Machine Learning and Decision Making for Sustainability

Stefano Ermon
Department of Computer Science
Stanford University

IJCAI - July 13

Vision

Big Data

Crowdsourcing

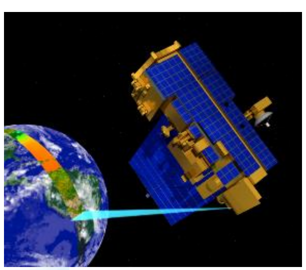
Technology

Push

Artificial Intelligence

Society

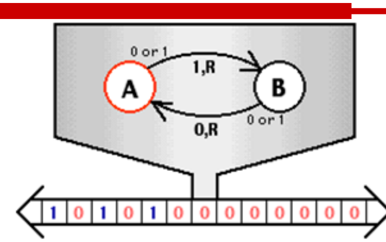
Pull



Sensing revolution

Challenges in reasoning about complex systems

Statistics



Computer Science

Engineering

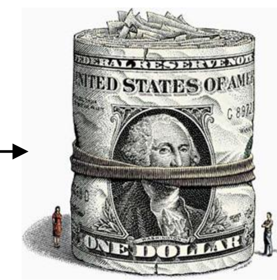
Operations Research

Three major challenges:

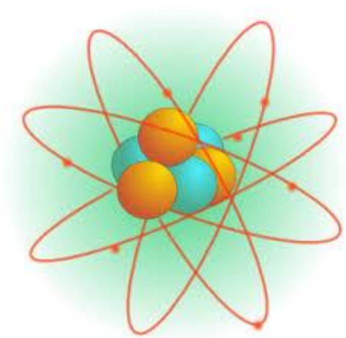
1. High dimensional spaces: need to consider many variables

2. Uncertainty: limited information, need to use stochastic models

3. Preferences and utilities: need to consider optimization criteria



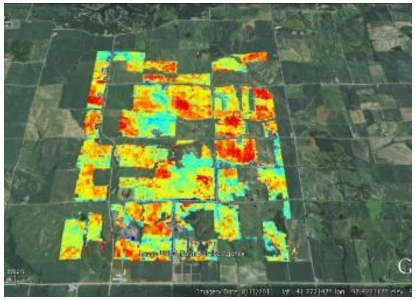
Economics



Physics

Management Science

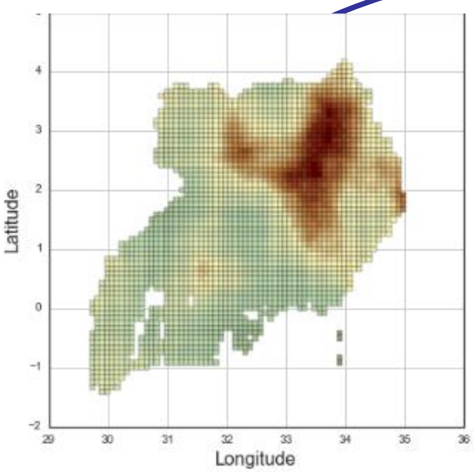
Computational Sustainability



Yield mapping
[Ongoing]

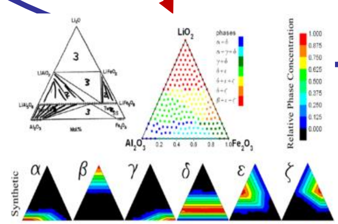
Preferences and utilities

Migrations
[AAAI-15]

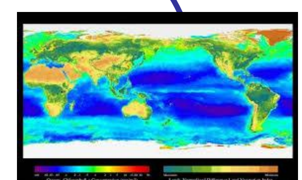


Large scale
Poverty mapping
[AAAI-16]

High dimensional spaces



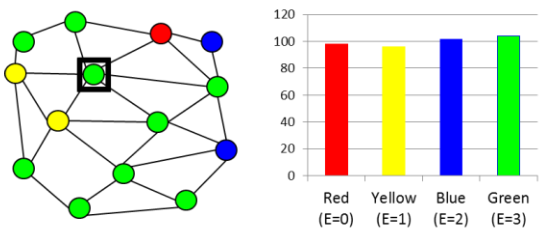
Materials discovery
[UAI-16, AAAI-15, SAT-12]



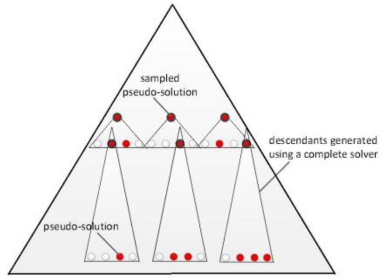
natural resources management
[UAI-10, IJCAI-11]

Uncertainty

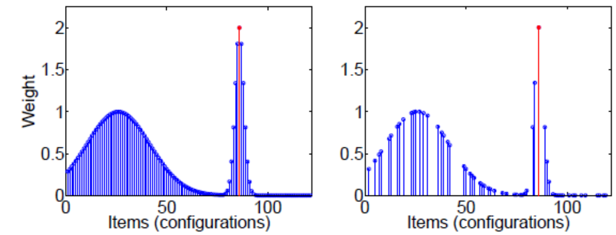
Research Agenda



bridging statistical physics and computer science
[CP-10, IJCAI-11, NIPS-11, NIPS-12]



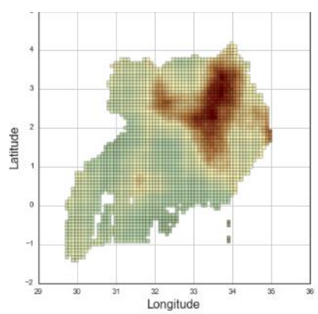
sampling
[NIPS-13, UAI-12, AAAI-16]



inference by hashing and optimization
[ICML-13, UAI-13, ICML-14, UAI-15, AISTATS-16, AAAI-16, ICML-16]

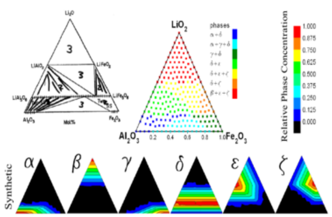
Foundations

Applications



Large scale Poverty mapping
[AAAI-16]

Migrations
[AAAI-15]

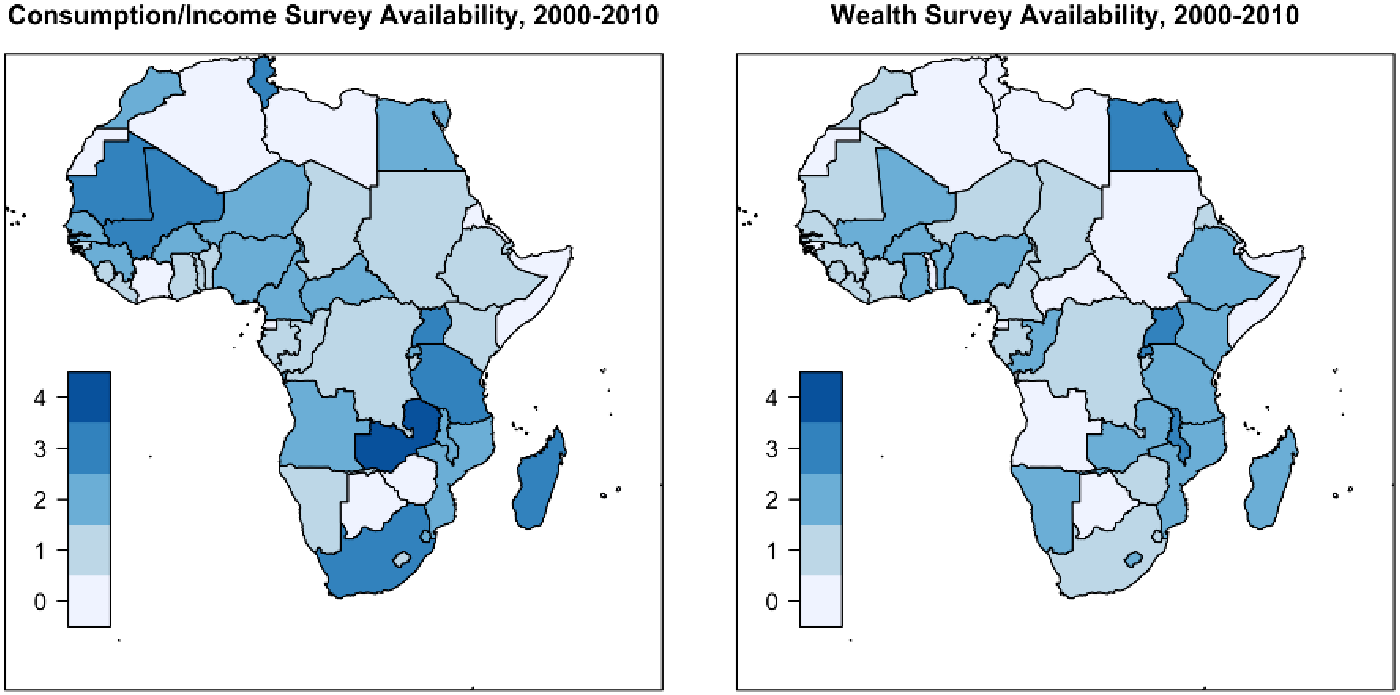


Materials discovery
[UAI-16, AAAI-15, SAT-12]

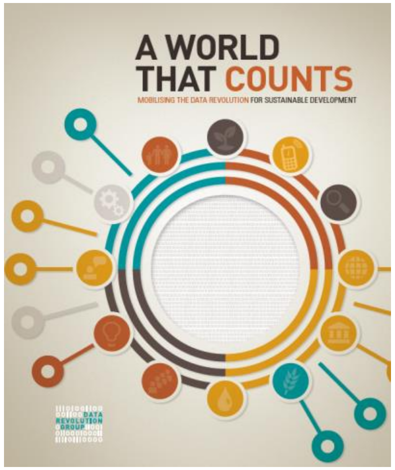
Why poverty?

- **#1** UN Sustainable Development Goal
  - Global poverty line: **$1.90/person/day**
- Understanding poverty can lead to:
  - Informed policy-making
  - Targeted NGO and aid efforts

Data scarcity

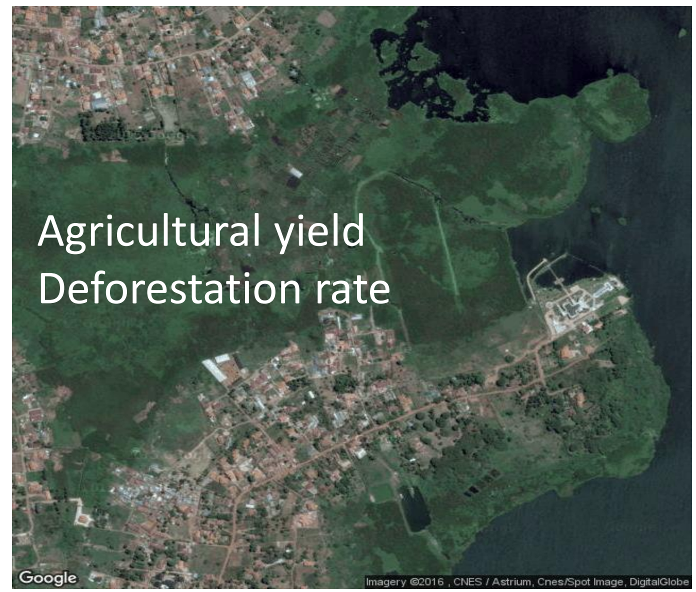


- **Expensive** to conduct surveys
- **Poor** spatial and temporal **resolution**
- **Questionable** data quality



Satellite imagery is low-cost and globally available

- Many cheap, unconventional data sources: remote sensing, phones/smartphones, crowdsourcing, …
- Remote sensing is becoming **cheaper** and **more accurate**



- Challenge: Lots of useful information, but data is unstructured

Standard supervised learning won't work

Input

Output

Model

Poverty, wealth,
child mortality, etc.

- Very little training data
- Nontrivial for humans

Transfer learning overcomes label scarcity

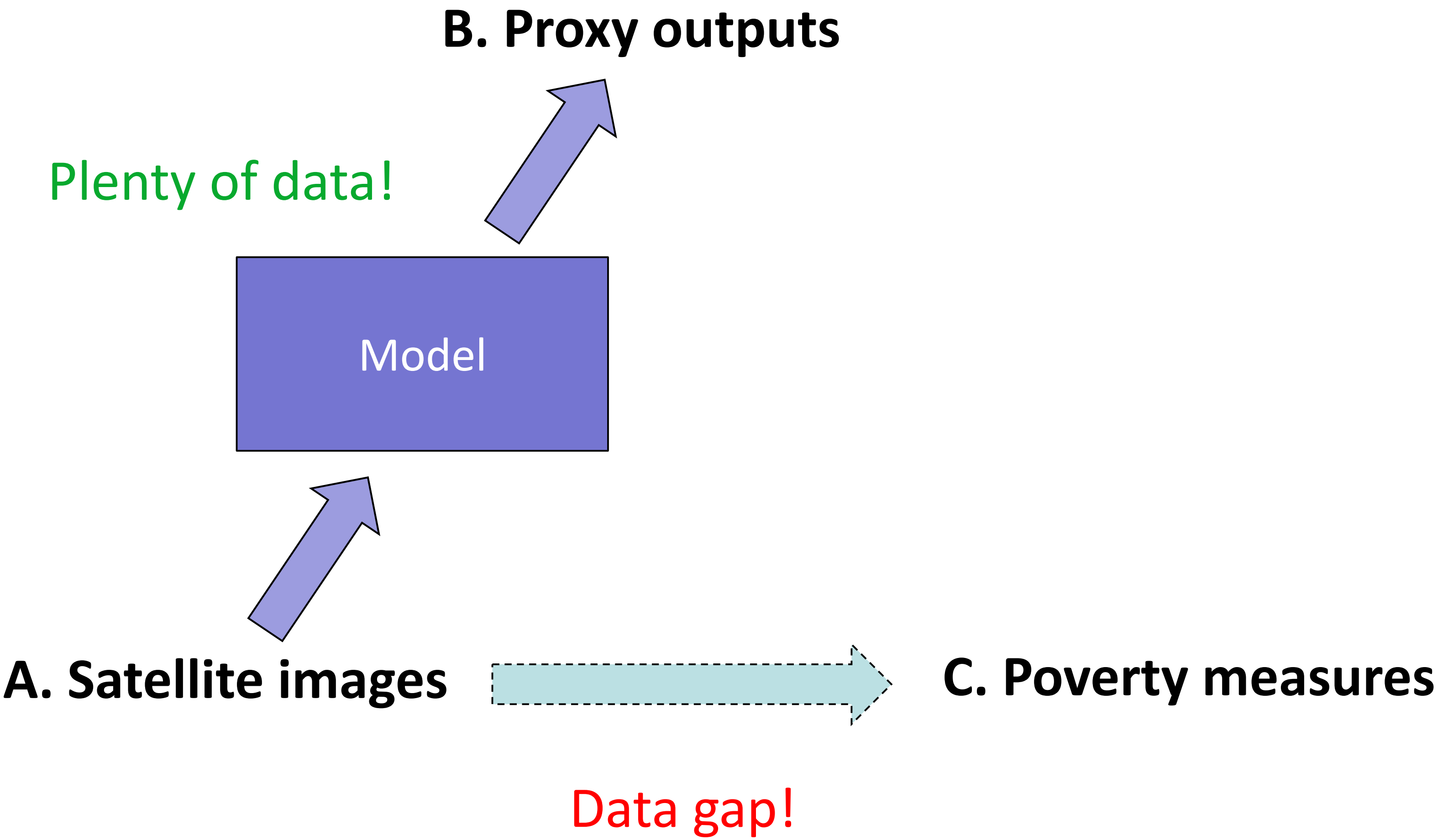
Transfer learning: Use knowledge gained from one task to solve a different (but related) task

Transfer learning bridges the data gap

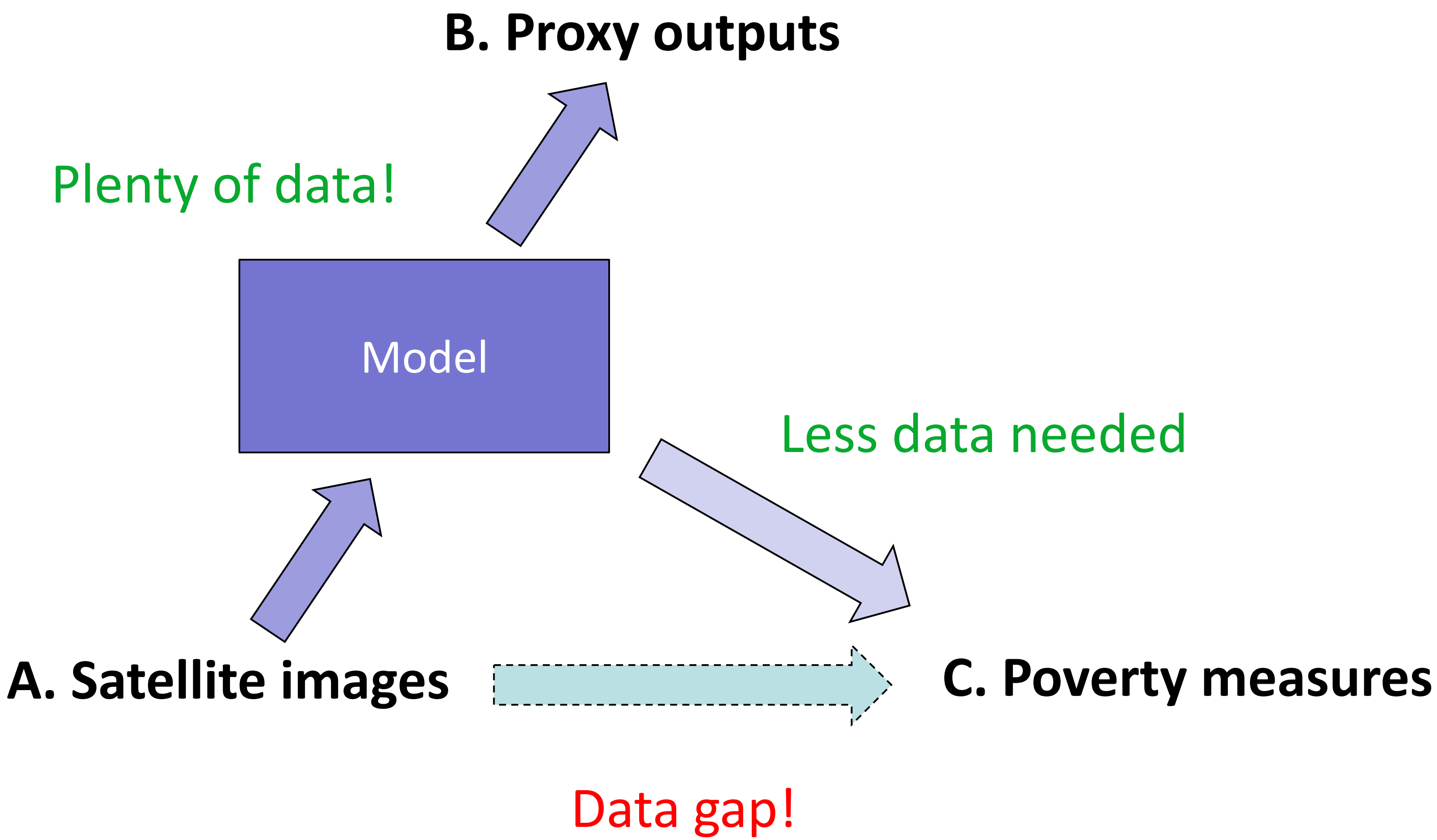
A. Satellite images → **C. Poverty measures**

Data gap!

Transfer learning bridges the data gap



Transfer learning bridges the data gap



Nighttime lights as proxy for economic development

Nighttime lights as proxy for economic development

Nighttime lights as proxy for economic development

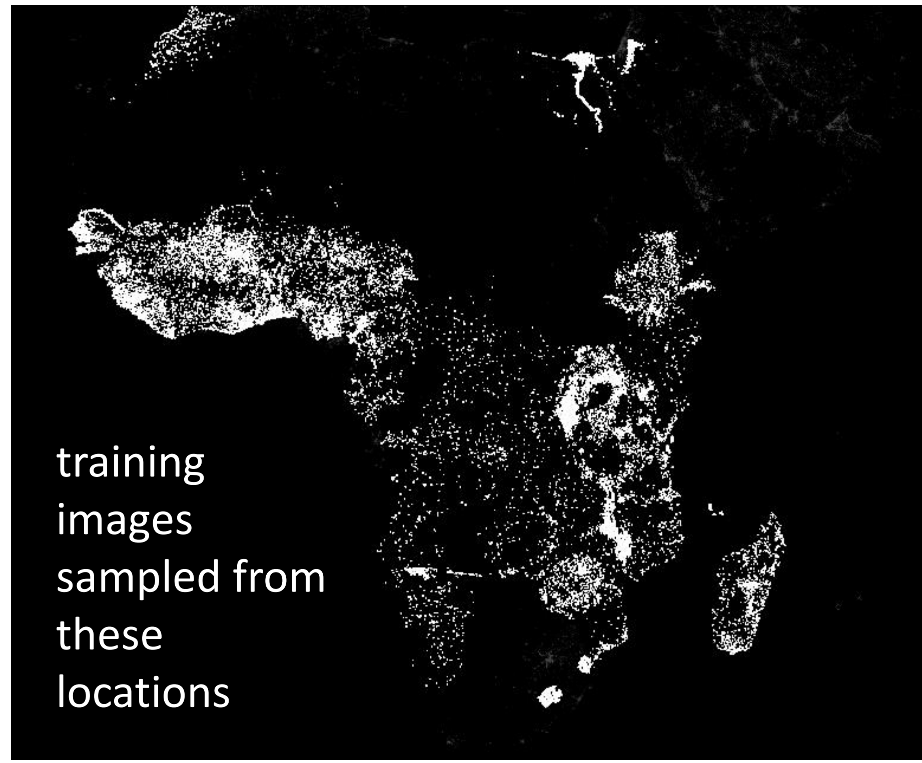
Training data on the proxy task is plentiful

Labeled input/output training pairs

( , Low nightlight intensity)

…

( , High nightlight intensity)

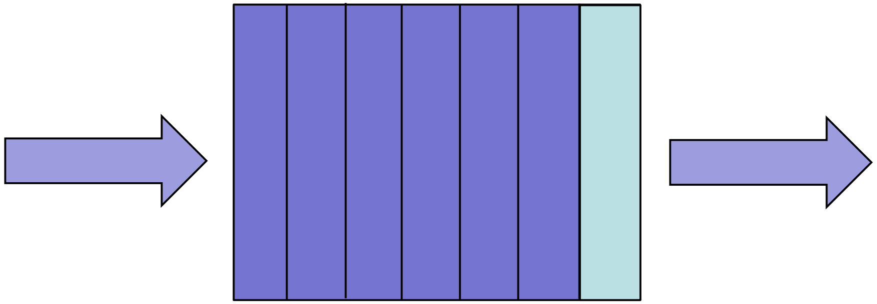


> 300,000 training images

Images summarized as low-dimensional feature vectors

Inputs: daytime satellite images

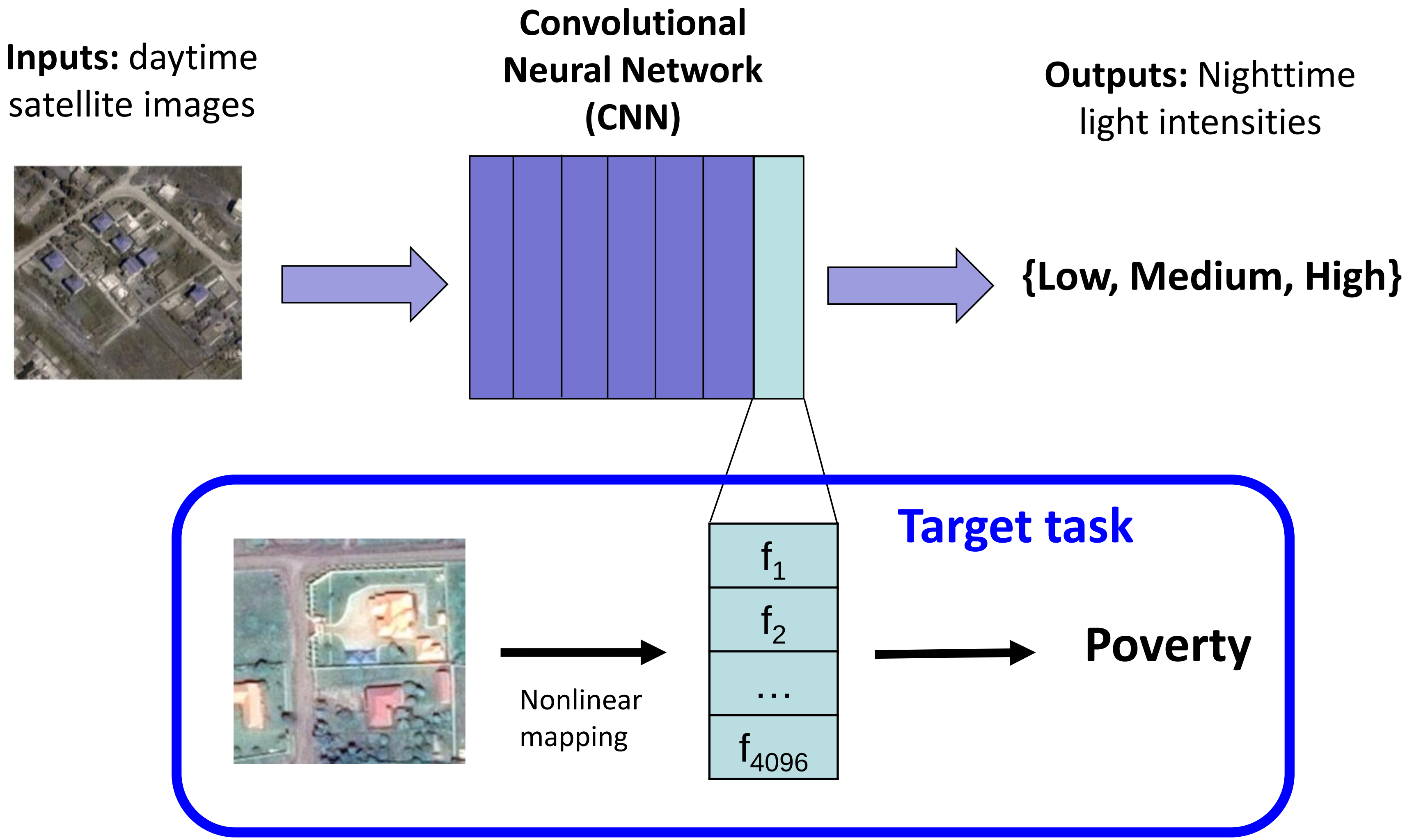
Convolutional Neural Network (CNN)



Outputs: Nighttime light intensities

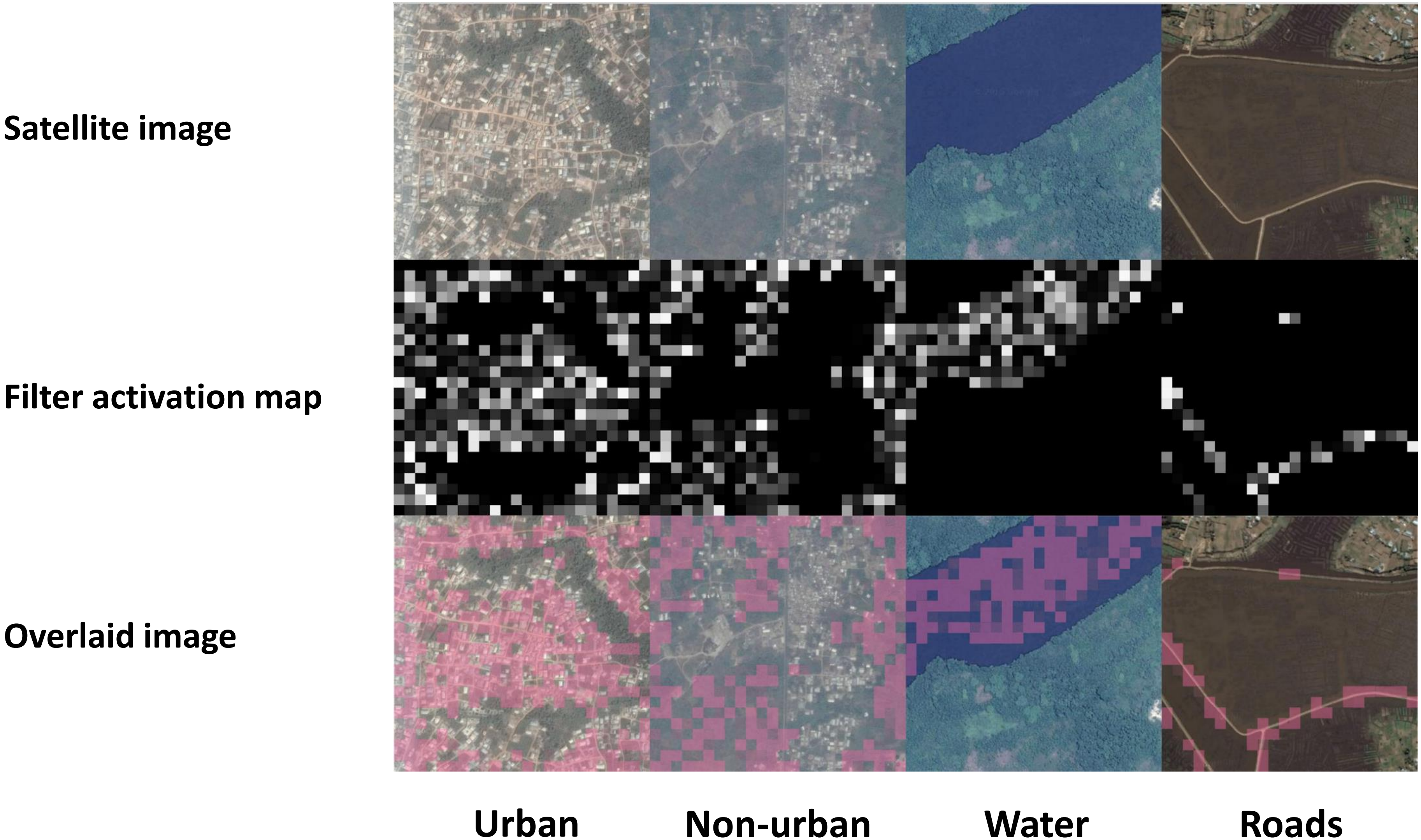
{Low, Medium, High}

Images summarized as low-dimensional feature vectors

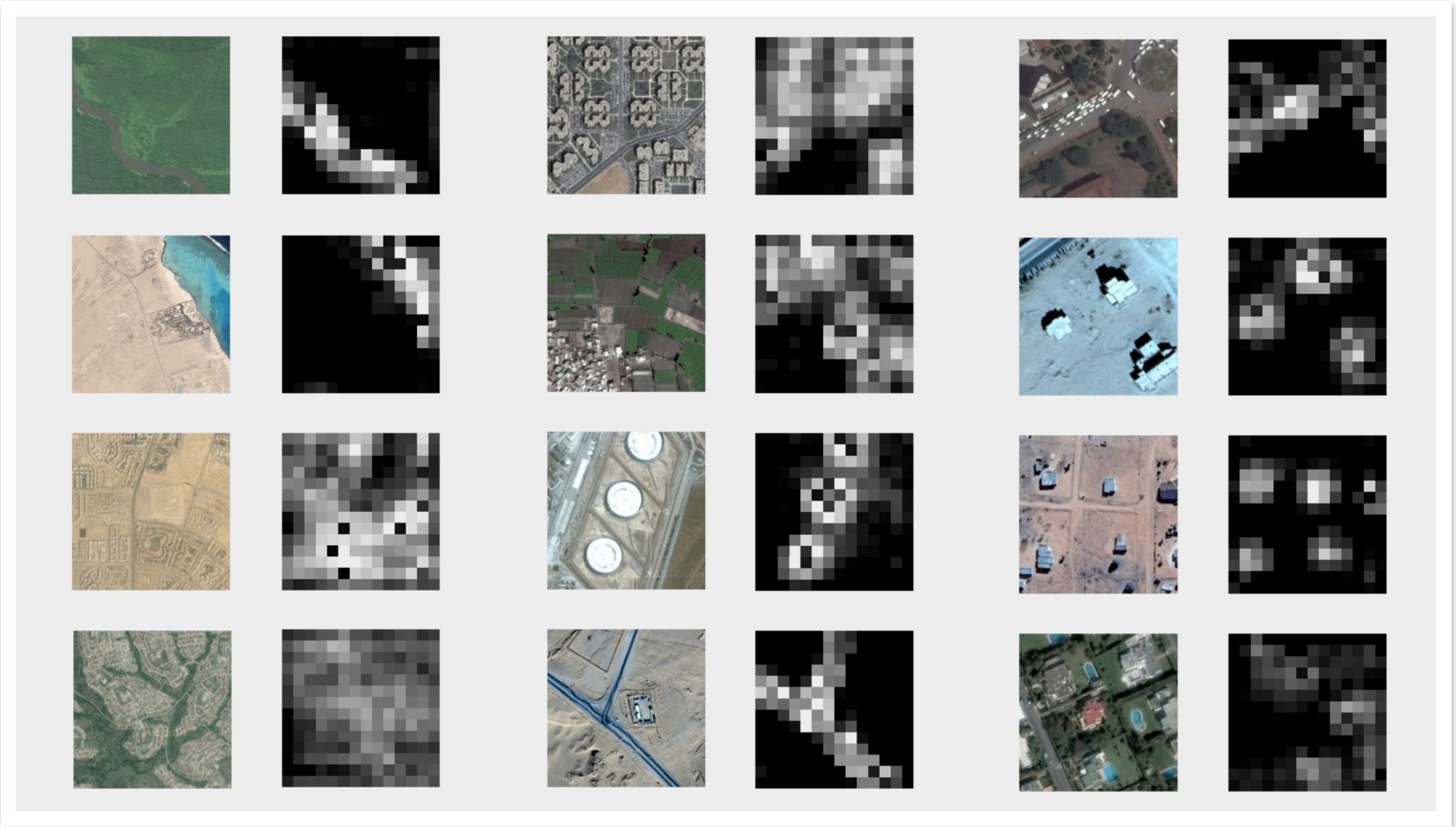


Have we learned to identify useful features?

Model learns relevant features automatically

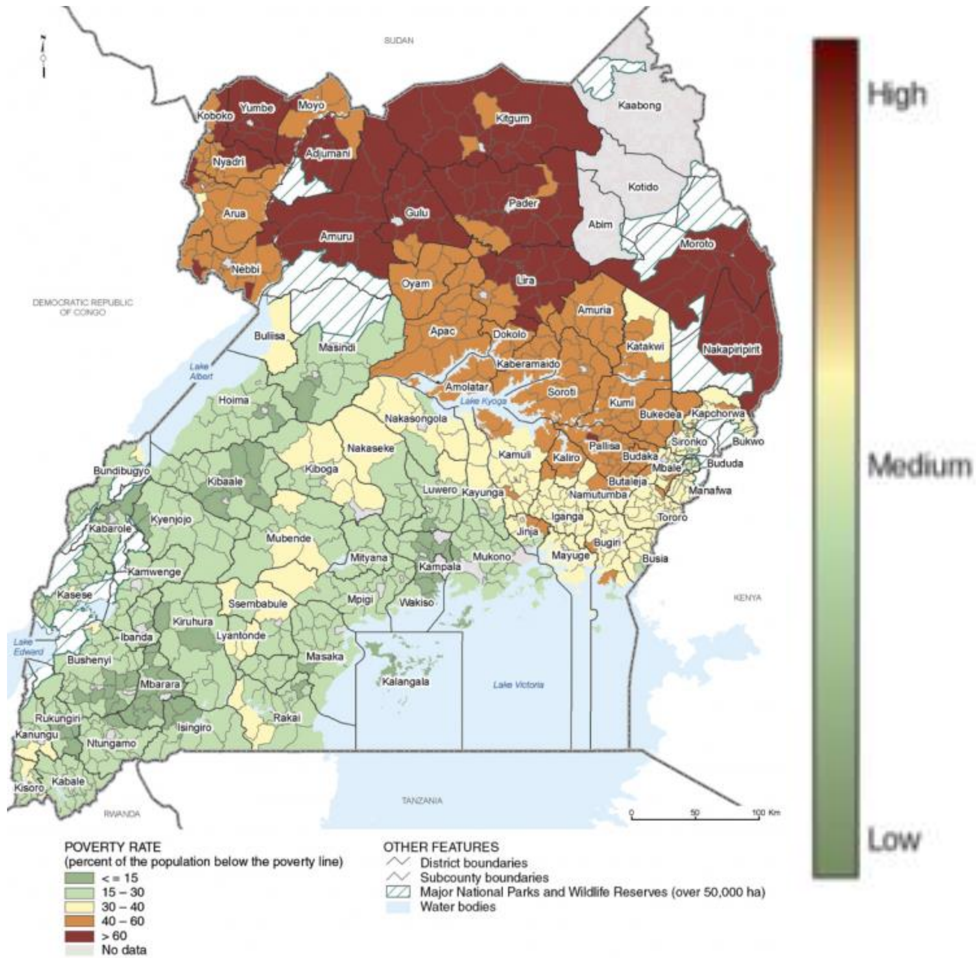


Model learns relevant features automatically



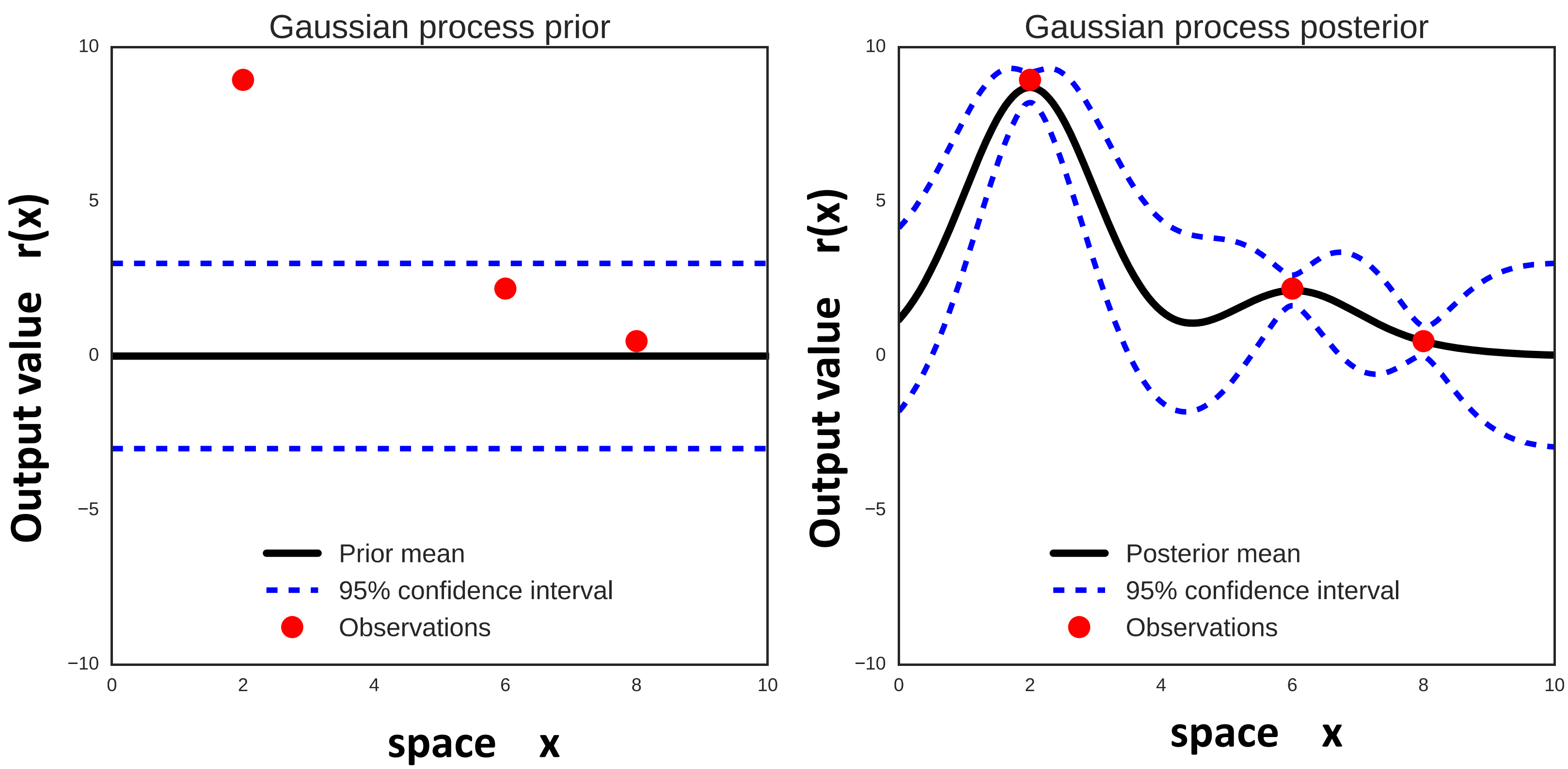
Poverty is NOT spatially independent

Uganda poverty rates (2005)



How to model spatial correlations?

Use Gaussian Process to model spatial correlations



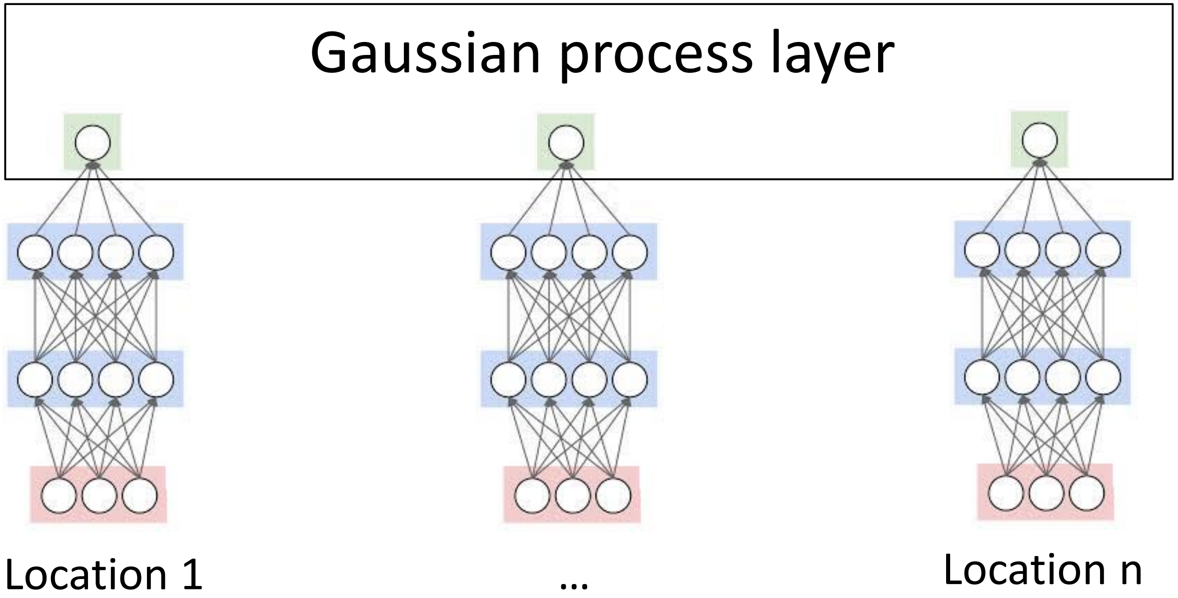
$$r(x) \sim \mathcal{GP}(0, k(x, x'))$$

How to take images into account?

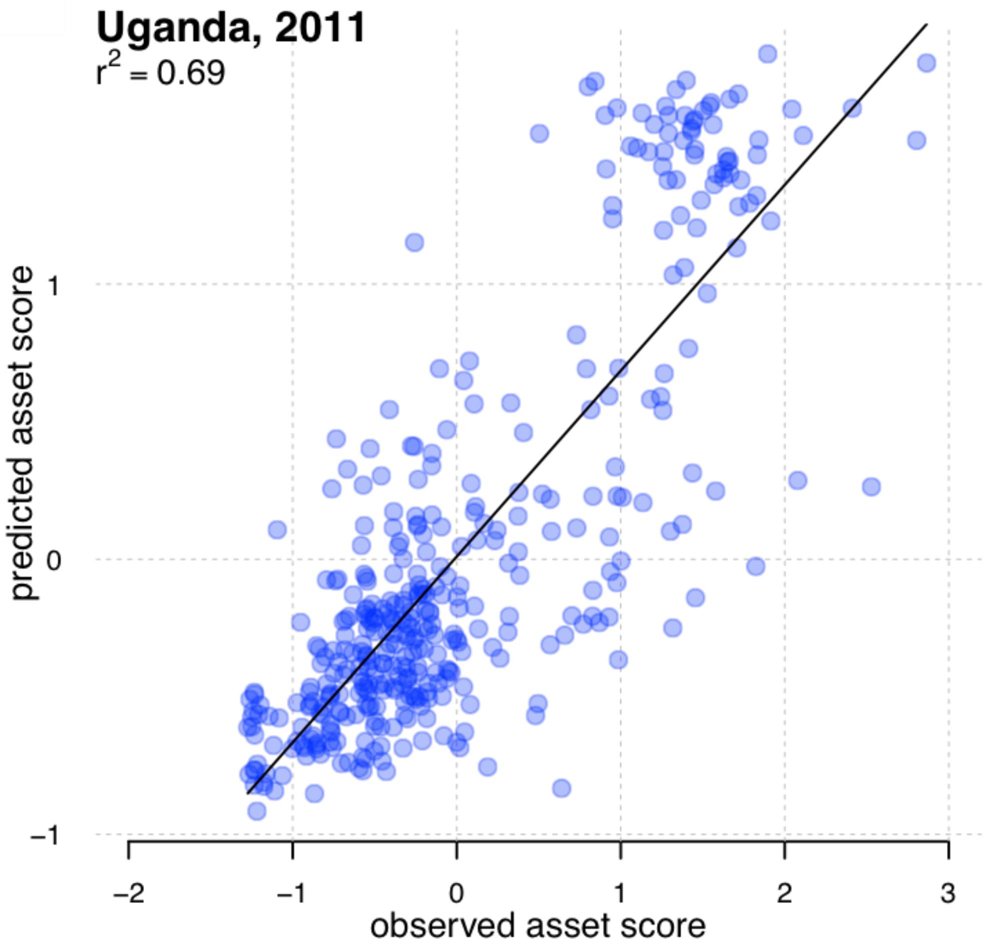
Linear GP model

$$f(x) = \underbrace{h(x)^T \beta}_{\text{Features from CNN}}$$

Key Idea: combine GP with CNN



Predicting household asset-based wealth

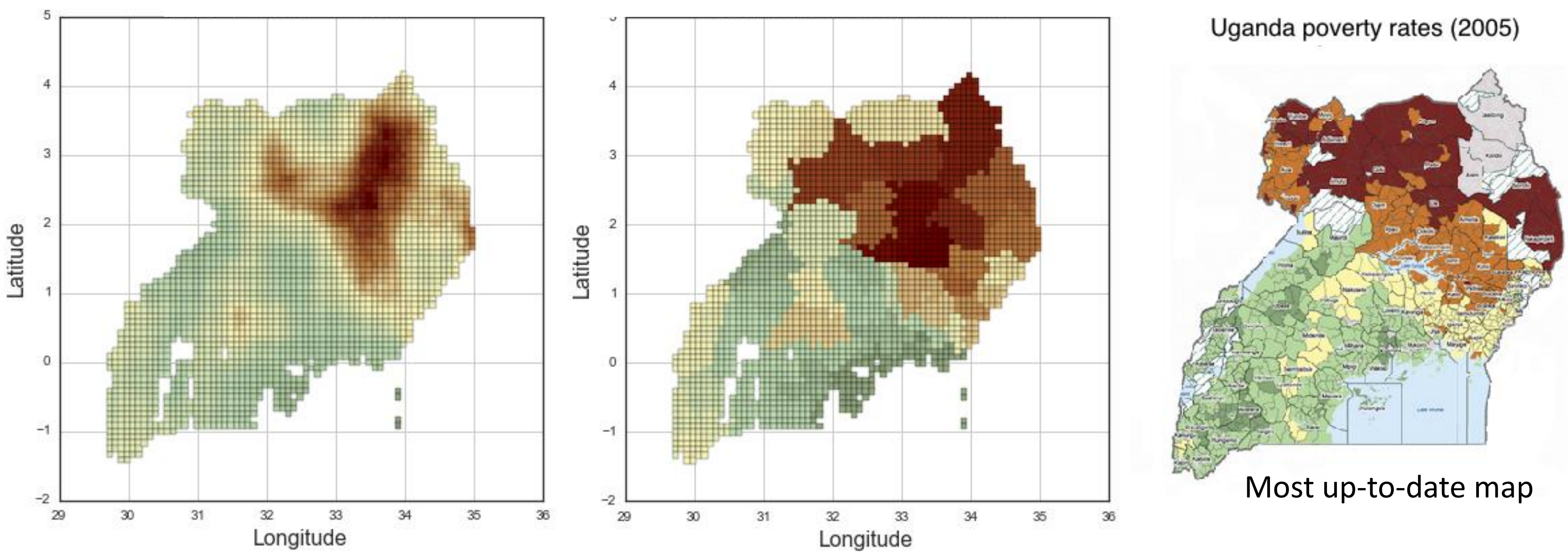


We outperform recent methods based on mobile call record data

Blumenstock et al. (2015) Predicting Poverty and Wealth from Mobile Phone Metadata, *Science*

Predicting Poverty from Space

Estimates from the model using about 500,000 images from Uganda:



Scalable and inexpensive approach to generate high resolution maps.

GiveDirectly

TheUpshot The New York Times

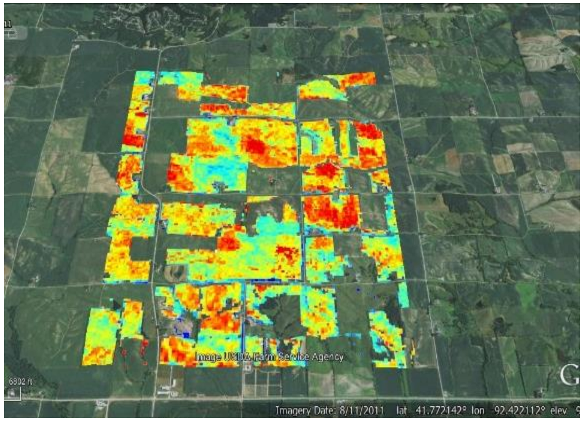
Satellite Images Can Pinpoint Poverty Where Surveys Can't

Economic View
By SENDHIL MULLAINATHAN APRIL 1, 2016

THE WORLD BANK
Working for a World Free of Poverty

Ongoing work: food security

- Mapping and estimating **crop yields**



- **Est. 2 billions more people to feed by 2050: information technologies** will have to play a role to **increase productivity**

MATT SIMON SCIENCE 05.25.16 1:26 PM

WIRED

THE FUTURE OF HUMANITY'S FOOD SUPPLY IS IN THE HANDS OF AI

Increasing productivity

Crop Challenge: which soybean varieties to plant to **maximize yield**, given knowledge about soil and climate?

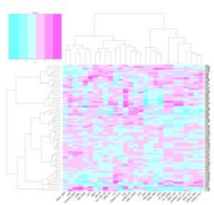
Data → Yield model → **Stochastic optimization** ← Uncertainty model

> 100 varieties

Ensemble of ML techniques

Monte Carlo simulations

syngenta
CROP CHALLENGE



$$\max_p \; \boxed{p^T \mu}$$

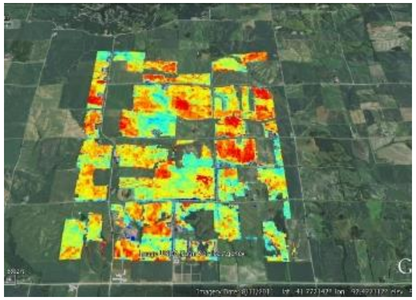
maximize **expected yield**

$$\text{s.t.} \quad p \in \mathcal{C}, \quad \boxed{p^T \Sigma p \leq \beta}$$

small **variance**

1st prize

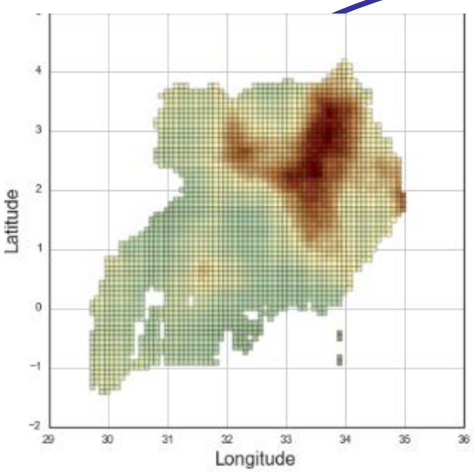
Computational Sustainability



Yield mapping
[Ongoing]

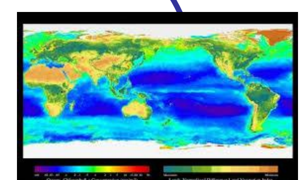
Preferences and utilities

Migratory pastoralism
[AAAI-15]



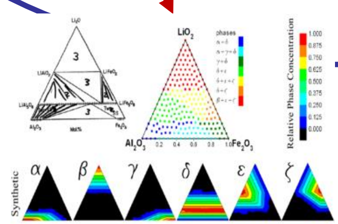
Large scale
Poverty mapping
[AAAI-16]

High dimensional spaces



natural resources management
[UAI-10, IJCAI-11]

Uncertainty



Materials discovery
[UAI-16, AAAI-15, SAT-12]

Motivation: migratory pastoralism

8 million pastoralists in Ethiopia and **3 million** in Kenya depend on livestock to make a living, relying on the vast **arid and semi-arid rangelands** of East Africa.

Motivation: migratory pastoralism

- Issues: **droughts, environmental degradation, climate change**

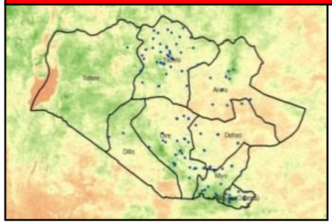
Understanding herding and grazing choices is critical to characterizing the impact of **policy interventions** on pastoralists' lives and on the ecosystem of the region

Motivation: migratory pastoralism

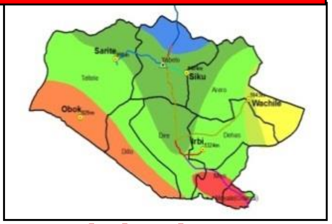
Develop a **generative** model to capture
the decision making processes of pastoralists

Can use the model to predict what would happen if we provide insurance, build new water points, climate changes, …

Environmental Remote Sensing

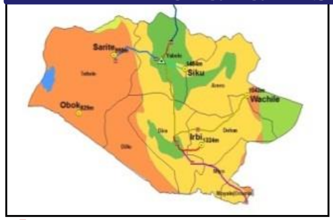


NDVI



Precipitation

Borana – southern Ethiopia



Agronomy

Land cover

\+

Socio-Economic Field Surveys

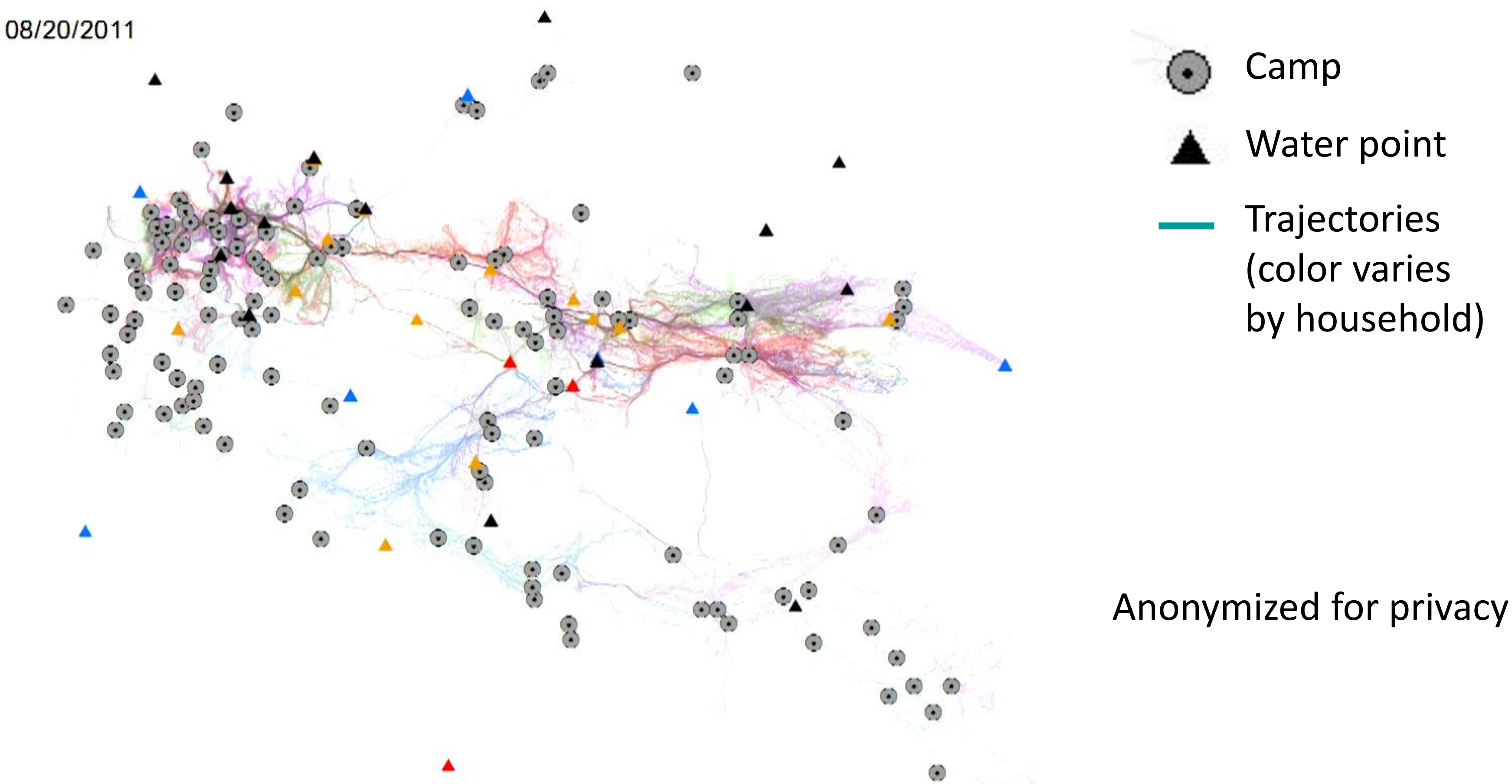
Herd size
Household size
Home location
Household assets
Education
Age

\+

GPS Collar Data
5 min intervals

Migratory pastoralism: Ethiopia data

GPS collar traces

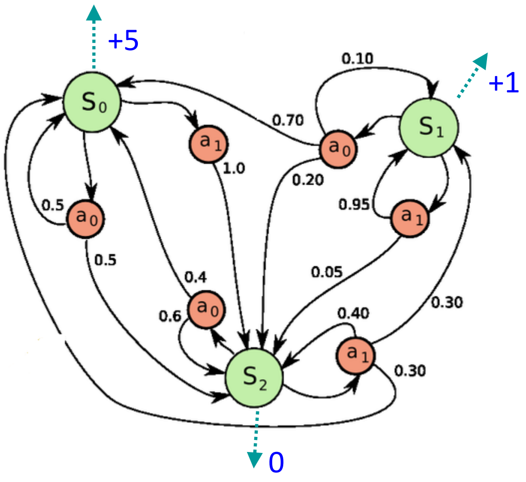


Anonymized for privacy

Markov Decision Process

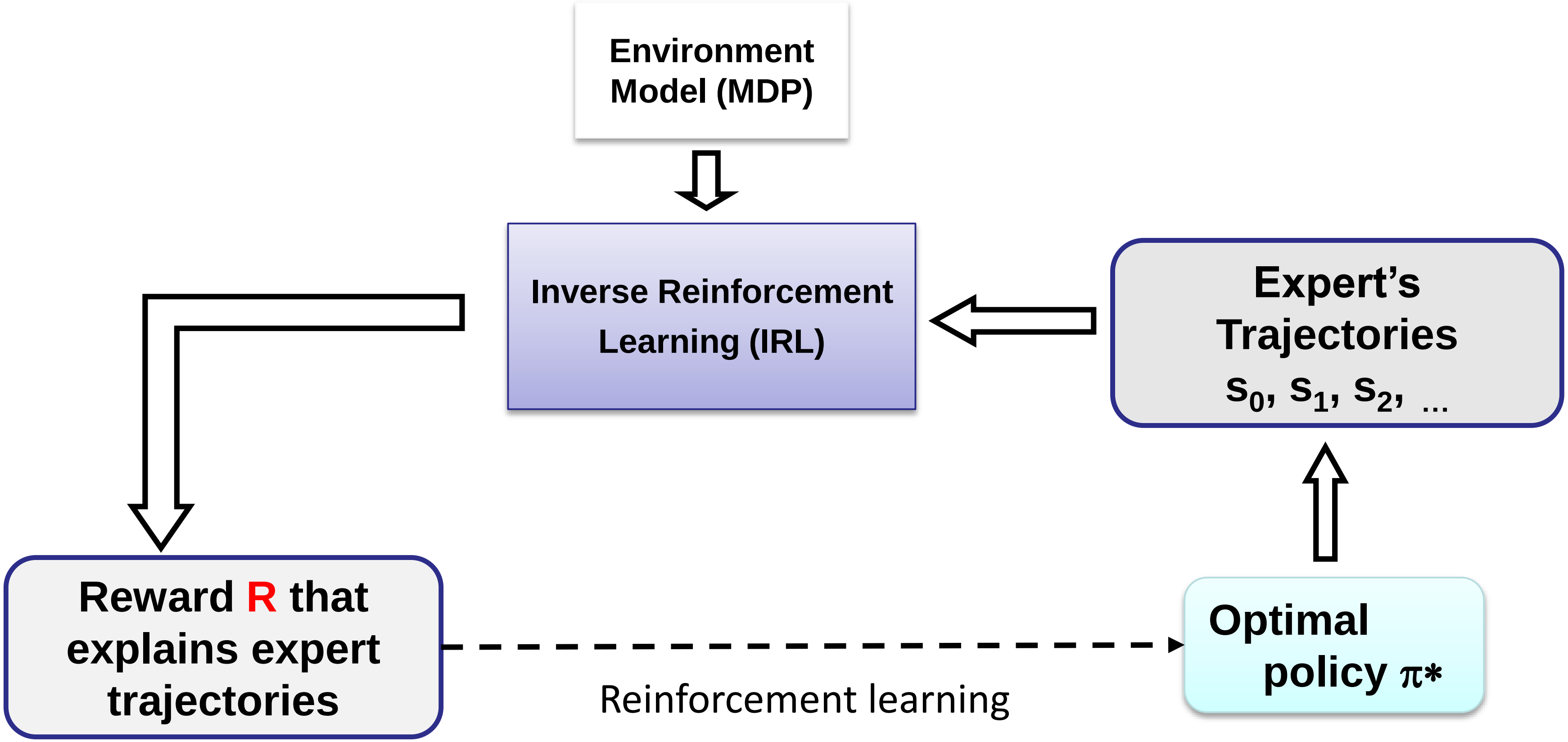
- **Markov Decision Process**

- States *S*
- Actions *A*
- **Reward** function (immediate): $\boxed{r : S \to \mathbf{R}}$
- Transition Probabilities: *P(s'|s,a)*



- Planning Problem/ Reinforcement learning: pick actions to maximize (expected discounted) total reward
  - Policy: sequence of decision rules that prescribe the action to be taken (depending on the current state)

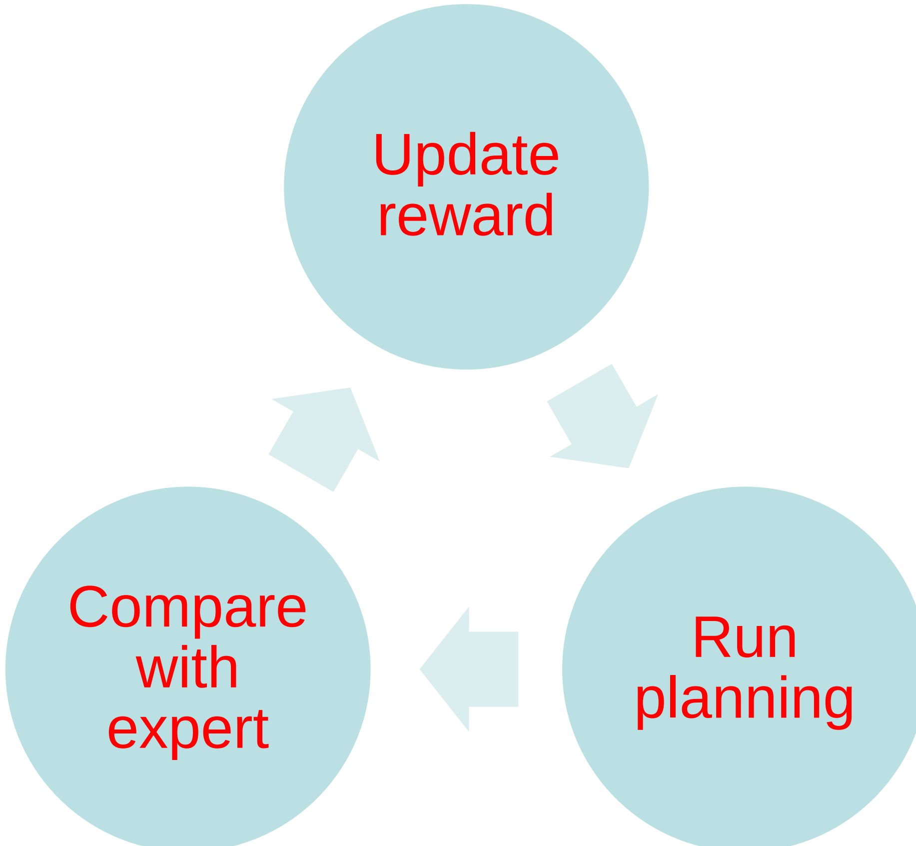
Estimation Problem



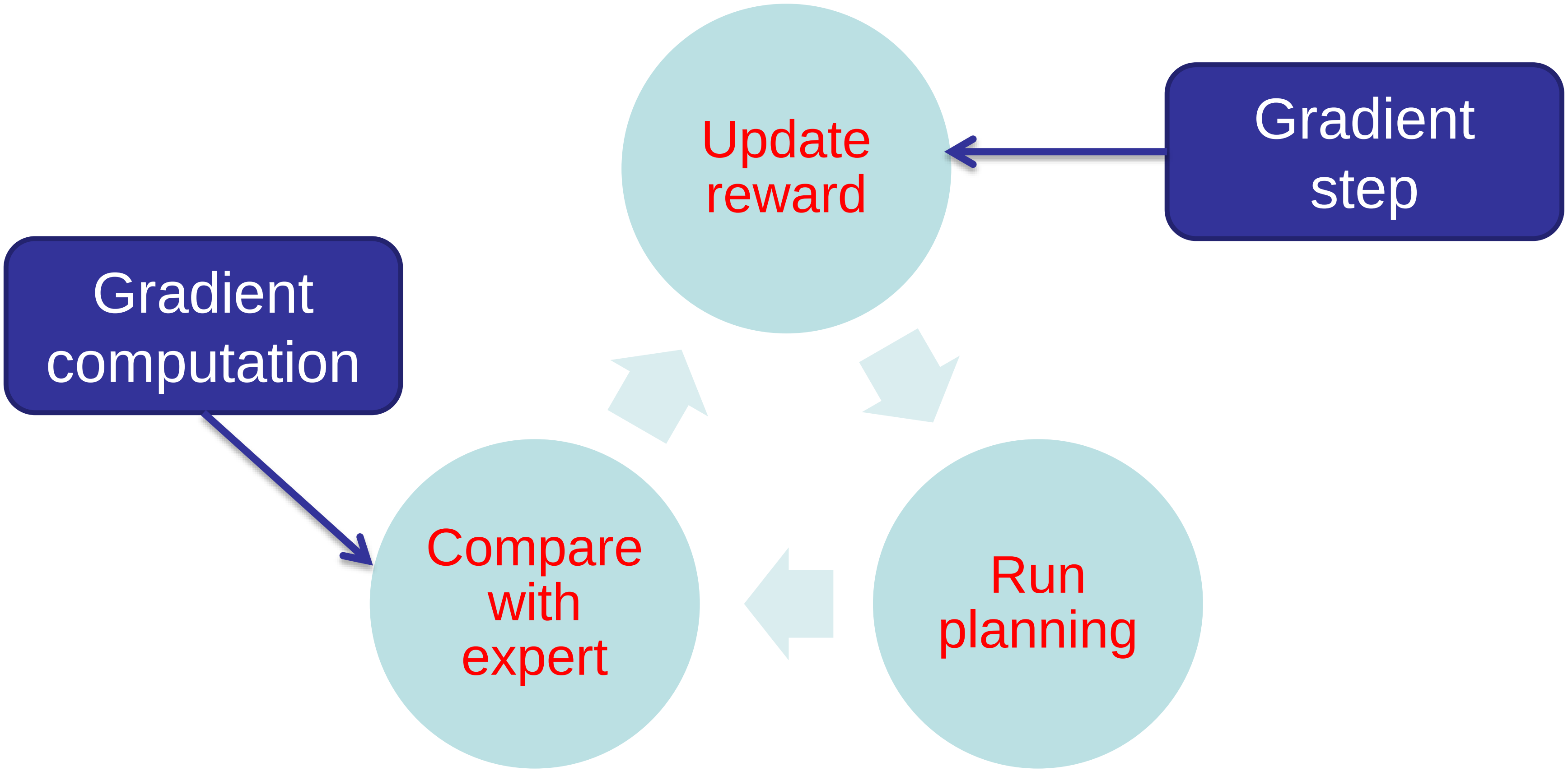
Assumptions:

- Agents (pastoralists) are following a policy
- **Rational**: Policy is optimal with respect to (unknown) reward **R**
- **Goal:** estimate **R** from the trajectories

Inverse RL

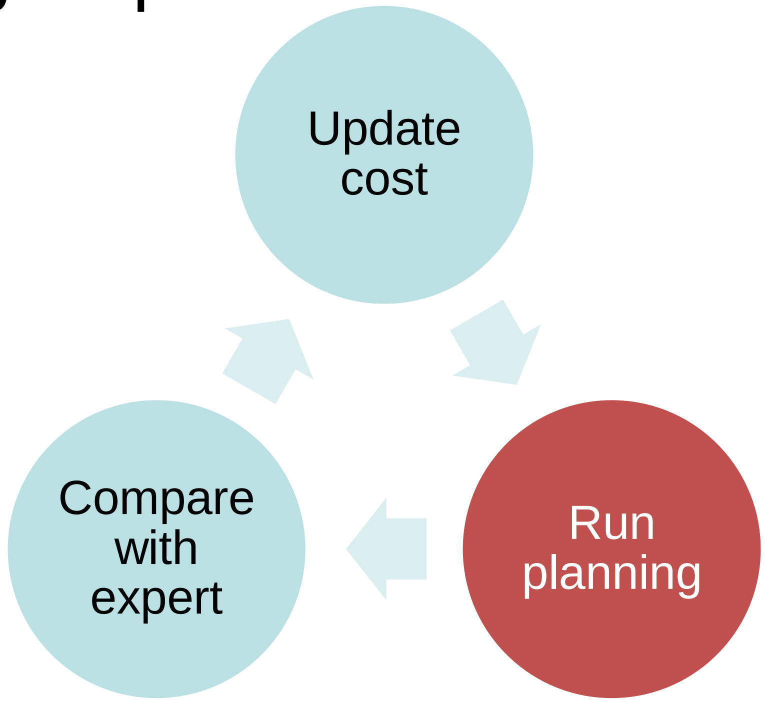


Inverse RL


Update reward
Gradient step
Gradient computation
Compare with expert
Run planning

Inverse RL

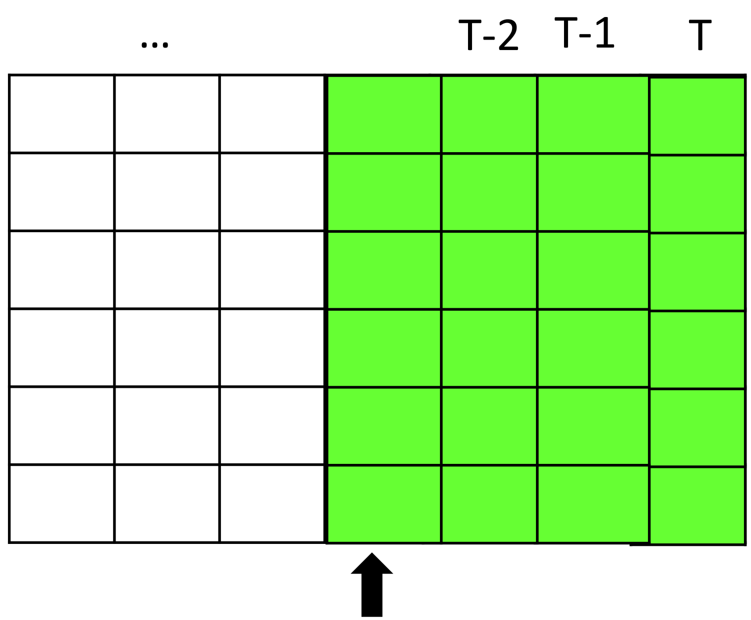
Expensive: have to solve a planning problem in a learning loop



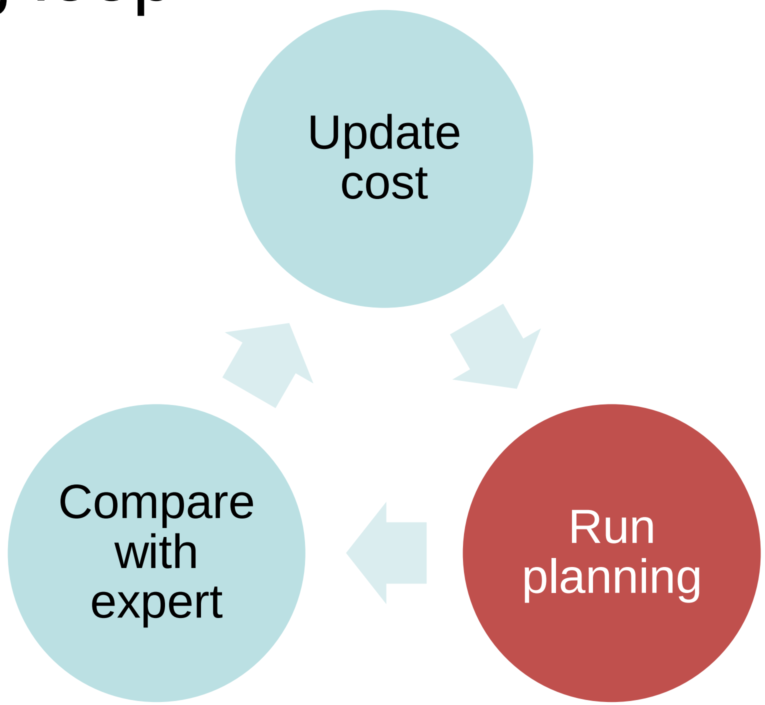
Simultenous Learning and planning

Expensive: have to solve a planning problem in a learning loop

Dynamic Programming Table:

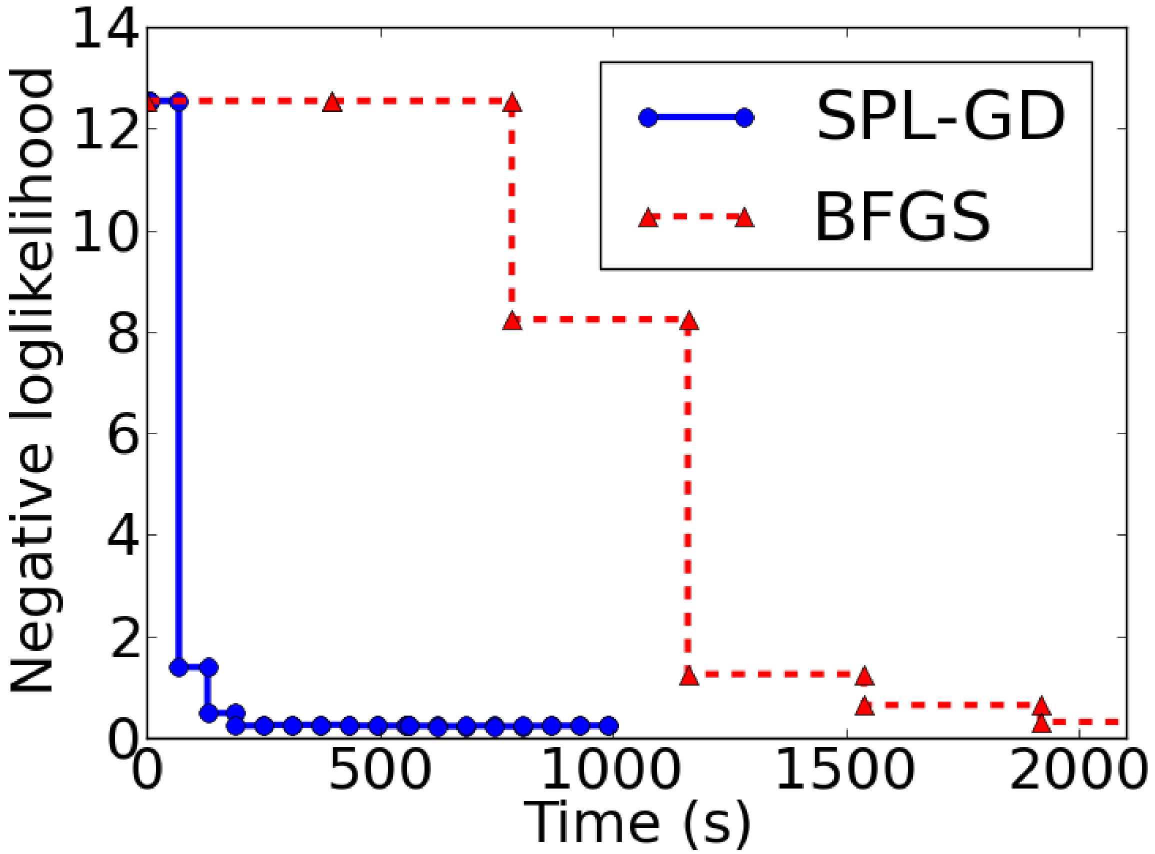


we have an optimal policy for the last 4 steps



Can update the cost as we fill the DP table

Convergence Rate



Our algorithm converges much faster (20x).

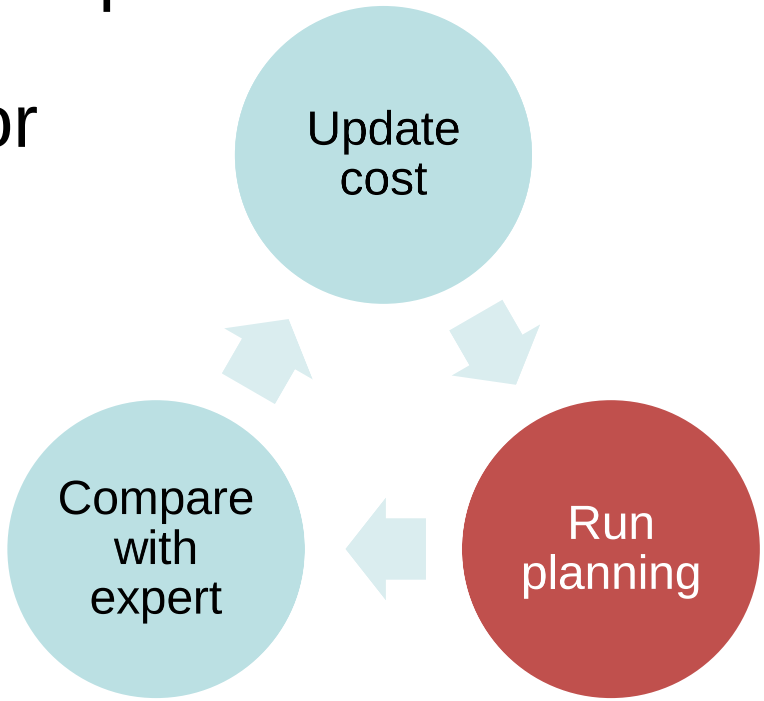
Policy gradient approach

Expensive: have to solve a planning problem in a learning loop

What if state space is too big for Dynamic Programming?

Policy gradient (with TRPO)

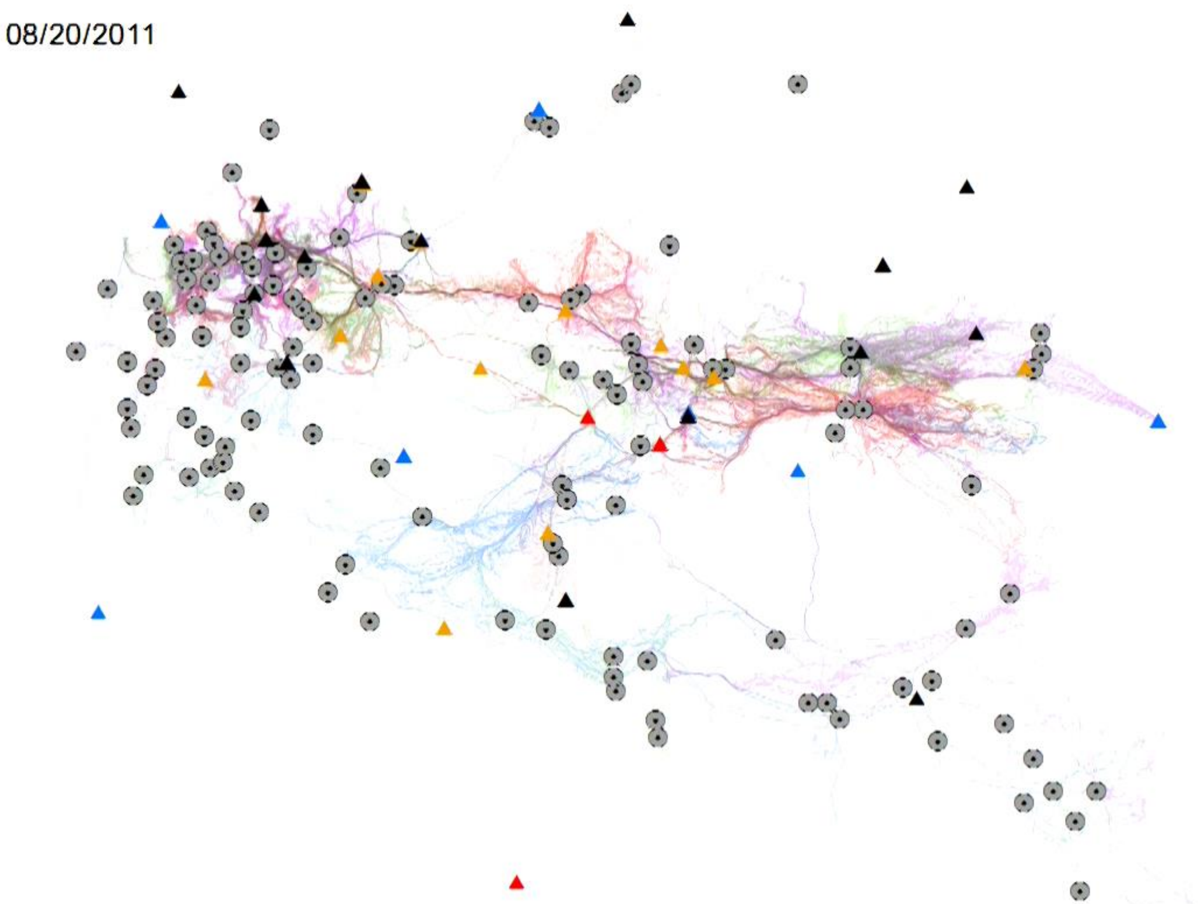
- Model free
- Fast
- Scales to high-dimensional, raw observations



Model-Free Imitation Learning with Policy Optimization" ICML-16

Migratory pastoralism: Ethiopia data

Features: distance between camps, greenness, dist. to water and village, distance to road, etc.



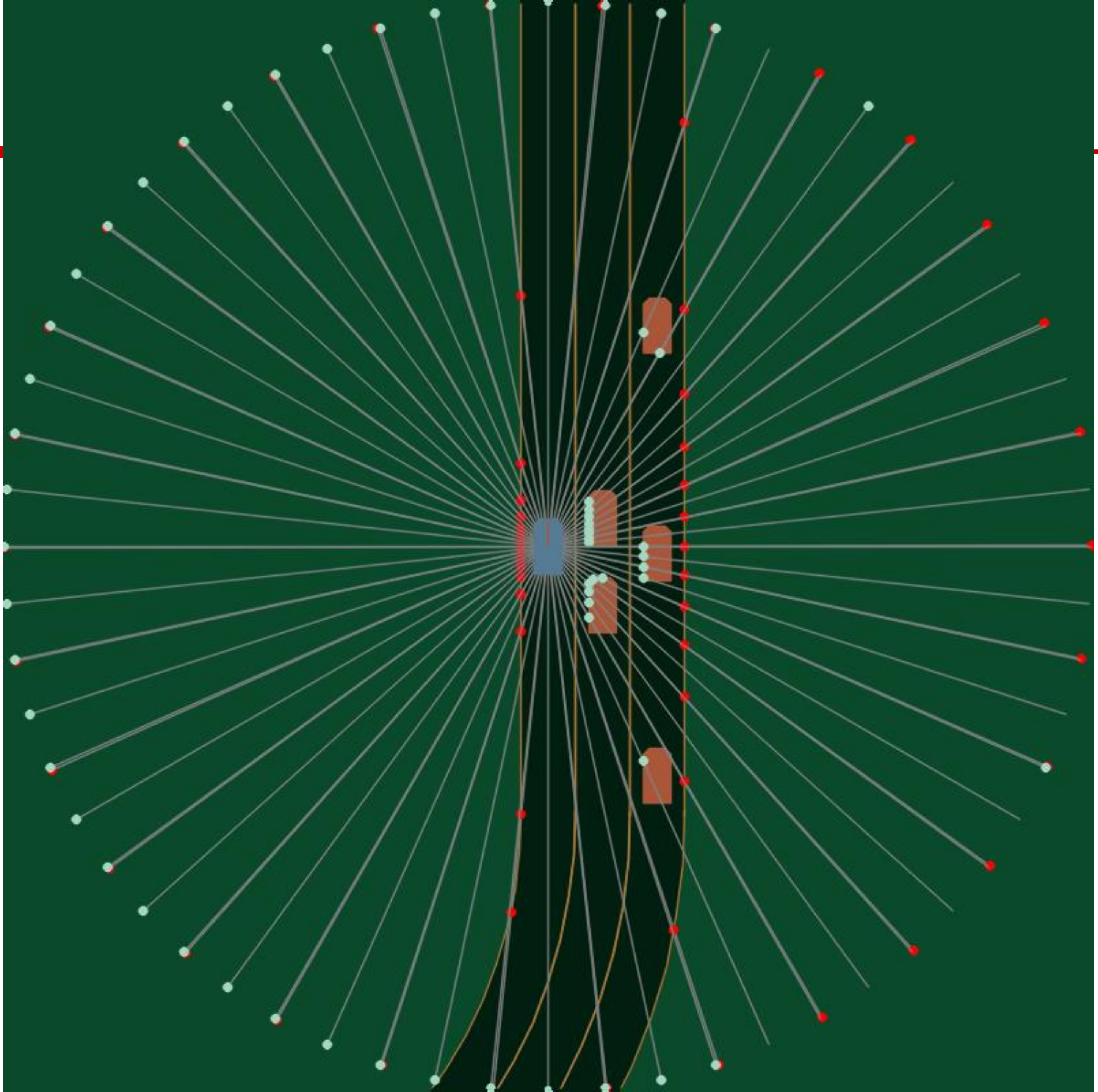
Anonymized for privacy

Evaluation

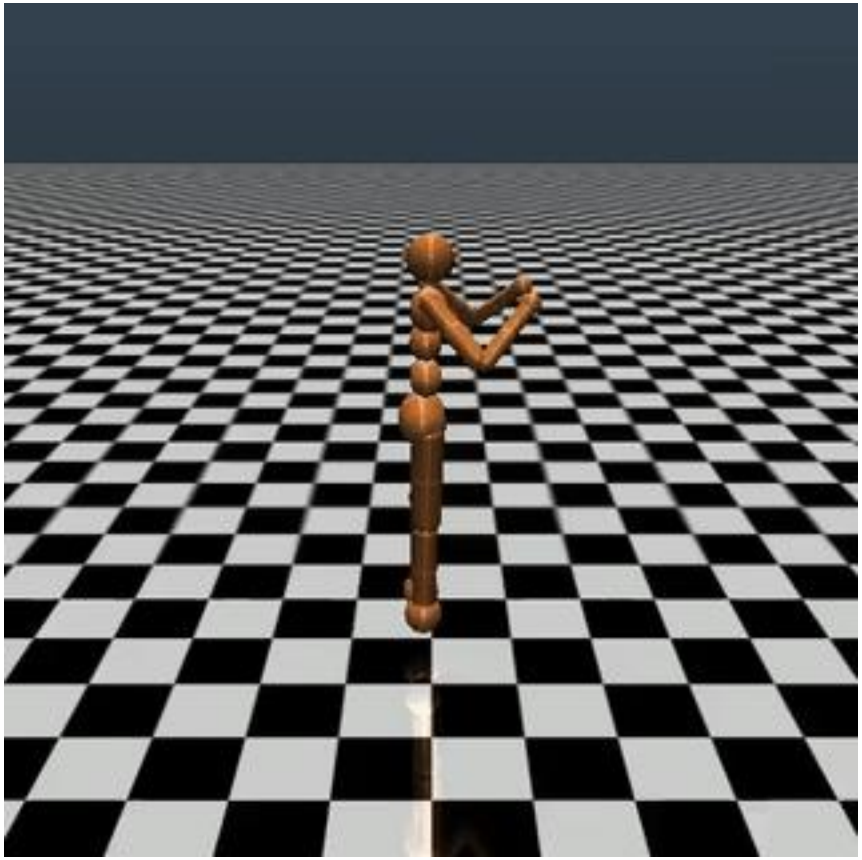
- 4 fold cross-validation: log-likelihood and number of predicted (movements)

Method	Fold 1 LogLik.	Fold 1 Moves (191)	Fold 2 LogLik.	Fold 2 Moves (85)	Fold 3 LogLik.	Fold 3 Moves (78)	Fold 4 LogLik.	Fold 4 Moves (116)
Markov	-8864.5	2209.8	-1807.8	372.2	-7265.7	1756.0	-4570.2	1214.1
MaxEnt IRL	-1524.4	424.6	-787.7	293.8	-796.7	339.7	-1004.2	299.4
Discrete Choice	**-1422.1**	**102.5**	**-657.8**	**104.9**	**-643.4**	**115.9**	**-911.3**	**94.9**

- Can fit well to the data
- The trained model recovers facts that are consistent with our intuition, e.g. herders prefer short travel distances
- Future work: compare what-if predictions with a randomized trial



Ongoing work

Generative Adversarial Imitation Learning, 2016 on Arxiv

Conclusions

- Growing concerns about the threats of AI to the future of humanity
- Recent advances in AI also create enormous opportunities for having deeply beneficial influences on society (healthcare, education, sustainability, …)
- New opportunities for CS research