Preface

The Deep Learning Era of Recommender Systems

In 1992, David Goldberg and his colleagues from Xerox Palo Alto Research Center created a recommender system using a collaborative filtering algorithm [1], marking the beginning of this 31-year history of recommender system development. In these 31 years, especially in the past eight years, this technology has been completely revolutionized by deep learning. In 2012, with the deep learning network AlexNet winning the famous ImageNet competition [2], deep learning detonated the fields of image, speech, natural language processing, and so on. And even in the fields of recommendation, advertising, and search, where internet commercialization has been the most successful and machine learning models are widely used, the wave of deep learning has also been sweeping through them. In 2015, with companies such as Microsoft, Google, Baidu, and Alibaba successfully applying deep learning models to recommendation, advertising, and other business scenarios, the field of recommender systems officially entered the era of deep learning.

As recommender system machine learning engineers (hereinafter referred to as recommendation engineers) in the era of deep learning, we are fortunate to witness the most profound and rapid technological changes. Yet we are also unfortunate because in this age of rapidly evolving technology and models, we may be on the edge of being eliminated. However, this era ultimately leaves ample room for passionate engineers to develop. While enthusiastic recommendation engineers build their own technology blueprints and enrich their technical reserves, it is hoped that this book can serve as a mental map, helping them build the technical framework of deep learning recommender systems.

The Origin of This Book

The motivation behind writing this book stems from two reasons. First, I have always had the desire to structure my knowledge of recommender systems. Second, I was invited by an editor from the Electronic Industry Press. In December 2018, after reading my technical columns and some articles from my blog, Editor Zheng Liujie contacted me and invited me to write a technical book on recommendation or advertising algorithms. At that time, I had just finished coauthoring *The Quest for Machine Learning: 100+ Interview Questions for Algorithm Engineers* with my colleagues

from Hulu. This book about machine learning interviews has received positive feedback from the market and helped many readers. This writing experience made me realize that taking something seriously and writing technical content earnestly could really benefit many readers. On the other hand, I have been working in the fields of recommendation and advertising for eight years and have witnessed the wave of deep learning in the recommender system field. Therefore, I chose the topic of "Deep Learning Recommender Systems" in the hope of sharing my limited knowledge and experience with those peers who are interested in this field.

Following the publication of the Chinese edition in 2020, the book received very positive reviews and quickly rose to become the most popular technical work in its field. Consequently, I accepted an invitation from Liu Yongchen, an editor at Cambridge University Press, to translate the book into English. I've been fortunate to collaborate with two other translators, Felice Wang and Chao Pu, both recognized experts in the recommender system domain. Together, we have not only worked on the translation but also incorporated the most recent advancements from the past three years in this rapidly evolving field.

Features of This Book

This book hopes to discuss the "classic" or "cutting-edge" technical content related to recommender systems, with a particular focus on the application of deep learning in the industry. It should be noted that this book is not an introductory book on machine learning or deep learning. Although the book will intersperse the introduction of basic machine learning knowledge, most of the content assumes that readers have some machine learning background. Additionally, this book is not a purely theoretical book, but a technical book that introduces the application methods of deep learning in the field of recommender systems and industry cutting-edge technology related to recommender systems from the perspective of engineers' practical experience.

Readers of This Book

The target audience for this book can be divided into two categories.

The first group consists of professionals in the internet industry, especially those in the fields of recommendation, advertising, and search. By studying this book, these peers can become familiar with the development of deep learning recommender systems and the details of each key model and technology. This could enable them to apply or even improve these technologies in their work.

The second group consists of enthusiasts and students with some background in machine learning who wish to enter the field of recommender systems. This book aims to introduce the relevant principles and application methods of recommender system technology from a practical perspective, starting with the details and using plain language, helping readers to build a cutting-edge and practical knowledge framework of recommender systems from scratch.

Discussions Are Welcomed

The field of deep learning recommender systems is evolving rapidly, but my knowledge is limited, and it is inevitable that there will be some omissions and mistakes. I sincerely hope to work together with readers to iterate on and improve our knowledge of deep learning recommender systems. Please feel free to provide feedback during the reading process. Whether it is to point out errors, make improvement suggestions, or discuss technical issues with me, you can contact me through the following methods:

Email: wzhe06@gmail.com

I will respond to valuable feedback as soon as possible.

Appreciations

The process of writing this book has not been easy. Aside from devoting almost all of my free time to writing, it also required a significant amount of time to review research papers, organize technical frameworks, and even communicate with peers and authors in various companies to discuss technical details and stay up to date with the latest technological applications in the industry. I am deeply grateful to the industry colleagues who provided invaluable assistance for this book.

During the writing process, the editor, Liu Yongchen, offered numerous valuable and constructive suggestions for the book and made many professional edits to address detailed issues. I am deeply grateful to Liu Yongchen and the editors at Cambridge University Press who have contributed to this book.

I also wish to express my gratitude to my two coauthors, who diligently and meticulously completed the translation of this book to a high standard. Simultaneously, they enriched the book by including numerous cutting-edge developments. It has been an honor to work with them, and I have learned a great deal from our collaboration.

Last but not least, I would also like to thank my wife and daughter for their immense support and understanding during the writing process. Your care for our family and support for my work has been my greatest motivation in completing this book.

Thank you all!

Foster City, San Francisco Bay Area, USA
Wang Zhe

References

[1] David Goldberg, et al. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 1992: 61–71.

[2] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*. 2012. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf