RADemics

Design and Implementation of Low Power ARM Cortex Microcontrollers for Edge AI Computing

Shaik Maheboob, Shaik Lal John Basha
MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT, VIT-AP UNIVERSITY

Design and Implementation of Low Power ARM Cortex Microcontrollers for Edge AI Computing

[1]Shaik Maheboob, Associate Professor, CSE Department, Marri Laxman Reddy Institute of Technology and Management, Hyderabad, Telangana, India. shaikmaheboob@mlritm.ac.in

[2]Shaik Lal John Basha, Research Scholar and Assistant professor(C), School of Electronics Engineering, VIT-AP University, Amaravati, Guntur, Andhra Pradesh, India. sk.laljohn@gmail.com

Abstract

The integration of Artificial Intelligence (AI) into embedded systems has accelerated the evolution of edge computing, particularly within power-constrained environments such as educational institutions and intelligent learning infrastructures. This book chapter presents a comprehensive examination of the design and implementation strategies for low-power ARM Cortex microcontrollers tailored for AI-driven edge applications. Emphasis is placed on architectural innovations, execution efficiency, real-time responsiveness, and memory management techniques essential for deploying lightweight machine learning models on resource-constrained devices. The chapter explores the capabilities of various Cortex-M cores, delving into their instruction set architecture, interrupt handling mechanisms, and power management features, while highlighting optimization practices that balance energy efficiency with computational performance. Use cases in AI-based learning analytics, student engagement systems, and adaptive educational technologies illustrate the practical deployment of such architectures. Through theoretical analysis and application-oriented perspectives, the chapter outlines how Cortex-M microcontrollers are enabling scalable, intelligent, and secure AI systems at the edge of educational ecosystems.

Keywords: Edge AI, ARM Cortex-M, Low Power Computing, Embedded Systems, Learning Analytics, Real-Time Processing

Introduction

The rapid advancement of Artificial Intelligence (AI) has catalyzed a significant transformation in embedded system design, particularly within edge computing environments where localized data processing and real-time responsiveness are crucial [1]. In educational technology, the shift toward intelligent, autonomous systems has placed a growing emphasis on developing solutions that can function with minimal energy consumption and computational overhead [2]. This demand has propelled the integration of AI into resource-constrained platforms, with ARM Cortex microcontrollers emerging as ideal candidates for implementing edge AI systems [3]. Their efficient instruction set architecture, low power footprint, and hardware support for deterministic execution make them suitable for executing lightweight AI models, enabling real-time analytics and decision-making within distributed learning ecosystems [4]. The relevance of these systems is especially evident in applications such as adaptive learning platforms, student monitoring devices, and classroom environmental sensing, where continuous operation and real-time feedback are essential [5].

ARM Cortex-M microcontrollers offer a scalable architecture that meets the performance and energy requirements of edge AI deployment [6]. The Cortex-M series, ranging from M0 to M85, supports a variety of applications by providing different levels of computational capabilities, memory access mechanisms, and integrated peripherals [7]. These cores are designed with a Harvard architecture, Thumb-2 instruction set, and advanced interrupt handling features, which collectively contribute to efficient execution and fast response times [8]. For AI workloads, these characteristics are vital in maintaining low-latency inference and conserving energy during idle states [9]. In the context of higher education, where edge devices are increasingly used to enhance personalized learning and institutional analytics, the ability to manage real-time data locally without relying on cloud connectivity enhances both reliability and privacy [10].