

PREFACE

Twenty-five years have passed since the publication of the Russian version of the book *Estimation of Dependencies Based on Empirical Data* (*EDBED* for short). Twenty-five years is a long period of time. During these years many things have happened. Looking back, one can see how rapidly life and technology have changed, and how slow and difficult it is to change the theoretical foundation of the technology and its philosophy.

I pursued two goals writing this Afterword: to update the technical results presented in *EDBED* (the easy goal) and to describe a general picture of how the new ideas developed over these years (a much more difficult goal).

The picture which I would like to present is a very personal (and therefore very biased) account of the development of one particular branch of science, *Empirical Inference Science*.

Such accounts usually are not included in the content of technical publications. I have followed this rule in all of my previous books. But this time I would like to violate it for the following reasons. First of all, for me *EDBED* is the important milestone in the development of empirical inference theory and I would like to explain why. Second, during these years, there were a lot of discussions between supporters of the new paradigm (now it is called the VC theory¹) and the old one (classical statistics). Being involved in these discussions from the very beginning I feel that it is my obligation to describe the main events.

The story related to the book, which I would like to tell, is the story of how it is difficult to overcome existing prejudices (both scientific and social), and how one should be careful when evaluating and interpreting new technical concepts.

This story can be split into three parts that reflect three main ideas in the development of empirical inference science: from the pure technical (mathematical) elements of the theory to a new paradigm in the philosophy of generalization.

¹VC theory is an abbreviation for Vapnik–Chervonenkis theory. This name for the corresponding theory appeared in the 1990s after *EDBED* was published.

The first part of the story, which describes the main technical concepts behind the new mathematical and philosophical paradigm, can be titled

Realism and Instrumentalism: Classical Statistics and VC Theory

In this part I try to explain why between 1960 and 1980 a new approach to empirical inference science was developed in contrast to the existing classical statistics approach developed between 1930 and 1960.

The second part of the story is devoted to the rational justification of the new ideas of inference developed between 1980 and 2000. It can be titled

Falsifiability and Parsimony: VC Dimension and the Number of Entities

It describes why the concept of VC falsifiability is more relevant for predictive generalization problems than the classical concept of parsimony that is used both in classical philosophy and statistics.

The third part of the story, which started in the 2000s can be titled

Noninductive Methods of Inference: Direct Inference Instead of Generalization

It deals with the ongoing attempts to construct new predictive methods (direct inference) based on the new philosophy that is relevant to a complex world, in contrast to the existing methods that were developed based on the classical philosophy introduced for a simple world.

I wrote this Afterword with my students' students in mind, those who just began their careers in science. To be successful they should learn something very important that is not easy to find in academic publications.

In particular they should see the big picture: what is going on in the development of this science and in closely related branches of science in general (not only about some technical details). They also should know about the existence of very intense paradigm wars. They should understand that the remark of Cicero, "Among all features describing genius the most important is inner professional honesty", is not about ethics but about an intellectual imperative. They should know that Albert Einstein's observation about everyday scientific life that "Great spirits have always encountered violent opposition from mediocre minds," is still true. Knowledge of these things can help them to make the right decisions and avoid the wrong ones. Therefore I wrote a fourth part to this Afterword that can be titled

The Big Picture.

This, however, is an extremely difficult subject. That is why it is wise to avoid it in technical books, and risky to discuss it commenting on some more or less recent events in the development of the science.

Writing this Afterword was a difficult project for me and I was able to complete it in the way that it is written due to the strong support and help of my colleagues Mike Miller, David Waltz, Bernhard Schölkopf, Leon Bottou, and Ilya Muchnik.

I would like to express my deep gratitude to them.

Princeton, New Jersey,
November 2005

Vladimir Vapnik



<http://www.springer.com/978-0-387-30865-4>

Estimation of Dependences Based on Empirical Data

Vapnik, V.

2006, XVIII, 505 p., Hardcover

ISBN: 978-0-387-30865-4