Online Research Seminar Syllabus

1. Overview

<table>
<tr><td>Title</td><td colspan="3">Computational Optimization for Machine Learning</td></tr>
<tr><td>Mode</td><td colspan="3">Online lectures and mentor sessions</td></tr>
<tr><td>Hours</td><td colspan="3">4*2 hours lecture +2*2 hours final project discussion session+ 1*2 hours final presentation session+ 6*1.5 hours mentor sessions (conducted by mentor)</td></tr>
<tr><td>Targeted Students</td><td colspan="3"><u>The course will be suitable for college or high-school students that have a basic knowledge in mathematics and statistics. Some Python programming will also be useful if the students want to implement the algorithms we will study.</u></td></tr>
<tr><td rowspan="2">Prerequisites</td><td rowspan="2">College Students</td><td>Required course/Knowledge</td><td>Vector calculus, linear algebra, statistics, probability theory</td></tr>
<tr><td>Recommended Materials for preparing for the course</td><td>Chapters 2,3,4,5,6 Mathematics for Machine Learning
M. Deisenroth, A Faisal, S. Ong
https://mml-book.github.io/book/mml-book.pdf</td></tr>
</table>

2. Program Introduction and Objectives

Course Description In a paragraph, please specify: What kind of program is it? What field is the program based in? What knowledge/concepts does the program include? What is the final outcome of the program (types of projects/ What did the students do to demonstrate their learning outcome, etc.)	Optimization algorithms lie at the core of many of the recent breakthroughs in machine learning and artificial intelligence. The objective of this course is to provide the students with the appropriate theoretical and practical background to understand and develop optimization algorithms for machine learning applications. The students will understand how to develop optimization algorithms that can scale to massive datasets. The students will be able to demonstrate their knowledge by creating their own algorithms and applying them to typical machine learning problems such as image classification, regression, and Bayesian optimization.
Software/Tools (if any)	**If the students want to implement any of the proposed algorithms they will need access to a python development environment (OPTIONAL)**

3. **Program Schedule**

<table>
<tr><th colspan="2">Week</th><th>Lecture</th><th>Mentor Session (lab/case study, etc.)</th><th>Assignment</th><th>Reading Materials</th></tr>
<tr><td rowspan="2">1</td><td>Topic</td><td>Introduction: Optimization and Dynamical Systems</td><td>Case Study: Supervised Learning as a Convex Optimization Problem</td><td rowspan="2">Exercises on convex analysis and dynamical systems</td><td rowspan="2">Boyd and Vandenberghe. Convex Optimization. (Ch 1,2,3,4)
Khali, Nonlinear Systems (Ch 4)
Aggarwcal, Linear Algebra and Optimization for Machine Learning (Ch 4)</td></tr>
<tr><td>Detail</td><td>In this lecture we introduce the basic definitions of optimization algorithms and discuss their connection to dynamical systems and Lyapunov Stability Theory</td><td></td></tr>
<tr><td rowspan="2">2</td><td>Topic</td><td>First Order Optimization</td><td>Case Study: Discretization of dynamical systems</td><td rowspan="2">Exercise: Analysis of preconditioned methods</td><td rowspan="2">Bubeck. Convex Optimization: Algorithms and Complexity.</td></tr>
<tr><td>Detail</td><td>We introduce and analyze the properties of gradient descent (stochastic and deterministic) and mirror descent</td><td></td></tr>
<tr><td rowspan="2">3</td><td>Topic</td><td>Advanced Optimization Algorithms</td><td>Case Study: Implementation of an ML model on the google cloud</td><td rowspan="2"></td><td rowspan="2">Aggarwcal, Linear Algebra and Optimization for Machine Learning (Ch 5)</td></tr>
<tr><td>Detail</td><td>In this lecture we discuss acceleration and other advanced algorithms such as Adagrad, RMSProp, Adam and momentum methods</td><td></td></tr>
<tr><td rowspan="2">4</td><td>Topic</td><td>Applications of Optimization to Machine Learning</td><td>Case Study: How to compute derivatives of ML models</td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>Detail</td><td>This is a practical lecture where we learn how to apply optimization to large scale ML models. We discuss algorithmic differentiation, cloud computing, finite difference methods, and Bayesian optimization.</td><td></td></tr>
<tr><td rowspan="2">5</td><td>Topic</td><td colspan="4">Final Project Phase I</td></tr>
<tr><td>Detail</td><td colspan="4">A) Final Project Milestone: Students are required to meet the following objectives before</td></tr>
</table>

<table>
<tr><td></td><td></td><td>attending the session in Week 5:
✓ Be able to model algorithms as dynamical systems
✓ Understand the role of convexity in optimization
✓ Be able to write simple ML models as optimization models

B) Things to do during the class and arrangement:
Present the result of Project Phase I in a form of mini presentation with few slides prepared.</td></tr>
<tr><td rowspan="2">6</td><td>Topic</td><td>Final Project Phase II</td></tr>
<tr><td>Detail</td><td>A) Final Project Milestone: Students are required to meet the following objectives before attending the session in Week 5:
✓ Clear idea of what the project topic will be
✓ Show mastery of main idea and theoretical foundations of the method
✓ Be able to explain the basic idea to a non-expert

B) Things to do during the class and arrangement:
Present the result of Project Phase II in a form of conversation using breakout rooms.</td></tr>
<tr><td>7</td><td colspan="2">Final Oral Presentation and Written Reporting</td></tr>
</table>

4. **Problem Sets/Written Assignments/Quizzes**

<table>
<tr><td>Total Number of Assignments</td><td colspan="2">__2_ times</td></tr>
<tr><td>Weekly Assignment Submission Deadlines</td><td colspan="2">__7__ Days after the distribution/announcement</td></tr>
<tr><td>Is Mentor needed to review and grade assignment?</td><td>Yes (X)</td><td>No ( )</td></tr>
<tr><td>Will a standard answer be provided?</td><td>Yes (X)</td><td>No ( )</td></tr>
<tr><td>Will there be Quizzes? How often/how many?</td><td colspan="2">No</td></tr>
<tr><td>Other Requirements (if any)</td><td colspan="2">N/A</td></tr>
</table>

5. Final Oral and Written Project

The students will be split into groups. The students will then perform a presentation on one of the following topics:

- Momentum Optimization Methods in Machine Learning
- Large scale distributed and federated learning
- Discretization methods for optimization algorithms
- Optimization and machine learning in financial applications

For all the topics a detailed description will be provided along with the key papers to study. The students will be able to ask questions regarding the topic and the papers before the final oral presentation.

5.1 Final Oral Presentation

- Oral Project Requirements : The students will perform a presentation and discuss their conclusions. There will be time for questions and suggestions for further work. The presentation should be 2 weeks after the end of the course.

5.2 Will you require a written final report as well?

The written report should include the following sections:

- **Background:** Introduce the topic
- **Related literature:** discuss the key papers in the field and briefly summarize their results.
- **Applications:** Discuss how the methods you described in the report could be applied to real world problems.
  Report length (approximately 10 pages)

6. Evaluation

Percentage basis of evaluation (must total 100%):

- Participation: __40______;
- Assignments/Quizzes: ___20_____;
- Final Project: ___40_____ (Oral: ___20_____; Written: ___20_____).

Joint scoring among faculty and mentors is possible

7. Suggested Future Research Fields/Direction/Topics

The students will have the opportunity to discuss future topics during their project presentation. All project topics can form the basis of a publication.

8. Instructor Introduction

8.1 Instructor Title Professor

8.2 Instructor Bio: Panos Parpas is a Professor of Computational Mathematics at the Department of Computing, Imperial College London. Before joining Imperial College, he was a postdoctoral fellow at MIT (2009-2011). Before that, he was a quantitative associate at Credit-Suisse (2007-2009). He is interested in the development and analysis of algorithms for large scale optimisation problems. He is also interested in exploiting the structure of large scale models arising in applications. His research has been published in leading journals such as SIAM Journal of Optimization, SIAM Journal of Scientific computing, SIAM Journal on Mathematical Finance among others. He has presented his research in several meetings, conferences and seminars and is frequently involved in the organisation of specialist workshops and meetings. He is an associate editor for two journals and was awarded a JP Morgan Faculty Award in 2019 and 2021. He currently supervises 3 PhD students and one postdoc

8.3 Instructor Profile Photo