

システム仕様書

本システムを使って，規則から決定木作成までに行う手順を「授業料免除制度の申請判定木」を例に説明する．プログラムの実行する際には，準備するものが2つあり，ひとつは文章で書かれた規則から抽出した抽象事例，もうひとつは，属性と属性値とクラスの関係を表した関係図である．

抽象事例の設定方法

抽象事例を記述するためには，文章で書かれた規則から特徴を抽出しなければならない．授業料免除制度の例の場合，表??のような業務の運用者が知識や経験から考えられる規則の特徴が考えられる．この特徴から属性を考え，属性と属性値とクラスを定義したものが表??である．

表 1 文章で書かれた規則の特徴

	申請可となる条件
1	新入生である，かつ，支払い困難な理由が1年以内
2	4回生以下，かつ，支払い困難な理由が半年以内
	申請不可となる条件
1	申請学期に休学する
2	免除される時，5回生以上
3	支払い困難な理由がない
4	申請時の学年が留年した年である

表 2 属性と属性値とクラスの関係

属性 1	休学 = [はい , いいえ]
属性 2	回生 = [4 以下 , 5 以上]
属性 3	理由 = [半年以内 , 1 年以内 , 数年 , なし]
属性 4	留年 = [はい , いいえ]
属性 5	新入生 = [はい , いいえ]
クラス	合否 = [申請可 , 申請不可]

表 1 の規則を表 2 を使い , 簡易表現する . それを表 3 のように表す .

表 3 抽象事例

休学	回生	理由	留年	新入生	合否
*	*	1 年以内	*	はい	申請可
いいえ	4 以内	半年以内	いいえ	*	申請可
はい	*	*	*	*	申請不可
*	5 以上	*	*	いいえ	申請不可
*	*	なし	*	*	申請不可
*	*	*	はい	いいえ	申請不可

この表 3 をテキストデータ (ここでは Rule.txt とする) として用意する . その際の注意点として , 「各属性値の間は半角スペースにすること . 」 「属性の名前である (休学 , 回生 , … , 合否) は , テキストデータで表す場合 , 省略して書くこと . 」 などがある .

表 4 属性と属性値のデータベース化 (attr)

attribute	value
test0	yes
test0	no
test1	Under-4
test1	Over-5
test2	Under-Half
test2	Under-Year
test2	Over-Year
test2	Not-happen
test3	yes
test3	no
test4	yes
test4	no
test5	yes
test5	no

次に用意するものとして，属性と属性値とクラスの関係を mysql を使用し，データベースに与える．表 2 のような属性と属性値とクラスの関係をデータベースに入力する場合，

- 1．データベース名は任意とすること．
 - 2．テーブル名は「attr」にすること．
 - 3．attr テーブルのフィールド名は「attribute」と「value」にすること．
- などの注意点がある．

表 4 のようなデータベース (ここでは db1 とする) に属性と属性値の関係を表したものを db1 内のテーブルに attr として格納する．

以上のように，表 3 の抽象事例のテキストと属性間の関係データベースの 2 つを使用すると，自動的に決定木を作成するプログラムを実行できる ..

実行例

表 3 の抽象事例と表 4 の属性値の関係を使って実際に実行する．今回の例の場合，「java BTree db1 Rule.txt」を入力することになる．

表 5 実行コマンド

≫ java BTree [データベース名] [テキストファイル名]

実行例

≫ java BTree db1 Rule.txt

DB より事例集合を読み込みます。

Arr-0:no Under-4 Under-Half no yes yes
Arr-1:no Under-4 Under-Half no no yes
Arr-2:yes Under-4 Under-Year yes yes yes
Arr-3:yes Under-4 Under-Year no yes yes
Arr-4:yes Over-5 Under-Year yes yes yes
Arr-5:yes Over-5 Under-Year no yes yes
Arr-6:no Under-4 Under-Year yes yes yes
Arr-7:no Under-4 Under-Year no yes yes
Arr-8:no Over-5 Under-Year yes yes yes
Arr-9:no Over-5 Under-Year no yes yes
Arr-10:yes Over-5 Under-Half yes no no
Arr-11:yes Over-5 Under-Half no no no
Arr-12:yes Over-5 Under-Year yes no no
Arr-13:yes Over-5 Under-Year no no no
Arr-14:yes Over-5 Over-Year yes no no
Arr-15:yes Over-5 Over-Year no no no
Arr-16:yes Over-5 Not-happen yes no no
Arr-17:yes Over-5 Not-happen no no no
Arr-18:no Over-5 Under-Half yes no no
Arr-19:no Over-5 Under-Half no no no
Arr-20:no Over-5 Under-Year yes no no
Arr-21:no Over-5 Under-Year no no no
Arr-22:no Over-5 Over-Year yes no no
Arr-23:no Over-5 Over-Year no no no
Arr-24:no Over-5 Not-happen yes no no
Arr-25:no Over-5 Not-happen no no no

Arr-26:yes Under-4 Under-Half yes no no
 Arr-27:yes Under-4 Under-Year yes no no
 Arr-28:yes Under-4 Over-Year yes no no
 Arr-29:yes Under-4 Not-happen yes no no
 Arr-30:yes Over-5 Under-Half yes no no
 Arr-31:yes Over-5 Under-Year yes no no
 Arr-32:yes Over-5 Over-Year yes no no
 Arr-33:yes Over-5 Not-happen yes no no
 Arr-34:no Under-4 Under-Half yes no no
 Arr-35:no Under-4 Under-Year yes no no
 Arr-36:no Under-4 Over-Year yes no no
 Arr-37:no Under-4 Not-happen yes no no
 Arr-38:no Over-5 Under-Half yes no no
 Arr-39:no Over-5 Under-Year yes no no
 Arr-40:no Over-5 Over-Year yes no no
 Arr-41:no Over-5 Not-happen yes no no
 Arr-42:yes Under-4 Under-Half yes yes no
 Arr-43:yes Under-4 Under-Half yes no no
 Arr-44:yes Under-4 Under-Half no yes no
 Arr-45:yes Under-4 Under-Half no no no
 Arr-46:yes Under-4 Under-Year yes yes no
 Arr-47:yes Under-4 Under-Year yes no no
 Arr-48:yes Under-4 Under-Year no yes no
 Arr-49:yes Under-4 Under-Year no no no
 Arr-50:yes Under-4 Over-Year yes yes no
 Arr-51:yes Under-4 Over-Year yes no no
 Arr-52:yes Under-4 Over-Year no yes no
 Arr-53:yes Under-4 Over-Year no no no
 Arr-54:yes Under-4 Not-happen yes yes no
 Arr-55:yes Under-4 Not-happen yes no no
 Arr-56:yes Under-4 Not-happen no yes no
 Arr-57:yes Under-4 Not-happen no no no
 Arr-58:yes Over-5 Under-Half yes yes no
 Arr-59:yes Over-5 Under-Half yes no no
 Arr-60:yes Over-5 Under-Half no yes no
 Arr-61:yes Over-5 Under-Half no no no
 Arr-62:yes Over-5 Under-Year yes yes no
 Arr-63:yes Over-5 Under-Year yes no no

Arr-64:yes Over-5 Under-Year no yes no
 Arr-65:yes Over-5 Under-Year no no no
 Arr-66:yes Over-5 Over-Year yes yes no
 Arr-67:yes Over-5 Over-Year yes no no
 Arr-68:yes Over-5 Over-Year no yes no
 Arr-69:yes Over-5 Over-Year no no no
 Arr-70:yes Over-5 Not-happen yes yes no
 Arr-71:yes Over-5 Not-happen yes no no
 Arr-72:yes Over-5 Not-happen no yes no
 Arr-73:yes Over-5 Not-happen no no no
 Arr-74:yes Under-4 Not-happen yes yes no
 Arr-75:yes Under-4 Not-happen yes no no
 Arr-76:yes Under-4 Not-happen no yes no
 Arr-77:yes Under-4 Not-happen no no no
 Arr-78:yes Over-5 Not-happen yes yes no
 Arr-79:yes Over-5 Not-happen yes no no
 Arr-80:yes Over-5 Not-happen no yes no
 Arr-81:yes Over-5 Not-happen no no no
 Arr-82:no Under-4 Not-happen yes yes no
 Arr-83:no Under-4 Not-happen yes no no
 Arr-84:no Under-4 Not-happen no yes no
 Arr-85:no Under-4 Not-happen no no no
 Arr-86:no Over-5 Not-happen yes yes no
 Arr-87:no Over-5 Not-happen yes no no
 Arr-88:no Over-5 Not-happen no yes no
 Arr-89:no Over-5 Not-happen no no no

(0) の種類と数: yes(60) no(30)

(1) の種類と数: Under-4(38) Over-5(52)

(2) の種類と数: Under-Half(18) Under-Year(24) Over-Year(16) Not-happen(32)

(3) の種類と数: yes(52) no(38)

(4) の種類と数: yes(33) no(57)

(5) の種類と数: yes(10) no(80)

事例集合より木を作成します。

(((中略)))

決定木を表示します。
(Decition Tree!)

```
Root-(test4:node:90)
  L yes(test2:node:33)
    L Under-Half(test0:node:5)
      L yes(no:4)
      L no(yes:1)
    L Under-Year(test0:node:12)
      L yes(no:8)
      L no(yes:4)
    L Over-Year(no:4)
    L Not-happen(no:12)
  L no(test0:node:57)
    L yes(no:36)
    L no(test1:node:21)
      L Under-4(test3:node:7)
        L yes(no:5)
        L no(test2:node:2)
          L Under-Half(yes:1)
          L Not-happen(no:1)
      L Over-5(no:14)
```