MATH 495: Special Topics — Mathematics of AI

Course Description

Artificial Intelligence (AI) relies on a rich set of mathematical ideas. This course introduces the linear-algebraic and probabilistic foundations most frequently used in modern machine-learning and AI systems. The material accessible to students from engineering, computer science, statistics, and the natural and social sciences.

Prerequisites

- A course in Linear Algebra.
- Prior experience with programming not required. All coding tutorials will use Python/NumPy.

Required Textbook

Deisenroth, Faisal & Ong. *Mathematics for Machine Learning.* Cambridge Univ. Press, 2020.

Supplementary References (Optional)

- Boyd & Vandenberghe, *Introduction to Applied Linear Algebra* (2018).
- Aggarwal, *Neural Networks and Deep Learning* (2018).
- Goodfellow, Bengio & Courville, *Deep Learning* (2016) — Chapters 2–3 for background reading.

Learning Outcomes

By the end of the term, students will be able to:

1. Explain and apply least-squares and ridge regression from a linear-algebra viewpoint.
2. Interpret probability distributions and expectation as tools for modeling data uncertainty.
3. Compute gradients of scalar- and vector-valued functions using matrix calculus.
4. Implement gradient-descent variants to train simple neural networks.
5. Analyze singular value decomposition (SVD) and principal component analysis (PCA) for dimensionality reduction.
6. Evaluate model complexity, overfitting, and regularization trade-offs mathematically.

Tentative Weekly Schedule

Week	Topics	Key Ideas / Skills
1	Course overview; linear algebra refresher	Vector spaces, inner products, norms.
2	Least–squares estimation	Normal equations, pseudoinverse.
3	Probability review	Random variables, expectation, covariance.
4	Multivariate Gaussians	Covariance matrices, Mahalanobis distance.
5	Optimization basics	Gradients, critical points, convexity.
6	Gradient descent methods	Batch vs. stochastic descent, convergence intuition.
7	Matrix calculus toolkit	Jacobians, chain rule in matrix form.
8	Singular Value Decomposition	Low-rank approximation, PCA motivation.
9	Principal Component Analysis	Variance maximization, data whitening.
10	Linear classifiers	Perceptron, softmax, logistic regression.
11	Feed-forward neural networks	Universal approximation, back-prop through layers.
12	Regularization & generalization	L2/L1 penalties, bias-variance trade-off.
13	Kernel tricks (intro)	Feature maps, RBF kernels, support-vector intuition.
14	Fairness & ethical AI math	Metrics, constraints, regularization for fairness.
15	Student project presentations	—

Assessment

- **Homework (40%)** – Bi-weekly problem sets combining theory & coding.
- **Midterm Exam (20%)** – Concepts from Weeks 1–8.
- **Final Project (30%)** – Team mini-research or application project, written report and presentation.
- **Participation (10%)** – In-class polls and peer reviews.

Software & Computational Resources

We will use Python 3, NumPy, and JupyterNotebooks. Setup guides and conda environment files will be posted on the course page.