# MAT 6932/23514 Special Topics in Applied Mathematics (Fall 2020)

# Introduction to Deep Learning and Its Mathematical Foundation

**Objective and Description of the Course:**

Deep learning (DL) is a novel methodology currently receiving much attention and has been successfully applied to a large variety of fields, such as data classification, image recognition, speech recognition, natural language understanding, precision medicine, and computational biology. The mathematical and computational methodology underlying deep learning models is very challenging, but it is extremely important for further development of deep learning techniques.

The aim of this course is to prepare our students this modern computational skill for his/her future research. This course will provide basic concepts on machine learning and basic statistical learning theory, introduce the basic structure of artificial neural network (ANN) and convolutional neural network (CNN), the techniques for network architecture design and training, and review several popular and efficient CNNs for supervised regression and classification and their application in computer vision. Moreover, to have better understanding for deep learning from mathematical aspect, the course will introduce some necessary knowledge on convex and noncovex analysis, sub-gradient, optimality characterization for first order optimization and several optimization algorithms for solving certain classes of convex/nonconvex and smooth/nonsmooth optimization problems. In the end we will study how to integrate deep leaning and optimization methods to design CNNs that have better performance with much less parameters and are more interpretable.

The topic of this course is one of the rapid developing fields. There is no textbook available. I will provide some references (recent papers). Students presentations, discussions and projects are expected.

**References:**


- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015).
- https://sites.google.com/site/deeplearningsummerschool/
- http://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf
- M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: A system for large-scale machine learning. In OSDI, volume 16, pages 265–283, 2016.

- J. Adler and O. Oktem, “Learned primal-dual reconstruction,” ¨ IEEE Transactions on Medical Imaging, vol. 37, no. 6, pp. 1322–1332, 2018.
- M. Borgerding, P. Schniter, and S. Rangan, “ AMP-inspired deep networks for sparse linear inverse problems,” IEEE Transactions on Signal Processing, vol. 65, no. 16, pp. 4293–4308, 2017.
- J. R. Chang, C.-L. Li, B. Poczos, and B. V. Kumar. One network to solve them all: solving linear inverse problems using deep projection models. In 2017 ICCV, pages 5889–5898. IEEE, 2017.
- X. Chen, J. Liu, Z. Wang, and W. Yin. Theoretical linear convergence of unfolded ista and its practical weights and thresholds. In NIPS, pages 9061–9071, 2018.
- D. Gilton, G. Ongie, and R. Willett, “Neumann networks for linear inverse problems in imaging,” IEEE Transactions on Computational Imaging, 2019.
- Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q Weinberger. Densely Connected Convolutional Networks. CVPR, pages 2261–2269, 2017.
- Jie Hu, Li Shen, and Gang Sun. Squeeze-and-Excitation Networks. CVPR, 2018.
- K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll. Learning a variational network for reconstruction of accelerated mri data. Magnetic resonance in medicine, 79(6):3055–3071, 2018.
- K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in The IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
- K. He, X. Zhang, S. Ren, and J. Sun. Identity mappings in deep residual networks. In European conference on computer vision, pages 630–645. Springer, 2016.
- T. Meinhardt, M. Moller, C. Hazirbas, and D. Cremers. Learning proximal operators: Using denoising networks for regularizing inverse imaging problems. In ICCV, pages 1781–1790, 2017.
- Saining Xie, Ross B Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated Residual Transformations for Deep Neural Networks. CVPR, pages 5987–5995, 2017.
- Y. Yang, J. Sun, H. Li, and Z. Xu. Deep admm-net for compressive sensing mri. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, NIPS 29, pages 10–18. Curran Associates, Inc., 2016.
- J. Zhang and B. Ghanem. Ista-net: Interpretable optimization-inspired deep network for image compressive sensing. In CVPR, 2018.
- J. Zhang, C. Zhao, and W. Gao. Optimization-inspired compact deep compressive sensing. IEEE Journal of Selected Topics in Signal Processing, 2020.
- K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep cnn denoiser prior for image restoration. In CVPR, pages 3929–3938, 2017.
- Z. Zhang, Y. Liu, J. Liu, F. Wen, and C. Zhu. Amp-net: Denoising based deep unfolding for compressive

- Z. Zhang, Y. Liu, J. Liu, F. Wen, and C. Zhu. Amp-net: Denoising based deep unfolding for compressive image sensing. arXiv preprint arXiv:2004.10078, 2020.
- S. Zhou, Y. He, Y. Liu, and C. Li. Multi-channel deep networks for block-based image compressive sensing. arXiv preprint arXiv:1908.11221, 2019.

- Amir Beck, First-Order Methods in Optimization, MOS-SIAM Series on Optimization, 2017.
- Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. Physica D, 60(1):259-268, 1992.
- A. Beck and M. Teboulle. A fast iterative shrinkage thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences, 2(1):183–202, 2009.
- M. R. Hestenes, Multiplier and gradient methods, J. Optim. Theory Appl., 4 (1969), 303–320
- A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging Vision, 40 (2011), 120–145.

**Meeting Time and Rooms***:*

MWF 4 online (Synchronous Course)

Office Hours: MWF 5 or by appointment

**Arrangement of the Course:***:*

Unit 1: Basic Concepts on Machine Learning (Tentatively week 1-3)

1.1 What is machine learning, types of machine learning, basic machine learning theory: under-fitting and overfitting, bias and variance, sample and computational complexity. (week 1)

1.2. Maximum likelihood estimator (MLE) based data fitting; Bayes rule and maximum a priori (MAP) estimation; (week 1-2)

1.3. Regression and classification as supervised learning: (Tentatively week 2-3)

1.3.1. Linear and nonlinear regression, model accuracy assessment, generalization;(week 2)

1.3.2. Classification: Linear models for classification, logistic regression, support vector machine. (week 3)

Unit 2:  Basic concepts on convex and nonconvex analysis and useful optimization algorithms: (Tentatively week 4-8)

2.1. Convex set, dual space and bidual space,
adjoint transformations, vector and matrix norms;(week 4)

2.2. Convex functions: Definition and basic properties of convex functions, L-smooth function, convex conjugate function,
property and calculation of conjugate function;(week 5)

2.3. Sub-gradient and sub-differential: Definition and basic properties, sub-gradient calculus, duality and optimality conditions, directional derivative; (week 6)

2.4. Several first order optimization algorithm widely used in deep learning: proximal gradient method, gradient decent method, accelerated gradient decent method, stochastic gradient decent method, adaptive stochastic gradient decent methods; (week 7-8)

Unit 3: Deep neural networks (DNNs) (Tentatively week 9-12)

3.1. Architecture and learning algorithm of artificial neural network (ANN): universality theory, computation in forward pass, network training-backward propagation; (week 9-10)

3.2. Architecture and learning algorithm of convolutional neuron network (CNN), convolutional operations, dropout, batch normalization, architecture designing, residual learning, network training (week 10-11);

3.3. Case study and applications. (week 12)

Unit 4: Integration of deep learning and variational methods (Tentatively week 13-16)

4.1. Advantage of integration of deep learning and variational methods, framework of learnable optimization algorithm (LOA); (week 13)

4.2. More first order optimization methods: Alternating direction method of multipliers (ADMM) for equality constrained convex optimization, primal-dual algorithms; (week 13-14)

4.3. Types of LOAs. Case study: Prox-net (ISTA-net); ADMM-net, Primal-dual-net. Gradient-decent-net (variational-net) and their applications. (week 15-16)

**Additional Information:**

**Grading:**
Students will be required to present one to two papers or projects related to the course content. The projects may be related to problems of particular interest to the individual student. Grades will be assigned on the basis of the presentations or projects. Current UF grading policies can be found from the following link https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx.

**Honor Code:** “UF students are bound by The Honor Pledge which states, “We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of

University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: “On my honor, I have neither given nor received unauthorized aid in doing this assignment.” The Honor Code specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.”

**Class Attendance:** “Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx.”

**Accommodations for Students with Disabilities:** “Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://www.dso.ufl.edu/drc/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.”

**Online Evaluations:** “Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at https://evaluations.ufl.edu. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at https://evaluations.ufl.edu/results/.”

**Contact information for the Counseling and Wellness Center:** https://counseling.ufl.edu/, 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

**Diversity:**
I and the department of Mathematics are committed to diversity and inclusion of all students in this course. I acknowledge, respect, and value the diverse nature, background and perspective of students and believe that it furthers academic achievements. It is our intent to present materials and activities that are respectful of diversity: race, color, creed, gender, gender identity, sexual orientation, age, religious status, national origin, ethnicity, disability, socioeconomic status, and any other distinguishing qualities.

UF UNIVERSITY of FLORIDA
The Foundation for The Gator Nation

© 2020 **University of Florida**, Gainesville, FL 32611; (352) 392-3261. Page Updated: August 22, 2020

This page uses **Google Analytics** (**Google Privacy Policy**)