Introduction to Machine Learning

Mark A. Austin

University of Maryland

austin@umd.edu

ENCE 688P, Fall Semester 2021

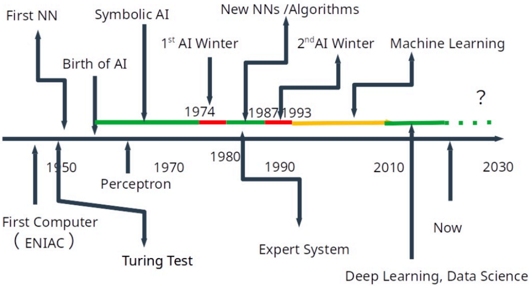
October 16, 2021

Overview

Part 02

Machine Learning Capabilities

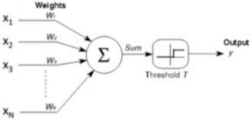
A Brief History



- 1943: First neural networks invented (McCulloch and Pitts)
- 1958-1969: Perceptrons (Rosenblatt, Minsky and Papert).
- 1980s-1990s: CNN, Back Propagation.
- 1990s-2010s: SVMs, decision trees and random forests.
- 2010s: Deep Neural Networks and deep learning.

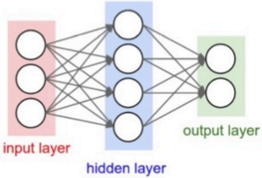
Machine Learning Capabilities (1980-1990)

Expressive Power of a Neural Network

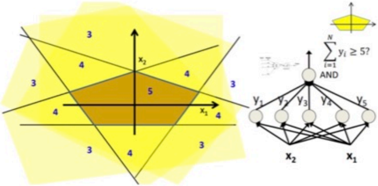


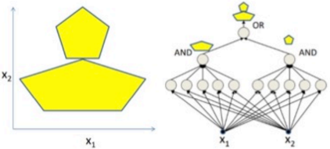
$$y = \begin{cases} 1 & \text{if } \sum_{i=1}^{d} w_i x_i \geq T \\ 0 & \text{else} \end{cases}$$

Neural Network with Single Hidden Layer



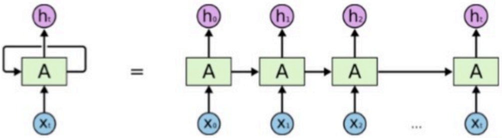
Approximation of Functions / Boolean Logic





Machine Learning Capabilities (1997-2014)

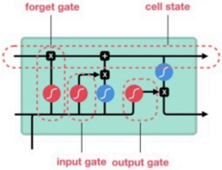
Recurrent Neural Networks (RNN): Learn sequences in data streams (text, speech)



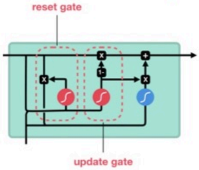
Hidden state "h" serves two purposes:

- Make an output prediction.
- Represent features in the previous steps ....

Long Short-Term Memory (1997)



Gated Recurrent Units (2014)



Key Features of LSTM:

- Standard RNN suffers from vanishing gradients for modeling of long-term dependencies.
- LSTM gives cells the ability to remember values for long periods of time.
- Gates regulate the flow of information in / out of the cell, and what should be remembered or discarded.

Applications:

- Time series prediction.
- Time-series anomaly detection.

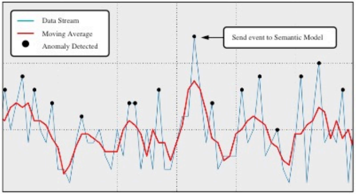
Machine Learning Capabilities (1997-2014)

Learning Streams of Text

- Download complete works of Shakespeare (5.4 million characters)
- Train machine to remember text.
- Write new Shakespeare!

```
Sample after 1000 Iterations
----------------------------

  OTHERS. Allow up, whe, Will a moutice,
  Reason thot boy Eckents alliem; comn jusgius,
  And encarnembover hath last)!
  TASTOR. When no; devires at though me beligo jody?
  MANCELOT. Trom juther and bur itnot spock,
  That as take have wendisho

Sample after 2000 Iterations
----------------------------

  VIRGTARD. Walk this boy and door as am 'stone!
  NTRURLEN. Being entainure af Haton.
  His that by much that, I mion's now, who make foll the kidl!
  CLOTE. Why which Hamm'd?
  QUICKLY. And stand quast of I; this Pi

Sample after 2500 Iterations
----------------------------

  FORD. Nay, You're, excount: and now did yet.
  PAROLLES. Take DUTBY
  This who is begin Gnoban a bows; but yet which that have be,
  Oll, thou stam, and me not ready withered gids
  And he in the pleasues or pardon us.
  Mer. I pray you, how can, and tu

Sample after 3000 Iterations
----------------------------

  Hume, 'tigoning, dear?
  My lam, an hour, chork'd more in my grain,
  I am thing forent innonion, madam! I thank I do speak you?
  ABESSITGBO. But his bosines, giving to know; foward to
  the distvesail.' The to you well know yea, my lovi.
  SECOND CMONBOR. He needs, for the reforeds are;
```

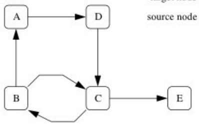
Time Series Anomaly Detection



Time Series Prediction

Machine Learning Capabilities (2014-present)

Traditional Approach to Graph Representation



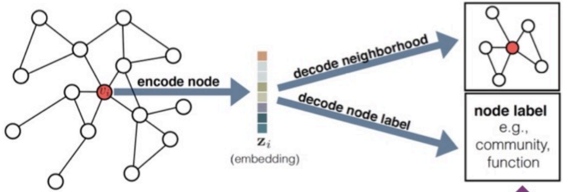
target node → A B C D E
source node → A

	A	B	C	D	E
A	0	0	0	1	0
B	1	0	1	0	0
C	0	1	0	0	1
D	0	0	1	0	0
E	0	0	0	0	0

Graph Analysis

- Connectivity / reachability analysis
- Cycle detection
- Traversal problems
- Shortest path problems
- Traceability problems (MBSE)
- Matching problems
- Topological sort problems
- .....

Graph Embedding Techniques



- Each node in the graph is mapped to a low-dimensional space.
- Goal is to preserve local linkage structure (not global structure).
- Each dimension corresponds to a community in the network.

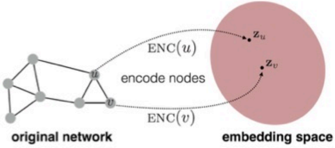
Graph Analytics

- Node Classification
- Node Clustering
- Anomaly Prediction
- Attribute Prediction
- Link Prediction
- Recommendation
- Etc ...

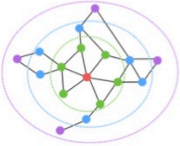
Captures semantics in domain application

Machine Learning Capabilities (2014-present)

Graph-to-Embedding Space Transformation



Graph Reachability



- **Red:** Target node
- **Green:** 1-hop neighbors
  - **A** (i.e., adjacency matrix)
- **Blue:** 2-hop neighbors
  - $\mathbf{A}^2$
- **Purple:** 3-hop neighbors
  - $\mathbf{A}^3$

Goal: Design encoder so that similarity in embedding space is closely approximates similarity in original network.

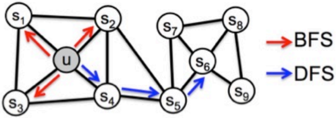
Graph Embedding Vector Design: node2vec, DeepWalk, ...

Node2vec: Combine two strategies:

BFS: Breadth First Search provides a local view of graph neighborhood.

DFS: Depth First Search provides a global view of the neighborhood.

Encoder is just a simple embedding vector lookup.



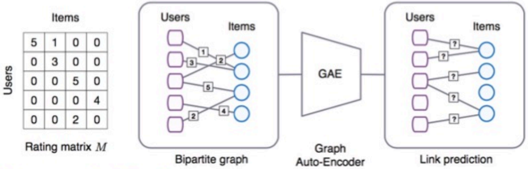
$N_{BFS}(u) = \{ s_1, s_2, s_3 \}$ Local microscopic view

$N_{DFS}(u) = \{ s_4, s_5, s_6 \}$ Global macroscopic view

Machine Learning Capabilities (2014-present)

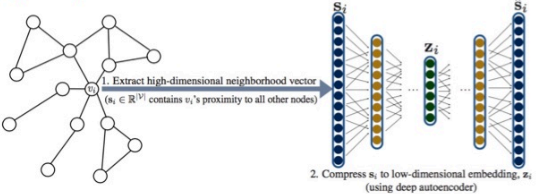
Graph Auto-encoder Link Prediction

Link prediction procedures can be used in new types of system validation / verification.



Deep Graph Auto-encoder Design

Requirements traceability needs arbitrarily large levels of reachability – first order neighbors, second-order neighbors, etc.



References

- Austin M.A., Delgoshaei P., Coelho M. and Heidarinejad M. , Architecting Smart City Digital Twins: Combined Semantic Model and Machine Learning Approach, Journal of Management in Engineering, ASCE, Volume 36, Issue 4, July, 2020.
- Coelho M., and Austin M.A. , Teaching Machines to Understand Urban Networks, The Fifteenth International Conference on Systems (ICONS 2020), Lisbon, Portugal, February 23-27, 2020, pp. 37-42.
- Bhiksha R., Introduction to Neural Networks, Lisbon Machine Learning School, June, 2018.
- Lu T., Fundamental Limitations of Semi-Supervised Learning, MS Thesis in Mathematics in Computer Science, University of Waterloo, Canada, 2009.
- Van Engelen J.E., and Hoos H.H., A Survey on Semi-Supervised Learning, Machine Learning, Vol. 109, 2020, pp. 373-440.