Modern Machine Learning: Simple Methods that Work

Adityanarayanan Radhakrishnan (Adit Radha)

January 11, 2022

Course Information

Description

Over the past decade, interest in machine learning research has spiked drastically, with advancements in deep learning being a significant driving force. Indeed, deep learning has transformed many areas in computer science including computer vision, natural language processing, and reinforcement learning. Unfortunately, given the rapid pace of progress in deep learning, a newcomer looking for a simple set of guiding principles for building machine learning applications can be easily overwhelmed by the nuances of training deep networks. Thus, motivated by recent developments in machine learning, we present a simple class of machine learning methods that are easy to implement and which achieve competitive performance in practice. In particular, our methods rely on the recently established equivalence between kernel regression and infinite width neural networks given by the neural tangent kernel (NTK). In addition to being a theoretical tool for understanding neural networks, we demonstrate that the NTK is a simple method for achieving competitive results in a variety of machine learning applications including regression, classification, and matrix/image completion. We provide problem sets containing both theoretical and coding exercises with the aim of (1) providing newcomers a simple toolkit for building effective machine learning models in practice and (2) preparing interested students for research in the area.

Recommended Pre-requisites

Knowledge of linear algebra (level of 18.06 or 18.700), analysis (level of 18.100), and probability (level of 6.041 or 18.600) is generally assumed. Familiarity with Python (in particular, NumPy) is also assumed. Knowledge of Fourier analysis (18.103), functional analysis (18.102), random matrix theory (18.338), and complex analysis (18.112) is suggested for students who want to pursue research in this area.

Readings and Extra Materials

This course will cover material from several recent papers such as:

1. Double Descent: arXiv:1812.11118
2. Neural Tangent Kernel (NTK): arXiv:1806.07572
3. Why the NTK is constant: arXiv:2010.01092
4. Matrix Completion with NTK: arXiv:2108.00131
5. Attractors in Overparameterized Autoencoders: pnas:27162, arXiv:2006.16540

For those interested in pursuing research in this area, we recommend the following materials.

1. Functional & Complex Analysis: Princeton Lectures in Analysis ; Measure, Integration, and Analysis
2. Random Matrix Theory: Lecture Notes and Videos by Roland Speicher
3. Linear Algebra & Introductory Analysis: Linear Algebra Done Right ; Understanding Analysis

Schedule

Tuesday, Jan. 18:

- Lecture 1: Course Overview and Preliminaries
- Problem Set 1: Review of Linear Algebra, Analysis, and Probability

Wednesday, Jan. 19:

- Lecture 2: Linear Regression
- Problem Set 2: Linear Regression and Kernel Regression

Thursday, Jan. 20:

- Lecture 3: Kernel Regression

Friday, Jan. 21:

- Lecture 4: Neural Network Gaussian Processes (NNGP), Dual Activations, and Over-parameterization
- Problem Set 3: NNGP Derivations

Monday, Jan. 24:

- Lecture 5: Neural Tangent Kernel (NTK) Origin and Derivation
- Problem Set 4: NTK and the Neural Tangents Library
- Project Proposals / Paper Review

Tuesday, Jan. 25:

- Lecture 6: NTK for Deep Neural Networks and Convolutional NTK (CNTK)

Wednesday, Jan. 26:

- Lecture 7: NTK Applications - Matrix Completion and Image Inpainting

Thursday, Jan. 27:

- Lecture 8: Additional Office Hours

Friday, Jan. 28:

- Project Proposal / Paper Review Presentations

Grading, Problem Sets, and Project Proposals

This course will be graded during MIT IAP for 6 units under P/D/F grading. Receiving a passing grade requires completion of the problem sets and the project proposal/paper review.

Problem Sets

Each problem set will consist of theoretical exercises followed by coding exercises. For each problem set, students are only required to complete either the theoretical exercises or the coding part, depending on their interest. The coding exercises will be provided in the form of iPython notebooks. The goal of these problem sets is to (1) provide a foundation for those interested in performing research in the area and (2) provide a simple and effective machine learning toolkit for those interested in applications.

Project Proposal/Paper Review

Students are encouraged to take part in the project proposal assignment in which they can either (1) pose interesting theoretical questions regarding class material or (2) suggest potential applications of the methods covered in class. Students should prepare a short slide deck (5 - 10 minutes of material) that frames their question/application, reviews related work, and then presents their proposed approach. The NTK is a very recent development in machine learning, and students in this class will have the unique opportunity of being some of the first to use this material in their research or applications. Ideally, students will be able to follow up on these proposals and turn them into academic papers or new applications in industry. To get students started, we provide some broad project proposal ideas below:

1. New derivations/modifications to the NTK to match training modern neural network architectures.
2. Application of the NTK to domain specific problems, e.g. problems in biology, chemistry, physics, etc.

Instead of a project proposal, students also have the option of preparing a slide deck reviewing a recent paper in the area. As some of the papers in this area are quite technical, the goal of these paper reviews is to distill the main concepts and analysis into a form that is suitable for a broad audience. We provide some initial suggestions below but encourage students to find other papers not included here.

1. Finite Versus Infinite Neural Networks: an Empirical Study: arXiv:2007.15801.
2. Graph NTK: arXiv:1905.13192.
3. Enhanced Convolutional Neural Tangent Kernels: arXiv:1911.00809
4. Neural Kernels Without Tangents: arXiv:2003.02237
5. Dynamics of Deep Neural Networks and Neural Tangent Hierarchy: arXiv:1909.08156
6. Infinite attention: NNGP and NTK for deep attention networks: arXiv:2006.10540
7. Any paper referencing the Neural Tangents Library: https://github.com/google/neural-tangents