

Deep Learning for Text 1

Applied Text Mining

Ayoub Bagheri

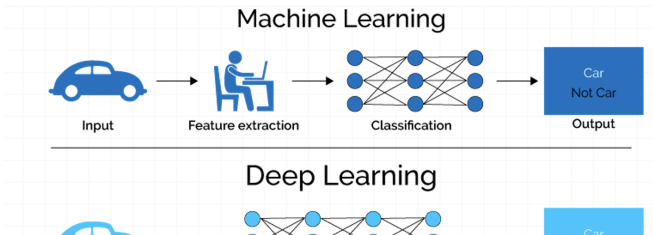
Lecture plan

1. Feed-forward neural networks
2. Recurrent neural networks
 - 2.1 SRN
 - 2.2 LSTM
 - 2.3 Bi-LSTM
 - 2.4 GRU

What is Deep Learning (DL) ?

A machine learning subfield of learning representations of data. Exceptional effective at learning patterns.

Deep learning algorithms attempt to learn (multiple levels of) representation by using a hierarchy of multiple layers.



Recurrent Neural Networks

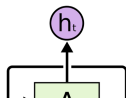
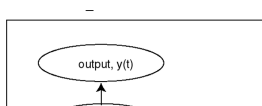
- ▶ Another architecture of NN
- ▶ RNN for LM

Recurrent Neural Network (RNN)

- ▶ Add feedback loops where some units' current outputs determine some future network inputs.
- ▶ RNNs can model dynamic finite-state machines, beyond the static combinatorial circuits modeled by feed-forward networks.

Simple Recurrent Network (SRN)

- ▶ Initially developed by Jeff Elman (*"Finding structure in time,"* 1990).
- ▶ Additional input to hidden layer is the state of the hidden layer in the previous time step.



LSTM

Vanishing gradient problem

Suppose we had the following scenario:

Day 1: Lift Weights

Day 2: Swimming

Day 3: At this point, our model must decide whether we should take a rest day or yoga. Unfortunately, it only has access to the previous day. In other words, it knows we swam yesterday but it doesn't know whether had taken a break the day before.

Therefore, it can end up predicting yoga.

- ▶ Backpropagated errors multiply at each layer, resulting in exponential decay (if derivative is small) or growth (if derivative is large).
- ▶ Makes it very difficult train deep networks, or simple recurrent networks over many time steps.
- ▶ LSTMs were invented, to get around this problem.

<https://towardsdatascience.com/>

Long Short Term Memory

Summary

Summary

- ▶ Deep learning can be applied for automatic feature engineering
- ▶ Recurrent neural networks are ideal for sequential data such as text

Practical 6