RADemics

Handling Uncertainty and Noise in Deep Learning Predictive Models

C P. Karthikeyan, Dr. S. Ranganathan, Arivu Pandeeswaran V

DHANALAKSHMI SRINIVASAN COLLEGE OF ENGINEERING AND TECHNOLOGY, ACADEMY OF MARITIME EDUCATION AND TRAINING - DEEMED TO BE UNIVERSITY, SETHU INSTITUTE OF TECHNOLOGY

Handling Uncertainty and Noise in Deep Learning Predictive Models

[1]C P. Karthikeyan, Professor & Head, Department of Robotics and Artificial Intelligence, Dhanalakshmi Srinivasan College of Engineering and Technology, Mamallapuram, Chennai – 603104. cpkarthik222.palani@gmail.com

[2]S. Ranganathan, Professor, Department of Marine Engineering, Academy of Maritime Education and Training - Deemed to be University, Kanathur, Chennai -603112. ranganathan.s@ametuniv.ac.in

[3]Arivu Pandeeswaran V, Assistant Professor, Department of Artificial Intelligence and Data Science, Sethu Institute of Technology, Pulloor, Kariapatti, Virudhunagar - 626 115. pandeeswaran2607@gmail.com

Abstract

Deep learning predictive models increasingly operate in complex environments characterized by noisy observations, incomplete information, uncertain labels, and changing data distributions. Effective management of these challenges remains essential for maintaining predictive accuracy, robustness, and decision reliability. This book chapter, Handling Uncertainty and Noise in Deep Learning Predictive Models, presents a comprehensive examination of methodological approaches for developing dependable prediction systems under imperfect data conditions. The discussion covers noise detection, preprocessing, hybrid neural network architectures, probabilistic distribution modeling, Bayesian deep learning, Monte Carlo dropout, and ensemble-based uncertainty estimation. Particular attention focuses on noise-robust training strategies, including curriculum learning, sample reweighting, regularization, and robust loss functions. Probabilistic decision thresholds and confidence estimation further demonstrate how uncertainty information can support risk-sensitive decisions across critical applications. Robustness assessment through benchmarking under noise-free and noisy data conditions provides a foundation for evaluating generalization, calibration, and performance stability. The chapter also considers distribution shifts, adversarial perturbations, and computational constraints affecting practical implementation. By integrating uncertainty quantification with systematic robustness evaluation, this work establishes a structured foundation for developing transparent, reliable, and context-aware deep learning models across healthcare, manufacturing, finance, and intelligent decision support environments.

Keywords: Deep Learning, Uncertainty Quantification, Noise-Robust Training, Predictive Reliability, Probabilistic Modeling, Robustness Evaluation

Introduction

The expansion of artificial intelligence and deep learning has significantly transformed predictive modeling across healthcare, manufacturing, finance, cybersecurity, and intelligent decision support systems [1]. Modern predictive models process large volumes of structured and

unstructured data to identify complex relationships, estimate future outcomes, and support evidence-based decisions [2]. In practical environments, data frequently contain measurement errors, incomplete observations, incorrect labels, and unpredictable variations that can affect model performance [3]. Such imperfections create challenges for reliable learning and accurate prediction [4]. Developing deep learning models capable of handling uncertainty and noise therefore represents an important research direction for improving predictive reliability and strengthening the practical value of artificial intelligence applications [5].

Uncertainty and noise influence predictive models through different mechanisms and require appropriate handling strategies [6]. Aleatoric uncertainty arises from inherent variability within observations, whereas epistemic uncertainty reflects limitations in the model's knowledge and training data [7]. Noise can originate from faulty sensors, inconsistent annotations, incomplete records, environmental disturbances, or inaccurate data collection procedures [8]. These factors can distort learned representations, increase prediction errors, and reduce the ability to generalize across unfamiliar situations [9]. In complex applications, misleading predictions can also affect resource allocation, operational planning, and risk assessment [10]. Distinguishing uncertainty from data corruption provides a necessary foundation for selecting suitable preprocessing methods, model architectures, and uncertainty estimation techniques.

The development of robust deep learning architectures provides several opportunities to reduce sensitivity to imperfect data [11]. Convolutional neural networks, recurrent neural networks, transformers, autoencoders, and hybrid architectures offer complementary capabilities for extracting meaningful features from complex inputs [12]. Noise-resistant training approaches, including data augmentation, regularization, robust loss functions, curriculum learning, and sample reweighting, further support reliable representation learning [13]. Probabilistic modeling and ensemble-based methods can estimate prediction uncertainty and reveal situations where model outputs require careful interpretation. The effectiveness of these techniques depends on the noise characteristics, dataset composition, network architecture, and intended application [14]. Coordinating data preparation, model design, and training procedures therefore supports more dependable predictive performance [15].