Introduction to Neural Networks for Natural Language Processing

Advanced Network Architectures

Advanced Network Architectures

- Recap: Natural Language in Neural Networks.
- Structured Problems in Neural Networks
- Convolutional Neural Networks.
- Recurrent Neural Networks.
- Notebook Introduction: Structure in Images and Text.

Recap: Natural Language in Neural Networks

- At this point we know **how to design neural networks** that can solve **binary and multi-class classification problems**.

- We have also learned of two ways to represent text: **Bag-of-Words and Word Embeddings**.

- We saw that word embeddings deal with natural language problems by **capturing word similarities and relations**.

Recap: Natural Language in Neural Networks

- However, the way we have used both Bag-of-Words and Word Embeddings **ignores the structured order of words**!

- In some cases, the existence of a general sense of meaning is captured by **those 'aggregate' representations**, which **are reasonably accurate and fast**!

- They are a great solution to try early on, if your problem is suited for it… **But what if it is not**?

Structured Problems in Neural Networks

- Natural Language problems can get pretty complicated, for instance, consider **identifying entities within sentences**:

As the dog calmly walked down the 7th, Bernard Fitzgerald whistled a song from The Doors.

The doors stayed open until the 7th on the Bernie Fitzgerald Memorial.

Structured Problems in Neural Networks

As the dog calmly walked down the 7th {place} [avenue], Bernard Fitzgerald {person} whistled a song from The Doors {organization}.

The doors stayed open until the 7th {date} [of May] on the Bernie Fitzgerald Memorial {place}.

- **The meaning of each word depends on the context**, some of which may appear beyond the sentence.

- The same words **may be tagged differently across sentences, or not tagged at all**.

Structured Problems in Neural Networks

- **Entity Tagging** is a problem in which **structure matters**: the order of words, one after the other, affect our outcome!

- How can a neural network deal with structure? **Neural networks always take inputs of the same size**, but sentences have different lengths!

- Standard Deep Neural Networks need to be ‘extended’ to manage structured inputs effectively: **enter convolutional and recurrent networks!**

Convolutional and Recurrent Neural Networks

- Both **convolutional and recurrent networks** tackle with structured inputs, though in different ways.

- The fundamental characteristic of both is that they **learn local properties that can be extracted across the whole structure**!

- Convolutional and recurrent networks are powerful, but **they can have many parameters and, thus, require more time and data to train**!

Convolutional and Recurrent Neural Networks

- Convolutional neural networks **take several inputs of the same type at a time** and pass them through a ‘standard’ neural network **to learn a ‘filter’**.

- The inputs taken by a convolutional network **follow some structure**: a patch of pixels, a group of contiguous words, etc.

- The learned filters are applied over each substructure in the instance and **extract higher level information** such as **shapes from every patch in an image** or **meaning in triads of words in a sentence**.

Convolutional and Recurrent Neural Networks

- Recurrent neural networks **take an input at a time, keeping a 'memory' of what it has seen before.**

- The inputs **seen at every step in time are always of the same type**: a word embedding, an audio spectrogram or a video frame!

- Recurrent networks learn (1) **how to best keep track of that it has seen** so it can (2) **produce an appropriate answer to the problem**.

Notebook Introduction: Structure in Images and Text

- At this point we have recapped **the issue of learning on structured data such as text**, and how **our previous models were too simplistic**.

- We have also introduced advanced architectures that work with, **namely convolutional and recurrent networks**. However, we still don't know to use them!

- This is the objective of the fourth and last notebook of the course: **to learn to use convolutional and recurrent networks to tackle problems… Let's go!**

Introduction to Neural Networks for Natural Language Processing