Datizrace: daži filosofiski secinājumi

Kārlis Podnieks

LU asociētais profesors

2005. gada 26. maijā

Atkāpe: vārdi un sejas

Kādas asociācijas Jums saistās ar Petri tīkliem? Vai tiem ir kāds sakars ar naftas produktiem? Kas ir Petri? Ja cilvēks – tad kā viņu sauc? Kad viņš ir dzimis un kā izskatās? Vai vēl dzīvs?

Man nepatīk matemātikā un datorzinātnē iesakņojies paradums ignorēt svarīgu ideju autoru personības! Var izlasīt veselu grāmatu par kādu nozari, un tā arī neuzzināt pat visu iesaistīto galveno cilvēku iniciāļus…

Niederländische Königin zeichnet Prof. Dr. Carl Adam Petri für sein Lebenswerk aus (attēlu var palielināt).

Carl Adam Petri – dzimis 1926.gadā Leicigā – Petri tīklus ieveda 1962.gadā savā doktora disertācijā “Kommunikation mit Automaten ” - laudatio

All models are wrong…

"All models are wrong, but some are useful."
***George E. P. Box* (1979, p. 202)**

Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer, & G. N. Wilkinson (Eds.), *Robustness in statistics* (pp. 201-236). New York: Academic Press.

George E. P. Box (dzimis 1919, viens no statistikas klasiķiem).

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association, 71,* 791-799. Merriam-Webster.

Attēls no ASQ

Solītie secinājumi

1. Datizrace radās daudz senāk par to brīdi, kad kāds to nosauca par data mining. Netriviālu “datu rakšanu” izmantoja jau senās Romas juristi.

2. Statistikas un datizraces prakse liecina, ka arī ļoti neprecīzs (“nepareizs”) modelis var būt sekmīgs (efektīvs, piemēram, naudas izteiksmē).

3. Tātad, droši vien, arī “nepareiza” (ne gluži “pareiza”) loģika var būt ļoti efektīva. Kā tad var rasties “absolūti pareiza” loģika? Nevar!

Avoti

[1] L. J. Maistrovs. *Varbūtības jēdziena attīstība*. “Nauka”, Maskava, 1980, 270 lpp. (krievu val.)

[2] M. Ptuha. *XVII-XVIII gadsimta statistikas vēstures apraksti*. “Gospolitizdat”, 1945, 350 lpp. (krievu val.)

[3] A. V. Aņikins. *Zinātnes jaunība. Domātāju-ekonomistu dzīve un idejas – līdz Marksam*. “Politizdat”, Maskava, 1979, 370 lpp. (krievu val.)

[4] D. Hand, H. Mannila, P. Smyth. *Principles of Data Mining*. MIT, 2000, 550 pp.

[5] S. Mitra, T. Acharya. *Data Mining: Multimedia, Soft Computing, and Bioinformatics*. Wiley&Sons, 2003, 400 p.

[6] E. Mendelsons. *Ievads matemātiskajā loģikā*. “Nauka”, Maskava, 1984, 330 lpp. (krievu val.).

Kapteinis Graunts (1662)

John Graunt (1620-1674) - galantērijas tirgotājs, Londonas milicijas kapteinis (vēlāk – majors), statistikas (un datizraces) "tēvs".

Attēls no University of York

J.Graunt. Natural and Political Observations upon the Bills of Mortality, 1662

Šī Graunta grāmata ir pirmais sistemātiskas datizraces gadījums vēsturē.

Graunta laikmets:
1641 – revolūcija, 1649 +, 1660 – restaurācija, 1666 – lielais Londonas ugunsgrēks, ik pa laikam – mēra epidēmijas (1563, 1593, 1603, 1625, 1665 u.c.)

Graunts II

Biežo mēra epidēmiju dēļ Londonas pašvaldība jau 1500-jos gados sāka vākt, apkopot un publicēt statistikas datus par mirušajiem (Mortality Bills), vēlāk šiem datiem pievienoja datus par laulībām, bērnu piedzimšanu un kristīšanu.

Šajos datos Graunts **pamanīja ("izraka") regularitātes**.
Piemēram, Londonā: 1) zēni un meitenes dzimst proporcijā 14/13, 2) vīrieši mirst vairāk nekā sievietes, 3) vīriešu ir vairāk nekā sieviešu, 4) cilvēka dzīves sākumā mirstība ir vislielākā, utt.

No regularitātēm Graunts (ne vienmēr korekti) atvasināja dažādus **vidējos lielumus**. Piemēram, katrā ģimenē ir vidēji 8 cilvēki, katrās 11 ģimenēs ik gadu nomirst vidēji 3 cilvēki ("kross-validācija" pa draudzēm?).

Graunts III

Avoti: Ian Johnston. A Handbook on the History of Modern Science, Section 4.
Ed Stephan. Official John Graunt Site of the 1996 Olympic Games.
Ed Stephan's Timeline of Demography

Atklātās sakarības Graunts izmantoja, lai “izraktu” atbildes uz tādiem jautājumiem kā:

Cik Londonā ir iedzīvotāju?
Cik Londonā ir militārajam dienestam derīgu vīriešu?
Vai poligāmija kā dzimstības veicinātājs (pēc mēra epidēmijas) būtu lietderīga?

[Vairāki Graunta rezultāti bija nepamatoti un pat nepareizi.]

Graunts IV

Graunts vairākos veidos (kāpēc?) aprēķināja (t.i. “izraka”) ka Londonas iedzīvotāju skaits ir ~384000 (tolaik daudzi domāja, ka tas ir vairāki miljoni…).

Lūk, viens no veidiem: Londonā gadā mirst vidēji 13000, tātad ģimeņu skaits ir 13000*11/3 jeb ~48000, bet iedzīvotāju skaits: 48000*8, jeb ~384000.

Reālais skaitlis varēja būt par ~10% lielāks, jo “nepareizas ticības” mirušos toreiz nereģistrēja…

Militārajam dienestam derīgu vīriešu skaits ir ~81233. Poligāmija nebūtu lietderīga, jo…

Politiskā aritmētika

Savās statistisko datu analīzēs Graunts balstījās uz tīri intuitīviem priekšstatiem, brīvi izmantojot vidējās vērtības. Varbūtības jēdzienu viņš vēl tikai apjauta, tomēr viņam jau bija izveidojusies nojauta par lielo skaitļu likumu.

Graunta draugs un sekotājs William Petty (1623-1685) Graunta metodi nosauca par **politisko aritmētiku** - "the art of reasoning by figures upon things relating to the government".

Termins "Statistik" tika ieviests 18.gs. Vācijā, atvasinot to no vārda "Staat" (valsts, t.i. "valstszinība", Staatskunde, valstu salīdzināšana).

Avots: Stuart Witt, Statistics and Political Science, 1993

Cilvēka dzīves ilgums

Mirstības statistikā visvairāk ieinteresēti ir dzīvības apdrošinātāji. Šis bizness bija pazīstams jau senatnē. Kādu mūža pensiju katru gadu maksāt dotajam cilvēkam? Kādu prēmiju katru gadu prasīt no dotā cilvēka, apdrošinot viņu nāves gadījumam? [Izskaidrot.]

Šim nolūkam nepieciešams iemācīties aprēķināt **doto vecumu sasnieguša cilvēka sagaidāmo dzīves ilgumu**. Un šo sakarību vajag "izrakt" no mirstības un dzimstības datiem. (Kaut vai no pieminekļu uzrakstiem kapsētās, sk. Death and the Romans).

Kā to darīja pats Graunts? Sk. arī
H. L. Seal. Early Uses of Graunt's Life Table, 1980

Kā to darīja Graunts

No mirstības datiem Graunts secināja, ka no 100 piedzimušiem līdz 6 gadu vecumam nomirst **36**, un ka 76 gadu vecumu pārsniedz labi ja viens ("perhaps but one surviveth 76"). Tā kā precīzu datu par mirušo vecumu viņam nebija, tad pārējo viņš atrod ar savdabīgu "interpolāciju": 6-15 gadu vecumā nomirst **24**, 16-25 gadu – **15**, 26-35 gadu – **9**, 36-45 gadu – **6**, 46-55 – **4**, 56-65 – **3**, 66-75 – **2**, 76-85 – **1**.

Summējot uz atpakaļ, Graunts secina, ka dzīvi ir: 76 gadu vecumā – 1, 66 – 3, 56 – 6, 46 – 10, 36 – 16, 26 – 25, 16 – 40, 6 – 64. Tas viņam nozīmē iedzīvotāju vecumu spektru.

Tātad armijai derīgā vecumā (16-56 gadi) ir 40-6=34, t.i. 34% no visiem iedzīvotājiem.

Domitius Ulpianus (228)

Graunta “interpolācija” (“guesstimation”) ir cienījams modelēšanas piemērs datizraces stilā. Bet pati ideja nebija jauna, par to jau sen bija rakstījis romiešu jurists Domitius Ulpianus (~170 - 228), sk. šo Ulpian life table:

G. Simon. History of Insurance (Presentation), 2000

Te periods bija 5 gadi un iedomātie skaitļi: 30 (līdz 20 gadiem), 27, 25, 22, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 7, 5 (summa 100).

Pa īstam pamatoti šo uzdevumu spēja atrisināt tikai matemātiķi…

Brāļi Heigensi

Christiaan Huygens (1629-1695), konsultējot savu brāli Ludwig Huygens 1662. gadā, izmantojot Graunta tabulas, parādīja kā aprēķināt doto vecumu sasnieguša cilvēka vidējo dzīves ilgumu:

36*3+24*11+15*21+9*31+6*41+4*51+3*61+2*71+1*81 = 1822 / 100 = 18
(jaundzimušā vidējais dzīves ilgums)

24*11+15*21+9*31+6*41+4*51+3*61+2*71+1*81 = 1714 / 64 = 27
(6 gadus veca bērna vidējais dzīves ilgums)

15*21+9*31+6*41+4*51+3*61+2*71+1*81 = 1450 / 40 = 36
(16 gadus cilvēka vidējais dzīves ilgums)

9*31+6*41+4*51+3*61+2*71+1*81 = 1135 / 25 = 45
(26 gadus cilvēka vidējais dzīves ilgums)

Uzdevums

K. Heigenss, vēstulē brālim, risina šādu uzdevumu:

[Londonā] 56 gadus vecs vīrietis apprecējis 16 gadu vecu sievieti. Cik ilgi viņi nodzīvos kopā? Kad nomirs pirmais no viņiem, man apsolīti 100 franki. Par kādu cenu es šodien varētu šo saistības rakstu pārdot?

Te jau savienojas statistika un azartspēles – divi varbūtību teorijas avoti.

Par īpašuma apdrošināšanu - sk. 1. nodaļu manā varbūtību teorijas grāmatiņā vidusskolām.

Beidzot - korekti dati (1692)

K. Heigenss pirmais saprata, ka korektas “dzīves ilguma līknes” aprēķināšana, balstoties uz reāliem datiem, ir ļoti sarežģīts uzdevums. (Arī pati līkne ir viņa ideja, sk. bildes: Roman Life Expectancy no Canadian Content)

Edmond Halley (1656-1742), astronoms un matemātiķis. Viņam pirmajam izdevās aprēķināt (kāpēc?) Breslaw pilsētas (vēlāk - Breslau, Wroclaw) 1687. – 1691. gada reālajai situācijai atbilstošas korektas mirstības un cilvēku vecuma tabulas.

Sk. Hellija 1692. gada rakstu un tabulas – no Pierre Marteau’s Publishing House

.

De Muavra “mirstības formula”

Abraham de Moivre (1667-1754), aplūkojot Hellija tabulu, konstatēja, ka, no 12 līdz 70 gadu vecumiem cilvēku skaita diferences ir no 6 līdz 10, tāpēc šiem vecumiem tabulu ar samērā lielu precizitāti var aizstāt formula 646-8(V-12), jeb: [Breslaw pilsētā] no 86 dzimušiem gadā nomirst viens (646/8 = 86).

De Moivre. Evaluation of Annuities on Lives, **1724**

Pats Hellijs līdz tam neaizdomājās, un sūdzējās par aprēķinu sarežģītību. Muavra formula tos vienkāršoja (precizitāte apdrošināšanas aprēķiniem izrādījās pietiekama).

Muavra piegājiens – lineāru sakarību meklēšana datos, ir viens no mūsdienu datizraces tipiskiem paņēmieniem (tiesa, šodien mēs varam rīkoties daudz brīvāk – mūsu datori tiek galā arī ar nelineārām sakarībām).

Secinājums Nr. 1

Datizrace radās daudz senāk par to brīdi,
kad kāds to nosauca par
data mining.

Netriviālu “datu rakšanu” izmantoja jau senās Romas juristi un 17.gadsimtā - kapteinis Graunts.

Captain John Graunt

Kas tad ir datizrace?

Termins “**data mining**” radies statistiķu aprindās kā nievājošs apzīmējums datorizācijas jauno iespēju izraisītajai “aklai” pārmeklēšanai bez iepriekš noformulētām pārbaudāmām hipotēzēm (dažreiz tas saukts arī par “**data dredging**” – datu bagarēšana).

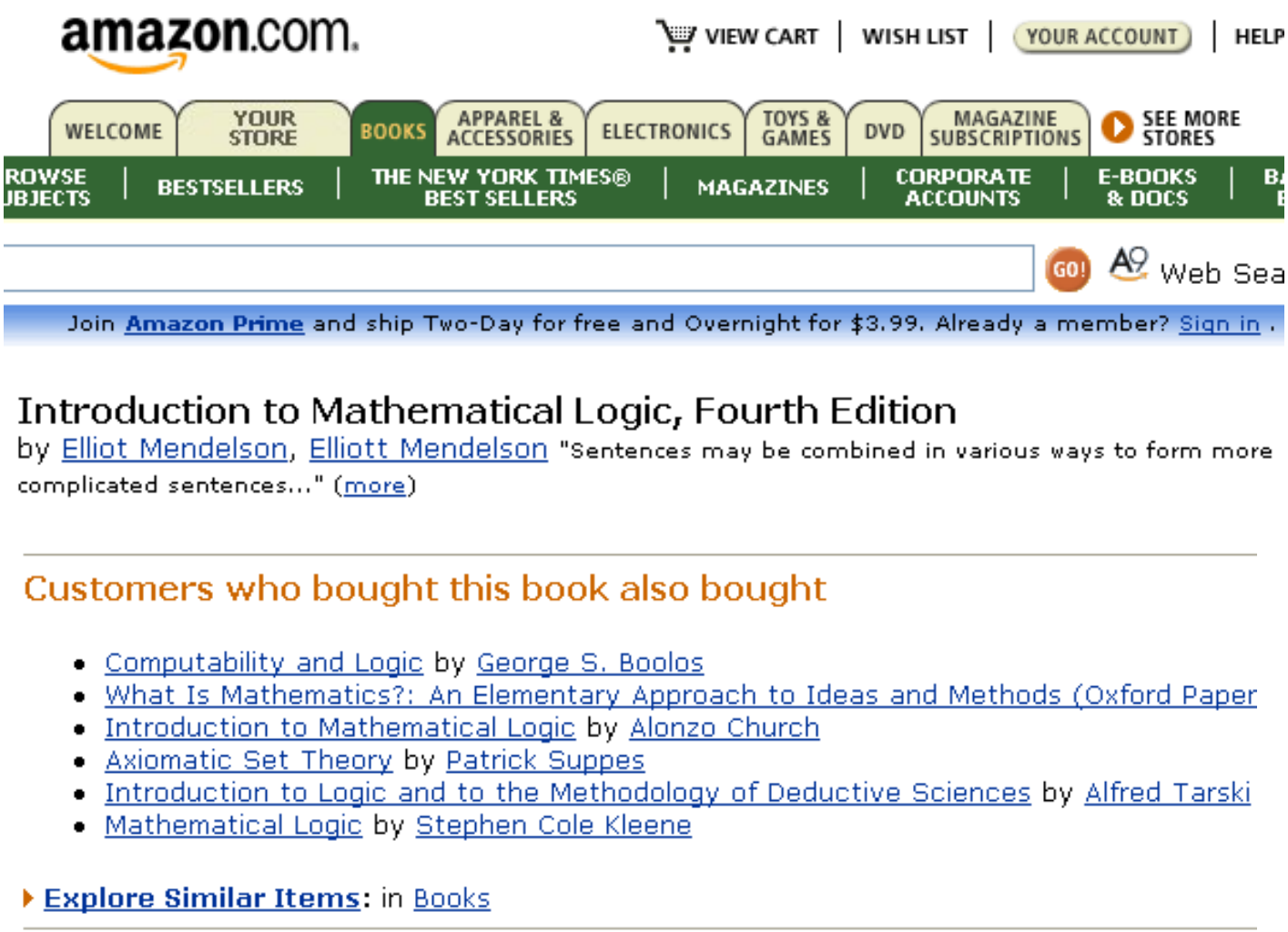
Sākumā – atsevišķi pētījumi, bet nozares pirmais konsolidācijas pasākums bija konference 1989. gada augustā, Detroitā, ar priekšlikumu saukt to par **knowledge discovery in databases (KDD).**

Gregory Piatetsky-Shapiro – galvenais organizators, sk. viņa “Knowledge Discovery in Databases: 10 years after”, SIGKDD Explorations, Vol 1, No 2, Feb 2000.

Divas pieejas terminoloģijai: a) identificēt datizraci ar KDD, b) definēt datizraci tikai kā vienu KDD soli – modeļu vai šablonu izguvi no attīrītiem un sagatavotiem datiem, bez talākas interpretācijas. Pamazām uzvaru gūst (a)…

Konflikts ar statistiķiem? “Why do statisticians “hate” us?” no Susan Imberman. Datizraces īpatnība – miljardiem objektu, tūkstošiem dimensiju.

Tirgus grozu uzdevums



Market-basket: meklējam un izmantojam likum-sakarības pircēju uzvedībā (ar un/vai bez krāpšanas!)

Asociāciju atklāšana: modelis

[4] Mūsu dati ir matrica, n rindas (pirkumi vai pircēji), p kolonas (preces), tikai 0 un 1 (1 nozīmē, ka šajā pirkumā šī prece ir nopirkta):

```
00011000100000000000001000000000000000000000
00000100000000000001100000000000010000000000
00000000000000000001000000000000010000000000
00001000000001000000000000000000000000110000
00000000000000000000000000000000000000001010
00011000000000000000000000000000000000000000
00000000000010000000100000000000010000000000
...
```

Tipiskos gadījumos n=10000..10000000, p=1000..100000, un matrica ir ļoti stipri izkliedēta (sparse, t.i. 1 ir ļoti maz).

Ticamība un atbalsts

Uzdevums ir sameklēt matricā šādas **asociācijas *(likumus, likumsakarības)***: ja pirkumā ir tādas un tādas preces, tad, ar ievērojamu varbūtību, šajā pirkumā būs arī noteiktas citas preces.

A(1), …, A(p) – mainīgais-pirkums.

Asociācija: $A(i_1)\&\ldots\&A(i_k)\text{->}A(j_1)\&\ldots\&A(j_s)$. Jeb, īsāk: f->g.

Piemēram, A(1)&A(2)->A(3).

Mūs interesē tādas f->g, kam ir liela nosacītā varbūtība P(g|f), t.i. ja zināms, ka A(1)=1 un A(2)=1, tad ar lielu varbūtību (50% vai tml.) arī A(3)=1. Šo varbūtību varētu saukt par asociācijas **ticamību** (confidence, accuracy):

confidence(f->g) = P(g|f) = P(f&g)/P(f) = freq(f&g) / freq(f) = count(f&g) / count(f).

Varbūtību P(f&g) sauc par asociācijas **atbalstu** (support):

support(f->g) = P(f&g) = freq(f&g) = count(f&g)/count(all).

Ticamas asociācijas, ja to atbalsts ir mazs, parasti nav interesantas.

Uzdevums

Tātad, vēl precīzāk, mūs interesētu atrast tādas **asociācijas, kuru atbalsts un ticamība ir ievērojama**. Piemēram, attiecīgi ne mazāki par 0,01% un 50%. Tad šo informāciju varētu kaut kā izmantot.

Vispārīgajā gadījumā (patvaļīga 0,1-matrica) asociāciju meklēšana ir bezcerīgs uzdevums: visu iespējamo preču apakškopu skaits ir 2^p, kur p=1000..100000.

Bet reālos datos, t.i. stipri izkliedētās matricās šis uzdevums, izrādās, ir reāli atrisināms.

APriori algoritms

Šis uzdevums reducējas uz **bieži sastopamu preču kopu** atrašanu: kopa B ir bieži sastopama, ja matricā freq(B)>=0,01% (vai tml.). Ja mēs prastu pietiekami ātri ģenerēt visas 0,01%-īgās preču kopas, tad arī asociāciju meklēšanas uzdevumu mēs varētu atrisināt.
Tā šeit ir ļoti svarīga ideja.

APriori algoritms

R. Agrawal, R. Srikant, Fast Algorithms for Mining Association Rules, IBM, 1994

Ideja: ja kopa ir bieži sastopama, tad tās apakškopas nav mazāk biežas. Tātad jāsāk ar viena elementa kopām, jānoskaidro, kuras no tām ir biežas (>=0,01%), pēc tam no tām jābūvē visas iespējamās divu elementu kopas, jānoskaidro to biežums, atlasot derīgās utt.

Pseidokods

APriori algoritma pseidokods (sk. arī kodu no Ferenz Bodon)

```
k:=1;
C(1) := { {j} | j = 1..p}; // kopas ar elementu skaitu 1
while C(k) nav tukša do
        matricas caurskate:
        katrai kopai no C(k) noskaidro, vai tā ir bieža;
        L(k) := visu biežo C(k) kopu kolekcija;
        kandidātu formēšana:
        C(k+1) := visas k+1 elementu kopas, kurām visas
                apakškopas ir no L(k);
        k:=k+1;
end;
```

Laika novērtējumi

APriori algoritms beidz darbu, kad lielāka apjoma biežas kopas vairs netiek atrastas.

C(k) kopu biežuma testēšanas laiks: O(n*p*|C(k)|).

C(k+1) formēšanas laiks nav atkarīgs no n. Tas ir $O(|C(k)|^3)$ - ņem L(k) kopu pārus, apvieno, ja apvienojumā ir tieši k+1 elements, tad to iekļauj kopā C(k+1). Praksē kubiska laika vietā iznākot lineārs laiks.

Kopējais laiks: O(n*p*sum(|C(k)|).

Vēl ātrāk…

C(k+1) formēšanas paātrināšana

Ideja: visas atrastās biežās kopas glabāt šādā kokā:

1 2 3 4

12 13 14 23 24 34

123 124 134 234

1234

Tad, veidojot, piemēram, 3 elementu kopas, jāskata cauri 2.līmenis. No 12 var veidot 123, ja 23 ir palicis neizmests 2.līmenī, un 124 – ja 24 ir palicis neizmests 2.līmenī.

Vispārīgi: no kopas i1, …, ik veidojot k+1 elementu kopas, ejot no koka virsotnes, k-jā līmenī jāņem visas tur esošās kopas (i2, .., ik, j) un jāveido kopas (i1, …, ik, j).

C(k) kopu testēšanas paātrināšana

Radikāls risinājums ir **iztveršana** (sampling) – pilnas matricas vietā sākumā izmantot stipri mazāku (bet pareizi atlasītu) apakškopu. Iztvērumam atrastās biežās kopas beigās vajag testēt uz pilnas matricas.

Kādas asociācijas ir interesantas?

APriori algoritms “rok” visas asociācijas, kam ir ievērojams atbalsts – kā lietotājam interesantās, tā neinteresantās.

Lietotāja intereses pakāpe par doto asociāciju var būt atkarīga no specifiskiem apsvērumiem, piemēram, no iesaistīto preču cenām (“dārgās” asociācijas varētu būt interesantākas par “lētajām”). Tādos gadījumos meklēšanu vajag speciāli “iestatīt”. [Asociācijas starp pircēja dzimšanas mēnesi un viņa “uzvedību”?]

Bet, izrādās, ir iespējami arī tādi “intereses mēri”, ko var atvasināt tieši no datu statistiskajām īpašībām. Un tos var iemācīt asociāciju meklēšanas algoritmam!

Intereses mēri

Ideja: asociācija f->g ir interesanta, ja f parādīšanās stipri ietekmē g parādīšanās varbūtību. T.i. vajag salīdzināt varbūtību P(g|f) ar varbūtību P(g). Ja tās ir vienādas vai tuvas, tad f un g ir maz-atkarīgi, un tāda asociācija nebūs interesanta. Tātad intereses mērs varētu būt

interest(f->g) = P(g|f) - P(g) = confidence(f->g) – freq(g).

Lai varētu meklēt interesantās asociācijas, APriori algoritms ir jāpārveido (sk. [4], Section 13.6).

Tiek izmantoti arī citi intereses mēri. Tie visi cenšas vērtēt varbūtību sadalījuma {P(g|f), P(~g|f)} "diverģenci" no sadalījuma {P(g), P(~g)}. Vispopulārākais ir t.s. J-measure:

J(f->g) = P(f)(P(g|f)log(P(g|f)/P(g)) + P(~g|f)log(P(~g|f)/P(~g)).

Izteiksme iekavās ir Kullback-Leibler divergence (no MathDaily), jeb H(g,f) - H(g).

Korelācija nenozīmē cēloņsakarību

Tas ir sen zināms statistikas likums. Un viens no cēloņiem, kāpēc statistiskie modeļi ir neprecīzi – tie **modelē parādību izpausmes, nemēģinot atklāt to mehānismu.**

Piemērs. Ja cilvēks lielveikalā pērk labus vīnus, tad viņš ļoti bieži pērk arī "dizaineru" apģērbus. Labu vīnu pirkšana neizraisa labu apģērbu pirkšanu. Cēlonis ir citur: labu vīnu pirkšana liecina par cilvēka pārticības augstāku pakāpi, un ar tādiem ienākumiem ir dabiski pirkt arī dārgākus apģērbus.

Bet tas nenozīmē, ka šādas korelācijas - pat, varbūt, **nemaz** neizprotot to cēloņus - lielveikali nevarētu izmantot savā labā.

Neironu tīkli

Neironu tīkli vēl viens tipisks "seklās" modelēšanas līdzeklis, ko izmanto datizracē. Nemēģinot atklāt objekta, struktūru un mehānismu, mēs cenšamies atveidot tā ārējo izturēšanos **viena un tā paša veida struktūrā**. Piemēram, viena līmeņa perceptronā:

m ieejas $x_1, \ldots, x_m$ (reāli mainīgie),

n neironi $y_1, \ldots, y_n$ starpslānī,

un viena izeja z,

$y_j = h(a_{i1}*x_1+a_{i2}*x_2+\ldots+a_{im}*x_m)$, kur $\{a_{ij}\}$ ir mxn matrica, bet h ir kāda fiksēta nelineāra funkcija, piemēram, $h(x)=1/(1+e^{-x})$,

$z = b_1*y_1+b_2*y_2+\ldots+b_n*y_n$.

Ja atmestu nelineāro funkciju h, tad z varētu būt tikai lineāra funkcija no $x_1, \ldots, x_m$. Bet ar h – izrādās, ka daudzos gadījumos "apmācības ceļā" izdodas izrēķināt tādu matricu $\{a_{ij}\}$ un vektoru $\{b_k\}$, ka z atkarība no $x_1, \ldots, x_m$ ir labs tuvinājums iepriekš dotai funkcijai, kuras analītiskā izteiksme nav zināma.

[Vai Furjē transformācijas darbība nav līdzīga neironu tīkliem?]

Pārlieka pielāgotība (overfitting)

Ja mūsu modelēšanas metode ir pietiekami elastīga, dažkārt nav grūti precīzi “nomodelēt” arī ļoti lielus reālu datu apjomus. Bet pēc tam izrādās, ka šāds modelis slikti prognozē, saņemot jaunus, modeļa būves laikā nezināmus ievaddatus.

Kāpēc? Cenšoties modelēt pārāk precīzi, esam “nomodelējuši” arī datos sastopamos “trokšņus”. Šo parādību sauc par **pārlieku pielāgotību** (overfitting).

Robustāka modelēšana parasti dod labākus rezultātus.
Kross-validācija – kā tā notiek.
Bias-variance trade-off – balansēšana starp ievirzi un novirzi.
Ar šīm problēmām datizraces teorija nopietni nodarbojas.

Mathematistry

25. maijā Google atrada tikai 4 avotus, kur šis termins pieminēts. Autors, šķiet, atkal ir Dž. Bokss:

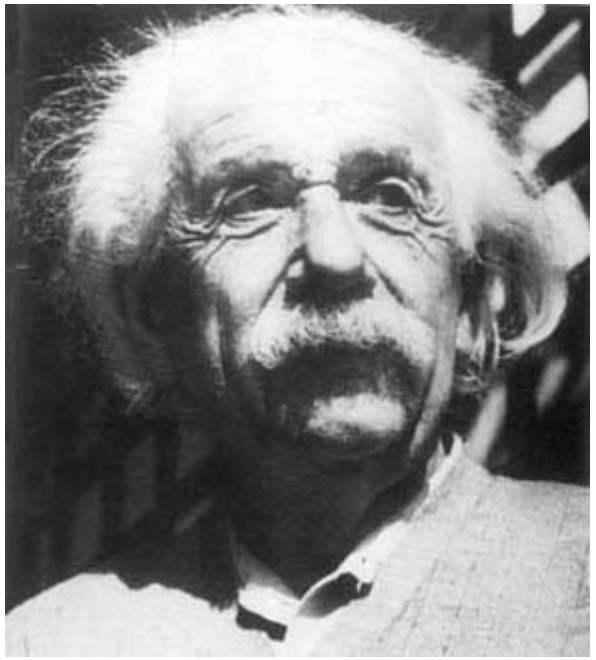
"Mathematistry is characterized by development of theory for theory's sake, which, since it seldom touches down with practice, has a tendency to **redefine a problem rather than to solve it**." **[Meklēsim tur, kur gaišāks…]**
Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association, 71,* 791-799. Merriam-Webster.

"The problem with mathematistry is that the statistical work is ignored by the scientific community."
"While evidence of mathematistry exists in various disciplines from engineering to psychology, it is nonetheless harmful because "valuable talents are wasted at a period in history when they could be put to good use."
[Pašā matemātikā matemātistrija netiek nosodīta… Bet statistikā iespējas pētīt reāli neeksistējošu sistēmu modeļus ir tik lielas, ka…]

Einšteins

Everything should be made as simple as possible, but not simpler.

Attributed to Albert Einstein.



Sk. What is Occam's Razor?
no Usenet Physics FAQ

Attēls no MacTutor History of Mathematics

Secinājums Nr. 2

Statistikas un datizraces prakse liecina,
ka arī ļoti neprecīzs (“nepareizs”)
modelis
var būt sekmīgs
(efektīvs, piemēram, naudas
izteiksmē).

“All models are wrong, but some are useful.”
George E. P. Box

Loģika

Ja pat ļoti aptuveni modeļi var dot labumu (vismaz, vairāk naudas), tad, varbūt, absolūti precīzi modeļi sare\z\g\itos gadījumos ne tikai nav vajadzīgi, bet pat – nemaz nav iespējami?

Tipisks gadījums, kad vairums cilvēku pret modeli izturas kā pret absolūti patiesību, ir loģika (precīzāk, tas loģikas variants, ko matemātiskajā loģikā sauc par **klasisko loģiku**).

Cilvēkus apmāca izmantot klasisko loģiku gan matemātikas, gan ikdienas dzīves spriedumos. Un cilvēks var sekmīgi nodzīvot visu mūžu, izmantojot šo loģiku, un tā arī neuzzinot, ka, stingri ņemot, šī loģika ir daļēji “nepareiza”.

Kā tas notika?

Vai atceramies, kā mums iemācīja cienīt tradicionālo klasisko loģiku kā kaut ko absolūti pareizu, kā vienīgi iespējamo garantēti pareizu secinājumu iegūšanas līdzekli matemātikā? Un kā kļūdu meklēšanas kritēriju “nepareizos” spriedumos?

Kā tas notika? Vidusskola, matemātika, matemātiskā loģika…

Vai (Ax in b)F(x) ir patiess, ja b ir tukša kopa?

Ir izveidojies paradums domāt, ka klasiskā loģika ir kaut kas absolūti “pareizs”, ko cilvēks nav izdomājis, bet tikai atklājis.

Bet vai tā ir?

Būla algebra

George Boole (1815-1864). Mēģināsim atcerēties…

Pieņemsim, ka 1 nozīmē “patiess”, bet 0 – “aplams”.
~0 = 1, ~1 = 0
0&0 = 0&1 = 1&0 = 0, 1&1=1
0v0 = 0, 0v1 = 1v0 = 1v1 = 1

Bet kādiem jābūt 0->0, 0->1, 1->0, 1->1?
Liekas, skaidrs, ka 1->0 = 0. Tad no A = 1 un A->B = 1 seko B = 1. Tā vajadzētu. Bet pārējie 3 gadījumi?

Vai atceraties, kā Jums iestāstīja, ka vajag

0->0 = 0->1 = 1->1 = 1 ?

“Ja V. Skots nav uzrakstījis nevienu romānu, tad ASV nav bijis pilsoņu kara.” [6]

Kopu algebra

Kopas K visas apakškopas. Vismazākā ir tukšā kopa 0, vislielākā - pati kopa K (apzīmēsim to ar 1).
Operācijas: x^y – šķēlums, xUy – apvienojums, x-y – starpība, -x = A-x – papildinājums.
Kopu algebras izteiksmes: (xUy)^z, tā ir = (xUz)^(yUz),
-(xUy) = (-x)^(-y), -(x^y) = (-x)U(-y), utt.

x->y definējam kā –xUy. Tad:
A->(B->A) = -AU(-BUA) = -AUAUB = 1,
-A->(A->B) = -AU(AUB) = -AUAUB = 1,
utt. visām citām klasiskās loģikas aksiomām.
Teorēma. Jebkurām izteiksmēm A, B, izteiksme A->B ir identiski vienāda ar 1, tad un tikai tad, ja identiski A<=B [identiski – tāpēc, ka izteiksmes satur mainīgos].
Teorēma. Izteiksme, kas satur tikai operācijas -, ^, U un ->, ir identiski vienāda ar 1 tad un tikai tad, ja tā ir identiski patiesa Būla algebrā [kas būtībā ir kopas K={k} apakškopu algebra].
Secinājums. Būla algebra atbilst ļoti dabiskai matemātiskai struktūrai. Vai tas liecina par labu loģikas “pareizībai”, vai pret to?

Implikācijas paradoksi

Visvienkāršākās problemātiskās (t.i. apšaubāmās) lietas klasiskās loģikas pamatos ir t.s. **materiālās implikācijas paradoksi.**

Pozitīvais paradokss: patiesa skaitās formula **A->(B->A)**, kur B ir patvaļīgs apgalvojums. T.i. ja A ir patiess, tad A seko arī no apgalvojumiem, kuriem ar A nav nekāda sakara!

Negatīvais paradokss: patiesa skaitās formula **A->(~A->B)**, kur atkal B ir patvaļīgs apgalvojums. T.i. no pretrunas seko jebkurš apgalvojums! Kā mēs to zinām?

Vēl: patiesas skaitās formulas (A->B)v(B->A), Av(A->B), …

Matemātiķiem domātās grāmatās šos paradoksus parasti nemaz nepiemin…

Vēl sliktāk…

… šie paradoksi seko no principiem, kurus neviens nekādās aizdomās netur.

Ja A, B |- C, tad A |- B->C.
Bet, ja ņemam C=A, tad iznāk, ka A |- B->A.
AvB, ~A |- B
A |- AvB.
Bet tad jau iznāk, ka A, ~A |- B.

Tātad, ja mēs gribētu implikācijas paradoksus novērst, tad būtu jāpārskata arī tādi principi, kurus apšaubīt nemaz negribētos…

Kā to dara? Sk. Relevance Logic no Stanford Encyclopedia of Philosophy.

Kā ir īstenībā?

Paradoksālie pieņēmumi A->(B->A), ~A->(A->B) padara mūsu loģisko aparātu vienkāršāku. Jo mēs atsakāmies prātot, kā ir "pareizi": ja 2*2=5, tad Visums izplešas, vai tieši otrādi – saraujas.

Mēs vienkārši postulējam, ka 1->1 = 0->1 = 1->0 = 1, un tā iegūstam ļoti ērtas unificējošas lietas, piemēram:

a) T.s. dedukcijas teorēmu: A, B |- C tad un tikai tad, ja A |- B->C.
b) ~AvB -> (A->B) utml.
c) Atbilstība kopu algebrai utml.

Un neko nezaudējam – izņemot ticību (vienai vienīgi "pareizajai" loģikai)!

Pareizā attieksme

Pareizā attieksme: klasiskā loģika ir ne gluži perfekts cilvēka **izgudrojums (nevis atklājums!)**, kas ietver dažus pilnīgi patvaļīgus vienkāršojumus, bet kas tomēr izcili labi "strādā" (un tas ir labi pārbaudīts eksperimentāls fakts!).

Tāpēc pat matemātiķis var nodzīvot visu mūžu, nekādas loģikas problēmas tā arī nepamanot…

Un tomēr, arī klasiskā loģika ir tikai robusta "spriešanas mašīna" (engine of reasoning – mans termins). Tā ne tik daudz modelē "pareizas" spriešanas likumus, cik piedāvā gatavu konkrētu spriešanas aparātu, kas izcili labi darbojas.

Un - ir iespējamas arī citas loģikas…

Platonisko pasauļu fizika?

Lasot Leo Corry rakstu The Origins of Eternal Truth in Modern Mathematics: Hilbert to Bourbaki and Beyond. *Science in Context* 12 (1998): 137-183…

Ja Eiklīda ģeometrija kā fiziskās telpas teorija izrādījās esam ne gluži precīza, tad kā varēja rasties pārliecība, ka tā ir perfekta (skolas izjūtas…, bet I. Kantam – "sintētiski apriora" struktūra, kas iebūvēta mūsu apziņā)?
Kā varēja rasties pārliecība, ka tikpat perfekta ir Ņūtona mehānika? Arī tā taču vēlāk izrādījās esam neprecīza…
Un kā visas šīs briesmīgās vēstures rezultātā var saglabāties ticība, ka dažas perfektas teorijas tomēr vēl eksistē? Pat kopu teorijas paradoksi…

Daudzi uzskata matemātika par tādu kā **platonisko pasauļu fiziku** (physics of platonist realms – mans termins). Tad kā rodas pārliecība, ka mūsu aksiomas, kas apraksta kādu no šīm pasaulēm, ir "patiesas" (tātad – bezpretrunīgas)? Vai aritmētikas aksiomu attiecības ar (platonisko) naturālo skaitļu pasauli nevarētu būt tādas pat kā Eiklīda ģeometrijai ar fizisko telpu? **T.i. šis aksiomātiskais apraksts, varbūt (vai – pilnīgi noteikti?) ir ne tikai nepilnīgs, bet arī neprecīzs, t.i. daļēji nepareizs?**

Secinājums Nr. 3

Arī “nepareiza” (ne gluži “pareiza”) loģika
var būt ļoti efektīva.

Kā tad var rasties “absolūti pareiza” loģika?
Nevar!

Viss!