SNS COLLEGE OF TECHNOLOGY

An Autonomous Institution

Coimbatore-35

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A++' Grade
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

19ECT303-ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

III YEAR/ V SEMESTER

1

UNIT 5 – DEEP LEARNING

5.5 Recursive Neural Networks

OUTLINE

- Recap
- Recursive Neural Networks
- Computational graph of a Recursive Network
- Advantage of Recursive over Recurrent Nets
- Need for Recursive nets in NLP
- Recurrent Model for NLP
- Recursive Models for NLP
- Activity
- Advantage of Recursive Model for NLP
- Structure of the Tree
- Variants of Recursive Net idea
- Recursive Neural Networks
- Unrolling Recurrent and Tree Nets
- Advantage of Recursive Nets
- Conclusion-Summary

2

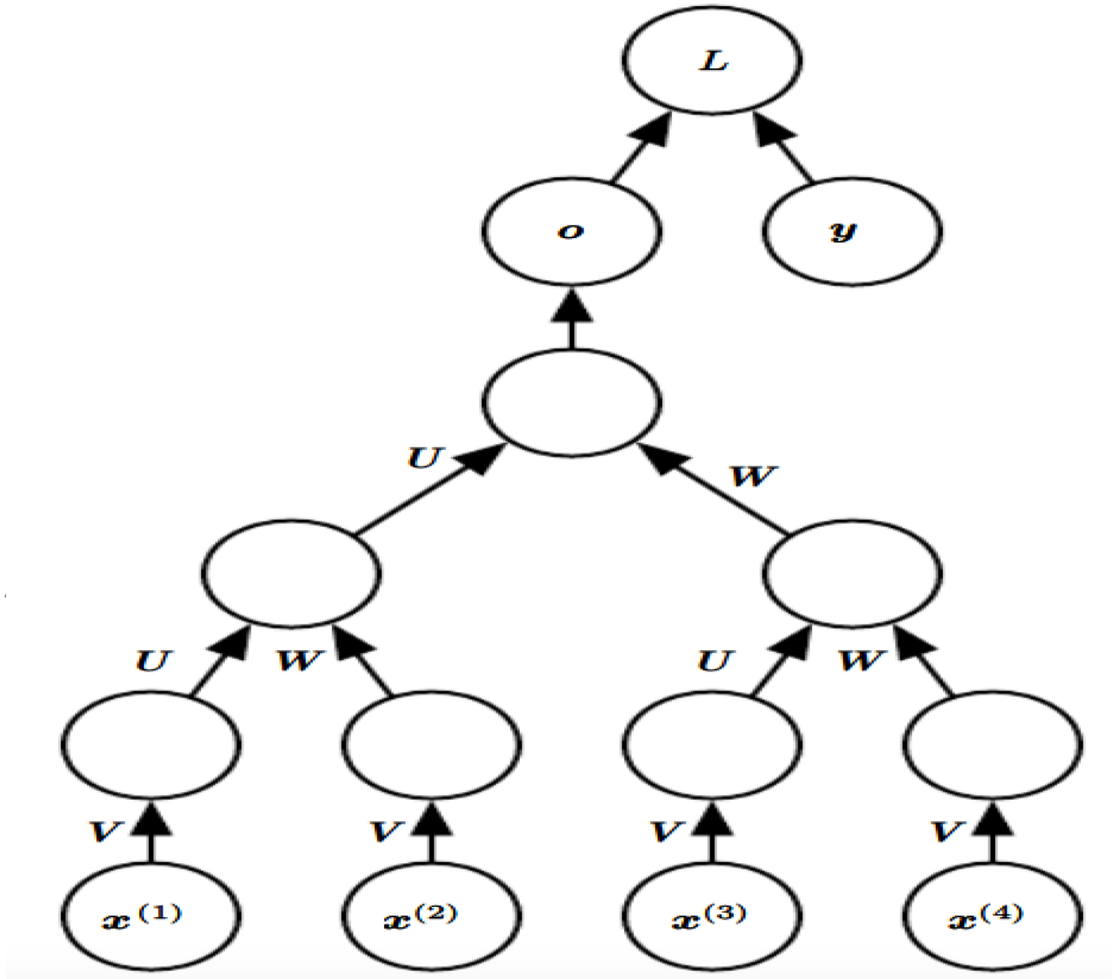
Recursive Neural Networks

- They are yet another generalization of recurrent networks with a different kind of computational graph
- It is structured as a deep tree, rather than the chain structure of RNNs
- The typical computational graph for a recursive network is shown next

Recursive Neural Networks

Computational graph of a Recursive Network

- It generalizes a recurrent network from a chain to a tree
- A variable sequence ***x(1),x(2),,x(t)*** can be mapped to a fixed size representation (the output ***o), with*** a fixed set of parameters (the weight matrices *U,V,W)*
- Figure illustrates supervised learning case in which target ***y is*** provided that is associated with the whole sequence



Recursive Neural Networks

Advantage of Recursive over Recurrent Nets

- For a sequence of the same length τ, the depth (measured as the no. of compositions of nonlinear operations) can be reduced from τ to *O(log τ), which might help deal with long-term* dependencies
- An open question is how best to structure the tree[5]

Recursive Neural Networks

Need for Recursive nets in NLP

- Deep learning based methods learn low-dimensional, real valued vectors for word tokens, mostly from a large data corpus, successfully capturing syntactic and semantic aspects of text
- For tasks where the inputs are larger text units,

e.g., phrases, sentences or documents, a compositional model is first needed to aggregate tokens into a vector with fixed dimensionality that can be used for other NLP tasks

- Models for achieving this fall into two categories: recurrent models and recursive models

Recursive Neural Networks

Recurrent Model for NLP

- Recurrent models deal successfully with time series data
- The recurrent models generally consider no linguistic structure aside from the word order
- They were applied early on to NLP by modeling a sentence astokens processed sequentially and at each step combining the current token with previously built embeddings
- Recurrent models can be extended to bidirectional ones from both left to right and right to left
- These models consider no linguistic structure aside from word order

Recursive Neural Networks

Recursive Models for NLP

- Recursive neural models (also referred to as tree models) by contrast are structured by syntactic parse trees
- Instead of considering tokens sequentially, recursive models combine neighbors based on the recursive structure of parse trees, starting from the leaves and proceeding recursively in a bottom-up fashion until the root of the parse tree is reached
- **Ex**: for the phrase the food is delicious, following the operation sequence ((the food) (is delicious)) rather than the sequential order (((the food) is) delicious)

ACTIVITY

9

Recursive Neural Networks

Advantage of Recursive Model for NLP

- They have the potential of capturing long-distance dependencies
- Two tokens may be structurally closer to each other even though they are far away in word sequence
- Ex: a verb and its corresponding direct object can be far away in terms of tokens if many adjectives lie inbetween, but they are adjacent in the parse tree
- However parsing is slow and domain dependent
- See performance comparison with LSTM on four NLP tasks at https://nlp.stanford.edu/pubemnlp2015_2_jiwei.pdf

Recursive Neural Networks

Structure of the Tree

- One option is to have a tree structure that does not depend on the data, such as a balanced binary tree
- In some application domains, external methods can suggest the appropriate tree structure
- Ex: when processing natural language sentences, the tree structure for the recursive network can be fixed to the structure of the parse tree of the sentence provided by a natural language parse

11

- Ideally, one would like the learner itself to discover and infer the tree structure that is appropriate for any given input

Recursive Neural Networks

Variants of Recursive Net idea

- Associate data with a tree structure and associate inputs and targets with individual nodes of the tree
- The computation performed for each node does not have to be the artificial neuron computation (affine transformation of all inputs followed by a monotone nonlinearity)

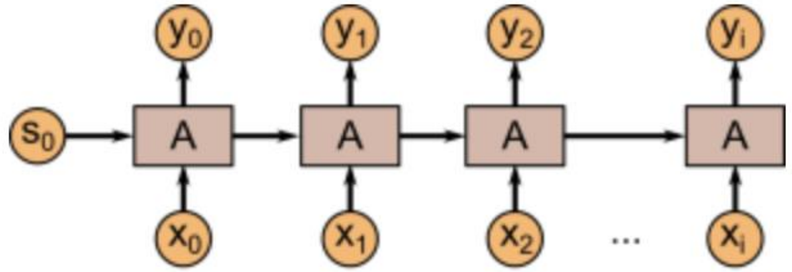
12

- Can use a tensor operations of bilinear forms
- Previously found useful to model linear relationships between concepts when the concepts are represented by continuous vectors

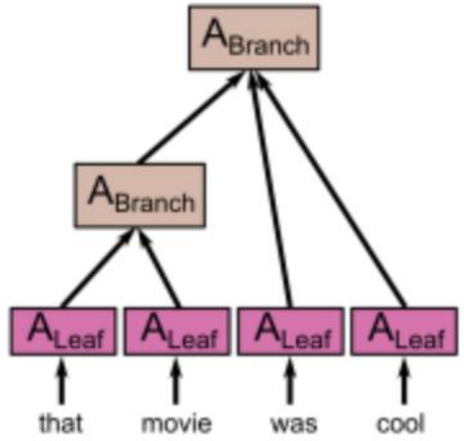
Recursive Neural Networks

Recursive Neural Networks

- Recursive neural networks are also called Tree Nets
- Useful for learning tree-like structures
- They are highly useful for parsing natural scenes and language



Recurrent Neural Net



Recursive Neural Net

Recursive Neural Networks

Unrolling Recurrent and Tree Nets

- In RNNs, at each time step the network takes as input its previous state s(t-1) and its current input x(t) and produces an output y(t) and a new hidden state s(t).
- TreeNets, on the other hand, don't have a simple linearstructure like that.
- With RNNs, you can 'unroll' the net and think of it as a large feedforward net with inputs x(0), x(1), …, x(T), initial state s(0), and outputs y(0),y(1),…,y(T), with T varying depending on the input data stream, and the weights in each of the cells tied with each other.
- You can also think of TreeNets by unrolling them – the weights in each branch node are tied with each other, and the weights in each leaf node are tied with each other.

Recursive Neural Networks

Advantage of Recursive Nets

- The advantage of Recursive Nets is that they can be very powerful in learning hierarchical, tree-like structure.
- The disadvantages are, firstly, that the tree structure of every input sample must be known at training time.

15

ASSESSMENT
SUMMARY & THANK YOU

16