Elective 3-NEXT GENERATION COMPUTING

Subject Code: CS35321CS	Subject Type: DE (Elective-3)	Evaluation Scheme: MSE (30), TA (20), ESE (50)	L-T-P: 3-0-0 (3)

Pre-Requisites: Computer Organization & Architecture, Linear Algebra, Complex Analysis, Probability Theory, Foundations of Machine Learning or Pattern Recognition, Proficiency in a High-Level Programming Language (Python, C/C++)

Course Objective: To analyze GPU architectures and design scalable parallel computing strategies for massive AI workloads. To establish the mathematical foundations of quantum mechanics, quantum algorithms, and Quantum Machine Learning (QML). To investigate neuromorphic computing paradigms, spiking neural networks, and biologically inspired learning.

Course Outcomes: At the end of the course, the student will be able to

CO	Course Outcomes
CO-1	Evaluate GPU architectures, optimize kernel-level execution, and design scalable distributed training pipelines for deep learning models.
CO-2	Analyze quantum states and circuits to implement foundational quantum algorithms and Parameterized Quantum Circuits (PQCs) for ML tasks.
CO-3	Design and simulate Spiking Neural Networks (SNNs) and critically evaluate neuromorphic hardware for energy-efficient cognitive computing.
CO-4	Apply advanced model compression techniques for Edge AI deployment and critically evaluate emerging optical, analog, and in-memory computing paradigms.

Syllabus:

Unit-1: High-Performance Computing and Distributed AI Architectures

Memory hierarchies, computational bottlenecks, and the Roofline performance model. GPU Architecture: Streaming multiprocessors, warp execution, and the SIMT paradigm. CUDA programming fundamentals: Thread hierarchy, memory coalescing, and synchronization mechanisms. Advanced tensor operations and mixed-precision optimization. Distributed Deep Learning: Strategies for data, tensor, and pipeline parallelism; gradient synchronization; and memory optimization frameworks (ZeRO, FSDP) for massive language models.

Unit-2: Quantum Computing Foundations and Machine Learning

Mathematical foundations: Qubits, superposition, entanglement, and Dirac notation. Universal quantum gate sets and circuit design. Foundational algorithms: Quantum Fourier Transform (QFT), Quantum Phase Estimation (QPE), and Grover's search algorithm. Quantum Machine Learning (QML): Mapping classical data to quantum states (Amplitude, basis encoding). Parameterized Quantum Circuits (PQCs), expressibility, and gradient computation (parameter shift rule). Optimization challenges in the NISQ era: Barren plateaus and error mitigation.

Unit-3: Neuromorphic Computing and Bio-Inspired Systems

Biological plausibility in computation: Membrane potential dynamics and Leaky Integrate-and Fire (LIF) neuron models. Neural coding schemes: Rate, temporal, and population coding. Synaptic plasticity and bio-inspired learning: Hebbian learning and Spike-Timing Dependent Plasticity (STDP). Training SNNs: Surrogate gradient methods and backpropagation through time (BPTT). Architectural overview and energy efficiency analysis of contemporary neuromorphic hardware accelerators.

Unit-4: Edge AI, Model Compression and Post-CMOS Paradigms

The Edge-Fog-Cloud continuum and latency-constrained AI deployment. Model compression techniques: Post-training quantization, quantization-aware training, unstructured/structured pruning, and knowledge distillation. Federated learning paradigms and communication efficiency. Survey of emerging post-CMOS architectures: Analog in-memory computing (crossbar arrays), photonic/optical neural networks, DNA computing concepts, and the theoretical thermodynamic limits of computational energy efficiency.

Learning Resources:

Text Books:

1. Programming Massively Parallel Processors, Kirk D.B., Hwu W.W. Morgan Kaufmann 2022 (4th Ed.).
2. TinyML: Machine Learning with TensorFlow Lite, Warden P., Situnayake D. O'Reilly Media 2020.
3. Deep Learning, Goodfellow I., Bengio Y., Courville A. MIT Press 2016.
4. Quantum Computation and Quantum Information, Nielsen M.A., Chuang I.L. Cambridge Univ. Press 2010 (10th Ed.)

Reference Books:

1. Computer Architecture: A Quantitative Approach, Hennessy J., Patterson D.
2. Introduction to Quantum Computing, Rieffel E., Polak W. MIT Press.
3. Spiking Neuron Models, Gerstner W., Kistler W.M. Cambridge Univ. Press.