

```

1  -- <img align=left width=150 src=head.png>
2  --
3  -- **[Repo](https://github.com/timm/lu) &M~@& [Issues](https://github.com/timm
4  -- /lua/issues) &M~@& [copy;2022] (LICENSE.md)** Tim Menzies
5  --
6  -- If we choose our AI tools not on their complexity, but
7  -- on their understandability, what would they look like?
8  -- To that end, I've been looking back over
9  -- common themes seen in my
10 -- AI graduate students (30+ students, over 20 years). What I was
11 -- after were the least lines of code that offer the most
12 -- AI functionality-- and which could be mixed and matched in
13 -- novel and interesting ways.
14 --
15 -- The result is this file. My standard "intro to AI" exercise is six
16 -- weeks of homeworks where students rebuild the following code, from
17 -- scratch, in any language they like (except LUA). After that,
18 -- students can review all the assumptions of this code, then read the
19 -- literature looking for other tools that challenge those assumptions.
20 -- That leads to a second a 4-6 week project using these tools as a baseline aga
21 -- inst
22 -- which they can compare other, more complex, approaches.
23 --
24 -- <hr>
25 --
26 -- The need for baselines. XXXX
27 --
28 -- Standard supervised learners assume that all examples have labels.
29 -- When this is not true, then we need tools to incrementally
30 -- (a) summarize what has been seen so far; (b) find and focus
31 -- on the most interesting part of that summary, (c) collect
32 -- more data in that region, then (d) repeat.
33 --
34 -- <a href="div.png"></a>
35 -- To make that search manageable, it is useful to exploit a
36 -- manifold assumption; i.e.
37 -- higher-dimensional data can be approximated in a lower dimensional
38 -- manifold without loss of signal [Ch05,Le05].
39 -- Manifolds lead to _continuity_
40 -- effects; i.e. if there are fewer dimensions, then there are more
41 -- similarities between examples.
42 -- Continuity simplifies _clustering_
43 -- (and any subsequent reasoning). More similarities means easier
44 -- clustering. And after clustering, reasoning just means reason about
45 -- a handful of examples (maybe even just one) from each cluster.
46 --
47 -- **ASSIGNMENTS**
48 -- -- **Instance selection**: filter the data down to just a few samples per
49 -- cluster, the reason using just those.
50 -- -- **Anomaly detection**
51 -- -- **Explanation**
52 -- Discretize the numeric ranges (\*) at each level of the recursion,
53 -- then divide the data according what range best selects for one half, or the o
54 -- ther
55 -- at the data at this level of recursion.
56 -- -- **Multi-objective optimization**: This code
57 -- can apply Zitzler's multi-objective ranking predicate [Zit04] to prune the
58 -- worst
59 -- half of the data, then recurs on the rest [Ch18]. Assuming a large over-gener
60 -- ation
61 -- of the initial population (to say, 10,000, examples), this can be just as eff
62 -- ective
63 -- as genetic optimization [Ch18], but runs much faster.
64 -- -- **Semi-supervised learning**: these applications require only the _2.log(N)
65 -- _ labels at
66 -- of the pair of furthest points seen at each level of recursion.
67 -- -- **Privacy**
68 -- -- **Planning**
69 -- -- **Monitoring**
70 --
71 -- local help = [[
72 --
73 -- 15 == a little lab of lots of LUA learning algorithms.
74 -- (c) 2022, Tim Menzies, BSD 2-clause license.
75 --
76 -- USAGE:
77 -- lua 15.lua [OPTIONS]
78 --
79 -- OPTIONS:
80 -- -cohen F Cohen's delta = .35
81 -- -data N data file = etc/data/auto93.csv
82 -- -Dump stack dump on assert fails = false
83 -- -furthest F far = .9
84 -- -Format s format string = %5.2f
85 -- -keep P max kept items = 512
86 -- -p P distance coefficient = 2
87 -- -seed P set seed = 10019
88 -- -todo S start up action (or 'all') = nothing
89 -- -help show help = false
90 -- -want F recurse until rows^want = .5
91 --
92 -- KEY: N=fileName F=float P=posint S=string
93 --
94 -- NOTES: This code uses Aha's distance measure [Aha91] (that can
95 -- handle numbers and symbols) to recursively divide data based on two
96 -- distant points (these two are found in linear time using the Fastmap
97 -- heuristic [Fa95]).
98 --
99 -- To avoid spurious outliers, this code use the 90% furthest points.
100 --
101 -- To avoid long runtimes, uses a subset of the data to learn where
102 -- to divide data (then all the data gets pushed down first halves).
103 --
104 -- To support explanation, optionally, at each level of recursion,
105 -- this code reports what ranges can best distinguish sibling clusters
106 -- C1,C2. The discretizer is inspired by the ChiMerge algorithm:
107 -- numerics are divided into, say, 16 bins. Then, while we can find
108 -- adjacent bins with the similar distributions in C1,C2, then
109 -- (a) merge then (b) look for other merges.
110 -- ]]
111 --
112 -- ## Namespace
113 --
114 -- Cache current globals, use at end to find rogue variables
115 -- local b4={}; for k,_ in pairs(_ENV) do b4[k]=k end
116 --
117 -- Defined local names.
118 -- local any, asserts, big, cli, csv, fails, firsts, fmt, goalp, ignorep, klassp
119 -- local lessp, map, main, many, max, merge, min, morep, new, nump, o, oo, per, pop, push
120 -- local r, rows, rnd, rnds, slots, sort, sum, thing, things, unpack
121 --
122 -- Classes have UPPER CASE names.
123 -- local CLUSTER, COLS, EGS, NUM, ROWS = {}, {}, {}, {}, {}
124 -- local SKIP, SOME, SPAN, SYM = {}, {}, {}, {}
125 --
126 -- ## Settings
127 -- Parse the help text for flags and defaults (e.g. -keep, 512).
128 -- Check for updates on those details from command line
129 -- (and and there,
130 -- some shortcuts are available;
131 -- e.g. _k N _&rArr; 'keep=N';
132 -- and _booleanFlag _&rArr; 'booleanFlag=not default').
133 -- local the={}
134 -- help:gsub("\n [^\n]*%s([^\n]*)", function(key, x)
135 -- for n, flag in ipairs(arg) do
136 -- if flag:sub(1,1)=="-" and key:find("^"..flag:sub(2).."..") then
137 -- x = x=="false" and true or x=="true" and "false" or arg[n+1] end end

```

```

130 if x=="false" then the[key]=false elseif x=="true" then the[key]=true else
131 the[key] = tonumber(x) or x end end )
132
133 -----
134 -- this code reads csv files where the words on line1 define column types.
135 function ignorep(x) return x:find"$" end -- columns to ignore
136 function klassp(x) return x:find"$" end -- symbolic goals to achieve
137 function lessp(x) return x:find"$" end -- number goals to minimize
138 function morep(x) return x:find"$" end -- numeric goals to maximize
139 function nump(x) return x:find"[A-Z]" end -- numeric columns
140 function goalp(x) return morep(x) or lessp(x) or klassp(x) end
141
142 -- strings
143 fmt = string.format
144
145 -- maths
146 big = math.huge
147 max = math.max
148 min = math.min
149 r = math.random
150
151 function rnds(t, f) return map(t, function(x) return rnd(x, f) end) end
152 function rnd(x, f)
153 return fmt(type(x)=="number" and (x~x//1 and f or the.Format) or "%s", x) end
154
155 -- tables
156 pop = table.remove
157 unpack = table.unpack
158 function any(t) return t[#t] end
159 function firsts(a,b) return a[1] < b[1] end
160 function many(t,n, u) u={}; for i=1,n do push(u, any(t)) end; return u end
161 function per(t,p) return t[ (#t*(p or .5))//1 ] end
162 function push(t,x) table.insert(t,x); return x end
163 function sort(t,f) table.sort(t,f); return t end
164
165 -- meta
166 function map(t, f, u) u={}; for k,v in pairs(t) do push(u, f(v)) end; return u end
167 function sum(t, f, n) n=0; for _,v in pairs(t) do n=n+f(v) end; return n end
168 function slots(t, u)
169 u={}
170 for k,v in pairs(t) do k=tostring(k); if k:sub(1,1)=="_" then push(u,k) end end
171 return sort(u) end
172
173 -- print tables, recursively
174 function oo(t) print(o(t)) end
175 function o(t)
176 if type(t)=="table" then return tostring(t) end
177 local key=function(k) return fmt("%s %s", k, o(t[k])) end
178 local u = #t>0 and map(t,o) or map(slots(t), key)
179 return '{ ' .. table.concat(u, " " ) .. " }" end
180
181 -- strings to things
182 function csv(file, x)
183 file = io.input(file)
184 return function()
185 x=io.read(); if x then return things(x) else io.close(file) end end end
186
187 function thing(x)
188 x = x:match("^(%s*)(-)%s*$")
189 if x=="true" then return true elseif x=="false" then return false end
190 return tonumber(x) or x end
191
192 function things(x, sep, t)
193 t={}
194 for y in x:gmatch(sep or "([^\n]+)") do push(t, thing(y)) end
195 return t end

```

```

196 -- CLASSES
197
198
199
200 function new(k,t) k.__index=k; k.__tostring=o; return setmetatable(t,k) end
201
202 -- COLS: turns list of column names into NUMs, SYMs, or SKIPs
203 function COLS.new(k,row, i)
204   i = new(k, {all={}, x={}, y={}, names=row})
205   for at,txt in ipairs(row) do push(i.all, i.col(at,txt)) end
206   return i end
207
208 function COLS.add(i,t)
209   for _,col in pairs(i.all) do col:add( t[col.at] ) end
210   return t end
211
212 function COLS.col(i,at,txt, col)
213   if ignorep(txt) then return SKIP:new(at,txt) end
214   col = (nump(txt) and NUM or SYM):new(at,txt)
215   push(goalp(txt) and i.y or i.x, col)
216   if klassp(txt) then i.klass = col end
217   return col end
218
219 -- NUM: summarizes a stream of numbers
220 function NUM.new(k,n,s)
221   return new(k, {n=0, at=n or 0, txt=s or "", has=SOME:new(), ok=false,
222     w=lessp(s or "") and -1 or 1, lo=big, hi=-big}) end
223
224 function NUM.add(i,x)
225   if x == "?" then
226     i.n = i.n + 1
227     if i.has:add(x) then i.ok=false end
228     i.lo, i.hi = min(x,i.lo), max(x,i.hi); end end
229
230 function NUM.dist(i,x,y)
231   if x=="?" and y=="?" then return 1
232   elseif x=="?" then y=i:norm(y); x=y<0.5 and 1 or 0
233   elseif y=="?" then x=i:norm(x); y=x<0.5 and 1 or 0
234   else x,y = i:norm(x), i:norm(y) end
235   return math.abs(x-y) end
236
237 function NUM.mid(i) return per(i:sorted(), .5) end
238
239 function NUM.norm(i,x)
240   return math.abs(i.hi-i.lo)<1E-9 and 0 or (x-i.lo)/(i.hi - i.lo) end
241
242 function NUM.sorted(i)
243   if i.ok==false then table.sort(i.has.all); i.ok=true end
244   return i.has.all end
245
246 -- ROWS: manages 'rows', summarized in 'cols' (columns).
247 function ROWS.new(k,init, i)
248   i = new(k, {rows={}, cols=nil})
249   if type(init)=="string" then for t in csv(init) do i:add(t) end end
250   if type(init)=="table" then for t in init do i:add(t) end end
251   return i end
252
253 function ROWS.add(i,t)
254   if i.cols then push(i.rows,i.cols:add(t)) else i.cols=COLS:new(t) end end
255
256 function ROWS.clone(i, j) j = ROWS:new(); j:add(i.cols.names); return j end
257
258 function ROWS.dist(i,row1,row2, d,fun)
259   function fun(col) return col:dist(row1[col.at], row2[col.at])^the.p end
260   return (sum(i.cols.x, fun) / #i.cols.x)^1/the.p end
261
262 function ROWS.furthest(i,row1,rows, fun)
263   function fun(row2) return (i:dist(row1,row2), row2) end
264   return unpack(per(sort(map(rows,fun),firsts), the.furthest)) end
265
266 function ROWS.half(i, top)
267   local some, top,c,x,y,tmp,mid,lefts,rights,_
268   some = many(i.rows, the.keep)
269   top = top or 1
270   _x = top:furthest(any(some), some)
271   _c,y = top:furthest(x, some)
272   tmp = sort(map(i.rows,function(r) return top:fastmap(r,x,y,c) end),firsts)
273   mid = #i.rows//2
274   lefts, rights = i:clone(), i:clone()
275   for at,row in pairs(tmp) do (at <=mid and lefts or rights):add(row[2]) end
276   return lefts,rights,x,y,c, tmp[mid] end
277
278 function ROWS.mid(i,cols)
279   return map(cols or i.cols.all, function(col) return col:mid() end) end
280
281 function ROWS.fastmap(i, r,x,y,c, a,b)
282   a,b = idist(r,x), idist(r,y); return {(a^2 + c^2 - b^2)/(2*c), r} end
283
284 -- SKIP: summarizes things we want to ignore (so does nothing)
285 function SKIP.new(k,n,s) return new(k, {n=0,at=at or 0,txt=s or ""}) end
286 function SKIP.add(i,x) return x end
287 function SKIP.mid(i) return "?" end
288
289 -- SOME: keeps a random sample on the arriving data
290 function SOME.new(k,keep) return new(k, {n=0,all={}, keep=keep or the.keep}) end
291 function SOME.add(i,x)
292   i.n = i.n+1
293   if #i.all < i.keep then push(i.all,x) ; return i.all
294   elseif r() < i.keep/i.n then i.all[r(#i.all)]=x; return i.all end end
295
296 -- SYM: summarizes a stream of symbols
297 function SYM.new(k,n,s)
298   return new(k, {n=0,at=n or 0,txt=s or "",has={},most=0}) end
299
300 function SYM.add(i,x,inc)
301   if x ~= "?" then
302     inc = inc or 1
303     i.n = i.n + inc
304     i.has[x] = inc + (i.has[x] or 0)
305     if i.has[x] > i.most then i.most,i.mode=i.has[x],x end end
306
307 function SYM.dist(i,x,y) return x=="?" and y=="?" and 1) or (x==y and 0 or 1) end
308 function SYM.mid(i) return i.mode end
309 function SYM.div(i, p)
310   return sum(i.has,function(k) p=-i.has[k]/i.n;return -p*math.log(p,2) end) end
311
312 function SYM.merge(i,j, k)
313   k = SYM:new(i.at,i.txt)
314   for x,n in pairs(i.has) do k:add(x,n) end
315   for x,n in pairs(j.has) do k:add(x,n) end
316   ei, ej, ejk = i:div(), j:div(), k:div()
317   if i.n==0 or j.n==0 or .99*ek <= (i.n*ei + j.n*ej)/k.n then
318     return k end end
319
320 -- CLUSTER
321
322
323
324 function CLUSTER:recursively divides data by clustering towards two distant points
325 function CLUSTER.new(k,egs,top)
326   local i,want,left,right
327   i = new(k, {here=egs})
328   top = top or egs
329   want = (#top.rows)*the.want
330   if #egs.rows >= 2*want then
331     left, right, i.x, i.y, i.c, i.mid = egs:half(top)
332     if #left.rows < #egs.rows then
333       i.left = CLUSTER:new(left, top)
334       i.right = CLUSTER:new(right, top) end end
335   return i end
336
337 function CLUSTER.show(i,pre, here)
338   pre = pre or ""
339   here=""
340   if not i.left and not i.right then here= o(i.here:mid(i.here.cols.y)) end
341   print(fmt("%s: %-30s %s", #i.here.rows, pre, here))
342   for _,kid in pairs(i.left, i.right) do
343     if kid then kid:show(pre .. "|. ") end end end
344
345 -- EXPLAIN
346
347 -- SPAN: keeps a random sample on the arriving data
348 function SPAN.new(k, col, lo, hi, has)
349   return new(k, {col=col, lo=lo, hi=hi or lo, has=has or SYM:new()}) end
350
351 function SPAN.add(i,x,y,n) i.lo, i.hi=min(x,i.lo),max(x,i.hi); i.has:add(y,n) end
352 function SPAN.merge(i,j)
353   local has = i.has:merge(j.has)
354   if now then return SPAN:new(i.col, i.lo, j.hi, has) end end
355
356 function SPAN.select(i,row, x)
357   x = row[i.col.at]
358   return (x=="") or (i.lo==i.hi and x==i.lo) or (i.lo <= x and x < i.hi) end
359
360 function SPAN.score(i) return i.has.n/i.col.n, i.has:div() end
361
362 function SPAN.good(i, sizes,divs)
363   size,div = i:score()
364   size,div = sizes:norm(size), divs:norm(div)
365   return ((1-size)^2 + (0 - div)^2)^.5 end
366
367 -- EXPLAIN:
368 function EXPLAIN.new(k,egs,top)
369   local i,no,yes,divs,sizes,top,div,best,want,size,left,order,right,spans
370   i = new(k, {here = egs})
371   top = top or egs
372   want = (#top.rows)*the.want -- if enough to recurse
373   if #top.rows >= 2*want then
374     left,right = egs:half(top) -- cluster in two
375     spans, divs, sizes = {}, Num(), Num()
376     for n,coll in pairs(i.cols.x) do -- for each x attribute ...
377       coll2 = j.cols.x[n] -- coll,coll2 is same col in either cluster
378       for _,span in pairs(coll:spans(coll2)) do -- spans are deltas between clust
379         push(spans, span) -- cache the span
380         size, div = span:score() -- remember the span's score (so
381         sizes:add(size) -- we can normalize it, later)
382         divs:add(div) end end
383     order = function(a,b) -- compare two spans, normalizing the scores
384       return a:good(sizes,divs) < b:good(sizes,divs) end
385     best, no = sort(spans, order)[1] -- best span is first in this sort
386     yes, no = egs:clone(), egs:clone()
387     for _,row in pairs(egs.rows) do --
388       (best:selects(row) and yes or no):add(row) end -- divide data in two
389     if #yes.rows<#egs.rows then -- make kids if kid size different to parent siz
390       e
391       if #yes.rows==want then i.yes=EXPLAIN:new(yes,top) end
392       if #no.rows >=want then i.no =EXPLAIN:new(no,top) end end end
393   return i end
394
395 function EXPLAN.show(i,pre)
396   pre = pre or ""
397   tmp = i.here:mid(i.here.y)
398   print(fmt("%s: %-30s %s", #i.yes.rows, pre, o(i.here:mid(i.here.cols.y))))
399   for _,pair in pairs((true,i.yes),{false,i.no}) do
400     status,kid = unpack(pair)
401     k:shpw(pre .. "|. ") end end
402
403 function SYM.spans(i, j)
404   local xys,all,one,last,xys,x,c,n = {},{}
405   for x,n in pairs(i.has) do push(xys, {x,"this",n}) end
406   for x,n in pairs(j.has) do push(xys, {x,"that",n}) end
407   for _,tmp in ipairs(sort(xys,firsts)) do
408     x,c,n = unpack(tmp)
409     if x <= last then
410       last = x
411       one = push(all, Span(i,x,x)) end
412     one:add(x,y,n) end
413   return all end
414
415 function NUM.spans(i, j)
416   local xys,all,lo,hi,gap,xys,one,x,c,n = {},{}
417   lo,hi = min(i.lo, j.lo), max(i.hi, j.hi)
418   gap = (hi - lo) / (6/the.cohen)
419   for x,n in pairs(i.has) do push(xys, {x,"this",1}) end
420   for x,n in pairs(j.has) do push(xys, {x,"that",1}) end
421   one = Span:new(i.lo,lo)
422   all = {one}
423   for _,tmp in ipairs(sort(xys,first)) do
424     x,c,n = unpack(tmp)
425     if one.hi == one.lo > gap then one = push(all, Span(i, one.hi, x)) end
426     one:add(x,y) end
427   all = merge(all)
428   all[1].lo = -big
429   all[#all].hi = big
430   return all end
431
432 function merge(b4, j,n,now,a,b,merged)
433   j,n,now = 0,#b4,{}
434   while j < #b4 do
435     a, b = b4[j], b4[j+1]
436     if b then
437       merged = a:merge(b)
438       if merged then a,j = merged, j+1 end end
439     push(now,a)
440     j = j+1 end
441   return #now == #b4 and b4 or merge(now) end
442

```

```

444 -- DEMOS
445 --
446 --
447 --
448 fails=0
449 function asserts(test, msg)
450   print(test and "PASS: " or "FAIL: ", msg or "")
451   if not test then
452     fails=fails+1
453     if the.dump then assert(test, msg) end end end
454
455 function EGS.nothing() return true end
456 function EGS.the() oo(the) end
457 function EGS.rand() print(r()) end
458 function EGS.some(s, t)
459   s=SOME:new(100)
460   for i=1,100000 do s:add(i) end
461   for j,x in pairs(sort(s.all)) do
462     --if (j % 10)==0 then print("") end
463     --io.write(fmt("%6s", x)) end end
464     fmt("%6s", x) end end
465
466 function EGS.clone( r, s)
467   r = ROWS:new(the.data)
468   s = r:clone()
469   for _,row in pairs(r.rows) do s:add(row) end
470   asserts(r.cols.x[1].lo==s.cols.x[1].lo, "clone.lo")
471   asserts(r.cols.x[1].hi==s.cols.x[1].hi, "clone.hi")
472   end
473
474 function EGS.data( r)
475   r = ROWS:new(the.data)
476   asserts(r.cols.x[1].hi == 8, "data.columns") end
477
478 function EGS.dist( r, rows, n)
479   r = ROWS:new(the.data)
480   rows = r.rows
481   n = NUM:new()
482   for _,row in pairs(rows) do n:add(r:dist(row, rows[1])) end
483   --oo(r.cols.x[2]:sorted()) end
484   o(r.cols.x[2]:sorted()) end
485
486 function EGS.many( t)
487   t={}; for j=1,100 do push(t, j) end
488   --print(oo(many(t, 10))) end
489   o(many(t, 10)) end
490
491 function EGS.far( r, c, row1, row2)
492   r = ROWS:new(the.data)
493   row1 = r.rows[1]
494   c, row2 = r:far(r.rows[1], r.rows) end
495   --print(c, "\n", o(row1), "\n", o(row2)) end
496
497 function EGS.half( r, c, row1, row2)
498   local lefts, rights, x, y, x
499   r = ROWS:new(the.data)
500   r:mid(r.cols.y)
501   lefts, rights, x, y, c = r:half()
502   lefts:mid(lefts.cols.y)
503   rights:mid(rights.cols.y)
504   asserts(true, "half") end
505
506 function EGS.cluster(r)
507   r = ROWS:new(the.data)
508   --CLUSTER:new(r):show() end
509   CLUSTER:new(r) end
510
511 -- start-up
512 if arg[0] == "slua" then
513   oo(the)
514   if the.help then print(help:gsub("\nNOTES:*$", "")) else
515     local b4={}; for k,v in pairs(the) do b4[k]=v end
516     for _,todo in pairs(the.todo=="all" and slots(EGS) or {the.todo}) do
517       for k,v in pairs(b4) do the[k]=v end
518       math.randomseed(the.seed)
519       if type(EGS[todo])=="function" then EGS[todo]() end end
520     end
521     for k,v in pairs(_ENV) do if not b4[k] then print("?", k, type(v)) end end
522     os.exit(fails)
523   else
524     return {CLUSTER=CLUSTER, COLS=COLS, NUM=NUM, ROWS=ROWS,
525            SKIP=SKIP, SOME=SOME, SYM=SYM, the=the, oo=oo, o=o}
526   end
527
528 -- git rid of SOME for rows
529 -- nss = NUM | SYM | SKIP
530 -- COLS = all:[nss]t, x:[nss]*, y:[nss]*, klass:col?
531 -- ROWS = cols:COLS, rows:SOME
532 -- ## References
533 -- - [Ah91]:
534 -- Aha, D.W., Kibler, D. & Albert, M.K. Instance-based
535 -- learning algorithms. Mach Learn 6, 37&M-^@M-^S66 (1991).
536 -- https://doi.org/10.1007/BF00153759
537 -- - [Boley, 1998]:
538 -- Boley, D., 1998.
539 -- [Principal directions divisive partitioning](https://www-users.cse.umn.edu/~boley/publications/papers/PDDP.pdf)
540 -- Data Mining and Knowledge Discovery, 2(4): 325-344.
541 -- - [Ch05]:
542 -- [Semi-Supervised Learning](http://www.molgen.mpg.de/3659531/MITPress--SemiSupervised-Learning)
543 -- (2005) Olivier Chapelle, Bernhard Sch&#252;l&#228;kopf, and Alexander Zien (eds).
544 -- MIT Press.
545 -- - [Ch18]:
546 -- [Sampling&M-^@M-^] as a Baseline Optimizer for Search-Based Software Engineer
547 -- ing](https://arxiv.org/pdf/1608.07617.pdf),
548 -- Jianfeng Chen; Vivek Nair; Rahul Krishna; Tim Menzies
549 -- IEEE Trans SE, (45)6, 2019
550 -- - [Ch22]:
551 -- [Can We Achieve Fairness Using Semi-Supervised Learning?](https://arxiv.org/p
552 -- df/2111.02038.pdf)
553 -- (2022), Joymallya Chakraborty, Huy Tu, Suvodeep Majumder, Tim Menzies.
554 -- - [Fal95]:
555 -- Christos Faloutsos and King-Ip Lin. 1995. FastMap: a fast algorithm for index
556 -- ing, data-mining and visualization of traditional and multimedia datasets. SIGMO
557 -- D Rec. 24, 2 (May 1995), 163&M-^@M-^S174. DOI:https://doi.org/10.1145/568271.223
558 -- 812
559 -- - [Le05]:
560 -- Levina, E., Bickel, P.J.: [Maximum likelihood estimation of intrinsic dimensi
561 -- on](https://www.stat.berkeley.edu/~bickel/mldim.pdf).
562 -- In:
563 -- Advances in neural information processing systems, pp. 777&M-^@M-^S784 (2005)
564 -- - [Pl04]:
565 -- Platt, John.
566 -- [FastMap, MetricMap, and Landmark MDS are all Nystrom Algorithms](https://www
567 -- .microsoft.com/en-us/research/wp-content/uploads/2005/01/nystrom2.pdf)
568 -- AISTATS (2005).
569 -- - [Zit04]:
570 -- [Indicator-based selection in multiobjective search](https://link.springer.co
571 -- m/chapter/10.1007/978-3-540-30217-9_84)
572 -- Eckart Zitzler, Simon K&#228;nzli
573 -- Proc. 8th International Conference on Parallel Problem Solving from Nature (P
574 -- PSN VIII

```