Deep Learning for Natural Language Processing

Cornelia Caragea

Computer Science
University of Illinois at Chicago

Logistics

Today

Lecture Structure

- Logistics
- What is this course about?

Logistics

- Instructor: Dr. Cornelia Caragea
  - Email: cornelia@uic.edu
  - WWW: https://www.cs.uic.edu/∼cornelia/
  - Office: 3-190E Daley Library
- Class structure:
  - Asynchronous: recorded lectures posted on Blackboard
- Office hours:
  - Synchronous: Tuesday 1:00pm - 2:00pm or by appointment.
  - Zoom Link:

    https://uic.zoom.us/j/91426290683?pwd=RzJqd3N4SnZJdzlXYytNZGkwM0FQUT09
- Prerequisites: Linear algebra and calculus, machine learning, basic probability and statistics.
- Class material: Available on Blackboard.

Recommended Textbooks

- **Neural Network Methods for Natural Language Processing** (Synthesis Lectures on Human Language Technologies)
  by Yoav Goldberg.
- **Deep Learning**
  by Ian Goodfellow, Yoshua Bengio and Aaron Courville.
  Online version available at:
  https://www.deeplearningbook.org/.
- **Introduction to Deep Learning**
  by Eugene Charniak.
- **Research papers.**

Course Description

The course will cover the foundations of deep learning, with focus on different types of modern neural network (such as convolutional neural networks, recurrent neural networks, long short-term memory networks, attention and transformers). Furthermore, the course will also cover tools and software available for building and training deep neural networks.

Course Objectives

- Introduce students to deep learning and discuss types of deep neural networks
- Study concepts such as regularization in the context of deep learning
- Identify applications where deep learning excels, and discuss what type of network is appropriate for what application
- Identify limitations of existing deep learning approaches
- Explore tools available for deep learning
- Teach students about how to stay up-to-date with the latest research in deep learning

Tentative Topics

- Basics of Machine Learning
- Word Embeddings (word2vec, Glove, FastText)
- Word Window Classification and Neural Networks
- Language Models and Recurrent Neural Networks
- Recurrent Neural Networks Variants (LSTMs and GRUs)
- Sequence-to-Sequence Models and Attention Mechanisms
- Natural Language Generation - Evaluation and Metrics
- Convolutional Neural Networks for Sentence Classification
- Transformers and Self-Attention
- Bidirectional Encoder Representations from Transformers
- Multi-Task Learning and Transfer Learning
- Adversarial Learning
- Auto-encoders

Applications

NLP Applications:

- Text Classification
- Sentiment Analysis and Emotion Detection
- Word Classification (Named Entity Recognition)
- Neural Machine Translation
- Twitter Mining
- Information Extraction - Keyphrase Extraction
- Question Answering

Software Resources

- Python (programming language)
  https://www.python.org/
- Machine Learning in Python - scikit-learn
  http://scikit-learn.org/stable/
- TensorFlow
  https://www.tensorflow.org/
- PyTorch
  https://pytorch.org/
- AllenNLP: An open-source NLP research library, built on PyTorch
  https://allennlp.org/
- Keras: The Python Deep Learning library
  https://keras.io/
- Gensim
  https://radimrehurek.com/gensim/

Computational Resources

- Amazon Web Services (AWS)
  http://aws.amazon.com
- Google Cloud
  https://cloud.google.com

Word Embeddings Resources

- word2vec
  https://code.google.com/archive/p/word2vec/
- Glove
  https://nlp.stanford.edu/projects/glove/
- FastText
  https://fasttext.cc/
- ELMo
  https://allennlp.org/elmo
- BERT
  https://arxiv.org/abs/1810.04805

Course Grading

Section	Weight
Reading Assignments	20%
Programming Assignments	30%
Quizzes	20%
Final Project	30%

Class Project

Goal: teach you how to apply deep learning to real-world tasks.

- **Application project**. Find a problem/application of interest to you, identify algorithms that are best suited for that problem, and use the algorithms to solve the problem.
- **Algorithmic project**. Find a problem/application of interest to you, identify limitations of existing algorithms, and develop a new learning algorithm, or a novel variant of an existing algorithm, to solve it.

Submission Policies

- Assignments will be submitted online through Blackboard.
- Late submissions are not encouraged. I will accept late submissions, but there will be grading penalties (and at most three day late submissions).

Collaboration Policies

- You are encouraged to discuss the course material, concepts, and assignments, but you must write your answers independently.
- For each assignment, you are required to list students with whom you have discussed the assignment.
- Your submission should reflect your own knowledge and you should be able to reproduce the material you turn in at any time.
- Sharing answers will not be tolerated.
- Plagiarism will not be tolerated either.
- Appropriate citations for any external sources used in your work are mandatory. Never use sentences or phrases taken directly from a paper you are reviewing.

Who Am I?

- Cornelia Caragea, Associate Professor in CS@UIC
  <cornelia@uic.edu>
- Artificial Intelligence, Information Retrieval, Natural Language Processing, Machine Learning
  - Information Extraction - Keyphrase Extraction
  - Mining Scientific Documents
    - Building Scholarly Knowledge Graphs
  - Mining Social Media in Disaster Events
    - Domain Adaptation approaches for Disaster Management
  - Image Privacy Prediction

Next...

- What is Natural Language Processing?