CptS 434/534 Neural Network Design and Applications (aka Deep Learning)
Instructor: Yan Yan (yan.yan1@wsu.edu)

Spring-2021 Logisics:
Class timings - MWF via Zoom (Jan 19 to April 30)
Instructor office hours - Tue 4-5pm
Class announcements - Piazza
Homework/Project submissions and grades - Blackboard

Course Overview:

This course provides an introduction to deep neural networks (DNNs) including mathematical, statistical, and computational challenges to learn good representations from raw data such as images, text, and graphs. Representation learning is one of the most fundamental challenges for both supervised learning (e.g., classifying an image as cat or dog) and reinforcement learning (e.g., learning sequential decision-making policies to play board games such as Go). The course will start with a motivation for DNNs over traditional machine learning based on hand-designed features, different abstraction of DNNs to fully understand their utility, computational challenges to train DNNs, and key principles for learning parameters of DNNs from data (inductive bias, backpropagation, regularization, batch normalization, stochastic gradient descent). Subsequently, specific DNNs for different modalities of data such as Convolutional Neural Networks (CNNs) for image data, Recurrent Neural Networks (RNNs) for sequence data, and Graph Neural Networks (GNNs) for graph data will be covered. Finally, the course will touch upon some advanced concepts such as Generative Adversarial Networks (GANs), Neural Architecture Search (NAS), deep reinforcement learning, and open research problems in deep learning.

Learning Objectives of the Course:

1) Understand the benefits of DNNs for learning representations from raw data over traditional ML algorithms.

2) Understand the computational and statistical challenges of learning DNNs from data.

3) Understand the key principles/ideas for efficiently learning parameters of DNNs from data.

3) Understand different types of DNNs (feed-forward, CNNs, RNNs, and GNNs) and their inductive biases.

4) Should be able to apply DNNs to solve supervised learning problems over images, text, and graphs.

5) Should be able to read research papers in deep learning and understand the issues raised by them.

Course Topics:

1. Supervised machine learning

- Supervised learning setup (training data, testing data)
- Abstract machine learning algorithm (representation, evaluation or loss function, optimization)
- Empirical risk minimization
- Maximum likelihood estimation
- Empirical error vs. Generalization error
- Overfitting vs. Underfitting
- Bias and variance trade-offs

2. Introduction to deep neural networks

- Motivation (compared to hand-crafted features)
- Why DNNs have become popular since 2010? (large data, computing hardware, and algorithmic innovations)
- Feedforward networks; multi-layer perceptron vs. DNNs
- Role of activation function and non-linearity
- Differentiable programming paradigm; deep learning as an instance of differentiable programming
- Different interpretation of DNNs

3. Key principles for learning DNNs from data

- Inductive bias in terms of architecture with example illustrations
- Supervised training loop and its instantiation for DNNs
- Backpropagation algorithm
- Automatic differentiation and illustrative examples; static versus dynamic computation graphs
- Optimization problem, challenges and SGD algorithms (e.g., RMSprop, Adam) and variants
- Regularization
- Batch normalization to address the challenge of exploding and vanishing gradients
- Hyper-parameters and hyper-parameter optimization tools (e.g., AutoML, Bayesian optimization)

4. Convolutional Neural Networks (CNNs)

- Why vanilla neural networks are not well-suited for images?

- Key insight behind CNNs and the corresponding inductive bias of convolution layer
- Convolution layer and illustrations
- Pooling layer and illustrations
- Fully-connected layer and illustrations
- Understanding CNNs via visualization (e.g., feature map, saliency maps, etc)

5. Recurrent Neural Networks (RNNs)

- Key insight behind RNNs and the corresponding inductive bias
- Backpropagation through time
- Issues with vanilla RNNs
- GRUs and LSTMs

6. Graph Neural Networks (GNNs)

- CNNs and RNNs work over regular structures (grid and sequence), but would not work for irregular structures such as graphs
- What is network representation learning and why is it important?
- Learning low-dimensional embeddings of nodes in complex networks (e.g., DeepWalk and node2vec)
- Techniques for deep learning on network/graph structed data. Key idea of message-passing and neighborhood aggregation.

7. Generative Adversarial Networks (GANs)

- What are generative models and why are they useful?
- What are GANs and why they are useful?
- Role of generator and discriminator
- Minimax game and training difficulties
- Practical guidelines

8. Neural Architecture Search (NAS)

- How to automatically find high-performing neural architectures?
- An instance of automated ML
- Key challenges: structures of varying sizes and evaluation of each candidate is expensive
- Non-differentiable NAS approaches
- Differentiable NAS approaches
- Practical tools and guidelines

9. Deep Reinforcement Learning (DRL) -- Guest lecture by Prof. Jana Doppa

- Supervised learning vs. Reinforcement learning
- Markov Decision Process (MDP) formulation

- Value function and policy
- Two computational questions: policy evaluation and policy optimization
- Exploration and exploitation trade-off for policy learning
- Large state spaces: tabular methods vs. function approximator
- DNNs as function approximators and advantages
- DNN based policy learning methods: policy gradient methods, approximate policy iteration, and advanced methods (e.g., trust-region policy optimization and proximal policy optimization)
- Offline model-based RL and variants

10. Open Challenges

- Requirement of large amounts of data. How can we use weak supervision and self-supervision ideas?
- Causal modeling for higher-level of reasoning
- Meta-learning and curricullum learning
- etc.

Books:
Ian Goodfellow, Yoshua Bengio, and Aron Courville. Deep Learning. MIT Press. Online book: https://www.deeplearningbook.org/
Dive into Deep Learning. Online book with code: https://d2l.ai/

Grading:

4 Assignments (4 x 10% = 40%)
Course project (30%)
1 Midterm Exam (30%)

Assignments will consist of a mix of analytical questions, writing summaries of research papers, and practical implementation of DNNs for different applications. 3 of the 4 assignments will contain practical aspects of CNNs, RNNs, and GNNs.

Students can bring their research to class and apply DNNs to solve those problems. Course projects can be done in teams of at most two students. A course project proposal needs to be submitted by Feb 28, 2021. Each team needs to submit a course project report in a format specified by the instructor (e.g., six pages in a conference paper format). All teams will make presentations on their course projects during the slot of final exam.

There will be an exam in mid-March on topics covered up to that point.