Using Neural Networks for Modeling and Representing Natural Languages

Tomas Mikolov
Facebook A.I. Research
One Hacker Way, Menlo Park
California, US
tmikolov@fb.com

Artificial neural networks are powerful statistical models that have been shown to provide excellent results in a number of domains. In the last few years, the computer vision and automatic speech recognition communities have been heavily influenced by these techniques. Applications to problems that involve natural language, such as machine translation or computational semantics, are becoming mainstream in the NLP research.

This tutorial aims to introduce the basic concepts and provide intuitive understanding of neural networks, including the very popular field of deep learning. This should help the researchers who are entering this field to quickly understand the major tricks of the trade.

The structure of the tutorial is as follows:

Basic machine learning applied to natural language

- n-grams and bag-of-words representations

- logistic regression, support vector machines

Introduction to neural networks

- architecture of neural networks: neurons, layers, synapses

- activation function

- objective function

- training: stochastic gradient descent, backpropagation, learning rate, regularization

- multiple hidden layers and intuitive explanation of deep learning

Distributed representations of words

- basic application of neural networks for obtaining vector representation of words

- linguistic regularities in the word vector space

- word analogy tasks with vector representations

- representations of phrases and sentences

- simple application to machine translation of words and phrases

Neural network based language models

- feedforward and recurrent neural net architectures for language modeling

- class based softmax, hierarchical softmax

- joint training with maximum entropy model

- recurrent model with slow features

- application to language modeling, speech recognition, machine translation

Tips for future research

- understanding the current research culture

- hints how to recognize good papers and ideas

- promising future directions

Resources

- introduction to open-source software: RNNLM toolkit, word2vec and other tools

- links to large text corpora, pre-trained models

- benchmark datasets for advancing the state of the art