

---

---

**Information technology — Database  
languages — SQL multimedia and  
application packages —**

**Part 6:  
Data mining**

**iTeh STANDARD PREVIEW**  
*Technologies de l'information — Langages de bases de données —  
Multimédia SQL et paquetages d'application —  
(standards.iteh.ai)  
Partie 6. Exploration de données*

ISO/IEC 13249-6:2002

<https://standards.iteh.ai/catalog/standards/sist/9673eece-9fcd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002>

**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC 13249-6:2002](https://standards.iteh.ai/catalog/standards/sist/9673eece-9fcd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002)

<https://standards.iteh.ai/catalog/standards/sist/9673eece-9fcd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002>

© ISO/IEC 2002

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

Contents	Page
Foreword .....	viii
Introduction .....	ix
1 Scope .....	1
2 Normative references .....	2
2.1 International standards .....	2
2.2 Publicly-available specifications .....	2
3 Definitions, notations, and conventions .....	3
3.1 Definitions .....	3
3.1.1 Definitions provided in Part 1 .....	3
3.1.2 Definitions provided in Part 6 .....	3
3.2 Notations .....	7
3.2.1 Notations provided in Part 1 .....	7
3.2.2 Notations provided in Part 6 .....	8
3.3 Conventions .....	8
4 Concepts .....	9
4.1 Data mining techniques .....	9
4.2 Computational patterns .....	9
4.2.1 Training phase .....	9
4.2.2 Application phase .....	10
4.2.3 Testing phase .....	10
4.3 Mapping on user-defined types .....	11
4.3.1 Target application .....	11
4.3.2 The relationships of the user-defined types .....	11
4.3.3 User-defined types that are not related to mining techniques .....	12
4.3.4 User-defined types that are related to mining techniques .....	13
5 Data Mining Model Types .....	15
5.1 DM_RuleModel Type and Routines .....	15
5.1.1 DM_RuleModel Type .....	15
5.1.2 DM_impRuleModel Method .....	16
5.1.3 DM_expRuleModel Method .....	17
5.1.4 DM_getNumRules Method .....	18
5.1.5 DM_getRuleBldTask method .....	19
5.2 DM_ClusteringModel Type and Routines .....	20
5.2.1 DM_ClusteringModel Type .....	20
5.2.2 DM_impClusModel Method .....	22
5.2.3 DM_expClusModel Method .....	23
5.2.4 DM_getNumClusters Method .....	24
5.2.5 DM_applyClusModel Method .....	25
5.2.6 DM_getClusTask method .....	26
5.2.7 DM_getClusMdlSpec method .....	27
5.3 DM_RegressionModel Type and Routines .....	28
5.3.1 DM_RegressionModel Type .....	28
5.3.2 DM_impRegModel Method .....	30
5.3.3 DM_expRegModel Method .....	31
5.3.4 DM_applyRegModel Method .....	32
5.3.5 DM_testRegModel Method .....	33

5.3.6	DM_getRegRSquared method .....	34
5.3.7	DM_getRegTask method .....	35
5.3.8	DM_getRegMdlSpec method .....	36
5.3.9	DM_getRegTarget method .....	37
5.4	DM_ClasModel Type and Routines .....	38
5.4.1	DM_ClasModel Type .....	38
5.4.2	DM_impClasModel Method .....	40
5.4.3	DM_expClasModel Method .....	41
5.4.4	DM_applyClasModel Method .....	42
5.4.5	DM_testClasModel Method .....	43
5.4.6	DM_getClasCostRate method.....	44
5.4.7	DM_getClasTask method .....	45
5.4.8	DM_getClasMdlSpec method.....	46
5.4.9	DM_getClasTarget method .....	47
5.5	DM_ClusAppITask Type and Routines .....	48
5.5.1	DM_ClusAppITask Type .....	48
5.5.2	DM_defClusAppITask Method .....	50
5.5.3	DM_getClusAppIMdl Method .....	51
5.5.4	DM_getClusAppIData Method.....	52
5.5.5	DM_getClusOutpData Method .....	53
5.5.6	DM_getClusResFld method .....	54
5.5.7	DM_applyClusMdl method .....	55
5.6	DM_RegTestTask Type and Routines.....	56
5.6.1	DM_RegTestTask Type .....	56
5.6.2	DM_defRegTestTask Method.....	57
5.6.3	DM_getRegTestData Method .....	58
5.6.4	DM_getRegTestMdl Method.....	59
5.6.5	DM_testRegMdl Method .....	60
5.7	DM_RegAppITask Type and Routines.....	61
5.7.1	DM_RegAppITask Type .....	61
5.7.2	DM_defRegAppITask Method .....	63
5.7.3	DM_getRegAppIMdl Method .....	64
5.7.4	DM_getRegAppIData Method.....	65
5.7.5	DM_getRegOutpData Method .....	66
5.7.6	DM_getRegResFld method .....	67
5.7.7	DM_applyRegMdl method .....	68
5.8	DM_ClasTestTask Type and Routines.....	69
5.8.1	DM_ClasTestTask Type .....	69
5.8.2	DM_defClasTestTask Method .....	70
5.8.3	DM_getClasTestData Method .....	71
5.8.4	DM_getClasTestMdl Method .....	72
5.8.5	DM_testClasMdl Method.....	73
5.9	DM_ClasAppITask Type and Routines .....	74
5.9.1	DM_ClasAppITask Type .....	74
5.9.2	DM_defClasAppITask Method .....	76
5.9.3	DM_getClasAppIMdl Method .....	77
5.9.4	DM_getClasAppIData Method.....	78
5.9.5	DM_getClasOutpData Method .....	79
5.9.6	DM_getClasResFld method .....	80
5.9.7	DM_applyClasMdl method .....	81
6	Data Mining Settings Types .....	82
6.1	DM_ClasSettings Type and Routines .....	82
6.1.1	DM_ClasSettings Type .....	82
6.1.2	DM_impClasSettings Method .....	85

STANDARD PREVIEW  
 (standards.iteh.ai)  
<https://standards.iteh.ai/catalog/standards/sist/9673eece-9fd-4498-8e3c-1dd0aa1b4677/iso-iec-13249-6-2002>

6.1.3	DM_expClasSettings Method .....	86
6.1.4	DM_setClasCostRate Method.....	87
6.1.5	DM_getClasCostRate Method.....	88
6.1.6	DM_setClasCost Method.....	89
6.1.7	DM_getClasCost Method .....	90
6.1.8	DM_useClasDataSpec Method .....	91
6.1.9	DM_getClasDataSpec Method.....	92
6.1.10	DM_setClasTarget Method .....	93
6.1.11	DM_getClasTarget Method.....	94
6.2	DM_ClusSettings Type and Routines.....	95
6.2.1	DM_ClusSettings Type.....	95
6.2.2	DM_impClusSettings Method.....	98
6.2.3	DM_expClusSettings Method .....	99
6.2.4	DM_setMaxNumClus Method .....	100
6.2.5	DM_getMaxNumClus Method .....	101
6.2.6	DM_setFldWeight Method.....	102
6.2.7	DM_getFldWeight Method.....	103
6.2.8	DM_setFldSimScale Method.....	104
6.2.9	DM_getFldSimScale Method.....	105
6.2.10	DM_useClusDataSpec Method .....	106
6.2.11	DM_getClusDataSpec Method .....	107
6.3	DM_RegSettings Type and Routines.....	108
6.3.1	DM_RegSettings Type.....	108
6.3.2	DM_impRegSettings Method.....	110
6.3.3	DM_expRegSettings Method .....	111
6.3.4	DM_setRegRSquared Method.....	112
6.3.5	DM_getRegRSquared Method .....	113
6.3.6	DM_useRegDataSpec Method.....	114
6.3.7	DM_getRegDataSpec Method.....	115
6.3.8	DM_setRegTarget Method.....	116
6.3.9	DM_getRegTarget Method.....	117
6.4	DM_RuleSettings Type and Routines.....	118
6.4.1	DM_RuleSettings Type.....	118
6.4.2	DM_impRuleSettings Method.....	121
6.4.3	DM_expRuleSettings Method .....	122
6.4.4	DM_setMinSupport Method .....	123
6.4.5	DM_getMinSupport Method.....	124
6.4.6	DM_setMinConf Method.....	125
6.4.7	DM_getMinConf Method.....	126
6.4.8	DM_setMaxLength Method .....	127
6.4.9	DM_getMaxLength Method .....	128
6.4.10	DM_useRuleDataSpec Method .....	129
6.4.11	DM_getRuleDataSpec Method .....	130
6.4.12	DM_setGroup Method.....	131
6.4.13	DM_getGroup Method.....	132
6.5	DM_ClasBldTask Type and Routines .....	133
6.5.1	DM_ClasBldTask Type .....	133
6.5.2	DM_defClasBldTask Method .....	135
6.5.3	DM_getClasTrnData Method.....	136
6.5.4	DM_getClasValData Method .....	137
6.5.5	DM_getClasSettings Method .....	138
6.5.6	DM_buildClasModel Method.....	139
6.6	DM_RuleBldTask type and Routines .....	140
6.6.1	DM_RuleBldTask Type.....	140
6.6.2	DM_defRuleBldTask Method .....	141

6.6.3	DM_getRuleTrnData Method	142
6.6.4	DM_getRuleSettings Method	143
6.6.5	DM_buildRuleModel Method	144
6.7	DM_RegBldTask type and Routines	145
6.7.1	DM_RegBldTask Type	145
6.7.2	DM_defRegBldTask Method	147
6.7.3	DM_getRegTrnData Method	148
6.7.4	DM_getRegValData Method	149
6.7.5	DM_getRegSettings Method	150
6.7.6	DM_buildRegModel Method	151
6.8	DM_ClusBldTask type and Routines	152
6.8.1	DM_ClusBldTask Type	152
6.8.2	DM_defClusBldTask Method	153
6.8.3	DM_getClusTrnData Method	154
6.8.4	DM_getClusSettings Method	155
6.8.5	DM_buildClusModel Method	156
7	Data Mining Application Result Types	157
7.1	DM_ClusResult Type and Routines	157
7.1.1	DM_ClusResult Type	157
7.1.2	DM_getClusterID Method	158
7.1.3	DM_getQuality Method	159
7.2	DM_ClasResult Type and Routines	160
7.2.1	DM_ClasResult Type	160
7.2.2	DM_getPredClass Method	161
7.2.3	DM_getConfidence Method	162
7.3	M_RegResult Type and Routines	163
7.3.1	DM_RegResult Type	163
7.3.2	DM_getPredValue Method	164
8	Data Mining Test Result Types	165
8.1	DM_ClasTestResult Type and Routines	165
8.1.1	DM_ClasTestResult Type	165
8.1.2	DM_getClasError Method	166
8.1.3	DM_getNumClasses Method	167
8.1.4	DM_getClass Method	168
8.1.5	DM_getClasPreds Method	169
8.2	DM_RegTestResult Type and Routines	170
8.2.1	DM_RegTestResult Type	170
8.2.2	DM_getPredError Method	171
9	Data Mining Data Types	172
9.1	DM_LogicalDataSpec Type and Routines	172
9.1.1	DM_LogicalDataSpec Type	172
9.1.2	DM_addDataSpecFld Method	174
9.1.3	DM_remDataSpecFld Method	175
9.1.4	DM_getNumFields Method	176
9.1.5	DM_getFldName Method	177
9.1.6	DM_setFldType Method	178
9.1.7	DM_getFldType Method	179
9.1.8	DM_isCompatible Method	180
9.2	DM_MiningData Type and Routines	181
9.2.1	DM_MiningData Type	181
9.2.2	DM_defMiningData Method	182
9.2.3	DM_setFldAlias Method	183

9.2.4	DM_genDataSpec Method.....	184
9.3	DM_ApplicationData Type and Routines .....	185
9.3.1	DM_ApplicationData Type .....	185
9.3.2	DM_genDataSpec Method.....	186
9.3.3	DM_impAppIData Method .....	187
10	Status Codes.....	188
11	Conformance.....	190
11.1	Requirements for conformance .....	190
11.2	Claims of conformance .....	190
Annex A	(informative) Implementation-defined elements .....	191
A.1	Implementation-defined Meta-variables .....	191
Annex B	(informative) Implementation-dependent elements.....	192
B.1	Implementation-dependent elements .....	192
Annex C	(informative) A scenario using the classification technique.....	193

## Figures

Figure 1	— The training phase .....	10
Figure 2	— The application phase.....	10
Figure 3	— The test phase .....	11
Figure 4	— Overview of user-defined types for definitions.....	11
Figure 5	— Overview of user-defined types for training, test and application.....	12

## Tables

Table 1	— Values for field types.....	178
Table 2	— SQLSTATE class and subclass values.....	188

## Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 13249-6 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 32, *Data management and interchange*.

ISO/IEC 13249 consists of the following parts, under the general title *Information technology — Database languages — SQL multimedia and application packages*:

- *Part 1: Framework*
- *Part 2: Full-Text*
- *Part 3: Spatial*
- *Part 5: Still Image*
- *Part 6: Data mining*

IT-**STANDARD-PREVIEW**  
(standards.iteh.ai)  
<https://standards.iteh.ai/catalog/standards/sist/9673eece-9fcd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002>

Annexes A, B and C of this part of ISO/IEC 13249 are for information only.



## Introduction

The purpose of this International Standard is to define multimedia and application specific types and their associated routines using the user-defined features in ISO/IEC 9075.

SQL/MM is structured as a multi-part standard. At present it consists of the following parts:

Part 1: Framework

Part 2: Full-Text

Part 3: Spatial

Part 5: Still Image

Part 6 : Data mining

The organization of this part of ISO/IEC 13249 is as follows:

- 1) Clause 1, "Scope", specifies the scope of this part of ISO/IEC 13249.
  - 2) Clause 2, "Normative references", identifies additional standards that, through reference in this part of ISO/IEC 13249, constitute provisions of this part of ISO/IEC 13249.
  - 3) Clause 3, "Definitions, notations, and conventions", defines the notations and conventions used in this part of ISO/IEC 13249.
  - 4) Clause 4, "Concepts", presents concepts used in the definition of this part of ISO/IEC 13249.
  - 5) Clause 5, "Data Mining Model Types", defines the user-defined types and associated routines for data mining models.
  - 6) Clause 6, "Data Mining Settings Types", defines the user-defined types and associated routines for data mining settings.
  - 7) Clause 7, "Data Mining Application Result Types", defines the user-defined types and associated routines for the results of the application of a data mining model.
  - 8) Clause 8, "Data Mining Test Result Types", defines the user-defined types and associated routines for the results of a test of a data mining model.
  - 9) Clause 9, "Data Mining Data Types", defines the user-defined types and associated routines for the data definitions of data mining.
  - 10) Clause 10, "Status Codes", defines the SQLSTATE codes used in this part of ISO/IEC 13249.
  - 11) Clause 11, "Conformance", defines the criteria for conformance to this part of ISO/IEC 13249.
- A) Annex A, "Implementation-defined elements", is an informative Annex. It lists those features for which the body of this part of ISO/IEC 13249 states that the syntax or meaning or effect on the database is partly or wholly implementation-defined, and describes the defining information that an implementor shall provide in each case.
- B) Annex B, "Implementation-dependent elements", is an informative Annex. It lists those features for which the body of this part of ISO/IEC 13249 states explicitly that the meaning or effect on the database is implementation-dependent.

C) Annex C, "A scenario using the classification technique", is an informative Annex. It presents explanatory material on how to use this part of 13249.

In the text of this part of ISO/IEC 13249, Clauses begin a new odd-numbered page, and in Clause 5, "Data Model Types", through Clause 9, "Data Mining Data Types", subclauses begin a new page. Any resulting blank space is not significant.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC 13249-6:2002](https://standards.iteh.ai/catalog/standards/sist/9673eece-9fcd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002)

<https://standards.iteh.ai/catalog/standards/sist/9673eece-9fcd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002>

# Information technology — Database languages — SQL multimedia and application packages —

## Part 6: Data mining

### 1 Scope

This part of ISO/IEC 13249:

- a) introduces the Data mining part of ISO/IEC 13249,
- b) gives the references necessary for this part of ISO/IEC 13249,
- c) defines notations and conventions specific to this part of ISO/IEC 13249,
- d) defines concepts specific to this part of ISO/IEC 13249,
- e) defines data mining user-defined types and their associated routines.

The data mining user-defined types defined in this part adhere to the following:

- A data mining user-defined type is generic to data mining data handling. It addresses the need to store, manage and retrieve information based on elements such as data mining models, data mining settings, and data mining test results.
- A data mining user-defined type does not redefine the database language SQL directly or in combination with another data mining data type.

## 2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 13249. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this part of ISO/IEC 13249 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

### 2.1 International standards

ISO/IEC 9075-1, *Information technology — Database languages — SQL — Part 1: Framework (SQL/Framework)*

ISO/IEC 9075-2:1999, *Information technology — Database languages — SQL — Part 2: Foundation (SQL/Foundation)*

ISO/IEC 9075-4:1999, *Information technology — Database languages — SQL — Part 4: Persistent Stored Modules (SQL/PSM)*

ISO/IEC 13249-1, *Information technology — Database languages — SQL multimedia and application packages — Part 1: Framework*

### 2.2 Publicly-available specifications

Extensible Markup Language (XML) 1.0 (Second Edition), 2 October 2000,  
<http://www.w3.org/TR/REC-xml>

Predictive Model Markup Language (PMML) 1.1, <http://www.dmg.org/pmml-v1-1.htm>

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

ISO/IEC 13249-6:2002

<https://standards.iteh.ai/catalog/standards/sist/9673eece-9fd1-4498-8e3c-1dd0aa1b4677/iso-iec-13249-6-2002>

### 3 Definitions, notations, and conventions

#### 3.1 Definitions

##### 3.1.1 Definitions provided in Part 1

This part of ISO/IEC 13249 makes use of all terms defined in part 1 of ISO/IEC 13249.

##### 3.1.2 Definitions provided in Part 6

For the purposes of this part of ISO/IEC 13249, the following definitions apply.

###### 3.1.2.1

###### **active field**

An input field that is used during the data mining run to compute a data mining model, evaluate or predict a single input row.

Note: Active fields are defined in the logical data specification.

###### 3.1.2.2

###### **application phase**

Phase, during which an input row is evaluated against a data mining model and one or more values are computed: the identification of the cluster for data mining clustering, the predicted value for data mining classification or data mining regression.

Note: The application phase is only used for data mining clustering, classification and regression now.

###### 3.1.2.3

###### **association rule**

Combination of field values, which appears frequently (according to a specified threshold), i.e. there is a significant number of groups containing the combination of the values.

Example of application: store layout.

###### 3.1.2.4

###### **association rule support**

The percentage of transactions that contain all items of an association rule.

Note: The association rule support of the association rule "X and Y implies Z" is 100 times the number of transactions that contain X, Y and Z divided by the total number of transactions.

###### 3.1.2.5

###### **categorical field type**

A field being of categorical field type supports only the <equal operator> as a comparison operator.

Note: There is no defined order, and no arithmetic operations are supported.

###### 3.1.2.6

###### **class label**

Target field used in data mining classification to create a classification model, whose application to data without a class label allows it to predict a value for this class label.

Note: The class label must have a categorical field type.

###### 3.1.2.7

###### **classification cost rate**

Maximal classification error tolerated during the validation phase of data mining classification, above which the classification model is considered not good enough.

**3.1.2.8**

**classification error**

Percentage of wrongly predicted values among the total number of values predicted during the validation phase of the classification technique.

Note: This error is returned at the end of the test phase.

**3.1.2.9**

**classification model**

The result of data mining classification.

**3.1.2.10**

**data mining application task**

Abstraction for the complete data needed to apply a data mining model, containing values of the data mining model, the data mining data for input and output, and the result field used for the operation.

**3.1.2.11**

**data mining classification**

Data mining technique that computes a classification model such that the class label value can be predicted by using the model and input field values without the class label value.

Note: In this document, the term “classification” is sometimes used as a synonym for “data mining classification” where the meaning is unambiguous from the context.

Example of application: insurance risk prediction.

**3.1.2.12**

**data mining cluster**

A set of input rows with common characteristics (see data mining clustering).

Note: In this document, the term “cluster” is sometimes used as a synonym for “data mining cluster” where the meaning is unambiguous from the context.

iTeh STANDARD PREVIEW

(its standard iTeh.ai)

<https://standards.itih.ai/catalog/standards/sist/9673eece-9fd-4498-8e3c-fdd0aa1b4677/iso-iec-13249-6-2002>

**3.1.2.13**

**data mining clustering**

A data mining technique that discovers sets of input rows with common characteristics - the *clusters*, so that rows are possibly homogeneous inside a cluster and possibly heterogeneous between two clusters.

Note: In this document, the terms “cluster” or “segment” are sometimes used as a synonym for “data mining cluster” where the meaning is unambiguous from the context.

Example of application: customer mailings.

**3.1.2.14**

**data mining data**

Description of data contained in tables that represents the metadata required for access to the data for data mining training, test or application runs.

**3.1.2.15**

**data mining data type**

A data mining data type represents an abstraction of a table. A data mining data type maps the input field names (i.e. column names of the table) to alias names. These alias names will be compared, before the data mining run, to the field names used in the specified logical data specification.

**3.1.2.16**

**data mining model**

The result of a data mining run of a data mining technique over a given set of data, which contains the usable correlations discovered in the data.

Note 1: The data mining model can also be used alone as a substitute of the original large amount of data to associate, classify or predict the behavior of additional data (in application and testing phases) .

Note 2: In this document, the terms “mining model” or “model” are sometimes used as a synonym for “data mining model” where the meaning is unambiguous from the context.

### 3.1.2.17

#### **data mining regression**

Data mining technique similar to data mining classification except for the numerical type of the target field, which computes a regression model allowing to predict a numerical value.

Note 1: The predicted value might not be identical with any value contained in the data used to build the model.

Note 2: In this document, the term “regression” is sometimes used as a synonym for “data mining regression” where the meaning is unambiguous from the context.

Example of application: customer ranking.

### 3.1.2.18

#### **data mining run**

Computation of a data mining model using data mining settings.

Note: According to the data mining technique, a data mining run may contain a training phase, a validation phase and a testing phase. A data mining run is called a training run when it contains a training phase, and a test run when it contains only a test phase.

### 3.1.2.19

#### **data mining segmentation**

see data mining clustering.

### 3.1.2.20

#### **data mining settings**

Settings used in the data mining run that specify the logical data specification and parameters depending on the data mining technique.

Note: In this document, the terms “mining settings” or “settings” are sometimes used as a synonym for “data mining settings” where the meaning is unambiguous from the context.

### 3.1.2.21

#### **data mining task**

Abstraction for the complete data needed to compute, test or apply a data mining model, containing the data mining settings and the data mining data type.

Note: In this document, the terms “mining task” or “task” are sometimes used as a synonym for “data mining task” where the meaning is unambiguous from the context.

### 3.1.2.22

#### **data mining technique**

Four types of data mining techniques are distinguished: discovering association rules, clustering/segmentation, classification and regression.

### 3.1.2.23

#### **data mining test result**

Statistical result returned at the end of the testing phase that contains an evaluation of the correctly and incorrectly predicted values in the test data.

Note 1: This result can be considered a quality criterion of the data mining model.

Note 2: In this document, the term “test result” is sometimes used as a synonym for “data mining test result” where the meaning is unambiguous from the context.