

Part A : Introduction to Some Pattern Recognition Methods

1. Basic Concepts	2
1.1. First Ideas of Pattern Recognition,	2
1.2. Pattern Space,	3
1.3. Binary Classifiers,	5
1.4. Training and Evaluation of Classifiers,	8
1.5. Additional Aspects,	11
1.6. Warning,	11
1.7. Applications of Pattern Recognition,	13
1.8. Literature,	14
1.8.1. General Pattern Recognition,	14
1.8.2. Pattern Recognition from the Chemist's Point of View,	15
2. Computation of Binary Classifiers	18
2.1. Classification by Distance Measurements to Centres of Gravity	18
2.1.1. Principle,	18
2.1.2. Centres of Gravity in a d-Dimensional Space,	19
2.1.3. Classification by Distance Measurements,	19
2.1.4. Classification by the Symmetry Plane,	21
2.1.5. Classification by Mean Vectors,	22
2.1.6. Evaluation,	23
2.1.7. Projection of Pattern Points on a Hypersphere,	24
2.1.8. Distance Measurements in Pattern Space (Overview),	25
2.1.9. Distance Measurements with Weighted Features (Generalized Distances),	27
2.1.10. Chemical Applications,	29
2.2. Learning Machine	30
2.2.1. Principle,	30
2.2.2. Initial Weight Vector,	32
2.2.3. Correction of the Weight Vector,	33
2.2.4. Methods of Training,	35
2.2.5. Restrictions,	37
2.2.6. Evaluation,	38
2.2.7. Dead Zone Training,	39
2.2.8. Chemical Applications,	41

2.3. Linear Regression (Least-Squares Classification)	42
2.3.1. Principle,	42
2.3.2. Mathematical Treatment,	43
2.3.3. Characteristics and Variations of the Method,	46
2.3.4. Chemical Applications,	48
2.4. Simplex Optimization of Classifiers	48
2.4.1. Principle,	48
2.4.2. Starting the Simplex,	50
2.4.3. Response Function,	51
2.4.4. Moving the Simplex,	52
2.4.5. Halting the Simplex,	55
2.4.6. Chemical Applications,	55
2.5. Piecewise-Linear Classifiers	56
2.6. Implementation of Binary Classifiers	57
2.6.1. Discrete or Continuous Response,	57
2.6.2. Classification by a Committee of Classifiers,	59
2.6.3. Multicategory Classification,	59
3. K - Nearest Neighbour Classification (KNN-Method)	62
3.1. Principle,	62
3.2. Maximum Risk of KNN-Classifications,	63
3.3. Characteristics and Variations of the KNN-Method,	64
3.4. Classification with Potential Functions,	65
3.5. KNN-Classification with a Condensed Data Set,	69
3.6. Chemical Applications,	69
4. Classification by Adaptive Networks	72
4.1. Perceptron,	72
4.2. Adaptive Digital Learning Network,	74
4.3. Chemical Applications,	75
5. Parametric Classification Methods	78
5.1. Principle,	78
5.2. Bayes- and Maximum Likelihood Classifiers,	78
5.3. Estimation of Probability Densities,	80
5.4. Bayes- and Maximum Likelihood Classifiers for Binary Encoded Patterns,	83
5.5. A Simple Sequential Classification Method Based on Probability Densities,	85
5.6. Chemical Applications,	87

6. Modelling of Clusters	88
6.1. Principle,	88
6.2. Modelling by a Hypersphere,	88
6.3. SIMCA-Method,	89
7. Clustering Methods	92
7.1. Principle,	92
7.2. Hierarchical Clustering,	93
7.3. Minimal Spanning Tree Clustering,	95
7.4. Chemical Applications,	96
8. Display Methods	97
8.1. Principle,	97
8.2. Linear Methods,	97
8.3. Nonlinear Methods,	100
8.4. Chemical Applications,	101
9. Preprocessing	102
9.1. Principle,	102
9.2. Scaling,	102
9.3. Weighting,	103
9.4. Transformation,	104
9.5. Combination of Features,	104
10. Feature Selection	106
10.1. Principle,	106
10.2. Feature Selection by Using Data Statistics,	107
10.2.1. Variance Weighting,	107
10.2.2. Fisher Weighting,	108
10.2.3. Use of Probability Density Curves,	109
10.2.4. Methods for Binary Encoded Patterns,	110
10.2.5. Elimination of Correlations,	111
10.3. Feature Selection by Using Classification Results,	112
10.3.1. Evaluation of the Components of a Single Weight Vector,	112
10.3.2. Feature Selection with the Learning Machine,	112
10.3.3. Other Methods,	113
10.4. Number of Intrinsic Dimensions and Number of Patterns (n/d - Problem),	114

11. Evaluation of Classifiers	118
11.1. Principle,	118
11.2. Predictive Abilities,	119
11.3. Loss,	123
11.4. A posteriori Probabilities,	123
11.5. Terminology Problems,	127
11.6. Application of Information Theory,	129
11.6.1. Introduction to Information Theory,	129
11.6.2. Transinformation,	131
11.6.3. Figure of Merit,	136
11.7. Evaluation of Classifiers with Continuous Response,	136
11.8. Confidence of Predictive Abilities,	138
11.9. Comparison with the Capability of Chemists,	140

Part B : Application of Pattern Recognition Methods in Chemistry

12. General Aspects of Pattern Recognition in Chemistry	142
13. Spectral Analysis	145
13.1. Mass Spectrometry,	145
13.1.1. Survey,	145
13.1.2. Representation of Mass Spectra as Pattern Vectors,	146
13.1.3. Determination of Molecular Formulas and Molecular Weights,	150
13.1.4. Recognition of Molecular Structures,	152
13.1.5. Chemical Interpretation of Mass Spectral Classifiers,	154
13.1.6. Simulation of Mass Spectra,	155
13.1.7. Miscellaneous,	156
13.2. Infrared Spectroscopy,	157
13.3. Raman Spectroscopy,	160
13.4. Nuclear Magnetic Resonance Spectroscopy,	161
13.5. Gamma-Ray Spectroscopy,	164
13.6. Combined Spectral Data,	165
14. Chromatography	166
14.1. Gas Chromatography,	166
14.2. Thin Layer Chromatography,	168

15. Electrochemistry	169
16. Classification of Materials and Chemical Compounds	171
16.1. Technology, 171	
16.2. Archaeology, 173	
16.3. Food, 174	
16.4. Biology, 175	
16.5. Chemistry, 175	
17. Relationships between Chemical Structure and Biological Activity	177
17.1. Pharmacological Activity, 177	
17.1.1. General Remarks, 177	
17.1.2. Sedatives and Tranquilizers, 178	
17.1.3. Cancer and Tumors, 179	
17.1.4. Miscellaneous, 180	
17.2. Odour Classification, 181	
17.3. Spectra - Activity Relationships, 182	
18. Clinical Chemistry	184
19. Environmental Chemistry	185
19.1. Petroleum Pollutants, 185	
19.2. Atmospheric Particulates, 186	
19.3. Miscellaneous, 187	
20. Classification of Analytical Methods	188
<u>Part C : Appendix</u>	
21. Literature	190
21.1. Pattern Recognition in Chemistry, 190	
21.2. General Pattern Recognition, 200	
21.3. Other Literature, 202	
21.4. List of Authors, 203	
22. Subject Index	212