Preface

Mixture models have been used for classification and clustering for over 50 years, and the topic has received increasing attention since the end of the last century. In fact, these mixture model-based approaches have burgeoned into an important subfield of classification. The volume of work on the subject and the breadth of its application justify its treatment in a monograph. However, the genesis of this book lies in my experiences introducing the topic to new Ph.D. students. The task of starting a Ph.D. student on a course of research on model-based approaches to clustering and classification previously involved assigning much, often cumbersome, reading material. This reading material consists of various media, mainly journal articles and book chapters, that use different approaches to make generally similar points. The different notation and arguments do not present a major problem in themselves; however, there is a distinct lack of consistency on some crucial points and this is difficult for a starting Ph.D. student. The origin of mixture models as a tool for clustering was not clearly understood, which deprived students of a natural beginning in understanding why this is a good idea. This matter is finally resolved at the beginning of Chapter 2 so that future students can understand that the initial question concerned how a cluster should be defined; it was ten years later that what we now call model-based clustering was first executed. A copy of a key text was provided to me by Lizette Royer Barton, a reference archivist at the Drs. Nicholas and Dorothy Cummings Center for the History of Psychology, University of Akron, Ohio, and I am very grateful to her in this regard. I am also most grateful to John Wolfe for providing a copy of his famous but also elusive master's thesis, from which I was able to get to the work of Tiedeman (1955) and others.

This monograph is laid out to progress naturally. Broadly, the three chapters immediately following the introduction cover Gaussian mixtures, mixtures of factor analyzers and extensions thereof, variable selection, and high-dimensional applications. Then the departure from the Gaussian assumption begins with consideration of mixtures of distributions that parameterize concentration, or tail weight. This departure continues with mixtures of skewed distributions, mixtures of distributions that parameterize both skewness and concentration, and mixtures of multiple scaled distributions. The remaining chapters mostly cover special topics, so to speak, including methods for clustering and classification of longitudinal data, cluster-weighted models, averaging mixture models, and so on. Notable exceptions are Section 9.1, wherein the definition of a cluster is explored, and Section 9.2, where the existence

of a "best" clustering and classification method is discussed. While the layout is intended to assist a beginning Ph.D. student, it will equally benefit anyone who is coming to the subject anew or afresh. More seasoned readers will find the chapter and section titles self-explanatory. While its genesis lies in experiences gained through supervising Ph.D. students, this is a book for anyone with an interest in, or a use for, mixture model-based approaches to classification and clustering. The question of how much mathematical detail to include is addressed in a practical manner: where including mathematical detail is considered helpful in achieving a broad understanding, it is included; otherwise, the reader is directed elsewhere. Moreover, there are already several very good books on mixture models (e.g., Everitt and Hand, 1981; Titterington et al., 1985; McLachlan and Peel, 2000a; Frühwirth–Schnatter, 2006) and this monograph does not seek to replicate their contents.

I am most grateful to David Grubbs of the Taylor & Francis Group for his suggestion, back in 2011, that I write a monograph and for his unerring patience as I repeatedly missed deadlines. I am thankful to Dr. Brendan Murphy for starting me on my ongoing journey through the field. A number of my current and former supervisees kindly provided comments on parts of this book, while others helped with figures and analyses. In this regard, I am thankful for the assistance of Dr. Jeffrey Andrews, Dr. Ryan Browne, Dr. Utkarsh Dang, Dr. Sanjeena Dang, Dr. Matthieu Marbac, Paula Murray, Yang Tang, Dr. Cristina Tortora, and Dr. Irene Vrbik. I am particularly indebted to Dr. Brian Franczak for his invaluable assistance with the figures and analyses in Chapters 6 and 7. I am most grateful to Dr. Douglas Steinley, who read a draft of this monograph and provided some very helpful comments. I also wish to thank Drs. Salvatore Ingrassia and Antonio Punzo, who provided very helpful feedback on parts of this monograph.

I have been very fortunate to work with several wonderful collaborators on real clustering problems; in this regard, I am particularly grateful to Dr. Chris Findlay and John Castura, both of Compusense Inc. I wish to express very special and heartfelt thanks to Professor Petros Florides, Eamonn Mullins, Dr. Donal O'Donovan, and Dr. Myra O'Regan, all either currently or formally of Trinity College Dublin, Ireland. Suffice it to say that without their respective interventions and kindnesses, I probably would not have completed even an undergraduate degree. A final word of thanks goes to my wife, Sharon; without her patience and support, this book would not have come to fruition.

Guelph Paul D. McNicholas