Optimization for AI

1. **Course number and name:** 020OAIES5 Optimization for AI

2. **Credits and contact hours:** 4 ECTS credits, 2x1:15 contact hours

3. **Name of course coordinator:** Khalil Hariss

4. **Instructional materials:** PowerPoint slides

   References:
   - E. Alpaydın, Introduction to Machine Learning, 4th ed. Cambridge, MA: MIT Press, 2020.
   - Mathematics of Neural Networks book, Bart M.N. Smets, November 12, 2022
   - Deep Learning book by Ian Goodfellow, Yoshua Bengio, Aaron Courville. www.deeplearningbook.org
   - Mathematics for Machine Learning book, Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong.

5. **Specific course information**
   a. **Catalog description:**
   This course aims to provide students with a solid theoretical and practical foundation in mathematical optimization techniques essential to the development and refinement of machine learning algorithms and artificial intelligence applications. Students will learn to analyze and implement optimization methods, including gradient-based algorithms, adaptive learning rate techniques (e.g., Adam, RMSProp), automatic differentiation, and backpropagation, while addressing critical training challenges such as vanishing and exploding gradients. The course also covers neural network initialization strategies, dimensionality reduction (PCA), density estimation, and support vector machines (SVM), along with both unconstrained and constrained optimization problems. By the end of the course, students will be equipped to apply these techniques to improve model performance and solve complex problems across various AI domains.

   b. **Prerequisites:** 020STAES1/020STTES1 Statistics

   a. **Required** for CCE Artificial Intelligence Option students; **Selected Elective** for students in the CCE Software Engineering and Telecommunication Networks Options.

6. **Educational objectives for the course**
   a. **Specific Course Learning Outcomes:**
   - Explain the role and importance of optimization in training machine learning and AI models.
   - Apply gradient-based and adaptive optimization algorithms (e.g., SGD, Adam, RMSProp) to train neural networks.

- Implement automatic differentiation and backpropagation in deep learning architectures.
- Identify and address optimization challenges such as vanishing and exploding gradients using appropriate strategies (e.g., initialization techniques, activation functions, gradient clipping).
- Solve both unconstrained and constrained optimization problems using mathematical techniques such as gradient descent and Lagrange multipliers.
- Analyze the impact of weight initialization and learning rate scheduling on training stability and model convergence.
- Formulate and solve optimization-based approaches to PCA, density estimation (e.g., EM algorithm), and support vector machines (SVM).
- Evaluate the effectiveness of optimization techniques in improving model accuracy, generalization, and training efficiency across different AI applications.

b. PI addressed by the course:

PI	1.1	1.2	1.3	2.3	2.4	2.5	4.2	7.1	7.2
Covered	x	x	x	x	x	x	x	x	x
Assessed	x	x	x		x	x			

7. Brief list of topics to be covered

- Introduction to Optimization in Machine Learning and AI (1 lecture
- Convex vs. Non-Convex, Constrained vs. Unconstrained Optimization Problems (2 lectures)
- Gradient Descent and Its Variants (SGD, Mini-Batch, Momentum) (3 lectures)
- Adaptive Optimization Algorithms: AdaGrad, RMSProp, and Adam (2 lectures)
- Automatic Differentiation and Backpropagation (2 lectures)
- Activation Functions and Their Role in Optimization (2 lectures)
- Neural Network Initialization and Training Challenges (Vanishing/Exploding Gradients) (2 lectures)
- Learning Rate Strategies and Gradient Clipping Techniques (2 lectures)
- Solving Constrained Optimization Problems: Lagrange Multipliers and KKT Conditions (2 lectures)
- Principal Component Analysis (PCA) and Dimensionality Reduction via Optimization (2 lectures)
- Density Estimation and the EM Algorithm (2 lectures)
- Support Vector Machines (SVM), Quadratic Programming, and Regularization (2 lectures)