Welcome to
instats

The Session Will Begin Shortly
(At the top of the hour, Eastern USA time)

1

START

2

Nonlinear Time Series Analysis, Part II: Modeling and Phenomenology

Barney Ricca
Lyda Hill Institute for Human Resilience
University of Colorado Colorado Springs

3

Seminar Overview

- Day 1
  - Session 1 – Overview of Phenomenology
  - Session 2 – Dynamical Systems Analysis
- Day 2
  - Session 3 – Sparse Identification of Nonlinear Dynamics
  - Session 4 – Dynamic Mode Decomposition
- Day 3
  - Session 5 – Hidden Markov Models
  - **Session 6 – Machine Learning Approaches**
- Day 4
  - Session 7 – Putting it All Together: Lorenz
  - Session 8 – Putting it All Together: Infectious Diseases

4

Session Outline

- K-means clustering
- Machine Learning Taxonomy
- Neural Networks
  - Terminology and structure
  - Tradeoffs
- Technology
  - Python (through R)
- Choosing a model

5

K-means Clustering

- We've seen this (session 3)
- Iterative approach
  - Pick random starts
    - Avoid overfitting and find optimal model
  - Choose updating rule
  - Choose loss function
  - Let it run

6

Machine Learning

- ML: A subfield of artificial intelligence
  - Roughly: anything that includes back-propagation
- Learning approaches
  - Supervised: Correct answers are included in training
  - Unsupervised: Correct answers are not included in training
  - Reinforcement: Achieve a goal (e.g., beat an opponent)
- All the usual statistical tasks
- Our focus: (Recursive) Neural networks, supervised and unsupervised learning, forecasting

7

Technology Choices

- Some (not-too-complicated) approaches in R
  - K-means clustering, support vector machines (SVM)
  - Single hidden layer neural networks using *forecast::NNETAR()*
- Most work done in Python
  - Tensorflow, Keras

8

Neural Networks

- Mimic human brain neurons
  - Lots of inputs
  - Combine
  - Threshold
  - Output
  - Feedback and back-propagation to evolve
- Networks
  - Lots of (simple) neurons
  - Connections

9

NN Pros

- Mimic (sort of) human brain
- Captures complex relationships
- Little up-front from user
  - Domain expertise is helpful
- Forecasting models

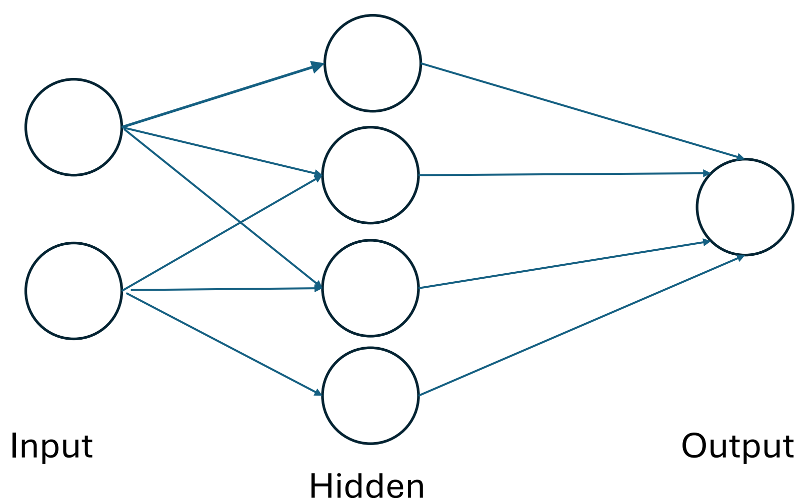
10

NN Cons

- Black box
  - Very big black box
- Not very phenomenological
  - Although, see *interpretable AI*
- Inconclusive performance
- Can mistake noise for signal
  - Test v. train (helps avoid overfitting)
- Long workflows

11

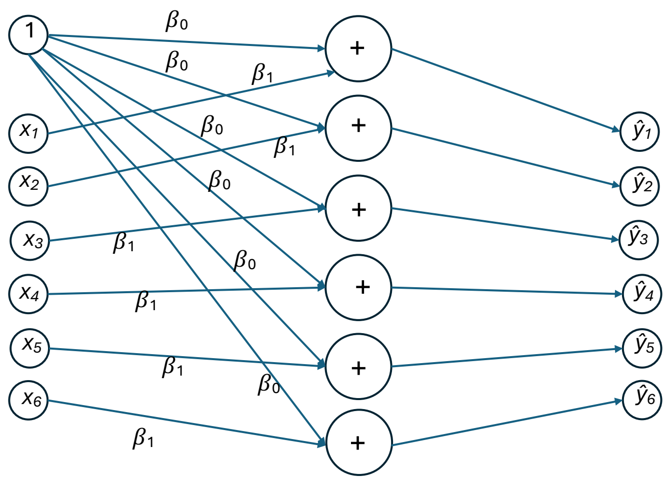
Models



12

Linear Regression (Machine Learning Version)

- Sample architecture
  - A gazillion types
- Usual nomenclature
  - β_1 are *weights*
  - β_0 is *bias* (a weight)
  - *x* are data
  - *1* is for the intercept
  - ŷ are the predictions
- Pick some random numbers for bias and weights
- *Feedforward* (a.k.a., calculate) to get $\hat{y}$ from *x*



13

Linear Regression (via ML)

- Almost certainly, $\hat{y}$ are wrong
- *Back propagate* to adjust β_1 and β_0
- How?
  - Iteration approach
    - Many possibilities…this is the second place with a gazillion possibilities
- *Gradient descent*
  - Mean-square error
  - Calculus to find the "slope" in the β_1 and β_0 directions
  - Solve resulting equations for change in β_1 and β_0
  - Update and try again
  - Repeat until convergence (we hope!)

14

Explication Examples

- Linear Regression
  - Without ordinary least squares
  - One hidden layer
- Classification
  - Try this with k-means clustering!

15

General Workflow

- Pre-processing
  - Derived predictors
  - List-wise remove NA (or impute)
  - Scale!
  - Remove trend (and seasonality)
  - Test v. train data
- Network Construction
  - Number of input nodes, hidden layers, nodes/hidden layer, activation function, output nodes
- Run It (and refine it)

16

Working with Time Series

- Pre-processed data?
  - Maybe
- CCM?
  - Maybe
- Use PCA embedding approach?
  - Create delayed time-series for many delays
  - Maybe?
- Can include individual traits

17

Working with Time Series: Deep Learning

- *Deep learning* = more than 3 layers
  - Tensorflow is the way to go
- Recurrent Neural Networks
  - Common for time series
  - But…
    - Recursion introduces memory which can produce a “vanishing gradient” (as in gradient descent)
  - Long Short-Term Memory
    - Overcomes the “vanishing gradient problem” of RNN

18

Examples

- Lorenz (Part 1)
- Lorenz (Part 2…oops!)
- Mayport (preprocessed)

19

Interpretable Machine Learning

- Can we get past the black box?
  - Interpret the parameters?
  - Closer to phenomenology?
- Potential Ideas
  - General post-hoc methods
  - TimberTrek
    - Rashomon set…ugh…

20

Questions

21

STOP

22

Next session @ UTC 1600

Tomorrow in most time zones

23