Math 565: Optimization for Machine Learning (Spring 2026)

Credits	3
Time	Tue+Thu, 1:30–2:45 PM
Location	Sloan 7 (P), VECS 120 (V), Zoom
Instructor	Bala Krishnamoorthy
Check-in Hours	Wed 10:30–11:30 AM, Thu 3-4 PM, in person and Zoom
Email	kbala@wsu.edu
Web page	https://bala-krishnamoorthy.github.io/Math565.html
Text	Charu C. Aggarwal: Linear Algebra and Optimization for Machine Learning Springer, ISBN: 978-3-030-40344-7; Full text available via WSU Libraries

Description of the Course

This course will offer a systematic treatment of optimization methods with a focus on applications in machine learning (ML). There are key differences between optimization as used in ML and traditional optimization—for instance, the performance of ML models are assessed based on how well they generalize to *test data* (rather than on the whole dataset). For instance, the widely used stochastic gradient-descent methods could have lower accuracy than gradient-descent on training data but often perform better on test data!

We will cover a selection of relevant topics from the book titled “Linear Algebra and Optimization for Machine Learning” by Charu Aggarwal, including gradient and stochastic gradient descent, Newton method in ML problems such as regression and support vector machines (SVM), Lagrangian relaxation and duality for SVM, penalty-based methods, optimization in computational graphs including neural networks (NNs), and backpropagation in NNs (see a Tentative List of Topics to be covered in the next Page).

Homework assignments will include proof-type problems as well as ones needing use of software. Students will also work on a computational project. No exams will be given. This course will (ideally) be a follow-up of Math 564: Nonlinear Optimization, which will be offered in Fall 2025. Students taking that course will be best prepared to do well in Math 565. Independent of Math564, prerequisites for Math565 are familiarity with analysis and linear algebra with proof-based work at the undergraduate (400-) level, **or obtain the permission of the instructor**. Familiarity with computer programming languages or packages such as Matlab or Python will also be expected.

Organization and Grading

The course will have around six homework assignments. These assignments will include theoretical problems as well as ones that involve the use of software packages. Apart from the homework assignments, there will a course project that will involve the use of software packages such as Matlab/Python (or another programming language). The total score for the course will calculated using the weights: homework: 70% and project: 30%. There will be no exams. This total score for the course will be curved to assign grades.

Software

As part of homework assignments, students will be asked to implement certain algorithms presented in class in Matlab or Python or another programming language. The project will involve the implementation of a selected set of optimization methods and compare their performance with that of standard tools on large datasets.

Tentative List of Topics Covered

1. Basic Problems in ML and Optimization in ML: 2 lectures
2. Optimization Basics with an ML View: 7 lectures
   - (uni/multi)variate optimization
   - gradient descent
   - stochastic gradient descent
   - optimization models for binary targets (classification)
     - support vector machine (SVM)
     - generalization to multiclass setting
3. Advanced Optimization Methods in ML: 7 lectures
   - Adjusting first-order derivatives for descent
     - AdaGrad
     - RMSProp
     - Adam
   - Newton Methods in ML
     - Newton method for SVM and regression
     - saddle-point problem
     - conjugate gradient method
     - quasi-Newton and BFGS methods
4. Constrained Optimization and Duality in ML: 7 lectures
   -
     - primal gradient and coordinate descent
     - Newton method with equality constraints
   - Lagrangian relaxation and duality
     - KKT conditions
     - SVM dual: formulation and algorithms
     - penalty-based and primal-dual methods
5. Optimization in Computational Graphs: 7 lectures
   - neural networks (NNs) as directed computational graphs
   - optimization in directed acyclic graphs
   - backpropagation in NNs

Academic Integrity

I encourage discussion of homework problems with others. But each student should submit their own (hand or type) written solutions and/or computer programs. You might search the internet for finding materials to enhance your understanding. If you use such material to assist in your homework submission, you **should** cite the relevant sources. Plagiarism or cheating will **not** be tolerated. In particular, do not copy blindly from internet sources! Such behavior is easy to detect, and will result in a zero grade for the item in question and possibly a failing grade for the entire course.

AI Use Permitted with Acknowledgment: You are welcome to try AI tools based on generative models such as CoPilot, Chat-GPT, Gemini, and others. Any such use in your homework assignments or project(s) should be properly acknowledged in your submission. One way to list such an acknowledgment is as follows:
"This response is based on the solution provided by CoPilot *when queried with* the exact text of the problem statement *on* Feb 12, 2026.*"* (change the entries in red appropriately).

University Syllabus

For information on the following WSU policies, please see the University Syllabus. Students are responsible for reading and understanding all university-wide policies and resources pertaining to all courses provided on this web page.

- Reasonable Accommodations
- Arrangements for Religious Reasons
- Emergencies on Campus (including active shooter and severe weather)
- Student Support Resources (including Student Care Network and Campus Resources and Support)