



Sinteza

BOOK OF PROCEEDINGS

INTERNATIONAL SCIENTIFIC CONFERENCE ON
INFORMATION TECHNOLOGY AND DATA
RELATED RESEARCH



Publishing of Conference Proceedings of the International Scientific Conference on Information Technology and Data Related Research - Sinteza 2022 has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Belgrade
April 16, 2022.
sinteza.singidunum.ac.rs



SCIENTIFIC COMMITTEE

- ◆ Milovan Stanišić, Singidunum University, Serbia
- ◆ Endre Pap, Singidunum University, Serbia
- ◆ Aleksandar Jevremović, Singidunum University, Serbia
- ◆ Andreja Stojić, Institute of Physics in Belgrade, Serbia
- ◆ Bratislav Milovanović, Singidunum University, Serbia
- ◆ Dragan Cvetković, Singidunum University, Serbia
- ◆ Đorđe Obradović, Singidunum University, Serbia
- ◆ Goran Šimić, Military Academy, University of Defence, Serbia
- ◆ Gordana Dobrijević, Singidunum University, Serbia
- ◆ Gordana Jovanović, Institute of Physics in Belgrade, Serbia
- ◆ Ivan Ćuk, Singidunum University, Serbia
- ◆ Jelena Filipović, Faculty of Philology, University of Belgrade, Serbia
- ◆ Miloš Antonijević, Singidunum University, Serbia
- ◆ Marina Marjanović Jakovljević, Singidunum University, Serbia
- ◆ Marko Tanasković, Singidunum University, Serbia
- ◆ Marko Šarac, Singidunum University, Serbia
- ◆ Milan Milosavljević, Singidunum University, Serbia
- ◆ Milan Tair, Singidunum University, Serbia
- ◆ Milan Tuba, Singidunum University, Serbia
- ◆ Miloš Stojmenović, Singidunum University, Serbia
- ◆ Miodrag Živković, Singidunum University, Serbia
- ◆ Mirjana Perišić, Institute of Physics in Belgrade, Serbia
- ◆ Miroslav Popović, Singidunum University, Serbia
- ◆ Mladen Veinović, Singidunum University, Serbia
- ◆ Mladan Jovanović, Singidunum University, Serbia
- ◆ Nebojša Bačanin Džakula, Singidunum University, Serbia
- ◆ Nemanja Stanišić, Singidunum University, Serbia
- ◆ Nina Dragičević, Singidunum University, Serbia
- ◆ Petar Spalević, Faculty of Technical Sciences in Kosovska Mitrovica, University of Priština, Serbia
- ◆ Predrag Popović, Vinča Institute, Serbia
- ◆ Radosav Pušić, Faculty of Philology, University of Belgrade, Serbia
- ◆ Sanja Filipović, Singidunum University, Serbia
- ◆ Saša Adamović, Singidunum University, Serbia
- ◆ Tijana Radojević, Singidunum University, Serbia
- ◆ Valentina Gavranović, Singidunum University, Serbia
- ◆ Zora Konjović, Singidunum University, Serbia
- ◆ Alexandru Nedelea, Stefan cel Mare University of Suceava, Romania
- ◆ Aurora Pedro Bueno, Department of Applied Economics, University of Valencia, Spain
- ◆ Chen Yudong, Communication University of China, Beijing, People's Republic of China
- ◆ Deasún Ó Conchúir, Scatterwork GmbH, Ireland
- ◆ Diego Andina De la Fuente, Technical University of Madrid, Spain
- ◆ Duško Lukač, Rheinische Fachhochschule Köln – University of Applied Sciences, Germany
- ◆ Egons Lavendelis, Riga Technical University, Latvia
- ◆ Gordana Pesaković, Argosy University, USA
- ◆ Hong Qi, Dalian University of Technology, China
- ◆ Ivan Bajić, Simon Fraser University, Canada
- ◆ Ina Bikuviene, PhD Kauno kolegija- University of Applied Sciences, Kaunas (Lithuania)
- ◆ Jesus Amador Valdés Díaz de Villegas, Iberoamericana University, Mexico
- ◆ Jovica V. Milanović, University of Manchester, United Kingdom
- ◆ Juan Ruiz Ramírez, University of Veracruz, Mexico
- ◆ Li Liwen, Beijing Foreign Studies University, Beijing, PR China
- ◆ Lorenzo Fagiano, Polytechnico di Milano, Italy
- ◆ Luis de la Torre Cubillo, UNED, Dpt. Informática y Automática, Madrid, Spain
- ◆ Luis Hernández Gómez, Technical University of Madrid, Spain
- ◆ Maarten De Vos, University of Oxford, United Kingdom
- ◆ María Magdalena Hernández Alarcón, University of Veracruz, Mexico
- ◆ Mike Dawney, Middlesex University, United Kingdom
- ◆ Moe Win, Massachusetts Institute of Technology, USA
- ◆ Mohammed Ismail Elnaggar, The Ohio State University, USA
- ◆ Nataša Vilić, Filozofski fakultet, Univerzitet u Banja Luci, BiH
- ◆ Nellie Swart, University of South Africa, Pretoria
- ◆ Nuno Gonçalo Coelho Costa Pombo, University Beira Interior, Portugal
- ◆ Nuno Manuel Garcia dos Santos, University Beira Interior, Portugal
- ◆ Riste Temjanovski, Goce Delčev University, Macedonia
- ◆ Roberta Grossi, Horizons University, France
- ◆ Simona Distinto, Department of Life and Environmental Sciences University of Cagliari, Italia
- ◆ Slobodan Luković, ALaRI, Switzerland
- ◆ Snezana Lawrence, Bath Spa University, United Kingdom
- ◆ Stanimir Sadinov, Technical University of Gabrovo, Bulgaria
- ◆ Vassilis S. Moustakis, Technical University of Crete, Greece
- ◆ Violeta Grubliene, Klaipeda University, Lithuania
- ◆ Vladimir Terzija, University of Manchester, United Kingdom
- ◆ Yipeng Liu, University of Electronic Science and Technology of China, China

ORGANIZING COMMITTEE

- ◆ Milovan Stanišić
- ◆ Nebojša Bačanin Džakula
- ◆ Marko Tanasković
- ◆ Mladen Veinović
- ◆ Valentina Gavranović
- ◆ Miroslav Popović
- ◆ Nina Dragičević
- ◆ Miodrag Živković
- ◆ Ivan Ćuk
- ◆ Miloš Antonijević
- ◆ Mladan Jovanović
- ◆ Srđan Marković
- ◆ Milan Tair
- ◆ Aleksandar Mihajlović
- ◆ Petar Jakić
- ◆ Uroš Arnaut
- ◆ Miloš Mravik
- ◆ Jelena Gavrilović
- ◆ Predrag Obradović
- ◆ Jovana Maričić
- ◆ Miloš Višnjić

INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION TECHNOLOGY AND DATA RELATED RESEARCH

Publisher: Singidunum University, 32 Danijelova Street, Belgrade
Editor-in-Chief: Milovan Stanišić, PhD
Prepress: Miloš Višnjić, Jovana Maričić
Design: Aleksandar Mihajlović
Year: 2022
Circulation: 75
Printed by: Caligraph, Belgrade
ISBN: 978-86-7912-800-3

Contact us:
Singidunum University
32 Danijelova Street, 11010 Belgrade, Serbia
Phone No. +381 11 3093220, +381 11 3093290,
Fax. +381 11 3093294
E-mail: sinteza@singidunum.ac.rs
Web: sinteza.singidunum.ac.rs

Copyright © 2022

All rights reserved. No part of this work covered by the copyright herein may be reproduced, transmitted, stored or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, Web distribution, information networks, or information storage and retrieval systems, without the prior written permission of the publisher.



ABOUT SINTEZA 2022

The 9th International Scientific Conference Sinteza 2022 was held online following the current epidemiological measures due to the pandemic caused by the Covid 19 virus.

International Scientific Conference SINTEZA provided an ideal platform for exchanging information and disseminating best practices, ideas, and advancements in the state-of-the-art and technical improvements in Information Technology and data-related research.

Rapid advances in Information Technologies (IT) in recent decades have had a considerable impact on numerous facets of everyday life and have created tremendous opportunities for economic, technological and social gains on a global scale. In particular, the advances in data science, blockchain technology, and optimisation techniques are becoming the driving force behind many changes in both technology and business. The emergence of new technologies has caused widespread expansion of the internet of things. At the same time, problems related to cybersecurity, security of communications, and security in the cloud are becoming essential topics.

New technologies and scientific breakthroughs have already altered the working and living environments making them safer, more convenient and more connected. These scientific advances are also used to solve some of the most pressing problems our society faces today, such as climate change and environmental issues.

The conference sought submissions from academics, researchers, and industry professionals presenting novel research on all practical and theoretical aspects of Information Technology and Data Related Research and their applications in a range of business, engineering, environmental and research fields. Traditionally held each year, the conference features several prominent keynote speakers and presentations organised in thematic sessions covering topics such as computer science, information systems, IT security, IT and data science in environmental engineering, education and sports. In addition, there was a student session reserved for research work done by students (undergraduate, master and PhD level).

Sincerely,

Sinteza 2022 Organising Committee



CONFERENCE SPEAKERS IN PLENARY SESSION

Opening Keynote speakers:

Goranka Knežević, PhD – *Acting Rector of Singidunum University, Serbia*

Nebojša Bačanin Džakula, PhD – *Vice-Rector for Scientific Research, Singidunum University, Serbia*

Keynote speakers:

Jelena Luković, PhD – *University of Belgrade, Faculty of Geography, Serbia*

Simona Distinto, PhD – *University of Cagliari – Faculty of Biology and Pharmacy Sardinia, Italy*

Luis de la Torre Cubillo, PhD – *UNED, Departamento de Informática y Automática*

Mirjana Perišić – *Environmental Physics Laboratory, Institute of Physics Belgrade, Serbia*

Bulent Bayram – *Yildiz University, Department of Geomatic Engineering, Istanbul, Turkey*



CONTENTS



THEORETICAL COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE SESSION

-
- 3 - 13 | ANALYSIS OF BAYESIAN SYMBOLIC REGRESSION APPLIED TO CRUDE OIL PRICE
Krzysztof Drachal
- 14 - 19 | ANALYSIS OF SOCIAL MEDIA POSTS RELATED TO POSTPARTUM DEPRESSION
Ulfeta Marovac, Aldina Avdić
- 20 - 28 | APPLICATION OF NEURAL NETWORKS IN VIDEO GAME SIMULATION
Biljana Tešić, Marko Marković, Katarina Marković
- 29 - 35 | CONVOLUTIONAL NEURAL NETWORKS FOR REAL AND FAKE FACE CLASSIFICATION
Natalija Perišić, Radiša Jovanović
- 36 - 43 | GREY WOLF OPTIMIZATION FOR POSITION CONTROL OF A DIRECT CURRENT MOTOR DRIVEN
BY FEEDBACK LINEARIZATION METHOD
Mitra Vesović, Radiša Jovanović
- 44 - 48 | ON QUASI-CYCLIC CODES FF INDEX $1\frac{1}{2}$
Biljana Radičić
- 49 - 57 | POINT OBJECT EXTRACTION FROM SCANNED TOPOGRAPHIC MAPS FOR THE DIGITAL TOPOGRAPHIC
MAPS PRODUCTION
Miloš Basarić, Marko Mrlješ, Saša Bakrač
- 58 - 64 | PREDICTION OF GOLD PRICE MOVEMENT CONSIDERING THE NUMBER OF INFECTED WITH THE COVID-19
Jovana Stokanovic Sevic, Ana Jovancai Stakic
- 65 - 73 | REGRESSION ANALYSIS APPLIED TO INFLUENCE OF TRAFFIC FLOW PARAMETERS ON EMISSIONS
OF COMBUSTION PRODUCTS
Milan Marković, Dragana Velimirović
- 74 - 80 | CFD ANALYSIS APPLIED TO SMOKE EXTRACTION IN CAR PARK CAUSED BY SINGLE FIRE SOURCE
Milan Marković, Andrijana Đurđević
- 81 - 88 | THE PROCESS OF TRAINING A GENERAL-PURPOSE AUDIO CLASSIFICATION MODEL
Petar Petrović, Nemanja Čoso, Sanja Maravić Čisar, Robert Pinter
- 89 - 95 | SOFTWARE FOR RELIABILITY TESTING OF THERMOMECHANICAL INSTALLATIONS
Milan Marković, Andrijana Đurđević



IT APPLICATIONS IN EDUCATION SESSION

- 98 - 103 | AUTONOMOUS GRADER – A NEW DIGITAL SUMMATIVE ASSESSMENT TOOL
Mladen Vidović, Radmila Suzić, Ivan Radosavljević, Nebojša Nešić
- 104 - 108 | HIGH SCHOOL STUDENTS' COMMON ERRORS IN PROGRAMMING
Davorka Radaković, William Steingartner
- 109 - 113 | IMPORTANCE OF STUDYING COMPUTER GRAPHICS AT FACULTIES FOR COMPUTER AND
INFORMATICS ENGINEERS
Dejan Viduka, Luka Ilić, Davor Vrandečić
- 114 - 120 | M-LEARNING IN SERBIAN SECONDARY EDUCATION – A COVID-19 SIDE EFFECT?
Veljko Aleksić, Dionysios Politis
- 121 - 128 | TEACHING TRIGONOMETRY ONLINE USING GEOGEBRA DYNAMICAL MATERIALS
Tanja Sekulić, Goran Manigoda, Valentina Kostić
- 129 - 136 | THE IMPORTANCE OF INTRODUCING INNOVATIONS IN THE TEACHING OF INFORMATICS AND
COMPUTER SCIENCE IN PRIMARY SCHOOL
Đorđe Šarčević
- 137 - 143 | UNIVERSITY STUDENTS' EXPERIENCES AND ATTITUDES TOWARD ONLINE LEARNING DURING
THE PANDEMIC COVID-19
Ilija Savić, Slavko Alčaković, Angelina Njeguš
- 144 - 151 | MULTIMEDIA STORYTELLING ON THE WEB
Andreja Samčović



ADVANCED ENGINEERING SYSTEMS AND SOFTWARE DEVELOPMENT SESSION

- 154 - 162 | AN APPROACH FOR SOFTWARE DESIGN AND DEVELOPMENT
Miloš Cvetanović, Zaharije Radivojević, Stefan Tubić
- 163 - 168 | A SUMMARY PROTOCOL ON HOW TO DEVELOP A REMOTE LABORATORY
Luis de la Torre, Milan Matijević, Đorđe Seničić, Maja Milojević, Marko Tanasković
- 169 - 175 | AN IMPLEMENTATION OF A WEB LABORATORY: CONVERTING OFF-LINE EXPERIMENTS INTO
REMOTELY ACCESSIBLE EXPERIMENTS
Đorđe Seničić, Milan Matijević, Marko Tanasković, Luis de la Torre
- 176 - 183 | AVAILABILITY OF NGA MARKET POTENTIAL INDICATORS IN THE WESTERN BALKANS
Slobodan Mitrović, Valentina Radojičić, Goran Marković
- 184 - 190 | COMPARATIVE STUDY OF TWO PROPOSED METHODS FOR IMAGE ENHANCEMENT OF ABNORMAL
MRI BRAIN IMAGES IN SPATIAL AND FREQUENCY DOMAIN
Jelena Cerovina, Ivana Milošević, Petar Spalević, Mile Petrović, Haider Rashid Swadi
- 191 - 198 | IMAGE PROCESSING COMING FROM AGRICULTURAL DRONES – IT SOLUTIONS, LEGAL OBSTACLES
Žaklina Spalević, Aleksandra Stojnev Ilić, Miloš Ilić, Petar Spalević
- 199 - 204 | PROTOTYPE OF AUTONOMOUS WATER MANAGEMENT SYSTEM IN MOUNTAINOUS HOMESTEAD
Miloš Dobrojević
- 205 - 210 | THE EDUCATION OF DIGITAL GAME TESTERS – PRACTICAL EXPERIENCE
Veljko Aleksić
- 211 - 220 | THE PERSPECTIVE OF RADIO BROADCASTING
Jovan Zlatanović, Marina Marjanović



INFORMATION SYSTEMS AND SECURITY SESSION

- | | |
|-----------|--|
| 224 - 229 | THE APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS FOR FINGERPRINT RECOGNITION:
A COMPARATIVE ANALYSIS
Srđan Barzut, Milan Milosavljević |
| 230 - 237 | A MODEL FOR DYNAMIC CYBER SECURITY RISK ASSESSMENT IN THE INDUSTRIAL IOT ENVIRONMENT
Mirjana D. Stojanović, Jasna D. Marković-Petrović |
| 238 - 245 | SELECTING EFFECTIVE COLLABORATIVE TOOLS FOR CREATING A NEW SOCIAL NETWORK IN
THE COMPANY FOR RAISING EMPLOYEE ENGAGEMENT DURING COVID-19 PANDEMIC
Melanija Vasilijević, Srđan Atanasijević, Monika Zahar, Tatjana Atanasijević |
| 246 - 253 | SOFTWARE SYSTEM FOR SIMILARITY DETECTION IN THE PICOCOMPUTER ASSEMBLY PROGRAMS
Vojislav Tomašević, Marko Mišić, Violeta Tomašević |



IT APPLICATIONS IN SPORT SESSION

- 256 - 261 | CLOUD BASED MEDICAL PATIENT TRACKING SYSTEM (MPTS) FOR MULTIPLE SIMULTANEOUS RUNNING EVENTS
Siniša Malinović, Zora Konjović
- 262 - 268 | DYNAMIC STABILITY ASSESSMENT FOR MONITORING RECOVERY FROM ACL RECONSTRUCTION: A HANDBALL CASE STUDY
Aleksandar Gadžić, Aleksandar Živković, Vladimir Banković, Nenad Trunić
- 269 - 274 | POWER OUTPUT IN RUNNING – ASSOCIATION WITH RUNNING SPEED, ALTITUDE AND HEART RATE
Ćuk Ivan, Marković Srđan, Denić Lazar
- 275 - 280 | SCOUTING IN BASKETBALL
Nenad Trunić, Miodrag Milovanović
- 281 - 286 | SOFTWARE SOLUTIONS FOR STUDYING MENTAL FATIGUE
Branislav Božović, Miloš Milošević
- 287 - 291 | THE INFLUENCE OF PHYSICAL ACTIVITY ON THE HEALTH AND PLAYING QUALITY OF THE E-SPORTS PLAYERS
Anastasija Kocić, Branislav Božović, Aleksandar Vićentijević, Jefimija Kocić, Miloš Milošević
- 292 - 299 | THE USE OF IT TECHNOLOGY IN SPORTS COMPETITIONS
Igor Radošević, Lazar Denić



IT APPLICATIONS IN LANGUAGE TEACHING SESSION

- 302 - 308 | **BLENDED LEARNING APPROACH TO TEACHING ESP AT TERTIARY LEVEL**
Nataša Stanišić, Tijana Gajić, Ivana Đerić, Maja Rončević
- 309 - 315 | **E-BOOKS IN FOREIGN LANGUAGE TEACHING**
Magdalena Duvnjak, Jovan Travica, Maja Rončević, Miloš Pupavac
- 316 - 321 | **MEDIA LITERACY AND CYBER SECURITY: THEORY AND PRACTICE IN EDUCATION**
Zsolt Csutak
- 322 - 328 | **STUDENTS' ATTITUDES TOWARDS A SPECIALLY DESIGNED LANGUAGE LEARNING PLATFORM
IN A RUSSIAN ONLINE LANGUAGE SCHOOL**
Ana Tripković
- 329 - 334 | **TEACHING SOCIOLINGUISTICS IN A HYBRID CLASSROOM**
Valentina Bošković Marković, Nina Pantelić
- 335 - 340 | **THE COMPARISON OF MOST POPULAR MOBILE APPLICATIONS FOR LEARNING FOREIGN
LANGUAGES**
Tijana Gajić, Neda Maenza
- 341 - 346 | **THE DISADVANTAGES OF ONLINE FOREIGN LANGUAGE TEACHING AT THE UNIVERSITY LEVEL**
Valentina Bošković Marković, Maja Veljković Michos, Darija Lunić, Katarina Nasradin, Miloš Pupavac
- 347 - 352 | **THE USE OF KAHOOT! IN FOREIGN LANGUAGE TEACHING**
Jelena Nikolić, Jovana Nikolić, Tijana Gajić, Aleksandra Gagić
- 353 - 357 | **UTILIZING LUDIC FOREIGN LANGUAGE PEDAGOGY AT THE TERTIARY LEVEL**
Valentina Gavranović, Maja Veljković Michos



COMPUTER SCIENCE IN DRUG DESIGN IN PHARMACY AND MEDICINE SESSION

- 360 - 364 | IN SILICO APPROACHES FOR DRUG DISCOVERY: FOCUS ON VIRTUAL SCREENINGS
Simona Distinto, Elias Maccioni
- 365 - 372 | ANTI-INFLAMMATORY SCREENING OF THIOUREA DERIVATIVES BASED ON MOLECULAR DOCKING
STUDIES
Miloš Nikolić, Marina Mijajlović, Nikola Nedeljković
- 373 - 378 | THE USE OF ASR TO MAKE CLINICAL DOCUMENTATION IN SERBIAN
Aldina Avdić, Ulfeta Marovac, Dragan Janković



DECIDE PROJECT SESSION

- | | |
|-----------|--|
| 382 - 388 | AUTONOMOUS DRONE CONTROL FOR VISUAL SEARCH BASED ON DEEP REINFORCEMENT LEARNING
Uroš Dragović, Marko Tanasković, Miloš S. Stanković, Aleksa Ćuk |
| 389 - 393 | DISTRIBUTED ON-POLICY ACTOR-CRITIC REINFORCEMENT LEARNING
Miloš S. Stanković, Marko Beko, Miloš Pavlović, Ilija Popadić, Srđan S. Stanković |
| 394 - 398 | IMAGE SEGMENTATION PROCESSING FOR THERMOGRAPHIC ANALYSIS
Ivana Walter, Marko Tanasković |
| 399 - 403 | MOBILE ROBOT PATH PLANNING OPTIMIZATION BY ARTIFICIAL BEE COLONY
Aleksa Ćuk, Jelena Gavrilović, Marko Tanasković, Miloš S. Stanković |



STUDENT SESSION

- 406 - 413 | ARITHMETIC OPTIMIZATION ALGORITHM FOR SPAM DETECTION
Luka Jovanović (Student), Bushra Al Barwani, Esraa Al Maani, Miodrag Živković (Mentor),
Nebojša Bačanić
- 414 - 420 | ARTIFICIAL INTELLIGENCE APPLICATIONS IN HUMAN RESOURCE MANAGEMENT
Teodora Čormarković (Student), Lazar Dražeta (Mentor)
- 421 - 426 | CYBER SECURITY AND DOMAIN NAME SYSTEMS DEPLOY AND PROTECT NETWORK WITH DNS
SINKHOLE/BLACKHOLE
Aleksandar Jokić (Student), Marko Šarac (Mentor)
- 427 - 434 | DIGITAL FORENSICS ARTIFACTS OF THE MICROSOFT PHOTOS APPLICATION IN WINDOWS 10
Luka Jovanovic (Student), Saša Adamović (Mentor)
- 435 - 442 | GENDER INEQUALITY PRESENT IN THE GAMING COMMUNITY
Anđela Drča (Student)
- 443 - 448 | PLANT CLASSIFICATION USING FIREFLY ALGORITHM AND SUPPORT VECTOR MACHINE
Mina Milovanović (Student), Aleksandar Petrović, Hothefa Shaker Jassim, Ivana Štrumberger,
Miodrag Živković (Mentor)
- 449 - 456 | SDN APPROACH IN DEVELOPMENT IOT ENVIRONMENTS
Stefan Biševac (Student), Marko Šarac (Mentor)
- 457 - 463 | UNDERSTANDING JOB REQUIREMENTS USING NATURAL LANGUAGE PROCESSING
Luka Aničin (Student), Miloš Stojmenović (Mentor)



ABSTRACT

- 465 | EXAMINING THE SEASONAL CYCLE IN THE MEDITERRANEAN CLIMATES
Jelena Luković



ANALYSIS OF SOCIAL MEDIA POSTS RELATED TO POSTPARTUM DEPRESSION A SUMMARY PROTOCOL ON HOW TO DEVELOP A REMOTE LABORATORY

Ulfeta Marovac*,
Aldina Avdić

Department for Technical Sciences,
State University of Novi Pazar,
Novi Pazar, Serbia

Abstract:

Analysing the content of posts from social networks related to specific health problems can contribute to improving the health of the general population. This study gives an analysis of posts related to postpartum depression, which was performed to automatically detect content that correlates with postpartum depression. Machine learning methods can be used to detect posts that correlate with postpartum depression. The specificity of the language in which the posts are written reduces the availability of training corpora and processing tools. In this paper, a topic analysis is provided and a model for the prediction of postpartum depression in posts using a corpus composed of posts from the Reddit and ana.rs forums is presented.

Keywords:

Social media, Postpartum depression, Machine learning, Topic analysis.

Correspondence:

Ulfeta Marovac

e-mail:

umarovac@np.ac.rs

INTRODUCTION

Caring for women in the postpartum period is very important, but due to the care of the baby, women often neglect their physical and mental health. Very often, health problems are tried to be solved through advice on social networks.

How do recognize the state of postpartum depression? This question should be answered by psychiatrists and psychologists. Postpartum depression signs and symptoms may include: depressed mood or severe mood swings, excessive crying, difficulty bonding with the baby, withdrawal from family and friends, loss of appetite or eating much more than usual, inability to sleep (insomnia) or sleeping too much, overwhelming fatigue or loss of energy, reduced interest and pleasure in usual activities, intense irritability and anger, fear of unsuccessful parenting, hopelessness, feelings of worthlessness, shame, guilt or inadequacy, etc.



To determine postpartum depression (PPD), the Edinburgh scale is used, with a score greater than 12 indicating postpartum depression. EPDS is one of the most commonly used scales for assessing depressive symptoms of women who have given birth. The respondent estimates the weight for ten different depressive symptoms in the past seven days on a scale of 0 to 3 [1]. The possible range of results is from 0 to 30, where higher scores refer to more difficult symptomatic.

Detection of posts in which the symptoms of postpartum depression are described is one of the goals of this paper. Such research has already been done, but for posts in English. However, people express themselves most sincerely in their mother tongue, so it is necessary to adapt these analyses to other languages as well. In this paper, we will perform an analysis of posts in English and posts in Serbian translated into English. The presented methodology uses the tools available for the English language and adapts their use to the Serbian language.

The second section presents an overview of similar research related to the application of machine learning in predicting postpartum depression in fasts. The third section provides an overview of the materials used and the method. An analysis of the topic of the posts as well as the results of the methods for predicting postpartum depression in the posts are given in the fourth section. Finally, a conclusion is given and ideas for further research are presented.

2. RELATED WORK

In the era of the Internet and the social network of Facebook, Twitter, and Reddit, conditions have been created to collect large amounts of textual data through which it is possible to monitor personality behavior in posts on social networks. Thus, the context of the analysis of depression-related chatter on Twitter to glean insight into social networking about mental health was performed [2]. In particular, the authors in [3] analysed the posts on social networks that correlated with postpartum depression and showed that very good results can be obtained in predicting postpartum depression in posts.

Social support in the postpartum period directly affects the birth rate. Through social networks and PPD support groups, opportunities are created for women to share their experiences and receive support [4]. An analysis of changes in the mood of mothers before and after childbirth on Twitter created a model for predict-

ing mood [5]. The possibility of applying machine learning to predict postpartum depression was investigated over the corpus created from Reddit posts in [6]. The characteristic of that most similar research is that they are related to the English language, and non-English social media are quite unexplored. A model for diagnosing postpartum depression via the crowdsourcing platform has been proposed for Serbian, which also includes an automated test for determining the degree of postpartum depression using the Edinburgh scale [7]. This model has been expanded with the detection system of posts that are correlated with postpartum depression by detecting potential users of the crowdsourcing PPD platform on various social networks [8].

3. DATA AND METHODS

For this research, two sources were used: the Reddit forum for posts in English and the ana.rs forum for posts in Serbian (Table 1). Data were collected from both sources according to two criteria:

1. posts related to postpartum depression;
2. posts related to pregnancy and the postpartum period that are not related to postpartum depression.

Data set	Label	Source	URL	Post language
1.	DS1	Reddit	https://www.reddit.com/	English
2.	DS2	ana.rs	https://www.ana.rs/forum/	Serbian

Table 1 – Data sources

The first data set (DS1) consists of Reddit posts from the postpartum depression section (community about: “A non-judgemental place for you to ask for help and vent your frustrations on anything related to issues postpartum, be they hormonal, parental or other mental health issues. PPD, PND, PPA, PPOCD, APD etc. ”) and happy sections (community about: “Too many depressing things on the main page, so post about what makes you warm and fuzzy inside! ”) which contain the word “pregnancy”. They were collected using the Pushshift API. 150 posts from the postpartum depression group were singled out, and 150 posts from the happy group.

The second set of data (DS2) is from the forum ana.rs, which is one of the largest women's forums in Serbian. This forum contains the topic of postpartum depression from which 150 posts have been selected.

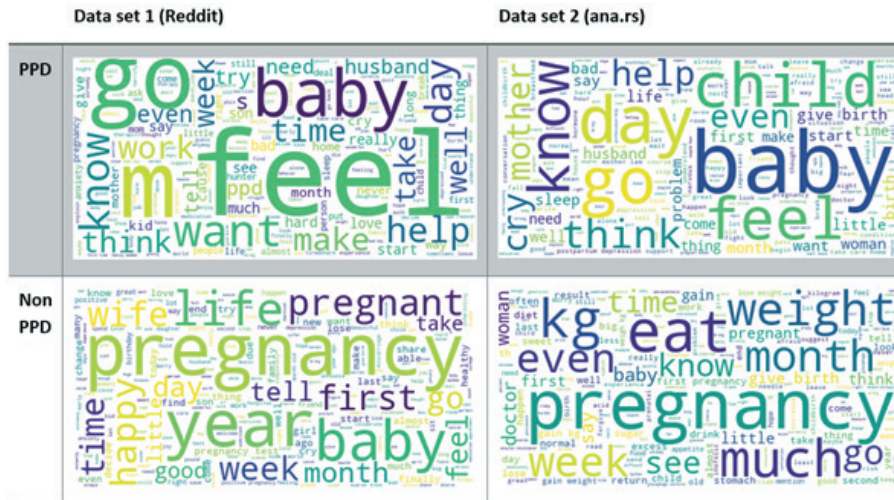


Figure 1 – Word clouds for contents of PPD post and non-PPD from datasets DS1 and DS2

Also, 150 posts unrelated to postpartum depression were singled out from the sections in “Pregnancy and Weight” and “Pregnancy Chat Corner”.

Posts are marked according to the section to which they belong to the group of PPD-related posts and the group of posts not related to PPD. Both sets were manually checked, and posts related to postpartum depression are marked.

3.1. DATA PRE-PROCESSING

Data collected from Reddit are in JSON format, while data from the ana.rs forum are manually processed to CSV format. To achieve data uniformity, both sets went through the preprocessing step.

1. JSON data set (DS1) is transformed into CSV format;
2. The set of data in Serbian (DS2) has been translated into English;
3. Punctuation marks are removed from the data
4. Removals are stop words;
5. Lemmatization

Before processing, it was done using python libraries (*nltk*, *pandas*, *gensim*, *spacy*...). A neural machine translation service¹ that is part of the Azure Cognitive Services family of REST APIs was used for the language translation Serbian post into English. The quality of Microsoft Translator’s machine translation outputs are evaluated using a method called the BLEU score.

1 api.cognitive.microsofttranslator.com

BLEU is a measurement of the differences between an automatic translation and one or more human-created reference translations of the same source sentence. In the medical domain that corresponds to the topic of Microsoft Translator’s posts, it has a BLEU score of approximately 50, which is considered high-quality translation.²

3.2. TOPIC ANALYSIS

To analyse the contents of the posts, the topic analysis was done for all four groups (two datasets DS1 and DS2 divided by groups if they are related to PPD or not). The topic analysis is performed by using the currently most used LDA (Latent Dirichlet Allocation) topic model developed by David Blei, Andrew Ng i Michael I. Jordan [9]. The python library - *gensim* to construct an LDA model is used.

3.3. CLASSIFICATION METHODS

For the classification of posts into PPD posts and non-PPD posts, methods of classification using Weka tools were performed. Tokenization was performed on the textual data. Sequential minimal optimization (SMO), J48, and RandomForest machine learning methods were applied. Accuracy (Equation 1) was used as the primary measure to gauge the performance of each model. In equation TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively.

2 microsoft.com/en-us/translator



$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Equation 1 - Accuracy

4. RESULTS

This paper aims to analyse the posts related to postpartum depression. Figure 1 contains visual results of the most common words in all four groups of data: PPD and non-PPD posts from DS1 and DS2.

The python *wordcloud* library was used to display the most common words. We can notice that translating from the Serbian data set into English gives a similar word set as in the corresponding word set of English posts from Reddit. The keywords that predominate in PPD-related posts are: feel, baby help think know... While posts that come from groups that have non-PPD refer to: baby, pregnancy...

Using the LDA topic analysis, we attempted to extract topics within datasets. The high degree of coherence (DS1: 0.49, DS2: 0.42) is obtained at high alpha and

beta parameters, which indicates a great connection of posts within the set. The best results (high coherence and separation of posts by topic) were obtained for 4 topics (Figure 2, Figure 3).

However, the words that appear in them are still elements of other topics. We obtained similar results at both sets for posts related to postpartum depression. Visualizations of the results show that the terminology is slightly different in posts written in Serbian and posts written in English. English posts contain abbreviations, so they have PPD, while Serbian posts have postpartum depression.

The alpha parameter has values of 0.31 in dataset DS1 and 0.61 for DS1 while the beta parameter has values of 0.9 in both cases. From this we can conclude that each topic will probably contain a mixture of most words.

On a given dataset of posts, classification models were made using methods SMO, J48, and RandomForest (RF). A set of 52 posts from ana.rs from the sections "Postpartum depression", "Pregnancy and Weight" and "Pregnancy Chat Corner" were taken for verification.

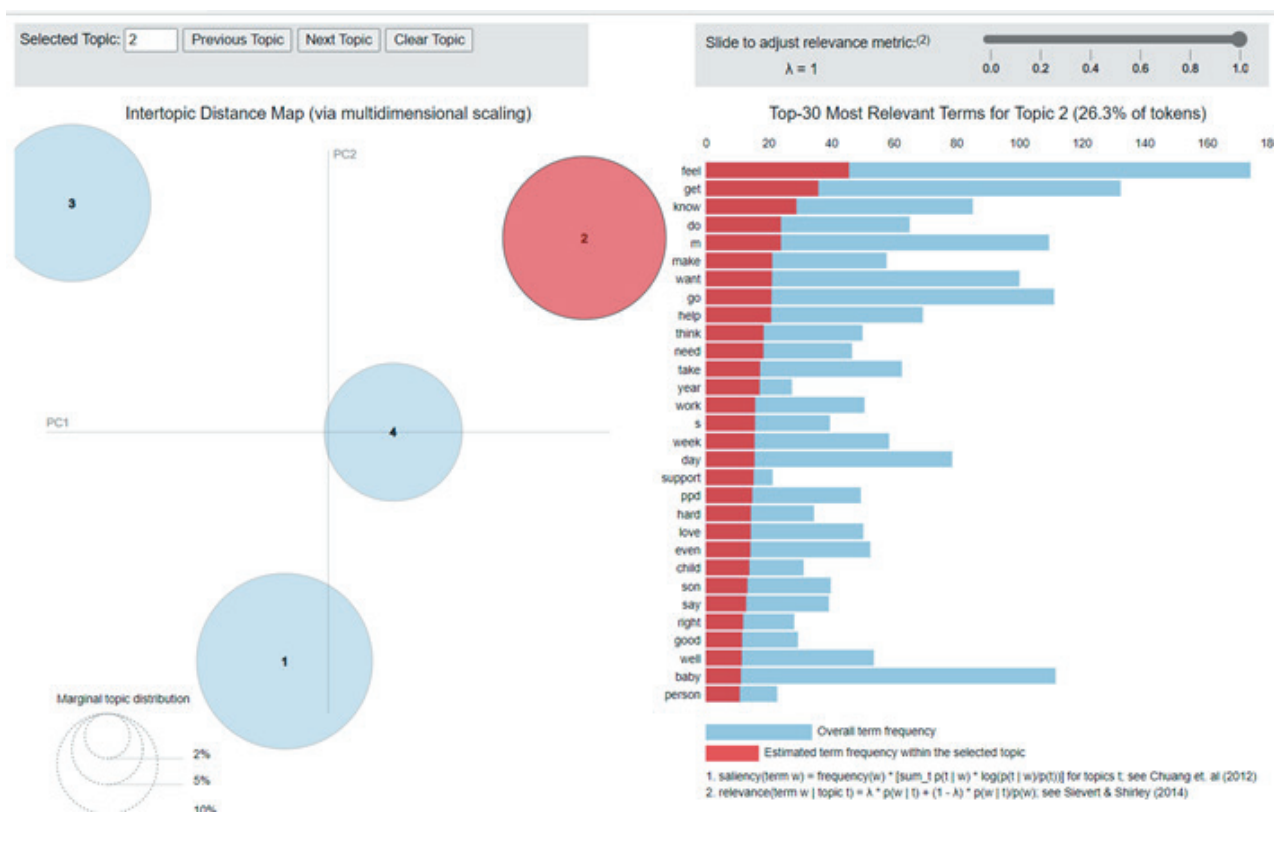


Figure 2 – 4 topics from PPD posts (dataset DS1)

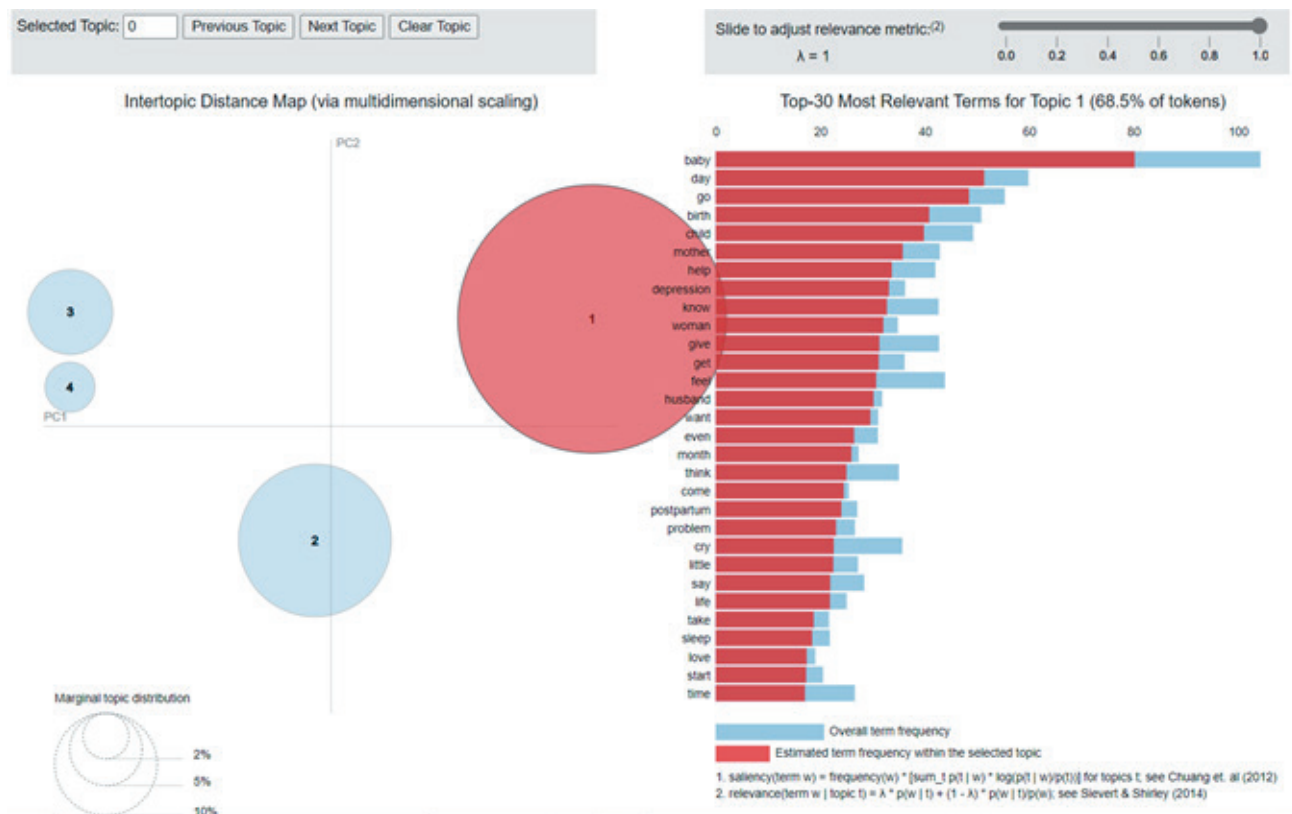


Figure 3 – 4 topics from PPD posts (dataset DS2)

The first model M1 was made over the posts of the Serbian forum ana.rs (DS2) while the second model M2 was made with posts from both forums (DS1 + DS2).

Model	Dataset	SMO	J48	RF
M1	ds2	92.16	86.27	90.20
M2	ds1+ds2	94.12	70.59	92.16

Table 2 – Classification models

Table 2 shows that by expanding the model with posts from Reddit, better results are obtained. This justifies the fact that for languages with fewer resources, English corpora and tools can alternatively be used. Translation errors affect the results of the classification and should be considered with caution. Anyway, the obtained accuracy is satisfactory and comparable to the results of PPD posts detection in English [3].

5. CONCLUSION

This paper presents an analysis of posts related to postpartum depression from the Serbian and English forums. Data processing was done in English. The datasets showed mutual similarity in content. The obtained classification models on a set of posts from the English and Serbian forums gave satisfactory results. The following research is related to the expansion of the set and the application of more detailed preprocessing methods to more accurately detect posts that are correlated with postpartum depression.

6. ACKNOWLEDGEMENTS

This paper is partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under project III44007.



7. REFERENCES

- [1] J. L. Cox, J. M. Holden and R. Sagovsky, "Detection of postnatal depression. De-," *Br J Psychiatry*, vol. 150, pp. 782-786, 1987.
- [2] P. A. Cavazos-Rehg, M. J. Krauss, S. Sowles, S. Connolly, C. Rosas, M. Bharadwaj and L. J. Bierut, "A content analysis of depression-related tweets.," *Computers in human behavior*, vol. 54, pp. 351-357, 2016.
- [3] I. Fatima, B. U. D. Abbasi, S. Khan, M. Al-Saeed, H. F. Ahmad and R. Mumtaz, "Prediction of postpartum depression using machine learning techniques from social media text.," *Expert Systems*, vol. 36, no. 4, p. e12409, 2019.
- [4] M. Evans, L. Donelle and L. Hume-Loveland, "Social support and online postpartum depression discussion groups: A content analysis.," *Patient education and counseling*, vol. 87, no. 3, pp. 405-410.
- [5] M. D. Choudhury, S. Counts and E. Horvitz, "Predicting postpartum changes in emotion and behavior via social media.," in *In Proceedings of the SIGCHI conference on human factors in computing systems*, 2013.
- [6] A. Trifan, D. Semeraro, J. Drake, R. Bukowski and J. L. Oliveira, "Social media mining for postpartum depression prediction," In *Digital Personalized Health and Medicine*, pp. 1391-1392, 2020.
- [7] U. Marovac, A. Ljajić, A. Avdić and A. Fazlagić, "Automation of psychological testing of stressful situations in the Serbian," in *ICIST 2019 Proceedings.*, Kopaonik, 2019.
- [8] U. Marovac and A. Avdić, "Detection of postpartum depression-related," in *The 14th International Symposium on Intelligent Distributed Computing*.
- [9] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, no. 3, pp. 993-1022, 2003.