

知識工学課題

岡山大学大学院

竹内孔一

課題について

- 提出期限
 - 5月19日(金曜日)講義の時間中まで
- 提出形式
 - pdfかword
 - 無理な場合は竹内まで連絡すること
 - 名前, 学籍番号を忘れないこと
 - WebClassで提出する
- 注意
 - わからなかったら質問すること

課題1

- 決定木学習モデルC5.0 (お試し版)を download し, windows または linux にインストールして
付属のデモファイルについて解説をする. さらに動かした結果を付与し説明を加えよ. いくつかのデモが付与されているが1つで良い.

C5.0のサイト

<http://www.rulequest.com/download.html>

課題2

- 売れているスマートフォン(simフリー)の人気機種
の属性を調べる
 - サイト価格.comの基本スペックを属性とする
 - 人気順にならべて上位は売れている
(例:1~3位まで)
 - 下位は売れていない(例:38~40位)
 - 属性は, OS種類, メモリ量, バッテリー容量, カメラ画素数など重要そうなものを選択(3つ以上)
 - 正例(売れている)3事例以上, 負例(売れていない)3事例以上を作成して, 決定木に入れる

課題2 説明(例)

- 例えば 下記で人気の順にsortした結果を表示してみよう
 - http://kakaku.com/keitai/smartphone/ranking_3147/
 - 1位から3位, 38位から40位の商品をクリック
 - 右上の「閲覧履歴」=>比較したい商品を選択=>「詳しく比較」ボタンをクリック => 属性の比較が一覧でできる
 - 一覧を見て下記のような決定木用のデータを作成する
 - (属性は適切なものを選ぶ)

機種	キャリア	OSの種類	通信速	...	売れた?
AQUOS	au	Android	150M	...	正
Xperia	docomo	Android	150M	...	正
.....					

項目が無い場合は nullとすれば良い

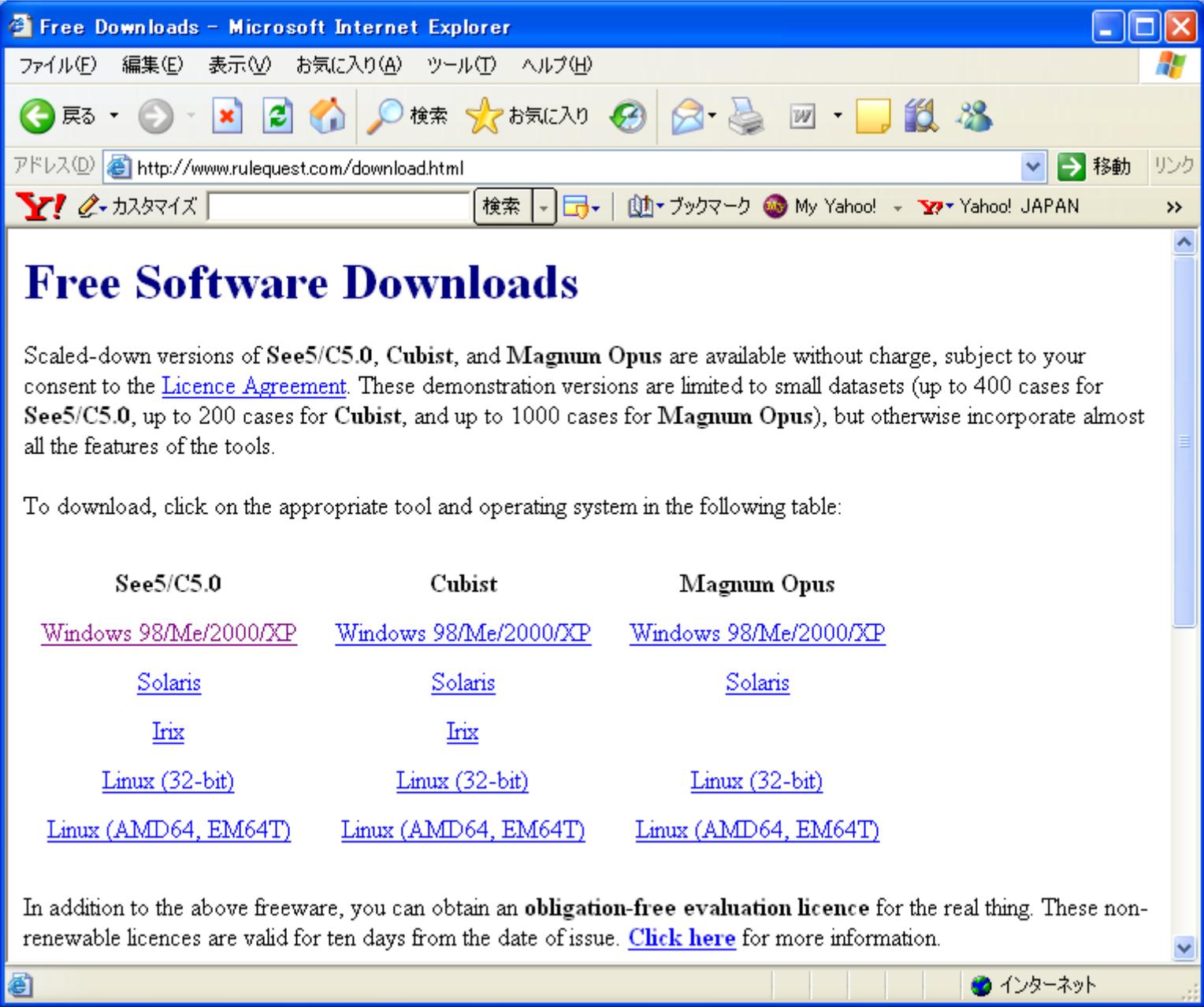
決定木は数字も属性値として使える

属性を選ぶ前にどこで差が出ててそうかよく見る

補足

- ソフトウェアについて機能をよく調べてるレポートを高く評価する.
- かならずソフトウェアを動かすこと. 分析をしない場合は評価しない
 - 分析: 得られた決定木が正しいかどうか, またどの程度信用できそうかなど説明すること. よりよいやり方などあれば考えること

補足: download するのは左側の See/C5.0



Free Downloads - Microsoft Internet Explorer

ファイル(E) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り

アドレス(D) <http://www.rulequest.com/download.html> 移動 リンク

Y! カスタマイズ 検索 ブックマーク My Yahoo! Yahoo! JAPAN

Free Software Downloads

Scaled-down versions of **See5/C5.0**, **Cubist**, and **Magnum Opus** are available without charge, subject to your consent to the [Licence Agreement](#). These demonstration versions are limited to small datasets (up to 400 cases for **See5/C5.0**, up to 200 cases for **Cubist**, and up to 1000 cases for **Magnum Opus**), but otherwise incorporate almost all the features of the tools.

To download, click on the appropriate tool and operating system in the following table:

See5/C5.0	Cubist	Magnum Opus
Windows 98/Me/2000/XP	Windows 98/Me/2000/XP	Windows 98/Me/2000/XP
Solaris	Solaris	Solaris
Irix	Irix	
Linux (32-bit)	Linux (32-bit)	Linux (32-bit)
Linux (AMD64, EM64T)	Linux (AMD64, EM64T)	Linux (AMD64, EM64T)

In addition to the above freeware, you can obtain an **obligation-free evaluation licence** for the real thing. These non-renewable licences are valid for ten days from the date of issue. [Click here](#) for more information.

インターネット

補足2

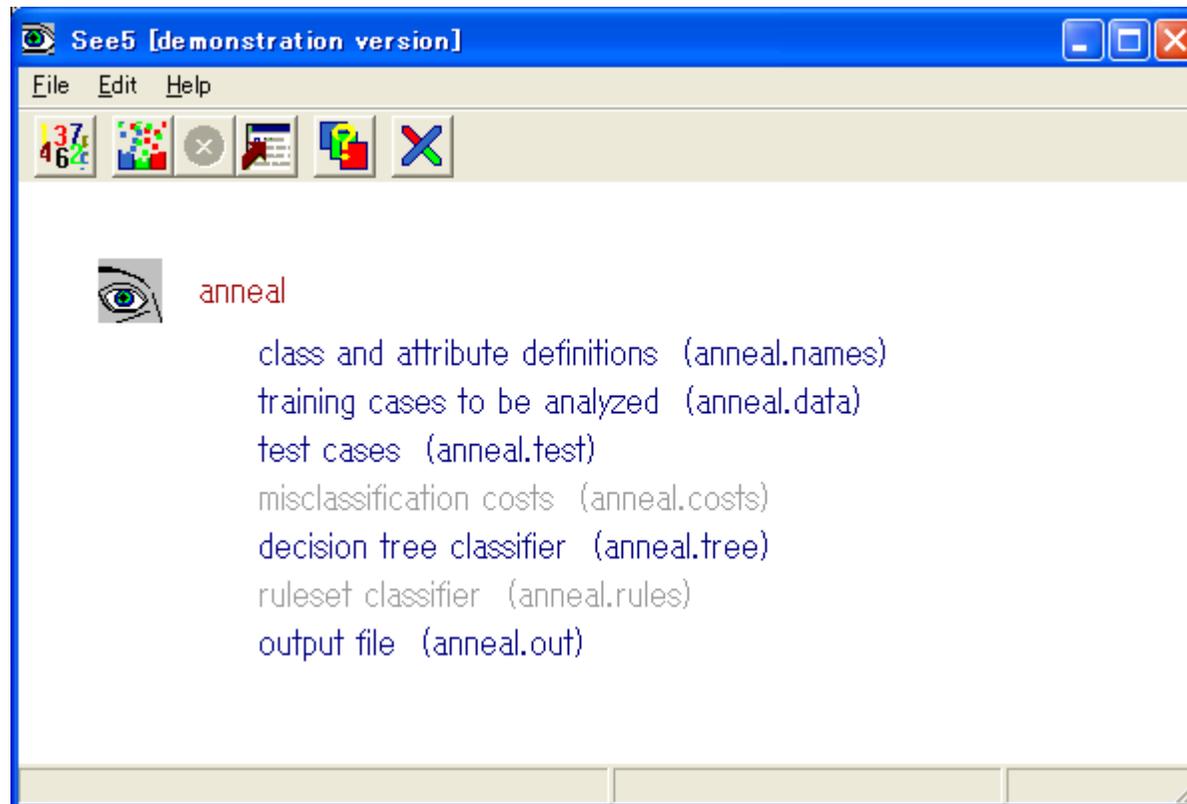
- windowsに入れる場合

download して展開する際, どこにinstallするか
決める C:\program files がdefault

c:\program files\See5-demo\See5-demo
を起動する

これが
起動画面

ヘルプも
参考に
なります



- 操作に関する補足

- 左上のfile のlocate data でデモファイルを読み込む

- デモファイルは \Samplesの下にある

- 例えば anneal.data を選んで

- File->Construct classifier を実行する

- すると結果のファイルが anneal.out,
anneal.tree として \Samples の中にできる

- 結果のファイルがいったい何を示しているのかを解釈するのがポイント

- ・annealとは鉄鋼を硬くするための「焼きなまし」のこと

補足

- 自分でデータを作るとき
 - anneal.data, anneal.names, anneal.setなどのファイル形式を参考に解析したいデータを作成する
 - どういうデータ形式にすればよいかを付属のサンプルデータから考えて推察するところがポイント
 - annealのデータをいろいろ書き換えてどう出力されるか探ってみると良い

Windows tips

- windows の場合以下の点に注意
 - Program Files 下が見れない。また *.dataなどの名前をうまく作成できない。
 - (対策) エクスプローラの[ツール]->[フォルダオプション]->[表示]タブで「すべてのファイルとフォルダを表示する」さらに「登録されている拡張子は表示しない」のを解除する
 - aneal.data など *.txtでないものは編集することができない。見れないなど。
 - (対策) [スタート]->[すべてのプログラム]->[アクセサリ]->[メモ帳] をつかって読み込む。読み込みの際は *.txt 以外のどんなファイルでもみれるようにする

Linuxにインストール

Windowsで入らなかったひとはこちらを試して下さい.



Free Software Downloads

Scaled-down versions of **See5/C5.0** and **Cubist** are available without charge, subject to [Agreement](#). These demonstration versions are limited to small datasets (up to 400 cases for **Cubist** but otherwise incorporate almost all the features of the tools.

To download, click on the appropriate tool and operating system in the following table:

これ

See5/C5.0

Cubist

[Windows 2000/Xp/Vista/7](#)

[Windows 2000/Xp/Vista/7](#)

[Linux \(32-bit\)](#)

[Linux \(32-bit\)](#)

In addition to the above freeware, organizations can obtain an **obligation-free evaluation licence**. These non-renewable licences are valid for ten days from the date of issue. [Click here](#)

RuleQuest also provides free **C source code** to read and interpret the classifiers and **See5/C5.0** and **Cubist**. Once you have run **See5/C5.0** or **Cubist** on your data, you can then incor

展開して終わり

(1) 展開する

```
tar xvfz C50Demo_linux.tgz
```

(2) プログラム

各プログラムは C50Release2/bin の下
c5.0demo というバイナリファイルが動かすプログラム
bin/c5.0demo -h でヘルプがみれる

(3) 使う

サンプルファイルは Data/の下にあるので
c5.0demo -f ../Data/anneal
とすると決定木の結果が stdout に出力される