Preface

> Talent wins games, but teamwork and intelligence wins championships.
>
> Michael Jordan

Ensemble analysis is a widely used class of meta-algorithms for many data mining problems such as classification and clustering. Numerous ensemble-based algorithms have been proposed in the literature for these problems. Compared to the clustering and classification problems, ensemble analysis has been studied in a limited way in the context of outlier detection. It is only in recent years that the problem of outlier ensembles has been recognized formally, and some techniques for these problems have been proposed. However, the material in the field has not been formally recognized, and it is expected that this book will play a significant role in creating a structured exposition of the topic. This book discusses a variety of methods for outlier ensembles and organizes them by the specific principles with which accuracy improvements are achieved. In addition, a discussion is provided on the techniques with which such methods can be made more effective. A formal classification of these methods has been provided, and the circumstances in which they work well are discussed. A discussion is also provided on how outlier ensembles relate (both theoretically and practically) to the ensemble techniques used commonly for other data mining problems such as classification. The similarities and (subtle) differences in the ensemble techniques for the classification and outlier detection problems are discussed. These subtle differences do impact the design of ensemble algorithms for the latter problem.

Ensemble techniques are increasingly finding their way into the curricula of many data mining courses. This book can be used for such courses. Many illustrative examples and exercises are provided in order to facilitate classroom teaching. A familiarity is assumed to the outlier detection problem and also to the generic problem of ensemble analysis in classification. This is because many of the ensemble methods discussed in this book are adaptations from their counterparts in the classification domain. Some techniques discussed in this book, such as

wagging, randomized feature weighting, and geometric subsampling, provide new insights that are not available elsewhere. We have also provided an analysis of the performance of various types of base detectors and their relative effectiveness. This book combines a textbook-style discussion of older topics with new insights. Therefore, we believe that the book will also be of interest to researchers and practitioners for leveraging ensemble methods into optimal algorithmic design.

Yorktown Heights, NY, USA Charu C. Aggarwal
January 2017 Saket Sathe

http://www.springer.com/978-3-319-54764-0

Outlier Ensembles
An Introduction
Aggarwal, C.C.; Sathe, S.
2017, XVI, 276 p. 55 illus., 9 illus. in color., Hardcover
ISBN: 978-3-319-54764-0