Machine learning applications in developing-world sustainability

John Quinn

AI-DEV Group
School of Computing and Informatics Technology
Makerere University, Kampala, Uganda

CompSust 2012
Overview

• Why machine learning in the developing world?
• Some current projects
  • Biosurveillance
  • Automated disease diagnosis
  • Traffic
  • Auctions
• Some failures
Why machine learning in the developing world?

Many reasons, but in particular:

- Compensating for a lack of human experts.
- Improved intelligence gathering.
New data collection/deployment opportunity: 84 million internet enabled phones in Africa.
Example application: biosurveillance

Use different forms of data with spatiotemporal modelling to infer disease risk across spatial field. As data points are seen, refine estimates/predictions of risk (here for malaria):

\[ x_t \text{ is underlying risk across field.} \]
\[ p(x_t|x_{t-1}) \text{ are the spatial dynamics.} \]
\[ p(y^{(i)}_t|x_t) \text{ is the } i\text{th observation process.} \]
There has been a rapid break out of malaria in my community. We the p’ple of Abim are severely suffering from the deadly malaria; Why can’t mosuitoe spraying be done in Abim? Is Abim not apart of Uganda?

ARSENAL VS FULHAM 03.00 PM EMIRATES STADIUM

I have 2 say, K.C.C should clear the potholes in the new taxi park coz its flooding.

Why is it that whenever am back for holidays and at school, several young and old girls want me 2 fall in love with them.

Elephants are destroying people’s crops in nwoya district.
Automated diagnosis of malaria with digital microscopy.
Automated diagnosis of malaria with digital microscopy.

More than half of Ugandan health centres have microscopes, but most do not have lab technicians trained to operate them.
Automated lab diagnosis

- 2703 blood sample fields of view annotated by lab technicians, bounding boxes around 49,900 parasites.
- Parasite/non-parasite patches close to decision boundary:
Automated lab diagnosis

Object detection software can also be run on a phone:

Though we find that using the phone’s built in camera degrades accuracy quite a lot – bigger CCD probably necessary.
Combined diagnosis and mapping

We can model joint uncertainty in the diagnosis and the risk map. This improves the accuracy of both.

\[ x_t: \text{underlying risk of disease} \]
\[ y_t: \text{infection counts from health centres} \]
\[ d_t: \text{disease status of an individual} \]
\[ s_t: \text{symptoms presented by an individual} \]
Crop disease monitoring
Smartphone crop disease diagnosis

Image classification implemented on $100 Android phone:

350,000 of these phones sold in Kenya by mid-2011\(^1\): significant potential for crowdsourcing.

\(^1\)Source: *Android Marches on East Africa*, MIT Technology Review 06/2011
Counting whitefly

Disease vector counts are needed: an infuriating task when done manually. We can also automate this count in software on the phone.
Disease vector counts are needed: an infuriating task when done manually. We can also automate this count in software on the phone.
Traffic monitoring

Tropical-strength traffic monitoring
Tropical-strength traffic monitoring
Tropical-strength traffic monitoring

Roadside monitoring units

1. phone, mounted to side panel
2. extended battery pack
3. solar panel
4. observation hole
5. locked, waterproof box
6. bolted to lamp-post

Vehicle/clutter classification

Flow vectors from video stream

Mapping, routing, prediction
Auction design

- Double auction design for trading agricultural commodities.
- Market inefficiency can be assessed by looking at arbitrage opportunities from past data.
- Computing challenges: market clearing algorithm, reputation system, bid/ask language constraints over SMS.
Some failures

Data fail
Cholera and meningitis outbreak detection; attempted to use anomaly detection models for early detection.
Problem: data unusably noisy, lots of it fabricated completely.

Economic analysis fail
Crowdsourcing water supply problem reporting.
Problem: cost of a single SMS is significant to potential users.

Advocacy fail
Combining remote sensing data with demographic survey data to make specific famine predictions.
Problem: lack of interest from decision makers.
Conclusions

- Many opportunities for AI in the developing world.
  - Can compensate for a lack of experts.
  - Can collect better data for decision making.
- Large scale deployment/crowdsourcing have recently become possible.
- Success depends on several factors, particularly...
  - Enthusiastic, influential domain experts.
  - Focussing on ideas that make money or save money.