Urania tables and integrating Weka to Java project

Bc. Peter Nosáľ
207773@mail.muni.cz
Faculty of Informatics
Masaryk University
April 2012

Content

- Urania tables
- Importing Weka into Java project
- Obtaining data from DB using Weka
- Using data in Java project or Weka

Urania tables

- Odpovedniky
- Odpovedi
- Otazky

- Sql for creation of these tables is found here:
  - *http://urania.fi.muni.cz/dev/db.sql*
- Project that is working with these tables
  - *http://urania.fi.muni.cz/documentation/importerProject.zip*

Odpovedniky

nameOfVariable	typeOfVariable
id	INT(11)
file_id	VARCHAR(32)
file_name	VARCHAR(50)
dr_name	VARCHAR(32)
passwrd	VARCHAR(8)
uco	VARCHAR(16)
datum	DATETIME
status	TINYINT(4)

Odpovedniky

select * from odpovedniky ×

Page Size: 20 | Total Rows: 186 Page: 1 of 10 | Matching Rows:

#	id	file_id	file_name	dir_name	passwd	uco	datum	status
1	1	2c0e5b7d690c2bb2d279df3e0eb6301b	Jak_funguje_e_learning_v_IS_MU.xml	FjkgxQBoIyiCVeTdWs1hENHq54b2MY87	C1DTlJ9E	2660	2008-02-12 12:00:08.0	0
2	2	cdbadfe70a803c5d3eae78dd1ae11590	Vyrokova_logika.xml	X0RUhbvHN95LmTzDtlEMYgkZPaOrF1qA	2Sb7mzUa	72543	2008-02-12 11:54:03.0	0
3	3	5ea141ab4d39888857ed0fc761c80887	novy.xml	XRFyt0CzZTKiOSVbHpPmGDBLUWcgjYx5	N7CFvz46	2Sb7mz Ua	2008-02-12 11:14:26.0	0
4	4	Iha	Bez_id.xml	nJRhFWQGdkDmOpxBC8wVy5lP2cKSsUv0	ha		2008-02-14 08:03:54.0	0
5	5	412fd44e8d1db66ea9dc015b6aaf3467	novy.xml	KtAR7sfhcE0U4m6xCQdPLlyYSoNIaB8D	N7CFvz46	143277	2008-02-15 11:19:12.0	0
6	6	244b6debf4d4f225e8238190f44bc02f	test_logika.xml	BAfyarLjZHiSclp2FRM53XOksowK9t7Y	d90WZPiR	1945	2008-02-16 17:14:17.0	0
7	7	all	Vyrokova_all.xml	TugN03D8mZbj65vhiFVXx4qrpwa1QCsf	all	72543	2008-02-17 15:35:56.0	0
8	8	4ec73df34d0048b77a126863193e90fc	Vyrokova_logika.xml	WBbje6dl4w5T71R0IMmKYLpvQGxVNg3c	logika	1945	2008-04-24 12:36:08.0	0
9	9	bf24224b1a0766b77bee52d4e0aee940	porovnavani_odp.xml	jChGFV5nXoP4Qd8Nfy30Kvkxrw2eYm9D	9wICMpjk	72543	2008-02-20 16:58:52.0	0
10	10	3ab83a2f7445c3c35a6f24935e58969f	porovnavani_odp.xml	w5dOrHvaIkoXR3Llbj2npWiVc9m0C7Sh	9wICMpjk	72543	2008-02-21 10:55:50.0	1
11	11	f55d54bc32df3db0cc80539b37074d85	porovnavani_odp.xml	lcQn3V4YL1SqR0gwHZPkF9jTUmhGBW8X	9wICMpjk	72543	2008-02-27 13:12:13.0	0
12	12	a893f21c0d6ad70fd05943c33d0cbec0	porovnavani_odp.xml	TpCVAs5SIMONxKc8lFP9ftDaLE3UZ02n	9wICMpjk	72543	2008-02-29 09:02:30.0	1

Otazky

nameOfVariable	typeOfVariable
otazkaId	INT UNSIGNED
hashOtazky	VARCHAR(32)
textOtazky	TEXT
bodyZaSpravnouOdpoved	DOUBLE
bodyZaChybnouOdpoved	DOUBLE
bodyZaNezodpovezeni	DOUBLE
odpovedniky_id	INT(11)

Otazky

select * from otazky ×

Page Size: 20 | Total Rows: 660 Page: 1 of 33 | Matching Rows:

#	otazkaId	hashOtazky	textOtazky	bodyZaSpravnouOdpoved	bodyZaChybnouOdpoved	bodyZaNezodpovezeni	odpovedniky_id
1	361	10ea5744cbb83f214e3c36534aa923e2	Uvažte gramatiku definitních klauzulí<pre>...	2.0	-1.0	0.0	1081
2	362	154a71d489e18749b0bb591c82c84a24	Máme následující DC gramatiku<pre>a --&g...	2.0	-1.0	0.0	1081
3	363	1a804763d21865962aefe61bf85fc834	Která z následujících klauzulí je generalizací ...	2.0	-1.0	0.0	1081
4	364	217a87a2f02163539b93fc3848998cb6	Pokud <m>B</m> je doménová znalost, <...	2.0	-1.0	0.0	1081
5	365	225a48f489d5785d4b9d582ef7228487	Prostor verzí pro data bez šumu ...	2.0	-1.0	0.0	1081
6	366	2522b0a43c8ba9762f6427f5759ed785	Nechť val(P) je pravdivostní hodnota výrok...	2.0	-1.0	0.0	1081
7	367	2cf6642c6409fc841b5b6fc259b835b3	Která z následujících klauzulí není specializac...	2.0	-1.0	0.0	1081
8	368	2d6945061ba9716ac506626c10bbf9a2	(General) resolution is for both propositiona...	2.0	-1.0	0.0	1081
9	369	30413a63d4002711d1675876cae05801	Specializační operátor <m>\rho</m> je <e...	2.0	-1.0	0.0	1081
10	370	31283dfea39544e7e76939b9a4a3df33	Máme následující DC gramatiku pro jazyk <...	2.0	-1.0	0.0	1081
11	371	3920426e6bd7ed08774dd981ac0d3f0b	The Prolog program a(b,c). a(X,Y) :- a(...	2.0	-1.0	0.0	1081
12	372	3e774a996eac35aaddabcbe7632f1fe8	Algoritmus pro výpočet prostoru verzí ...	2.0	-1.0	0.0	1081
13	373	40b2abee958c9626c75340ed708b7c46	V deskripční logice <m>\mathcal{ALC}</m...	2.0	-1.0	0.0	1081

Odpovedi

nameOfVariable	typeOfVariable
answerId	INT UNSIGNED
studentId	INT
pruchod	MEDIUMINT
konecneUlozeni	CHAR
ulozeni	MEDIUMINT
odpoved	TEXT
hodnoceni	VARCHAR(255)
body	DOUBLE
celkovaSpravnost	VARCHAR(4)
otazky_otazkaId	INT UNSIGNED

Odpovedi

select * from odpovedi ×

Page Size: 20 Total Rows: 5082 Page: 1 of 255 Matching Rows:

#	answerId	studentId	pruchod	konecneUlozeni	ulozeni	odpoved	hodnoceni	body	celkovaSpravnost	otazky_otazkaId
1	2773	272490	1	T	1	:r3a	ok	2.0	ok	361
2	2774	139764	1	T	1	:r1a	nok	-1.0	nok	361
3	2775	173360	1	T	1	:NULL	null	0.0	null	361
4	2776	143390	1	T	1	:NULL	null	0.0	null	361
5	2777	207724	1	T	1	:NULL	null	0.0	null	361
6	2778	207730	1	T	1	:r3a	ok	2.0	ok	361
7	2779	173174	1	T	1	:r3a	ok	2.0	ok	361
8	2780	255880	1	T	1	:r1a	nok	-1.0	nok	361
9	2781	172734	1	T	1	:r3a	ok	2.0	ok	361
10	2782	173340	1	T	1	:r1a	nok	-1.0	nok	361
11	2783	98854	1	T	1	:r1a	nok	-1.0	nok	361
12	2784	207519	1	T	1	:NULL	null	0.0	null	362
13	2785	172885	1	T	1	:NULL	null	0.0	null	362

Importing Weka into Java project

1. Download and install the newest Weka from

   http://sourceforge.net/projects/weka/files/weka-3-7/3.7.5/

2. Use weka.jar or weka-src.jar
3. Modify DatabaseUtils.prop
   - DatabaseUtils.prop is located in folder .\weka\experiment
   - Need to change driver and database name
   - Add types of attributes used in tables
   - More info: *http://weka.wikispaces.com/Databases*

Importing Weka into Java project

4. Add into system Enviroment Variables in Classpath path to your database driver
    - *http://weka.wikispaces.com/CLASSPATH*
5. Import weka.jar or weka-src.jar into your project as dependency

Obtaining data from DB using Weka

- Imports for creating Instances and ARFF file
  - import weka.core.Instances
  - import weka.experiment.InstanceQuery
- ```
  InstanceQuery query = new InstanceQuery();
  query.setUsername(String userName);
  query.setPassword(String userPassword);
  query.setQuery("select * from table");
  Instances data = query.retrieveInstances();
  ```
- *http://weka.wikispaces.com/Use+WEKA+in+your+Java+code*
  ```

Obtaining data from DB using Weka - 2

- Once you have data
  - Use data in program
  - Create ARFF file from data
- Creating ARFF file

```
BufferedWriter writer = new BufferedWriter(new
FileWriter("name.arff"));

writer.write(data.toString());

writer.flush();

writer.close();
```

- Don't use unique attributes

  http://weka.wikispaces.com/Remove+Attributes

Using data in Java project or Weka

- In Java project
  - import weka.gui.treevisualizer.PlaceNode2;
  - import weka.gui.treevisualizer.TreeVisualizer;
  - import weka.classifiers.trees.J48;
  - Create instance of J48 tree
  - Build classifier with selected attribute
  - Create a graph using TreeVisualizer
  - Display in Jframe
- More info:
  http://weka.wikispaces.com/Visualizing+a+Tree

Using data in Java project or Weka

- In Weka
  - Open created ARFF file in Weka
  - Do whatever you want :-)

Any questions?

Thank you for your attention