1. **Course Number and Course Title:**
   ML 506 – Hardware Architectures for Machine Learning

2. **Credits Hours:** 3–0–3

3. **Prerequisites and/or Co-Requisites:**
   Prerequisite: Approval of the CSE Head of Department Competencies: Undergraduate course in Computer Architecture

4. **Name and Contact Information of Instructor:**
   Dr. Assim Sagahyroon
   Office: ESB-2001
   Email: asagahyroon@aus.edu
   Phone: (06) 515-2952
   Office Hours: as posted on *iLearn* or by appointment

5. **Course Description (Catalog Description):**
   Covers the basics of machine learning (ML) algorithms. Introduces the design and implementation of hardware architectures for efficient processing of ML algorithms. Covers the topics of ML on programmable platforms, deep learning accelerators, design alternatives and optimization for ML algorithms, training for deep learning. Discusses the characteristics, advantages and disadvantages of popular hardware platforms including CPUs, GPUs, and FPGAs in designing efficient learning networks.

6. **Textbook and other Supplemental Material:**
   Textbook:
   - L. Chen, D. Penney and D. Jiminez, AI for Computer Architecture: Principles, Practice and Prospects, Morgan and Claypool, 2021.

   References:
   - P. Whatmough, G.Y Wei, D. Brooks, *Deep Learning for Computer Architects*, Morgan and Claypool, 2017.
   - M. Vladutiu, *Computer Arithmetic: Algorithms and Hardware Implementations,* Springer, 2012.
   - N. Hemsoth and T. Prickett Morgan, *FPGA Frontiers: New Applications in Reconfigurable Computing,* NextPlatform Press, 2017.
   - C. Kachris, B. Falsafi, D. Soudris, *Hardware Accelerators in Datacenters, Springer*, 2019.

7. **Learning Outcomes:**
   Upon completion of the course, students will be able to:
   1. Discuss the computing characteristics of ML algorithms
   2. Analyze tradeoffs between various hardware architectures and platforms including CPUs, GPUs and FPGAs in running ML applications
   3. Contrast between the key design considerations for efficient ML processing
   4. Assess the utility of various design techniques for efficient processing of machine learning algorithms.

5. Evaluate future implementation trends and opportunities in deep learning systems.

8. Teaching and Learning Methodologies:

The course will involve a mix of lectures interspersed with paper reading and discussions. A term paper and a semester long project (broken into mini projects) will focus on developing a hardware accelerator for ML, and prototyping it on a FPGA.

9. Course Topics and Schedule:

Topic	Weeks
Introduction to machine learning –Part I	Week 1
Introduction to machine learning –Part II	Week 2
Application examples of machine learning	Week 3
Arithmetic circuits for machine learning processing	Week 4
Available hardware platforms and their characteristics: GPUs	Week 5
Available hardware platforms and their characteristics: FPGAs and ASICs	Week 6
Training versus inference : differences	Week 7
Training versus inference : computational needs	Week 8
Hardware accelerators and accelerated systems architecture	Week 9
Energy efficient acceleration	Week 10
Machine learning on Specialized Hardware	Week 11
Metrics for evaluating machine Learning Systems	Week 12
ML hardware usage in datacenters	Week 13
Discussion on emerging trends in Machine Learning hardware architectures	Week 14
Presentations	Week 15
Final Exam	Week 16

10. Schedule of Laboratory and other Non-Lecture Sessions:

Term Paper

the term paper will require extensive literature review and analysis of existing literature to propose an improvement on an existing architectural approach to accelerate the execution of machine learning algorithms on hardware platforms. It is meant to be a research-based paper that requires students to present their findings and analysis in a structured format.

Papers Critique

At the offering of the course, selected readings from journals and conferences on the subject will be assigned. Students are required to complete an analysis and evaluation of these scholarly papers or research articles. The task involves critically examining the content, structure, methodology, and overall quality of the paper. Student are to submit written critique that provides feedback and identifies strengths and weaknesses of the paper.

11. Out-of-Class Assignments with Due Dates:

Assignment	Due Date (tentative)
Homework 1	Week #3
Homework 2	Week #5
Homework 3	Week #7
Mini-Project #1	Week #9
Homework 4	Week #10
Homework 5	Week #11
Mini-Project #2	Week #12

The Project

The course project involves building a hardware accelerator. This task is broken into two mini-projects as follows:

Mini-Project # 1

The first mini-project will require the student to complete the following tasks:
--Define the requirements
--Design the architecture: Based on the requirements, design the high-level architecture of the hardware accelerator
--Implement the design: Create a detailed design specification based on the architecture. This step involves selecting and designing individual components such as processing units, memory elements, interfaces, and control logic. Utilize hardware description languages VHDL or Verilog to describe the design.
--Verification and simulation: Verify the correctness and functionality of the design through simulation using tools like ModelSim or VCS

Mini-Project # 2

The second mini-project will require the student to complete the following tasks:
--Prototype and debugging: Build a physical prototype of the hardware accelerator using an FPGA
--Integrate the hardware accelerator into the target system or platform. This may involve connecting it to a host processor or other components, ensuring proper communication and data transfer between the accelerator and the rest of the system
--Evaluate the performance of the hardware accelerator by running benchmark applications or test cases. Measure factors such as speedup, throughput, and power efficiency to assess the effectiveness of the accelerator.
--Based on the evaluation results, analyze any performance and identify areas for improvement.

12. Student Evaluation:

Assessment	Weight	Due Date (tentative)
Homework	10 %	biweekly
Mini-projects	10 %	week# 9 & week #12
Midterm Exam	30 %	week #7
Papers Critique	10 %	week #13
Term Paper	10 %	week #14

Final Exam	30 %	week #16

13. Assessment Instruments

Assessment	Course Learning Outcomes
Homework	O1 – O4
Mini-projects	O1 – O4
Midterm Exam	O1 – O3
Papers critique	O2 – O4
Term Paper	O4 - O5
Final Exam	O1 – O4

14. Contribution of Course to Program Outcomes

This course contributes to the accomplishment of the following program outcomes:

MSML Program Outcomes	Emphasis in this course	Course Learning Outcomes
1. Perform research emphasizing creativity, independent learning, and scientific methods in the field of Machine Learning.	●	O1-O5
2. Apply advanced mathematics, computer science knowledge, and software tools in identifying, formulating, and solving real world problems.	○	O2-O4
3. Demonstrate an in-depth understanding of modern Machine Learning approaches, algorithms, and tools.	○	O1
4. Select and use techniques, skills, and modern tools necessary for research or professional practice.	◑	O1-O5
5. Communicate effectively through technical presentations and reports.	○	O5
6. Recognize the need for, and engage in, lifelong learning in professional areas.	○	O5
7. Attend to professional and ethical responsibilities within global and societal contexts.		

Emphasis: ● High; ◑ Medium; ○ Low; Blank – Nothing Specific Expected

15. Letter Grade Policy

Total (T)	Letter Grade
$90 \leq T$	A
$85 \leq T < 90$	A-
$80 \leq T < 85$	B+
$75 \leq T < 80$	B
$70 \leq T < 75$	B-
$65 \leq T < 70$	C+
$60 \leq T < 65$	C
$T < 40$	F