

INFCON22



Infcon 2022

Zbornik sažetaka znanstvenog skupa
doktorskih studenata informatike



ZBORNIK SAŽETAKA

1. ZNANSTVENI SKUP DOKTORSKIH
STUDENATA INFORMATIKE: INFCON22

Rijeka, studeni 2022.

Impressum

Naslov: ZBORNIK SAŽETAKA - ZNANSTVENI SKUP DOKTORSKIH STUDENATA INFORMATIKE: INFCON22

Nakladnik: Sveučilište u Rijeci, Fakultet informatike i digitalnih tehnologija

Urednici:
prof. dr. sc. Sanda Martinčić-Ipšić
izv. prof. dr. sc. Marina Ivašić-Kos
dr. sc. Slobodan Beliga

Naziv skupa: 1. ZNANSTVENI SKUP DOKTORSKIH STUDENATA INFORMATIKE:
INFCON22

Vrijeme održavanja konferencije: 11. i 12. studenoga 2022.

Mjesto održavanja konferencije: Rijeka

Učestalost izdavanja: jednom godišnje

Postavljeno na mrežu: 10. studenog 2022.

Dizajn i oblikovanje: dr. sc. Slobodan Beliga

ISBN: 978-953-7720-67-4 (online)

1. ZNANSTVENI SKUP DOKTORSKIH STUDENATA INFORMATIKE: INFCON22

Organizator skupa:

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci

Programski i znanstveni odbor:

prof. dr. sc. Sanda Martinčić-Ipšić, Sveučilište u Rijeci, Hrvatska

prof. dr. sc. Mirjana Kljajić Borštnar, Sveučilište u Mariboru, Slovenija

prof. dr. sc. Gorgji Madjarov, Sveučilište sv. Ćirila i Metoda u Skopju, Makedonija

izv. prof. dr. sc. Marina Ivašić-Kos, Sveučilište u Rijeci, Hrvatska

dr. sc. Slobodan Beliga, Sveučilište u Rijeci, Hrvatska

Organizacioni odbor:

prof. dr. sc. Sanda Martinčić-Ipšić, Sveučilište u Rijeci, Hrvatska

izv. prof. dr. sc. Marina Ivašić-Kos, Sveučilište u Rijeci, Hrvatska

dr. sc. Slobodan Beliga, Sveučilište u Rijeci, Hrvatska

dr. sc. Maja Gligora Marković, Sveučilište u Rijeci, Hrvatska

Matea Turalija, Sveučilište u Rijeci, Hrvatska

Uvodno slovo

Dragi doktorandi, mentori, sudionici i kolege!

Neizmjerno me raduje veliki odziv doktoranada na 1. Znanstvenom skupu doktorskih studenata informatike: INFCON22 koji se održava na Fakultetu informatike i digitalnih tehnologija Sveučilišta u Rijeci, 11. i 12. studenog 2022. godine.

Doktorski studenti prijavili su 20 sažetaka s pregledom trenutnih rezultata vezanih uz teme doktorskih istraživanja. Spektar tema o kojima doktorandi istražuju je izrazito širok, no ipak prevladavaju dva glavna usmjerena - umjetna inteligencija (računalni vid, računalna analiza prirodnog jezika, analiza kompleksnih mreža, strojno učenje iz senzorskih podataka) te e-učenje (igrifikacija i proširena stvarnost). Preostala istraživanja vezuju se uz izgradnju informacijskih sustava.

Zahvalujem se dr.sc. Ananada Schoembaru, s Fakulteta primjenjenih znanosti iz Utrechtu u Nizozemskoj, što se odazvao našem pozivu da održi pozvano predavanje koje odlično povezuje različite generacije doktorskih studenta i njihovih diversificiranih znanstvenih interesa i stavlja ih pod zajednički nazivnik, jednako važan za sve metodologije znanstvenih istraživanja, a uključuje dobru formulaciju istraživačkih pitanja i hipoteza. Nadam se da će smjernice istaknute na predavanju doprinijeti boljem usmjeravanju metodologije istraživanja.

Osim predstavljanja znanstvenih rezultata i područja istraživanja naših doktoranada, INFCON želi povezati različite generacije doktorskih studenta te osigurati kontinuirane susrete na kojim se razmjenjuju iskustva i ideje, započinju nove suradnje i otvaraju nove prilike za transfer tehnologija i znanja među doktorandima, industrije i zajednice.

Povezivanje i umrežavanje sa znanstvenicima u zemlji i inozemstvu važno je za razvoj znanosti i podizanje kvalitete znanstvenog rada. Zahvalujem vanjskim evaluatorima prof. dr.sc. Slobodan Ribarić sa Sveučilišta u Zagrebu te prof. dr. sc. Ljupčo Todorovski sa Sveučilišta u Ljubljani koji su kroz niz konstruktivnih savjeta, za daljnja istraživanje te motivacijom doktorandima za daljnji rad.

Zahvalu upućujem i predsjedavajućima pojedinih sekcija : prof. dr. sc. Ani Meštrović, izv. prof. Marini Ivašić Kos, dr. sc. Slobodanu Beligi te dr.sc. Maji Gligora. Naposljetu iskrena hvala i svim članovima organizacijskog i međunarodnog programskog odbora na pomoći oko organizacije i provedbe INFCON-a.

P. S. Sve organizacijske propuste planiramo popraviti zato se vidimo i dogodine!

prof. dr. sc. Sanda Martinčić-Ipšić
predsjedavateljica programskog i organizacijskog odbora INFCON22 i
voditeljica doktorskog studija Informatike

Rijeka, 5. studenog 2022.

Sadržaj

IMPRESSUM	I
UVODNO SLOVO	III
SADRŽAJ	IV
ANAND SHEOMBAR:	
THE OFFLINE & ONLINE RESEARCH JOURNEY OF THE ACADEMIC SCHOLAR: WHAT CAN PHD STUDENTS DO?	1
DINO ALJEVIĆ:	
MULTI-LABEL DOCUMENT CLASSIFICATION IN THE MESOC MATRIX	2
KRISTINA BLAŠKOVIĆ:	
DEVELOPMENT METHOD FOR REAL-TIME CLOSED-LOOP CONTROL SYSTEMS	4
AMATIJA BURIĆ:	
DIJAGNOZA BOLESTI OKA PASA I MAČAKA SEGMENTACIJOM SLIKA	6
ANTO ČABRAJA:	
COMBINE MACHINE LEARNING AND PRE-MATCH PUBLIC OPINION TO PREDICT OUTCOME IN TEAM SPORTS: THE METHODOLOGY OF DATA COLLECTION	8
VANJA ČOTIĆ POTURIĆ:	
IDENTIFIKACIJA PREDIKTIVNIH ČIMBENIKA ZA (NE)USPJEH STUDENATA U STEM ORIJENTIRANOM KOLEGIJU	11
VLATKA DAVIDOVIĆ:	
PRELIMINARY REPORT ON SEQUENCE-TO-SEQUENCE NEURAL MODEL FOR CROATIAN ABSTRACTIVE SUMMARIZATION	13
KRISTINA HOST:	
AUTOMATSKO RASPOZNAVANJE AKCIJA U SPORTU I PRIMJENA NA RUKOMETU	16
INGRID HRGA:	
EVALUACIJA SUSTAVA ZA AUTOMATSKO OPISIVANJE SLIKA	18
KARLO BABIĆ:	
PRELIMINARY RESULTS: PYRAMIDAL RECURSIVE NEURAL NETWORK	20
DEJAN LJUBOBRATOVIĆ:	
PREDVIĐANJE ZRELOSTI BRESKVI KORIŠTENjem MODELA STROJNOG UČENJA	22
GORAN PAULIN:	
PRIMJENA SINTETIZIRANE GRAFIKE ZA AUTOMATSKU DETEKCIJU U RAZLIČITIM SCENARIJIMA	24
MILAN PETROVIĆ:	
STATIC AND DYNAMIC ANALYSIS OF <i>DROSOPHILA MELANOGASTER</i> SOCIAL INTERACTION NETWORKS	27
MARKO PRIBISALIĆ:	
TOWARDS SURVEY ON AUTOMATIC GENERATION OF TEST CASES FROM SOFTWARE SPECIFICATION DOCUMENTS	28
IVAN ŠIMAC:	
IZRADA SKUPA PODATAKA ZA AUTOMATSKU DETEKCIJU PLIVAČA	32
ROMEO ŠAJINA:	
SKELETON-BASED HUMAN ACTION SCORING AND COMPARISON USING MONOCULAR CAMERA	33
IGOR ŠKORIĆ:	
MODEL VREDNOVANJA RAZVOJNIH OKRUŽENJA KOJA SE KORISTE U UČENJU PROGRAMIRANJA	35
MATEA TURALIJA:	
DESETLJEĆE RAZVOJA METODA SIMULACIJA MOLEKULSKE DINAMIKE U PARALELIZACIJI NA HETEROGENIM SUPERRAČUNALIMA	38
ANA VRCELJ:	
KORIŠTENJE IGRIFIKACIJE U OSNOVNOŠKOLSKOM I SREDNJOŠKOLSKOM OBRAZOVANJU	41
ĐURĐICA VUKIĆ:	
OKVIR ZA ANALIZU OBRAZOVNIH PROGRAMA TEMELJEN NA KOMPLEKSnim MREŽAMA	43
MARINA ŽUNIĆ:	
PRIMJENA PROŠIRENE STVARNOSTI U KOMBINACIJI S IGROM I IGRIFIKACIJOM S CILJEM RAZVOJA RAČUNALNOG RAZMIŠLJANJA U NIŽIM RAZREDIMA OSNOVNE ŠKOLE – PREGLED PODRUČJA	45



PLENARNO PREDAVANJE

Anand Sheombar

HU University of Applied Sciences in Utrecht, The Netherlands



The Offline & Online Research Journey of the Academic Scholar: What can PhD Students do?

The talk will discuss the academic research lifecycle and on which moments scholars may seek offline and online interaction to improve their research question, design, execution, analysis and knowledge dissemination. The presenter will argue that this applies to PhD students as well.

Anand Sheombar is a postdoc researcher & lecturer at the HU University of Applied Sciences in Utrecht, The Netherlands. He has a PhD in Information systems, a masters in Research methods, a master in Business Informatics, and a bachelor type of degree in Astrophysics. He is online present, for example as AnandsTweets on Twitter.



IZLAGANJA

Dino Aljević

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia

dino.aljevic@uniri.hr

Multi-label Document Classification in the MESOC Matrix

In this paper we describe multi-label document classification method which aims to classify documents in the MESOC matrix, i.e. up to three social impact domains and up to ten cultural sectors as defined by 2009 UNESCO Framework of Cultural Statistics [1]. Novelty of the method emerges from the highly specific context in which it was developed, is used and the way it tackles challenging multi-label classification problem by reducing it to multi-class classification problem. This classification is a crucial step in the analysis performed by the MESOC Toolkit [2] and attempts to describe a document within the framework of the MESOC matrix while taking into consideration the scarcity of domain-specific documents on which the model can be fitted and a high number of target classes.

The dataset consists of 573 open-access scientific papers on the topic of arts and culture retrieved from academic databases. The length of all 573 documents ranged from 859 characters or 172 word tokens to 696,182 characters or 124,976 word tokens, with an average of 49,894 characters and 8,936 word tokens per document. The dataset is unbalanced in terms of cultural sectors and societal impact domains, i.e., the classes used in the multi-label classification. To mitigate this problem, the minority classes were oversampled using the SMOTE technique [3].

Goal of the method is a multi-label classification task in 30 classes (10 cultural sectors and 3 domains of the MESOC matrix). Multi-label classification is inherently a challenging problem, but the difficulty is further compounded by the relatively small number of documents in the dataset [4] [5] [6].

To overcome this problem, we train two Random Forest classifiers [7] on a multi-class problem with the objective to differentiate between 10 cultural sectors (rows of the matrix) and 3 domains (columns of the matrix) respectively. Vectors of obtained class probabilities are multiplied together to compute final probability for each of the 30 classes. This way, multi-label classification problem is decomposed into two multi-class classification problems, one with 10 classes and the other with 3.

Formally, let $P_r \in \mathbb{R}^{10}$ be the vector of probabilities for each domain and $P_c \in \mathbb{R}^3$ be the vector of probabilities for each cultural sector. The probabilities for each of the 30 classes are then calculated as $P = P_r P_c$. Before computing the probability matrix P , we set a threshold to filter out columns and rows with low probabilities and redistribute these values evenly among the remaining classes.

The models are conditioned on TF-IDF vectors computed from keywords which can be manually annotated or extracted from input documents using YAKE algorithm [8]. Before computing TF-IDF values, input text is preprocessed in order to removing unnecessary tokens, e.g. common abbreviations, hyperlinks, email addresses, phone numbers and unprintable characters and so on. Additionally, input text undergoes tokenization, removal of stop words, normalization and lemmatization. Results are evaluated by 10-fold cross validation using F1 macro average as an evaluation metric.

Keywords – *multi-label classification, text classification, random forest, natural language processing*

References

- [1] UNESCO Framework for Cultural Statistics (FCS) (2009).
<http://uis.unesco.org/sites/default/files/documents/measuring-cultural-participation-2009unesco-framework-for-cultural-statistics-handbook-2-2012-en.pdf>
- [2] P. K. Bogović, D. Aljević, B. Kovačić and S. Martinčić-Ipšić, "The NLP Powered BI Toolkit: The Case of MESOC," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), 2022, pp. 1191-1196, doi: 10.23919/MIPRO55190.2022.9803434.
- [3] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique" Journal of Artificial Intelligence Research 16 321–357. 2002. doi: <http://dx.doi.org/10.1613/jair.953>
- [4] D. Jurafsky and J. H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics. 2nd edition. Prentice-Hall. 2009.
- [5] C.D. Manning, P. Raghavan and H. Schütze, "Introduction to Information Retrieval"; Cambridge University Press: New York, NY, USA, 2008.
- [6] S. Martinčić-Ipšić, T. Miličić and L. Todorovski, The Influence of Feature Representation of Text on the Performance of Document Classification. Applied Science. 9, 743. 2019.
- [7] L. Breiman, Random forests. Machine learning 45.1: 5-32. 2001.
- [8] R. Campos, V. Mangaravite, A. Pasquali, A. Jorge, C. Nunes and A. Jatowt, "YAKE! Keyword extraction from single documents using multiple local features." Information Sciences 509 257-289.2020.

Kristina Blašković

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
kikablaskovic@gmail.com

Development Method for Real-time Closed-loop Control Systems

Embedded systems are part of our daily life. Their complexity arises from the tight connection between software and hardware parts of the system. For this reason, the development process of these systems must consider software and hardware specifics simultaneously. Real-time closed-loop control system [1], [6] is a special case of embedded systems.

Modern trends in the marine industry imply the increasing presence of hybrid and electric propulsion systems that comply with current and future environmental regulations. These environmentally friendly solutions are usually based on premium efficiency alternating current (AC) synchronous electric motors and variable speed generators in permanent magnet technology. Since these are electronically commutated machines, they must be coupled with very specific accompanying electronics that ensures optimal functionality of the system.

Development of this special case of a real-time closed-loop control system in our research focus implies development of an application-specific hardware part on which management software with extremely compact code, high execution speed and code modularity is implemented. To meet the defined requirements and achieve full compatibility of software and hardware, a methodological approach for the development of these systems must be proposed that unifies the development steps for both software and hardware parts of the system.

As a first step, a systematic literature review of existing development methodologies for embedded systems has been conducted [4]. On obtained results cluster analysis was performed according to the following criteria: coverage/scope, type of contribution, thematic focus, and implications criterion. Thereat, k-means algorithm was used, where the distance was determined by the squared Euclidian distance formula. This cluster analysis resulted in three clusters, two of which were identified as significant, based on the number of elements/articles. The results of this cluster analysis will serve as a framework for identifying desirable characteristics which will provide guidelines for defining new development methodologies. Additionally, to gain a deeper understanding of the relationship between the defined classification criteria, relational and non-relational data model was created and populated using the Naive-Bayes algorithm in the validation process [5]. The data model was graphically represented by a multilayer conceptual network, showing the connections between the criteria of the adjacent layers. To identify further causalities in the criteria relationships, specific queries were defined for each data model.

Based on the obtained research results, a first version of a software development methodology for special case of real-time closed-loop control systems (DEM4RTS) [3] has been proposed, aiming at meeting the defined specific software requirements, ensuring software and hardware compatibility, providing an adequate testing environment, and ensuring a well-documented development process with good traceability.

The defined methodology was further used in process of development of Power Management System (PMS) for electric and hybrid marine vessels [2]. In complex marine propulsion systems PMS collects and interprets system component data and based on

closed-loop reaction in feedback manner adjusts parameters in real-time for optimal system functionality.

By following the steps defined by DEM4RTS, in development process of the PMS, ad-hoc approach was replaced by systematic activities. Furthermore, full software and hardware compatibility was achieved by establishing parallel and intertwined software and hardware development processes. Additionally, better team collaboration is achieved, followed by a better documented development process and traceability. The limitations of the proposed methodology are that it focuses primarily on the development of software part of the system and does not fully cover the development process of the application-specific hardware.

Future work should define a similar compatible systematization method for the hardware development process, analyze the impact of the DEM4RTS methodology on the development of a real-time closed-loop control system, and propose and develop an information system to support the DEM4RTS methodology.

Keywords – real-time closed loop control system, embedded system, development methodology

References

- [1] T. Barbera, M. L. Fitzgerald, and J.S. Albus, "Concepts for a real-time sensory-interactive control system architecture", In Fourteenth Southeastern Symposium on System Theory, April 1982.
- [2] K. Blašković, "Power Management System For Electric And Hybrid Marine Vessels," 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), pp. 902–907, 2021
- [3] K. Blašković, and S. Čandrić, "DEM4RTS: Software Development Methodology for Special Case of Real-Time Closed-Loop Control Systems," Proceedings of the 29th DAAAM International Symposium, pp. 993–1000, 2018.
- [4] K. Blašković, S. Čandrić, and A. Jakupović, "Systematic Review of Methodologies for the Development of Embedded Systems," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 12(1), pp. 410–420, 2021.
- [5] K. Blašković, S. Čandrić, A. Jakupović, M. Ašenbrener Katić, and P. Poščić, "Object-oriented data model for the classification of development methodologies for embedded systems," unpublished
- [6] R. Quintero, and T. Barbera, "A Real-Time Control System Methodology for Developing Intelligent Control Systems", No. NIST Interagency/Internal Report (NISTIR)-4936, 1992.

Matija Burić

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
matija.buric@hep.hr

Dijagnoza bolesti oka pasa i mačaka segmentacijom slika

Kućni ljubimci, posebno psi i mačke, imali su važnu ulogu u ljudskoj povijesti radi zaštite koju su pružali ljudima i pomoći u lovu, a suvremenom životu se sve više tretiraju kao članovi obitelji i dobivaju puno veću ljubav i pažnju. Zbog privrženog odnosa prema psima i mačkama, sve se više pažnje pridaje njihovom zdravlju pa se paralelno s tim razvija sve bolja dijagnostika i sve kompleksniji medicinski tretmani i liječenje.

Svrha ovog istraživanja je postići što bržu i točniju dijagnostiku bolesti očiju pasa i mačaka koja će biti dostupna njihovim vlasnicima i veterinarima. Iako postoje istraživanja koja su se bavila analizom oka kod kućnih ljubimaca [1] [2] ona nisu fokusirana na bolesti oka prvenstveno zbog nedostatka označenih podataka od strane liječnika koji bi se mogli koristiti za učenje modela za automatsku detekciju promjena na oku i bolesti. Za određivanje bolesnog područja i segmentaciju slika u posljednje vrijeme se uobičajeno koriste varijante duboke neuronske mreže U-Net [3] koje su pokazale i dobre rezultate u medicinskoj obradi slika [4] [5]. Za treniranje U-Net modela koriste se segmentirane slike objekata i njihove maske na kojima se svakom pikselu definira klasa kojem pripada. Rezultat mreže je segmentirano područje gdje se svakom pikselu dodjeljuje vrijednost klase što se opet može prezentirati maskom za lakšu vizualnu evaluaciju rezultata.

U provedenom istraživanju biti će obuhvaćeno 4 bolesti (klase) i 150 ulaznih slika. Distribucija slika po klasama nije jednaka tako da odražava učestalost pojavljivanja bolesti. Kako bi se dobili što bolji rezultati U-Net modelom napravljena je augmentacija skupa podataka te su određeni parametri modela i učenja koji daju najbolje rezultate.

Evaluacija modela je izvedena korištenjem IoU (Intersection-over-Union) metrike. Postignuti rezultati su dosegnuli srednju vrijednosti od 84.1% za 4 klase pri čemu najbolje prepoznata klasa ima 98.3% IoU a najlošije prepoznata klasa doseže maksimum od 37.9%. Dobiveni rezultati su obećavajući i za očekivati je da bi se istom metodom mogli dobiti bolji rezultati kada bi se povećao broj slika u skupu za učenje i podesili parametri neuronske mreže.

U konačnici cilj je da se definira što pouzdaniji model koji bi pomagao u otkrivanju ranog stadija bolesti oka kod pasa i mačaka i davao korisne savjete vlasnicima pasa vezano za dijagnozu i tretiranje bolesti te bio pomoćni alat stručnom osoblju za trijažu bolesnih životinja.

Ključne riječi – dijagnoza bolesti očiju, oftalmologija pasa i mačaka, U-Net, augmentacija, IoU metrika

Literatura

- [1] S. Park, K. Holmqvist, D. Niehorster, L. Huber i Z. Virányi, "How to improve data quality in dog eye tracking," Behavior Research Methods, Springer, doi.org/10.3758/s13428-022-01788-6, 2022.
- [2] S. Park, C. Bacelar i K. Holmquist, "Dog eye movements are slower than human eye movements," National Library of Medicine, svez. doi: 10.16910/jemr.12.8.4. PMID: 33828775; PMCID: PMC7881887., 2020.

- [3] O. Ronneberger, P. Fisher i B. Brox, "U-Net: Convolutional networks for biomedical image segmentation," Conf. on Medical Image Computing & Computer-assisted Intervention, svez. Springer, p. pp. 234–241, 2015.
- [4] X. Yin, L. Sun, Y. Fu, R. Lu i Y. Zhan, "U-Net-Based Medical Image Segmentation," National Libray of Medicine, svez. doi: 10.1155/2022/4189781. PMID: 35463660; PMCID: PMC9033381., 2022.
- [5] H. Lu, Y. She, J. Tie i S. Xu, "Half-UNet: A Simplified U-Net Architecture for Medical Image Segmentation," Frontiers in Neuroinformatics, svez. <https://doi.org/10.3389/fninf.2022.911679>, 2022.

Anto Čabralja

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia
anto.cabralja@student.uniri.hr

Combine Machine Learning and Pre-match Public Opinion to Predict Outcome in Team Sports: The Methodology of Data Collection

In the domain of sports, we come across a large number of variables that need to be considered to obtain a quality prediction model. The problem becomes even more sensitive in the field of team sports because we have a new scope of descriptive variables related to the condition of the players, the status of the club, and interpersonal relationships. This presents a few significant computational challenges for existing and well-known machine learning techniques.

Predictions are the most attractive area of research from a market and business perspective described by Beal and Middleton [1]. There are a significant number of reported research about different machine learning techniques used to predict match outcomes or strategies to win. We can even find benchmarks for the accuracy of those predictions also presented in [2] but majority of conducted research is focused on statistics and machine learning based on historical data for individual teams.

The goal of our research is to validate if we can improve the accuracy of the traditional machine learning models for match outcomes of team sports by involving public and expert opinion. Still, this is quite a big scope to make a valuable first conclusion that's why we had to focus on one prediction type, one sport, and one top-level league. To be more precise in our research we are focusing on pre-match public opinion in football for one of the most popular football leagues, the English Premier League. The reason why research is primarily focused on pre-match analysis is that it's the most attractive problem for which we expect a lot of data available since fans, relevant news, and sports analysts comment before the match. Football attracted our attention because of the 54% accuracy of prediction models according to benchmarks [2]. This is much lower than any other popular team sport such as basketball, handball, or baseball according to [2]. One reasonable explanation why the accuracy percentage is lower in comparison to the other sports is the fact that in football we have a significant number of draw results. English Premier League was profiled as a perfect candidate because of the English-speaking area for which we can find data set and predefined models to validate our hypothesis. In this research we are explaining methodology and reason of using Twitter as the data source of public opinion.

Twitter has been used to predict and explain a various of real-world scenarios, including opinion pools [3], elections [4] and the stock market [5]. It's as a strong indicator that the twitter data, especially messages, contains time series information that can be combined with statistical methods. We are analyzing tweets in period between 2016 and 2022. The reason why this period is ideal for our research is because of several important events. First, we had World Cup 2018, after that COVID-19 concur the world, then we had Euro Cup 2020 and finally in 2022 we have a new World Cup 2022 in winter season which is the first time in the history for European football.

Our goal is to extract pre-match comments written in English and that requires advanced Twitter querying, we decide to use twitter v2 API which allows us to build very robust and customized queries documented in [6].

Our research relies on finding relevant tweets and assigning them to the game during the season. To ensure quality we consider the top 5 teams form the English Premier League: Manchester United (Man Utd), Chelsea (Che), Liverpool (Liv), Manchester City

(Man City), and Arsenal (Ars) according to [7] but also confirmed from various web sources. Each of these teams has more than 10M fans on Twitter also evaluated in [7].

We can easily assign tweets with the match time by the timestamp of the tweet but linking them to the team is a more difficult task. One simple approach is linking fan accounts. According to our research, this method didn't show a significant result because of rivalry between teams and subjectivity. That's why we chose a simple high-precision approach based on hashtags and keywords. We manually created a keywords pool associated with the teams and after that, we did key extraction on the obtained results to update the keywords pool. We did the focus on the comments associated at least with one of the teams rather than comments generally talking about the match. This way we created a truly featured base on tweets focused on teams. As a result of the analysis, we discovered the following patterns for the hashtag keywords as the most valuable:

```
#{teamName}{matchDay}, #{teamName1}{teamName2}, #{teamName}{PL},  
#{teamName}{win}, #{teamName}{vote}
```

Filtering was not complete just by matching perfect hashtags or their combinations because there is a lot of noises and fake news as well. Luckily, using Twitter negations explained at [6] and key extraction we were able to detect fake news terms and build better optimized queries.

Query example: "Arsenal matchday premier league -funzone -fans -news -is:retweet lang:en"

This approach helps us increase tweets quality for 15% and reduce number of non-relevant data for 20% which is presented in the following table:

# 100 000 tweets	Quality Tweets	Non-relevant	Unsorted
Before High Quality Extraction	~ 60%	~ 30%	~ 10%
After High Quality Extraction	~ 75%	~ 10%	~ 15%

We presented methodology how to build strategy over the parameters by researching the field of interest and combining them with the powerful API like Twitter API. We were focused on creating a powerful filter pattern that helps us extract high quality dataset for our future research with the minimum noises and outliers. As we have limited resource access on APIs this present one of the most challenging tasks in any researches especially when we had to extract data from one of the largest data source collections.

The next step in our research is to build a classifier to detect tweets that contain the opinion about the match outcome. We plan to train and compare the results of several binary classifiers based on different algorithms and techniques such as Naive Bayes, SVM, and Random Forest.

Keywords – public opinion, team sports, twitter data, key extraction, collection optimization

References

- [1] R.J. Beal, S. Middleton, T. Norman and S.D. Ramchurn, "Combining Machine Learning and Human Experts to Predict Match Outcomes in Football: A Baseline Model", The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21), pp. 15447-51, Feb. 2021
- [2] R.J. Beal, T. Norman and S.D. Ramchurn, "Artificial intelligence for team sports: a survey", The Knowledge Engineering Review 34, 2019

- [3] O'Connor, Brendan, et al. "From tweets to polls: Linking text sentiment to public opinion time series.", Fourth international AAAI conference on weblogs and social media, 2010
- [4] Tumasjan, Andranik, et al. "Predicting elections with twitter: What 140 characters reveal about political sentiment.", Proceedings of the International AAAI Conference on Web and Social Media. Vol. 4. No. 1., 2010
- [5] Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter mood predicts the stock market." Journal of computational science 2.1, pp. 1-8, 2011
- [6] Twitter Developer Platform, Building Queries, [Online], 2022, Available: <https://developer.twitter.com/en/docs/twitter-api/tweets/counts/integrate/build-a-query>
- [7] FootGoal, Top 10 Premier League Football Clubs With the Most Fans in the World, 2022, <https://footgoal.pro/top-10-premier-league-football-clubs-with-most-fans-in-the-world-in-2021>

Vanja Čotić Poturić

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
vcotic@uniri.hr

Identifikacija prediktivnih čimbenika za (ne)uspjeh studenata u STEM orientiranom kolegiju

Kolegij Matematika 2 izvodi se u ljetnom semestru prve godine preddiplomskih studija brodogradnje, elektrotehnike, računarstva i strojarstva na Tehničkom fakultetu Sveučilišta u Rijeci. Aktivnosti poput pružanja nastavnih materijala, dodatnih video materijala i testova za uvježbavanje gradiva odvijaju se u sklopu e-inačice ovog kolegija sustava za e-učenje Merlin. U izvješćima sustava Merlin dostupni su podaci o aktivnostima studenata u e-kolegiju poput podataka temeljenih na klikovima koji opisuju da li, kada i koliko često studenti pristupaju resursima. Ta su izvješća prvenstveno deskriptivne prirode, govore sudionicima što se dogodilo, ali ne i zašto se dogodilo te ne predviđaju ishode niti savjetuju sudionike kako poboljšati svoje rezultate.

Ovim problemima bavi se područje rudarenja obrazovnih podataka (engl. Educational Data Mining, EDM). Cilj EDM-a je poboljšati nastavu i proces učenja poboljšanjem procesa vrednovanja, razumijevanjem obrazovnih problema i planiranja intervencija [1] te koristi tehnike i metode kao što su klasifikacija, grupiranje, regresija i vizualizacija. EDM se koristi za prepoznavanje izazova učenja, za proučavanje i predviđanje uspješnosti učenika [2], [3], [4] te za evaluaciju integracije tehnologija u proces učenja [5].

Prediktivne (nadzirane) metode rudarenja podataka predviđaju vrijednosti izlaznih varijabli na temelju ulaznih podataka. Model se razvija pomoću podataka za obuku u kojima su definirane vrijednosti ulaznih i izlaznih varijabli. Model generalizira odnos između ulaznih i izlaznih varijabli i koristi ga za predviđanje drugih skupova podataka gdje su poznati samo ulazni podaci [6]. Dva glavna modela nadziranih metoda rudarenja podataka su klasifikacijski i regresijski model. Klasifikacijski algoritmi određuju kojoj od unaprijed definiranih kategorija pripadaju izlazni podaci, dok je zadatak regresijskih algoritama predvidjeti numeričku vrijednost izlazne varijable.

U ovom radu koriste se regresijski i klasifikacijski algoritmi s ciljem predikcije studenata koji su u opasnosti da neće položiti kolegij Matematika 2. Želi se odrediti model pomoću kojeg bi se u narednim akademskim godinama čim ranije u semestru mogli identificirati studente koji su u opasnosti da neće položiti kolegij kako bi im navrijeme mogla biti ponuđena određena pedagoška pomoć.

Kolegij Matematika 2 traje 15 tjedana, a prvi kolokvij polaže se u sedmom tjednu. Broj bodova postignutih na kolokviju je izlazna (ciljna) varijabla. Izlazna varijabla je kontinuirane vrijednosti i stoga može poslužiti samo kao ciljna varijabla za regresijski problem. Za klasifikacijski problem definira se dihotomna (0 ili 1) ciljna varijabla. Rizičnim studentima smatraju se oni studenti koji su na kolokviju postigli 40% bodova ili manje i pridružuje im se vrijednost 0, dok se ostalim studentima pridružuje vrijednost 1. Ulazne varijable su podaci temeljeni na klikovima koji su prikupljeni iz Merlin sustava u četvrtom nastavnom tjednu. Prikupljeno je šest različitih podataka (broj preuzetih nastavnih materijala i materijala za samostalno učenje, bodovi ostvareni na tjednim kvizovima, broj pogledanih video lekcija, ukupna aktivnost studenta na e-kolegiju te broj bodova na kolegiju Matematika 1) za 249 studenata.

Podaci su podijeljeni u omjeru 80:20 u trening set (200 uzoraka) i test set (49 uzoraka), a za zadatke prediktivne analize korišten je paket Caret programskog jezika R.

Za predikciju broja bodova na kolokviju korišten je generalizirani linearni model, dok je za klasifikacijski problem korištena logistička regresija. Bolji rezultati dobiveni su logističkom regresijom, točnost modela iznosi 86%, osjetljivost je 87%, specifičnost 85% te AUC 0.85.

Brojni autori bave se ovom problematikom. Autori rada [7] koriste logističku regresiju za predviđanje neuspjeha studenata u matematičkom kolegiju preddiplomskog studija. Predstavljaju tri logistička regresijska modela koji su postigli preko 70% točnosti, preko 70% osjetljivosti, preko 60% specifičnosti i preko 0,75 AUC. Autori rada [8] naglašavaju da interpretabilnost modela postaje sve važnija pri radu s velikim skupovima podataka i složenim modelima. Ističu da su izvedene analitike i modeli prihvativi i vjerodostojni samo ako su interpretabilni.

Ovo istraživanje pokazalo je da su značajni prediktori prolaznosti kolegija Matematika 2 uspjeh u prethodnom kolegiju Matematika 1, redovitost rada izražena rješavanjem tjednih testova, preuzimanje nastavnih materijala i gledanje video lekcija. Najveći utjecaj na prolaznost kolegija Matematika 2 ima uspjeh na kolegiju Matematika 1. Drugi važan aspekt je redovitost rada, odnosno rješavanje tjednih testova, a tu informaciju treba podijeliti sa studentima na početku nastave kako bi ih potaknuli na redovit rad. U kontekstu mogućih pedagoških intervencija najvažniji prediktori su bodovi postignuti na tjednim testovima, broj preuzetih nastavnih materijala i broj pogledanih video lekcija, jer se na njih može utjecati tijekom nastave. Drugim riječima, redovito praćenje ove tri varijable omogućilo bi individualizirane pedagoške intervencije za koje se može pretpostaviti da će povećati stopu uspješnog polaganja kolegija Matematika 2.

Ključne riječi – generalizirani linearni model, logistička regresija, predikcija, (ne)uspjeh

Istraživanje "Identification of predictive factors for student failure in stem oriented course" autora V. Čotić Poturić, I. Dražić i S. Čandrić bit će predstavljeno na konferenciji ICERI2022 koja će se održati od 7.-9. studenog 2022. u Sevilli te objavljeno u ICERI2022 zborniku.

Literatura

- [1] G. Siemens, R. S. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in ACM International Conference Proceeding Series, 2012.
- [2] R. Asif, A. Merceron, S. A. Ali, N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," Computers&Education, pp. 177-194, 2017.
- [3] D. Gasevic, C. Rose, G. Siemens, A. Wolff, Z. Zdrahal, "Learning analytics and machine learning," in Fourth International Conference on learning analytics and knowledge, 2014.
- [4] G. Kostopoulos, S. Kotsiantis, C. Pierrakeas, G. Koutsonikos, G. A. Gravvanis, "Forecasting students' success in an open university," Learning Technology, vol. 1, 2018.
- [5] C. Angeli, S. K. Howard, J. Ma, J. Yang, P. A. Kirschner, "Data mining in educational technology classroom research: can it make a contribution?," Computers&Education, pp. 226-242, 2017.
- [6] H. Witten, E. Frank, M. A. Hall, C. J. Pal, "Part II: More advanced machine learning schemes," in Data mining: Practical machine learning tools and techniques, 4 ed., Morgan Kaufmann, 2017.
- [7] S. F. Costa, M.M. Diniz,, "Application of logistic regression to predict the failure of students in subjects of a mathematics undergraduate course," Education and Information Technologies, 2022.
- [8] M. Chitti, P. Chitti and M. Jayabalan, "Need for Interpretable Student Performance Prediction," 13th International Conference on Developments in eSystems Engineering (DeSE), pp. 269-272, 2020

Vlatka Davidović

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia
Polytechnic of Rijeka, Rijeka, Croatia
v david@veleri.hr

Preliminary Report on Sequence-to-Sequence Neural Model for Croatian Abstractive Summarization

Abstractive text summarization task generates short report from document so that main information and meaning of the text is preserved. Nowadays, state of the art in abstractive text summarization uses deep learning methods to create summary. Two main deep neural architectures are: sequence-to-sequence model and transformer-based models.

Sequence-to-sequence architecture relies on recurrent neural networks (RNN) to predict next tokens in the sequence based on previous sequences. In RNN each step incorporates knowledge of the previous input, adds a new knowledge and generates output. Sequence-to-sequence models encountered two problems: how to capture long-term dependencies between sequences and how to convert input sequences to output, when they are different length. RNN model with Long Short-Term Memory [1] and Gated Recurrent Unit architectures [2] became capable of learning from much longer sequences, by adding the gates that decides which information will be kept, which will be sent as output and which will be forgot. The sequence-to-sequence architecture advanced to encoder-decoder architecture [3], [2] consisting of two networks: the first one encodes the input sequence into fixed length vector representation, and the second decodes the vector into output sequence. Using fixed-length vector was a bottleneck in encoder-decoder performance. It decreases rapidly as the length of the input sentence increases [4]. The solution is introduction of the attention mechanism [5] that attends a part of a source sequence with the most relevant information. Further small improvement of the architecture came with the large vocabulary ‘trick’, which restricts vocabulary to the most frequent words [6].

The second architecture - transformer was a big breakthrough in text generation tasks [7]. It uses parallelization to speed up processing of the large amount of data and includes the attention mechanism. Most of the transformer models work in two stages: the first is generative unsupervised pre-training on a very large amount of data (i.e. trains the language model) and the second is discriminative fine-tuning on a much smaller supervised datasets to solve specific task (e.g. sentiment prediction, summarization, etc.). GPT, GPT 2 and 3 [8]-[10] models utilize decoder only layers to predict next word. While GPT model use pre-training and fine-tuning, GPT 2 and 3 learn tasks from the raw text, without task-specific training data. Encode-decoder transformers like BART [11], XSum [12], PEGASUS [13] achieve good results in generating abstractive summaries. BigBird [14] has a sparse attention mechanism with capability to handle long context. Transformer models perform well on large amount of texts, but they require a lot of resources for training which is their main drawback.

In this preliminary report we use sequence-to-sequence architecture: Encoder and decoder map a source sequence into the target sequence. First, encoder takes information from the source sequence and captures contextual information in a single vector - the context vector. The context vector is passed to the decoder to produce the output sequence. Model has two phases: training and inference phase. During training phase encoder reads the input sequence one token at a time, encapsulates information and store it in its final internal states (context vector): hidden state h and cell state c . Then internal

states are passed to decoder. Decoder reads the target sequence one token at a time and predicts next token y_{t+1} given the previous one y_t and internal states from encoder h, c . Output of decoder is token with maximum probability. Inference phase is phase of testing model on new source sequences.

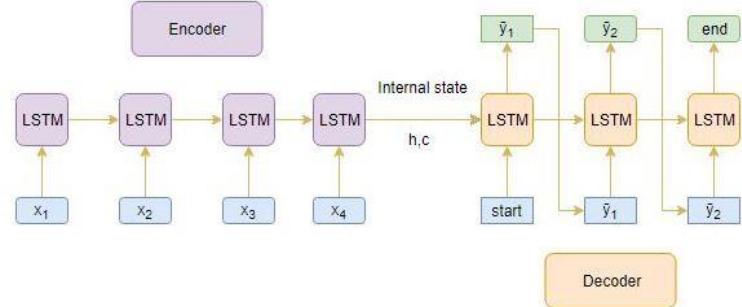


Figure 1 Encoder- Decoder architecture [15]

The performance of the encoder-decoder deteriorates rapidly as the length of an input sentence increases. To overcome that problem, solution is to add attention mechanism which looks only at specific part of the input sequence only, rather than the entire sequence. Attention mechanism gets information from all the hidden states and it is processed to get attention score. Attention score highlight important information to make a prediction. At each prediction, produced hidden state is fed back into the context vector.

Preliminary research on abstractive summarization for Croatian language reported here, uses sequence-to-sequence model. The model consists of LSTM layers and attention mechanism. The used dataset consists of bachelor theses from Croatian higher education institutions, available through open access and collected from DABAR (Digital Academic Archives and Repositories) [16]. Texts are retrieved in pdf format and converted into plain text format. Each text is divided into two parts: the first part includes the text body from theses without front pages, titles, authors, contents and literature - used for training; the second part captures only abstract with keywords- used for testing. The model is trained and tested on very small dataset, with texts that consists of different vocabulary sizes. Preliminary results are very limited and show that further research will require larger dataset and more epochs during training. The results are calculated with ROUGE evaluation metrics.

Keywords - abstractive summarization, sequence-to-sequence, Croatian dataset

References

- [1] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in Neural Computation, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [2] K. Cho et al., "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation", Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014. Available: 10.3115/v1/d14-1179
- [3] I. Sutskever, O. Vinyals, and Q. V. Le, 'Sequence to Sequence Learning with Neural Networks', in Advances in Neural Information Processing Systems, 2014, vol. 27.
- [4] K. Cho, B. van Merriënboer, D. Bahdanau and Y. Bengio, 'On the Properties of Neural Machine Translation: Encoder-Decoder Approaches', Proceedings of SSST 2014 - 8th Workshop on Syntax, Semantics and Structure in Statistical Translation, pp. 103-111., 2014., Doha; Qatar; 25 October 2014

- [5] D. Bahdanau, K. Cho and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate", 3rd International Conference on Learning Representations, ICLR 2015, 2014. [Online]. Available: <https://arxiv.org/abs/1409.0473>.
- [6] R. Nallapati, B. Zhou, C. dos Santos, Ç. Gülcühre, and B. Xiang, 'Abstractive Text Summarization using Sequence-to-sequence RNNs and Beyond', in Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning, 2016, pp. 280–290.
- [7] A. Vaswani et al., "Attention is all you need," Advances in neural information processing systems, vol. 30, 2017.
- [8] A. Radford and K. Narasimhan, 'Improving Language Understanding by Generative Pre-Training', 2018. [Online]. Available: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf. [Accessed: 2-Oct-2022].
- [9] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, 'Language Models are Unsupervised Multitask Learners', OpenAI blog, vol. 1, no. 8, p. 9, 2019.
- [10] T. Brown et al., 'Language models are few-shot learners', Advances in neural information processing systems, vol. 33, pp. 1877–1901, 2020.
- [11] M. Lewis et al., 'BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension', in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 7871–7880.
- [12] S. Narayan, S. B. Cohen, and M. Lapata, 'Don't Give Me the Details, Just the Summary! Topic-Aware Convolutional Neural Networks for Extreme Summarization', in EMNLP, 2018, pp. 1797–1807.
- [13] J. Zhang, Y. Zhao, M. Saleh, and P. Liu, 'Pegasus: Pre-training with extracted gap-sentences for abstractive summarization', in International Conference on Machine Learning, 2020, pp. 11328–11339.
- [14] M. Zaheer et al., "Big Bird: Transformers for Longer Sequences," in Advances in Neural Information Processing Systems, 2020, vol. 33, pp. 17283–17297. [Online]. Available: <https://proceedings.neurips.cc/paper/2020/file/c8512d142a2d849725f31a9a7a361ab9-Paper.pdf>
- [15] Adarsh, "Text Summarization with Attention based Networks", Medium.com, <https://cdn.analyticsvidhya.com/wp-content/uploads/2019/05/82.jpg> (accessed 2.10.2022.)
- [16] DABAR repository, Digital Academic Archives and Repositories, 2020, [Online]. Available: <https://dabar.srce.hr/browse/field>

Kristina Host

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
kristina.host@inf.uniri.hr

Automatsko raspoznavanje akcija u sportu i primjena na rukometu

Automatsko raspoznavanje akcija (eng. *Human action recognition* - HAR) koristi se u raznim sportovima poput odbojke, košarke ili tenisa kako bi se pratilo performanse igrača tijekom treninga, raznih priprema, zagrijavanja ili utakmica [1].

Osnovna ideja HAR-a je detekcija igrača na sceni, praćenje njihovih pokreta i prepoznavanje akcija koje izvode kako bi se vršila analiza i usporedba različitih akcija i taktika ili napravila automatska statistička analiza. Takve analize moguće bi pomoći trenerima da poboljšaju stil ili učinak igrača i momčadi te omogući medicinskom osoblju da predviđe moguće ozljede kod igrača [2].

Složenost akcija i aktivnosti sportaša može se značajno razlikovati, počevši od osnovnih akcija poput trčanja i skakanja, malo složenijih akcija poput blokiranja u odbojci ili zakucavanja lopte u košarci pa do složenih aktivnosti koje uključuju interakciju više sportaša, poput križanja u rukometu. Obzirom na navedeno predlaže se sistematizacija na tri glavne kategorije [2]: individualne akcije (eng. *Individual action*), zajedničke aktivnosti (eng. *Joint activity*) i timske aktivnosti (eng. *Team activity*). Individualne akcije uključuju geste (npr. oznaka za *time-out*), jednostavne akcije u kojima sportaš pomiče različite dijelove tijela (npr. trčanje), interakcije s objektima (npr. bacanje lopte) i složene akcije koje mogu biti sastavljene od kombinacije gesta, jednostavnih akcija i interakcija (npr. servis u tenisu prilikom kojega sportaš prvo baca lopticu jednom rukom, zatim zamahuje drugom te pogađa lopticu reketom). U timskim sportovima, osim individualnih akcija, imamo i zajedničke aktivnosti koje uključuju kombinacije više individualnih akcija, ali s dvoje ili više sportaša u interakciji (npr. dodavanje u košarci). Također, imamo i aktivnosti na razini cijelog tima koje se mogu sastojati od raznih kombinacija individualnih akcija i zajedničkih aktivnosti (npr. napad u nogometu ili rotacija igrača u smjeru kazaljke na satu u odbojci) [2].

Važno je naglasiti da su sportske scene dinamične scene u kojima više sportaša istovremeno izvodi različite akcije i aktivnosti, pri čemu često mijenjaju položaje, udaljenost i kut prema kamери, pokrivaju jedni druge, ulaze i izlaze iz vidnog polja kamere i slično, što dovodi do pretrpanih scena i okluzije igrača [3]. Takve situacije značajno otežavaju prepoznavanje akcija i aktivnosti. U posljednje vrijeme istraživači sve više pridaju pažnju ovom problemu, stoga je napravljen pregled primjene HAR-a u različitim sportovima s loptom kako bi se dobio uvid u najčešće sportove koje istraživači istražuju. Primjećeno je da istraživači najviše pažnje posvećuju nogometu, stolnom tenisu, odbojci i hokeju. Vidljiv je porast u broju radova u području razvitkom dubinskog učenja, a posebno u posljednje četiri godine u kojima je napisano 75% od ukupnog broja pregledanih istraživanja [2]. Istraživači su uglavnom koristili metode dubinskog učenja temeljene na konvolucijskim mrežama i LSTM-u, ali također postoji interes za metode strojnog učenja poput SVM-a i značajka kretanja kao što je optički tok. U posljednje vrijeme se sve češće koriste i mehanizmi pažnje i transformeri [2]. Iako se vidi pozitivan pomak u istraživanjima HARa u sportu, postojeći modeli još uvijek nisu dovoljno dostatni i stabilni da bi se u potpunosti odgovorilo na zahtjevne izazove ovog zadatka. No, s obzirom na dosadašnji interes, očekuje se da će daljnja istraživanja u tom području nastaviti eksponencijalno rasti.

U slučaju rukometa, timskog sporta s 14 igrača u kojem je cilj postignuti zgoditak koristeći ruke, tokom igre svaki igrač se kreće kroz teren i izvodi različite akcije poput trčanja, driblanja, bacanja, hvatanja ili aktivnosti poput dodavanja ili križanja. U tu je svrhu stvorena baza podataka s 11 različitih klasa na kojoj se uče razni modeli za HAR kako bi se postigla što bolja točnost [4].

Kao početna točka korišten je InceptionV3 [5] model kojem se kasnije dodala metoda „glasanja“ kako bi se uzela u obzir vremenska komponenta. LSTM [6] i MLP [7] modeli, koji uzimaju vremensku komponentu u obzir, trenirani su na izlučenim značajkama početnog modela [1]. Budući da se prosjek trajanja različitih akcija ili aktivnosti znatno razlikuje i u stvarnome životu i na snimkama, a za korištenje algoritama koji uzimaju u obzir vremenski period potreban je određeni broj okvira (eng. Frame) snimke kao ulazni podatak, testirani su navedeni modeli s različitim ulazom od 20 do 80 okvira [1]. Ukoliko snimka ima manje okvira nego što mreža očekuje, kopije postojećih okvira ubacuju se između postojećih kako bi se dobio odgovarajući broj. Također, budući da bitna obilježja odredene akcije mogu biti u različitim dijelovima snimaka, koriste se različite strategije biranja ulaznih okvira: s početka, sa sredine, s kraja ili decimacijom, odnosno preskakanjem okvira u regularnim intervalima [4]. Sve postignute točnosti modela bolje su od početnog gdje se ne uzima vremenska komponenta u obzir, čime se može zaključiti da ona igra veliku ulogu. Najbolji rezultati postignuti su s MLP modelom na prvih 45 okvira. Otkriveno je da povećanje broja okvira na ulazu u MLP model poboljšava rezultate, dok istovremeno ne utječe na performanse LSTM modela [1]. U dalnjem radu, najviše pažnje biti će posvećeno pronalasku najbolje strategije odabira ulaznih okvira i raspoznavanju akcija na temelju razina predložene sistematizacije.

Ključne riječi – automatsko raspoznavanje akcija, HAR, rukomet, LSTM, MLP

Literatura

- [1] K. Host, M. Ivasic-Kos, i M. Pobar, „Action Recognition in Handball Scenes“, Lecture Notes in Networks and Systems, sv. 283, str. 645–656, 2022, doi: 10.1007/978-3-030-80119-9_41.
- [2] K. Host i M. Ivašić-Kos, „An overview of Human Action Recognition in sports based on Computer Vision“, Heliyon, sv. 8, izd. 6, str. e09633, lip. 2022, doi: 10.1016/J.HELION.2022.E09633.
- [3] K. Host, M. Ivašić-Kos, i M. Pobar, „Tracking handball players with the Deepsort algorithm“, u ICPRAM 2020 - Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods, 2020, str. 593–599. doi: 10.5220/0009177605930599.
- [4] M. Ivasic-Kos, K. Host, i M. Pobar, „Application of Deep Learning Methods for Detection and Tracking of Players“, u Artificial Neural Networks and Deep Learning - Applications and Perspective [Working Title], IntechOpen, 2021. doi: 10.5772/intechopen.96308.
- [5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, i Z. Wojna, „Rethinking the Inception Architecture for Computer Vision“, u Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pros. 2016, sv. 2016-December, str. 2818–2826. doi: 10.1109/CVPR.2016.308.
- [6] H. Sak, A. Senior, i F. Beaufays, „Long short-term memory recurrent neural network architectures for large scale acoustic modeling“, USA, 2014.
- [7] M. W. Gardner i S. R. Dorling, „Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences“, Atmos Environ, sv. 32, izd. 14–15, str. 2627–2636, kol. 1998, doi: 10.1016/S1352-2310(97)00447-0.

Ingrid Hrga

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
ingrid.hrga@gmail.com

Evaluacija sustava za automatsko opisivanje slika

Osnovni cilj sustava za automatsko opisivanje slika je generiranje jedne ili više rečenica na prirodnom jeziku koje doslovno opisuju vizualni sadržaj određene scene. Da bi sustav mogao proizvesti ispravan opis potrebno je: 1) detektirati sadržaj slike u smislu glavnih objekata, atributa i odnosa među njima [1] te dokučiti što je na slici novo i zanimljivo, 2) izraziti semantički sadržaj slike pravilno formuliranom rečenicom [2].

Tijekom posljednjih nekoliko godina, sustavi bazirani na dubokim neuronskim mrežama ostvarili su značajan napredak u generiranju opisa na tečnom jeziku i to primarno generičkih slika. Takvi sustavi uobičajeno slijede arhitekturu kodera i dekodera pri čemu se koder koristi za stvaranje vektorske reprezentacije slike, a dekoder zatim generira opis uvjetovan takvim vektorom. Uobičajeno se na strani kodera koristi konvolucijska neuronska mreža, a na strani dekodera varijanta povratne mreže s mehanizmom pažnje [3] ili, sve češće, mreže bez povratnih veza koje su u cijelosti bazirane na pažnji [4].

Međutim, unatoč ostvarenom uspjehu, jedan od osnovnih izazova automatskog opisivanja slika i dalje ostaje evaluacija rezultata [5]. Istu je sliku moguće opisati na različite načine, ovisno o tome kojem se aspektu pridaje više pažnje, a da ipak svaki od tih opisa bude točan i prikladan. Također, dva opisa mogu sadržavati gotovo jednake riječi, a prenositi sasvim drugačije značenje, što evaluaciju čini posebno izazovnom. Generirane opise mogu evaluirati ljudi, međutim, takav je angažman spor, skup te podložan subjektivnosti [6]. Zato se prednost daje metrikama za automatsku evaluaciju.

Cilj istraživanja je analizirati postojeće metrike za evaluaciju sustava za automatsko opisivanje slika kako bi se identificirale one metrike koje u najvećoj mjeri koreliraju s ocjenama kakve inače dodjeljuju evaluatori. Odabir metrike koja će „voditi“ sustav tijekom učenja direktno utječe na naknadne sposobnosti sustava. Stoga je primjena odgovarajućih metrika ključna kako bi se razvoj mogao ispravno usmjeriti.

I dok se većina postojećih istraživanja odnosi se na poboljšanje određenih aspekata samih sustava, samo se manji dio radova bavi problemom evaluacije. Analizom literature izdvojili smo relevantne radove koji se direktno bave evaluacijom sustava za opisivanje slika kako bi utvrdili tijek razvoja metrika, usporedili njihove prednosti i mane te ukazali na buduće smjerove istraživanja.

Uobičajeni skup metrika za evaluaciju opisa bazira se na primjeni pravila za kvantificiranje podudaranja generiranih i referentnih opisa. Neke su metrike preuzete iz strojnog prevođenja (BLUE [7], METEOR [8]), ROUGE [9] je originalno zamišljena za ocjenu automatske sumarizacije tekstova, dok su CIDEr [10] i SPICE [5] posebno osmišljene za evaluaciju automatski generiranih opisa. Međutim, predložene metrike pokazuju određene slabosti. Standardne metrike primarno su orijentirane na podudaranje n-grama dok se slika pri izračunu ocjene uopće ne uzima u obzir, što u konačnici rezultira time da slabo koreliraju s ljudskom prosudbom [6]. Osim toga, pokazalo se da sustavi koji tijekom učenja optimiziraju npr. CIDEr, imaju sklonost generiranju suviše općenitih opisa s nedovoljnom razinom detalja [11].

Neki autori predlažu metrike bazirane na učenju, poput LEIC [12] koja se oslanja na model učen razlikovati strojno generirane opise od onih koje su napisali ljudi. Međutim, takvo učenje može iziskivati znatnije računalne resurse što primjenu čini nepraktičnom.

Recentni radovi baziraju se na korištenju unaprijed treniranih modela za preslikavanje generiranih i referentnih opisa ili opisa i slika u zajednički vektorski prostor u kojem se jednostavno može izračunati kosinusna sličnost. Na primjer, [13] uspoređuje prikladnost opisa posredstvom SCAN [14] modela za određivanje semantičke sličnosti slika i teksta. Važna prednost sastoji se u tome što takve metrike pokazuju jaču korelaciju s ocjenama evaluadora. Osim toga, takav je pristup otvorio put istraživanjima u smjeru metrika koje iz procesa izostavljaju referentne opise, koje je teško prikupiti te postoje samo za manji broj javno dostupnih skupova, već se uspoređuju samo slike i generirani opisi. S druge strane, u slučaju standardnih metrika, evaluacija ne bi niti bila moguća bez referentnih opisa.

Preliminarni rezultati eksperimentalne usporedbe nekoliko sustava za opisivanje slika primjenom različitih metrika dodatno ukazuju na drugačije rangiranje ovisno o primjenjenoj metrići, što sugerira da evaluacija i dalje ostaje otvoreni istraživački problem.

Ključne riječi – automatsko opisivanje slika, evaluacija, metrike

Literatura

- [1] M. Ivašić-Kos, I. Ipšić, S. Ribarić, "A knowledge-based multilayered image annotation system.", Expert systems with applications, 42 (2015), 2015; 9539-9553.
- [2] O. Vinyals, A. Toshev, S. Bengio, i D. Erhan, "Show and tell: A neural image caption generator," u CVPR, 2015, pp. 3156–3164.
- [3] K. Xu et al., "Show, attend and tell: Neural image caption generation with visual attention," u ICML, 2015, pp. 2048–2057.
- [4] A. Vaswani et al., "Attention is all you need". Advances in neural information processing systems, 2017, 30.
- [5] Anderson, B. Fernando, M. Johnson, i S. Gould, "Spice: Semantic propositional image caption evaluation," u ECCV 2016, 2016, pp. 382–398.
- [6] M. Hodosh, P. Young, i J. Hockenmaier, "Framing image description as a ranking task: Data, models and evaluation metrics," JAIR, vol. 47, pp. 853–899, 2013.
- [7] K. Papineni, S. Roukos, T. Ward, i W. J. Zhu, "BLEU: a method for automatic evaluation of machine translation," u Proceedings of the 40th annual meeting on association for computational linguistics, 2002, pp. 311–318.
- [8] M. Denkowski i A. Lavie, "Meteor universal: Language specific translation evaluation for any target language," u Proceedings of the ninth Workshop on Statistical Machine Translation, 2014, pp. 376–380.
- [9] C. Y. Lin, "Rouge: A package for automatic evaluation of summaries," u Text summarization branches out: Proceedings of the ACL-04 workshop, 2004, vol. 8.
- [10] R. Vedantam, C. L. Zitnick, i D. Parikh, "Cider: Consensusbased image description evaluation," u CVPR, 2015, pp. 4566–4575.
- [11] Q. Wang i A. B. Chan, „Describing like humans: on diversity in image captioning”, u Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019. pp. 4195-4203.
- [12] Y. Cui et al., „Learning to evaluate image captioning”, u Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. pp. 5804-5812.
- [13] M. Jian et al., „Tiger: Text-to-image grounding for image caption evaluation”. arXiv preprint arXiv:1909.02050, 2019.
- [14] K. H. Lee et al., „Stacked cross attention for image-text matching”, u Proceedings of the European conference on computer vision. 2018., pp. 201-216.

Karlo Babić

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia
karlo.babic@uniri.hr

Preliminary Results: Pyramidal Recursive Neural Network

Neural text representation models learn continuous vector spaces that are rich in information (e.g., similar words will be close and dissimilar words far from each other) and range from simpler ones (shallow) to deeper ones (often called language models). The shallow models have shown what is possible with neural representations of text, and the larger language models have improved neural representation [1] and brought new ways a language model can be used.

The proposed model (pyramidal recursive neural network - PyRvNN) builds representations of phrases recursively starting from subwords, but can start with letters or tokens as well. The model recursively combines N representations (each neighboring pair) on one level to produce N-1 representations on the level above that one, which generates a pyramidal hierarchy of representations. Each higher level of this constructed pyramid has higher-level representations of the input text abstracting it from letters or subwords to words, phrases, and eventually perhaps sentences. PyRvNN will be evaluated on downstream tasks (e.g., word similarity, plagiarism, sentiment classification). Preliminary evaluations show promising results for machine-paraphrased plagiarism and a constructed memorization experiment where PyRvNN is compared against FastText and its variant.

Throughout PyRvNN training the most used evaluation is a qualitative one; how well PyRvNN decodes a phrase from a single representation. PyRvNN is currently able to recursively embed subwords pair-wise level-by-level into a single token representation, and then recursively embed token representations pair-wise level-by-level into a single representation of a phrase up to four tokens long, and decode that representation down to subwords again.

Quantitative evaluations we tried, or will try, are: word similarity, memorization, plagiarism, and sentiment classification (each is explained below).

The first (and simplest) task we evaluated PyRvNN on is word (and bigram) similarity and relatedness with datasets WordSim353 [2] and BiRD [3]. The measured correlations between PyRvNN's scoring and the labeled scores are currently positive but low. On WordSim353 the Pearson correlation is 0.17 (Spearman: 0.16), and on BiRD the Pearson correlation is 0.28 (Spearman: 0.33). The correlation results are not high because PyRvNN currently leans more on the decoding of the inputs than on the prediction of neighboring words while training. Further evaluations use tasks better suited for PyRvNN.

To test the potential of PyRvNN representations, one could train a simple language model to predict the next token in a small dataset (effectively memorizing the entire dataset). We constructed such an experiment with a recurrent neural network (as the language model) with fully connected layers as the first and the last layers, and multiple LSTM layers in-between. This simple language model is trained on 500 random paragraphs from Wikipedia.

The original FastText [5] embeddings (with 300 dimensions) performed very poorly (the language model trained with FastText as input could barely memorize 2 paragraphs). To have a leveled ground we trained FastText with 2000 dimensional vectors (same as PyRvNN) on the same dataset (and the same number of epochs, one) as PyRvNN was trained on, and will refer to it as "FastText2k".

In the table below are shown the current results of that experiment (after 400 epochs of training). Each language model generated the first 100 paragraphs (from 500 memorized in total). The generated paragraphs are generated token by token with all the previous tokens (from the original paragraph) as input. The two measures used are: the percentage of tokens in the generated paragraph that are identical as in the original paragraph, and the average Levenshtein distance between the generated tokens and the original tokens.

Embedding model	% of identical tokens	Avg Levenshtein
PyRvNN	95.09	0.45
FastText2k	74.34	1.71

Another task that can be used for evaluation of PyRvNN representations is machine-paraphrased plagiarism classification. Authors in [4] constructed a dataset for such a task with 200,767 paragraphs (50% of which are paraphrased using the SpinBot API) extracted from Wikipedia (English) articles. They evaluate some embedding models (GloVe, Word2Vec, Doc2Vec, and FastText) on that task by training classical machine learning algorithms (Naive Bayes, Logistic Regression, and Support Vector Machine). Each paragraph is represented as an average of token embeddings, which are produced with the mentioned embedding models. The best performing machine learning algorithm for every embedding model is SVM (support vector machine), so we will use SVM for further comparisons.

Their evaluations with SVM are as follows (F1 score): GloVe 89.55, Word2Vec 87.27, Doc2Vec 83.04, FastText-rw (without subwords) 86.15, FastText-sw (with subwords) 82.57. PyRvNN is evaluated with F1 of 89, which is better than all evaluated embeddings but one (GloVe), and is better than the only other evaluated subword embedding model (FastText-sw). When the SVM's regularization parameter C is set to 500 (instead of 100, which is the highest value of C tested by Wahle et al.), SVM with PyRvNN embeddings reaches F1 of 90.

Keywords – pyramidal, recursive, neural, representations

References

- [1] Babić, K., Martinčić-Ipšić, S., & Meštrović, A. (2020a). Survey of neural text representation models. *Information*, 11(11), 511.
- [2] Agirre, E., Alfonseca, E., Hall, K., Kravalova, J., Pasca, M., & Soroa, A. (2009). A study on similarity and relatedness using distributional and wordnet-based approaches.
- [3] Asaadi, S., Mohammad, S., & Kiritchenko, S. (2019, June). Big BiRD: A large, fine-grained, bigram relatedness dataset for examining semantic composition. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 505-516).
- [4] Wahle, J. P., Ruas, T., Foltynek, T., Meuschke, N., & Gipp, B. (2022, February). Identifying machine-paraphrased plagiarism. In International Conference on Information (pp. 393-413). Springer, Cham.
- [5] Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5, 135-146.

Dejan Ljubobratović

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska

dejan.ljubobratovic@uniri.hr

Predviđanje zrelosti breskvi korištenjem modela strojnog učenja

Breskva (*Prunus persica* (L.) Batsch) je popularno voće u Europi i Hrvatskoj. Zrelost pri berbi ima presudan utjecaj na kvalitetu ploda breskve, vijek trajanja, a posljedično i prihvatanje kod potrošača. Stoga bi točno predviđanje zrelosti plodova omogućilo odabir optimalnog vremena za berbu. Parametar kojim se obično predstavlja zrelost ploda je tvrdoća breskve. Kako su tradicionalne metode za određivanje tvrdoće breskve destruktivne, u radu su se koristile alternativne metode za određivanje zrelosti breskve na temelju impedancije breskve i boje breskve, a prema preporuci stručnjaka za to područje.

Korištena su dva skupa podataka na kojima su se provodile analize, a koji su sadržavali mjerena dobivena iz nekoliko stotina plodova voća uključujući morfološka svojstva breskvi, njihovu boju, kiselost, udio šećera i električnu impedanciju. U analizi podataka korišteni su modeli strojnog učenja visoke točnosti, koje karakterizira niska interpretabilnost, tzv. modeli crne kutije, pa su se za interpretaciju dobivenih rezultata koristile objašnjive metode strojnog učenja.

Primarni cilj ove studije bio je razviti model strojnog učenja koji bi s visokom preciznošću predviđao zrelost breskvi na osnovi ulaznih, senzorski prikupljenih podataka [1].

U tu svrhu izgrađen je model slučajne šume (Random forest model) kojim je predviđana zrelost breskvi, a za interpretaciju modela i otkrivanje varijabli koje su najviše utjecale na predviđanja zrelosti koristile su se objašnjive metode strojnog učenja: permutacijska važnost značajki permutacije (eng. Permutation Feature Importance, PFI), važnost značajki (eng. Variable Importance, VI) i LIME metoda interpretabilnosti.

Model je pokazao visoku točnost predviđanja na testnim podacima s preko 97% točnosti, a korištenjem objašnjivih metoda strojnog učenja pokazalo se da je najjači prediktor za predviđanje zrelosti breskvi električna impedancija (varijable Zs i Theta).

Pokazalo se da se tvrdoća breskve, koja prema stručnjacima predstavlja njezinu zrelost, može vrlo točno predvidjeti iz električne impedancije ploda.

Sekundarni cilj istraživanja bio je detektirati ostale značajke ploda koje su važne za uspješno predviđanje zrelosti breskve u određenom stadiju njezina sazrijevanja [2]. Na ubranim breskvama izmjerene su ukupno 33 značajke ploda i kreirana tri skupa podataka nad kojima se zatim trenirao model strojnog učenja. Drugi skup podataka nije sadržavao električnu impedanciju.

Dakle, nad dobivenim skupom podataka se prvo trenirao modeli strojnog učenja slučajne šume, a zatim su se koristile ocjene važnosti značajki tih modela u svrhu otkrivanja nelinearnih (i linearnih) odnosa među njima.

PFI je pokazao da su parametri osnovne boje (h° , a^* i COL indeks osnovne boje), omjer slatkoće i kiselosti breskve (SSC/TA) i kiselost ploda (TA) među prvih deset varijabli koje najviše doprinose u sva tri modela. S druge strane slične rezultate je dobio i model važnosti varijabli (VI) koji je pokazao da su u predviđanju najvažniji parametri osnovne boje (a^* parametar osnovne boje, COL i CCL indekse osnovne boje) te omjer slatkoće i kiselosti breskve (SSC/TA).

Pokazalo se opravdanim korištenje interpretabilnih tehnika strojnog učenja na modelima strojnog učenja u svrhu otkrivanja najvažnijih značajki za predviđanje zrelosti

breskve. Očekivano, različite značajke pokazale su se važnima za različite pragove zrelosti. Opravdanje za takve rezultate nalazi se u velikoj mjeri i u postojećoj literaturi. Uz standardne značajke koje se koriste za predviđanje zrelosti breskve, ova studija naglašava potencijalnu upotrebu nekih dodatnih značajki, kao što su COL, CCL i CIRG1 indeksi osnovne boje. Najistaknutiji među njima je indeks osnovne boje COL, jer se pokazao kao značajka koja je najviše doprinosila predviđanjima. Stoga bi njegov potencijal trebalo dodatno istražiti.

Zaključno, zrelost breskve može se vrlo točno predvidjeti iz električne impedancije ploda, a važnost pojedinih značajki u predviđanju mijenja se ovisno o fazi zrelosti breskve.

Prihvaćanje i potvrda ove ideje otvara mogućnost izgradnje preciznog, ne destruktivnog i brzog modela za određivanje zrelosti ploda pomoću električne impedancije, koji bi se mogao koristiti u automatiziranom branju voća gdje bi mehanička ruka mogla izmjeriti impedanciju breskve neposredno prije branja voća te odlučiti treba li plod brati ili ne. Također, gubitak plodova truljenjem tijekom transporta može se značajno smanjiti podešavanjem određenih parametara teretnog prostora, pa bi predviđanje zrelosti breskvi tijekom transporta imalo važnu ulogu u njegovom optimiziranju.

U drugom istraživanju pokazalo se da se modelom strojnog učenja umjetne neuronske mreže (eng. Artificial Neural Networks, ANN), dobivaju čak i bolji rezultati predviđanja zrelosti breskvi na skupu dobivenom ne destruktivnim metodama [3], tj. metodama koje ne uništavaju breskvu prilikom mjerena svojstava.

Ključne riječi – predviđanje zrelosti breskvi, slučajne šune, strojno učenje, impedanca

Literatura

- [1] D. Ljubobratović, G. Zhang, M. B. Bakarić, T. Jemrić, and M. Matetić, “Predicting peach fruit ripeness using explainable machine learning,” Annals of DAAAM and Proceedings of the International DAAAM Symposium, vol. 31, no. 1, pp. 717–723, 2020, doi: 10.2507/31st.daaam.proceedings.099.
- [2] D. Ljubobratović, M. Vuković, M. Brkić Bakarić, T. Jemrić, and M. Matetić, “Utilization of Explainable Machine Learning Algorithms for Determination of Important Features in ‘Suncrest’ Peach,” Electronics (Basel), vol. 10, no. 24, p. 3115, 2021, doi: <https://doi.org/10.3390/electronics10243115>.
- [3] D. Ljubobratović, M. Vuković, M. Brkić Bakarić, T. Jemrić, and M. Matetić, “Assessment of Various Machine Learning Models for Peach Maturity Prediction Using Non-Destructive Sensor Data,” Sensors, vol. 22, no. 15, 2022, doi: 10.3390/s22155791.

Goran Paulin

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
Kreativni odjel d.o.o., Rijeka, Croatia
gp@kreativni.hr

Primjena sintetizirane grafike za automatsku detekciju u različitim scenarijima

Uvjerljivo najveći broj istraživanja u domeni detekcije objekata na statičnim slikama i videu usmjeren je na optimizaciju modela strojnog učenja, danas većinom baziranih na dubokim konvolucijskim neuronskim mrežama, te primjenu tih modela za različite zadatke detekcije. Aktualni modeli detekcije postižu vrlo dobre rezultate kad god se detekcija provodi na slikama koje su snimljene u sličnim uvjetima kao i slike korištene za treniranje modela. To je pristup koji traži pripremu velikih skupova podataka potrebnih za treniranje modela, pri čemu se najčešće prednost daje kvantiteti nad kvalitetom takvog skupa.

Kvalitetni skup podataka za treniranje modela za detekciju ima 3 glavne karakteristike: podaci u njemu moraju biti učinkoviti, primjereni zadatku i vizualno slični [1]. Učinkovitost proizlazi iz istovremenog smanjivanja količine podataka, kako bi se štedjelo na resursima potrebnim za učenje, uz povećanje raznovrsnosti uzorka [2]. U smislu primjerenošću zadatku (eng. *task-aware*) moraju sadržavati primjere koji pomažu poboljšati ciljani model [3]. Pritom moraju biti vizualno realistični kako bi minimizirali jaz domene prisutan između slika iz više različitih izvora, te na taj način poboljšali generalizaciju [4].

Tradicionalno, skupovi podataka nastaju prikupljanjem podataka, dobivenih senzorima ili izravnim mjeranjem. S obzirom da detektori objekata često koriste nadzirano učenje, tako prikupljen skup slika potrebno je manualno anotirati kako bi se definiralo činjenično stanje (eng. *ground truth*) potrebno za treniranja i evaluaciju modela. Anotiranje stvarnih slika je sporo, podložno ljudskog pogrešci i često stručan proces koji zahtijeva iznimnu pozornost anotatora. Sve navedeno čini izradu anotiranog skupa podataka dugotrajnim i stoga skupim [5].

Potencijalno rješenje ovog problema je sintetizirana grafika odnosno sintetički skupovi podataka (eng. *synthset*). Sintetički podaci su podaci koji se ne dobivaju izravnim mjeranjem ili radom senzora već generiranjem koje omogućava prikupljanje dovoljne količine podataka za treniranje nadziranih modela strojnog učenja i nudi mogućnost automatske anotacije. Sintsetovi nisu novost, u računalnom vidu koriste se od 1989. godine [6], ali značajan razvoj metoda i tehnika za njihovo generiranje pripada posljednjem desetljeću [7].

Sintetički podaci mogu biti neograničeni izvor uravnoteženih i raznolikih podataka. Mogu simulirati rijetke događaje ili situacije s kojima se još nismo susreli te omogućavaju prevladavanje ograničenja upotrebe stvarnih podataka uvjetovano poštivanjem privatnosti ili drugim propisima [8].

Jedan od najvećih problema korištenja isključivo stvarnih podataka je što u skupu za učenje vrlo vjerojatno nisu sadržani svi primjeri koji mogu postojati. Posljedično, treniramo model bez takvih primjera. Kada se novi primjer pojavi, možemo ga pokušati koristiti za dotrenirati model, ali izgledno je da takav postupak neće imati željeni učinak na preoblikovanje naučene funkcije detekcije zbog neuravnoteženosti (eng. *bias*) podataka, zbog čega model može zapravo izignorirati taj usamljeni novi, ali ipak mogući slučaj. Kako nabaviti još takvih (stvarnih) podataka koje smo sada prvi put sreli, a prethodno nismo ni znali da postoje? Nikako - barem dok ne najđemo na sljedeći sličan primjer. Sintetički podaci ovaj problem mogu riješiti inženjeringom: kada postanemo svjesni postojanja

novog slučaja, možemo generirati N potrebnih primjera pri čemu sami možemo odrediti koliku zastupljenost tih novih primjera u ukupnom skupu želimo (ravnopravnu ili veću/manju), čime garantiramo učinkovitost dotreniranog modela.

Kako bismo ocijenili mogućnost primjene sintetizirane grafike za automatsku detekciju u različitim scenarijima osmislili smo seriju eksperimenata rastuće složenosti. Započeli smo s generiranjem sintseta u domeni rukometa [9], proceduralno generirajući 3D scenu s prethodno pripremljenim 3D elementima (rukometna dvorana, manualno animirani igrači i fizikalno simulirana lopta). Koristeći takav sintset za augmentaciju skupa stvarnih podataka unaprijedili smo detekciju (YOLO) za 17,24% (mAP). Potom smo praksu korištenja gotovih 3D elemenata zamijenili proceduralnim generiranjem 3D geometrije na bazi 2D tekstura, a za generiranje slika koristili smo metodu komponiranja polažeći 3D sjenčane objekte na fotografije pozadina. Ovu metodu primijenili smo na složeniji problem, segmentaciju (Mask R-CNN) leptira [10] te postigli 27,7% (mAP) koristeći isključivo sintetičke podatke za treniranje modela. Unatoč dobrim rezultatima, nedostajala nam je informacija koje se sve metode i tehnike generiranja sintetičkih skupova mogu koristiti, zbog čega smo istražili i sistematizirali dosadašnju, 30-godišnju praksu [7] i izlučili 9 jedinstvenih, potencijalno isprepletenih metoda generiranja sintsetova, te 17 pojedinačnih procesa koje kreatori sintsetova mogu slijediti primjenjujući raspoložive tehnike unutar njih. Poučeni tim spoznajama, proširili smo analizu mogućnosti primjene sintetizirane grafike koristeći tehniku praćenja zrake (eng. raycasting) u svrhu lokaliziranja nestalih osoba u misijama traganja i spašavanja (eng. search and rescue, SAR) u neurbanim područjima, uz podršku bespilotnih letjelica (eng. unmanned aerial vehicle, UAV), i postigli visoku točnost (predviđena udaljenost do tražene osobe uz pogrešku manju od 6 m), koristeći *in silico* metodu i vlastiti simulator izgrađen za tu priliku [11]. Izvodeći prethodno opisani eksperiment *in vivo* i dodatno optimizirajući raycast metodu [12] postigli smo jednako dobre, a u nekim slučajevima i bolje rezultate (najbolji postignuti rezultat bila je predviđena udaljenost do tražene osobe uz pogrešku od 0,7 m).

Ohrabreni dosadašnjim rezultatima, razvijamo metodu za detekciju i lokalizaciju objekata u UAV podržanim SAR misijama baziranu u cijelosti na sintetiziranoj grafici te, očekujući dobre rezultate, u budućem radu planiramo dotičnu metodu prenijeti u domenu medicinskih istraživanja.

Ključne riječi – računalni vid, detekcija objekata, sintetički skup podataka, sintset

Literatura

- [1] S. Tripathi, S. Chandra, A. Agrawal, A. Tyagi, J. M. Rehg, i V. Chari, „Learning to generate synthetic data via compositing“, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., sv. 2019-June, str. 461–470, 2019.
- [2] F. E. Nowruzi, P. Kapoor, D. Kolhatkar, F. Al Hassanat, R. Laganiere, i J. Rebut, „How much real data do we actually need: Analyzing object detection performance using synthetic and real data“, 2019.
- [3] A. Prakash i ostali, „Structured domain randomization: Bridging the reality gap by context-aware synthetic data“, Proc. - IEEE Int. Conf. Robot. Autom., sv. 2019-May, str. 7249–7255, 2019.
- [4] A. Rozantsev, V. Lepetit, i P. Fua, „On rendering synthetic images for training an object detector“, Comput. Vis. Image Underst., sv. 137, izd. 0, str. 24–37, 2015.
- [5] G. Paulin, „Pregled i analiza metoda generiranja sintsetova za primjenu u domeni detekcije osoba i raspoznavanja akcija“, Sveučilište u Rijeci, 2020.
- [6] D. A. Pomerleau, „Alvinn: An autonomous land vehicle in a neural network“, Adv. Neural Inf. Process. Syst. 1, str. 305–313, 1989.

- [7] G. Paulin i M. Ivašić-Kos, „Review and analysis of synthetic dataset generation methods and techniques for application in Computer Vision“, Artifical Intell. Rev., prihvaćen za tisk
- [8] D. B. Rubin, „Discussion of statistical disclosure limitation“, J. Off. Stat., sv. 9, izd. 2, str. 461–468, 1993.
- [9] M. Burić, M. Ivašić-Kos, i G. Paulin, „Object Detection Using Synthesized Data“, u ICT Innovations 2019 Web proceedings, 2019, str. 110–124.
- [10] G. Paulin, M. Burić, K. Host, J. Liu, i L. P. L. Nguyen, „Project Butterfly“, 2020. [Na internetu]. Dostupno na: <https://github.com/khost95/SSIP-2020-Project-Butterfly>. [Pristupljeno: 02-lis-2022].
- [11] G. Paulin, S. Sambolek, i M. Ivasic-Kos, „Person localization and distance determination using the raycast method“, u 2021 6th International Conference on Smart and Sustainable Technologies (SpliTech), 2021, str. 1–5.
- [12] G. Paulin, S. Sambolek, i M. Ivasic-Kos, „Application of Raycast Method for Person Localization and Distance Determination in Real-World Search and Rescue Scenarios“, neobjavljen

Milan Petrović

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia

milan.petrovic@uniri.hr

Static and Dynamic Analysis of *Drosophila Melanogaster* Social Interaction Networks

This paper presents a method for analyzing the group behavior of *Drosophila melanogaster*. Quantification of activity was performed using complex networks. The first part of the paper describes the static analysis of networks at local, middle and global levels. In the second part, an analysis of the dynamics of behavior is presented. Addiction is a multifactorial biological and behavioral disorder studied using animal models based on simple behavioral responses in isolated individuals. A few decades ago, it was shown that *Drosophila melanogaster* can serve as a model organism for behaviors related to alcohol, nicotine, and cocaine (COC) addiction. Evaluating COC-induced behaviors in a large group of flies was a technological challenge. Therefore, we applied a local, middle and global level of network-based analyses to examine social interaction networks (SINs) in a group of 30 untreated males compared to those orally administered 0.50 mg/mL of COC for 24 hours [1], [2]. In this study, we confirmed the previously described increase in locomotion after COC feeding. We isolated new network-based measures associated with COC that were influenced by group on individual behavior. COC-fed flies showed longer duration of local-level interactions and formed larger, more densely populated, and more compact mid-level communities. Untreated flies had a higher number of interactions with other flies in a group at the local level, and at the middle level, these interactions resulted in the formation of separate communities. Although network density at the global level is higher in COC-fed flies, modularity at the middle level is higher in untreated flies. One COC-specific behavior we isolated was an increase in the proportion of individuals that did not interact with the rest of the group, which is thought to be an individual difference in COC-induced behavior and/or a consequence of the influence of the group on individual behavior. Our approach can be extended to other classes of drugs with the same acute response as COC to determine drug-specific network-based measures and could serve as a tool to determine genetic and environmental factors that influence both drug addiction and social interaction. The analysis of network dynamics was performed on two different populations of adult *D. melanogaster* males, the first being a wild type (CTRL) and the second being a population consuming a psychostimulant (COC). Prior to the experiment, individuals were separated under cold anesthesia and placed in vials in which a psychostimulant was added to the food for the needs of the COC group. Analysis of behavioral dynamics was performed in several ways. The first is a quantitative analysis of the time in which the interaction occurred from the beginning of the experiment to the end. The second method is to analyze the triads in the networks and measure the time it takes three flies (nodes) to form an open triad, and then how long it takes for that triad to change from open to closed.

Keywords- *Drosophila melanogaster*, complex networks, social networks analysis, results

References

- [1] M. Petrović, A. Meštrović, R. Andretić Waldowski, A. Filošević Vujnović, “A network-based analysis detects cocaine-induced changes in social interactions in *Drosophila melanogaster*”, PLOS ONE, accepted for publication
- [2] M. Petrović, A. Meštrović, R. Andretić Waldowski, A. Filošević, “Dynamics of “*Drosophila melanogaster*” social interaction networks”, Complex Networks 2022, extended abstract

Marko Pribisalić

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia

marko.pribisalic@gmail.com

Towards Survey on Automatic Generation of Test Cases from Software Specification Documents

Software testing is usually conducted from specification documents written in natural language and involves intensive human interaction and potential errors or at least suboptimal results. We hypothesize that it is possible to automatically infer test cases from natural language specifications, thus to reduce the manual effort, to improve test coverage, to speed up and to raise quality of test cases. The purpose of the research is to develop a methodology for automatic generation of cases for software testing from software specifications. For that purpose, we made systematic literature review by analyzing papers indexed in Google Scholar and Scopus databases. The goal of study is to identify existing state-of-the-art approaches and prospective future research directions for creating test cases in agent-based modeling, with a focus on their potential real-world applications. We used Leximancer tool, which automatically analyses text documents to identify high-level concepts and provide key ideas and insights from the text.

The method used in the paper is a Systematic Literature Review (SLR) and Systematic Literature Mapping (SLM). A SLR is a method in Evidence-Based Software Engineering (EBSE). Doing SLR and SLM investigations as well, as following known protocols in conducting SLR as well as SLM experiments in Software Engineering (SE) [1].

Software testing is a time-consuming and expensive operation in the software development process. The bulk of the expenditure is spent during the test case development process. Requirement-based testing is an approach for building test cases independently of the implementation's internal structure. Requirement-based testing incorporates both functional and nonfunctional requirements. Clearly, there is a need for tools and techniques that make software testing accessible to academics and practitioners from a wide variety of fields. The topic of test optimization has grown in popularity in recent decades, and many researchers have presented strategies, approaches, and methods that could aid in the optimization of the testing process. Test automation, test case prioritization [2], and test suite minimization [3] are all techniques to improve the software testing process. Furthermore, artificial intelligence approaches such as machine learning, deep learning, and natural language processing (NLP) have garnered a lot of interest in the testing sector. Artificial intelligence has been found to assist in achieving more accurate results and saving time in software testing. Artificial intelligence-driven testing is expected to be the dominating technology for quality assurance work in the near future, according to recent studies [4].

One technique to optimize software testing and increase test effectiveness is to improve the process of producing test cases. The IEEE Standard Glossary of Software Engineering Terminology (IEEE Std610.12-199) [5] defines a test case as a set of test inputs, execution conditions, and expected outcomes created for a specific goal, such as exercising a certain program route or verifying compliance with a given requirement. The procedure for creating test cases can be done manually, semi-automatically, or automatically. Manual labor is still one of the most common means of generating test cases nowadays. However, there is a demand for manual testing to be replaced with automated testing. Manual testing entails a team of testers manually executing a set of test case requirements [6]. Manual software testing is still an important strategy for computer system validation [7], but it is a time-consuming process that is prone to human errors [8]. Manual labor can be minimized

by automating the generation of test cases, resulting in reduced test time and software development and maintenance costs [9].

Natural language processing has already been employed in the realm of software testing [10]. It is often used in software testing for two goals, requirement engineering and specification mining. Requirements engineering is a process of discovering the purpose, by identifying stakeholders and their needs and documenting these in a form amenable to analysis, communication, and implementation [11]. Specification mining, also known as specification discovery, often aims at discovering behavioral patterns that are either unknown to the user or missed during a design and verification cycle. The lack of software specifications and poor documentation, a prominent problem in both hardware and software engineering, can be very costly due to the difficulty ensued on maintaining these systems. Specification mining is used as a tool to aid program understanding [12]. Numerous prior research [13] explored the process of creating test cases from various forms of specification documents. Some authors conduct an evaluation of numerous research articles demonstrating how to build test cases from natural language requirements documents using a variety of natural language processing approaches [14]. Numerous natural language processing (NLP) techniques are often used to analyze text data. Text segmentation, which is considered the first stage in natural language processing, is the process of splitting a text into linguistically relevant pieces [15]. The text segmentation step includes both sentence segmentation and tokenization. The process of segmenting a text into sentences is called sentence segmentation. Tokenization, also called word segmentation refers to the process of subdividing text into discrete words (tokens) on the basis of a preset set of criteria, which may include the use of white-space as a separator [16]. Text segmentation is critical in any natural language processing system because the words and sentences that are identified at this stage are later passed on to next processing stages. The practice of applying a morphosyntactic tag for every word as well as punctuation marks in a text is called part-of-speech (POS) tagging. By applying morphological and syntactic criteria to words, morphosyntactic analysis defines their POS categories [17]. After that, the NLP process may move to text lemmatization. This is the process of mapping several variants of a word (e.g., helps, helped) to its most fundamental form (help), called a lemma [16]. Additionally, natural language phrases often include many filler words such as "and," "to," and "the," which may introduce noise into text analysis. These words are referred to as stop words since they are often omitted from analysis [18]. Named Entity Recognition (NER) is another often used NLP technique that is sometimes called chunking. It extracts information such as locations, companies, and human names from text [16]. Several different forms of specification papers, mostly requirements, were employed in the reported study at various testing phases, most notably system- and acceptability testing. The most frequently utilized NLP methods in research were tokenization, parts - of - speech, chunking, and parsing [2].

This study maps and explores the corpus of published research papers on NLP-assisted software testing by finding state-of-the-art as well as best practices in NLP-assisted software testing. Article acts as an "index" to the large amount of literature accessible on this subject by summarizing what we know and brings systemizing of the existing methods, techniques and ways of NLP-assisted software testing. While we conducted a high-level synthesis for approach accuracy, a more comprehensive systematic review of the literature (SLR) is required to objectively examine and compare the different NLP-assisted software testing approaches discovered in this initial SLM study. The SLR may be strengthened by more empirical research in which a limited number of Systems Under Test (SUTs) are selected and then subjected to the NLP-based techniques. This would permit a critical evaluation of the strengths and limitations of the observed methods. Given that the first

phase (a systematic classification of the literature) has been completed, we will proceed to future efforts conduct in-depth systematic literature review in this domain.

Keywords – software testing, test case generation, software specification documents, natural language processing

References

- [1] A. Goffi, A. Gorla, M.D. Ernst, and M. Pezzè, "Automatic generation of oracles for exceptional behaviors," International Symposium on Software Testing and Analysis, pp. 213–224, 2016.
- [2] S. Kobayashi, K. Fukuda, and H. Esaki. "Towards an NLP-based log template generation algorithm for system log analysis." In Proceedings of The Ninth International Conference on Future Internet Technologies, pp. 1-4, 2014.
- [3] S.D.R. Konreddy, "The Impact of NLP on Software Testing.", Journal of University of Shanghai for Science and Technology, ISSN: 1007-6735, Volume 23, Issue 8, pp. 296-304, 2021.
- [4] M. Shahbaz, P. McMinn, and M. Stevenson. "Automated discovery of valid test strings from the web using dynamic regular expressions collation and natural language processing." International Conference on Quality Software, pp. 79–88, 2012.
- [5] G. Vahid, S. Bauer, and M. Felderer. "NLP-assisted software testing: A systematic mapping of the literature." Information and Software Technology 126, 2020.
- [6] M.W. Esser, and P. Struss. "Obtaining models for test generation from natural-language-like functional specifications." International Workshop on Principles of Diagnosis, pp. 75–82, 2007.
- [7] A.L.L. de Figueiredo, W.L. Andrade, and P.D.L. Machado. "Generating interaction test cases for mobile phone systems from use case specifications." ACM SIGSOFT Software Engineering Notes 31, no. 6, pp. 1–10, 2006.
- [8] B. Shajulin, R. Badami, and M. Bhagyalakshmi. "APM Bots: An Automated Presentation Maker for Tourists/Corporates Using NLP-Assisted Web Scraping Technique." In International Conference on Advanced Network Technologies and Intelligent Computing, Springer, Cham, pp. 647-659, 2021.
- [9] A. Ansari, M.B. Shagufta, A.S. Fatima, and S. Tehreem. "Constructing Test cases using Natural Language Processing." International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatic, pp. 95–99, 2017.
- [10] P. Kulkarni, and Y. Joglekar. "Generating and analyzing test cases from software requirements using nlp and hadoop." International Journal of Current Engineering and Technology, vol.4, no.6, pp. 3934–3937, 2014..
- [11] B. Nuseibeh, S. Easterbrook, in Encyclopedia of Physical Science and Technology (Third Edition), 2003.
- [12] L. Erlkh. Leveraging legacy system dollars for e-business. IT Professional, 2(3):17–23, 2000.
- [13] D. Angus, S. Rintel, and J. Wiles. "Making sense of big text: a visual-first approach for analysing text data using Leximancer and Discursis." International Journal of Social Research Methodology 16, no. 3, pp pp. 261-267, 2013., 2013.
- [14] B.K. Aichernig, D. Ničković, and S. Tiran. "Scalable incremental test-case generation from large behavior models." In International Conference on Tests and Proofs, pp. 1–18, 2015.
- [15] S. Nogueira, H.L.S. Araujo, R.B.S. Araujo, J. Iyoda, and A. Sampaio. "Automatic generation of test cases and test purposes from natural language."Brazilian Symposium on Formal Methods, pp. 145–161, 2015.
- [16] R. Sharma, and K.K. Biswas. "Natural Language Generation Approach for Automated Generation of Test Cases from Logical Specification of Requirements." International Conference on Evaluation of Novel Approaches to Software Engineering, pp. 125–139, 2014.

- [17] F. Fabrizio, M. Fusani, S. Gnesi, and G. Lami. "The linguistic approach to the natural language requirements quality: benefit of the use of an automatic tool." In Proceedings 26th Annual NASA Goddard Software Engineering Workshop, IEEE, pp. 97-105, 2001.
- [18] E. Sarmiento, J.C. Sampaio do Prado Leite, and E. Almentero. "C&L: Generating model based test cases from natural language requirements descriptions." International Workshop on Requirements Engineering and Testing, pp. 32–38, 2014.

Ivan Šimac

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
isimac@veleri.hr

Izrada skupa podataka za automatsku detekciju plivača

Danas se u svijetu sporta, kako profesionalnog tako i amaterskog, te u rehabilitaciji, sve više koriste razne vrste senzora za praćenje i analizu aktivnosti i performansi sportaša. Često se koriste razni gadgeti sa senzorima, od vitalnih koji prate rad srca i mišića do senzora za kretanje, posebice u sportovima koji donose "profit".

Danas je moguće pronaći mnoge otvorene baze podataka slika različitih senzorskih i vizualnih podataka koji se mogu koristiti za otkrivanje i analizu različitih radnji u sportovima poput nogometa, rukometa, hokeja i biciklizma. Problem je što nedostaju slike plivačkog sporta, osim što su neki od datasetova skup slika televizijskih prijenosa natjecanja ili su preuzeti s YouTube videa. Dakle, za analizu tehnika i stilova plivanja u slikama ili videima, prvo je potrebno snimiti i prikupiti podatke plivača u stvarnim situacijskim uvjetima treninga i natjecanja snimljene akcijskim kamerama iz različitih kutova snimanja.

No, snimanje u blizini vode ima mnogo otvorenih problema kao što su kapi, valovi i odsjaj svjetlosti. Također je nemoguće pratiti snimku u realnom vremenu zbog prekida veze između kamere uronjene u vodu i prijemnika na suhom.

Tijekom istraživanja, pokušat ćemo se nositi s tim izazovima dok detektiramo plivače, prepoznajemo njihove tehnike i analiziramo njihove stilove plivanja. Sljedeći korak bit će identificiranje plivača na slikama ili videozapisima na temelju njihovog držanja i tehnike plivanja. Planiramo također istražiti mogućnost automatske video pretrage i pronalaženja svih isječaka određenih tehnika ili svih isječaka u kojima je prisutan pojedini plivač.

Naposljeku, važan korak u obradi videa bit će detekcija najmanje jedinice akcije i automatsko rezanje videa kako bi se samo ona obuhvatila, kao i automatsko uklanjanje "idle" snimaka na kojima nema plivača.

Keywords – dataset, plivanje, računalni vid, detekcija akcije

Literatura

- [1] I. Šimac, M. Ivašić-Kos, M. Pobar, „Baza slika za strojno učenje modela za detekciju plivača“, Zbornik Veleučilišta u Rijeci, prihvaćeno za tisk
- [2] S.M. Safdarnejad, X. Liu, L. Udupa, B. Andrus, J. Wood, D. Craven, Sports Videos in the Wild (SVW): A Video Dataset for Sports Analysis Proc. International Conference on Automatic Face and Gesture Recognition (FG 2015), Ljubljana, Slovenia, 2015
- [3] M. Ivašić-Kos, M. Pavlić, M. Pobar, Analyzing the semantic level of outdoor image annotation, Proceedings of MIPRO 2009 - 32nd International convention on information and communication technology, electronics and microelectronics, Opatija, 2009
- [4] K. Soomro and A. R. Zamir, Action Recognition in Realistic Sports Videos, Computer Vision in Sports. Springer International Publishing, 2014
- [5] M. Ivašić-Kos, M. Pobar, „Building a labeled dataset for recognition of handball actions using mask R-CNN and STIPS, 7th European Workshop on Visual Information Processing (EUVIP), 10.1109/EUVIP.2018.8611642, 2018

Romeo Šajina

Faculty of Informatics and Digital Technologies, University of Rijeka, Rijeka, Croatia

Faculty of Informatics in Pula, Pula, Croatia

rsajina@unipu.hr

Skeleton-based Human Action Scoring and Comparison using Monocular camera

In the field of sports, computer vision can be used to detect and track players or the ball, detect player actions, detect objective score change, player pose estimation, etc. In this work, we will focus on action evaluation, i.e. comparison of poses while the player is executing certain action (e.g. handball shot) against the template sequence performed by a professional. By comparing player poses to the template poses we can provide them with the information of the needed corrections to improve their action execution. This application can be especially useful to beginners in the sport or in rehabilitation exercises purposes.

Comparison of poses is the final step in action evaluation which requires that we have collected players' poses in 3D space. Previous research on action evaluation (such as [1]) mostly used markers and sensors to capture players' poses that provide accurate poses, but require a controlled environment that is often expensive and time-consuming. For that reason, our system is based on a monocular camera, making it simple and fast to use, but obtaining accurate players' poses becomes a challenging task. To tackle this task we use a 2D pose-estimation method like MaskRCNN [2] or UDP-Pose [3], that produces 2D points forming a pose. We then feed obtained 2D poses in a 3D pose-estimation method like EvoSkeleton [4] or VideoPose3D [5], which finally produces poses in 3D space. Given that the distance of the player from the camera can vary, we need to standardize pose sequences before comparison. A simple normalization technique that scales all poses to the same height of 1 is sufficient to solve the problem. Another technique that we can use for normalization is retargeting, a method that translates pose from the source skeleton to the target one, as in [6]. This technique also solves the common problem of pose-estimation methods where keypoints are not consistently detected in the same place (e.g. foot can be detected directly on the ankle or even on the middle of the lower leg), producing bones of different lengths across the sequence. The final step before the comparison is sequence alignment, a task where the two sequences are aligned so that actions are performed at the same time in both sequences. For sequence alignment, we can use a simple sliding window approach where we slide the template sequence across the evaluated sequence to find the action by finding the part with the smallest overall distance (e.g. euclidean distance, cosine distance, etc.). However, more advanced approaches like Dynamic Time Warping (DTW) [7] provide better results because they can handle the difference in speed of movements within the action sequence. Once the sequences are aligned we can inspect each frame and visualize the difference between poses.

Most of the time we don't need information on every single pose in the sequence as we can evaluate the action sequence based on a couple of keyframes within the sequence. Naturally, keyframes are different for each template action sequence, and determining them requires time and expertise, as in [8]. To tackle this problem we developed a simple algorithm that finds keyframes based on the overall pose movement. The algorithm first computes the absolute difference between two subsequent poses which produces a 1D series of overall pose movement across the sequence. Next, we are looking for keyframes within the 1D series where the person was moving the least (i.e. valleys) and keyframes where the momentum is starting to drop down (i.e. hills). Valleys usually represent

preparing for a big movement, while hills usually represent the moment when the movement reached its peak. Finally, we have a full system that visualizes players' action sequences and visualizes the differences compared to the template action sequence.

Keywords – action comparison, action scoring, human pose estimation, sequence alignment

References

- [1] M. Oshita, T. Inao, S. Ineno, T. Mukai, S. Kuriyama, "Development and evaluation of a self-training system for tennis shots with motion feature assessment and visualization", in The Visual Computer, vol. 35, 2019, pp. 5700-5709.
- [2] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask r-cnn", in Proceedings of the 16th IEEE International Conference on Computer Vision, Venice, Italy, 22-29 October 2017, pp. 2980–2988.
- [3] J. Huang, Z. Zhu, F. Guo, G. Huang, "The devil is in the details: Delving into unbiased data processing for human pose estimation", in Proceedings of the 2021 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Virtual, Online, 14-19 June 2020, pp. 5699–5708.
- [4] S. Li, L. Ke, K. Pratama, Y.W. Tai, C.K. Tang, K.T. Cheng, "Cascaded Deep Monocular 3D Human Pose Estimation with Evolutionary Training Data", in Proceedings of the 2020 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Virtual, Online, 14-19 June 2020, pp. 6172-6182.
- [5] D. Pavllo, C. Feichtenhofer, D. Grangier, M. Auli, "3d human pose estimation in video with temporal convolutions and semi-supervised training", in Proceedings of the 2019 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 16-20 June 2019, pp. 7745–7754.
- [6] Monzani, J.S.; Baerlocher, P.; Boulic, R.; Thalmann, D. "Using an intermediate skeleton and inverse kinematics for motion retargeting", in Computer Graphics Forum, vol. 19, 2000, pp. 11–19.
- [7] D. Berndt, J. Clifford, "Using dynamic time warping to find patterns in time series", in AAAI-94 Workshop on Knowledge Discovery in Databases, Seattle, Washington, 1994, pp. 359-370.
- [8] A. Voulodimos, I. Rallis, N. Doulamis, "Physics-based keyframe selection for human motion summarization", in Multimedia Tools and Applications, vol. 79, 2020, pp. 3243–3259.

Igor Škorić

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
skoric.igor@gmail.com

Model vrednovanja razvojnih okruženja koja se koriste u učenju programiranja

Evaluacija informacijskih sustava (IS) kontinuirano se istražuje već nekoliko desetljeća. Vremenom je u istraživačkoj zajednici sazrela spoznaja o potrebi višekriterijskog modela vrednovanja, te su DeLone i McLean [1] 1992. godine na osnovu prethodnih istraživanja kreirali model uspješnosti informacijskog sustava (D&M model). Model su izvorni autori kasnije dodatno usavršili [2] kao odgovor na povratne informacije dobivene korištenjem početne inačice modela. Poboljšani model uspješnosti IS-a sastoji se od šest dimenzija: kvaliteta informacija, kvaliteta sustava, kvaliteta usluge, namjera korištenja/korištenje sustava, zadovoljstvo korisnika i koristi od sustava. Kvaliteta informacija su poželjne karakteristike izlaznih podataka sustava (npr. razumljivost, prikladnost ili upotrebljivost). Kvaliteta sustava odnosi se na obilježja sustava kao što su jednostavnost korištenja, fleksibilnost ili sveobuhvatnost. Kvaliteta usluge opisuje podršku koju dobivaju korisnici sustava (npr. dostupnost ili brzina odaziva). Zadovoljstvo predstavlja mjeru do koje sustav ispunjava očekivanja korisnika, a koristi predstavljaju dodane vrijednosti koje informacijski sustav pruža korisniku ili organizaciji. Ovih šest dimenzija zajedno daju cjelovitu sliku uspjeha informacijskog sustava [3]. Prilikom primjene modela, za svaku od navedenih dimenzija određuju se mjere koje je opisuju. Odabir mjera ovisi o prirodi i svrsi sustava, kao i o kontekstu u kojem se sustav upotrebljava. D&M model je testiran u mnogim istraživanjima [4] i primijenjen je u različitim kontekstima, te je postao jedan od najviše korištenih i citiranih modela procjene uspješnosti informacijskog sustava.

U obrazovanom sustavu se koriste brojni alati sa ciljem poboljšanja procesa poučavanja i učenja, pa se pojavljuje i problem njihovog vrednovanja. Dosadašnja istraživanja koristila su . D&M model za evaluaciju uspješnosti sustava poput e-portfolija [5], inteligentnog tutorskog sustava za učenje [6], virtualne okoline za učenje [7], ili sustava za upravljanje učenjem [8], a u pojedinim studijama model je i proširen na način da je prilagođen specifičnostima obrazovnog konteksta. Problem kojim se bavi ovaj rad je vrednovanje razvojnih okruženja koja se koriste u učenju programiranja primjenom proširenog D&M modela. Korištenje razvojnog okruženja sastavni je dio svakog učenja programiranja Brojni radovi posvećeni su razvoju novih okruženja [9]. Svaka priprema poučavanja programiranja nužno uključuje i odabir prikladnog razvojnog okruženja. To naglašava potrebu evaluacije takvih okruženja. Dosadašnja istraživanja pokrivaju i vrednovanje razvojnih okruženja koja se koriste u učenju programiranja, ali velika većina njih se bavi samo pojedinim aspektima uspješnosti, poput utjecaja vizualizacije programa na motiviranost studenata [10], učinke korištenja alata tijekom programiranja na anksioznost, razinu kognitivnog opterećenja i postignuće učenika [11], ili percipirane koristi od web baziranih razvojnih okruženja iz perspektive studenata [12]. Nedostatak spomenutih istraživanja je što ne daju sveobuhvatno vrednovanje razvojnih okruženja što otežava njihovu usporedbu. Malobrojna istraživanja koja koriste D&M model u ovom kontekstu [13] ograničavaju se na primjenu originalne inačice modela bez prilagodbe odabranom kontekstu. Izostanak prilagodbi na temelju konteksta upotrebe, razine analize i svrhe sustava može rezultirati izostavljanjem relevantnih obilježja predmeta vrednovanja.

Razvojna okruženja koja se koriste u učenju programiranja međusobno se razlikuju po mnogočemu. Ovi alati su prilagođeni različitim platformama, podržavaju različite programske jezike, imaju različite funkcionalnosti i obilježja. To znači da model treba biti

dovoljno općenit, a s druge strane prilagođen da što bolje obuhvati proučavani problem. Dodatno ograničenje na model je činjenica da se razvojna okruženja u učenju programiranja koriste za različite oblike nastave i uz različite pedagoške intervencije. Posljedično, razvojno sučelje koje je jednostavno ili sveobuhvatno (dimenzija kvalitete sustava) i koje ima poželjne karakteristike izlaznih podataka (dimenzija kvalitete informacija) ne mora nužno biti prikladno. Navedeni problem adresirao bi se proširenjem D&M modela dimenzijom prikladnosti tehnologije zadatku. Ovo proširenje u skladu je sa dosadašnjim istraživanjima [14,15] koja dovode u vezu prikladnost tehnologije i njenu uspješnost.

Praktična primjena razvijenog modela demonstrirala bi se razvojem obrazovnog sustava preporučivanja baziranog na tom modelu. Obrazovni sustavi preporučivanja pri kreiranju preporuka uzimaju u obzir i pedagoške principe [16], što ih čini logičnim izborom u edukacijskom kontekstu. Već postoji određen broj sustava preporučivanja namijenjenih korištenju u kontekstu učenja programiranja, poput sustava koji su razvili Cabada i sur. [17] koji preporuča studentu odabir vježbe iz programiranja na osnovu ocjena prethodnih vježbi. Svi ovi sustavi namijenjeni su studentima jer su oni korisnici sustava, ali ne postoji sustav koji omogućava nastavnicima odabir prikladne razvojne okoline. Nastavnik najčešće bira alat za cijelu grupu studenata uspoređujući samo tehničke karakteristike dostupnih alata, nakon kratke probe alata ili preko preporuke drugih nastavnika. Kako danas postoji velik broj razvojnih okolina i pri odabiru treba uzeti u obzir i brojne obrazovne aspekte zaključujemo da postoji potreba za ovakvim sustavom preporučivanja. Temeljenje sustava preporučivanja na novorazvijenom modelu omogućiti će da sustav prilikom preporučivanja u obzir uzme različite dimenzije (uključujući i one obrazovne) i ponudi sveobuhvatnu preporuku, te tako omogući nastavnicima kvalificiran odabir prikladne radne okoline.

Ključne riječi – model uspješnosti informacijskog sustava, vrednovanje razvojnih okruženja, obrazovni sustav preporučivanja

Literatura

- [1] W. H. DeLone, i E. R. McLean, "Information systems success: The quest for the dependent variable," u Information systems research, vol. 3(1), 1992, pp. 60-95.
- [2] W. H. DeLone, i E. R. McLean, "The DeLone and McLean model of information systems success: a ten-year update," u Journal of management information systems, vol. 19(4), 2003, pp. 9-30.
- [3] S. Petter, W. H. DeLone, i E. R. McLean, "Measuring information systems success: Models, dimensions, measures, and interrelationships," u European Journal of Information Systems, vol. 17(3), 2008, pp. 236–263.
- [4] A. Jeyaraj, "DeLone & McLean models of information system success: Critical meta-review and research directions," u International Journal of Information Management, vol. 54, 2020, pp. 102139.
- [5] I. Balaban, I., E. Mu, i B. Divjak, "Development of an electronic Portfolio system success model: An information systems approach," u Computers & Education, vol. 60(1), 2013, pp. 396-411.
- [6] A. Yuce, A.M. Abubakar, i M. Ilkan, "Intelligent tutoring systems and learning performance: Applying task-technology fit and IS success model," u Online Information Review, 2019.
- [7] R. Halonen, H. Thomander, i E. Laukkanen, "DeLone & McLean IS success model in evaluating knowledge transfer in a virtual learning environment," International Journal of Information Systems and Social Change (IJISSC), vol. 1(2), 2010, pp. 36-48.
- [8] B. Soledad Fabito, L. R. Rodriguez, O. A. Trillanes, G. J. Lira, Z. D. Estocada, i Q. P. M. Sta Ana, u "Investigating the Factors influencing the Use of a Learning Management System (LMS):

- An Extended Information System Success Model (ISSM)," In 2020 The 4th International Conference on E-Society, E-Education and E-Technology, Taipei, Taiwan, 2020, pp. 42-46.
- [9] A. Luxton-Reilly, I. Albluwi, B.A. Becker, M. Giannakos, A. N. Kumar, , L. Ott, i C. Szabo, "Introductory programming: a systematic literature review," In Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education, 2018, pp. 55–106.
 - [10] J. Á. Velázquez-Iturbide, I. Hernán-Losada, i M. Paredes-Velasco, "Evaluating the effect of program visualization on student motivation," u IEEE Transactions on Education, vol. 60(3), 2017, pp. 238-245.
 - [11] A. Unal, i F. B. Topu, "Effects of teaching a computer programming language via hybrid interface on anxiety, cognitive load level and achievement of high school students," u Education and Information Technologies, vol. 26(5), 2021, pp. 5291-5309.
 - [12] M. Valez, M. Yen, M. Le, Z. Su, i M. A. Alipour, "Student Adoption and Perceptions of a Web Integrated Development Environment: An Experience Report," u Proceedings of the 51st ACM Technical Symposium on Computer Science Education, 2020, pp. 1172-1178.
 - [13] C. J. Costa, i M. Aparicio, "Evaluating success of a programming learning tool," u Proceedings of the International Conference on Information Systems and Design of Communication, 2014. pp. 73-78.
 - [14] T. M. Ahmed i M. E. Seliaman, "Investigating the adoption and impact of e-learning in KSA: Prince Sattam Bin Abdulaziz university case study," u Journal of Theoretical and Applied Information Technology, vol. 95(11), 2017.
 - [15] W. S. Lin, i C. H. Wang, "Antecedences to continued intentions of adopting e-learning system in blended learning instruction: A contingency framework based on models of information system success and task-technology fit," u Computers & Education, 58(1), vol. 2012, pp. 88-99.
 - [16] G. Đurović, M. Holenko Dlab, i N. Hoić-Božić, "Educational recommender systems: An overview and guidelines for further research and development," u Croatian Journal of Education: Hrvatski časopis za odgoj i obrazovanje, vol. 20(2), 2018, pp. 531-560.
 - [17] R. Z. Cabada, M. L. B. Estrada, F. G. Hernández, R. O.Bustillos, i C. A. Reyes-García, "An affective and Web 3.0-based learning environment for a programming language," u Telematics and Informatics, vol. 35(3), 2018, pp. 611-628.

Matea Turalija

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
matea.turalija@inf.uniri.hr

Desetljeće razvoja metoda simulacija molekulske dinamike u paralelizaciji na heterogenim superračunalima

Posljednjih desetljeća računalne simulacije postale su nezamjenjiv alat za proučavanje i predviđanje fizikalnih i kemijskih procesa. Fenomeni koji se na ovaj način mogu proučavati kreću se od najmanjih mjerila udaljenosti, kao što je kvantna mehanika materije na nanoskali, do najvećih, poput svemira, čiji se fenomeni proučavaju na udaljenostima do nekoliko desetaka milijardi svjetlosnih godina. Slično tome, vremenske skale procesa kreću se od femtosekundi do nekoliko milijardi godina.

Širok raspon opisanih pojava pokazuje da se eksperimenti ne mogu uvijek izvesti na željeni način. U astrofizici, primjerice, modeli koji dovoljno dobro opisuju prirodu često su toliko složeni da se ne može pronaći analitičko rješenje. Štoviše, mogu se promatrati pojave koje je teško ili nemoguće izmjeriti, ili eksperimenti traju predugo ili prekratko da bi bili vidljivi.

Ovo su samo neki od razloga zašto se računalna simulacija u posljednjim desetljećima nameće kao treći pristup u znanosti, uz eksperimentalni i teorijski pristup. Tako simulacija omogućuje proučavanje pojava koje su prije bile nedostupne eksperimentu.

Sve veća složenost simulacija molekulske dinamike (MD) dovode do velikih zahtjeva za memorijom i računalnim vremenom. Međutim, brzi razvoj računalne tehnologije nadilazi te poteškoće i danas nam omogućuje izvođenje sve realističnijih simulacija. Stoga su trenutna istraživanja [1] uglavnom usmjerena na razvoj simulacijskih metoda i povezanih algoritama koji omogućuju najbrži mogući izračun dijelova pojedinog koraka simulacije, poput metode brzih multipola, Ewaldove mreže čestica, brzih Fourierovih transformacija i sl.

Paralelno programiranje na heterogenim sustavima također je privuklo veliku pozornost u posljednjih 15 godina. U paralelnim računalima, od nekoliko desetaka do nekoliko desetaka tisuća moćnih procesora može računati istovremeno i neovisno uz međusobnu komunikaciju za izvođenje operacija u razumnom vremenu. Međutim, heterogenost sustava dovodi do poteškoća u komunikaciji i balansiranju opterećenja između osnovnih procesora i koprocesora, među koje spadaju i grafički procesori. S obzirom na ogroman broj mogućih kombinacija osnovnih i grafičkih procesora u računalnim sustavima, optimalna raspodjela izračuna ostaje vruća tema istraživanja i danas [2].

Čak i uz korištenje najsuvremenijih algoritama i učinkovitih paralelnih implementacija, cijena izračuna sila između svih čestica velikih sustava (106 atoma) predstavlja novi izazov za računarstvo visokih performansi. Izuzetak su superračunala posebne namjene, poput Antona i MDGRAPE-a koja su masivna paralelna superračunala namijenjena isključivo za simulacije MD [3], [4]. Nedavni napredak u algoritmima, softveru i hardveru omogućio je superračunalima da probiju prag od 1 exaFLOPS-a, započinjući eru eksaskalarnih superračunala. Upravo ove godine predstavljeno je najjače superračunalo ikad, Frontier, koje je ujedno i prvo pravo superračunalo iz eksaskalarne ere i daje znanstvenicima bolji alat za rješavanje problema u širokom rasponu znanstvenih područja.

Važan i računski zahtjevan dio simulacija MD je izračun dugodometnih elektrostatskih interakcija. Danas prevladavajuća metoda za izračunavanje ovih interakcija je Ewaldova mreža čestica (engl. Particle Mesh Ewald, kraće PME). Metoda PME je prilagođena za paralelizaciju na skali superračunala prošle generacije i na njima omogućuje vrlo brzi

izračun navedenih interakcija. Međutim, za velike paralelne simulacije s više čvorova, PME postaje usko grlo skaliranja jer zahtijeva komunikaciju između svih čvorova. Kao posljedica toga, broj razmijenjenih poruka kvadratno raste s brojem uključenih čvorova. Kako bi se omogućile učinkovite i skalabilne biomolekulske simulacije na eksaskalarnim superračunalima, jasno je da je potrebna metoda s boljim svojstvom skaliranja. Metoda brzih multipola (engl. Fast Multipole Method, kraće FMM) je takva metoda [1]. Međutim, postojeće ograničenje ove metode je korištenje isključivo kockastih simulacijskih kutija. Naš budući rad na ovu temu uključio bi podršku za druge vrste simulacijskih kutija, poput rombičkog dodekaedra. Prednost takvih "krnjih" kocaka je smanjenje količine vode koju je potrebno simulirati, čime se poboljšavaju performanse simulacije. Kao prvi korak prema toj implementaciji razmatramo mogućnost simulacije kvadarskih simulacijskih kutija korištenjem nepravilnih podjela kutija u podkutije. U slučaju pravilne podjele u svakom se koraku kutija dijeli na 8 podkutija. Sustav je moguće generalizirati tako da se može podijeliti u 6 ili 4 podkutije u slučajevima kada to odgovara simuliranom sustavu. To bi omogućilo da se iz kvadra, koji se uvelike razlikuje od kocke, u nekoliko koraka može doći do podkutija u obliku kvadra koja su dobra aproksimacija kocaka.

U doba rastućeg obujma podataka kojima se bave znanstvenici u području računarstva, kartice mrežnog sučelja (engl. Network Interface Card, kraće NIC) koje povezuju računalo s računalnom mrežom mogu se dodatno koristiti za obradu podataka koje prenose. Stoga NIC pruža mogućnost smanjenja opterećenja osnovnog procesora tako da se funkcije obrade primljenog ili poslanog mrežnog prometa koje tradicionalno obavlja osnovni procesor sada izvode na samom NIC-u, nazvanom SmartNIC (engl. Smart Network Interface Card). Kako se obrada podataka proširila izvan poslužitelja i osobnih računala na velike, centralizirane podatkovne centre čije se mrežne brzine nastavljaju povećavati, trenutno dostižući brzine od 400 GB u sekundi, potreban je treći stupanj rasterećenja. Upravo su ovi zahtjevi za velikim podacima doveli do razvoja jedinica za obradu podataka ili kraće DPU-ova (engl. Data Processing Unit). Kako bismo omogućili prijenos MD na DPU-ove, zainteresirani smo za daljnji rad na razvoju nove MPI (engl. Message Passing Interface) biblioteke koja bi omogućila pretvorbu postojećih alata za simulaciju MD u alate koji dio izračuna izvode na DPU-ima. Motivacija za ovaj razvoj je učiniti ubrzanje molekulske dinamike, prethodno ograničeno na superračunalo Anton 2, dostupnim širokom rasponu računala visokih performansi.

Mogućnosti molekulske simulacije proširuju se s porastom korištenja eksaskalarnih superračunala. S trilijunom (10^{18}) izračuna u sekundi, takva superračunala moći će brzo analizirati goleme količine podataka i realističnije simulirati medicinske procese, klimatske učinke, seizmičke fenomene, neviđenu fiziku, temeljne sile svemira i mnogo više. Eksaskalarno računarstvo sigurno će imati dubok utjecaj na razumijevanje svijeta i života u nadolazećim desetljećima, a naš je doprinos u njegovom utjecaju na efikasnije i brže simulacije molekulske dinamike.

Ključne riječi – simulacija molekulske dinamike, tehnike paralelizacije, heterogeno računarstvo, superračunala

Literatura

- [1] B. Kohnke, C. Kutzner i H. Grubmüller, "A GPU-Accelerated Fast Multipole Method for GROMACS: Performance and Accuracy", Journal of Chemical Theory and Computation, vol. 16, no. 11, pp. 6938-6949, 2020.
- [2] S. Páll i dr., "Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS", The Journal of Chemical Physics, vol. 153, no. 13, p. 134110, 2020.

- [3] J. P. Grossman, B. Towles, B. Greskamp i D. E. Shaw, "Filtering, Reductions and Synchronization in the Anton 2 Network," 2015 IEEE International Parallel and Distributed Processing Symposium, pp. 860-870, 2015.
- [4] G. Morimoto i dr., "Hardware acceleration of tensor-structured multilevel ewald summation method on MDGRAPE-4A, a special-purpose computer system for molecular dynamics simulations", Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2021.

Ana Vrcelj

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
Građevinska tehnička škola Rijeka, Rijeka, Hrvatska
ana.vrcelj@student.uniri.hr

Korištenje igrifikacije u osnovnoškolskom i srednjoškolskom obrazovanju

Igrifikacija (engl. *gamification*) u obrazovanju predstavlja primjenu elemenata dizajna igara i principa igre u nastavi s ciljem povećanja motivacije i angažmana učenika što naknadno doprinosi i uspješnjem ostvarivanju ishoda učenja [1]. Iako je važno motivirati učenike tijekom nastave u školama, to je još važnije u online ili hibridnim modelima nastave koja je sve češća u suvremenom poučavanju [2].

Jedan od najčešćih opisa pojma igrifikacije je primjena elemenata igre i principa dizajna igre u kontekstu izvan igre. Važno je istaknuti da primjena igrifikacije nije dizajniranje edukativne digitalne igre ili uključivanje postojećih igara u proces učenja i poučavanja, već igrificiranje aktivnosti u kojima učenici sudjeluju koristeći razne elemente igre. To uključuje elemente kao što su bodovi, ljestvice poretka, bedževi, avatari, priče, nagrade, povratne veze i razine, a koje je moguće implementirati jednostavnim digitalnim alatima ili u sustavima za upravljanje učenjem. Iako se elementi igrifikacije mogu uvesti u nastavni proces bez digitalnih alata, napredak tehnologije omogućio je širu primjenu igrifikacije u obrazovanju pa se tako elementi igrifikacije osim u tradicionalnom okruženju učenja, mogu primjeniti i u *online* učenju ili u hibridnom učenju koje kombinira učenje licem u lice i *online* učenje [2].

Igrifikacija se može koristiti u različitim kontekstima i na različitim razinama obrazovanja od osnovnih i srednjih škola do sveučilišta te programa cjeloživotnog učenja. Iako bi se moglo očekivati da će se elementi igre češće koristiti u aktivnostima za mlađe učenike, analiza literature pokazuje da je igrifikacija prisutnija u sveučilišnom obrazovanju nego u osnovnom i srednjem obrazovanju [1]. Unatoč navedenim činjenicama, iskustvo pokazuje da se igre i brojni digitalni alati uspješno koriste u školama u svrhu igrifikacije, iako to uglavnom nije potkrijepljeno odgovarajućim istraživanjima objavljenim u radovima.

Vlastitim pregledom literature žele se pronaći i analizirati relevantni radovi o primjeni igrifikacije u osnovnim i srednjim školama kako bi se dao doprinos istraživanjima u području igrifikacije u obrazovanju. Istraživanje predstavlja sustavni pregled literature (engl. Systematic Literature Review – SLR) prema metodologiji autora Okoli i Schabram ([3],[4]) koja je prikladna za kombinirana istraživanja iz društvenih i tehničkih znanosti, što je vrlo često u području e-učenja.

Svrha ovog SLR je istražiti područje igrifikacije u obrazovanju s naglaskom na osnovne i srednje škole kako bi se dale preporuke za buduća istraživanja. Svrha će se postići odgovorima na sljedeća istraživačka pitanja:

Q1: Za koju se razinu obrazovanja, model izvođenja nastave, predmet te metode i aktivnosti učenja i poučavanja koristi igrifikacija?

Q2: Koji se digitalni alati koriste kod igrifikacije?

Q3: Koji se elementi dizajna igre koriste za igrifikaciju?

Q4: Kojeg je tipa provedeno istraživanje o uspješnosti igrifikacije i koji je njegov cilj?

Q5: Ima li igrifikacija pozitivan utjecaj na učenike i na koji način?

Upit nad bazama podataka Scopus i Web of Science rezultirao je s ukupno 291 radom, a nakon primjene kriterija uključivanja i isključivanja, osnovnog pregleda i procjene kvalitete, za detaljnu analizu u ovom istraživanju uključeno je 20 radova [2].

Rezultati su pokazali da se igrifikacija većinom koristila u osnovnim školama za predmete iz prirodnog područja, ali vrlo često i za učenje stranog ili materinjeg jezika. Koristila se tijekom f2f nastave ili kod hibridnog modela te najčešće u svrhu formativnog vrednovanja. Nastavnici su u podjednakoj mjeri koristili javno dostupne digitalne alate i vlastite alate odnosno igrificirane sustave za e-učenje. Iako su korišteni različiti elementi dizajna igre, česta je bila kombinacija bodova s ljestvicama poretkom. Osim jednog, sva provedena istraživanja su bila kvantitativna (ponekad uz dodatak kvalitativnih) i ispitanici su većinom bili učenici. U gotovo svim radovima došlo se do zaključka kako igrifikacija ima pozitivan utjecaj, većinom na motivaciju učenika, ali i na uspješnije usvajanje ishoda učenja, a valja istaknuti kako niti jedna studija nije pokazala negativne utjecaje igrifikacije na učenike [2].

Zaključak je da bi se trebalo i dalje nastaviti s istraživanjima igrifikacije kako bi se predložili odgovarajući pedagoško-tehnološki okviri koji bi nastavnicima u školama olakšali primjenu igrifikacije.

Ključne riječi – igrifikacija u obrazovanju, elementi dizajna igre, digitalni alati za igrifikaciju, motivacija učenika, sustavni pregled literature (SLR)

Literatura

- [1] C. Dichev i D. Dicheva, „Gamifying education: what is known, what is believed and what remains uncertain: a critical review,“ International Journal of Educational Technology in Higher Education, 14:9 (1), 2017.
- [2] A. Vrcelj, „Korištenje igrifikacije u osnovnoškolskom i srednjoškolskom obrazovanju: sustavni pregled literature“ (neobjavljen kvalifikacijski rad), Sveučilište u Rijeci, 2022.
- [3] C. Okoli, „A Guide to Conducting a Standalone Systematic Literature Review,“ Communications of the Association for Information Systems: Vol. 37 , Article 43., 2015.
- [4] C. Okoli i K. Schabram, »Guide to Conducting a Systematic Literature Review of Information, Systems research,« Ssrn Electronic Journal, 10-26, 2010.

Đurđica Vukić

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
durdica.vukic@student.uniri.hr

Okvir za analizu obrazovnih programa temeljen na kompleksnim mrežama

Kvaliteta obrazovnog procesa postiže se efikasnim planiranjem nastavnog plana i programa, poravnanjem sadržaja poučavanja, metoda vrednovanja i ishoda učenja. U obrazovnom kontekstu ishodi učenja primjenjuju se unutar kvalifikacijskih okvira, kvalifikacijskih standarda, i kurikuluma (na razini programa, modula, kolegija ili nastavne jedinice) te dovode do kompetencija tj. sposobnosti studenta da samostalno primjenjuje znanja i vještine u praksi [6]. Uvođenje kompetencijskog pristupa i naglašavanje ishoda učenja ima ključnu ulogu u planiranju i programiranju odgojno-obrazovnog procesa, a predstavlja osnovu za uvođenje svake pojedine kvalifikacije [2]. Europski kvalifikacijski okvir (EQF) i svi nacionalni kvalifikacijski okviri (NQF) koji se na njega pozivaju slijede pristup ishoda učenja koji omogućava transparentnost kvalifikacije i poboljšava njezinu usporedivost između zemalja i unutar zemalja. Nadalje, na ovaj način podupire se bolja usklađenost između potreba za vještinama na tržištu rada i pružanja obrazovanja i ospozobljavanja, te olakšava vrednovanje učenja stečenog u različitim okruženjima i otvaraju mogućnosti za širi izbor putova učenja i iskustava [1].

Posljednjih godina Znanost o mrežama pronalazi sve veći značaj u istraživanjima u području obrazovanja. Fokus je uglavnom usmjeren na analizu strukture znanja učenika - unutarnje strukture konceptualnih reprezentacija koje učenici stječu, analizu društvenih relacija između učenika/sudionika u obrazovnom okruženju, analizu topološke strukture znanja i otkrivanje razlika u strukturalnim svojstvima reprezentacije znanja stručnjaka i učenika, primjenu mrežne analize u kontekstu izrade kurikuluma [3]-[5][8]. Cilj ovog istraživanja je razviti okvir za reviziju kvalitete strukture obrazovnih programa u pogledu organizacije znanja na razini tematskih jedinica znanja, kolegija/modula i obrazovnog programa (npr. cjeloživotnog, studijskog itd.), a koji bi podržao pristup ishoda učenja u transparentnosti kvalifikacije i njezinoj usporedivosti. Razvojni okvir temeljiti će se na primjeni teorije kompleksnih mreža za pregled i analizu strukture organizacije jedinica znanja predviđenih kurikulumom obrazovnog programa. U svrhu navedenog istraživanja izrađen je model podataka koji ilustrira ključne entitete procesa organizacije znanja unutar kurikuluma obrazovnog programa te odnos između istih. Na taj način omogućava se dohvaćanje relacijskih informacija, olakšava planiranje strukture podataka koja će podržati razvoj mrežnog modela i osigurava postavljanje korisne baze podataka za analizu podataka neovisno koji model primjenjujemo za reprezentaciju znanja. U razvoju mrežnog modela ključnu ulogu imaju koncepti (koncepti tipa prethodnik i koncepti tipa sljedbenik) koji uvjetuju usmjerenošć mreže te ishodi i razina kognitivnih procesa.

U ovoj fazi istraživanje temeljimo na analizi bipartitne mreže internacionalnog obrazovnog programa Veleri-OI IoT School [7] pri čemu je fokus na analizi epistemološki dobro definirane organizacijske strukture znanja navedenog programa i pronalaženju ključnih koncepata.

Također, u sljedećoj fazi istraživanja definirat će se metodologija za:

- istraživanje strukturalnih značajki mrežnog modela koje odražavaju kvalitetan dizajn kurikuluma.
- primjenjivost različitih metrika u analizi kohezivnosti te identifikaciji poželjnih putanja

- identificiranje zajednica specifičnih za domenu znanja.

U posljednjoj fazi istraživanja predstaviti će se postavke razvojnog okvira ključne za formiranje organiziranih i koherentnih struktura znanja u predmetnim domenama. Preslikavanje problema organizacije znanja unutar dizajna kurikuluma u okvir pogoduje posebnim oblicima strukturne revizije kvalitete i otkrivanja različitih koherentnih zajednica jedinica znanja u različitim disciplinama koje bi inače bilo teško ili uopće ne bi bilo moguće razotkriti.

Ključne riječi – obrazovni program, kurikulum, ER model, bipartitna mreža

Literatura

- [1] European Commission, Directorate-General for Employment, Social Affairs and Inclusion, The European qualifications framework: supporting learning, work and cross-border mobility: 10th anniversary, Publications Office, 2019, Available: <https://data.europa.eu/doi/10.2767/385613>
- [2] R. Beljo Lučić et al., Hrvatski kvalifikacijski okvir - Uvod u kvalifikacije, Ministarstvo znanosti, obrazovanja i sporta, 2009. Zagreb
- [3] Ireland, J., Mouthaan, M., "Perspectives on Curriculum Design: Comparing the Spiral and the Network Models," Research Matters, n30 pp. 7-12 Aut. 2020. [Online]. Available: <http://files.eric.ed.gov/fulltext/EJ1286824.pdf>
- [4] O'Meara, J., Ashwin V.,.. "A Network Theory Approach to Curriculum Design." Entropy (Basel, Switzerland) vol. 23, no.10, p.1346. 15 Oct. 2021, doi:10.3390/e23101346
- [5] P. Vasyl and Y. Holovatch. "Bipartite graph analysis as an alternative to reveal clusterization in complex systems." 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP). IEEE, 2018. pp. 84-87.
- [6] Ministarstvo znanosti, obrazovanja i sporta. Primjena ishoda učenja: četvrta publikacija iz serije o Europskom kvalifikacijskom okviru (EQF) [tisk], Zagreb, 2011.
- [7] Razvoj internacionalnog obrazovnog programa Veleri-OI IoT School (2021) [Online] Available: <https://iot-school.veleri.hr/>
- [8] Siew CSQ. "Applications of Network Science to Education Research: Quantifying Knowledge and the Development of Expertise through Network Analysis," Education Sciences. vol. 10, no. 4, pp. 101-117, April 2020. doi <https://doi.org/10.3390/educsci10040101>

Marina Žunić

Fakultet informatike i digitalnih tehnologija, Sveučilište u Rijeci, Rijeka, Hrvatska
marina.zunic@inf.uniri.hr

Primjena proširene stvarnosti u kombinaciji s igrom i igrifikacijom s ciljem razvoja računalnog razmišljanja u nižim razredima osnovne škole – pregled područja

Nova generacija učenika odrasta u digitalnom svijetu te se u vrlo ranoj dobi susreću s tehnologijom. Kako bi se procesi učenja i poučavanja učinkovito prilagodili karakteristikama novih generacija potrebno je kombinirati različite strategije poučavanja, a nastavne materijale i aktivnosti osmislići tako da odgovaraju različitim stilovima učenja učenika. Dodatno, valja razmislići kako prilagoditi okruženje za učenje različitim karakteristikama učenika s ciljem poticanja njihovog aktivnog sudjelovanja.

Recentna istraživanja u ovom području pokazuju da primjena tehnologije proširene stvarnosti (engl. *Augmented Reality*, AR) u nastavnom procesu omogućuje stvaranje efikasnog, a učenicima zanimljivog okruženja za učenje. Glavne prednosti ove suvremene tehnologije su povećanje motivacije učenika tijekom procesa učenja što rezultira njihovim aktivnim sudjelovanjem, boljom izvedbom i ostvarenjem ishoda učenja [1]. Također, primjećeno je da ova tehnologija utječe na poboljšanje vještina poput računalnog razmišljanja, vještine rješavanja problema, kritičkog razmišljanja, kreativnosti, komunikacije i suradnje. Radi se o vještinama koje su poznate kao vještine 21. stoljeća i koje je u današnje vrijeme potrebno razvijati među učenicima [2]. Među navedenim vještinama naglasak se stavlja na računalno razmišljanje, vještini koja predstavlja proces razmišljanja, rješavanja problem i izražavanje njegovog rješenja na strojno razumljiv način [3]. Ova vještina se nerijetko navodi kao vještina današnjice i budućnosti te se njen razvijanje promovira od najranije razine obrazovanja.

Pored navedenih prednosti valja napomenuti kako primjena proširene stvarnosti kod učenja uključuje većinu osjetila što povećava učinkovitost procesa učenja, pogotovo kod učenika koji preferiraju vizualni i kinestetički stil učenja [4]. Uzimajući navedeno u obzir, očekuje se da će proširena stvarnost kao inovativna tehnologija biti sve češće korištena u edukativne svrhe u nadolazećem razdoblju. Ova tehnologija je pogodna i za kombiniranje s pristupima učenju temeljenima na igri (engl. *Game Based Learning*, GBL) i igrifikaciji (engl. *gamification*) koji su zanimljivi za sve uzraste, a posebno za one najmlađe [5].

Ovaj rad daje pregled istraživanja o korištenju proširene stvarnosti u kombinaciji s igrama i igrifikacijom u obrazovanju te analizira i uspoređuje postojeće primjene na temelju nekoliko istraživačkih pitanja. Željelo se istražiti u kojoj mjeri se koristi proširena stvarnost u obrazovnom procesu za učenike nižih razreda osnovne škole. Zatim, u kojoj mjeri se potiče primjena računalnog razmišljanja u obrazovanju uz pomoć tehnologije proširene stvarnosti u kombinaciji s igrom i igrifikacijom te za koje predmete su istraživanja namijenjena.

Istraživanjem se utvrdilo da se u posljednjih nekoliko godina primjećuje porast istraživanja s ciljem korištenja proširene stvarnosti u obrazovanju, što ukazuje na važnost ove tehnologije i njezinog uključivanja u obrazovni proces. Dobiveni podatci su se analizirali te se utvrdilo da postoje istraživanja koja uvode proširenu stvarnost u niže razrede osnovne škole, no najveći broj istraživanja je ipak napravljen za akademsku razinu obrazovanja. Nadalje, najčešći glavni ciljevi uvođenja proširene stvarnosti u kombinaciji sa igrom i igrifikacijom u obrazovanju bili su angažman i motivacija učenika, njihovo aktivno sudjelovanje, efikasnije ostvarivanje ishoda učenja odnosno stjecanja znanja i vještina.

Važno je istaknuti kako se poticanje računalnog razmišljanja spomenulo u vrlo malom broju istraživanja. Pregled postojećih istraživanja prikazuje i da je najveći broj istraživanja proveden za školske predmete Priroda, Matematika i učenje jezika, dok su ostali predmeti pomalo zanemareni. U dijelu studija su korištene isključivo kvantitativne metode, u dijelu isključivo kvalitativne, dok su rezultati kod nekih studija prezentirani na temelju primjene kombinacije kvalitativnih i kvantitativnih metoda. No, iako su mnoge studije provedene na ovu temu, zanemarivo mali broj njih je za cilj imao utvrditi mogućnosti korištenja proširene stvarnosti u kombinaciji s igrama u nižim razredima osnovne škole s ciljem razvoja računalnog razmišljanja. Navedeno će se iskoristiti u usmjeravanju budućeg istraživanja.

U budućem radu će se provesti sustavni pregled literature o korištenju igri u kombinaciji s tehnologijom proširene stvarnosti za razvoj računalnog razmišljanja u nižim razredima osnovne škole. Rezultati će se iskoristiti kako bi se osmislio pedagoško-tehnološki okvir za promicanje primjene ove tehnologije u kombinaciji s igrama u različitim predmetima nižih razreda osnovne škole. Time se želi pridonijeti podršci nastavnicima da integriraju u svoju nastavnu praksu inovativne pristupe i tehnologije.

Ključne riječi – proširena stvarnost, igra, igrifikacija, računalno razmišljanje, niži razredi osnovne škole

Literatura

- [1] Phil Diegmann, Manuel Schmidt-Kraepelin, Sven Eynden, and Dirk Basten, “Benefits of Augmented Reality in Educational Environments – A Systematic Literature Review,” in 12th International Conference on Wirtschaftsinformatik, 2015, pp. 1542–1556.
- [2] Pei-Hsuan Lin and Shih-Yeh Chen, “Design and Evaluation of a Deep Learning Recommendation Based Augmented Reality System for Teaching Programming and Computational Thinking,” IEEE Access - Spec. Sect. Data Min. Internet Things, vol. 8, pp. 45689–45699, Jan. 2020.
- [3] Paul Curzon, “cs4fn and computational thinking unplugged,” in Proceedings of the 8th Workshop in Primary and Secondary Computing Education on - WiPSE ’13, 2013, pp. 47–50.
- [4] Jia Zhang, Amy Ogan, Tzu-Chien Liu, Yao-Ting Sung, and Kuo-En Chang, “The Influence of using Augmented Reality on Textbook Support for Learners of Different Learning Styles,” in IEEE International Symposium on Mixed and Augmented Reality, 2016, pp. 107–114.
- [5] R. Hammady, M. Ma, and N. Temple, “Augmented Reality and Gamification in Heritage museums,” in Joint International Conference on Serious Games, 2016, pp. 181–187.



ISBN 978-953-7720-67-4