



Neumann János
Számítógép-tudományi Társaság
Orvos-biológiai Szakosztály



Orvosi informatika

A XXXIV. Neumann Kollokvium konferencia-kiadványa

Pannon Egyetem, Veszprém,
2021. december 3-4.

Szerkesztők:

Vassányi István, Fogarassyné Vathy Ágnes

Orvosi informatika — XXXIV. Neumann Kollokvium

Szerkesztők: Vassányi István, Fogarassyné Vathy Ágnes

Kiadta a Neumann János Számítógép-tudományi Társaság

Veszprém, 2021.

ISBN 978-963-396-217-6

XXXIV. Neumann Kollokvium

Veszprém, 2021. december 3-4.

© Neumann János Számítógép-tudományi Társaság.

Minden jog fenntartva.

© John von Neumann Computer Society. All Rights Reserved

Beköszöntő

Nagy örömmel rendeztük meg a XXXIV. Neumann Kollokviumot a hagyományoknak megfelelően ezúttal Veszprémben, a Pannon Egyetemen, bár sajnos megint csak tele-konferencia formájában, mivel a személyes részvételről a járványveszély miatt idén is le kellett mondanunk. Az érdeklődés ettől függetlenül nagy volt, ami bizonyítja, hogy a Kollokvium továbbra is a hazai egészségügyi informatikai kutatás-fejlesztés meghatározó eseménye és fóruma.

A Kollokvium alapvető jellegzetessége a sokszínűség. Bár a hazánkban évtizedek óta működő tudományos műhelyek adottak, mégis a témák változatossága évről-évre szembetűnő. Az idei nagy téma az egészségügyi adatmodellezés, adatelemzés volt, ez a téma jellemezte az előadások több mint felét, ez alkotta a konferencia első napjának a programját. Hallhattunk beszámolókat biológiai alaputatásról, régebb óta futó, országos szintű adatelemzési projektek újabb eredményeiről, az új eGYSE szolgáltatásról, az ICD-11-re való átállás informatikai támogatásáról, illetve az igazán forró kérdésről, a Covid járvány modellezéséről is. Ez utóbbi volt a témája a konferencia nyitóeseményeként rendezett kerekasztal-beszélgetésnek is, melyre a terület legismertebb szakértőit hívtuk meg. A beszélgetés végére kibontakozó kép meglepte a laikus hallgatókat azzal, hogy a járvány lefolyása ilyen sokféleképpen és főleg—megfelelő pontosságú bemeneti adatok birtokában—ilyen pontosan leírható és előre is jelezhető.

Örökzöld téma az egészségügyi jelfeldolgozás és diagnosztika, hiszen folyamatosan újabb eszközök, modalitások és diagnosztikai algoritmusok illetve számítási architektúrák jelennek meg, illetve új ötletekkel a meglévő módszerek hatékonysága is javítható, új típusú ellátások alakíthatók ki. A konferencia második napjának jelfeldolgozási és diagnosztikai szekciójában értesültünk egy komoly telemedicinális kísérlet első és igazán biztató eredményeiről, melyek alapján állíthatjuk, hogy az itt kialakított ellátási modellnek hosszabb távon és szélesebb (országos) körben is jövője kell, hogy legyen. A klasszikus egészségügyi jelfeldolgozást egy vérnyomásméréssel kapcsolatos, egy EEG és egy EKG témájú előadás

képviselte, és nagy érdeklődéssel hallgattuk az intenzív osztályos vércukorszint-szabályozás és a képregisztráció továbbfejlesztett módszereiről szóló beszámolókat is. A konferencia utolsó szekciója különféle egészségügyi szolgáltatásokról, webes rendszerekről szólt. Két előadás vette célba a weblapok akadálymentesítését, egy az online oktatás minden egyetemen érezhető hatásait, egy az autizmussal élők támogatását, végül megismertük egy folyamatban lévő, idős emberek revitalizálását célzó európai projekt célkitűzéseit is.

A bemutatott kutatások egy része egyéni PhD projektekhez kapcsolódik, de megjelentek a maguk területén nemzetközileg is elismert kutatócsoportok legfrissebb eredményei is.

A XXXIV. Neumann Kollokvium az NJSzT Orvos-biológiai Szakosztálya, a Pannon Egyetem Műszaki Informatikai Kara és a Veszprémi Akadémiai Bizottság Egészségügyi Informatikai Munkabizottságának szakmai összefogásával valósult meg. Külön szeretnénk megköszönni a meghívott előadóknak a részvételt és a Tudományos Bizottság tagjainak a szakmai előkészítésben nyújtott segítségét.

Veszprém, 2021. december

VASSÁNYI ISTVÁN ÉS FOGARASSYÉ VATHY ÁGNES

a kötet szerkesztői

Tudományos bizottság

A bizottság elnöke:

SURJÁN GYÖRGY, Országos Kórházi Főigazgatóság
NJSzT OBSz elnök

A bizottság tagjai:

BALKÁNYI LÁSZLÓ, Pannon Egyetem

BARI FERENC, Szegedi Tudományegyetem

BERTALAN LÓRÁNT, Semmelweis Egyetem

FOGARASSY NÉ VATHY ÁGNES, Pannon Egyetem

KÓSA ISTVÁN, Szegedi Tudományegyetem/Pannon Egyetem

KOZMANN GYÖRGY, Pannon Egyetem

NAGY ISTVÁN, Gottsegen György Országos Kardiológiai Int.

NYÁRI TIBOR, Szegedi Tudományegyetem

RÁROSI FERENC, Szegedi Tudományegyetem

TOLNAI JÓZSEF, Szegedi Tudományegyetem

VASSÁNYI ISTVÁN, Pannon Egyetem

Konferenciaprogram és tartalomjegyzék

I. szekció: Egészségügyi adatelemzés

Járványmodellezés kerekasztal

Meghívott előadók: *Csikász-Nagy Attila (PPKE), Ferenci Tamás (ÓE),
Szederkényi Gábor (PPKE).*

Moderátor: *Surján György (OKFŐ)*

Az orvosi szöveg, mint adat - esettanulmány: lelevezés	1
<i>Balkányi László</i>	
Új központi elektronikus egészségügyi fejlesztés - eGYSE Gyógyászati segédeszközök elektronikus vényen	6
<i>Bertalan Lóránt</i>	
A Covid-19 Sars-Cov-2 járvány alatti D3-vitamin forgalmának alakulása Magyarországon	13
<i>Batta Balázs, Hercegné Havasi Ildikó, Merth Gabriella</i>	
Hallgatói kérdőívek szabad szöveges válaszainak értékelése szöveg- bányászati módszerekkel	19
<i>Boda Krisztina, Szűcs Mónika, Rárosi Ferenc</i>	
Akut stroke betegutak kockázati tényezői	24
<i>Vassányi István, Kováts Tamás, Surján György, Nagy Zoltán</i>	
Egészségügyi indikációk és ICD-11 diagnóziskódok illesztését támogató alkalmazás	31
<i>Takács Krisztián, Vassányi István, Balkányi László</i>	

II. szekció: Egészségügyi adatmodellezés

ICD-11 kódok validálását végző webszolgáltatás fejlesztése	35
<i>Kukoda Balázs, Vassányi István, Balkányi László</i>	
Graphletek funkcionális eloszlása a C. elegans konnektómban.....	43
<i>Vassy Zsolt, Fenyves Bánk Gábor, Szilágyi Gábor Sándor, Csermely Péter</i>	
A kardiológiai ellátórendszer kiemelkedő befolyással rendelkező intézményei: Az Egyetemi Klinikák.....	51
<i>Vassy Zsolt, Vassányi István, Kósa István</i>	
Magyarországi járványadatok elemzése rendszerelméleti megközelítéssel.....	58
<i>Csutak Balázs, Polcz Péter, Szederkényi Gábor</i>	

Az antraciklinnel kezelt betegek szívelégtelenségi rizikófaktorainak analízise (csak előadás)

Fogarassy György, Fogarassyné Vathy Ágnes

A térbeli készséget fejlesztő virtuális valóság alapú teszteken töltött idők meghatározása64

Guzsvinecz Tibor, Mihálykóné Orbán Éva, Sikné Lányi Cecília, Perge Erika

III. szekció: Egészségügyi jelfeldolgozás és diagnosztika szekció

Első tapasztalatok kardiometabolikus betegek kombinált tréning és táplálkozás monitorozásával71

Máthéné Köteles Éva, Ábrahám Judit, Vágvölgyi Anna, Korom Andrea, Szálka Brigitta, Barnai Mária, Bilicki Vilmos, Lengyel Csaba, Kósa István

Automatizált vérnyomásmérés bal kamrai keringéstámogatással élő betegek esetén77

Nagy Péter, Sax Balázs, Osztheimer István, Jobbágy Ákos

A kamrai repolarizációs heterogenitás QRST integrál alapú osztályozási lehetőségei végtagi I. elvezetési EKG mérések alapján83

Tuboly Gergely, Kozmann György

Improved insulin sensitivity prediction method using sex information of the patients87

Bálint Szabó, Ákos Szlávecz, Béla Paláncz, Geoffrey Chase, Balázs Benyó

Geometric modelling and co-registration based approach for defining multiple patient parameters in medical diagnostics95

Bazsó Sándor, Viola Árpád, Hajdu Angéla, Benyó Balázs

Towards a High-Performance Independent Component Analysis Implementation on GPUs102

Zeyu Wang, György Kozmann, Zoltán Juhász

IV. szekció: Egészségügyi szolgáltatások, webes rendszerek szekció

A covid-19 járvány miatt bevezetett online oktatás hatásai a hallgatók vizsgán elért eredményére108

Rárosi Ferenc, Szűcs Mónika, Boda Krisztina

Idős emberek egészség-megőrzése és vitalizálása informatikai eszközökkel: a Turntable projekt tapasztalatai	113
<i>Vassányi István, Szakonyi Benedek</i>	
InfoBázis—felhasználóbarát online adatbázis tervezése autizmussal élő emberek támogatásához.....	118
<i>Havasi Ágnes, Kancsó Martin, Imre Anna, Kis-Jakab Éva, Sikné Lányi Cecília</i>	
Európai COVID-19 információt tartalmazó WEB oldalak akadálymentességi tesztelése	126
<i>Jinat Ara, Sikné Lányi Cecília</i>	
Weblapokra vonatkozó esélyegyenlőséget fejlesztő ajánlások megfogalmazása a Semmelweis Egyetem weblapjainak akadálymentességi vizsgálata alapján.....	134
<i>Bede-Fazekas Eszter, Sándor Zoltán</i>	

Hallgatói kérdőívek szabad szöveges válaszainak értékelése szövegbányászati módszerekkel

Boda Krisztina, Szűcs Mónika, Rárosi Ferenc
SZTE SZAOK Orvosi Fizikai és Orvosi Informatikai Intézet

boda.krisztina@med.u-szeged.hu,

szucs.monika@med.u-szeged.hu

rarosi.ferenc@med.u-szeged.hu

6720 Szeged Korányi fasor 9.

Összefoglaló: A biostatisztika alapjainak oktatásában rendszeresen alkalmazunk kérdőíveket, melyet a hallgatók önkéntesen és név nélkül töltenek ki. Ezekkel egyrészt adatbázist teremtünk a statisztika oktatásához, másrészt pedig a hallgatók véleményéről is értesülünk. A kérdőívek szabad szöveges válaszait szövegbányászati módszerekkel is feldolgoztuk. Az előadás bemutatja az R szoftver alkalmazási lehetőségeit angol és magyar nyelvű szövegek elemzésére, valamint a „miért akar orvos lenni”, illetve az „egyéb vélemény a biostatisztika oktatásáról” kérdésekre adott válaszok eredményét.

Bevezető

A biostatisztika oktatásában mindig is alkalmaztunk kérdőíveket, melyet a hallgatók saját magukról töltöttek ki nélkül. Az első hetekben kitöltött kérdőívek adatai szolgáltak a tanított módszerekhez adatbázisként (pl. testmagasság és testtömeg közötti korreláció vizsgálata). Másrészt a hallgatók véleményét is igyekeztünk megtudni az oktatásunkról, pl. a szoftver használatáról, vagy az előadás szükségességéről. Mindegyik kérdőívben lehetőséget adtunk szabad szöveges válaszoknak is. Az előadásban ezeknek a szövegeknek szövegbányászati módszerekkel történő feldolgozását és az R szoftver alkalmazását mutatjuk be.

Módszer

A szövegbányászat feladata struktúrálatlan vagy gyengén stuktúrált szöveges állományokból információk nyerése, összefüggések felderítése. Széles körben alkalmazzák pl. Internetes keresésekben, az államigazgatásban, e-kormányzatban, gyógyszerkutatásban, stb. A szövegbányászati módszerekről már magyar nyelven is egyre több irodalom található [1], [2], [3]. Már az orvosi irodalomban is megjelent a kérdőívek szabad szöveges válaszainak ilyen módszerekkel történő feldolgozása [4]. Rengeteg szoftver,

közöttük a statisztikai szoftvercsomagok is tartalmaznak szövegbányászati modulokat. Mi az R nyelvet alkalmaztuk, mivel ez a szoftver szabad hozzáférésű, ezt a szoftvert használjuk az oktatásban is, és ebben találtunk használható programkódokat is [5], [6].

Az általunk használt módszerek rövid összefoglalása

Korpusz. A feldolgozandó dokumentumok összességét *dokumentumgyűjteménynek* vagy *korpusznak* nevezzük. Lehet pl. pdf-file-ok gyűjteménye, de lehet egyetlen szövegfile is, melynek sorai a kérdőívet kitöltők szabad válaszai (esetünkben az utóbbi). Ezt a korpuszt kell egységes, feldolgozható formára hozni, amely letisztított formában tartalmazza a szöveget és esetleges meta-adatokat (a dokumentum egyéb jellemzőit); az eredmény az ún. *szó-dokumentum mátrix* (document-term matrix). A mátrix soraiban a vizsgálandó elemek (kifejezések vagy szavak), oszlopaiban a válaszadók (dokumentumok vagy hallgatók) szerepelnek, a mátrixban annak gyakorisága, hogy az adott kifejezés az adott válaszadónál hányszor fordul elő. Egyes esetekben gyakoriságok helyett csak 0-k és egyesek szerepelnek, vagy a szó fontosságát valahogyan súlyozhatjuk.

Előfeldolgozás.

A szó-dokumentum mátrix előállításához nélkülözhetetlen néhány előfeldolgozás. Ezek a következők lehetnek:

- Speciális karakterek (pl. /, @, |) eltávolítása vagy helyettesítése szóközzel, felesleges szóközök, számok eltávolítása, nagybetűk kisbetűkké alakítása.
- Stopszavak eltávolítása (pl. „and”, „the”, „on”, stb.). Angol nyelven van egy alapszótár, magyarul egyesével be tudjuk írni.
- Stemming (szótövezés): szavak csonkolása, pl. „fishing”, „fisher”, „fished” -> „fish”. Ez a rész különösen veszélyes lehet, hogy hogy álljunk meg a szótövezéssel, az eredmény függ az eljárástól és nyelvtől.

Szó-dokumentum mátrix alapján végzett eljárások

- Leggyakoribb szavak megkeresése
 - Szófelhő készítés (látványos)
 - Szavak közötti kapcsolatok felderítése – mely szavak milyen szavakkal állnak kapcsolatban
 - A korrelációk alapján clusteranalízis, faktoranalízis végezhető
- Egyéb eljárások, lehetőségek: érzelem detektálás (szentiment analízis)*

Eredmények

2020-ban a korábban is alkalmazott változók (neme, kora, testmagassága, testtömege) azt is megkérdeztük a hallgatóktól, hogy miért szeretne orvos

lenni? Angol nyelven 53, magyar nyelven 140-en válaszoltak a kérdésre, természetesen szabad szöveges formában. A kérdőív egy részét mutatja az 1. Tábl., melynek tehát a „Why do you want to be a doctor” kérdésre adott válaszait dolgozzuk fel.

1. sz. táblázat

nick	sex	height	mass	why_do_you_want_to_be_a_doctor	where_to_work	learnt_statist
light	female	160	53	the love of service	hospital	no
AMIR	male	175	85	i love it	hospital	no
578	female	168	63	Work with people but with theoretical background	clinics	yes
Ofir	male	170	61	to help people	clinics	no
Silvia	female	180	70	I like learning about the human body and diseases.	clinics	yes

A szövegbányászati eljárásokat az R (<https://www.r-project.org/>) ‘tm’ package segítségével végeztük. A tisztítás eredményeként előálló szó dokumentum mátrix egy részét mutatja a 2. Tábl. Ebben a részletben minden szó minden embernél csak 1-szer fordul elő, ezért láthatók csak 0-k és 1-esek a mátrixban. Pl. „help” szó előfordul a Ofir, MHG és fm jeliségű hallgatónál. Ha sor-összegeket képezünk, megkapjuk az egyes szavak gyakoriságait. Az oszlop-összegek azt mutatják, egy ember hány szót használt fel. Az első 10 leggyakoribb szót a 3. Tábl. tartalmazza, az összes szóból készült szófelhőt az 1. ábra. Ez már majdnem olyan mátrix, mint ami a statisztikához kell, csak transzponálva.

2. sz. táblázat

	light	AMIR	578	Ofir	Silvia	amirhossein	kilua	MHG	fm
love	1	1	0	0	0	0	0	0	0
service	1	0	0	0	0	0	0	0	0
background	0	0	1	0	0	0	0	0	0
people	0	0	1	1	0	0	0	1	1
theoretical	0	0	1	0	0	0	0	0	0
work	0	0	1	0	0	0	0	0	0
help	0	0	0	1	0	0	0	1	1

3. sz. táblázat

szó	help	people	like	want	others	love	field	service	body	human
gyakoriság	23	18	8	7	6	5	4	3	3	3

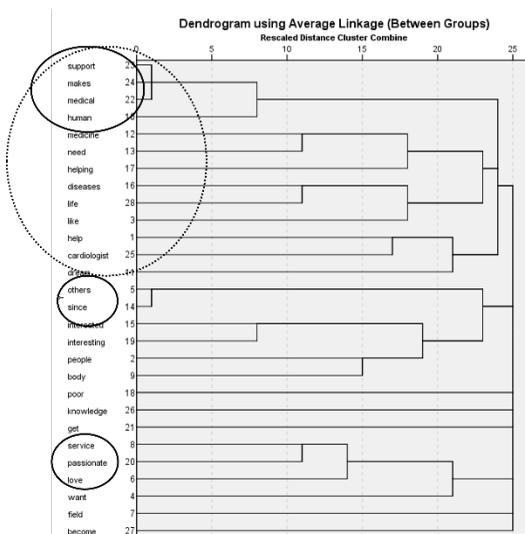
Ha a 2. sz. táblázatot transzponáljuk, akkor a statisztikai feldolgozásokhoz szokásos elrendezést kapjuk: a változók a szavak, a sorok pedig a válaszadók. Így kiszámíthatjuk a szavak közötti korrelációt, amely tehát a két változóban álló két számsor (egyes szavakra adott gyakoriságok) közötti Pearson-korreláció. Pl. a „help” és az „others” szavak közötti korreláció $r=0.42$. Az R programmal azt is megnézhetjük, hogy mely szavak mely más szavakkal állnak kapcsolatban, a korreláció mértékét mi magunk szabjuk meg. (1. ábra)

1. sz. ábra

```
findAssocs(TextDoc_dtm, terms=c("help", "people", "like", "want"),
corlimit = 0.2)
$help
others poor get
0.42 0.23 0.23
$people
helping poor get support
0.28 0.28 0.28 0.28
...
```

A korrelációk alapján ezután pl. cluster analízissel szemléltethetjük ugyanezt (2. ábra), vagy számszerűsíthetjük faktoranalízissel is.

2. sz. ábra.



- [4] D. Nair, K.A.Pivert, A. Baudi IV, C.V. Thakar. „Perceptions of nephrology among medical students and internal medicine residents: a national survey among institutions with nephrology exposure.” in *BMC Nephrology*, 2019, 20:146
- [5] S. Mhatre. „Text Mining and Sentiment Analysis: Analysis with R.” 2020, <https://www.red-gate.com/simple-talk/databases/sql-server/bi-sql-server/text-mining-and-sentiment-analysis-with-r/>
- [6] G. Pósfai. „Szövegbányászati gyakorlat” 2015. [Szövegbányászat gyakorlat Órai segédlet \(studio-pubs-static.s3.amazonaws.com\)](http://studio-pubs-static.s3.amazonaws.com)