Contents

Page

Foreword ix

Introduction x

1 Scope 1

2 Normative references 2
2.1 International standards 2
2.2 Publicly-available specifications 2

3 Terms, definitions, notations and conventions 3
3.1 Terms and definitions 3
3.2 Notations 10
3.3 Conventions 11

4 Concepts 12
4.1 Data mining techniques 12
4.2 Computational phases 12
4.2.1 Training phase 12
4.2.2 Model introspection phase 13
4.2.3 Application phase 15
4.2.4 Testing phase 16
4.3 Mapping on user-defined types 17
4.3.1 Target application 17
4.3.2 The relationships of the user-defined types 17
4.3.3 User-defined types that are not related to mining techniques 19
4.3.4 User-defined types that are related to mining techniques 19

5 Data Mining Data Types 22
5.1 DM_LogicalDataSpec Type and Routines 22
5.1.1 DM_LogicalDataSpec Type 22
5.1.2 DM_addDataSpecFld Method 24
5.1.3 DM_remDataSpecFld Method 25
5.1.4 DM_getNumFields Method 26
5.1.5 DM_getFldName Method 27
5.1.6 DM_setFldType Method 28
5.1.7 DM_getFldType Method 30
5.1.8 DM_isCompatible Method 31
5.2 DM_MiningData Type and Routines 32
5.2.1 DM_MiningData Type 32
5.2.2 DM_defMiningData Method 34
5.2.3 DM_setFldAlias Method 35
5.2.4 DM_genDataSpec Method 36
5.3 DM_ApplicationData Type and Routines 37
5.3.1 DM_ApplicationData Type 37
5.3.2 DM_genDataSpec Method 38
5.3.3 DM_impApplData Method 39

6 Association Rules 40
6.1 DM_RuleSettings Type and Routines 40
6.1.1 DM_RuleSettings Type 40
6.1.2 DM_impRuleSettings Method 43

6.1.3 DM_expRuleSettings Method 44
6.1.4 DM_setRuleFilter Method 45
6.1.5 DM_getRuleFilter Method 46
6.1.6 DM_setMinSupport Method 47
6.1.7 DM_getMinSupport Method 48
6.1.8 DM_setMinConf Method 49
6.1.9 DM_getMinConf Method 50
6.1.10 DM_setMaxLength Method 51
6.1.11 DM_getMaxLength Method 52
6.1.12 DM_useRuleDataSpec Method 53
6.1.13 DM_getRuleDataSpec Method 54
6.1.14 DM_setGroup Method 55
6.1.15 DM_getGroup Method 56
6.1.16 DM_setSequence Method 57
6.1.17 DM_getSequence Method 58
6.2 DM_RuleBldTask type and Routines 59
6.2.1 DM_RuleBldTask Type 59
6.2.2 DM_defRuleBldTask Method 60
6.2.3 DM_getRuleTrnData Method 61
6.2.4 DM_getRuleSettings Method 62
6.2.5 DM_buildRuleModel Method 63
6.3 DM_RuleModel Type and Routines 64
6.3.1 DM_RuleModel Type 64
6.3.2 DM_impRuleModel Method 67
6.3.3 DM_getRuleMdlType Method 68
6.3.4 DM_expRuleModel Method 69
6.3.5 DM_applyRuleModel Method 70
6.3.6 DM_getNumItemSets Method 71
6.3.7 DM_getNumSequences Method 72
6.3.8 DM_getNumRules Method 73
6.3.9 DM_filterRules Method 74
6.3.10 DM_getNumTransacts Method 75
6.3.11 DM_getItemSets Method 76
6.3.12 DM_getRules Method 77
6.3.13 DM_getRuleBodies Method 78
6.3.14 DM_getSequences Method 79
6.3.15 DM_getSeqDetails Method 80
6.3.16 DM_getSeqRules Method 81
6.3.17 DM_getRuleBldTask method 83
6.4 DM_RuleFilter Type and Routines 84
6.4.1 DM_RuleFilter Type 84
6.4.2 DM_addRangeConstr Method 86
6.4.3 DM_remRangeConstr Method 88
6.4.4 DM_getRangeConstrs Method 89
6.4.5 DM_setMaxNumRules Method 90
6.4.6 DM_getMaxNumRules Method 91
6.4.7 DM_addCountConstr Method 92
6.4.8 DM_remCountConstr Method 93
6.4.9 DM_getCountConstrs Method 94
6.4.10 DM_setItemConstr Method 95
6.4.11 DM_getItemConstrs Method 97
6.5 DM_RuleApplTask Type and Routines 98
6.5.1 DM_RuleApplTask Type 98
6.5.2 DM_defRuleApplTask Method 100
6.5.3 DM_getRuleMdl Method 102
6.5.4 DM_getRuleApplData Method 103
6.5.5 DM_getRuleOutpData Method 104
6.5.6 DM_getRuleGroupFld method 105
6.5.7 DM_applyRuleMdl method 106

7 Clustering 107
7.1 DM_ClusSettings Type and Routines 107

7.1.1 DM_ClusSettings Type 107
7.1.2 DM_impClusSettings Method 109
7.1.3 DM_expClusSettings Method 110
7.1.4 DM_setMaxNumClus Method 111
7.1.5 DM_getMaxNumClus Method 112
7.1.6 DM_setFldWeight Method 113
7.1.7 DM_getFldWeight Method 114
7.1.8 DM_setFldSimScale Method 115
7.1.9 DM_getFldSimScale Method 116
7.1.10 DM_useClusDataSpec Method 117
7.1.11 DM_getClusDataSpec Method 118
7.2 DM_ClusBldTask type and Routines 119
7.2.1 DM_ClusBldTask Type 119
7.2.2 DM_defClusBldTask Method 120
7.2.3 DM_getClusTrnData Method 121
7.2.4 DM_getClusSettings Method 122
7.2.5 DM_buildClusModel Method 123
7.3 DM_ClusteringModel Type and Routines 124
7.3.1 DM_ClusteringModel Type 124
7.3.2 DM_impClusModel Method 126
7.3.3 DM_expClusModel Method 127
7.3.4 DM_getNumClusters Method 128
7.3.5 DM_getClusters Method 129
7.3.6 DM_getMdlQuality Method 130
7.3.7 DM_applyClusModel Method 131
7.3.8 DM_getClusTask method 132
7.3.9 DM_getClusMdlSpec method 133
7.3.10 DM_getFields 134
7.4 DM_ClusApplTask Type and Routines 135
7.4.1 DM_ClusApplTask Type 135
7.4.2 DM_defClusApplTask Method 137
7.4.3 DM_getClusMdl Method 138
7.4.4 DM_getClusApplData Method 139
7.4.5 DM_getClusOutpData Method 140
7.4.6 DM_getClusResFld method 141
7.4.7 DM_applyClusMdl method 142
7.5 DM_ClusResult Type and Routines 143
7.5.1 DM_ClusResult Type 143
7.5.2 DM_getClusterID Method 144
7.5.3 DM_getQuality Method 145

8 Classification 146
8.1 DM_ClasSettings Type and Routines 146
8.1.1 DM_ClasSettings Type 146
8.1.2 DM_impClasSettings Method 149
8.1.3 DM_expClasSettings Method 150
8.1.4 DM_setClasCostRate Method 151
8.1.5 DM_getClasCostRate Method 152
8.1.6 DM_setClasCost Method 153
8.1.7 DM_getClasCost Method 154
8.1.8 DM_useClasDataSpec Method 155
8.1.9 DM_getClasDataSpec Method 156
8.1.10 DM_setClasTarget Method 157
8.1.11 DM_getClasTarget Method 158
8.2 DM_ClasBldTask Type and Routines 159
8.2.1 DM_ClasBldTask Type 159
8.2.2 DM_defClasBldTask Method 161
8.2.3 DM_getClasTrnData Method 162
8.2.4 DM_getClasValData Method 163
8.2.5 DM_getClasSettings Method 164
8.2.6 DM_buildClasModel Method 165
8.3 DM_ClasModel Type and Routines 166

8.3.1 DM_ClasModel Type ... 166
8.3.2 DM_impClasModel Method ... 169
8.3.3 DM_expClasModel Method ... 170
8.3.4 DM_applyClasModel Method ... 171
8.3.5 DM_testClasModel Method ... 172
8.3.6 DM_getClasCostRate Method ... 173
8.3.7 DM_isValidated Method ... 174
8.3.8 DM_getRankQuality method ... 175
8.3.9 DM_getPredAccuracy Method ... 176
8.3.10 DM_getReliability Method ... 177
8.3.11 DM_getMdlQuality Method ... 178
8.3.12 DM_getGainsChart method ... 179
8.3.13 DM_getClasTask method ... 180
8.3.14 DM_getClasMdlSpec method ... 181
8.3.15 DM_getFields ... 182
8.3.16 DM_getClasTarget method ... 183
8.4 DM_ClasTestTask Type and Routines ... 184
8.4.1 DM_ClasTestTask Type ... 184
8.4.2 DM_defClasTestTask Method ... 185
8.4.3 DM_getClasTestData Method ... 186
8.4.4 DM_getClasTestMdl Method ... 187
8.4.5 DM_testClasMdl Method ... 188
8.5 DM_ClasTestResult Type and Routines ... 189
8.5.1 DM_ClasTestResult Type ... 189
8.5.2 DM_getClasError Method ... 191
8.5.3 DM_getRankQuality Method ... 192
8.5.4 DM_getPredAccuracy Method ... 193
8.5.5 DM_getReliability Method ... 194
8.5.6 DM_getMdlQuality Method ... 195
8.5.7 DM_getGainsChart Method ... 196
8.5.8 DM_getNumClasses Method ... 197
8.5.9 DM_getClass Method ... 198
8.5.10 DM_getClasPreds Method ... 199
8.6 DM_ClasApplTask Type and Routines ... 200
8.6.1 DM_ClasApplTask Type ... 200
8.6.2 DM_defClasApplTask Method ... 202
8.6.3 DM_getClasMdl Method ... 203
8.6.4 DM_getClasApplData Method ... 204
8.6.5 DM_getClasOutpData Method ... 205
8.6.6 DM_getClasResFld method ... 206
8.6.7 DM_applyClasMdl method ... 207
8.7 DM_ClasResult Type and Routines ... 208
8.7.1 DM_ClasResult Type ... 208
8.7.2 DM_getPredClass Method ... 209
8.7.3 DM_getConfidence Method ... 210

9 Regression ... 211
9.1 DM_RegSettings Type and Routines ... 211
9.1.1 DM_RegSettings Type ... 211
9.1.2 DM_impRegSettings Method ... 213
9.1.3 DM_expRegSettings Method ... 214
9.1.4 DM_setRegRSquared Method ... 215
9.1.5 DM_getRegRSquared Method ... 216
9.1.6 DM_useRegDataSpec Method ... 217
9.1.7 DM_getRegDataSpec Method ... 218
9.1.8 DM_setRegTarget Method ... 219
9.1.9 DM_getRegTarget Method ... 220
9.2 DM_RegBldTask type and Routines ... 221
9.2.1 DM_RegBldTask Type ... 221
9.2.2 DM_defRegBldTask Method ... 223
9.2.3 DM_getRegTrnData Method ... 224
9.2.4 DM_getRegValData Method ... 225

9.2.5 DM_getRegSettings Method 226
9.2.6 DM_buildRegModel Method 227
9.3 DM_RegressionModel Type and Routines 228
9.3.1 DM_RegressionModel Type 228
9.3.2 DM_impRegModel Method 231
9.3.3 DM_expRegModel Method 232
9.3.4 DM_applyRegModel Method 233
9.3.5 DM_testRegModel Method 234
9.3.6 DM_getRegRSquared method 235
9.3.7 DM_isValidated Method 236
9.3.8 DM_getRankQuality Method 237
9.3.9 DM_getPredAccuracy Method 238
9.3.10 DM_getReliability Method 239
9.3.11 DM_getMdlQuality Method 240
9.3.12 DM_getRMSE Method 241
9.3.13 DM_getGainsChart method 242
9.3.14 DM_getRegTask method 243
9.3.15 DM_getRegMdlSpec method 244
9.3.16 DM_getFields method 245
9.3.17 DM_getRegTarget method 246
9.4 DM_RegTestTask Type and Routines 247
9.4.1 DM_RegTestTask Type 247
9.4.2 DM_defRegTestTask Method 248
9.4.3 DM_getRegTestData Method 249
9.4.4 DM_getRegTestMdl Method 250
9.4.5 DM_testRegMdl Method 251
9.5 DM_RegTestResult Type and Routines 252
9.5.1 DM_RegTestResult Type 252
9.5.2 DM_getPredError Method 254
9.5.3 DM_getRankQuality Method 255
9.5.4 DM_getPredAccuracy Method 256
9.5.5 DM_getReliability Method 257
9.5.6 DM_getMdlQuality 258
9.5.7 DM_getRMSE Method 259
9.5.8 DM_getGainsChart Method 260
9.6 DM_RegApplTask Type and Routines 261
9.6.1 DM_RegApplTask Type 261
9.6.2 DM_defRegApplTask Method 263
9.6.3 DM_getRegMdl Method 264
9.6.4 DM_getRegApplData Method 265
9.6.5 DM_getRegOutpData Method 266
9.6.6 DM_getRegResFld method 267
9.6.7 DM_applyRegMdl method 268
9.7 DM_RegResult Type and Routines 269
9.7.1 DM_RegResult Type 269
9.7.2 DM_getPredValue Method 270

10 Status Codes 271

11 Conformance 273
11.1 Requirements for conformance 273
11.1.1 Support for user-defined types and functions 273
11.1.2 Support for data mining phases 273
11.1.3 Support for basic user-defined types 273
11.1.4 Support for data mining techniques 273
11.2 Claims of conformance 274

Annex A (informative) Implementation-defined Elements 276

Annex B (informative) Implementation-dependent elements 278

Annex C (informative) A Scenario Using the Classification Technique 279

Index **281**

Figures Page Figure 1 -- The training phase **13**

Figure 2 -- Gains chart **15**

Figure 3 -- The application phase **15**

Figure 4 -- The testing phase **17**

Figure 5 -- Overview of user-defined types for the training phase **17**

Figure 6 -- Overview of user-defined types for training, testing and application **18**

Figure 7 -- Overview of user-defined types for bulk application **19**

Tables Page Table 1 -- Model interpretation **13**

Table 2 -- Gains chart information **14**

Table 3 -- Items inferred from an association rule model **16**

Table 4 -- Groups of items inferred from an association rule model **16**

Table 5 -- Values for field types **28**

Table 6 -- Values for operators **86**

Table 7 -- Values for restriction criteria **87**

Table 8 -- Values for constraint types **95**

Table 9 -- SQLSTATE class and subclass values **272**