# Preface

This book offers a transverse mathematical perspective of deep machine learning in artificial intelligence. It develops a framework of generalized transforms called multiserial and hyperserial decompositions for the sake of unifying standard and recent data representation spaces. The generalization consists of integrating expressions of several variants of convolutional neural networks and wavelet filter banks in the same analytical framework. The expressions are derived recursively, from downstream to upstream layers, for the sequences of features returned at the nodes of a general form of network architecture. The inspiring framework for the derivation of these expressions is that of the so-called M-band convolution filter banks. In addition to the inter-layer inter-node expressions, activation sequences of convolutional neural networks are also mathematically described by suitable algebraic path representations. The organization of this book is associated with the following observations.

*An intelligence constantly solves optimization problems to reach desirable objectives.* In this respect, Chapter 1 introduces the most useful and fundamental concepts for understanding the mathematics of optimization.

*Before being able to minimize its objective, an intelligence will break down the information into essential characteristics or features.* This decomposition is obtained for artificial intelligence through convenient measures and projections (convolutional integrals for instance). This is the field of functional analysis and, to be able to understand or explain the decision-making process of artificial intelligence, it is necessary to familiarize ourselves with functional analysis. The essentials of functional analysis are presented in Chapter 2.

*A decision is often made because its objective outweighs (often very slightly) the objectives of the alternative decisions.* Probabilities, because of their normalized nature, are a very suitable tool for associating a measure with a set of events related to the decision process. Entropy can then be used to measure the uncertainty (a priori) or the information (a posteriori) in the occurrence of a series of probabilistic events. Chapter 3 illustrates how we can derive parametric forms of relative information measures from probabilities and entropies. Then, it briefly describes the first (in the pioneering sense) statistical models that are historically and analytically the closest to convolutional neural networks. Chapter 4 then presents an analytical framework for unifying convolutional neural network expansions.

January 2023