a lot of labeled training data for an end-to-end deep learning approach

Model

Poverty Predictions

Wealth Index, Education, Access to Water, Electricity, Child Mortality

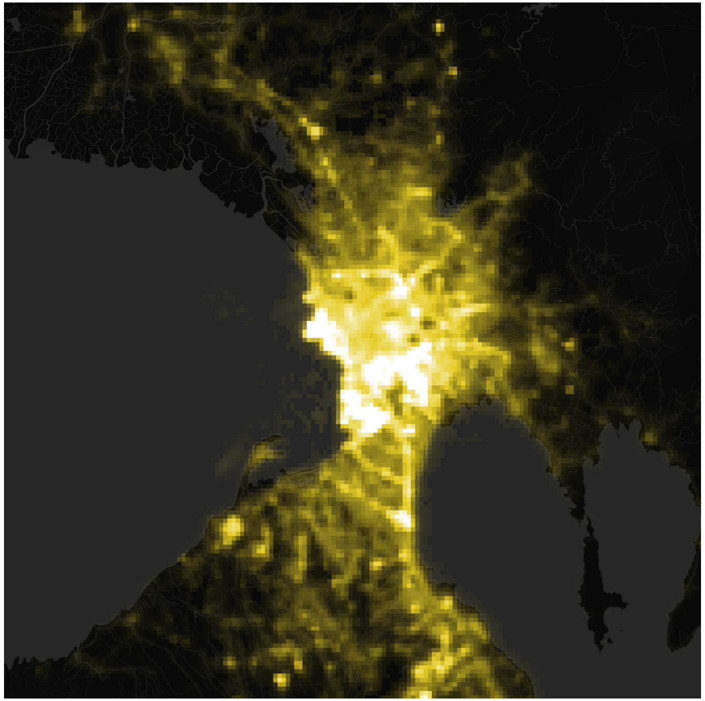
Data Gap: Not enough labeled training data!

SOLUTION

Nighttime lights can be used as a proxy for economic development

Metro Manila (Daytime)

Metro Manila (Nighttime)



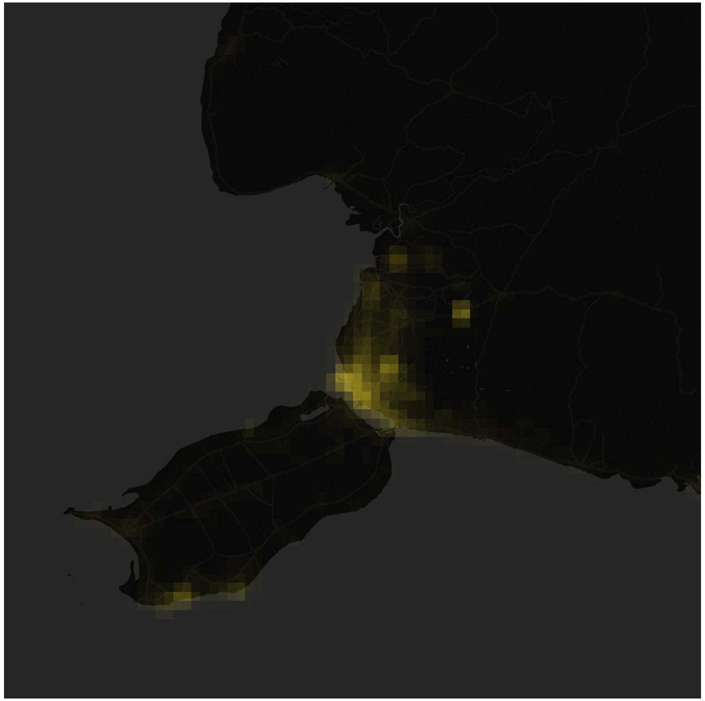
Thinking Machines
Data Science

SOLUTION

Nighttime lights can be used as a proxy for economic development

Tagbilaran, Bohol (Daytime)

Tagbilaran, Bohol (Nighttime)

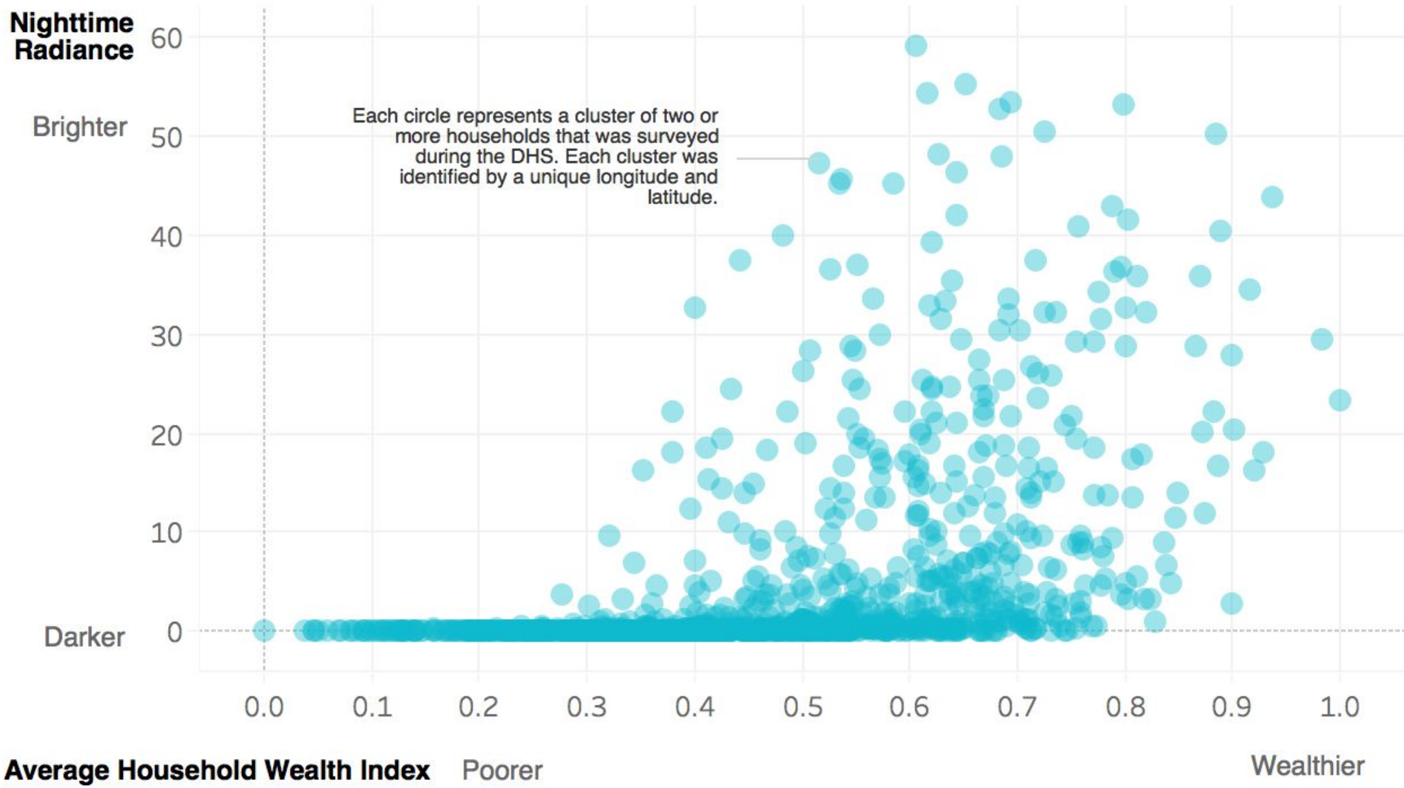


Thinking Machines
Data Science

SOLUTION

Wealthier places are brighter at night

For each of the 1,200+ sampled locations or "clusters" in the 2017 Demographic and Health Survey (DHS), there was a positive correlation between nightlight luminosity and average household wealth index. *(p=0.75, r=0.49)*

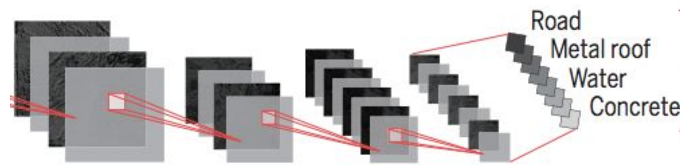


Thinking Machines Data Science

SOLUTION

Predict nighttime light intensity as a proxy task

Daytime Satellite Images



Convolutional Neural Network

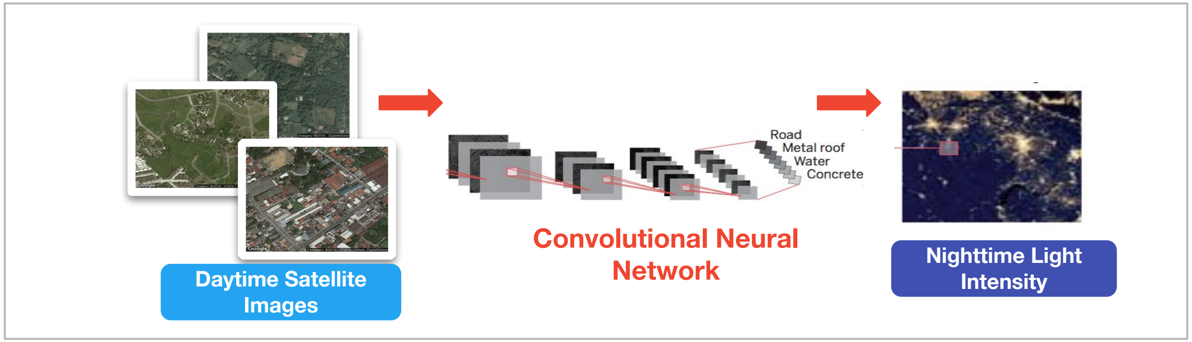
Nighttime Light Intensity

Low, Medium, High
based on pixel brightness

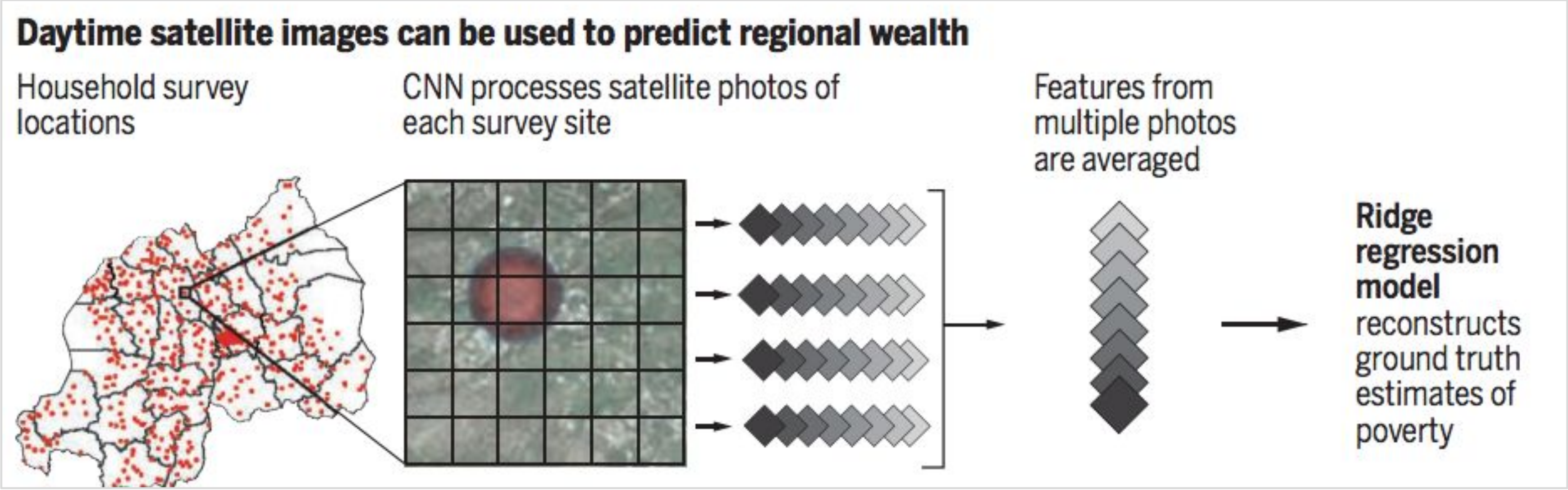
*Jean, Neal, et al. "Combining satellite imagery and machine learning to predict poverty." *Science* 353.6301 (2016): 790-794.

SOLUTION

Step 1: Predict Nighttime light intensity from Daytime Satellite Images



Step 2: Wealth Prediction using Cluster-level Feature Embeddings

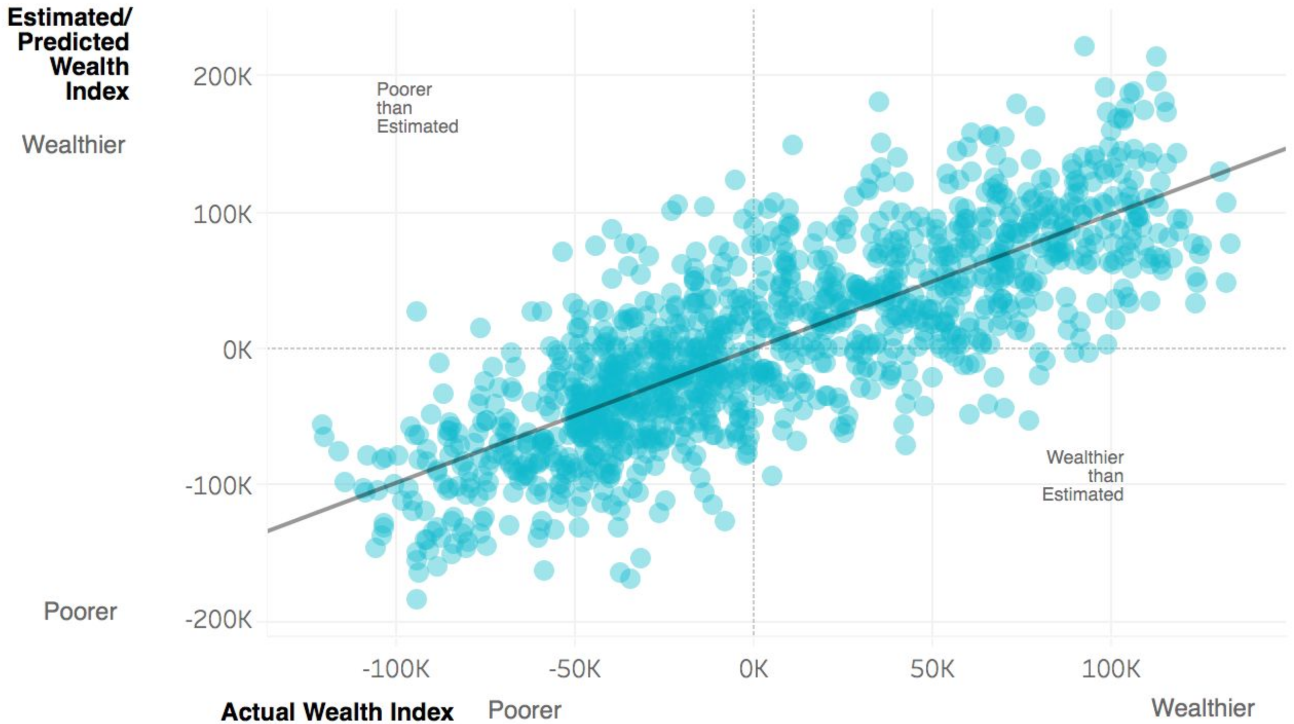


*Jean, Neal, et al. "Combining satellite imagery and machine learning to predict poverty." *Science* 353.6301 (2016): 790-794.

RESULTS

How accurate are our wealth predictions?

The chart below compares the actual versus estimated or predicted average household wealth index for each of the 1,200+ clusters surveyed in the 2017 Demographic and Health Survey. *(r2=0.625)*



The model is able to explain **62.5%** of the variance

*Using CNN feature embeddings with regional indicators

Thinking Machines Data Science

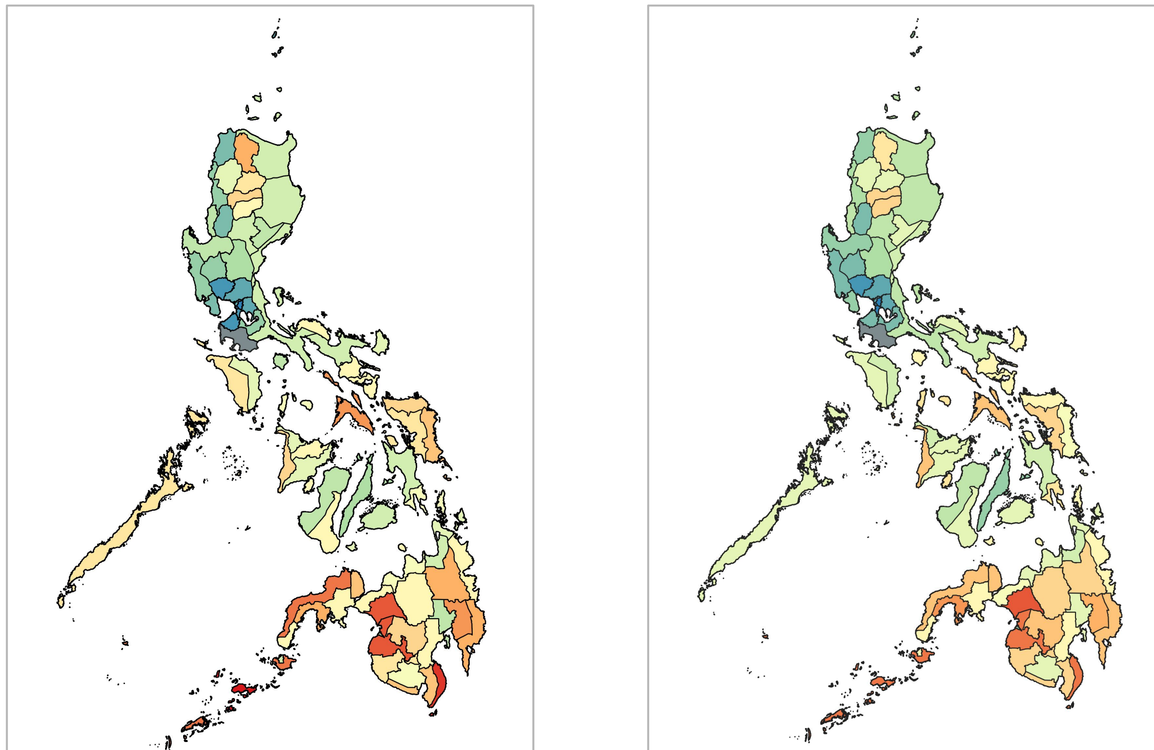
Predictions and reported r^2 values are from five-fold nested cross-validation.

Reconstructing Provincial-level Maps

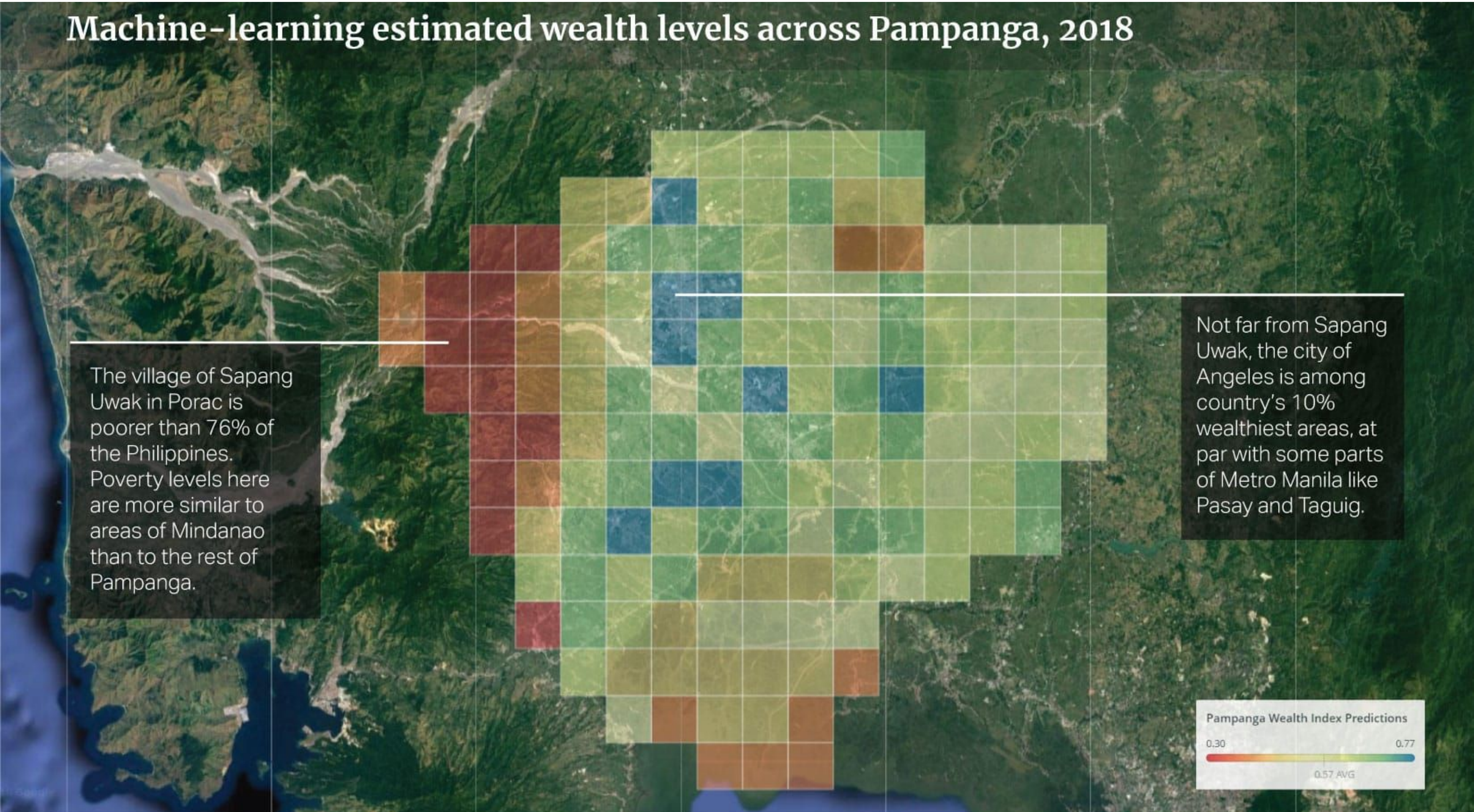
Normalized Wealth Index

0 (poor) 1 (wealthy)

Actual Wealth Index | Predicted Wealth Index



Predictions and reported r^2 values are from five-fold nested cross-validation.


Machine-learning estimated wealth levels across Pampanga, 2018
The village of Sapang Uwak in Porac is poorer than 76% of the Philippines. Poverty levels here are more similar to areas of Mindanao than to the rest of Pampanga.
Not far from Sapang Uwak, the city of Angeles is among country's 10% wealthiest areas, at par with some parts of Metro Manila like Pasay and Taguig.
Pampanga Wealth Index Predictions
0.30
0.77
0.57 AVG

CAVEATS

Limitations

- Difficult to distinguish between different levels of extreme poverty for areas with little to no electrification
- Difficult to predict other aspects of human development (e.g. access to water, child mortality) using nighttime lights
- Poverty estimates are not meant to replace on-the-ground surveys.

NEXT STEPS

Moving Forward

- **Research Collaboration**: Partner with research institutes on exploring alternative methodologies and data sources for poverty prediction:
    - Using social media, traffic data, and other geospatial data sources
    - Explore unsupervised representation learning methods
- **Data Sharing**: Model additional indicators including
    - Socioeconomic resilience to natural disasters
    - Child malnutrition rates
    - *Challenge: Lack of reliable granular ground-truth data*
- **Real-world Impact**: Collaborate with the public sector (NGOs, LGUs) for possible applications in humanitarian aid and resource allocation, both in the Philippines and other countries

Thinking Machines
Data Science

Special Thanks

Dohyoung Kim (UNICEF)
Vedran Sekara (UNICEF)
Priscilla Moraes (Google)
Ingmar Weber (QCRI)
Masoomali Fatehkia (QCRI)
Neal Jean (Stanford Sustainability & AI Lab)
Sherrie Wang (Stanford Sustainability & AI Lab)

Thank you!

issa@thinkingmachin.es

Thinking Machines

✦ **Read more data stories on our blog**
stories.thinkingmachin.es

✦ **Follow Us**
/thinkdatasci
@thinkdatasci

References

- Philippine Statistical Authority (PSA) 2015 Survey
- National Economic Development Authority (NEDA)
- Jean, Neal, et al. "Combining satellite imagery and machine learning to predict poverty." *Science* 353.6301 (2016): 790-794.
- Xie, Michael, et al. "Transfer learning from deep features for remote sensing and poverty mapping." *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. AAAI Press, 2016.
- Head, Andrew, et al. "Can Human Development be Measured with Satellite Imagery?." Proceedings of the Ninth International Conference on Information and Communication Technologies and Development. ACM, 2017.
- Perez, Anthony, et al. "Poverty Prediction with Public Landsat 7 Satellite Imagery and Machine Learning." *arXiv preprint arXiv:1711.03654* (2017).

Thinking Machines
Data Science