Carnegie Mellon University Africa
Certificate I: Understanding AI and Machine Learning

Course AIML02: AI and Machine Learning in Africa

Module 2: Application Case Studies
Lecture 5: Socioeconomics

Welcome to Lecture 4 of Module 2. We continue to look at applications of AI and machine learning in an African context, focusing here on a case study in socioeconomics.

We begin by introducing the case study target article by Yeh et al. (2020) and its significance for socioeconomic application.

We will explain the importance of geographical information to human well-being and how it influences economic decisions.

We then explore different options for acquiring data to generate economic indicators, focusing in particular on the approach used by Yeh et al. (2020), i.e., deep learning with satellite images.

We finish up by summarizing what we have covered and identifying the articles that you should read to consolidate what you have learned.

We have three learning objectives, so that, after studying the material covered in this lecture, you should be able to do the following.

1. Understand one role of AI and machine learning in socioeconomics in Africa.

2. Explain the use of deep learning with satellite images to infer the level of regional and local poverty, particularly in rural areas.

3. Understand how to evaluate AI and machine learning strategies, technologies, and solutions.

Slide 1 Welcome to Lecture 5 of Module 2. We continue to look at applications of AI and machine learning in an African context, focusing here on a case study in socioeconomics.

Slide 2 Socioeconomics deals with the relationship between economic activities and social processes and how they shape one another.

This case study is based on an article written in 2020 by Christopher Yeh and seven co-authors, entitled: “Using publicly available satellite imagery and deep learning to understand economic well-being in Africa”.

It demonstrates the application of machine learning, specifically deep learning, in predicting asset wealth across approximately 20,000 villages in Africa, using satellite images as training data.

Slide 3 For stakeholder poverty intervention programs from government agencies, non-government agencies (NGOs), and private sector organizations to be effective,

there is a need for up-to-date economic indicators of the level of human well-being at local and regional levels across a country.

Slide 4 Asset wealth surveys provide information on key economic indicators.

However, they are carried out infrequently, as the graphic illustration shows.

Most African countries carry out surveys every four or more years.

This makes it difficult to accurately measure variations in well-being over time.

The low frequency is mainly due to the cost of carrying out such surveys, requiring an estimated annual investment of $1 billion dollars.

Slide 5 The challenge, therefore, is to acquire the geographical information on economic indicators and human well-being in a timely and cost-effective manner.

The target article addresses this challenge by using data from satellite imagery to provide information on

local economic and well-being indicators across geographical regions as well as measurements of change in well-being over time.

This information is important when taking economic decisions on interventions to alleviate poverty.

Slide 4 Two approaches have previously been used for obtaining data to generate economic indicators.

These are traditional data collection methods and the use of machine learning using satellite imagery.

Slide 7 Traditional data collections methods involve the use of surveys and questionnaires such as the Demographic and Health Surveys (or DHS, for short).

The DHS is used to estimate the economic status of households from responses to questions such as such as whether the home has electricity, water supply, toilet, the quality of house floors, and ownership of a phone, radio, TV, car, or motorbike.

Slide 8 Unfortunately, such surveys are carried out infrequently, without necessarily revisiting the same individual locations.

For instance, a given household may feature in one survey but not in previous or subsequent surveys.

This makes it difficult to track economic changes and assess the impact of intervention programs over a period of several years.

Surveys are also expensive to conduct.

Slide 9 To overcome the problems associated with traditional data collection methods, alternative approaches have been explored to determine economic economic indicators.

For example, coarse 1 km/pixel nighttime light imagery has been used to train deep learning models to predict the asset wealth index.

Slide 10 Models can also be trained using high-resolution daytime images.

In this example, the satellite images have a resolution of less than one metre per pixel.

Slide 11 Transfer learning has also been used. We mentioned transfer learning in AIML01, Module 3, Lecture 1. It refers to the practice of first training a model, typically a deep neural network model, using a large general purpose dataset and then tuning the trained model using a smaller application-specific dataset.

In this case, nighttime lights are used as intermediate labels for training a convolutional neural network (CNN) feature extractor to find features relevant to predicting wealth in multi-spectral Landsat imagery.

We introduced convolutional neural networks in the micro course AIML01, Module 2, Lecture 2: AIML01-02-02.

Slide 12 The approach explored in this case study focuses on using cloud-free imagery from multiple satellite-based sensors, specifically nighttime and daytime imagery,

to model both spatial and temporal differences in economic well-being at a local level across sub-Saharan Africa.

The authors demonstrate that their results could be used to help target social programs and understand the factors that influence well-being across the developing world, including Africa.

To accomplish this, they use a deep CNN to predict asset-wealth distribution across villages and districts in countries without reliable survey data.

Slide 13 Although there are other welfare measurement indices, the target article uses the asset-wealth index which is also commonly used by development practitioners around the world to target social programs.

It is considered a less noisy indicator of longer-run economic well-being,

It is widely found in publicly available geo-referenced African survey data.

And it is also a common component of the multi-dimensional poverty index.

Slide 14 Model evaluation was accomplished using an out-of-sample technique. This just means the model was evaluated using data that was not used during training.

This is done to avoid overfitting, a problem that occurs when a model performs well on the training data but not on new unseen data. In such cases, we say the model generalizes poorly.

This ability to generalize makes the model robust in practical settings where predictions need to be made in the absence of ground-truth survey data, as is the case in a lot of African countries.

Slide 15 Let's turn our attention now to the way in which this system was implemented.

In Step 1, survey responses were obtained from forty-three DHS surveys carried out across twenty-three African countries between 2009 to 2016.

To protect the privacy of the surveyed households, DHS randomly displaces the coordinates reported in surveys by 2 km for urban clusters and 10 km for rural clusters.

Slide 16 In Step 2, the data was cleaned by removing invalid coordinates which might have occurred during the random displacement of coordinates.

Slide 17 In Step 3, the data was processed by converting the responses from each household into an asset quality score between a range of 1 to 5.

Slide 18 The asset-wealth index was then constructed in Step 4, using the first principal component of the survey responses,

adjusting it to so that it has a normal distribution with a mean of 0 and standard deviation of 1.

(As an aside, the concept of principal components refers to a mathematical technique that identifies and ranks the most important factors that contribute to some multi-factorial set of measurements, signal, or image. The first principal component, then, is the most important factor.)

Slide 19 Data validation is performed in Step 5 to avoid the bias of subjective and country-specific questions such as housing quality scores.

For example, a house considered a 5 in some countries might be a 3 in another country.

This validation is accomplished by comparing a pooled principal component analysis index across countries to an index derived only from the sum of all assets owned like phones, radio, TV, car, or motorbike,

(As another aside, principal component analysis, or PCA for short, is the mathematical technique that is used to identify the principal components.)

The comparison showed a high correlation between both indices with r^2 value between 0.80 and 0.98.

This implies that the index obtained from the PCA is not significantly affected by country-specific factors.

(As a final aside, The r^2 value is the coefficient of determination: a measure of how well observed outcomes are replicated by a model. Higher values indicate that the model is good at predicting the phenomenon being modelled. The coefficient of determination normally ranges from 0 to 1.)

Slide 20 The final training data is generated in Step 6. This data contains responses and asset-wealth indices from over 500,000 households spread across 19,669 African villages.

Slide 21 Additional validation was carried out to the second administrative level by comparing the asset wealth index obtained from the DHS surveys to two external datasets:

First, the country's census data and

Second, the Living Standards Measurement Study (LSMS) which is conducted by the World Bank.

The country census data focused on countries that have asset ownership questions in their census report and had carried out census within 4 years of the DHS survey year.

The comparison showed high correlation between the DHS surveys and both external datasets with r^2 values of 0.99 and 0.97 for the census and LSMS survey, respectively.

Slide 22 Two models were trained to establish baseline models.

First, a simple k-nearest neighbor model (labelled KNN scalar NL) on nightlights images. KNN is a simple machine unsupervised learning technique.

The parameter k is tuned using cross-validation.

The second model is a regularized linear regression on scalar nightlights (labelled Linear scalar NL).

Linear regression is a procedure by which we model a process that relates an independent variable, e.g., time or location, and a dependent variable, e.g., wealth, by fitting a line to the data that characterize the process. This model then allows one to predict values of the dependent variable based on the value of the independent variable, e.g., the wealth value at some required time or place.

Slide 23 To train the model, temporally and spatially matched multi-spectral daytime images from 30 m/pixel Landsat and less than 1 km/pixel nightlight images were used as inputs to train the CNN with ResNet-18 architecture.

The ResNet-18 architecture is trained independently on nighttime light intensity imagery and on multispectral daytime imagery

…and then concatenated in the final fully connected layer as shown in image.

This allows the network to learn the features in both the daytime and nighttime imagery that are predictive of asset-wealth distribution.

Slide 24 The models were trained with an Adam optimizer, a mean squared-error loss function with a batch size of 64.

Training was done for 150 epochs with in-country data and 200 epochs for DHS out-of-country data.

The learning rate is decayed by a factor of 0.96 after every epoch and grid search is used to determine the best hyper parameters for optimal learning rate and weight regularization.

Finally, geometric augmentation with random horizontal and vertical flipping as well as color-based augmentations like contrast and brightness adjustments were also applied to augment the images.

You should refer to the case study document for more information on these training details.

Slide 25 To compare the performance of the dual-input model with other models, three other models were also trained.

The first on only daylight multispectral model (CNN MS),

…the second on only the nightlight model (CNN NL)

…and the final model was based on a transfer learning approach on nightlights with Landsat imagery (CNN transfer)

Experiments were also carried out with the amount of training data fed to the models in the different approaches

Slide 26 Now, let's look at the results.

When considering wealth distribution over space, the dual-input model performed better than the transfer learning approach of stacking the nightlights and Landsat bands together in a single model.

The dual-input model was able to account for 70% of the variation in ground-based wealth management with an improvement of up to 83% when the clusters are aggregated and observed from a district level.

Slide 27 This graphic shows that the dual-input model differentiates reliably between different wealth levels across countries.

It shows that the model performs better when aggregated at a district level with higher average r-squared values across the countries.

It can also be seen that performance on the individual held-out countries which were used for out-of-sample evaluation was never below 50%, with a median of 70.4% of explained variation.

This shows the ability of the model to differentiate both wealth levels among countries and within countries.

Slide 28 The results also show that, based on the amount of training data used in training the model, the CNN MS + NL model performs consistently better in predicting asset-wealth index than all other approaches tested, including the baseline models.

The dual-input model performs best across held-out country years as well as held-out villages within the same country as seen in Figure a. and Figure b, respectively.

Figure c. shows performance as a function of amount of training data used. It illustrates the superior performance of the CNN ML+NL model, i.e., the model that is the subject of this case study.

Slide 29 The CNN ML+NL model has the following strengths:

1. It can be used to efficiently predict the wealth distribution across different countries, complementing the more standard survey-based collection of data on wealth distribution.

2. The high accuracy of the model also means that it can be applied to downstream research or policy tasks such as intervention programs by the government, social protection programs or impact assessment tracking for stakeholders

3. The solution is also scalable and can predict wealth asset values for locations where ground-truth information is not available.

4. The robustness of the approach was demonstrated in a study of Southern Nigeria that showed its ability to ignore spurious data that don't indicate wealth, such as the very bright nightlight data caused by burning oil flares.

Slide 30 However, it does have some weaknesses

1. It performs poorly in settings where within-village wealth variation is high, possibly as a result of the model's difficulty in making accurate predictions in locally heterogeneous environments.

This weakness is also worsened by the presence of random noise added to village-level geo-coordinates to protect the privacy of the surveyed households.

2. Model explainability.

While deep learning models are powerful, they are essentially black-box systems and their results are not as amenable to explanation as simpler AI solutions. This may impede their adoption by policy makers.

Slide 31 To improve the performance of the model, higher-resolution images can be combined with other wealth indicators from passive sensors such as mobile phones and social media to obtain more timely and accurate data.

This data may improve the model's performance on within-village wealth variations.

Concerning the issue of the lack of explainability of deep learning neural network architectures, this remains an open research question.

To summarize:

1. Reliable and precise measures of economic well-being are essential inputs for effectively allocating resources based on needs, but this data is often unavailable at the local level in most developing countries.

2. The approach described in this case study trains deep learning models to predict survey-based estimates of asset wealth across approximately 20,000 African villages.

3. The model outperforms previous benchmarks from high-resolution imagery and transfer learning, exhibiting a performance that is comparable with independent wealth measurements from censuses, by accounting for 70% of the variation in ground-measured village wealth.

4. For temporal changes in wealth aggregated across districts, the satellite-based estimates from the model explains 50% of the variation, with daytime imagery being especially helpful in this task.

5. This solution shows that AI and machine learning can be used to address the issue of lack of reliable and up-to-date data when making economic decisions in African countries.

Here is the article on which this lecture is based. It is quite dense so we have prepared a case study précis of it. Please take the time to read the article and the précis, and then review this lecture again.

Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., . . . Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nature Communications, 11 (1), 2583. https://www.nature.com/articles/s41467-020-16185-w

Here are the references cited to support the main points in what we covered in the lecture.

Babenko, B., Hersh, J., Newhouse, D., Ramakrishnan, A., & Swartz, T. (2017). Poverty mapping using convolutional neural networks trained on high and medium resolution satellite images, with an application in Mexico. In Proc. NIPS 2017 Workshop on Machine Learning for the Developing World (pp. 994–1028-794). https://arxiv.org/abs/1711.06323

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. The American Economic Review, 102 (2), 994–1028. https://doi.org/10.1257/aer.102.2.994

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. Science, 353 (6301), 790-794. https://www.science.org/doi/abs/10.1126/science.aaf7894