RADemics

Hybrid Deep Learning Models for Multi Source Data Integration Tasks

Dr. Vijayalakshmi. S, Dr.Anisha.K, Mrs. R. Mahalakshmi
ASAN MEMORIAL COLLEGE OF ENGINEERING AND TECHNOLOGY, SRI CHANDRASEKHARENDRA SARASWATHI VISWA MAHAVIDYALAYA

Hybrid Deep Learning Models for Multi Source Data Integration Tasks

[1]Vijayalakshmi.S, Associate Professor, Department of Electrical and Electronics Engineering, Asan Memorial College of Engineering and Technology, Chengalpattu-603105. vijirech@gmail.com

[2]Anisha.K, Assistant Professor, Department of Electrical and Electronics Engineering, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Enathur, Kanchipuram – 631561. anishakscsvmv@gmail.com

[3]R. Mahalakshmi, Research Scholar, Department of Electrical and Electronics Engineering, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Enathur, Kanchipuram – 631561. balavishua12@gmail.com

Abstract

The growing diversity and complexity of contemporary data environments have intensified the need for intelligent computational frameworks capable of integrating heterogeneous information and uncovering meaningful relationships across multiple sources. This book chapter, Hybrid Deep Learning Models for Multi-Source Data Integration Tasks, explores advanced architectures and learning strategies for combining structured, unstructured, and multimodal data into unified representations for reliable prediction and knowledge discovery. The discussion examines hybrid deep learning architectures, ensemble prediction, multimodal representation learning, cross-modal alignment, attention mechanisms, transformer-based fusion, and graph neural networks. Particular emphasis rests on cross-source correlation, semantic alignment, dependency modeling, and knowledge graph integration for identifying hidden relationships within interconnected datasets. The chapter also addresses data preprocessing, feature compatibility, missing modalities, conflicting information, computational scalability, and model evaluation. Applications across healthcare analytics, industrial monitoring, financial risk assessment, environmental observation, and intelligent transportation demonstrate the practical relevance of integrated learning approaches. Robustness, interpretability, privacy protection, and reliable validation receive particular consideration for responsible implementation. By connecting architectural principles with application-oriented challenges, this chapter establishes a comprehensive foundation for developing scalable and context-aware deep learning systems that transform fragmented data into coherent representations and actionable analytical insights.

Keywords: Hybrid Deep Learning, Multi-Source Data Integration, Multimodal Representation Learning, Feature Fusion, Cross-Modal Alignment, Knowledge Discovery

Introduction

The expansion of digital technologies, intelligent sensing systems, and data-driven applications has generated enormous volumes of information across diverse domains [1]. Modern analytical environments frequently incorporate structured databases, textual documents,

medical images, sensor measurements, and time-series observations to support complex prediction and decision-making tasks [2]. Processing these sources independently can limit the identification of meaningful relationships and reduce the effectiveness of comprehensive data analysis [3]. Hybrid deep learning models offer a promising approach by combining complementary neural network architectures to extract, align, and integrate heterogeneous information [4]. This capability supports the development of unified representations, enabling intelligent systems to discover hidden patterns and generate more contextually informed analytical outcomes [5].

Multi-source data integration presents several technical challenges arising from differences in data formats, feature distributions, sampling frequencies, semantic structures, and information quality [6]. Structured records typically contain well-defined attributes, whereas images, audio signals, and textual documents require specialized representation techniques [7]. Missing observations, inconsistent terminology, duplicated records, and conflicting information further complicate integration [8]. Effective preprocessing, feature extraction, normalization, and semantic alignment establish the foundation for reliable learning across heterogeneous sources [9]. Deep neural networks can transform diverse inputs into informative representations, while feature alignment mechanisms facilitate meaningful comparisons between corresponding observations [10]. These procedures improve data compatibility and support the extraction of complementary information from interconnected datasets.

Hybrid deep learning architectures strengthen multi-source integration by combining the distinctive capabilities of different neural network models [11]. Convolutional neural networks extract spatial features from visual data, recurrent architectures capture temporal dependencies, and transformer models identify contextual relationships across complex sequences [12]. Autoencoders facilitate compact latent representation learning, while graph neural networks model relationships among interconnected entities [13]. Ensemble architectures integrate predictions from multiple models, potentially improving predictive stability when component models capture complementary information [14]. Attention mechanisms further support selective feature weighting by emphasizing information relevant to specific analytical objectives [15]. Appropriate architecture selection depends on source characteristics, task requirements, computational resources, and the expected balance between predictive performance and interpretability.