# Two Courses and Six Questions

Sai Ravela

Massachusetts Institute of Technology

# The Questions

- How can Machine Learning Help Me?
  - Before: I don’t care about the methods, I just want the tool to apply
  - After: I have to worry about similar, if not the same, fundamental issues.
- Is the model efficacious?
  - Before: Look how it is reducing the error!
  - After: May be we should look at the bias-variance tradeoff.
- What is a good feature?
  - Before: I “feel that…”
  - After: Quantitative Performance and Verification essential
- Is the learning explainable?
  - Before: Who cares, the problems we have are very hard.
  - After: This would really help me understand
- Is there any physics in here?
  - Before: I don’t really care. We are desperate
  - After: It takes two to Tango, We can do better than either source alone…
- Is this all just statistics?
  - Before: I don’t know
  - After: I don’t know, I don’t know…

People typically get excited
by all that unsupervised learning
seems to throw at them

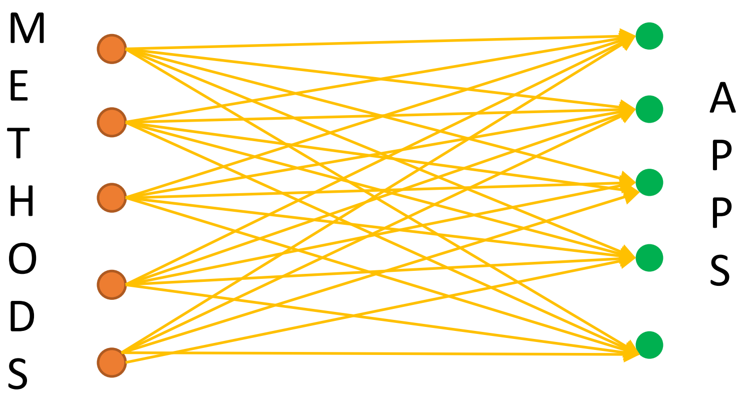
Only to become dejected,
then recognize some value and
ask themselves how this tool
could be honed further.

That’s the sweet spot

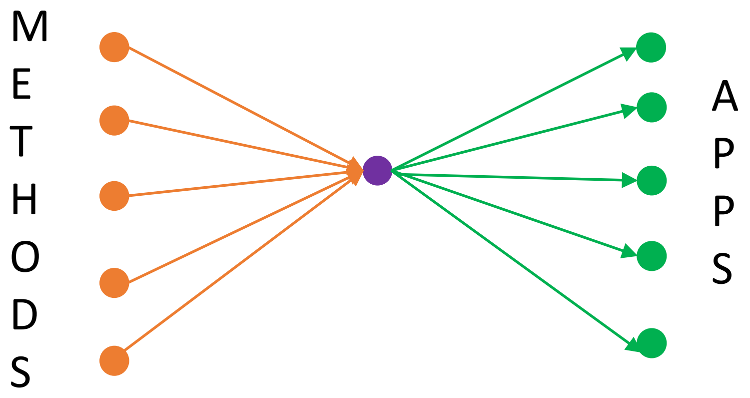
People typically get excited by the
error reduction in supervised learning
only to realize the “natural statistics”
are constrained.

They ask themselves how to do
this adaptively, and solutions emerge

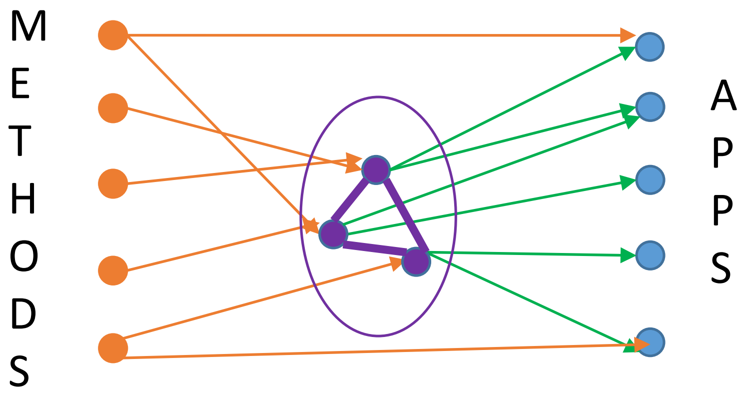
# Transformation




Too complex to learn everything about each other.




Too much for one person to parcel fundamental problems and deliver methods: 12 years




Emergence of a critical mass of internal nodes to sustain value

Energize a critical mass of interdisciplinary researchers.
Some who would like to meld both areas into a single whole.

# Course 2: Quantifying Uncertainty

## A little Ad-hoc

- Density Estimation: Exponential Family, Mixture Models, Kernels, Markov Chain Monte Carlo
- Model Selection: Jacknife, Bootstrap, Cross-validation and Information Criteria
- Dimensionality Reduction: PCA, ICA, and nonlinear modes
- Model Reduction: POD / EOF, Krylov Methods
- Response Surface Models, Polynomial Chaos
- Inference: Hierarchical Bayes, Graphical Models
- Time-dependent Inference: Linear, Ensemble, Mixture, Kernel, Mutual Information, and Particle Filtering and Smoothing
- Statistical Models: Regression Machines, Gaussian Processes, Markov Models
- Manifold Learning
- Information Theoretic Estimation, Control and Learning
- Stochastic Dynamic Programming Reinforcement Learning, Uncertain Probabilities

## Impact

Several research topics and (at least) one PhD thesis in the last three years motivated to apply material learned in this course

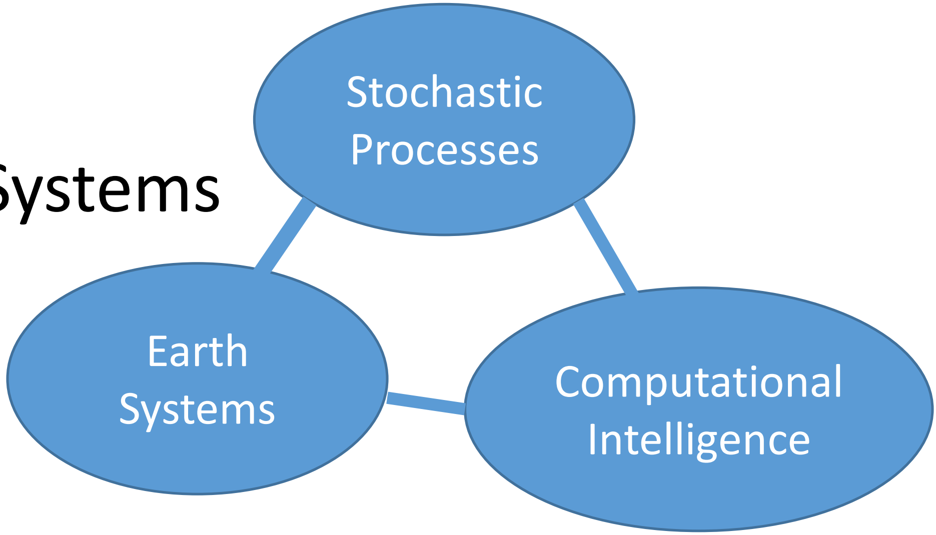
## Lesson

Reduce: the amount of material

Reuse: examples from other apps.

Recycle: Get past students into the pipe

# Course 1: Machine Learning Foundations for Natural Systems




- Distinction between Natural and Engineered Systems
- The breadth of Nature
- Not a glorified “toolbox” course
- Develop “Machine Learning Thinking”
- Collaborative
  - Students and other participants bring problems, augmenting ones I include
    - Variable from term to term
  - Include people from industry: Insurance, Energy
  - Include people from Labs: Lincoln, etc.
- Two part course.

# Machine Learning Foundations

Please

**Part 1**

- Foundational topics:
  Supervised vs. Unsupervised, Dimensionality, Bias-Variance Dilemma, Loss Functions, Generalization, Regularization, Predictability, Learnability, Explainability, Uncertainty, Validation, Sampling, Rigged Spaces
- Graphical Models
- Kernel Machines
- Ensemble Learning
- Manifold Learning
- Deep Learning
- Transfer Learning
- Physically-based Learning

**Part 2**

- Foundation topics
- Non-parametric Bayesian Inference
- Dictionary, Operator Learning
- Markov Decision Processes
- Reinforcement Learning
- Incremental Online Learning
- Causal Learning
- Information Theoretic Learning
- Physically-based Learning

# Special Emphasis

- The interaction between Learning and Physics
  - Learning Physical Models from Data
  - Learning Models from Data produced by Physical Models
  - Embedding Physical Models within Learning as Constraints
  - Augmenting Physics-based Models with Data

- These issues are not new
  - Replace the word Learning with Statistics
  - Explored by Data Assimilation, Uncertainty Quantification, Experimental Design and Atmosphere/Ocean/Climate communities for a long time!
  - E.g. Hurricane intensity forecasts using a mix of statistical and physical models; physical constraints in data assimilation.

- A key aspect of this ML course is to give appropriate credit to concepts and application that hitherto existed.

# Problems Solved

**Focus on a few**
**driven by applications**

**Many Problems**

- Sensing
- Computational Imaging
- Detection
- Classification
- **Segmentation**
- Morphology

- Estimation/Data Assimilation
- Control/Autonomy
- Planning Mitigation
- Model Reduction
- Model Abstraction
- **Uncertainty Quantification**
- Adaptive Sampling

- Predictability
- Characterization
- Model Abstraction
- Downscaling
- Parameterization
- Teleconnection
- **Non-local Operators**

Of special interest:

Nonlinearity, dimensionality,
Uncertainty
Extreme, Rare, Transient Events

# Some Application Areas

**Lots of Topics: Scary?**

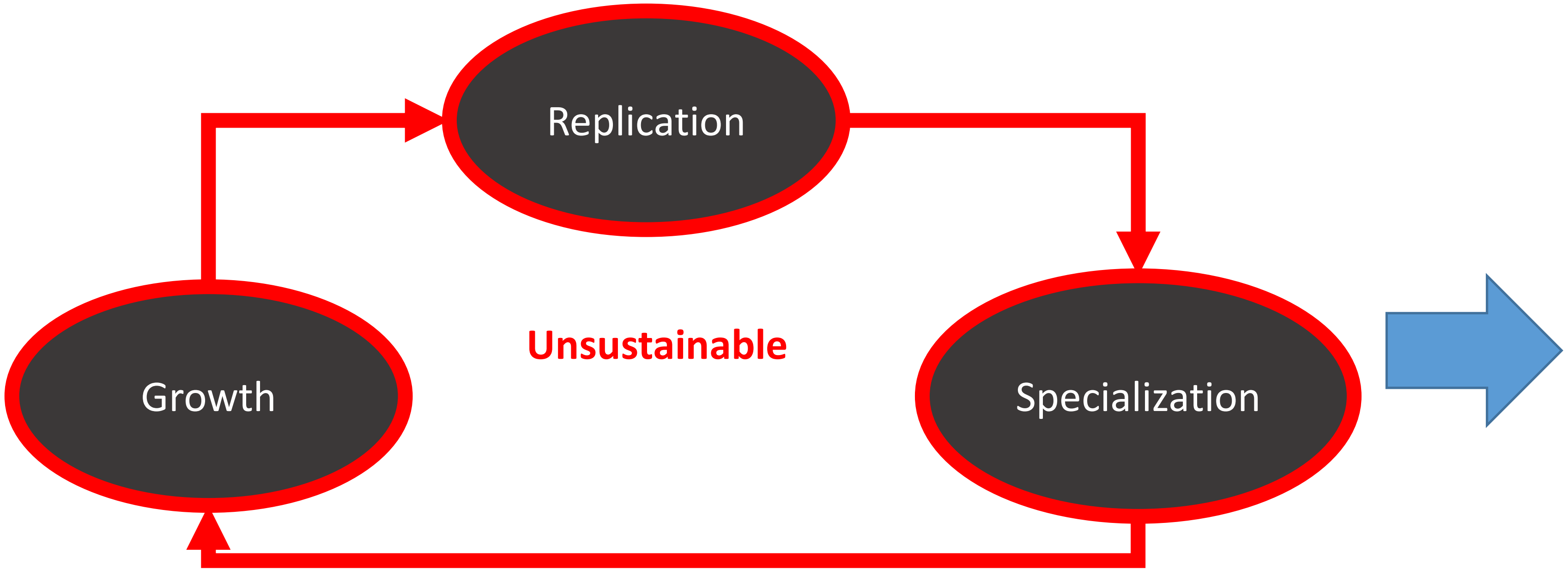
- **Deep Convection**
- Radiative Convective Equilibrium
- Aerosols and Climate
- Biogeochemical Cycles
- Ecological Monitoring
- **Environmental Monitoring**
- Reservoir Modeling
- Exoplanets
- Geomorphology
- Seismic Imaging

- General Circulation
- Climate Dynamics
- Volcanoes and Climate
- Seismic Sensor Networks
- Wind Energy
- **Turbulence**
- Climate Insurance
- Weather Derivatives
- **Hurricanes Storm Surges**
- Earthquakes, Tsunamis

**We pick a few no more than four, depending on participant interest, and drill down**

# A Cautionary Note

- Sustainability Principles and Practice (Selin, Ravela et al. MIT Summer course)
  - Complex Systems Science approach




Increase redundancy in the ecosystem, aim for a lower specialization fraction to improve sustainability