

```

1
2
3
4
5
6
7
8
9
10
11 ; Semi-supervised multi-objective explanation facility.
12 (defpackage :tiny (:use :cl) (:nicknames "tm"))
13 (in-package :tiny)
14 (mapc #'load '("lib/macros" "lib/matha" "lib/strings"
15               "lib/settings" "lib/structs" "lib/demos" ))
16
17 (defvar my
18   (settings "TOYIN: do stuff
19             (c)2022 Tim Menzies BSD-2 clause license "
20             '(("file" "h" "help file" " ./Data/auto93.lisp")
21               (help "h" "show help" " nil)
22               (keep "k" "items to keep" " 256)
23               (k "k" "ab low attributes classes" " 1)
24               (m "m" "ab low frequency classes" " 2)
25               (seed "s" "random number seed" " 10019)
26               (go "g" "start up action" " ls))))))
27
28 (mapc #'load '("sample" "row" "sym" "num" "about" "data"))
29
30
31 ; Keep up to "max" numbers (after which, replace any old with new).
32 (defstruct+ sample
33   (_kept ; where to keep
34     (make-array 2 :fill-pointer 0 :adjustable t))
35   max ; how many to keep
36   ok) ; nil if items added and list not resorted yet
37
38 (defun make-sample (&optional (max (? my keep-))) (%make-sample :max max))
39
40 (defmethod add ((i sample) (x number))
41   (incf (? i n))
42   (let (size (length (? i _kept)))
43     (cond
44       (setf (? i ok) nil)
45       (vector-push-extend x (? i _kept)))
46     ((< (randf) (/ (? i n) (? i max)))
47       (setf (? i ok) nil)
48       (setf (elt (? i _kept) (randi size)) x))))
49
50 (defmethod sorted ((i sample))
51   (unless (? i ok)
52     (sort (? i _kept) #'<))
53     (setf (? i ok) t))
54   (? i _kept))
55
56
57
58
59 ; Hold one record.
60 (defstruct+ row cells ; cells
61   _about ; pointer to someone who can say what are (e.g.) lo,hi
62
63 (defun make-row (about l) (%make-row :cells l :_about about))
64
65
66
67
68 ; Summarize symbolic columns
69 (defstruct+ sym (txt "") ; column name
70   (at 0) ; column position
71   (n 0) ; #items seen
72   kept) ; symbol counts of the items
73
74 (defun make-sym (&optional s n) (%make-sym :txt s :at n))
75
76 (defmethod add ((i sym) (lst cons)) (dolist (x lst i) (add i x)))
77 (defmethod add ((i sym) x)
78   (unless (eq x #\?)
79     (incf (? i n))
80     (incf (getf x (? i kept)))))
81
82 (defmethod adds ((i sym) x inc)
83   (incf (? i n) inc)
84   (incf (getf x (? i kept)) inc))
85
86 (defmethod div ((i sym))
87   (labels ((fun (p) (* -1 (* p (log p 2)))))
88     (loop for _ . n in (? i kept) sum (fun (/ n (? i n)))))
89
90 (defmethod mid ((i sym))
91   (loop for (key . n) in (? i kept) maximizing n return key))
92
93
94
95
96 ; Summarize numeric columns.
97 (defstruct+ num (txt "") ; column name
98   (at 0) ; column position
99   (n 0) ; #items seen
100   (w 1) ; (1,-1) = (maximize, minimize)
101   kept) (make-some))) ; items seen
102
103 (defun make-num (s n) (%make-num :txt s :at n :w (if (eq #\~ (charn s)) -1 1)))
104
105 (defmethod add ((i num) (lst cons)) (dolist (x lst i) (add i x)))
106 (defmethod add ((i num) x)
107   (unless (eq x #\?)
108     (incf (? i n))
109     (add (? i kept) x)))
110
111
112
113
114 ; Factory for making nums or syms.
115 (defstruct+ about names ; list of column names
116   all ; all the generated columns
117   x ; just the independent columns
118   y ; just the dependent columns
119   klass) ; just the klass col (if it exists)

```

```

120
121 (defun make-about (lst)
122   (let (all x y kl (at -1))
123     (dolist (str lst (%make-about :names lst :x x :y y :klass kl
124                                   :all (reverse all))))
125     (incf at)
126     (let (col (if (eq #\$ (char str 0)) (make-num str at) (make-sym str at)))
127       (push col all)
128       (unless (eq #\~ (charn str))
129         (if (member (charn str) '(! ! #\~ #\+)) (push col y) (push col x))
130         (if (eq #\! (charn str)) (setf kl col))))))
131
132 (defmethod add ((i about) (lst cons)) (add i (make-row i lst)))
133 (defmethod add ((i about) (r row))
134   (dolist (cols '(! (? i x) (? i y) r)
135     (dolist (col cols)
136       (add col (elt (? r cells) (? col at))))))
137
138
139
140
141 ; Place to hold rows, and their summaries.
142 (defstruct+ data rows ; all the rows
143   about ; summaries of all the columns
144
145 (defun make-data (names &optional src (i (%make-data :about (make-about names)))
146   (if (stringp src)
147     (with-lines src (lambda (line) (add i (cells line))))
148     (dolist (row src) (add i row))))
149   i)
150
151 (defmethod clone ((d data) &optional src) (make-data (? d about names) src))
152
153
154
155
156 ; Simple alist access
157 (defmacro | (l x) `(cdr (assoc 'x ,l)))
158
159 ; ? obj x y z == (slot-value (slot-value (slot-value obj 'x) 'y) 'z)
160 (defmacro ? (s x &rest xs)
161   (if (null xs) `(slot-value ,s 'x) `(? (slot-value ,s 'x),@xs)))
162
163 ; Endure lst has a slot for 'x'. If missing, initialize it with 'init'.
164 (defmacro geta (x lst &optional (init 0))
165   `(cdr (or (assoc ,x ,lst :test #'equal)
166             (car (setf ,lst (cons (cons ,x ,init) ,lst)))))
167
168
169
170
171 ; round
172 (defun rnd (number &optional (digits 3))
173   (let* ((div (expt 10 digits))
174         (tmp (/ (round (* number div)) div)))
175     (if (zerop digits) (floor tmp) (float tmp)))
176
177 ; Random number control (since reseeding in LISP is... strange).
178 (defvar *seed* 10013)
179
180 (defun randf (&optional (n 1.0))
181   (setf *seed* (mod (* 16807.0d0 *seed*) 2147483647.0d0))
182   (* n (- 1.0d0 (/ *seed* 2147483647.0d0))))
183
184 (defun randi (&optional (n 1)) (floor (* n (/ (randf 1000000000.0) 1000000000))))
185
186
187
188
189 ; Last thing from a string
190 (defun charn (x) (char x (1- (length x))))
191
192 ; Kill leading trailing whitespace.
193 (defun trim (x) (string-trim '(#\Space #\Tab #\Newline) x))
194
195 ; Turn 'x' into a number or string or "?"
196 (defun thing (x &aux (y (trim x)))
197   (if (string= y "??") #?
198       (let ((z (ignore-errors (read-from-string y))))
199         (if (numberp z) z y))))
200
201 ; Divide 'str' on 'char', filtering all items through 'filter'.
202 (defun splits (str &key (char #\,) (filter #'identity))
203   (loop for start = 0 then (1+ finish)
204         for finish = (position char str :start start)
205         collecting (funcall filter (trim (subseq str start finish)))
206         until (null finish))
207
208 ; String to lines or cells of things
209 (defun lines (string) (splits string :char #\Newline))
210 (defun cells (string) (splits string :filter #'thing))
211
212 ; Call 'fun' for each line in 'file'.
213 (defun with-lines (file fun)
214   (with-open-file (s file)
215     (loop (funcall fun (or (read-line s) nil) (return))))))
216
217
218
219
220 ; Update 'default' from command line. Boolean flags just flip defaults.
221 (defun setting (key flag help default) (key flag help default
222   (destructuring-bind (key flag help default) key flag help default
223     (let* ((args #clisp sb-ext:args)
224            #+sbcl sb-ext:posix-argv)
225       (it (member flag args :test 'equalp))
226       (cons key (cond
227                 ((not it) (not it) default)
228                 ((equal default t) nil)
229                 ((equal default nil) t)
230                 (t (thing (second it)))))))
231
232 ; Update settings. If 'help' is set, print help.
233 (defun settings (header options)
234   (let ((tmp (mapcar #'setting options)))
235     (when (! tmp help)
236       (format t "%&-%(a-a-)%~%OPTIONS::~%~" (lines header))
237       (dolist (one options)
238         (format t " -a -a-%~%~" (second one) (third one) (fourth one))))
239     tmp))

```

```

239
240
241
242
243 ; Creates &x for constructor, enables pretty print, hides slots with "_" prefix.
244 (defmacro defstruct+ (x &body body)
245   (let* ((slots (mapcar (lambda (x) (if (consp x) (car x) x)) body))
246         (public (remove-if (lambda (x) (eq #\_ (char (symbol-name x) 0))) slots)))
247     `('progn
248       (defstruct ,x (:constructor , (intern (format nil "%MAKE--a" x))) ,@body)
249       (defmethod print-object ((self ,x) str)
250         (labels ((fun (p) (format nil "~{~a~}~" a" y (slot-value self y))))
251           (format str "~a" (cons 'x (mapcar #'fun ',public)))))
252
253
254
255
256 ; Define one demos.
257 (defvar *demo* nil)
258 (defmacro defdemo (what arg doc &rest src)
259   (push (list 'what 'doc (lambda ,arg ,@src)) *demos*))
260
261 ; Run 'one' (or 'all') the demos. Reset globals between each run.
262 ; Return to the operating systems the failure count (so fails=0 means "success").
263 (defun demos (settings all &optional one)
264   (let ((fails 0)
265         (resets (copy-list settings)))
266     (dolist (trio all)
267       (destructuring-bind (what doc fun) trio
268         (setf what (format nil "~{~a~}~" what))
269         (when (member what (list 'all one) :test 'equalp)
270           (loop for (key . value) in resets do
271             (setf (cdr (assoc key settings)) value))
272           (setf *seed* (or (cdr (assoc 'seed settings)) 10019))
273           (unless (eq t (funcall fun))
274             (incf fails)
275             (format t "~&FAIL-[a]-a-%~" what doc))))
276       #+clisp (ext:exit fails)
277       #+sbcl (sb-ext:exit :code fails)))
278
279
280
281
282 ; test suite
283 (load "tiny")
284 (in-package :tiny)
285
286 ; (print (make-row 12 '(1 2 3 4)))
287 ; (print (make-about '("$aa" "bb"- "cc+")))
288 ; (print (! my *seed*))
289 ; (dotimes (i 20) (print (randi 200)))
290 ; ; (defmethod clone ((d data) &optional src) (make-data (? d about names) src))
291 ; ; (reads ".../Data/auto93.lisp" 'print)
292
293 (defdemo my () "show options" (pprint my) t)
294
295 (defdemo div () "num divs"
296   (let ((s (add (make-sym) ' (a a a b b c))))
297     (and (= 1.379 (rnd (div s))) (eq 'c (mid s)))))
298
299 (demo my *demos* (! my go))

```