

Avtomatsko generiranje konceptualnih zemljevidov znanja

Jovana Otašević¹, Andrej Košir¹

¹Fakulteta za elektrotehniko, Univerza v Ljubljani
E-pošta: andrej.kosir@fe.uni-lj.si

Automatic generation of conceptual knowledge maps

Concept maps are effective visual and structural representations of information. In the field of teaching and learning, it is the base of knowledge presentation to students and, at the same time, a structural representation of knowledge for algorithms such as recommender systems of student activities. In this paper, we present a procedure for automatic concept maps of knowledge generation based on natural language processing algorithms and a comparison of the automatic and manually generated concept maps. 29 students evaluated these two maps and no statistically significant differences were found.

1 Uvod

Proces poučevanja in učenja na univerzitetnem nivoju doživlja velike in nagle spremembe. Te med drugim vključujejo spremljanje podatkov o napredku študentov skozi celoten študijski proces in tehnološko podporo poučevanju, kot so priporočilni sistemi za aktivnosti študentov. Poučevanja in učenje postajata skupni proces učitelja in študenta, delež samostojnega dela študentov se povečuje in v večji meri seli v domače okolje študenta, vsebine postajajo bolj interdisciplinarne, spreminja se tudi način ocenjevanja [1], [2].

Priporočilni sistem za aktivnosti študentov je vpeljal in ovrednotil projekt *An Intelligent System to Learn Mathematics - iMath* [3]. Gre za priporočilni sistem za aktivnosti študentov, ki jih spodbuja k samoovrednotenju znanja v celotnem poteku semestra. Eden od algoritmov za priporočanje teh aktivnosti temelji na konceptualnih zemljevidih znanja.

Brez dvoma je predstavitev znanja s konceptualnimi zemljevidi učinkovit opis strukture tega znanja [4]. Konceptualni zemljevidi lahko služijo kot osnova za samo strategijo učenja [5]. Avtomatsko generiranje konceptualnih zemljevidov znanja med drugim pomeni korak k večji učinkovitosti uporabe učiteljevega časa. Konceptualni zemljevidi so se izkazali tudi v generiranju učnih poti študentov [6], [7] in [8].

Cilj članka je ovrednotiti postopek za avtomatsko generiranje konceptualnih zemljevidov znanja z namenom uporabe v tehnološki podpori poučevanja.

2 Trenutno stanje

2.1 Konceptualni zemljevidi znanja

Konceptualni zemljevidi [9], [10] so grafični prikazi, ki vizualno organizirajo in povezujejo koncepte ter njihove medsebojne odnose. Primer konceptualnega zemljevida s področja linearne algebre je prikazan na sliki 1. Sestavljeni so iz vozlišč, ki predstavljajo ključne besede ali koncepte, in usmerjenih puščic, ki označujejo povezave med njimi, pri čemer puščice lahko vključujejo povezovalne fraze, ki opisujejo naravo odnosa. Uporabni so v učnem procesu, saj omogočajo jasno razumevanje kompleksnih odnosov, spodbujajo aktivno učenje, poglobljeno razumevanje snovi ter razvoj metakognitivnih spretnosti.

2.2 Avtomatsko generiranje konceptualnih zemljevidov

Algoritmčni pristopi so v delu že več kot 10 let. Avtorji so razvili več metod na osnovi teorije grafov [11], [12]. Te metode so nadgradili z nevronskimi mrežami [6]. Še kasneje so znatne izboljšave prinesli algoritmi procesiranja naravnih jezikov [13], [8] in veliki jezikovni modeli [14]. Možnost avtomatskega generiranja prinaša tako prihranek časa učitelju kot bolj samostojno delo študenta. Poleg tega avtomatsko generirani služijo kot osnova za tehnološko podporo učenju in poučevanju.

2.3 Orodja za avtomatsko generiranje konceptualnih zemljevidov

Za avtomatsko generiranje konceptualnih zemljevidov smo uporabili metode obdelave naravnega jezika (NLP) [15]. Postopek je vključeval čiščenje podatkov, ekstrakcijo ključnih fraz z metodo TF-IDF, analizo sintaktičnih in semantičnih povezav ter generiranje zemljevida na podlagi identificiranih konceptov in njihovih relacij.

2.4 Merjenje kvalitete konceptualnih zemljevidov

Konceptualni zemljevidi so kompleksna struktura za podajanje znanja in merjenje njene kvalitete ni preprosta naloga. Zato se pri tem zanašamo na subjektivno oceno človeških ocenjevalcev, ki oceno podajo na podlagi skrbno načrtovanih kategorij. Izbrali smo metodologijo ocenjevanja, kot jo podaja raziskava [5]. Osnova ocene so subjektivna mnenja človeških ocenjevalcev, ki so lahko

ali učitelji ali študenti. Ta metodologija predlaga 5 kriterijev, ki so: 1. Koncepti in ideje, 2. Odnosi med koncepti, 3. Hierarhija, 4. Predlogi, in 5. Črkovanje. Vsak od teh kriterijev je dovolj natančno pojasnjen. Za vsakega od kriterijev ocenjevalec izbere eno od možnosti: Odlično (20 točk), Zadovoljivo (15 točk), Povprečno (10 točk) in Potrebno izboljšati (5 točk). V omenjenem članku [5] je dovolj metodoloških opomb, da je na njihovi podlagi mogoče pripraviti ustrezna navodila za ocenjevalce. Skupno oceno dobimo tako, da dobljene točke ocenjevalca seštejemo. Veljavnost tega instrumenta ni bila posebej preverjena, zanesljivost preverjamo v vsaki študiji posebej.

3 Materiali in metode

3.1 Podatki

Testne podatke predstavljajo vsebine, na katerih smo gradili konceptualne zemljevide, ter ročno in avtomatsko zgrajena konceptualna zemljevida.

3.1.1 Vsebine za gradnjo konceptualnih zemljevidov

Za testne podatke smo izbrali vsebine poglavja linearne algebre in matričnega računa, ki so bile sestavljene v projektu iMath [3] ter vsebine, ki se za omenjeno vsebino nahajajo na Wikipedija strani [16]. Vsebina je v obliki tekstovnega opisa, brez enačb in slik. Polega samega opisa vsebin smo uporabili tudi 40 vprašanj iz te vsebine, primernih za samoovrednotenje znanja študentov, ki spoznavajo te vsebine. Vse vsebine so v angleškem jeziku, s čimer smo izločili vpliv ročnega ali avtomatskega prevajanja med slovenščino in angleščino.

3.2 Ekstrakcija ključnih besed z naravnim procesiranjem signalov

Za primerjavo konceptualnih zemljevidov smo uporabili metodologijo članka [5], glejte podpoglavje 2.4. Ta metodologija temelji na subjektivni oceni človeških ocenjevalcev, ki po predpisani metodologiji točkujejo posamezne vidike konceptualnih zemljevidov. S tem namenom smo k sodelovanju povabili 29 študentov, ki so ocenili ročno generiran in avtomatsko generiran konceptualni zemljevid iste vsebine (linearna algebra in matrični račun).

3.3 Merjenje razdalje med teksti

Na rezultat generiranja konceptualnih zemljevidov znatno vpliva način merjenja razdalje med dvema tekstovnim nizoma. Glede na raziskave [17], [18] smo v tem delu izbrali razdaljo TF-IDF. TF-IDF meri, kako pomembna je beseda v dokumentu glede na zbirko dokumentov. Postopek vključuje vektorizacijo besedila, kjer se besede pretvorijo v pomembnostne številke. TF (Term Frequency) predstavlja frekvenco pojavnosti besede v dokumentu, medtem ko IDF (Inverse Document Frequency) predstavlja redkost besede med dokumenti v korpusu. Končna ocena TF-IDF se izračuna z množenjem TF in IDF, kar omogoča identifikacijo pomembnih konceptov v besedilu.

3.4 Predlagani postopek

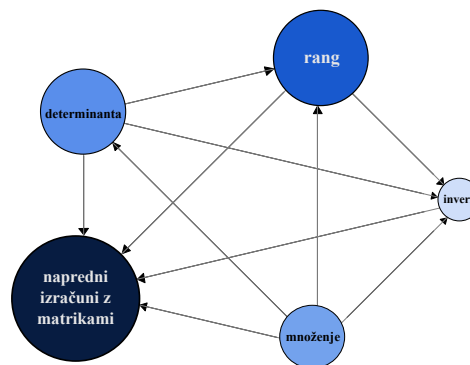
Predlagani postopek za generiranje konceptualnih zemljevidov vključuje štiri ključne korake. Prvi korak je predobdelava podatkov, kjer se besedilni podatki očistijo, odstranijo nepomembne besede in oblikuje korpus besedil. Naslednji korak je ekstrakcija konceptov, pri čemer se identificirajo ključne fraze in uporabljajo metode, kot je TF-IDF, za določanje glavnih konceptov v besedilu. Tretji korak je identifikacija relacij, kjer se analizira sintaktične in semantične povezave med koncepti za določitev njihovih medsebojnih relacij. Nazadnje, generiranje zemljevida vključuje ustvarjanje konceptualnega zemljevida na podlagi identificiranih konceptov in relacij, kar vizualno predstavlja informacije ter njihove povezave. Ta postopek omogoča ustvarjanje bogatih in informativnih konceptualnih zemljevidov, ki olajšajo razumevanje in vizualizacijo vsebine.

4 Eksperimentalni rezultati

Za primerjavo točk ocen kvalitete konceptualnih zemljevidov smo uporabili najmočnejši statistični test značilnosti glede na porazdelitev podatkov pri stopnji tveganja $\alpha = 0.05$. Ker gre za odvisne vzorce, smo izbrali t-test za odvisne vzorce in uporabili njegov neparametrični ekvivalent Wilcoxonov test znakov (dobljene razlike točk niso normalno porazdeljene). Pripadajočo aposteriori statistično moč smo določili z orodjem Gpower [19].

4.1 Ročno generiran konceptualni zemljevid

Pomembna ocena kvalitete avtomatsko generiranega konceptualnega zemljevida je njegova primerjava z ročno generiranim konceptualnim zemljevidom. Ročno generiran zemljevid, ki ga izdelava strokovnjak na zadevanem področju, še vedno smatramo kot ustrezen in tako z njim primerjamo ročno generirane konceptualne zemljevide. S tem namenom smo ročno izdelali konceptualni zemljevid izbrane učne vsebine (linearna algebra in matrični račun).

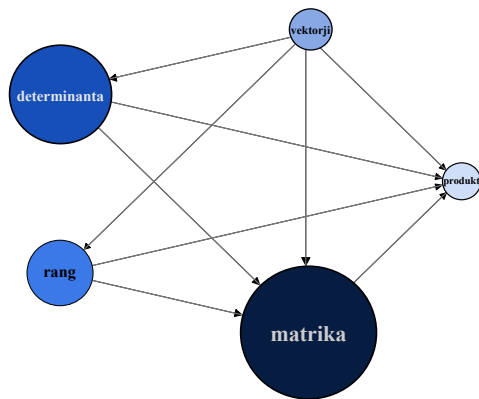


Slika 1: Ročno generiran konceptualni zemljevid poglavja linearne algebre.

4.2 Avtomatsko generiran zemljevid

Za avtomatsko generiranje konceptualnega zemljevida smo uporabili 40 vprašanj s področja linearne algebre, dopoljenih z dodatnim besedilom, da bi razširili nabor

potencialnih besed in izboljšali kakovost rezultatov. Poleg tega smo vključili tekst iz Wikipedije, ki je bil osredotočen na ključne pojme iz vprašanj. Ta kombinacija besedil je zagotovila obsežen nabor besed, iz katerega je program generiral ključne besede.



Slika 2: Avtomatsko generiran konceptualni zemljevid poglavja linearne algebre.

Sprva smo pričakovali, da bo večji obseg besed dal boljše rezultate, vendar se je izkazalo, da je bila potrebna dodatna korekcija, da bi dosegli želene rezultate. Med analizo se je namreč pojavilo veliko pogosto ponavljanih besed, kar je povzročilo, da je program prepoznal nekaj besed kot ključne, čeprav te niso prispevale k smiselni interpretaciji konceptualnega zemljevida. Zato je bilo treba prilagoditi seznam prepovedanih besed in odstraniti tiste, ki niso bile pomembne za končni zemljevid. Kljub temu pa je vključitev dodatnega besedila izboljšala kakovost ključnih besed in končnega zemljevida. Pridobili smo bogatejšo predstavitev.

4.3 Primerjava ročno in avtomatsko generiranega konceptualnega zemljevida

Ker sta oba predstavljena zemljevida zgrajena na istih vsebinah (linearna algebra), ju lahko primerjamo med sabo. Vsakega od obeh konceptualnih zemljevidov je po metodologiji iz poglavja 2.4 ocenilo $n = 29$ študentov, v okviru iste seanse v naključnem vrstnem redu.

Zanesljivost meritve kvalitete smo ocenili s koeficientom Cronbach α , ki je uporaben za enodimenzionalne konstrukte, glejte [20]. Vrednosti smo določili s Python knjižnico Pingouin [21], rezultate podaja tabela 1.

Generiranje	n	Cron. α	Int. zaup.
avtomatsko	29	0,85	[0, 74, 0, 92]
ročno	29	0,73	[0, 54, 0, 86]

Tabela 1: Koeficient zanesljivost Cronbachov α glede na način generiranja konceptualnega zemljevida, n je število ocen za posamezen konceptualni zemljevid.

Zanesljivosti obeh ocen sta različni, a obe zadovoljivo visoki glede na mejo koeficienta Cronbach $\alpha \geq 0.7$. Sklepamo, da so ocene veljavne in dovolj zanesljive za uporabo v tej študiji.

Z namenom primerjave smo testirali ničelno hipotezo $H_0 = [\text{Povprečje točk ročno in avtomatsko generiranega zemljevida je enako.}]$. Povprečje točk avtomatsko generiranega zemljevida je 85.17, ročno generiranega pa 86.38. Ker je Lillieforsov test pokazal, da točke niso normalno porazdeljene, smo uporabili Wilcoxonov test znakov in pri $n = 29$ dobili velikosti učinka $es = 0.089$ in p vrednost $p = 0.94$. Torej smo ničelno hipotezo sprejeli in razlike med zemljevidoma nismo zaznali. Dosežena statistična moč je le 0.11.

Ker je isti študent ocenil oba konceptualna zemljevida, bi lahko vrstni red ocene vplival na same točke. Zato smo s testom za razliko povprečij za odvisne vzorce testirali ničelno hipotezo, da razlik med točkami iz ocenjevanja v zaporedju "ročno, avtomatsko" in "avtomatsko, ročno" ni. Točke niso bile normalno porazdeljene in za prvo zaporedje ocenjevanja smo dobili $p = 0.52$, za drugo pa $p = 0.98$. Tako sklepamo, da zaporedje ocenjevanja ni značilno vplivalo na dodeljene ocene.

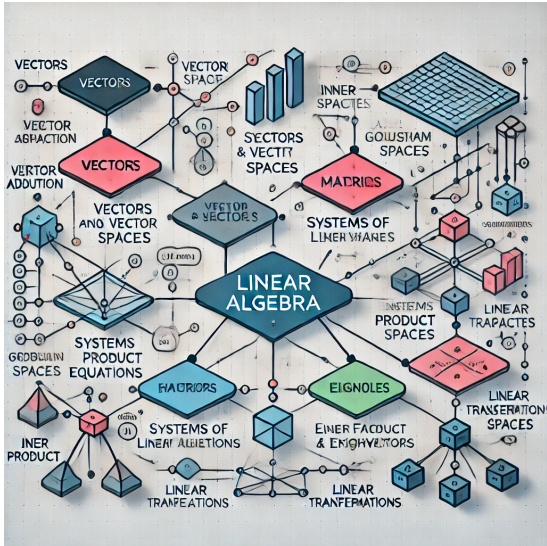
5 Zaključek in razprava

V članku smo predstavili postopek za avtomatsko generiranje konceptualnih zemljevidov znanja s pomočjo algoritmov obdelave naravnih jezikov. Primerjali smo avtomatsko generirane konceptualne zemljevide z ročno generiranimi, ki so jih izdelali človeški strokovnjaki.

Oba pristopa smo ovrednotili z metodologijo, kjer je 29 študentov ocenilo kvaliteto zemljevidov. Rezultati so pokazali, da med obema ocenjevanima zemljevidma ni bilo statistično značilnih razlik, kar nakazuje na uporabnost avtomatskega pristopa v izobraževalnem okolju. Zaradi nereprezentativne množice ocenjevalcev tega sklepa ne moremo posplošiti.

Kljub pozitivnim rezultatom smo opazili potrebo po nadaljnji prilagoditvi in izboljšanju algoritmov, zlasti pri filtriranju nepomembnih besed, da bi izboljšali interpretacijo in kakovost končnih zemljevidov. Prihodnje raziskave bi morale vključevati širše nabor podatkov in večjo skupino ocenjevalcev za pridobitev bolj reprezentativnih rezultatov. Raziskati bi bilo potrebno tudi možnosti integracije avtomatsko generiranih konceptualnih zemljevidov v različne izobraževalne tehnologije, kar bi lahko še dodatno izboljšalo učinkovitost in personalizacijo učnega procesa.

V času orodij na osnovi velikih jezikovnih modelov smo preverili sposobnost orodja ChatGPT 4.0 [22] za risanje konceptualnih zemljevidov. Rezultat za zahtevo, podano v angleškem jeziku "Can you draw me a conceptual map of Linear algebra for students" in nato z dodatkom "I would like a visual diagram of a conceptual map of Linear algebra" je prikazan na sliki 3. Opazimo, da je dobljen konceptualni zemljevid smiseln, a precej zapleten. Nakazane relacije med pojmi so ustrezne, vendar nismo uspeli nadzorovati ključnih besed, ki naj bi jih konceptualni zemljevid vključeval. Dodatne zahteve, kot je naštevanje konceptov, ki bi bili pomembni, niso izboljšale rezultatov. Poleg tega to orodje izdelava sliko, strukturo konceptualnega zemljevida pa je v obliki označenega grafa potrebno izdelati ročno.



Slika 3: Konceptualni zemljevid generiran z orodjem ChatGPT.

Naš pristop k avtomatskemu generiranju konceptualnih zemljevidov predstavlja korak naprej v smeri tehnološke podpore izobraževanju, ki lahko bistveno prispeva k izboljšanju kakovosti in dostopnosti znanja v učnem procesu. Ugotovili smo, da je predlagani pristop uporaben in obetajoč, a zahteva dodatne vložke v optimizacijo parametrov postopka in izbor ustreznih postopkov obdelave naravnih jezikov kot je razdalja med teksti.

Zahvala

Raziskavo je podprl program P2-0246 ICT4QoL - Informacijske in komunikacijske tehnologije za kakovost življenja.

Literatura

- [1] Michael Gaebel and Thérèse Zhang. *Learning and teaching in the European higher education area*. European University Association asbl, 2018. OCLC: 1142675866.
- [2] Dominic Orr, Maren Luebcke, J. Philipp Schmidt, Markus Ebner, Klaus Wannemacher, Martin Ebner, and Dieter Dohmen. *Higher Education Landscape 2030: A Trend Analysis Based on the AHEAD International Horizon Scanning*. SpringerBriefs in Education. Springer International Publishing, Cham, 2020.
- [3] iMath consortium. iMath, an intelligent system to learn mathematics, 2021. <https://imath.pixel-online.org/>, dostopano 24.07.2024.
- [4] Felix Krieglstein, Sascha Schneider, Maik Beege, and Günter Daniel Rey. How the design and complexity of concept maps influence cognitive learning processes. *Educational technology research and development*, 70(1):99–118, February 2022.
- [5] Carmen Romero, Moisés Cazorla, and Olga Buzón. Meaningful learning using concept maps as a learning strategy. *Journal of Technology and Science Education*, 7(3):313, September 2017.
- [6] Yuwen Zhou, Changqin Huang, Qintai Hu, Jia Zhu, and Yong Tang. Personalized learning full-path recommendation model based on LSTM neural networks. *Information Sciences*, 444:135–152, May 2018.
- [7] Yancong Li, Zengzhen Shao, Xiao Wang, Xuechen Zhao, and Yanhui Guo. A Concept Map-Based Learning Paths Automatic Generation Algorithm for Adaptive Learning Systems. *IEEE Access*, 7:245–255, 2019.
- [8] Vinicius Dos Santos, Érica Ferreira De Souza, Kátia Romero Felizardo, Willian Massami Watanabe, Nandanmudi Lankalapalli Vijaykumar, Sandra Maria Aluizio, and Arnaldo Cândido Júnior. Conceptual Map Creation from Natural Language Processing: a Systematic Mapping Study. *Revista Brasileira de Informática na Educação*, 27(03):150–176, December 2019.
- [9] B. Schwendimann. Concept mapping. In *Encyclopedia of the Sciences of Learning*. Springer, 2014.
- [10] T. Solonchak and S. Pesina. Concept and its Structure. *Procedia - Social and Behavioral Sciences*, 192:352–358, June 2015.
- [11] Shyi-Ming Chen and Shih-Ming Bai. Using data mining techniques to automatically construct concept maps for adaptive learning systems. *Expert Systems with Applications*, 37(6):4496–4503, June 2010.
- [12] Shyi-Ming Chen and Po-Jui Sue. Constructing concept maps for adaptive learning systems based on data mining techniques. *Expert Systems with Applications*, 40(7):2746–2755, June 2013.
- [13] Graziella Martins Caputo and Nelson Francisco Favilla Ebecken. Concept map construction applying natural language processing on text extracted from e-commerce web pages. In *2011 Third World Congress on Nature and Biologically Inspired Computing*, pages 409–414, Salamanca, Spain, October 2011. IEEE.
- [14] Wagner De A. Perin, Davidsom Cury, Camila Zacché De Aguiar, and Crediné S. De Menezes. From Text to Maps: Automated Concept Map Generation Using Fine-tuned Large Language Model. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação (SBIE 2023)*, pages 1317–1328, Brasil, November 2023. Sociedade Brasileira de Computação - SBC.
- [15] Mohana Murugan. *Natural Language Processing (NLP)*. April 2024.
- [16] Wikipedia contributors. Linear algebra — Wikipedia, the free encyclopedia, 2024. dostopano 24.07.2024.
- [17] William W. Cohen, Pradeep Ravikumar, and Stephen E. Fienberg. A comparison of string distance metrics for name-matching tasks, 2003.
- [18] Wen Zhang, Taketoshi Yoshida, and Xijin Tang. A comparative study of tf*idf, lsi and multi-words for text classification. *Expert Systems With Applications*, 38(3):2758–2765, 2011.
- [19] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2):175–191, May 2007.
- [20] Eunseong Cho. Making reliability reliable: A systematic approach to reliability coefficients. *Organizational Research Methods*, 19(4):651–682, 2016.
- [21] Raphael Vallat. Pingouin: statistics in python. *Journal of Open Source Software*, 3(31):1026, 2018.
- [22] OpenAI. Chatgpt - citation bibtext converter. <https://chat.openai.com>, 2024. Version GPT-4.