RADemics

Performance Metrics and Evaluation Techniques for IoT Machine Learning Models

P. R. SUKANYA SRIDEVI, R. ARIVUKKODI
MEENAKSHI ACADEMY OF HIGHER EDUCATION AND RESEARCH,
MEENAKSHI ACADEMY OF HIGHER EDUCATION AND RESEARCH

16. Performance Metrics and Evaluation Techniques for IoT Machine Learning Models

P. R. SUKANYA SRIDEVI, Assistant Professor, Computer Science and Applications, Meenakshi Academy Of Higher Education And Research, sukan.lec@gmail.com

R. ARIVUKKODI, Assistant Professor, Computer Science and Applications, Meenakshi Academy of Higher Education and Research, nkulali@gmail.com

Abstract

The integration of ML with the IOT has transformed data analysis and decision-making across a myriad of applications. As IoT systems continue to expand and evolve, ensuring the effective performance of ML models becomes increasingly critical. This chapter provides a comprehensive examination of performance metrics and evaluation techniques specifically tailored for ML models within IoT environments. It delves into context-specific metrics, including latency and throughput, to address the unique challenges posed by high-dimensional and heterogeneous data sources. Special emphasis was placed on evaluating error recovery mechanisms to maintain Quality of Service (QoS) amidst errors and disruptions. Key insights into model complexity, real-time performance constraints, and resource utilization are discussed to provide a holistic approach to performance evaluation. By exploring these aspects, this chapter aims to bridge the existing gaps in assessing and optimizing ML models in dynamic and resource-constrained IoT settings. The findings offer valuable guidance for researchers and practitioners striving to enhance the reliability and efficiency of IoT-driven ML applications.

Keywords: Machine Learning, Internet of Things, Performance Metrics, Latency, Throughput, Error Recovery.

Introduction

The intersection of ML and the IOT represents a transformative advancement in the field of data analytics and intelligent systems [1]. As IoT devices proliferate, generating enormous volumes of diverse data, the integration of ML algorithms offers unprecedented opportunities for real-time insights and automated decision-making [2]. This convergence has significant implications across various domains, including smart cities, healthcare, industrial automation, and environmental monitoring [3]. Despite these advancements, the performance and reliability of ML models in such contexts are subject to rigorous demands and constraints that require specialized evaluation techniques [4].

One of the primary challenges in evaluating ML models for IoT applications was managing high-dimensional data [5]. IoT systems often generate data with numerous features, making it essential to develop metrics that can effectively assess model performance in these complex environments [6]. High-dimensional data poses difficulties such as increased computational demands and potential overfitting, which necessitate the use of specific performance metrics to ensure accurate and generalizable models [7]. Metrics addressing feature relevance, dimensionality reduction effectiveness, and model complexity are critical for navigating these challenges and optimizing model performance [8].

Latency and throughput are crucial performance metrics in resource-constrained IoT environments [9]. Latency measures the delay between data input and model response, while throughput assesses the volume of data processed per unit of time [10]. In scenarios where computational resources and network bandwidth are limited, minimizing latency and maximizing throughput become essential for maintaining real-time system functionality and user satisfaction [11]. Effective evaluation of these metrics helps in balancing the trade-offs between model performance and resource usage, ensuring that the system can handle high data loads efficiently without compromising on responsiveness [12-16].

The evaluation of error recovery mechanisms was another vital aspect of maintaining Quality of Service (QoS) in IoT systems [17]. Errors and disruptions are inevitable in complex and distributed environments, making it crucial to have robust recovery strategies in place [18]. Error recovery metrics, including detection accuracy, recovery time, and impact assessment, provide insights into how well a system can handle and recover from failures [19]. These metrics are essential for ensuring that the system can continue to deliver reliable service and meet performance expectations despite encountering operational challenges [20-25].