



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Informační brožura k edukačnímu semináři

Zpracování, vizualizace a dolování dat z pohledu rozhodovacích procesů

13.-14. září 2017
Telč



Editoři: Martin Komenda, Jakub Gregor



Ústav zdravotnických informací a statistiky České republiky
Institut biostatistiky a analýz Lékařské fakulty Masarykovy univerzity
Společné pracoviště

Ústav zdravotnických informací a statistiky České republiky
Institut biostatistiky a analýz Lékařské fakulty Masarykovy univerzity
Společné pracoviště

Zpracování, vizualizace a dolování dat z pohledu rozhodovacích procesů

13.-14. září 2017
Telč

Editori:
Martin Komenda
Jakub Gregor

Konání semináře a vydání této publikace byly podpořeny
Evropským sociálním fondem, Operační program Zaměstnanost,
projekt „Centrum pro rozvoj technologické platformy registrů
Národního zdravotnického informačního systému,
modernizace vytěžování jejich obsahu a rozšíření jejich informační kapacity:
Rozvoj technologické platformy NZIS“ — CZ.03.4.74/0.0/0.0/15_019/0002748.



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Všechna práva vyhrazena. Žádná část této elektronické knihy nesmí být reprodukována nebo šířena v papírové, elektronické či jiné podobě bez předchozího písemného souhlasu vykonavatele majetkových práv k dílu, kterého je možno kontaktovat na adrese – Nakladatelství Masarykovy univerzity, Žerotínovo náměstí 9, 601 77 Brno.

© 2017 Masarykova univerzita

ISBN 978-80-210-8693-7

ISBN 978-80-210-8692-0 (brož. vaz.)

OBSAH

Úvodní slovo	4
Místo konání	5
Program semináře	6
Dolování dat a projektové řízení: pohled z praxe	7
Jazyk R, dolování dat a interaktivní vizualizace	8
Od nestrukturovaných dat k informacím a znalostem	9
Osvětový seminář o formální kvalitě odborných výstupů	10
Analýza a vizualizace dat	11
Přehled webových portálů pod záštitou IBA LF MU & ÚZIS ČR	25

ÚVODNÍ SLOVO



Milí čtenáři, účastníci tohoto semináře, je pro mne ctí otevřít tento sborník úvodním slovem. V omezeném počtu slov se ale tato možnost zvedá ne zcela lehce. Pokud tedy člověk nechce upadnout do frází, že data jsou nutná, nezbytná a jejich analýza je poselstvím, bez kterého není nic. Jakkoli tomu jistě je, není to sdělení, které by někoho překvapilo nebo někam posunulo. Osobně si myslím, že data jsou sice nutná, ale daleko nutnější jsou informace. A od dat k využitelné informaci bývá hodně dlouhá a pracná cesta, na které je sběr dat samotných jen jedním z mnoha kroků. A právě na cestu od dat k informacím se zaměřuje tato letní škola.

A je toho mnoho, co se učít, neboť dnešní doba daty nešetří. Tzv. svět analýzy „big data“ již vstoupil téměř do všech oborů lidské činnosti. Pod slovem „big“ já osobně nevidím nic příjemného – velmi často jde jen o kvantitu, ze které musí znalý analytik „vydolovat“ podstatné jádro, toto strukturovat do logických vazeb mezi parametry a tyto zviditelnit např. v prezentaci, grafu. Tak, aby lidské oko, bazálně zvyklé na 2D grafy a schopné přijmout 3D zobrazení, ale nic moc více, rychle dostalo konečnou informaci, kterou zpracuje. Může jít o oko manažera, správce letového provozu, vyšetřujícího neurologa či operujícího chirurga. Všechny tyto příklady spojuje dnešní době tak typický nedostatek času, který mají na konečné rozhodnutí. A právě proto je analýza dat pro rozhodovací procesy jednou z nejtěžších disciplín zpracování dat vůbec. Nejde jen o grafický výzor, ale opravdu o sdělení nejpodstatnějších informací na velmi malém prostoru; analytik separuje z primárních dat jakýsi minipříběh, který sděluje konečnému uživateli. Dobře připravit informace pro závažná rozhodnutí (zastavit – nezastavit dopravu?, operovat – neoperovat pacienta?) je umění, které dělá analýzu dat krásnou. A tímto osobně poznaným faktem bych se s úvodním slovem rozloučil a přeji vám všem krásný a přínosný pobyt.

doc. RNDr. Ladislav Dušek, Ph.D.

ředitel ÚZIS ČR & IBA LF MU



Ústav zdravotnické informatiky a statistiky České republiky (ÚZIS ČR) se jakožto organizační složka státu specializuje mimo jiné také na oblast analýzy zdravotnických dat, a to na lokální i mezinárodní úrovni. Samotné téma získávání hodnotných informací z rozsáhlých nestrukturovaných datových souborů je v dnešní době velmi aktuální. I proto byl pod garancí ÚZIS ČR byl připraven odborný edukační seminář zaměřený především na zpracování, vizualizaci a dolování dat, které přináší nové relevantní výstupy pro podporu rozhodování ve zdravotnictví jako takovém. Cílem dvoudenního semináře je poskytnout posluchačům komplexní pohled na proces vytěžování informací z různých datových zdrojů a v praxi upozornit na možná úskalí při řešení konkrétních úkolů. Jednotlivé tematické bloky jsou koncipovány tak, aby přinesly nejen teoretické pozadí pro celkovou orientaci v problematice dolování znalostí, ale také demonstrovaly různé přístupy použité v reálných projektech. V neposlední řadě je pozornost soustředěna na týmovou spolupráci, která vždy vyžaduje správné nastavení komunikačních kanálů při předávání informací mezi jednotlivými odděleními, odborným garantem a zadavatelem.

RNDr. Martin Martin Komenda, Ph.D.

odborný garant semináře, ÚZIS ČR & IBA LF MU

MÍSTO KONÁNÍ

UNIVERZITNÍ CENTRUM TELČ

<http://www.uct.muni.cz/>

Univerzitní centrum Telč (UCT) poskytuje příjemné a komplexní zázemí pro konání seminářů, konferencí a studijních kurzů. Vzdělávací centrum Masarykovy univerzity, které prošlo v letech 2009–2011 rozsáhlou rekonstrukcí, je tak ideálním místem pro konání rozličných vzdělávacích, odborných a společensko-kulturních akcí.

UCT je zasazeno do atraktivního prostředí historického jádra města Telč, které bylo v roce 1992 zapsáno do Seznamu světového kulturního a přírodního dědictví UNESCO. Prolínají se v něm prakticky všechny stavební slohy uplynulého tisíciletí a nalezneme zde řadu pamětihodností a staveb. Dominantou Telče je renesanční zámek, na nějž navazuje unikátní architektonický komplex nacházející se zejména na náměstí Zachariáše z Hradce a okolí.

Vznik města samotného se dle pověsti datuje do přelomu 11. a 12. století, nicméně první dochovaná zmínka pochází až z roku 1315. Největ-

ší rozvoj města je spojen s rodem Vítkovců od poloviny 14. století a se Zachariášem z Hradce v polovině 16. století. V následujících stoletích je město postupně ve vlastnictví rodů Slavatů, Lichtensteinů–Kastelkornů a Podstatkých–Lichtensteinů. Počátek 19. století znamená nástup industrializace města, významným milníkem je stavba železnice spojující Telč s Jihlavou, Slavonicemi a rakouským Schwarzenau (rok 1898).

Navzdory průmyslovému rozvoji, dvěma světovým válkám a 40 rokům totalitního režimu si Telč zachovala svůj historický ráz až do dnešních dnů a společně s dalšími městy v okolí (Slavonice, Jindřichův Hradec a další) a překrásnou přírodou je jedním z nejatraktivnějších turistických cílů nejen v rámci Kraje Vysočina, ale i celé České republiky.

Zpracováno s využitím informací z webových stránek města Telč (www.telc.eu) a Univerzitního centra Telč (www.uct.muni.cz)



Univerzitní centrum Telč.
Zdroj: www.uct.muni.cz

PROGRAM SEMINÁŘE

13. září 2017
10:00–11:00 Ubytování
11:00–13:00 Dolování dat a projektové řízení: pohled z praxe <i>Interaktivní blok</i>
13:00–14:00 Oběd
14:00–15:15 Jazyk R, dolování dat a interaktivní vizualizace <i>Přednáškový blok</i>
15:15–15:30 Přestávka a občerstvení
15:30–19:00 Od nestrukturovaných dat k informacím a znalostem I <i>Workshop</i>
19:00 Večeře
14. září 2017
07:30–08:30 Snídaně
08:30–09:45 Osvětový seminář o formální kvalitě odborných výstupů <i>Edukační seminář</i>
09:45–10:00 Přestávka a občerstvení
10:00–11:30 Analýza a vizualizace dat <i>Přednáškový blok</i>
11:30–12:15 Oběd
12:15–17:00 Od nestrukturovaných dat k informacím a znalostem II <i>Workshop</i>
17:00 Večeře

DOLOVÁNÍ DAT A PROJEKTOVÉ ŘÍZENÍ: POHLED Z PRAXE

13. září 2017, 11:00–13:00

Interaktivní blok

GARANCE

P. Brabec, M. Komenda & Webstudio IBA LF MU & ÚZIS ČR

CÍL

- Představit životní cyklus projektu a jednotlivé pozice, které se na řešení projektu v rámci celého týmu podílejí
- Na praktickém příkladu nastínit základy projektového řízení

TÉMATA

- Týmová práce jako základ úspěchu
- Představení rolí v týmu:
projektový manažer, odborný garant, systémový analytik, grafik, kodér, databázový expert, analytik, programátor/vývojář, redaktor
- Jak efektivně řešit reálný projekt – pohled z praxe

ČASOVÝ PLÁN

- 11:00–11:45 Projektový tým
- 12:00–13:00 Jak efektivně řešit projekt

JAZYK R, DOLOVÁNÍ DAT A INTERAKTIVNÍ VIZUALIZACE

13. září 2017, 14:00–15:15

Přednáškový blok

GARANCE

M. Karolyi, O. Klempíř, J. Ščavnický, M. Komenda

CÍL

- Formou prezentací s reálnými výstupy představit možnosti jazyka R z pohledu online interaktivních datových vizualizací

TÉMATA

- Programovací jazyk R jako nástroj pro vytěžování informací a následnou vizualizaci dat
- Platforma OpenCPU
- Interaktivní vizualizace dat na webu a možnosti nastavení uživatelský pohledů na dostupná data

ČASOVÝ PLÁN

- 14:00–14:25 MEDCIN – Medical Curriculum Innovations
- 14:25–14:50 NeuroEDA – Interactive Web Tool for Neuroinformatics Data Analysis
- 14:50–15:15 CCCN – Comprehensive Cancer Care Network pilot model

OD NESTRUKTUROVANÝCH DAT K INFORMACÍM A ZNALOSTEM

13. září 2017, 15:30–19:00; 14. září 2017, 12:15–17:00

Workshop

GARANCE

J. Bilík, J. Jarkovský, M. Karolyi, M. Komenda, T. Pavlík, D. Schwarz, R. Vyškovský

CÍL

- Ukázat význam a důležitost jednotlivých kroků metodiky CRISP-DM při dolování znalostí z nestrukturovaných a složitě organizovaných dat: porozumění doméně, pochopení struktury dat, příprava a zpracování dat, modelování, evaluace, nasazení v praxi

KONCEPT WORKSHOPU

- Samostatné týmy (2 až 3 studenti), které řeší prostřednictvím ověřené metodiky CRISP-DM pod vedením mentora zadaný úkol
- Součástí práce jsou tzv. milníky, kde si studenti společně s mentorem zhodnotí průběh práce v souladu s jednotlivými etapami metodiky
- Na závěr workshopu proběhne prezentace vybraných týmu a diskuze nad průběhem řešení úkolů

ČASOVÝ PLÁN

Den 1

- 15:30–15:45 Představení konceptu a cílů
- 15:45–16:15 Porozumění doméně
- 16:15–16:30 Milník 1
- 16:30–17:00 Pochopení struktury dat
- 17:00–17:15 Milník 2
- 17:30–18:30 Příprava a zpracování dat
- 18:30–18:45 Milník 3

Den 2

- 12:15–12:30 Souhrn předchozího dne
- 12:30–13:30 Modelování
- 13:30–13:45 Milník 4
- 13:45–14:15 Evaluace & nasazení v praxi
- 14:15–14:30 Milník 5 & 6
- 14:45–17:00 Prezentace týmů

14. září 2017, 8:30–9:45

Edukační seminář

GARANCE

J. Gregor, M. Komenda

CÍL

- Seznámit posluchače se základními jazykovými a typografickými pravidly, které zásadním způsobem přispějí ke zvýšení kvality připravovaných výstupů
- Připravit posluchače na budoucí spolupráci při přípravě podkladů pro odborné publikace

TÉMATA

- Jazykový a typografický koutek – jak zajistit formální a jazykovou kvalitu našich výstupů: přehled nejčastějších jazykových a typografických chyb
- Jednou budeme publikovat – jak připravit podklady pro odborný článek: účel vědeckých publikací, výběr časopisu, databáze, impakt faktor, příprava výstupů (grafy, obrázky)
- Predátorské časopisy a konference – jak ze sebe neudělat hlupáka: co to je, jak je poznáme, jak nenaletět a nevyhodit 10⁴ CZK oknem

ANALÝZA A VIZUALIZACE DAT

14. září 2017, 10:00–11:30

Přednáškový blok

GARANCE

M. Komenda & hosté

CÍL

- Formou krátkých přednáškových sdělení představit dosažené analytické výsledky v oblasti zpracování medicínských dat

ČASOVÝ PLÁN

- 10:00–10:15: Barbora Bučková – Image classification using machine learning
- 10:15–10:30: Ondřej Klempíř – Analýza aktivity jednotlivých neuronů u lidských bazálních ganglií metodami strojového učení
- 10:30–10:45: Lubomír Štěpánek – Webová aplikace poskytující podporu správného výběru a realizace statistické metody
- 10:45–11:00: Oliver Velich – Dátami k lepším exponátom

IMAGE CLASSIFICATION USING MACHINE LEARNING

Barbora Bučková

Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University

Objective

The purpose of this work was to identify the weather on the pictures obtained from the traffic cameras, using machine learning methods and R software. This could serve to automatic reporting of bad weather conditions on the road or adjusting the regime of automatically controlled cars.

Methods

The original dataset consisted of total 9,089 colour images divided into 4 weather categories, namely normal (2,666), fog (2,016), rain (2,042) and snow (2,365). The original size of the majority of the images was 320 × 180 pixels which translates to 5–8 kB in jpg format. However, after storing just one image into R memory with all the parameters, it became about 1,5 MB. Consequently, it was not possible to process all images at once. Considering given problems, Mogrify program, which is easily accessible via command line, was used for pre-processing. The dimension of the images was set to a third of the original size, and only the upper halves of the images were stored. After all the transformations, the images were still able to capture the appropriate weather information. The final size of the images was 107 x 30. In order to perform the analysis, images were

converted from RGB model to HSV, from which only the S parameter was extracted and thus prepared data were stored in arff. format.

Each of the four categories of adjusted images was divided into three sections in an approximate 6:2:2 ratio that, when combined, respectively formed the training, validation and test dataset that kept the balanced proportion of weather groups. The h2o package [1] was used for machine learning processes.

Neural networks

Image classification is one of the main problems neural networks (NN) are used for. The purpose of this work was to explore the impact of the architecture on the classification performance. Eight different architectures were used (shallow and deep NN). The rudimentary design consisted of 3,210 input neurons (one for every pixel of the image), four neurons in the inner layer and four output neurons. 10-fold cross-validation was performed. After determining the performance of this simplest design (94.5 % accuracy), neurons were added into the inner layer, as well as some new layers (Table 1).

The performance of the network started to stagnate after having 32 neurons in the (single) inner layer (confirmed by McNemar’s chi-square test). As a result, the final NN model had 3,210 input neurons, 32 neurons in the inner layer, and 4 output neurons. The accuracy was 98.0 %. Investigation of the importance of the input neurons (for learning) showed that the network uses all of the pixels equally, in other words, it did not prefer any particular part of the image (Figure 1). Pixels with the biggest importance were situated in the centre of the image, which is in consideration to the original image the upper quarter. The most of the misclassified images were in the fog category, which was misinterpreted as snow.

Random forests

Random forests is the method sometimes used for image classification. Their main disadvantages are problems while dealing with the unweighted groups. Because this matter was taken care of at the beginning of the analysis (when dividing images into groups), use of this method was appropriate. There are fewer parameters to tune, the main task was to decide upon a number of trees that will be constructed and the maximal depth of the tree. 10-fold cross-validation was performed.

Table 1. Overall accuracy of all neural network models constructed

ID	Number of neurons in inner layers				Overall accuracy	McNemar’s chi-square test (p-value)*	p-values adjusted by Bonferroni correction*
	1.	2.	3.	4.			
1.	4				94.50 %	0.00000	0.00000
2.	16				96.17 %	0.00001	0.00005
3.	32				98.00 %	-	-
4.	1,024				97.94 %	1.00000	0.08344
5.	8	8			97.17 %	0.01192	0.23723
6.	16	16			97.11 %	0.03389	0.01338
7.	32	16			97.00 %	0.00191	0.00000
8.	8	8	8	8	82.83 %	0.00000	0.00000

*the McNemar’s test between the best model (3.) and the model in question

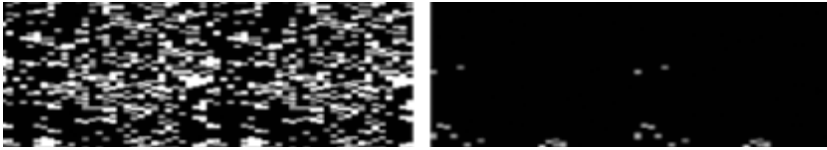


Figure 1. Pixels with scale importance > 0.5 in the Neural network (left) and Random forest (right)

The simplest model was constructed using automatic parameters to give an idea about the basic performance. The number of trees was set on 50, maximum depth on 20 and the number of pixels entering the classification in each tree was set on the third (1,070). The overall accuracy of this model was 97.2 %, similar to neural networks. Interestingly, the distribution of mistake, in this case, was much more uniform between the groups. The biggest problem of neural networks was to classify fog correctly (the most common mistake was labelling it as snow), whereas in this case, the error seemed distributed more equally (not tested).

Again, multiple models were constructed (Table 2). Their performance did not differ significantly. As the first model had the best overall accuracy, it was decided to explore its results. 14 out of 50 images that were wrongly classified by the first random forest model, were the same images as those mislabelled by the best neural net. Moreover, 9 out of this images were misplaced into the same category. Forest uses fewer variables (partly because it was told so) and fewer pixels are significant. The bottom part of the image, which corresponds to the centre part of the original image, is more important for classification.

Comparison of classifiers

In this project, multiple models were constructed in order to classify images by weather. Neural net with 32 neurons in the inner layer and random forest with 50 decision trees (first one) were picked as the best in their category. In order to compare them, McNemar's chi-square test was conducted (Table 3).

Conclusion

The goal of this project was to design and evaluate the performance of several models for image classification based on the weather into four categories: normal, rain, snow and fog. Neural networks and random forest were used.

One of the biggest challenges was to decide upon technique to be used to pre-process the data. Because there were problems concerning editing images when using R software, the dimensionality was reduced and the images cropped using Mogrify tool via command line. Images were converted into HSV model, from which the S parameter was extracted and they were transformed into vectors. Dataset was divided into training, validation and test subsets in 6 : 2 : 2 ratio.

Four shallow and four deep neural networks were constructed. For this particular task shallow neural nets seemed to be more appropriate. Experimentation was done purely on the architecture of the networks, other parameters were not changed after tuning. After comparing overall accuracy between the models, shallow neural net with 32 neurons in the inner layer was the most successful. Importance of variables indicated that neural net did not prefer any part of the image and that the most important pixels for classification were uniformly distributed in the middle of the picture. The biggest confusion in labelling the images was observed in the fog class (a lot of times mislabelled as snow) and the rain class (also mislabelled as snow).

In random forests, experimentation was done on the number of trees, the number of variables entering the classification, and the maximum depth of the tree. Four models were constructed that despite different parameters had very similar overall accuracy. The best model had 50 trees, max-

Table 2. Overall accuracy of all random forest models constructed

ID	No. of trees	No. of predictors	Max. depth	Overall accuracy	McNemar's chi-square test (p-value)*
1.	50	1,070	20	97.22 %	-
2.	200	1,070	50	97.06 %	0.248213
3.	50	2,000	50	96.67 %	0.164915
4.	50	535	50	96.78 %	0.256145

*5 % level of significance

Table 3. Comparison of the best neural network and the best random forest

		Random forest		McNemar's chi-square test with continuity correction
		correct	mistake	
Neural net	correct	1,728	36	p-value* = 0.08783
	mistake	22	14	

* 5% level of significance

imum depth of 20 and 1070 predictors. In contrast with neural networks, the scale importance of pixels was lower in decision trees. The most important variables were uniformly distributed at the bottom of the image.

When comparing the best neural net and the best random forest, the overall accuracy did not differ significantly (level of significance 5 %). An overlap between wrongly classified images was 14 (neural net mistakes: 36 / 1,800; random forest mistakes: 50 / 1,800). Consequently, it is not possible to mark one method as superior. It might be argued that more experimentation with the parameters of the neural network might lead to higher accuracy, but there is no way to tell in this state of research. To summarize, both methods are robust enough to reach more than 95 % accuracy and might be considered suitable for solving the initial problem.

Literature

1. KRALJEVIC, Tom. h2o package | R Documentation [online]. [Accessed 21 June 2017]. Available from: <https://www.rdocumentation.org/packages/h2o/versions/3.10.4.6>.

ANALÝZA AKTIVITY JEDNOTLIVÝCH NEURONŮ U LIDSKÝCH BAZÁLNÍCH GANGLIÍ METODAMI STROJOVÉHO UČENÍ

Ondřej Klempíř¹, Radim Krupička¹, Robert Jech²

¹ Katedra biomedicínské informatiky, Fakulta biomedicínského inženýrství, České vysoké učení technické v Praze

² Neurologická klinika a Centrum klinických neurověd, 1. lékařská fakulta Univerzity Karlovy a Všeobecné fakultní nemocnice v Praze

Úvod

Světová zdravotnická organizace zaznamenává zvyšující se výskyt neurologických onemocnění. Více než 50 milionů Evropanů trpí neurologickou nemocí, což vyžaduje náklady téměř 400 miliard euro [1]. Současně také roste počet dat získaných při studiu (pato)fyziologických procesů mozku. Aplikací analytických výpočetních metod jsou z dat získávány nové poznatky o funkci nervového

systemu či např. mechanismech neurologických onemocnění. Ukazuje se, že počítačové učení a predikce dokáží u některých nemocí na základě extrahovaných příznaků a biomarkerů rozlišit zdravou normu od patologie přesněji než expert. Dále např. s úspěchem predikovat odpověď na léčbu.

Ve studii se zaměřujeme na neurologická onemocnění, dystonie. Dystonie je heterogenní skupina syndromů, které se projevují křečovitým mimovolným pohybem jedné nebo více částí těla, přičemž etiologie mohou být různé [2]. Jednou z metod, která se pro léčbu symptomů pokročilých forem této nemoci používá, je hluboká mozková stimulace (DBS), jež se implantuje během neurochirurgické operace do specifických jader v mozku (bazálních ganglií). Při proudové stimulaci dochází k potlačení patologické komunikace v subkortikálních neuronových sítích, které odpovídají za motorické funkce. Data, která byla naměřena na Neurologické klinice 1. LF a VFN UK, jsou v světovém kontextu unikátní z důvodu vyšetření pacientů za různých podmínek, které dříve měřeny nebyly, např. při experimentech s nastavením DBS globus pallidus interna (GPi). Datové soubory byly získány elektrofyziologickou vyšetřovací technikou, která snímá na úrovni aktivity jednotlivých nervových buněk. Cílem probíhajícího výzkumu je definovat a realizovat řetězec zpracování mikroelektrodových záznamů (MER), a dále s pomocí metod strojového učení nalézt biomarker, který bude prediktorem úspěšnosti léčby DBS GPi.

Metody

Mikroelektrodové záznamy se používají pro diagnostiku správného umístění DBS elektrody. Technologicky jde o sadu mikroelektrod (průměr hrotu měří okolo 5 μm), které prochází skrz mozek a nahrávají aktivitu (tzv. pálení) okolních několika neuronů. Dosažení optimální pozice DBS je klíčové z hlediska terapeutického efektu. V současné době je klinický efekt posuzován primárně neurologem, a to vizuálně či akusticky, na základě specifického projevu akčních potenciálů neuronů.

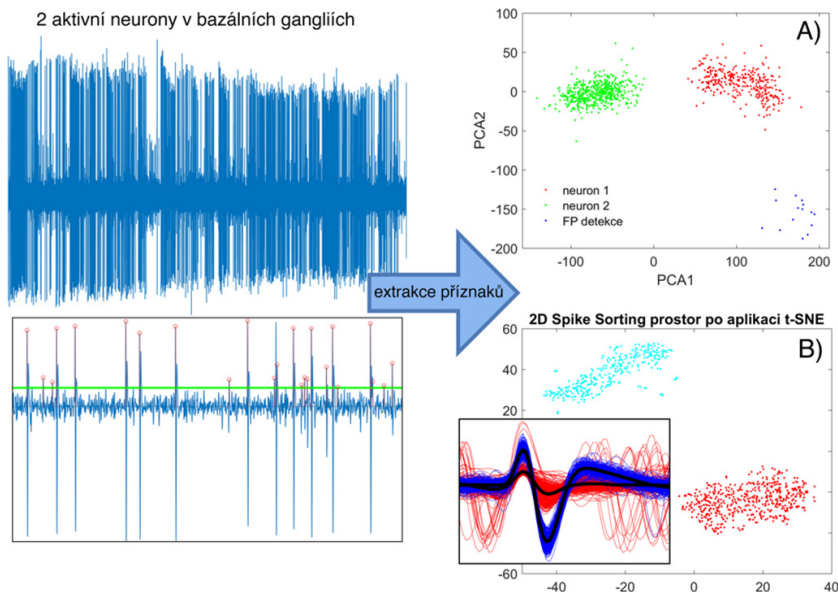
Stěžejní úlohou následující po základním předzpracování (EDA, filtraci či detekci artefaktů) je automatický spike sorting (třídění spiků). Neboli, automaticky v microEEG detekovat namodulované projevy neuronů (spikes) v okolí mikroelektrody a roztřídit je s ohledem na individuální charakteristiky včetně extrakce dalších užitečných informací o jednotlivých neuronech.

Vyvinutá spike sorting procedura sestává z 5 fází:

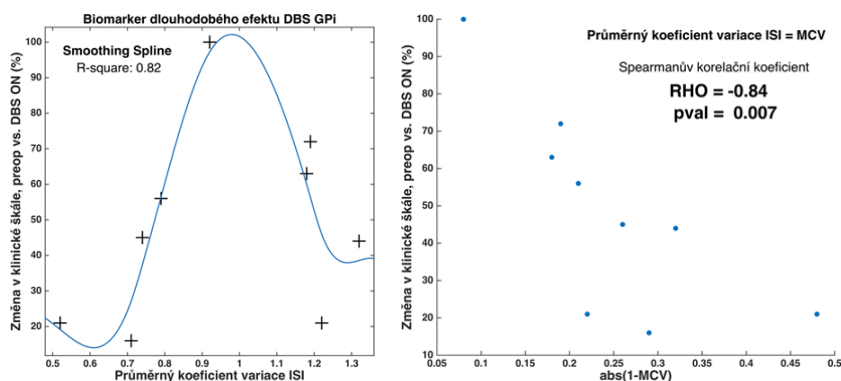
1. stanovení prahu pro odlišení aktivity od šumu pozadí (pravidlo 3σ),
2. extrakce mnohorozměrných tvarově založených příznaků pro detekce z bodu 1,
3. redukce dimensionality pomocí PCA a t-sne do 2D/3D prostoru,
4. shluková analýza K-means pro rozlišení jednotlivých neuronů,
5. výpočet odvozených charakteristik (např. průměrná frekvence pálení/s).

Metoda t-sne zachycuje v porovnání s PCA i nelineární strukturu v datech [3]. Na konkrétním experimentu (obrázek 1A) lze vidět, že PCA vede k rozpoznání 3 neuronů, kdy jeden shluk představuje falešnou detekci, která nesouvisí se skutečným charakterem a musel by být analytikem poloautomaticky vyřazen. T-sne může fungovat lépe při úplné automatizaci, na obrázku 1B s chybou 1.6 %. T-sne přiřadí málo četné hodnoty do bližšího shluku za cenu poměrně malé chyby odhadu výstupních parametrů skutečných neuronů. Ukázalo se také, že t-sne vede k neproblematickému automatickému určení počtu shluků.

Pro 9 pilotních, náhodně vybraných pacientů ve studii, s různorodou etiologií dystonie byla provedena intraoperační analýza jednotkové aktivity a další expertní vyšetření dle standardizovaných neurologických škál. Byly vybrány signály, které měly zjevnou maximální neuronovou aktivitu a nejednalo se o artefakty. Byl proveden spike sorting a výpočet koeficientu variability pálení Inter-Spike-Intervalů (ISI). Z nich byl odvozen parametr MCV = průměr variačních koeficientů ISI pro jednotlivé detekované neurony pacienta.



Obrázek 1. Metodika detekce jednotlivých neuronů v microEEG záznamů bazálních ganglií



Obrázek 2. Graf ukazuje vztah, že s rostoucí vzdáleností od ideálního poměru variability a průměru ISI, klesá klinické pooperační zlepšení. Výsledek vyšel s korelačním koeficientem $RHO = -0.84$

Výsledky

Na základě odvozené znalosti ze spline regrese (obrázek 2, vlevo) byla provedena transformace vypočteného parametru a definován biomarker (obrázek 2, vpravo). Nález neodporuje hypotéze, že úspěšný dlouhodobý efekt DBS souvisí s umístěním stimulační elektrody do okolí neuronů pálících s $CV = 1$. Hodnoty do $CV = 1$ vykazovaly rostoucí trend a od $CV = 1$ trend klesající ($CV = 1 ==$ hranice). Maximum klinického efektu je dosaženo, když směrodatná odchylka a střední hodnota ISI jsou porovnatelné.

Shrnutí

Metodami učení bez učitele byla analyzována MER data pacientů s dystonií. Pilotní analýza nabádá k důkladnému testování hypotéz o parametrech variability pálení jednotlivých neuronů v bazálních gangliích pro všechny ve studii dostupné pacienty. Data nevyvrací hypotézu o vztahu, že s rostoucí vzdáleností od ideálního poměru variability a průměru ISI, klesá klinické pooperační zlepšení. Při analýze zatím nebyly dostupné operační protokoly s přesným umístěním pozic nahrávání MER.

Poděkování

Studie byla podpořena projektem SGS17/114/OHK4/1T/17 Zpracování a analýza heterogenních neuroinformatických dat na KBI FBMI ČVUT v Praze.

Literatura

1. PTÁČEK, Radek, BARTŮNĚK, Petr et al. Etické problémy medicíny na prahu 21. století. Praha: Grada Publishing a.s., 2014. ISBN: 978-80-247-5471-0.
2. JECH, Robert. Hluboká mozková stimulace u dystonií. Neurologie pro praxi. 2013, 14(5), 232-236.
3. DIMITRIADIS, George, NETO, Joana, KAMPF, Adam. T-SNE visualization of large-scale neural recording. BioRxiv. 2016. DOI 10.1101/087395.

WEBOVÁ APLIKACE POSKYTUJÍCÍ PODPORU SPRÁVNÉHO VÝBĚRU A REALIZACE STATISTICKÉ METODY

Lubomír Štěpánek

*Katedra biomedicínské informatiky, Fakulta biomedicínského inženýrství,
České vysoké učení technické v Praze*

Úvod

Problematika výběru korektní statistické metody vhodné pro analýzu dat výzkumníka je stále kruciólní součástí přípravy a realizace vědecké práce; navíc správnost volby významně ovlivňuje to, zda bude nakonec publikace redakčním řízením přijata, či nepřijata k otisknutí do vědeckého časopisu. Ve snaze zjednodušit celý proces byla navržena a implementována webová aplikace, jež podává pomocnou ruku badateli, který čelí otázce výběru a vypracování statistického modelu ve své analýze. Grafické rozhraní aplikace je založeno na použití vývojového diagramu, který badatele naviguje k výběru a použití vhodné statistické metody.

Nástrojů pro statistickou analýzu dat je na trhu dostupná celá řada [1]. Jde buďto „klikací“ software typu IBM SPSS [2], Dell Statistica [3], SAS [4], Minitab [5] či do jisté míry i MS Excel [6] a mnohé další, anebo o pokročilejší nástroje umožňující vlastní skriptování, programování a řešení rozsáhlých a komplexních úloh – například prostředí Matlab [7] a Stata [8] či jazyk R se svými základními knihovnami [9] a jazyk Python s knihovnami SciPy a NumPy [10], případně mnohé další.

Výhodou „klikacích“ nástrojů je jejich uživatelská přívětivost, nevýhodou pak relativní uzavřenost daná konečným počtem nabízených metod a postupů, a tedy nemožnost řešit komplexnější a různé individuální úlohy. Část nástrojů je vyvíjena komerčně; zpoplatnění licencí může být limitací pro širší použití, zejména mezi uživateli-jednotlivci, kam patří i vysokoškolští studenti.

Snaha o integraci nějaké konkrétnější podoby podpory správného výběru vhodné statistické metody, která se nejlépe hodí pro badatelova data a hypotézy, je naopak relativně vzácná. Produkty obvykle pouze nabízejí manuál k ovládní uživatelského rozhraní – ten však sám o sobě může poskytovat kvalitní doménové informace přímo o statistických a analytických metodách a příkladech jejich použití.

Cílem autorova snažení je poskytnout dobře dostupný, nekomerční nástroj pro co možná nejsnazší výběr a následnou aplikaci statistické metody, která je vhodná pro výzkumný záměr a hypotézy badatele. Cestou k takovému nástroji je dle autora vývoj aplikace dostupné online a nevyžadující instalaci. Repertoárem implementovaných a nabízených metod je aplikace zaměřena na oblast biomedicíny. Práce s aplikací by měla být intuitivní a neměla by vyžadovat nutnost pročitání dlouhých manuálů, dodávaných ke komerčním statistickým balíkům. Těmto ideám se snaží aktuální verze vyvíjené aplikace co nejvíce přiblížit.

Materiál a metody

Aplikace je online a je zdarma a volně dostupná. Jádro aplikace je napsáno v jazyce R. Jazyk R je „free-as-in-beer“ a „free-as-in-speech“, tj. je zdarma ke stažení a je možné svobodně upravovat jeho zdrojový kód. R je jednak programovací jazyk, jednak prostředí pro statistické výpočty a grafiku a je široce využíváno mezi statistiky, ekonometrií či biologií. Užití jazyka R zajišťuje reliabilitu výpočtů prováděných zmíněnou aplikací.

Offline napsaný kód v jazyce R byl s použitím balíčku Shiny nahrán online na server dedikovaný pro výpočty v R. Balíček Shiny je také napsán v R a poskytuje online framework pro skripty v jazyce R. Aplikace je dostupná na R-serveru 1. lékařské fakulty Univerzity Karlovy na adrese http://shiny.statest.cz:3838/statisticke_nastroje/.

Aplikace používá češtinu jako jazyk volby, neboť v tuto chvíli je cílíme na česky mluvící výzkumníky a studenty.

Vývojový diagram integrovaný do aplikace pomáhá učinit správná statistická rozhodnutí v závislosti na použitých datech a zamýšlených statistických hypotézách; design diagramu vychází jednak ze zásad „dobré praxe“ problematiky, jednak z autorových zkušeností v této oblasti; korektnost závěrů vývojového diagramu lze porovnat s doporučeními, která pro daný typ dat a hypotéz navrhuje např. [11].

Statistické metody, které jsou implementovány, pokrývají jednak parametrické testy jako t-testy, F-testy nebo rodinu ANOVA a další; jednak neparametrické metody jako Wilcoxonovy testy, Kruskal-Wallisův test nebo Friedmanův test a další. Metody jsou přímo linkovány z koncových bodů diagramu k záložkám, které pak výpočet dané metody na vložených datech provádí.

Výsledky

Klíčovou komponentou aplikace je zmíněný vývojový diagram, viz obrázek 1. Umožňuje výzkumníkovi, který již nasbíral data a formuloval výzkumnou hypotézu, vybrat pokud možno nejvíce odpovídající statistickou metodu, a to tak, že výzkumník projde diagramem v naznačeném směru (shora dolů) na základě odpovědi na pokládané otázky. Ve chvíli, kdy se výzkumník dostane postupným odpovídáním na otázky v diagramu až k jeho koncovému bodu, je zde připraven odkaz na záložku, která vede k dané statistické metodě umožňující provést celou analýzu.

Jsou-li do aplikace nahrána online adekvátní data, analýza proběhne a uživateli je poté nabídnuta možnost stáhnout offline výstupy analýzy – tabulky se sumáři, diagramy s infografikou.

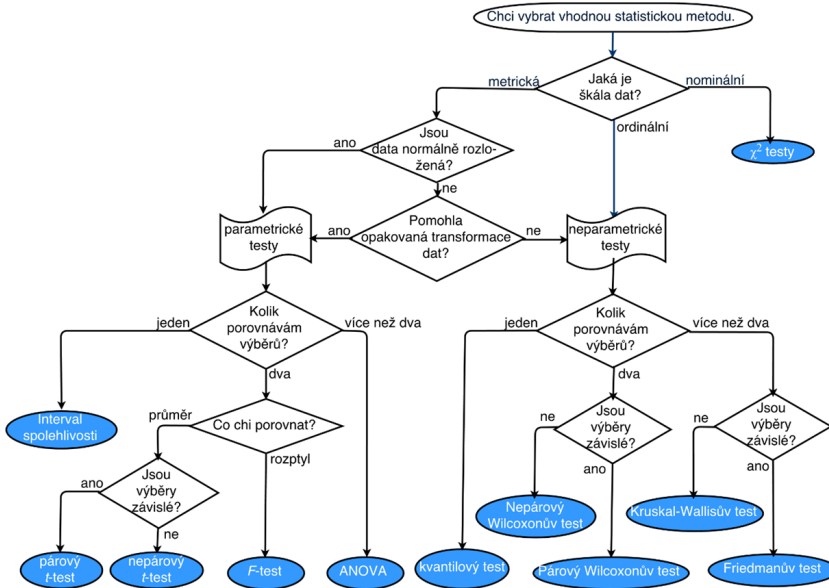
Hlavní rozhodovací uzly vývojového diagramu vyžadují zodpovězení otázek týkajících se typu a struktury nasbíraných dat, empirického rozložení dat, počtu porovnávaných výběrů a toho, zda by analýza měla porovnávat střední hodnoty nebo míry variability dat.

Uživatelské rozhraní aplikace odpovídá současným nárokům na webové utility. Roli back-endu hraje v aplikaci již zmíněné R, roli front-endu pak R-kový balíček Shiny, dále HTML, CSS a javascript, díky kterým je aplikace uživatelsky komfortnější (obrázek 2). Posledně jmenovaný umožňuje mít v aplikaci busy indikátor (symbol přesýpacích hodin při zaneprázdnění aplikace) a další funkcionality.

Diskuze

Přínos aplikace je momentálně předpokládán zejména pro uživatele z řad bakalářských a magisterských studentů různých oborů (bio)medicíny, kteří hledají relativně jednoduchý, přesto použitelný nástroj pro vypracování statistické analýzy do závěrečné práce. Výhodou pro ně může být, že je aplikace zcela volně přístupná, není omezena žádnou licencí. Dalším plusem může být relativně velký prostor pro rychlé doplnění (doimplementování) chybějící, nicméně žádané metody.

Obrázek 1.
Vývojový diagram



Statistické nástroje

Výběr metody Uplodnění dat Testování normality t-test F-test Wilcoxonův test Friedmanův test Kruskal-Walli

Alternativní hypotéza H_A

oba soubory mají různý průměr
průměrné

Typ t-testu

dvouvýběrový

párový

Zobrazit originální výstup z R?

Výsledky F-testu

parametr	hodnota
t-statistika	-5.889
počet stupňů volnosti	36.519
p-hodnota	< 0.00001

Originální výstup z R

```

Welch Two Sample t-test

data: my_data()[, 1] and my_data()[, 2]
t = -5.8889, df = 36.519, p-value = 9.298e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.797949 -2.828521
sample estimates:
mean of x mean of y
10.47129 14.78453
                    
```

Diagram

ženy muži

Obrázek 2.
Uživatelské rozhraní aplikace

Příkladem uveďme dva klasické problémy, se kterými studenti „bojují“ – (i) zpracovává-li student analýzu pouze s využitím MS Excelu® (když mu např. jiný nástroj není z nějakého důvodu dostupný), brzy zjistí, že v MS Excelu® není možné vykreslit krabicový diagram. Mnohdy (ii) rovněž studenti v rámci analýzy kategoriálních dat potřebují analyzovat pomocí testu hypotézy (zde χ^2 -testu nebo Fisherova přesného testu) kontingenční tabulku o rozměrech daleko větších než pouze 2×2 pole, jak je nabízí tradiční komerční statistické balíčky. Vznikl tak i mýtus, že Fisherův přesný test existuje pouze pro čtyřpólní tabulku 2×2 a pro větší není vůbec definován. Oběma nesnázím se lze s použitím předvedené aplikace celkem snadno vyhnout.

Projekt je stále ve vývoji a může být interaktivně updatován v návaznosti na zpětnou vazbu uživatelů. Do budoucna je v plánu implementovat (semi)automatické doporučení statistické metody už na bázi charakteru dat, které uživatel aplikaci nabídne a eventuálně (před)připraví.

Závěr

Kvalita statistického aparátu použitého v odborném sdělení má bezesporu vliv na celkový dojem z publikace. Prvním krokem každé statistické analýzy je volba správného a odpovídajícího statistického modelu. Tento krok je však trikový a mnohdy netriviální i pro zkušeného badatele. Uži-

vatelsky přátelská aplikace, která byla vyvinuta, se snaží tento krok zjednodušit. Propojení mezi doporučením konkrétní statistické metody, vyplývající z vývojového diagramu, a záložkou, která analýzu danou metodou provede, může zlepšit uživatelský komfort.

Literatura

1. ALI, Zulfiqar a SBala BHASKAR. Basic statistical tools in research and data analysis. Indian Journal of Anaesthesia [online]. 2016, 60(9), 662.
2. IBM, Corp. IBM SPSS Statistics for Windows, Version 22.0. B.m.: Armonk, NY: IBM Corp. 2013.
3. DELL SOFTWARE. Statistica, Version 13.0. 2015.
4. SAS INSTITUTE INC. SAS 9.4 Language Reference: Concepts. Cary, NC, USA: SAS Institute Inc., 2013. ISBN 1612905641, 9781612905648.
5. MINITAB, Inc. Minitab 17 Statistical Software. B.m.: State College, PA. 2010.
6. MICROSOFT. MS Office Excel, version 16. 2016.
7. THOMPSON, C. M. a L. SHURE. Global Optimization Toolbox: User's Guide. B.m.: MathWorks. 2011.
8. STACORP. Stata Statistical Software: Release 14 [online]. 2015. ISBN 1881228541. Dostupné z: doi:10.2307/2234838.
9. R CORE TEAM. R: A Language and Environment for Statistical Computing [online]. Vienna, Austria: R Foundation for Statistical Computing, 2016. Dostupné z: <https://www.R-project.org/>.
10. FOUNDATION, Python Software. Python Language Reference, version 2.7 [online]. 2017. Dostupné z: <http://www.python.org>.
11. HENDL, Jan. Přehled statistických metod: analýza a metaanalýza dat. Praha: Portál, 2015. ISBN 978-80-262-0981-2.

DÁTAMI K LEPŠÍM EXPONÁTOM

Oliver Velich

VIDA! science centrum; Fakulta informatiky, Masarykova univerzita

Úvod

Hlavnou motivací projektu je automatická analýza interakce návštěvníků s exponátem vrstevnice vo VIDA! vědeckom centre a následná evaluácia prínosu exponátu. Na základe výsledkov by sa mohli doplniť rôzne varianty hier, vylepšit edukatívnuosť exponátu, premiestniť exponát v priestoroch centra alebo inak propagovať.

Cieľom projektu je vytvoriť nástroj, ktorý bude vedieť odhadnúť počet návštevníkov aktuálne interagujúcich s exponátom a navrhnúť vizualizáciu, ktorá bude opisovať intenzitu interakcie návštevníkov s exponátom, a to v závislosti na móde, čase a dni v týždni. Počet návštevníkov bude odhadovaný pomocou regresie a/alebo analýzy videa.

Dátový súbor a metodika

Pôvod dátového súboru

Dátový súbor pochádza z vedecko-zábavného parku VIDA! (vida.cz), konkrétne z exponátu s názvom Vrstevnice. Jedná sa o „nádobu“ o veľkosti približne pol metra x pol metra, v ktorej sa nachádza približne 10 cm výšky piesku. Pomocou projektoru sa na piesok zobrazujú rôzne typy vizuali-

zácií, ktoré predstavujú geografické javy a označenia. Súčasťou exponátu je aj dotyková obrazovka, kde si návštevník môže vybrať rôzne módy týchto vizualizácií.

Technologicky je exponát riešený prepojením dotykovej obrazovky, počítača, Kinectu a projektoru. Kinect z oblasti s pieskoviskom vytvára hĺbkovú mapu, na základe ktorej počítač generuje scény a posielajú grafický výstup do projektoru. Úlohou návštevníka teda je meniť výšku piesku v rôznych bodoch, na čoho základe sa následne bude meniť príslušná vizualizácia. K dispozícii je 8 rôznych obrazových módov: lávový, čistý vrstevnicový, krajina, vyplnený vrstevnicový, spádové oblasti riek, voda – dážď, psychadelický a výukový mód.

Exponát je schopný aj exportu dát. Ten spočíva v exportovaní udalostí zmena módu (tá mohla nastať automaticky alebo na podnet výberu užívateľom) a udalosti, že v hĺbkovej mape sa nachádza veľa miest s extrémnymi hodnotami (prítomnosť objektov/ludí v pieskovisku). Takisto je k dispozícii video záznam z kamery exponátu.

Príprava dát

Surové dáta z exponátu je v prvom rade nutné spracovať pre potreby regresnej analýzy a následne aj pre vizualizáciu. Údaje boli predspracované pomocou jazyka Python, 2017 (© Python Software Foundation, dostupný zdarma na <https://www.python.org/>, použitá verzia 3.6). Dáta z exponátu boli pôvodne generované vo forme: (počet_bodov, Timestamp) a (mód, Timestamp). Pre potreby regresnej analýzy boli premenné rozšírené na formu: (počet_bodov, Timestamp, čas_od_predchádzajúceho, čas_do_konca) a (mód, Timestamp, dĺžka_predchádzajúceho, udalosť). Pre rýchle generovanie vizualizácie boli dáta predspracované a zjednotené na nasledujúce formáty: (dátum, sekunda, počet_bodov, mód, click), (dátum, 30sekúnd, počet_bodov, mód) a (dátum, začiatok_session, koniec_session, dĺžka_session). Pre vizualizáciu boli teda dáta rozšírené tak, aby obsahovali všetky zobrazované záznamy a žiadne hodnoty sa nedopočítavali. Naopak pri regresnej analýze sa počíta iba so záznamami, kedy nastala udalosť.

Video sa nenahrávalo s pevným počtom snímkov za sekundu a preto bolo potrebné vypočítať priemerný počet snímkov za sekundu (z dostupných informácií o reálnej dĺžke videa a celkovom počte snímkov), aby bolo jednoduchšie v ňom vyhľadávať informácie zachytené v ostatných súboroch.

Pre odhad počtu návštevníkov interagujúcich s exponátom v danom čase som použil 2 rôzne metódy: regresnú analýzu a analýzu obrazu. Regresnú analýzu som aplikoval na dáta získané z hĺbkovej mapy Kinectu (počet nadhraničných bodov, teda počet bodov presahujúcich hraničnú hodnotu 300) a analýzu obrazu som zase aplikoval na dáta získané z kamery (prítomnosť osoby vo výreze záznamu). Pre vytvorenie samotných klasifikátorov bolo nutné vytvoriť najprv tréningové súbory, prostredníctvom, ktorých prebehlo naučenie klasifikátorov. Následne boli klasifikátory vyvinuté na príslušnom tréningovom súbore, použité na predikciu premennej počet ľudí interagujúcich s exponátom na zvyšnom dátovom súbore a v poslednom bode bola otestovaná ich správnosť predikcie pomocou vizualizácií.

Celkový dátový súbor obsahuje vyše 3 miliónov vzoriek, pričom na natrénovanie regresného modelu bolo použitých 12 649 vzoriek. Klasifikátor detekcie objektov v obraze bol natrénovaný na 101 pozitívnych vzorkách a 970 negatívnych.

Diskusia a výsledky

Regresná analýza

Ako prvú štatistickú metódou pre odhad počtu návštevníkov interagujúcich s exponátom som použil regresnú analýzu. Tréningový súbor bol vytvorený tak, že som podľa časovej známky vo videu

pre jednotlivé session (nepretržitá práca s exponátom bez prestávky dlhšej ako 3 sekundy) vyhľadal ich záznam vo videu. S presnosťou na jednu sekundu som určil, koľko ľudí v daný moment interaguje s exponátom. Jedna sekunda mohla obsahovať 1 až 8 záznamov v pôvodných dátach. Celkovo som teda náhodne do tréningového súboru ovzorkoval zo všetkých videí dokopy približne 45 minút z celkových 17 220 minút.

Predikovanou premennou teda je počet interagujúcich návštevníkov ako jedna z hodnôt 1, 2, 3 alebo 4. Pri počte 4 a viac už nebolo možné odlišiť, koľko ľudí naozaj s exponátom interaguje. Jedným zmysluplným prediktorom teda je počet extrémnych (nadhraničných) bodov v hĺbkovej mape. Počet sekúnd do konca som takisto použil ako prediktor s predpokladom, že počet interagujúcich návštevníkov je závislý na dennej dobe.

Keďže je predikovaná premenná kategoriálna, regresnú analýzu som vykonal pomocou logistickej regresie a to v troch formách: multinomiálna logistická regresia [1], ordinálna multinomiálna logistická regresia [2], binárna logistická regresia. Pre binárnu logistickú regresiu som premennú rozdelil na kategórie 1 a 2 (jednotlivec, skupina aspoň dvoch ľudí).

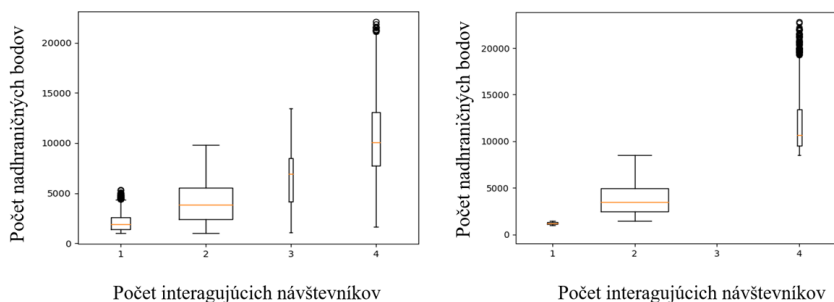
Pokiaľ bola interakcia súvislá, následné záznamy (nie sú od seba ďalej ako 3 sekundy) som zjednotil do tzv. sessions. Pre každú session som vytvoril premenné, ktoré definovali priemerný počet nadhraničných bodov, dĺžku session a priemerný počet interagujúcich návštevníkov. Keďže priemerný počet interagujúcich návštevníkov je spojitá premenná, bolo možné pre tieto premenné použiť lineárnu regresiu.

Výstupy všetkých regresných analýz som spracoval do formy krabicových grafov, aby bolo možné vizuálne skúmať výsledky regresie. Obrázok 1 znázorňuje vľavo rozloženie testovacích dát ručne označovaných pomocou záznamu z kamery a vpravo výsledky ordinálnej regresie na zvyšnom súbore. Výsledky tejto ordinálnej regresnej analýzy sú najbližšie pôvodnému rozloženiu z tréningovej sady. Prediktor počet bodov hĺbkovej mapy je štatisticky významný na hladine významnosti 0.001 pre všetky kategórie premennej počet účastníkov. Zaujímavým zistením je, že pridanie prediktora čas do záverečnej hodiny VIDA! centra zlepšilo výsledky predikcie modelu podľa Akaikovho informačného kritéria. Pričom, čím kratší čas do zatvorenia, tým väčší počet návštevníkov interaguje s exponátom. S interpretáciou je však nutné zaobchádzať opatrne, hlavne pri pohľade na krabicové grafy predikcií. Všetky výsledky avšak trpeli na veľmi nízky počet tréningových hodnôt v kategórii 3 ľudia.

Spracovanie obrazu

V ďalšom bode som pre predikciu počtu návštevníkov interagujúcich s exponátom použil analýzu obrazu. Vo výstupe z kamery je možné detegovať ľudí pomocou metód tzv. object detection, pričom v projekte som ako deskriptor obrazu použil Histogram of Oriented Gradients (HOG) [3] a ako klasifikátor lineárny Support Vector Machine (SVM). Obe metódy boli implementované v jazyku Python pomocou existujúcich knižníc (scikit-learn verzia 0.18.1, mord verzia 0.4, scikit-image verzia 0.13, OpenCV-Python verzia 3.0).

HOG deskriptor je vypočítaný pre 150x150 px výrez obrázku. Z obrázkov bolo ručne vytvorených 101 výrezov, ktoré obsahovali postavu a automaticky viac ako 1000 výrezov, ktoré postavu neobsahovali. Z týchto tréningových dát bol natrénovaný SVM a otestovaný na testovacom obrázku. Klasifikátor vykazoval vysokú mieru falošne pozitívnych výsledkov. Preto som vykonal jednu iteráciu tzv. hard negatives mining [3]: klasifikátor bol využitý na nájdenie výrezov v obrázkoch, ktoré sú určite negatívne a tieto výrezy boli pridané do množiny negatívnych vzoriek a klasifikátor bol znovu natrénovaný na rozšírenej množine. Po tejto procedúre sa zlepšil výkon klasifikátoru, stále však vykazuje niekoľko falošne pozitívnych výsledkov.



Obrázok 1. Rozloženie pôvodných dát a výsledky ordinálnej regresnej analýzy

Celkovo však pristupujem k výkonu modelu pozitívne, a predpokladám, že po rozšírení tréningovej sady bude spracovanie obrazu spoľahlivou metódou detekcie návštevníkov, ktorá bude môcť byť rozšírená aj o ďalšie sledované vlastnosti a použiteľná na iné scény.

Vizualizácia

Vizualizácia dát, ktorá slúži na prehľadné zobrazenie interakcie návštevníkov, je dostupná na odkaze: <http://vidatest.wz.cz/>. Aplikácia bola vyvíjaná v jazyku javascript pomocou knižnice D3.js (verzia 4.0, voľne dostupná na <https://d3js.org/>) a testovaná na Mozilla FireFox 53.0.3. Prvá vizualizácia predstavuje na osi X deň návštevy a na osi Y čas v priebehu dňa. Jednotlivé módy sú od seba farebne rozlíšené. Časový údaj 30 sekúnd predstavuje 1 pixel. Farba pixelu označuje mód, ktorý bol počas daných 30 sekúnd najčastejší. Svetlosť a tmavosť farieb znázorňuje priemerný počet bodov nad 300 jednotiek hĺbkovej mapy. Tieto body totiž predstavujú osoby, teda návštevníkov. Môžeme teda konštatovať, že čím svetlejšia je farba, tým je intenzívnejšie využívaný. Čierne oblasti predstavujú čas, kedy pri exponáte nebol žiadny návštevník. Takisto je možnosť odfiltrovať iba jeden mód záujmu.

Ďalšou súčasťou je zobrazenie po kliknutí na ľubovoľný deň. Na novej karte sa otvorí návštevnosť exponátu v daný deň, pričom čierne oblasti - bez návštevníkov už zobrazené nie sú (okrem víkendu, kedy sa VIDA! otvára neskôr, a teda sa aj neskôr zapína exponát a v plnom zobrazení je čierny Off mód). Jednotlivé časové oblasti, kedy sa s exponátom nepretržite (s prestávkou menšou ako 3 sekundy) pracuje nazývame session. Pomocou posuvníka si môžeme meniť zobrazenie podľa toho, aký chceme, aby bol minimálny čas strávený pri exponáte, teda dĺžku session. Na osi Y sa opäť nachádza čas v priebehu dňa a na osi X je tento krát pomocou žltej čiary zobrazená informácia o presnom počte extrémnych bodov v hĺbkovej mape v danej sekunde, teda o počte návštevníkov v danom čase. Pomocná interpretácia je, že hodnoty do 5000 väčšinou znamenajú interakciu jedného návštevníka, nad 5000 potom dvojicu a nad 10 000 už skupinu. Farebnosť jednotlivých úsekov opäť predstavuje mód. Zaujímavým doplnením je takisto biela čiara, ktorá znázorňuje to, či bol mód vygenerovaný automaticky, alebo si ho návštevník sám vybral pomocou dotykovej obrazovky.

Neinteraktívne zobrazenie istého časového úseku je možné pomocou výberu na prehľadovej obrazovke. Výber je zobrazený priehľadným bielym obdĺžnikom, ktorého veľkosť a umiestnenie je možné meniť. Dvojklik na danú oblasť otvorí novú kartu s priblíženým výberom.

Záver

V projekte som riešil predikciu počtu návštevníkov pri interaktívnom exponáte z dostupných dát a zároveň vizualizáciu dostupných dát pre evaluáciu interakcie v rámci dňa a pre jednotlivé módy. Hoci spoľahlivá predikcia bude vyžadovať pridanie ďalších tréningových vzoriek, je opodstatnené, že ako ordinálna regresia, tak aj analýza obrazu budú vhodným nástrojom na automatickú evaluáciu exponátu. Vytvorená vizualizácia aj napriek množstvu dát funguje interaktívne a príná-

ša vhlad do vytaženosti exponátu. Získané informácie budú slúžiť na úpravu exponátu zo strany výrobcu aj zadávateľa.

Literatúra

1. KWAK, Chanyeong a Alan CLAYTON-MATTHEWS. Multinomial logistic regression. *Nursing research*. 2002, 51(6), 404-410.
2. MCCULLAGH, Peter. Regression models for ordinal data. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1980, 42(2), 109-142.
3. DALAL, Navneet a Bill TRIGGS. Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego: IEEE, 2005, s. 886-893. ISBN 0-7695-2372-2. ISSN 1063-6919.

PŘEHLED WEBOVÝCH PORTÁLŮ POD ZÁŠTITOU IBA LF MU & ÚZIS ČR

Společné pracoviště IBA LF MU (Institut biostatistiky a analýz Lékařské fakulty Masarykovy univerzity) a ÚZIS ČR (Ústavu zdravotnických informací a statistiky České republiky) se dlouhodobě věnuje přípravě hodnotných analýz, reportů a publikací, které jsou zaměřeny na širokou odbornou i laickou veřejnost, zástupce krajské správy a poskytovatele zdravotní péče. Nedílnou součástí těchto aktivit je také návrh, vývoj a implementace webových portálů, které zpřístupňují finální data koncovému uživateli prostřednictvím Internetu. Od roku 2016 byla v rámci řešení konkrétních projektů spuštěna celá řada webů, které prostřednictvím moderních technologií a přístupů prezentují ucelené výstupy pod garancí IBA LF & ÚZIS ČR. Tabulka níže ukazuje seznam klíčových projektů včetně celkového doménového uspořádání.

Tabulka 1. Přehled projektů a jejich domén

Název projektu	Doména
Klasifikační systém CZ-DRG	http://drg.uzis.cz/klasifikace-pripadu/
Konference EDCD: European Digestive Cancer Days	http://www.cancer-days.eu/
Konference PREVON: Programy prevence, screeningu a časně diagnostiky vážných onemocnění ve světle současných poznatků a praktických zkušeností	http://prevon.uzis.cz/
Metodická optimalizace a zefektivnění systému úhrad nemocniční péče v ČR (DRG Restart)	http://drg.uzis.cz/
Národní datová základna paliativní péče	http://ndzpp.uzis.cz/
Národní koordinační centrum prevence časného záchytu onemocnění	http://nkc.uzis.cz/
Národní portál Systém hlášení nežádoucích událostí	http://shnu.uzis.cz/
Národní registr poskytovatelů zdravotních služeb	http://nrpzs.uzis.cz/
Regionální zpravodajství NZIS	http://reporting.uzis.cz/
Rozvoj technologické platformy NZIS	http://rozvojnzis.uzis.cz/
Výběrová šetření o zdraví	http://ehis.uzis.cz/

KLASIFIKAČNÍ SYSTÉM CZ-DRG

Odborný garant: RNDr. Tomáš Pavlík, Ph.D.

<http://drg.uzis.cz/klasifikace-pripadu/>



Webový portál Klasifikační systém CZ-DRG obsahuje sadu klasifikačních pravidel pro zařazení hospitalizačních případů do jednotlivých úrovní systému CZ-DRG. Primární motivací je snaha přehledně zpřístupnit tato propracovaná pravidla širokému spektru uživatelů v řad odborné i laické veřejnosti. CZ-DRG je nově vyvíjený patientský klasifikační systém navržený pro akutní lůžkovou péči v ČR. Je založen na hierarchickém třídění jednotek klasifikace, kterými jsou tzv. hospitalizační případy, do výsledných skupin systému, tzv. Diagnosis Related Groups, volně přeloženo skupin klinicky příbuzných diagnóz. Třídění probíhá podle předem stanovených kritérií, tedy známých klasifikačních pravidel. Pro klasifikaci využívá systém CZ-DRG známé charakteristiky hospitalizačního případu popsané s využitím primárních klasifikačních systémů, kterými v případě CZ-DRG jsou Mezinárodní klasifikace nemocí 10. revize (MKN-10) a Seznam zdravotních výkonů s bodovými hodnotami vydávaný Ministerstvem zdravotnictví ČR. Vývoj nového klasifikačního systému CZ-DRG lze charakterizovat následujícími principy:

- Návaznost struktury systému CZ-DRG na obecně platné strukturální prvky mezinárodních DRG systémů – struktura CZ-DRG respektuje společné prvky mezinárodních DRG systémů, zejména hierarchické uspořádání, členění systému dle MDC, využití rutinně sbíraných informací o hospitalizaci pacienta, apod.
- Rozšíření hierarchické struktury klasifikačního systému CZ-DRG o jednu úroveň – pro lepší klinickou srozumitelnost vnitřní struktury systému CZ-DRG jsme rozšířili jeho hierarchickou strukturu (proti systému IR-DRG, který má tři úrovně) o jednu úroveň, tzv. DRG kategorii, která je v taxonomii podřazena MDC a nadřazena DRG bázi, což umožní specifikovat DRG báze na míru konkrétním klinickým stavům.
- Vytvoření komplexních klasifikačních pravidel – klasifikační systém CZ-DRG využívá pro klasifikaci HP na jednotlivých úrovních další proměnné (atributy), které jsou relevantní vzhledem k jednotlivým onemocněním a klinickým stavům a s nimi souvisejícími náklady na hospitalizaci.

Jeden ze základních požadavků na DRG klasifikační systémy hospitalizačních případů je využití rutinně sbíraných údajů o hospitalizaci pacienta. Informace o klinickém stavu pacienta a poskytnuté zdravotní péči je potřeba popsat pomocí kódů, proto klasifikační systémy DRG jsou ve skutečnosti sekundárními klasifikačními systémy, neboť pro svá pravidla využívají jiné klasifikační systémy, tzv. primární klasifikace. Postup návrhu DRG bází zahrnuje kromě samotné definice a parametrizace DRG bází i přípravu předchozích pater hierarchické struktury klasifikačního systému CZ-DRG, což zahrnuje následující kroky: revize seznamu a obsahu MDC, návrh DRG kategorií, oponentura DRG kategorií, návrh DRG bází, oponentura DRG bází.

Klasifikační systém CZ-DRG

Úvod

Struktura systému CZ-DRG

Definiční listy

Interaktivní klasifikátor (grouper)

Analýzy a publikace

Projekt DRG Restart

Datové zdroje

Kontakty

Vítejte na webovém portálu klasifikačního systému CZ-DRG!

Vytvoření webového portálu s klasifikačními pravidly pro zařazení hospitalizačních případů do jednotlivých úrovní systému CZ-DRG je primárně motivováno snahou zpřístupnit tato pravidla širokému spektru uživatelů a zájemců.

V budoucnu bude možné využít webový portál i pro online zařazení hospitalizačního případu do DRG báze a skupiny dle algoritmu CZ-DRG.

MDC 05 Nemoci a poruchy oběhové soustavy

[zpět na výběr MDC](#)

Postavení MDC ve struktuře klasifikačního systému CZ-DRG

CZ-DRG

↪ MDC 05 Nemoci a poruchy oběhové soustavy

Definiční list MDC

MDC 05 představuje jednu z orgánově specifických MDC a zahrnuje hospitalizační případy pacientů s nemocemi a poruchami týkajícími se oběhové soustavy, tedy srdce, tepen a žil. Hospitalizační případ je klasifikován do MDC 05 tehdy, pokud se jeho hlavní diagnóza nachází v seznamu přípustných diagnóz pro tuto MDC (viz seznam na konci stránky). Výjimku tvoří hospitalizační případy pacientů splňující kritéria pro polytrauma (tyto případy jsou zařazeny do MDC 25), novorozenci (tyto případy jsou zařazeny do MDC 15), hospitalizační případy, v rámci kterých došlo k transplantaci orgánu resp. krevetvorných buněk nebo hospitalizační případy, kdy byl pacient více jak 5 dní na umělé plicní ventilaci (tyto případy jsou zařazeny do MDC 00). V systému CZ-DRG je pro MDC 05 navrženo celkem 19 DRG kategorií definovaných na základě typu a lokalizace onemocnění a 113 DRGází zohledňujících jednotlivé modality léčby.

Tato MDC zahrnuje následující DRG kategorie:

[rozbalit vše](#)

[sbalit vše](#)

- | 05-X01 Mechanická srdeční podpora
- | 05-Z01 Nemoci myokardu
- | 05-Z02 Nemoci perikardu
- | 05-F01 Akutní koronární syndrom
- | 05-F02 Chronická ischemická choroba srdeční
- | 05-F03 Poruchy srdečního rytmu
- | 05-F04 Srdeční zástava a šok
- | 05-F05 Srdeční selhání
- | 05-F06 Funkční a strukturální poruchy chlopní a zánět endokardu
- | 05-F07 Funkční a strukturální poruchy aorty

Odborný garant: doc. RNDr. Ladislav Dušek, Ph.D.

<http://www.cancer-days.eu/>



Mezinárodní konference European Digestive Cancer Days (EDCD) navazuje na tradici každoročních setkání European Colorectal Cancer Days a rozšiřuje její tematický rozsah. Jedním z jejích cílů byla podpora a propagace screeningu kolorektálního karcinomu podle mezinárodních doporučení a standardů. To bylo důležité zejména pro Českou republiku, neboť její screeningový program v původní podobě nezahrnoval adresné zvaní občanů. Tato konference EDCD přispěla mimo jiné právě i k zavedení adresného zvaní a tím pádem k přeměně českého screeningového programu na populační v roce 2014. Nová konference EDCD rozšiřuje rozsah témat původní konference na nádorová onemocnění celého trávicího traktu a související aktivity – prevenci, včasný záchyt, léčbu a dlouhodobou péči o pacienty s těmito onemocněními.

Konference EDCD poskytne prostor nejen k diskuzi nad aktuálními odbornými otázkami, ale i k prezentaci stanovisek politiků a pacientů k implementaci důležitých aktivit vedoucích k úspěšnému boji s nádory trávicího traktu. Pozornost bude věnována zejména následujícím otázkám:

- Existují strategie ke zmírnění rostoucí zátěže zhoubnými nádory slinivky břišní?
- Jaké jsou priority při boji s nádory trávicího traktu?
- Jaký je současný stav implementace programů screeningu kolorektálního karcinomu v Evropě?
- Jak začít program screeningu kolorektálního karcinomu, jak ho sledovat a hodnotit?
- Jak zvýšit účast ve screeningu kolorektálního karcinomu, usnadnit informované rozhodování a snížit rozdíly v dostupnosti vyšetření a pokrytí populace?

Konference 2015 Konference 2014 Konference 2013 Konference 2012 en

Úvod

Program

Registrace

Místo konání

Ubytování

Organizátoři

Sponzoři

Programový výbor

Organizační výbor

Pro média

European Digestive Cancer Days

PRAHA 2017 – PREVENCE A SCREENING

26.–27. září 2017

REGISTRACE

Úvod

KONFERENCE PREVON: PROGRAMY PREVENCE, SCREENINGU A ČASNÉ DIAGNOSTIKY VÁŽNÝCH ONEMOCNĚNÍ VE SVĚTLE SOUČASNÝCH POZNATKŮ A PRAKTICKÝCH ZKUŠENOSTÍ

Odborný garant: doc. RNDr. Ladislav Dušek, Ph.D.

<http://prevon.uzis.cz/>

Konference PREVON je organizována v rámci projektu Národní koordinační centrum programů časného záchytu onemocnění I, který je spolufinancován Evropskou unií. Důležitým cílem akce je také představit aktivity nově vznikajícího Národního koordinačního centra programů časného záchytu onemocnění. Diskutovány budou stávající screeningové programy, jejich metodika a výsledky, ale také nově plánované programy časného záchytu onemocnění. Dále budou představeny návrhy konceptů metodik pro zavádění programů časného záchytu onemocnění do praxe a jejich hodnocení. V rámci konference budou rovněž diskutovány zahraniční zkušenosti v oblasti časné diagnostiky závažných onemocnění.

Konference chce otevřít diskusi nad problematikou časného záchytu onemocnění v mezinárodním srovnání. Vzhledem k významné variabilitě programů prevence v různých evropských zemích jde jistě o témata aktuální a zajímavá nejen pro lékaře a vědce, ale také pro pacienty a politické zástupce občanů.

Diskuse výše uvedených otázek je důležitým krokem k posílení screeningových programů a k dosažení jejich co možná největší účinnosti i nákladové efektivity. Pevně doufáme, že setkání všech stran zainteresovaných na tématu časného záchytu onemocnění – ať už se jedná o zástupce státní správy, odborných lékařských společností, patientských organizací, vědeckých institucí a v neposlední řadě i politiků – vytvoří podnětnou atmosféru k plodným diskuzím o všech zmíněných problémech.

Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Ústav zdravotnických informací a statistiky ČR ÚZIS

Úvod

Program

Registrace

Místo konání a ubytování

Sponzoři

Programový výbor

Organizační výbor

Pro média

PREVON 2017 - 25. září 2017

Programy prevence, screeningu a časné diagnostiky vážných onemocnění ve světle současných poznatků a praktických zkušeností

Výzkum → Implementace → Kvalita → Bezpečnost → Přínos → Nákladová efektivita

REGISTRACE

Úvod

Konference *PreVOn 2017 – Programy prevence, screeningu a časné diagnostiky vážných onemocnění ve světle současných poznatků a praktických zkušeností* se koná 25. září 2017 a představí aktuální informace v oblasti časné diagnostiky onemocnění.

Tato konference je organizována v rámci projektu *Národní koordinační centrum programů časného záchytu onemocnění I*, který je spolufinancován Evropskou unií.

Důležitým cílem konference je také představit aktivity nově vznikajícího Národního koordinačního centra programů časného záchytu onemocnění.

METODICKÁ OPTIMALIZACE A ZEFEKTIVNĚNÍ SYSTÉMU ÚHRAD NEMOCNIČNÍ PÉČE V ČR (DRG RESTART)

Odborný garant: Ing. Jan Linda

<http://drg.uzis.cz/>



Projekt DRG Restart sleduje jako nosný cíl vybudování dlouhodobě udržitelné datové, informační a personální základny pro optimalizaci a průběžnou kultivaci systému úhrad lůžkové péče v ČR. Dosažení tohoto cíle předpokládá ustavení multioborového týmu klinických specialistů, odborníků na ekonomiku ve zdravotnictví, ICT a statistickou analýzu dat.

Základní principy, na kterých projekt DRG restart plánuje vybudovat systém optimalizace úhrad lůžkové péče, jsou následující:

- respekt k funkčním klinickým standardům, k standardizovaným diagnostickým a léčebným algoritmům
- vybudování reprezentativní a pravidelně aktualizované datové základny
- exaktní a reprodukovatelná metodika
- princip odborné oponentury
- transparentnost komunikace

Globálním cílem projektu DRG Restart je vybudování dlouhodobě udržitelné datové, informační a personální základny pro optimalizaci a průběžnou kultivaci systému úhrad lůžkové péče v ČR a zvýšení prediktivní schopnosti a efektivity úhradových mechanismů pro tento segment zdravotní péče. Dílčí cíle projektu představují:

- vybudování Kompetenčního centra a reprezentativní sítě referenčních nemocnic
- vytvoření metodik pro oceňování hospitalizačních případů, elektronických a databázových nástrojů pro implementaci tohoto systému a ověření těchto metodik v reprezentativní síti referenčních nemocnic
- tvorbu nové klasifikace hospitalizačních procedur, redefinice klasifikačního systému DRG a vytvoření nového kodérského manuálu a jejich implementace do úhradových mechanismů zdravotní péče v ČR
- přípravu podkladů pro převedení oceňování nákladů nemocniční péče do elektronické podoby a příprava k praktické a plošně realizovatelné implementaci těchto nástrojů ve zdravotnických zařízeních
- realizaci modelové studie v síti referenčních nemocnic, která budou oceňovat klinické případy lůžkové péče a generovat kalibrační data z reálné klinické praxe, jež budou využita k aktualizaci výstupů projektu
- realizaci podpůrných edukačních aktivit pro cílové skupiny



DRG RESTART

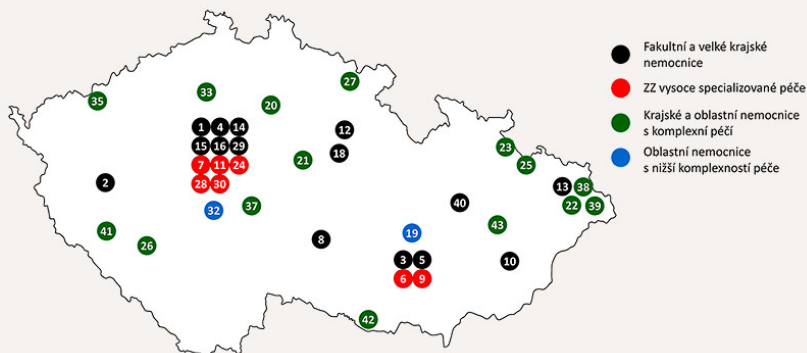
Metodická optimalizace a zefektivnění systému úhrad nemocniční péče v ČR

- Úvod
- Výstupy
- Referenční síť**
- Odborní asistenti
- Mapování péče
- Diskuzní fórum
- Aktuality
- Konference, semináře
- Kontakt

Referenční síť nemocnic

V roce 2015 byla na základě vícekritériální analýzy všech nemocnic v ČR vybudovaná referenční síť nemocnic. Vybrané nemocnice pokrývají 4 kategorie z hlediska velikosti a komplexnosti péče. Zástupci páté kategorie vzešlé z metodiky výběru referenčních nemocnic (ZZ s úzkým zaměřením) nejsou v tuto chvíli členy referenční sítě.

V roce 2016 referenční nemocnice podepsaly dodatek ke smlouvě, kde smluvní strany prohlašují, že budou pokračovat ve spolupráci na projektu DRG Restart, a to jejím rozšířením na společnou realizaci projektu „Metodická optimalizace a zefektivnění systému úhrad nemocniční péče v ČR“.



- Úvod
- Výstupy
- Referenční síť
- Odborní asistenti
- Mapování péče**
- Diskuzní fórum
- Aktuality
- Konference, semináře
- Kontakt

- Dostupné zdroje dat
- Metodika mapování péče
- Prezentace
- Interaktivní prohlížeč dat

Základní charakteristiky zdravotnických zařízení (ZZ) akutní lůžkové péče v ČR

II.1. Počet oddělení s akutní lůžkovou péčí a počet lůžek akutní péče v ZZ

Typ ZZ

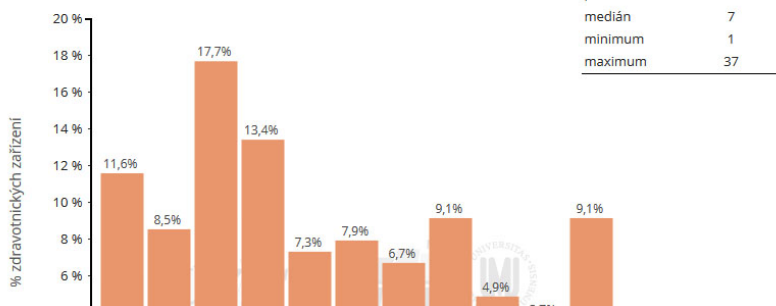
Všechny typy zdravotnických zařízení

Zobrazit

Zpět

Vybraná skupina: N = 164 | za období 2012

Počet oddělení akutní lůžkové péče v ZZ




NÁRODNÍ DATOVÁ ZÁKLADNA PALIATIVNÍ PÉČE


Odborný garant: Mgr. Jan Švancara

<http://ndzpp.uzis.cz/>



Cílem Národní datové základny paliativní péče je na základě dostupných datových zdrojů Národního zdravotnického informačního systému (NZIS) zmapovat a popsat závěr života a umírání v České republice. Data NZIS obsahují cenné údaje o hospitalizační historii, hospitalizačních i úmrtních diagnózách, jakož i základní socio-demografické údaje. Systematická analýza těchto již existujících dat může přinést užitečné informace o možnostech a slabinách českého zdravotnického systému. Cílem tohoto projektu je poskytnout komplexní pohled na problematiku závěru života a přispět tak k práci zdravotníků snažících se o zmírnění utrpení a zlepšení kvality života desetitisíců umírajících ročně. Právě tato robustní datová základna umožňuje mapovat, jak probíhá závěr života osob umírajících v České republice, a slouží jako podklad pro plánování paliativní péče v České republice.



 cs / en

[Úvod](#) - [Mortalitní data ČR](#) - [Indikátory kvality péče](#) - [Poslední rok života](#) - [Posledních 30 dní života](#) - [Poskytovatelé](#)

Poskytovatelé

Paliativní péče představuje široce definovaný a diverzifikovaný segment zdravotní péče. Symbolem paliativní péče je lůžkový hospice. Na hospicovém lůžku však dochází jen k necelým 3 % úmrtí v ČR. Jedná se převážně o onkologické pacienty s dobře predikovatelnou dobou dožití. Majoritní část úmrtí tak probíhá v nemocnicích či léčebnách dlouhodobě nemocných, často s nemalou potřebou paliativní péče. Proto jakékoli konceptní uvažování a monitorování paliativní péče musí být zohledňovat potřeby umírajících v různých zdravotnických zařízeních, stejně jako v domácím prostředí.

Popsat poskytování paliativní péče z dostupných datových zdrojů je značně složité a výsledný obrázek je spíše orientační. Mapa poskytovatelů paliativní péče totiž monitoruje pouze část celého spektra paliativní péče. Na mapě můžeme vidět jednak lůžkové hospice, ale také jiná zdravotnická zařízení, která vykazují výkonu paliativní péče, dle daného sazebníku výkonů. Mapa zcela opomíjí péči mobilní a ambulantní.

Na zdokonalení dostupnosti informací o regionální distribuci paliativní péče dále pracujeme.



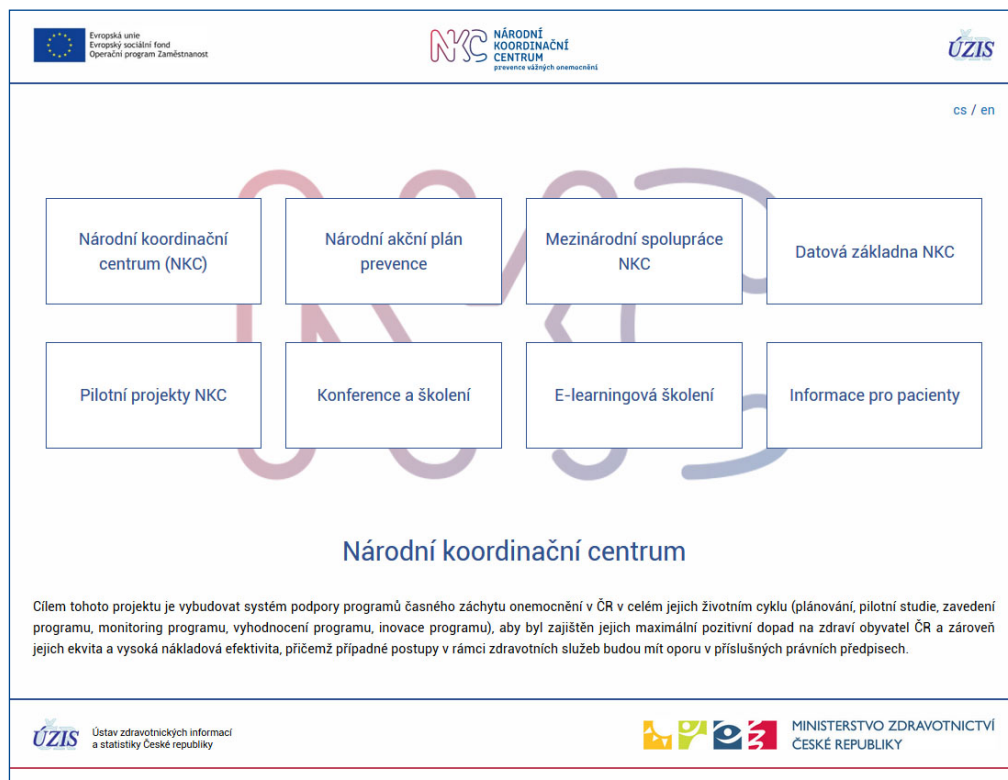
NÁRODNÍ KOORDINAČNÍ CENTRUM PREVENCE ČASNÉHO ZÁCHYTU ONEMOCNĚNÍ

Odborný garant: RNDr. Ondřej Májek, Ph.D.

<http://nkc.uzis.cz/>



Cílem projektu je vytvořit koordinační řídicí strukturu a metodické zázemí pro zavádění realizaci a evaluaci screeningových programů v ČR. V rámci projektu vznikne Národní koordinační centrum složené z Řídící rady a pracovních skupin, které budou navrhovat realizaci pilotních studií a vyhodnocovat jejich dopad a zavede postupy pro monitoring, evaluaci a inovace jednotlivých screeningů. Tím dojde ke zvýšení dopadu screeningů na zdraví obyvatel a optimalizaci jejich nákladové efektivity.



NÁRODNÍ PORTÁL SYSTÉM HLÁŠENÍ NEŽÁDOUCÍCH UDÁLOSTÍ

Odborný garant: doc. PhDr. Andrea Pokorná, Ph.D

<http://shnu.uzis.cz/>



Systém hlášení nežádoucích událostí je metodicky orientovaným projektem, který je koncipován jako systém hodnocení sloužící primárně:

- k umožnění sdíleného učení jednotlivých smluvně zapojených poskytovatelů zdravotních služeb (PZS) v unifikovaném způsobu identifikace rizik, hodnocení a vypořádávání nežádoucích událostí (NU) s využitím jednotných metodických pokynů;
- k vyhodnocování anonymizovaných agregovaných dat o hlášených hlavních typech NU (dle klasifikace Světové zdravotnické organizace);
- k porovnávání jednotlivých poskytovatelů zdravotních služeb v odpovídající kategorii mezi sebou, dle předem stanovených kritérií a denominátorů (např. počtu hospitalizovaných pacientů, počtu lůžek, počtu konkrétních nežádoucích událostí přepočtenému na 1000 pacientů, počtu nelékařských zdravotnických pracovníků přepočtenému na pacienta apod.).

SHNU je primárně platformou pro sdílené učení v oblasti identifikace a prevence rizik při poskytování zdravotních služeb. Metodické návody jsou určeny jako pomoc zdravotníkům v klinické praxi v jednotné interpretaci požadavků na zadávání údajů o nežádoucích událostech na lokální úrovni a jejich hlášení na centrální úrovni v jednotné agregované podobě (na základě unifikované terminologie). Sekundárně tak systém slouží k umožnění sběru agregovaných dat o počtu hlášení nežádoucích událostí v Centrálním úložišti výkazů, aktuálně v rámci pilotních sběrů u 78 smluvně zapojených poskytovatelů zdravotních služeb.

Aktuálně je v rámci SHNU vypracováno 5 obecných metodických dokumentů v rozsahu 176 stran, dále 16 metodických dokumentů dle typu NU v rozsahu 384 stran a z nich vyplývajících 9 algoritmů prevence a 8 algoritmů opatření po vzniku NU. Metodické pokyny k hlavním nežádoucím událostem jsou provedeny v několika vzájemně obsahově konzistentních dokumentech, které se liší svým rozsahem (plná metodika, zkrácená verze, algoritmus prevence a algoritmus opatření po vzniku NU). Prostřednictvím přípravy jednotných dokumentů a na základě jednotné metodiky pro identifikaci rizik, vypořádání NU a jejich systematickou prevenci s využitím potenciálu zástupců jednotlivých PZS a lokálního know-how by mělo být podpořeno racionální nakládání s lidskými a materiálními zdroji.



NÁRODNÍ PORTÁL Systém hlášení nežádoucích událostí

Úvod

Metodické materiály

Srovnání dle typu NU

Srovnání dle PZS

Přihlášení

Hledat



Kontaktní osoby

Aktuality a novinky

Informace o SHNU

Organizační struktura

Pracovní skupina

Odkazy

FAQ

Interní část webu

Vítejte na nově připravených stránkách Národního portálu pro Systém hlášení nežádoucích událostí (SHNU).

Jedním z hlavních cílů portálu je zajištění komunikační odborně orientované platformy pro spolupracující poskytovatele zdravotních služeb (PZS) v identifikaci rizik spojených s poskytováním zdravotních služeb, prevenci nežádoucích událostí, jejich rozpoznání a klasifikaci a jednotném posouzení na lokální úrovni, které umožní sledování a hlášení na centrální úrovni.

Portál SHNU je určený:

- o **k prezentaci metodických materiálů** s přesnou definicí pro zadávání dat hlášených událostí a jejich vypořádání (autorizované a odborně recenzované metodické pokyny)
- o **k podpoře edukačních procesů** cílových skupin pracovníků ve zdravotnictví o **sledování, hlášení a vypořádání nežádoucích událostí** (informace o seminářích, workshopech, celostátních setkáních a odborných symposiích)
- o **k posílení edukačních procesů** cílových skupin pracovníků ve zdravotnictví v **prevenci nežádoucích událostí** a zavádění bezpečnostních opatření (formou sdíleného učení) v rámci metodiky proaktivních a retroaktivních



NÁRODNÍ PORTÁL Systém hlášení nežádoucích událostí

Úvod

Metodické materiály

Srovnání dle typu NU

Srovnání dle PZS

Přihlášení

Hledat



Obecná metodika

Dekubitus(y)

Dieta/výživa

Dokumentace

Chování osob

Klinická administrativa

Klinický výkon

Medicínální plyny

Medicínské
přístroje/vybavení

Medikace / IV roztoky

Nehody a neočekávaná
zranění

Neočekávané zhoršení
klinického stavu

Pády

Transfuze / krevní deriváty

Metodické materiály / Medicínské přístroje/vybavení

Medicínské přístroje/vybavení

Technika a vybavení poskytovatelů zdravotních služeb používané lékaři/sestrami/jinými zdravotníky při poskytování péče o pacienty. Vymezení zdravotnické techniky a zdravotnických prostředků upravuje legislativní předpis (Zákon č. 268/2014). NU se tak týkají všech zákonem definovaných zdravotnických prostředků (nástroj, přístroj, zařízení, programové vybavení včetně programového vybavení určeného jeho výrobcem ke specifickému použití pro diagnostické nebo léčebné účely a nezbytného ke správnému použití zdravotnického prostředku, materiál nebo jiný předmět, určené výrobcem pro použití u člověka za účelem diagnostiky, prevence a léčby).

Spadají sem události v souvislosti s používanými zařízeními pro diagnostiku, léčbu a péči, vč. jednorázových pomůcek, jako jsou injekční stříkačky a katétrů, místa, vybavení laboratoří, zubní techniky a pomůcek pro zdravotně postižené, jako jsou postele, invalidní vozík, nosítka, chodítka a berle. Např. problém s nefunkčním dávkovačem léčiva, infuzní pumpou apod.

Metodické materiály ke stažení

Plná metodika

Zkrácená metodika

Algoritmus prevence

NÁRODNÍ REGISTR POSKYTOVATELŮ ZDRAVOTNÍCH SLUŽEB

Odborný garant: RNDr. Daniel Klimeš, Ph.D.
<http://nrpzs.uzis.cz/>



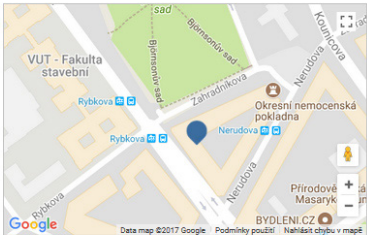
Národní registr poskytovatelů zdravotních služeb (NRPZS) poskytuje kompletní přehled o všech poskytovatelích zdravotních služeb v České republice, bez ohledu na jejich zřizovatele. Poskytuje údaje o profilu a rozsahu péče poskytované jednotlivými zdravotnickými zařízeními, obsahuje kontaktní údaje na daná zdravotnická zařízení a další podrobnější informace. Správcem a provozovatelem registru je Ústav zdravotnických informací a statistiky ČR. Zveřejňovány jsou pouze záznamy o aktuálně platných poskytovatelích, o historické a archivní záznamy je možné požádat správce registru. Poskytovatelé zdravotních služeb si mohou na portále ověřit správnost svých údajů v registru a naleznou zde i odkaz na neveřejnou část registru, kde po přihlášení mohou doplnit či opravit své kontaktní údaje a ordinační doby.

ABC alergologie s.r.o.

NÁZEV POSKYTOVATELE:
ABC alergologie s.r.o.

IČO:
29370973

ADRESA POSKYTOVÁNÍ:
Zahradníková 494/2
60200 Brno



Přehled poskytovaných zdravotních služeb

Forma péče	Obor péče
Ostatní	alergologie a klinická imunologie

Detailní záznamy Poskytovatele Rozhodnutí Detailní záznamy ÚZIS ČR

NÁZEV POSKYTOVATELE: ABC alergologie s.r.o. (právnícká osoba) IČO: 29370973

SÍDLO/MÍSTO PODNIKÁNÍ: Kupkova 712, 63800 Brno PRÁVNÍ FORMA: Společnost s ručením omezeným

Kontakty: pro poskytovatele: [editace detailu poskytovatele](#)

Oprávnění v krajích:

Název kraje	Platnost od	Platnost do
Jihomoravský kraj	1. 1. 2013	

Příjmení	Jméno	Tituly	Rozsah zastupování
Miklánková	Hana	MUDr.	
Pola	Josef	MUDr.	

Národní registr poskytovatelů zdravotních služeb

- Novinky a akce
- Informace o NRPZS
- Rada NRPZS
- Metodické dokumenty
- Kontakty
- Nápověda

Národní registr poskytovatelů zdravotních služeb (NRPZS) poskytuje kompletní přehled o všech poskytovatelích zdravotní péče v České republice. Poskytuje údaje o profilu a rozsahu péče poskytované jednotlivými zdravotnickými zařízeními, obsahuje kontaktní údaje na daná zdravotnická zařízení a další podrobnější informace.

Pro veřejnost

Pro správní orgán

Pro poskytovatele zdravotních služeb

Národní registr poskytovatelů zdravotních služeb

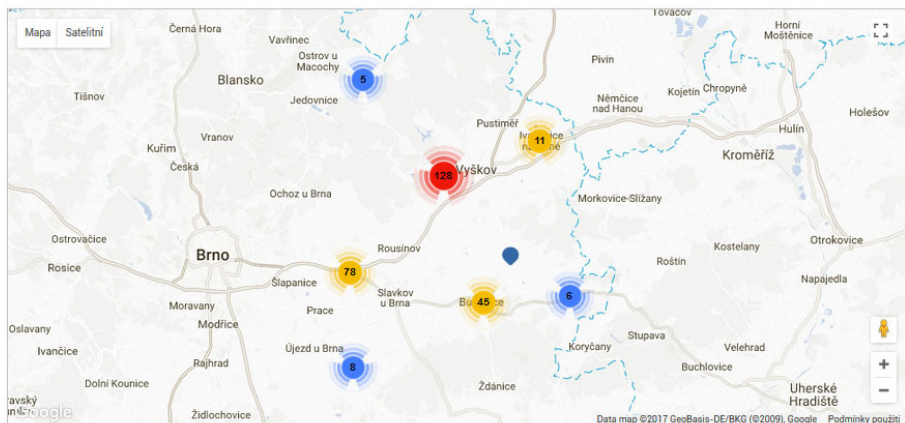
Úvod > Vyhledávání poskytovatele > **Mapa poskytovatelů** > Regionální přehledy > Analýzy a publikace

Mapa poskytovatelů

KRAJ: OKRES: OBEC:

OBOR PÉČE: FORMA PÉČE: SPECIALIZOVANÉ CENTRUM:

Jihomoravský kraj



Počet nalezených záznamů: 283

Export

méněčetné vícečetné

REGIONÁLNÍ ZPRAVODAJSTVÍ NZIS

Odborný garant: RNDr. Jan Mužík, Ph.D.

<http://reporting.uzis.cz/>



Portál regionálního zpravodajství Národního zdravotnického informačního systému (NZIS) nabízí uživatelům souhrn základních ukazatelů charakterizujících daný region (případně Českou republiku) společně se statistickými a analytickým výstupy ve formě interaktivní datových vizualizací. NZIS je jednotný celostátní informační systém, určený:

- ke zpracování údajů o zdravotním stavu obyvatelstva, o činnosti poskytovatelů a jejich ekonomice, o zdravotnických pracovnících a jiných odborných pracovnících ve zdravotnictví za účelem získání informací o rozsahu a kvalitě poskytovaných zdravotních služeb, pro řízení zdravotnictví a tvorbu zdravotní politiky,
- k vedení národních zdravotních registrů a zpracování údajů v nich vedených,
- k vedení Národního registru poskytovatelů a Národního registru zdravotnických pracovníků a zpracování údajů v nich vedených,
- k realizaci a zpracování výběrových šetření o zdravotním stavu obyvatel, o determinantách zdraví, o potřebě a spotřebě zdravotních služeb a spokojenosti s nimi a o výdajích na zdravotní služby,
- pro potřeby vědy a výzkumu v oblasti zdravotnictví,
- ke zpracování sledovaných údajů pro statistické účely a k poskytování údajů a statistických informací včetně poskytování informací pro mezinárodní instituce.

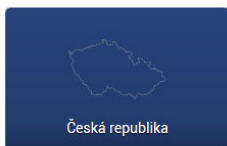
Účelem regionálního zpravodajství NZIS je poskytovat údaje o zdravotnictví a zdravotní situaci v krajích České republiky. Informace jsou podle způsobu zpracování rozděleny do těchto tří kategorií:

- Statistické výstupy:
zpravodajství formou tabulek, grafů a map
rozsahem navazuje na Zdravotnickou ročenku České republiky a krajů a Kardexy vydávané ÚZIS ČR
- Analýzy a publikace:
komentované analýzy a prezentace vybraných témat zdravotnické problematiky v krajích
přehledné prezentace indikátorů zdraví a zdravotní péče



Regionální zpravodajství NZIS Přehled regionů

Zpravodajství pro ČR



Zpravodajství pro kraje



Zpravodajství pro města

město Brno

Ústav zdravotnických informací a statistiky České republiky. Regionální zpravodajství Národního zdravotnického informačního systému [online]. Praha: ÚZIS ČR, 2016 [cit. 2017-8-23].

Dostupné z: <http://reporting.uzis.cz>.

Financováno z projektu ESF - Rozvoj technologické platformy NZIS (CZ.03.4.74/0.0/0.0/15_019/0002748)

6.1.7 Incidence – zhoubné novotvary žaludku (C16)

Tabulka

Graf

Mapa

Další výstupy

Náзор

Zpět

Název	6.1.7 Incidence – zhoubné novotvary žaludku (C16): I. Vývoj incidence v čase
Klíčová slova	nádory žaludku, incidence
Zdroj	Národní onkologický registr ČR
Pohlaví	obě pohlaví
Období	1994–2014
Věk	celkem
Region	Česká republika
Vybraná skupina:	N = 37 175

Pohlaví

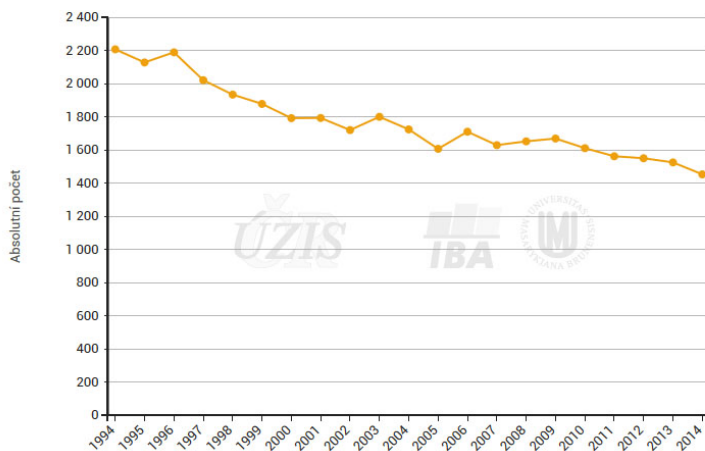
Věk

Období

Region/ČR

Nastavení
analýzy

Zobrazit tabulku



ROZVOJ TECHNOLOGICKÉ PLATFORMY NZIS

Odborný garant: doc. RNDr. Ladislav Dušek, Ph.D.
<http://reporting.uzis.cz/>



Národní zdravotnický informační systém (NZIS) je základní informační systém shromažďující dle zákona parametrické informace resortu zdravotnictví, přičemž data obsažená v NZIS jsou dále využívána pro řízení segmentu zdravotnictví. Obsahuje velké množství dat za dlouhé časové období, jejichž kvalita však nebyla nikdy systematicky prověřována. Data se často sbírala duplicitně, neefektivně, v řadě případů jsou sbírána data nepotřebná nebo dokonce zkreslená. Díky tomu je limitován správný výkon správy segmentu zdravotnictví, neboť data nemají potřebnou vypovídající schopnost a jejich sběr neadekvátně zatěžuje jednotlivé subjekty zapojené do systému reportingu. Rovněž vyhodnocení dat, nutných pro efektivní řízení segmentu zdravotnictví, je značně omezené.

Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Ústav zdravotnických informací a statistiky ČR
Institut biostatistiky a analýz LF MU
společně pracujeme

ÚZIS IBA UM

Centrum pro rozvoj technologické platformy registru Národního zdravotnického informačního systému, modernizace vytěžování jejich obsahu a rozšíření jejich informační kapacity.
Rozvoj technologické platformy NZIS – CZ.03.4.74/0.0/0.0/15_019/0002748

Rozvoj technologické platformy NZIS

Úvod **O projektu** Výstupy projektu Synergické projekty

O projektu / Cílová skupina

Cíle projektu

- Cílová skupina
- Plánované aktivity

Cílová skupina

Cílovou skupinu v projektu Rozvoj technologické platformy NZIS tvoří zejména obce a kraje, organizační složky státu včetně justice a státní příspěvkové organizace. Součástí jsou samozřejmě zaměstnanci všech zmíněných institucí.

Obce a kraje a jejich zaměstnanci

Zahrnuje všech 14 krajů včetně hl. m. Prahy, statutární města a zaměstnance zdravotnických zařízení zřízených kraji a městy. Díky projektu se zlepši efektivita kontroly a plánování těchto poskytovatelů zdravotní péče, procesy alokace finančních zdrojů do jednotlivých segmentů zdravotní péče, nastavení struktury zdravotní péče, standardizace indikátorů kvality a efektivity zdravotní péče a optimalizace sítě zdravotních služeb.

Organizační složky státu včetně justice a jejich zaměstnanci

Jedná se o Ministerstvo zdravotnictví ČR, ÚZIS a jejich zaměstnance. Díky zlepšení kvality analytických výstupů získaných z NZIS bude zajištěna vyšší efektivita kontroly a plánování poskytovatelů zdravotní péče, která se odrazí ve vyšší efektivitě výkonu státní správy a také v kvalitě poskytovaných zdravotních služeb. Výstupy projektu rovněž umožní zlepšení procesů alokace finančních zdrojů do jednotlivých segmentů zdravotní péče a nastavení struktury zdravotní péče i standardizaci indikátorů kvality a efektivity zdravotní péče a optimalizace sítě zdravotních služeb. Dále umožní postupný přechod k efektivní elektronizaci a automatizaci sběru dat v NZIS.

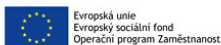
VÝBĚROVÁ ŠETŘENÍ O ZDRAVÍ

Odborný garant: RNDr. Ondřej Májek, Ph.D.

<http://ehis.uzis.cz/>



Výběrová šetření jsou nedílnou součástí zdravotnické statistiky. Jejich hlavní význam spočívá v tom, že poskytují data, která jsou z jiných zdrojů nedostupná, navíc umožňují provázat několik typů informací navzájem. Existuje celá řada výběrových šetření se zdravotní tematikou, všechny se však shodují v základních bodech a zároveň se odlišují od běžné zdravotnické statistiky. Své místo ve zdravotnické statistice však mají právě z důvodu poskytnutí chybějících dat za některé problematiky či podrobnější třídění dat za témata již pokrytá.



Ústav zdravotnických informací a statistiky ČR
Institut biostatistiky a analýz LF MU
společné pracoviště



Výběrová šetření o zdraví

cs / en

Úvod ▾ Výběrová šetření o zdraví v ČR ▾ HIS 1993–2002

HIS 1993–2002

V letech 1993–2002 probíhala výběrová šetření každé 3 roky, tzn. v letech 1993, 1996, 1999 a 2002.

Organizace šetření

Tato šetření se prováděla na vzorku populace ve věku 15 a více let s trvalým bydlištěm v ČR. Výběr respondentů byl náhodný, většinou několikastupňový, stratifikovaný, kdy se vybíraly nejprve obce, v rámci obcí pak lokality a následný výběr osob byl prováděn z Centrálního registru obyvatel.

Tazatelé v těchto šetřeních byli zaměstnanci regionálních pracovišť ÚZIS ČR, kteří byli na provádění šetření speciálně vyškoleni. Rozhovory byly prováděny osobně (face-to-face) a s respondenty byl vyplňován dotazník v papírové podobě.

Počet respondentů v jednotlivých letech šetření HIS:

1993

1 600

1996

3 396

1999

2 356

2002

1 955

Dotazník

Podkladem pro tvorbu obsahu dotazníku byly mezinárodní metodiky a zkušenosti Holandska a WHO.

Základními tématy zahrnutými v šetření byly především charakteristiky zdravotního stavu, jako jsou subjektivní hodnocení zdraví, výskyt konkrétních zdravotních potíží, emoční pohoda, výskyt dlouhodobých nemocí, dlouhodobé omezení běžných aktivit. Sledovaly se však také zdravotní determinanty jako např. sociální podpora, fyzická aktivita, kouření, konzumace alkoholu, stravovací návyky. V roce 1993 a 1996 navíc dotazník sledoval také názory populace na transformační změny ve zdravotnictví po roce 1989, hodnocení systému zdravotního pojištění a ochotu připlácet na zdravotní služby. V roce 1999 byl navíc zahrnut samostatný podrobný nástroj na úrazy v domácnosti, dále se nově sledovalo užívání drog a kvalita života.

Zpracování, vizualizace a dolování dat z pohledu rozhodovacích procesů

13.-14. září 2017, Telč

Editoři: RNDr. Martin Komenda, Ph.D., Mgr. Jakub Gregor, Ph.D.

Obálka: Mgr. BcA. Markéta Soukupová

Grafické zpracování a sazba: Mgr. Jakub Gregor, Ph.D.

Vydala Masarykova univerzita, Žerotínovo nám. 617/9, 601 77 Brno

První elektronické vydání / 2017

ISBN 978-80-210-8693-7



muni
PRESS