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

Course AIML02: AI and Machine Learning in Africa

Module 1: The Potential of AI and Machine Learning in Africa
Lecture 2: Computational Sustainability and AI in the Developing World

Welcome to Module 1, Lecture 2, of Understanding AI and Machine Learning in Africa.

In this lecture, we will explore the key messages in an article entitled "Computational sustainability and artificial intelligence in the developing world" (Quinn et al. 2014). It addresses how technogy, in general, and AI, in particular, can play an important role in sustainable development in Africa, and other developing economies. The article in question refers to this topic as computational sustainability.

We highlight again the importance of understanding the local social and cultural behaviors when designing and deploying solutions. We identify three broad areas where AI can contribute to computational sustainability:

1. Intelligence gathering.
2. Compensating for a lack of human experts.
3. Choosing how to allocate scarce resources.

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

After watching and listening to this lecture, you should read the original article, take notes, and then go through this lecture again to consolidate the messages in the article.

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

1. Explain how technology has created new opportunities to address poor management of resources and improve human well-being.

2. Identify several examples of the use of AI in Africa in sectors such as

    - Health, for point-of-care diagnosis & disease surveillance.
    - Food security, for monitoring crop disease and predicting food insecurity.
    - Transport, for monitoring road traffic congestion.
    - Socioeconomics, for gathering information to guide policy decisions.

Slide 1 Welcome to Module 1, Lecture 2, of Understanding AI and Machine Learning in Africa.

In this lecture, we will explore the key messages in an article entitled "Computational sustainability and artificial intelligence in the developing world", written by John Quinn and co-authors in 2014.

It addresses how technogy, in general, and AI, in particular, can play an important role in sustainable development in Africa, and other developing economies.

The article in question refers to this topic as computational sustainability.

Slide 2 In the previous lecture, we emphasized the importance of trust, and the consequent need to be aware of local and regional social and cultural behaviors, in the adoption of any new technology.

In this lecture, we begin by reinforcing this with a practical example, road traffic congestion, that illustrates this general principle.

Slide 3 There are many causes of congestion

Inadequate infrastructure, such as outdated road layouts, unsealed roads, pot-holes, and

Lack of resources to monitor and control, such as scarce traffic police and inoperative traffic lights.

Some possible computational solutions include

Gathering data cheaply in real time.
Providing advice on optimal routes.
Dynamically redeploying traffic.
Redesigning road networks.

We will return to this topic later in this lecture, when we will give an example of how data can be gathered cheaply in real time.

Slide 4 For now, we just want to emphasize that any solution must take into account the unique nature of local traffic.

Assumptions made in developed-world intelligent transport systems,
such as that drivers travel in the correct lane and only on the road
might not be valid.

Slide 5 In the words of the authors of the article we are studying,

"finding the right set of assumptions can also be difficult: in an engineering approach to abstract away the nonessential parts of a problem, we often find that subtle, yet crucial social factors are lost in the process."
(Quin et al., 2014).

Slide 6 There is another complication.

In a fragile economy, technology deployments generally need to be immediately cost saving or profit making in order to survive.

Slide 7 There are three areas where AI can contribute to computational sustainability

1. Intelligence gathering.
2. Compensation for lack of human experts.
3. Choosing how to allocate scarce resources.

Slide 8 Data is the key to contemporary machine learning and informed decision making.

Very often, though, the data available for developing-world applications is both noisy and scarce.

Policies on socioeconomic intervention, health, or agriculture might be formed on the basis of poor data.

This data is often gathered through expensive surveys or personal interviews that only sparsely sample a region.

AI techniques make it possible to generate better inferences from existing data sources, combining many weak sources of data into a much strong one, or taking advantages of new data collection possibilities such as satellite imagery.

We study this possibility more closely in the socioeconomic case study in Module 2, Lecture 6.

Slide 9 Decision-making processes can compensate for a lack of skilled personnel.

For example, a shortage of laboratory technicians in poor countries can make it difficult for people to get reliable diagnoses of diseases.

A shortage of experienced agricultural workers who can recognize viral plant infections and advise farmers on the best course of action can mean that farmers might not be able to plan effectively.

AI systems can automate laboratory tests and providing personalized advice to farmers.

We'll say more about both these examples shortly.

AI can also augment and enhance the abilities of an expert, rather than merely substitute for them, or replace them (recall Licklider's prediction in Lecture AIML01-01-01).

Slide 10 Developing countries have very limited resources.

It is a challenge to allocate them optimally.

Specific cases can be framed computationally as optimization problems with different types of constraint

For example, at an agent level, an inspector traveling between pharmacies checking for counterfeit drugs ideally needs a travel schedule that is both cost-effective and difficult to predict.

At a macroscopic level, data- and learning-driven policies lead to evidence-based decision-making.

Slide 11 The standard test for malaria requires the analysis of a blood smear under a microscope.

The first step is to collect a small blood sample, usually by finger prick.

The second step is to prepare a glass slide using a suitable staining solution.

The third step is to examine the red blood cells through a microscope to see whether the characteristic shapes of malaria parasites are visible.

While the first two steps can be done by somebody with little training, the microscopic analysis requires significant experience.

Slide 12 Researchers in Uganda have investigated automating the diagnostic process using computer vision.

A data set of labeled images taken under field conditions from a Ugandan hospital was collected.

This was used to train a classifier using a set of morphological (that is, shape-based) image features.

Artifacts and poor staining added to the complexity of the parasite detection problem.

Nevertheless, diagnostic performance is superior to antigen-based rapid diagnostic tests.

Slide 13 Real time diagnosis was possible even on low-powered Android devices.

Slide 14 The tasks of mapping disease density over space and time and of diagnosing individual cases are usually carried out separately.

However, a “risk map” can be used to provide prior knowledge in the diagnosis of an individual with a known location.

In turn, the results of individual diagnoses can be used to update the map in a way that is more effective than simply using summary count data.

Slide 15 This combination of mapping and diagnosis is possible due to networked location-aware diagnostic devices.

Researchers in Uganda and the United Kingdom showed that the accuracy in each case could be improved using this approach, compared to carrying out the tasks in isolation.

Slide 16 The processes of food production and supply in developing countries can be easily affected by changes in climate or economic factors.

It is important to be able to anticipate threats to sustainability.

For example, by measuring the spread of viral diseases in staple crops or by tracking the degradation of farmland.

Slide 17 The economies of many developing countries are dominated by an agricultural sector in which smallholder and subsistence farmers are responsible for most production, utilizing relatively low levels of agricultural technology.

Disease among staple crops presents a serious risk, with potentially devastating consequences.

Monitoring the spread of crop disease is essential for targeted interventions.

Slide 18 The usual approach of sending teams of trained agriculturalists to visit farms and assess of crop health is expensive and inefficient.

Suitably-trained trained staff are scarce.

Transport is difficult and expensive.

Coordination of paper reports is time-consuming

Slide 19 Diagnosis of plant disease can be automated using images taken by a camera phone, enabling data collection by survey workers with only basic training

Slide 20 Such a system has been deployed in Uganda for monitoring viral disease in cassava.

The classification uses a simple set of color and shape features that are feasible to extract on a mobile device.

Slide 21 Using this approach, a real-time map of crop disease can be constructed.

Furthermore, an active learning approach can be used in which survey teams can be dynamically directed to the most informative areas.

Slide 22 Features derived from satellite images, such as the Normalized Difference Vegetation Index (NDVI) and Rainfall Estimate Data (RFE), have long been used for early warning of food shortages.

They give an overall prediction of food insecurity in an area, though in a heterogeneous population it does not directly predict which sectors of society or households are most at risk.

Slide 23 Using both demographic information about households and satellite observation data gives better accuracy in making predictions at a household level.

Researchers used information on 3094 households across Uganda collected between 2004–2005 combined with remote sensing images taken at 10-day intervals in the same period to model probabilistically the relationship between:

Calorific intake per person in a household;

Satellite Normalized Difference Vegetation Index;

Rainfall estimate data;

and demographic features such as land size, household size, and livestock ownership.

Slide 24 Due to poorly planned road networks, a common feature of many developing regions is the presence of small critical areas that are common hot spots for congestion.

Poor traffic management around these hot spots often results in major traffic jams.

Collecting real-time congestion information with current technologies is often prohibitively expensive in many developing countries.

The congested and chaotic nature of traffic in these regions can invalidate some conventional approaches, for example, any that make assumptions that vehicles travel in fixed lanes.

Slide 25 In a project deployed in Uganda, solar-powered units built around camera phones were found to be effective in collecting traffic flow data.

The use of such hardware drastically cuts the cost of collecting congestion information compared to conventional roadside CCTV systems or other traffic sensors such as induction loops.

Slide 26 The speed of traffic flow is calculated by calibrating the camera

and then using keypoint matches to identify motion,

distinguishing vehicles from non-vehicles.

Slide 27 Socioeconomic maps contain important indicators about the status of households at urban, regional, and national scales.

Many policy decisions made by governments and institutions are based on information captured in these maps.

National statistical institutes (NSIs) conduct censuses every 5 to 10 years and typically require a large number of enumerators to carry out interviews and gather the required information.

Conducting these surveys is often challenging and comes at a high cost.

AI and machine learning provides ways of generating the same information at much lower cost.

We discuss these issues in detail in the next module when, in Lecture 6, we explore a case study based on an article entitled "Using publicly available satellite imagery and deep learning to understand economic well-being in Africa".

To summarize:

1. It is essential to factor an understanding of local social and cultural behaviors into AI systems.

2. AI systems can contribute to computational sustainability in Africa by:

    - Gathering information efficiently.
    - Compensating for a lack of human experts.
    - Allocating scarce resources effectively.

    in areas such as

    - Health, for disease diagnosis & disease surveillance.
    - Food security, for monitoring crop disease and predicting food insecurity.
    - Transport, for monitoring road traffic congestion.
    - Socioeconomics, for informing government policies.

Here is some recommended reading. It is the article on which this lecture is based. Please take the time to read it and then review this lecture again.

Quinn, J., Frias-Martinez, V., and Subramanian, L. (2014). Computational sustainability and artificial intelligence in the developing world. AI Magazine, 35(3). https://ojs.aaai.org/index.php/aimagazine/article/view/2529

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

Aduwo, J.; Mwebaze, E.; and Quinn, J. 2010. Automated Vision-Based Diagnosis of Cassava Mosaic Disease. In Proceedings of the Workshop on Data Mining in Agriculture (DMA). Berlin: Ibai Publishing. http://cit.mak.ac.ug/staff/jquinn/papers/dma2010.pdf

Mubangizi, M.; Ikae, C.; Spiliopoulou, A.; and Quinn, J. A., 2012. Coupling Spatiotemporal Disease Modeling with Diagnosis. In Proceedings of the 26th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press. https://ojs.aaai.org/index.php/AAAI/article/view/8180/8038

Nakibuule, R.; Ssenyange, J.; and Quinn, J. A. 2013. Low Cost Video-Based Traffic Congestion Monitoring Using Phones as Sensors. In Proceedings of the 3rd Annual Symposium on Computing for Development (ACM DEV). New York: Association for Computing Machinery http://jquinn.air.ug/files/Nakibuule_2013_TrafficMonitoring.pdf

Quinn, J. A.; Leyton-Brown, K.; and Mwebaze, E. 2011. Modeling and Monitoring Crop Disease in Developing Countries. In Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press https://ojs.aaai.org/index.php/AAAI/article/view/7811

Quinn, J. A.; Okori, W.; and Gidudu, A. 2010. Increased- Specificity Famine Prediction Using Satellite Observation Data. In Proceedings of the 1st Annual Symposium on Computing for Development (ACM DEV). New York: Association for Computing Machinery. http://jquinn.air.ug/files/Quinn_2010_ACMDEV.pdf