Mixture Models for Clustering and Dimension Reduction
Verbeek, J.J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CONTENTS

1. Introduction .......................................................... 1

2. A review of clustering and dimension reduction techniques .............. 7
   2.1 Clustering .......................................................... 7
       2.1.1 Hierarchical clustering ...................................... 8
       2.1.2 Partitional clustering ........................................ 11
       2.1.3 Spectral clustering .......................................... 13
       2.1.4 Comparison of clustering methods .......................... 16
   2.2 Dimension reduction .............................................. 17
       2.2.1 Principal components and generalizations .............. 20
       2.2.2 Neural network based methods ............................. 26
       2.2.3 Similarity based methods ................................. 29
       2.2.4 Comparison of dimension reduction methods ............ 37

3. Mixture density estimation ........................................... 41
   3.1 The EM algorithm and Gaussian mixture densities .................. 41
       3.1.1 Mixture densities ........................................... 42
       3.1.2 Parameter estimation with the EM algorithm ............ 43
       3.1.3 Model selection ............................................ 47
       3.1.4 Gaussian mixture models .................................. 48
   3.2 Efficient greedy learning of Gaussian mixtures ................... 53
       3.2.1 Greedy learning of Gaussian mixtures .................... 53
       3.2.2 Efficient search for new components ....................... 56
       3.2.3 Related work .............................................. 59
       3.2.4 Experimental results ....................................... 61
       3.2.5 Conclusions ............................................... 65
   3.3 A greedy approach to k-means clustering .......................... 66
       3.3.1 The global k-means algorithm .............................. 67
       3.3.2 Speeding up execution of global k-means ............... 68
       3.3.3 Experimental results ..................................... 69
       3.3.4 Conclusions ............................................... 75
   3.4 Accelerating the EM algorithm for large data sets ................... 75
       3.4.1 An accelerated EM algorithm for Gaussian mixtures ....... 76
## Contents

3.4.2 Related work .................................................. 78
3.4.3 Experimental results ...................................... 79
3.4.4 Conclusions .................................................. 83

4. Self-organizing mixture models .................................. 85
   4.1 Self-organizing maps ...................................... 85
   4.2 Self-organizing mixture models ......................... 88
      4.2.1 A constrained EM algorithm for self-organizing maps ..... 88
      4.2.2 Modelling data with missing values .................. 92
      4.2.3 The adaptive-subspace self-organizing map .......... 93
   4.3 Comparison with related work ........................... 94
   4.4 Experimental results ..................................... 98
   4.5 Conclusions ................................................ 105

5. Combining local linear models to form global non-linear models .... 107
   5.1 Introduction ............................................... 107
   5.2 Combining local linear models ......................... 109
      5.2.1 Mixtures of linear models ............................ 110
      5.2.2 Aligning local linear models ..................... 111
      5.2.3 Closed-form update equations .................... 115
      5.2.4 Parameter initialization ........................... 117
      5.2.5 Experimental results .............................. 122
   5.3 Learning mappings between manifolds ................... 127
      5.3.1 The probabilistic model ............................ 129
      5.3.2 Parameter initialization ........................... 131
      5.3.3 Experimental results .............................. 134
   5.4 Conclusions ................................................ 140

6. Conclusion and discussion .................................... 145
   6.1 Summary of conclusions ................................ 145
   6.2 Discussion of directions for further research ........ 147

Bibliography .......................................................... 149

Summary ............................................................. 161

Curriculum vitae ..................................................... 163