

*pursuing and exploiting intrinsic low-dimensional structures of data
distributions embedded in high-dimensional spaces.*

To this end, the book will take us through a complete journey from theoretical formulation to mathematical deduction, then to computational implementation and practical applications.

Preface to Version 2.0

As an open-source book, the first version, entitled “*Learning Deep Representations of Data Distributions*”, released on August 18, 2025 has been frequently and constantly updated in the past six months. Meanwhile, based on our experience from using the book for a new course taught at the University of Hong Kong in Fall 2025 and feedback that we have received from colleagues, teaching assistants and students, we have identified numerous points throughout the book that merit further revision and expansion.

Hence, we have decided to make some substantial changes and upgrades of the content and organization of the book with a new *Version 2.0* and changed its title to:

*Principles and Practice of Deep Representation Learning
or A Mathematical Theory of Memory*

The new version also allows us to explicitly reveal strong conceptual and technical connections among materials across different chapters and sections within the book, so that the overall pedagogical value of the book, we believe, has been improved significantly over the first version.

In our opinion, the new version gives a much more unified, complete, and stream-lined presentation of this subject. The new version also incorporates newly developed theoretical insights as well as a growing number of practical applications for real-world data that have not been properly documented and systematically explained elsewhere. The Version 2.0 serves as a timely remedy to this situation.

Major Changes in Version 2.0

- We have split Chapter 3 in the first version into two chapters, now Chapter 3 and 4. The new Chapter 3 focuses on the denoising process for learning low-dimensional distributions. We have added a new section that characterizes conditions under which the process leads to generalization or memorization. The new Chapter 4 focuses on representation learning based on a lossy coding approach and the principle of maximizing information gain. Since the old version only illustrated how to apply the principle to learn image representations in the supervised setting, the new version has added a case study with the unsupervised setting. It also provides

theoretical justification for popular unsupervised learning methods such as contrastive learning and DINO in particular.

- In the updated Chapter 5 on designing deep representations via unrolled optimization, we have added a principled derivation of a causal version for the white-box transformer architecture CRATE, which is important for processing sequential data such as texts as we often see in applications.
- In the updated Chapter 6 on consistent representation learning, we added a new section that further elucidates the relationships between distribution learning and representation learning. This provides theoretical clarification and justification for popular practical autoencoding methods such as variational autoencoding and representation autoencoders.
- In the updated Chapter 7, we have added a section that provides a principled explanation of representation learning with paired data and conditioned generation through the perspective of mutual information. It provides theoretical justification for popular practical methods such as CLIP and cross attention.
- The expanded application Chapter 8 now features many more detailed implementations of distribution learning and representation learning for real-world data, including natural 2D images, natural 3D objects, human body motions, and natural languages, under almost all popular and practical settings, supervised, weakly supervised, unsupervised, and conditioned.
- The final Chapter 9 about open directions has also been significantly rewritten. We attempt to give a clearer taxonomy of different levels of intelligence so that we can clarify what we have done and understood about intelligence (with this book) and what we have not. We believe that open problems associated with more advanced forms of intelligence can be well posed so that they can be studied qualitatively and quantitatively via scientific and mathematical means, instead of remaining as a mysterious subject or merely a philosophical topic.

Contributors for Version 2.0

Besides the authors, many students and colleagues have joined this project and contributed valuable content to different parts of the book during the preparation of *Version 2.0*. Below is an incomplete list of people and their specific contributions, in alphabetical order:

- Tianzhe Chu: Experiments in Section 8.4, and AI tooling for the book.
- Prof. Shenghua Gao: Sections 8.8 and 8.9.
- Bingbing Huang: Sections 8.8 and 8.9.
- Kerui Min: Chinese translation.