BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Kibertámadások előre jelzése historikus adat alapú előre jelző rendszerrel és Fuzzy logika alapú kockázatbecsléssel

Prediction of cyber attacks with historical data-based forecasting system and Fuzzy logic-based risk estimation

Krasnyánszki Brúnó*, Prof. Dr. Zlatko Čović**, Prof. Dr. Rajnai Zoltán***
* Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar, Budapest, Magyarország
** Szabadkai Műszaki Szakfőiskola, Szabadka, Szerbia
*** Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar, Budapest, Magyarország
brunokrasnyanszki@stud.uni-obuda.hu, chole@vts.su.ac.rs , rajnai.zoltan@bgk.uni-obuda.hu

***Összefoglalás* — Kutatásunkban az EuRepoC adatbázisát használtuk a jövőbeli kibertámadások előrejelzésére. Neurális hálózatokat, támogató vektorgép modelleket és Fine Tree algoritmust használtunk. Emellett egy Fuzzy alapú kockázatelemzőt is fejlesztünk az államilag támogatott kibertámadásokhoz. Kutatásunk folytatódik, így csak részeredmények vannak a cikkben.**

***Kulcsszavak*:** EuRepoC, Kibertámadások előrejelzése, predikció, AI, MI

***Abstract* — In our research paper we used the EuRepoC's database to make predictions for the future cyberattacks. We used neural networks, support vector machine modells and Fine Tree algorythm. We are also developing a Fuzzy based risk analyzer for state sponsored cyber attacks. Our research is continuing so there are only partial results in the paper.**

***Keywords*:** EuRepoC, Prediction of cyber attacks, prediction, AI

1 BEVEZETÉS

1.1 Motiváció

A kutatás ihletét az adta, hogy az idei félévben az Erasmus+ program során Prof. Dr. Zlatko Čović által részt vehettünk egy cég látogatáson a Studio Present cégnél Szabadkán. Ahol bemutatták, hogy ők fejlesztették a European Repository of Cyber Incidents (EuRepoC) weboldalát. Ezen weboldalon rengeteg korábbi támadást dokumentáltak és gyűjtöttek össze. Prof. Dr. Zlatko Čović-al való beszélgetésünk során azon kezdtünk el gondolkodni, hogy mit lehetne ezekkel az adatokkal kezdeni.

1.2 Kutatás hasznosulásának lehetőségei

Amennyiben kutatásunk eredményei hosszabb távon pozitív eredményt hoznak abban az esetben a korább mintázatok adatai alapján képesek lehetünk bizonyos napokat vagy időszakokat kockázatosabbnak prediktálni ezáltal előre tudunk készülni humán erőforrássokkal és egyéb intézkedésekkel

2 EUREPOC

„A European Repository of Cyber Incidents (EuRepoC) egy független kutatókonzorcium, amelynek célja a kiberfenyegetések környezetének jobb megértése az Európai Unióban és azon kívül. A 2022 novemberében elindított fő célunk az adatközpontú megbeszélések és politikaalkotás előmozdítása a kiberbiztonság területén, valamint a kiberbiztonsági fenyegetések tudatosítása. Ezt úgy érjük el, hogy elemzési keretet biztosítunk a kiberincidensek „életciklusának" értékeléséhez és összehasonlításához, a technikai, politikai és jogi szempontokra összpontosítva. Adataink és kutatásaink az érintettek széles köre számára relevánsak – beleértve a kormányzati tisztviselőket, a civil társadalom képviselőit, az üzleti élet vezetőit, az újságírókat, az oktatókat, a diákokat és a nagyközönséget." [1]

2.1 EuRepoC tevékenysége

2.1.1 Adatbázis készítése a kiberincidensekről

Az EuRepoC napi rendszerességgel dokumentálja a nyilvános incidenseket, valamint elemzéseket készít, amit elérhetővé tesz adatbázisában a weboldalukon.[1]

2.1.2 Kutatások és elemzések készítése

Az EuRepoc a kiberkonfliktusok trendjeit empirikus módszerekkel elemzi és akadémiai igényességgel készít cikkeket. Napi szinten is készít rövid jelentéseket, valamint konferenciákon megosztják az aktualitásokat a kiberincidenserőL.[1]

2.1.3 Európai hálózat építése

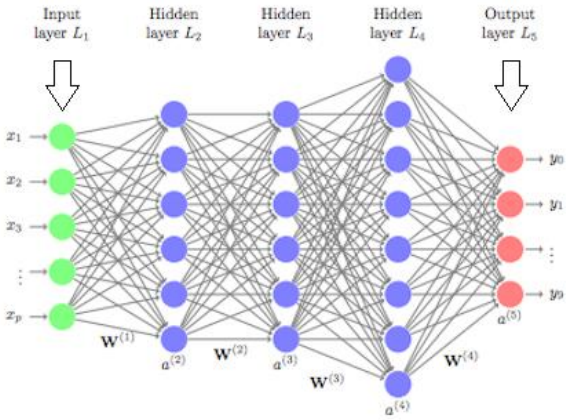
Speciális oktatási programokkal és tudományos kerekasztal-beszélgetésekkel erősítik a független kiberbiztonsági kutatószervezetek és kutatók európai hálózatát. [1]

3 TECHNOLÓGIAI BEVEZETŐ

Az alábbiakban kívánjuk ismertetni azokat a technológiákat, amelyeket felhasználtunk a kutatásunk során.

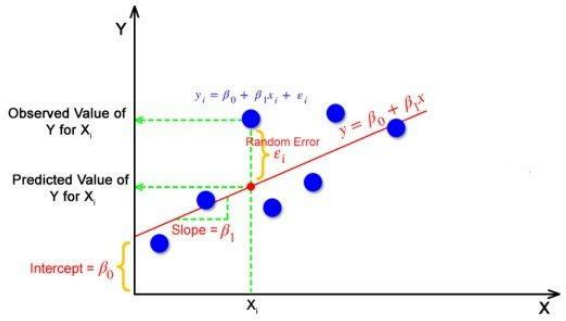
3.1 Gépi taunlás

A gépi tanulás az olyan algoritmusok és statisztikai modellek fejlesztésére összpontosító mesterséges intelligencia (AI/MI) egyik alágaként definiálható, amelyek lehetővé teszik a számítógépek számára, hogy konkrét feladatokat végezzenek el kifejezett utasítások nélkül. Ehelyett ezek a rendszerek a tapasztalatokból tanulnak és javulnak, az adatokban rejlő minták elemzésével és azonosításával. [2]



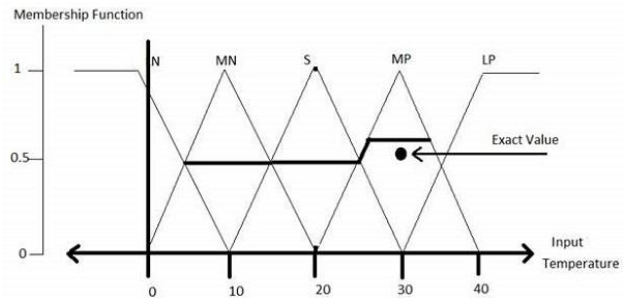
3.2 Predikció

"A predikció a gépi tanulásban egy olyan folyamat, amely során a modell a bemeneti adatok alapján megpróbálja előrejelezni a kimeneti értékeket. Ez a modell tanulási fázisában szerzett tapasztalatokra és a bemeneti adatokban található mintákra épül."[2]



3.3 Fuzzy Logika

"A fuzzy logika olyan logikaforma, amely közelítő, nem pedig rögzített és pontos következtetésekkel foglalkozik. Ez a Boole-logika kiterjesztése, amely a részleges igazság fogalmát kezeli – az 'teljesen igaz' és 'teljesen hamis' közötti igazságértékeket." [3]



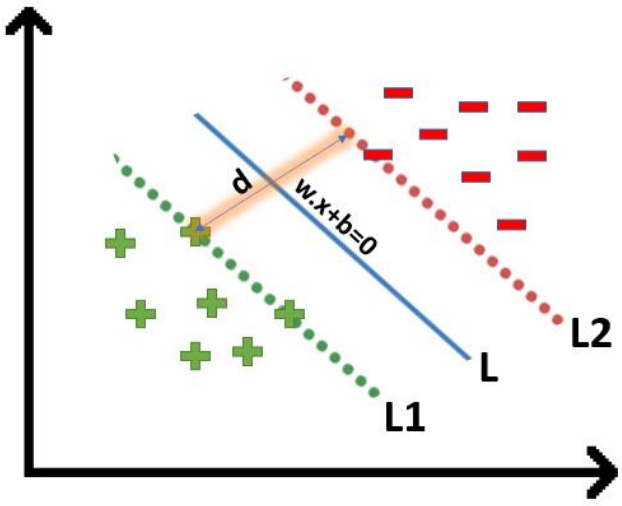
3.4 Fuzzy Logika alapú kockázatbecslés

A fuzzy alapú kockázatelemzés olyan módszer, amely a fuzzy logikát alkalmazza a kockázatok felmérésére és kezelésére, különösen bizonytalanságok és pontatlan adatok esetén. Ez a megközelítés a fuzzy halmazelméletet használja a kockázatok kvantifikálására és elemzésére a homályosság és kétértelműség kezelése révén, lehetővé téve a rugalmasabb és átfogóbb kockázatértékelést a hagyományos módszerekhez képest. [4]

Például a fuzzy kockázatelemzés egy esemény kockázatát nem egyszerűen "magasnak" vagy "alacsonynak" értékeli, hanem ezekbe a kategóriákba való tagság mértékeként, lehetővé téve a finomabb döntéshozatalt. [4]

3.5 Support Vector Machine

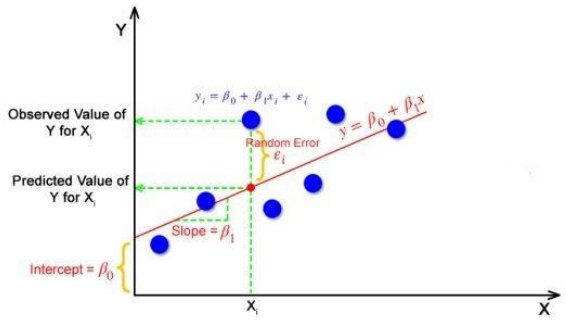
A támogatott vektorgép (Support Vector Machine, SVM) egy felügyelt tanulási algoritmus, amelyet elsősorban osztályozási és regressziós problémák megoldására használnak. Az SVM célja egy optimális választóvonal (hipersík) megtalálása, amely maximális távolságot (margin) tart fenn a különböző osztályokhoz tartozó adatok között. Az optimális hipersík az, amely a legjobban elválasztja az osztályokat, és a legközelebbi adatok (támogató vektorok) minimális távolságát maximalizálja a hipersíktól. [13]



3.6 Lineáris Regressió

A lineáris regresszió az egyik legismertebb és leggyakrabban használt statisztikai módszer az adatelemzésben. Célja, hogy két vagy több változó közötti kapcsolatot modellezze és kvantifikálja. Az alapvető cél

az, hogy egy egyenes vonallal (lineáris modellel) írja le az adatok közötti kapcsolatot. A lineáris regresszió matematikailag az alábbi egyenlettel írható le: [6][7][8]



3.7 Neurális hálók

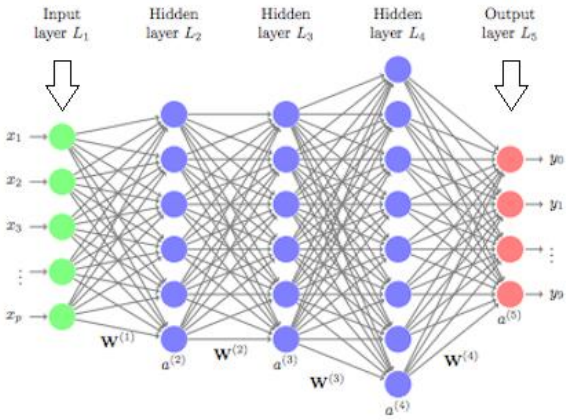
A neurális hálók a gépi tanulás egy speciális formája, amely az emberi agy működését mintázza. Ezek a hálók az adatfeldolgozási és mintafelismerési feladatokhoz kiválóan alkalmazhatók. A neurális hálók alapegysége a neuron, amelyet mesterségesen modelleznek, hogy a bemeneti jelekre adott kimeneteket képezzen. A neurális háló több rétegből állhat, ezek közül a legfontosabbak a bemeneti réteg, a rejtett rétegek és a kimeneti réteg. [9][10]

A neurális hálók matematikai modellje alapvetően az alábbiak szerint írható le:

Bemeneti réteg: Az adatok ebbe a rétegbe lépnek be.

Rejtett rétegek: Ezek a rétegek az adatok bonyolultabb feldolgozását végzik, itt történik az adatokból történő minták felismerése.

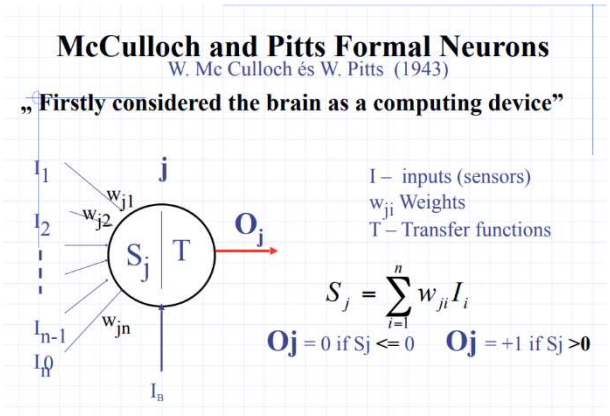
Kimeneti réteg: Az utolsó réteg adja meg a végső előrejelzést vagy osztályozást. [9][10]



1. ábra Neuron háló ábrázolása [11]

3.8 Neuronok

McCulloch és Pitts formális neuronjának modellje a mesterséges intelligencia alapvető építőköve, amely a biológiai neuronok működését imitálja. A modell bemeneti jelekből áll, amelyek mindegyike egy súllyal van összekötve a neuronhoz, és a bemeneti jelek súlyozott összegét egy transzferfüggvény (gyakran lépcsőfüggvény) segítségével alakítja át kimeneti jellé. A kimenet akkor aktiválódik, ha a súlyozott összeg meghalad egy előre meghatározott küszöbértéket, ami az alapvető logikai műveletek megvalósítását teszi lehetővé.



2. ábra Neuron felépítése [12]

4 ADATBÁZIS BEMUTATÁSA

Az EuRepoC 2000-es évekig visszamenőleg napjainkig nyújt adatok, amik nyílt forrásból elérhetőek.[5]

4.1 Az adatfájlok szerkezete

A letölthető Excel fájl három lapra tagolódik, a CSV letöltés pedig három külön fájlt tartalmaz. [5]

4.2 Fő adatkészlet

Ez a fájl vagy lap tartalmazza az egyes incidensekhez kódolt összes változót, úgy rendezve, hogy egy sor egy eseménynek feleljen meg – a mi fő vizsgálati egységünknek. Ha egyetlen eseményhez egyetlen változóhoz több kód is tartozik, ezeket pontosvesszővel választják el ugyanazon a cellán belül. [5]

4.3 Vevő adatkészlet

Ebben a fájlban vagy lapon az érintett entitások és személyek (fogadók) adatai átstrukturálva vannak az elemzés megkönnyítése érdekében. Minden cella csak egyetlen kódot tartalmaz, az adatok több sorban vannak „kicsomagolva”. Így egyetlen incidens több sort is átívelhet, amelyek az egyes eseményekhez rendelt egyedi azonosítók révén azonosíthatók. [5]

4.4 Hozzárendelési adatkészlet

Ez a lap vagy fájl a vevő adatkészletéhez hasonló megközelítést követ. A hozzárendelési adatok több sorban „kicsomagolva” vannak, így minden cella csak egy kódot tartalmazhat. Itt is egyetlen incidens több sort is elfoglalhat, és az egyedi azonosító lehetővé teszi az egyes események egyszerű nyomon követését. [5]

4.5 Gyűjtött adatok

Az EuRepoC az alábbi adatokat gyűjtötte:

1) ID
2) name
3) description
4) start_date
5) end_date
6) inclusion_criteria
7) inclusion_criteria_subcode
8) source_incident_detection_disclosure

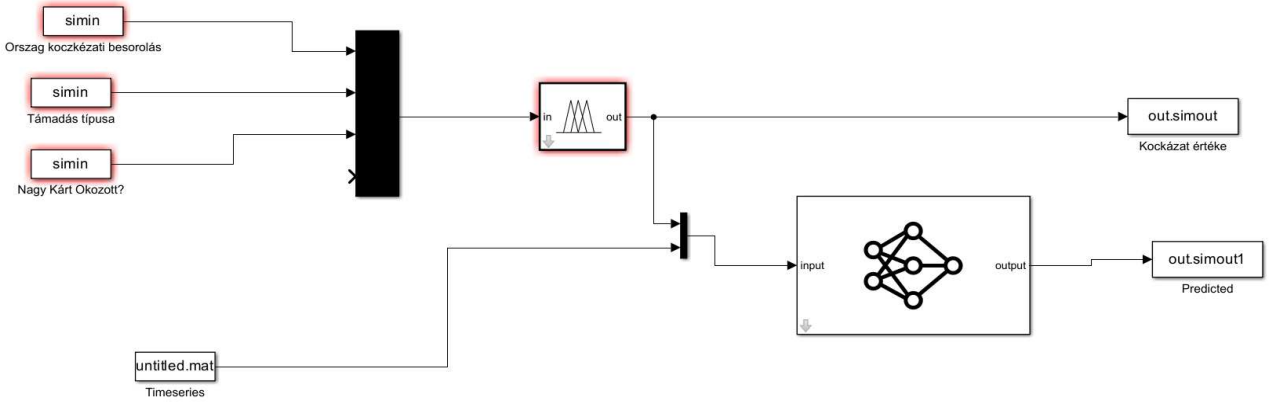
9) incident_type
10) receiver_name
11) receiver_country
12) receiver_region
13) receiver_category
14) receiver_category_subcode
15) initiator_name
16) initiator_country
17) initiator_category
18) initiator_category_subcode
19) number_of_attributions
20) attribution_ID
21) attribution_date
22) attribution_type
23) attribution_basis
24) attributing_actor
25) attribution_it_company attributing_country
26) attributed_initiator
27) attributed_initiator_country
28) attributed_initiator_category
29) sources_attribution cyber_conflict_issue
30) offline_conflict_issue
31) offline_conflict_issue_subcode
32) offline_conflict_intensity
33) offline_conflict_intensity_subcode
34) number_of_political_responses
35) political_response_date
36) political_response_type
37) political_response_type_subcode
38) political_response_country
39) political_response_actor zero_days
40) zero_days_subcode
41) MITRE_initial_access
42) MITRE_impact
43) user_interaction
44) has_disruption
45) data_theft
46) disruption
47) hijacking
48) physical_effects_spatial
49) physical_effects_temporal
50) unweighted_cyber_intensity
51) target_multiplier
52) weighted_cyber_intensity
53) impact_indicator
54) impact_indicator_value
55) functional_impact
56) intelligence_impact
57) political_impact_affected_entities
58) political_impact_affected_entities_exact_value
59) political_impact_third_countries
60) political_impact_third_countries_exact_value
61) economic_impact
62) economic_impact_exact_value
63) economic_impact_currency
64) state_responsibility_indicator
65) IL_breach_indicator
66) IL_breach_indicator_subcode
67) evidence_for_sanctions_indicator
68) number_of_legal_responses
69) legal_response_date
70) legal_response_type
71) legal_response_type_subcode
72) legal_response_country legal_response_actor
73) legal_attribution_reference
74) legal_attribution_reference_subcode
75) legal_response_indicator
76) casualties
77) sources_url
78) added_to_DB
79) updated_at

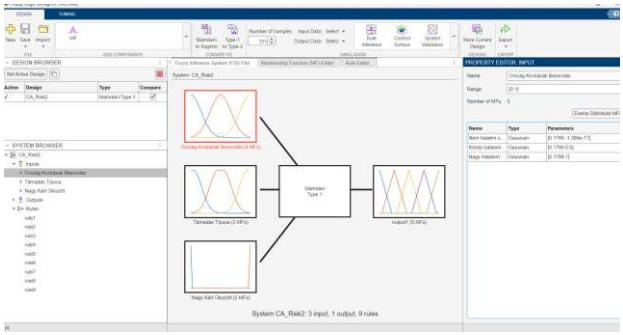
5 FUZZY LOGIKA ALAPÚ KOCKÁZAT BECSLŐ RENDSZER

Készítettünk egy Fuzzy logika alapú kockázat becslő rendszert a Matlab Fuzzy Logic Designer Toolbox használatával. A rendszerünk 3 értéket vesz figyelembe az adathalmazból. Első az ország kockázati besorolása. Ez azért fontos tényező, mivel egy nagyobb kibertámadás esetén a nagy és középhatalmaknak lehetősége van nem csak a kibertérben, hanem például pénzügyi, jogi vagy diplomáciai szankciókkal sújtani a vélt támadót (azért csak véltet, mivel a támadások vissza nyomozása sokszor technikai okok miatt lehetetlen). 3 kategóriát alkalmaztunk, az egyszerűség kedvéért. Nagyhatalmak, középhatalmak és kisebb országok. Nagyhatalmaknak az alábbiakat csoportosítottuk: USA, Kína, Oroszország, Egyesült Királyság, Franciaország, Németország. Középhatalmak esetében: India, Japán, Brazília, Kanada, Ausztrália, Olaszország, Dél-Korea, Törökország, Spanyolország, Mexikó. Minden más esetben kisebb ország csoportot alkalmaztuk. A csoportosítás során figyelembe vettük az országok gazdasági teljesítményét, katonai erőit, politikai befolyását és a kulturális befolyásosságuk. A tagsági függvényeket Gaussi típusúra állítottuk és egyenlően osztottuk el őket.

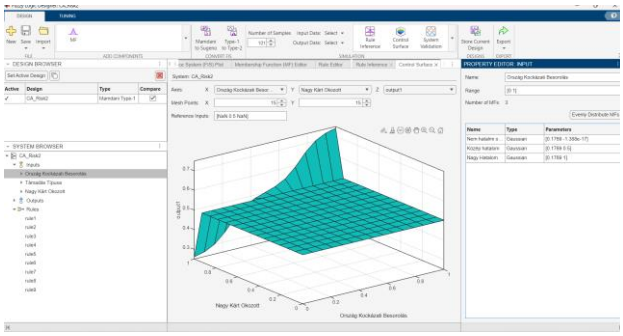
A második bemeneti érték a támadás típusa. Itt is három részre osztottuk fel a bemenetet: kis kiterjedésű támadás, közepes kiterjedésű támadás, nagy kiterjedésű támadás. Tagsági függvénynek két oldalú Gaussi függvényeket használtunk egyenletesen elosztva a tartományon.

Az utolsó bemeneti értékünk egy bool változó ahol annyi információt hordoz a bemenet, hogy maga a támadás kiterjedtségtől függetlenül nagy kárt okozott-e? Az EuRepoC szerint érhetik a kormányzati oldalakat, kritikus infrastruktúrákat, médiát, politikai pártokat és még másokat is a támadások. Ezek mind kihatással lehetnek az ország állampolgárainak életére.

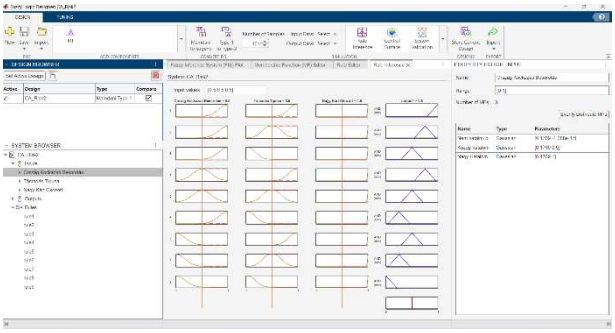




3. ábra Kép a Fuzzy Interfaceről | Szerző által szerkesztett



4. ábra Kép a Fuzzy Interfaceről Controll Surface| Szerző által szerkesztett



5. ábra Kép a Fuzzy Interfaceről Rule Inference| Szerző által szerkesztett

6 PREDIKCIÓS RENDSZER

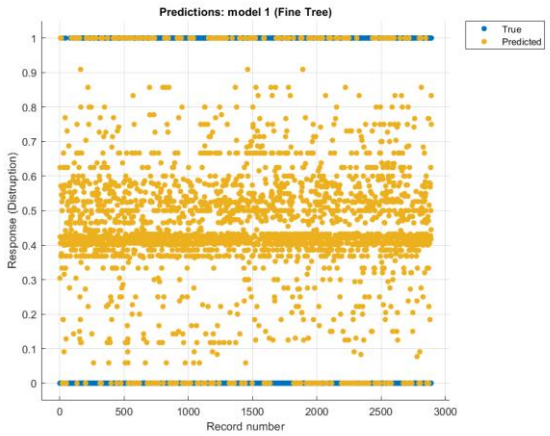
6. ábra Predikciós rendszer | Szerző által szerkesztett

6.1 Modell és Mérések bemutatása

Predikciók készítéséhez a Matlab Regression Learner Toolboxát használtuk. A mérések közül az alábbi 3-at szeretném ismertetni.

6.1.1 Fine Tree

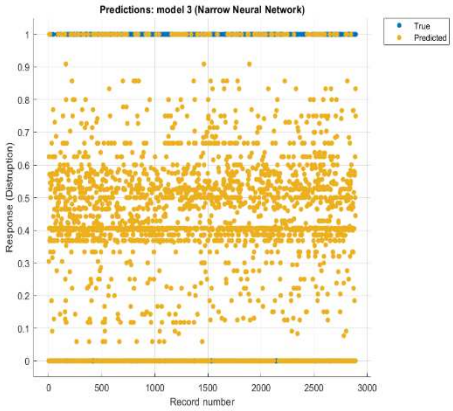
A Fine Tree algoritmust 4-es minimum levél mérettel futattuk le. Az RMSE validáció: 0,49035



7. ábra Fine Tree Response Plot| Szerző által szerkesztett

6.1.2 Narrow Neural Network

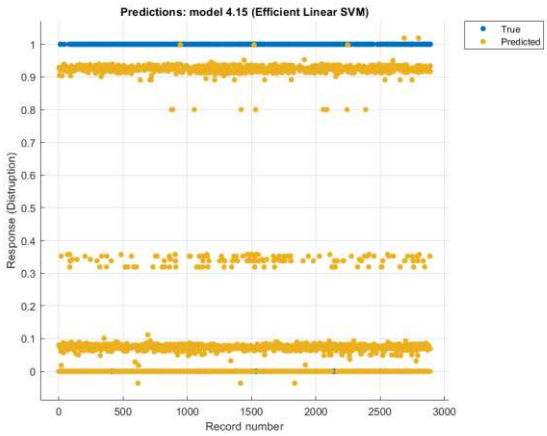
A Narrow Neural Networkot 1 teljesen csatlakoztatott réteggel. 10.-es bemeneti réteggel ReLU aktiváció függvénnyel és 1000-es iterációs limittel futtatok le. Az RMSE validáció értéke 0,49623.



8. ábra Narrow Neural Network Plot| Szerző által szerkesztett

6.1.3 Efficent Linear SVM

Az Efficient Linear SVM-et automata üzemmódban futattuk le. RTC értéke: 0,0001. RMSA validáció: 0,5787.



9. ábra Efficient Linear SVM Response Plot | Szerző által szerkesztett

IRODALOMJEGYZÉK

[1] https://eurepoc.eu/about-us/ Szerző által fordított

[1] Murphy, Kevin P.. "Machine learning - a probabilistic perspective." Adaptive computation and machine learning series (2012).

[2] J. M. Mendel, "Fuzzy logic systems for engineering: a tutorial," in Proceedings of the IEEE, vol. 83, no. 3, pp. 345-377, March 1995, doi: 10.1109/5.364485.

[3] Omidvar, M., Zarei, E., Ramavandi, B., Yazdi, M. (2022). Fuzzy Bow-Tie Analysis: Concepts, Review, and Application. In: Yazdi, M. (eds) Linguistic Methods Under Fuzzy Information in System Safety and Reliability Analysis. Studies in Fuzziness and Soft Computing, vol 414. Springer, Cham. https://doi.org/10.1007/978-3-030-93352-4_3

[4] https://eurepoc.eu/database/

[5] *European Repository of Cyber Incidents (EuRepoC) (2024) "Global Dataset of Cyber Incidents V.1.2". doi: 10.5281/zenodo.11108195.*

[6] Montgomery, D.C., Peck, E.A., & Vining, G.G. (2012). Introduction to Linear Regression Analysis. John Wiley & Sons.

[7] Weisberg, S. (2005). Applied Linear Regression. John Wiley & Sons.

[8] Kutner, M.H., Nachtsheim, C.J., & Neter, J. (2004). Applied Linear Regression Models. McGraw-Hill Irwin.

[9] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[10] Bishop, C.M. (2006). Pattern Recognition and Machine Learning. Springer.

[11] https://www.researchgate.net/figure/A-sample-neural-network-layout-with-three-hidden-layers_fig5_368540750

[12] Dr. Kutor László előadása OE NIK 2023 https://elearning.uni-obuda.hu/main/pluginfile.php/1132079/mod_resource/content/0/IS%202023-2-2.pdf

[13] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297.

BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Using Challenge-based Tasks for Testing the Security of Web Applications

Zlatko Čović
Subotica Tech – College of Applied Sciences, Subotica, Serbia
chole@vts.su.ac.rs

Abstract — **This paper explores the integration of challenge-based tasks in the education of cybersecurity and software engineers for the testing of the security of web applications. It provides an overview of security education and web application security, highlights the importance of challenge-based tasks, and details their implementation in the curricula. Specific tasks and their solutions are presented, along with future plans for improving these educational methods. The study emphasizes the critical role of hands-on, practical approaches in bridging the gap between theoretical knowledge and practical skills.**

***Keywords*:** challenge-based tasks, web security, web applications, owasp, code-entry challenges

1 INTRODUCTION

The rapid expansion and deep integration of web applications into everyday life have made securing these platforms more crucial than ever. Web applications are now omnipresent, enabling everything from online shopping and social interactions to banking services. This widespread use, however, also exposes them to numerous cyber threats, including data breaches and the exploitation of security vulnerabilities.

Traditional methods of security testing often fail to meet the demands posed by the complex and ever-evolving nature of modern web applications. These conventional approaches are typically static and predictable, lacking the ability to replicate real-world attack scenarios effectively. As a result, there is a pressing need for more innovative testing strategies that can more accurately assess and improve web application security.

In the education of software and/or cyber engineers, a gap often arises between theoretical and practical knowledge. It is essential to define and implement innovative educational methods that offer hands-on experience with real-world problems. The field of information security is one of the most critical aspects of modern life and communication. The vast amount of information we access and share through various services and tools forces the education of professionals who can create and maintain these systems securely.

Authors in the study [1] focus on the applicability of the hackathon method in training software engineers. The hackathon method, which is student-centered, uses a pedagogical approach based on constructivist theory. During a hackathon, programmers and other software development professionals collaborate intensively on projects with students. This method allows students to gain new knowledge and skills and develop key competences. The hackathon method is integrated into the Web programming course curriculum at Subotica Tech – College of Applied Sciences, aiming to assist students in building web applications. Upon completing the course, students will be able to work independently with multiple programming languages and technologies and manage databases in a client-server environment using appropriate security methods and techniques.

The application of the Challenge-Based Learning (CBL) methodology to cybersecurity education has been described in paper [2]. Challenges were formulated based on students' interest in securing information and systems, and solutions to meet these challenges were devised collaboratively by students working as a team. The knowledge learned was practiced by students in two cybersecurity competitions. Formative assessments showed that great benefits were derived from the CBL approach, though the extent of benefit varied among students. Computer skills, security knowledge, ability to teach others, and interest in cybersecurity were improved by the students. Although additional support resources and irregular meeting hours may be required, the increase in student learning justifies the extra effort [2].

Challenge-Based Learning (CBL) involves an external stakeholder (e.g., company, community, NGO) presenting a real-life complex problem. While CBL is gaining popularity in engineering education, its application in engineering ethics education lags due to specific design requirements. Although ethics CBL courses share similarities with general CBL courses, they necessitate unique design considerations [3]

A structured approach is needed to apply CBL in engineering ethics education, from both educational research and support perspectives. In [3] authors address this gap by exploring the particularities of CBL pedagogies in engineering ethics education using van den Akker's spider-web curriculum model and the Ethics Goal Model. The specificities for various course components in CBL engineering ethics courses are described, concluding that the balance between structured and open CBL approaches is crucial and warrants further research [3].

The design of defensive challenges for Capture the flag (CTF) in the industry is addressed in this [4] work. A challenge structure and six different challenge types were derived based on semi-structured interviews with security experts, a two-part survey, and informal discussions. Results indicate that security experts prefer traditional challenge types, such as Single-Choice and Multiple-Choice Questions, while Text-Entry Challenges are the least preferred. Additionally, Association Left-Right, Code-Entry Challenge, and Code-Snippet Challenge types were discussed. The Code-Entry Challenge, where players

submit code to the backend for validation, emerged as an unexpected yet promising type for secure coding challenges [4].

Further investigation is needed to detail the creation of Code-Entry Challenges. The results are based on feedback from security experts, and further work is required to validate the challenge structure and types in real CTF events within an industrial setting. Future publications will provide concrete examples of implementing the derived challenge types and will refine the structure based on feedback from CTF players [4].

For a web application or information system to be secure, the creators of these systems must test them using appropriate tests. It is also desirable to raise awareness about potential vulnerabilities and threats. The OWASP foundation created the OWASP Top 10 as a standard awareness document for developers and web application security. It reflects a wide consensus on the most critical security risks facing web applications [5].

This paper describes the usage of challenge-based tasks in the education of cyber security and software engineers. These tasks were implemented in two curricula: Security in e-business systems and Web programming. The challenges are defined to follow the OWASP Top 10 standard awareness document for developers and web application security.

This paper is organized as follows. Section 2 shortly introduces the education of security aspects and gives basic information about web applications security. In Section 3, the significance of challenge-based tasks is elucidated through several key arguments. The implementation of challenge-based tasks in education is described in Section 4. Some of the tasks and their possible solutions are also presented in detail. Future plans are outlined in the following section. The conclusion at the end of the paper summarizes the key points from the preceding sections.

2 Education of Security Aspects

In the education of cybersecurity and software engineers, it is very important to introduce new teaching methods to bridge the gap between practical and theoretical knowledge. Although practical topics are covered in lab exercises following theoretical knowledge, the traditional approach to assignments is not always sufficient.

At Subotica Tech – College of Applied Sciences, in most specialized subjects, students are required to complete homework assignments and a practical project, which often involves team collaboration. During the project, aspects of project-based learning (PBL) are applied. Digital tools are used for development as well as for collaboration. A newer method that has been tested after PBL is hackathon-based learning (HBL). This learning model also emphasizes teamwork but incorporates elements found in hackathons held outside of the university. Depending on the type of information system, whether it is a desktop program, mobile application, web application, or integrated web system students are required to test their software both functionally and for security within their projects. A strong emphasis is placed on simulating real-life situations.

During the hackathon, collaboration with local IT companies is encouraged, and these companies delegate their own programmers who join the teams as external mentors. Their task is to define additional functionalities for the project based on the project description provided by the course professors.

A strong focus is placed on various aspects of security within professional courses of study. In some of these courses, protection methods and techniques are taught through laboratory exercises, using diverse approaches to test information systems, web applications, mobile applications, and other types of software and networks.

It is crucial to provide students with opportunities to test their knowledge by assigning tasks in laboratory exercises and homework that are related to information security.

2.1 Web applications security

As part of information security, web application security specifically addresses the security of web applications, websites, web systems, and web services. Web application security is based on the principles of application security, combining and applying them separately to web systems and the internet.

Core issues in the development and use of web applications include insufficient knowledge and weak awareness of potential threats. Developers often fail to implement appropriate security techniques and methods, and do not conduct security checks of their applications.

Secure code has two aspects: develop code without bugs and security holes and develop code resistant to abuses and attacks. Current software engineering technology doesn't guarantee meeting both criteria. The best approach is dynamic or static testing to ensure code behavior matches user requirements. Functional testing checks if the code behaves as expected in various scenarios but can't guarantee the absence of errors. It can only show that errors are present. Additionally, traditional functional testing cannot confirm code security or immunity to attacks [6].

Developers and engineering students need to be able to identify security threats and vulnerabilities. The OWASP Top 10 standard document is the best resource for understanding these topics and for defining guidelines for security testing of software. It is also useful for creating challenge-based tasks in the education of software engineers and cybersecurity engineers.

3 Challenge-based tasks

Challenge-based tasks for assessing the security of web applications hold significant importance due to several key reasons:

- Realistic Cyberattack Simulations
- Thorough Vulnerability Testing
- Skill Verification
- Ongoing Skill Improvement
- Reducing Risks
- Meeting Compliance
- Affordable Security Testing

3.1 Realistic Cyberattack Simulations

These tasks mimic real cyberattacks, helping to evaluate how well an application's security can handle actual threats. By simulating authentic attack scenarios, these challenges provide a practical assessment of an application's defenses, enabling developers to identify and address potential weaknesses before they can be exploited by malicious actors.

3.2 Thorough Vulnerability Testing

These challenges cover various security issues like injection attacks (e.g., SQL injection, XSS), weak authentication, privilege escalation, and more, allowing testers to find weaknesses in different parts of the application.

3.3 Skill Verification

Overcoming these challenges requires a strong understanding of web security concepts and techniques, demonstrating testers' ability to identify and fix security risks effectively.

3.4 Ongoing Skill Improvement

These challenges can be set at different difficulty levels, helping testers advance from basic to advanced security testing skills and promoting continuous learning and skill growth within security teams.

3.5 Reducing Risks

By finding and fixing vulnerabilities through challenges, organizations can lower the chances of security breaches, protecting against data theft, system hacks, and damage to their reputation.

3.6 Meeting Compliance

Many industry standards and regulations require regular security testing. Challenge-based tasks help meet these requirements in an organized way.

3.7 Affordable Security Testing

Challenge-based tasks provide a cost-effective method to comprehensively assess an application's security posture. By simulating various threat scenarios in a controlled environment, organizations can efficiently identify and prioritize security weaknesses, thereby optimizing resource allocation for targeted remediation efforts. This approach not only helps mitigate potential risks but also promotes a proactive security culture, enabling continuous improvement of defenses without exceeding budgetary constraints.

4 IMPLEMENTATION OF CHALLENGE-BASED TASKS IN EDUCATION

Challenge-based tasks were first implemented in the education of cyber security and software engineers. These tasks were applied in two curricula: Security in e-business systems and Web programming.

The initial tasks were carried out with cybersecurity students from Obuda University who were participating in an Erasmus exchange program at Subotica Tech - College of Applied Sciences during one semester. These tasks were part of the mandatory curriculum within the subject of Security in e-business systems.

After successfully implementing these tasks and based on feedback from the students, the tasks were slightly modified and integrated into the Web Programming course with informatics students at Subotica Tech – College of Applied Sciences as the last class in the semester.

In the domain of information security and within various courses, students at Subotica Tech learn theoretically and practically about following terms:

- Filtering and validating user input data
- Secure file uploads and user management (registration, login)
- Secure use of sessions, cookies, and API endpoints
- Web server and directory protection
- Cryptography, crypto algorithms, and token authentication
- Security of IoT systems and networks
- Database protection and SQL injection prevention
- Mobile application security
- Restricting unsafe dynamic function calls and access to sensitive components
- Analyzing meta data in media files
- Avoiding legacy approaches, classes, and methods

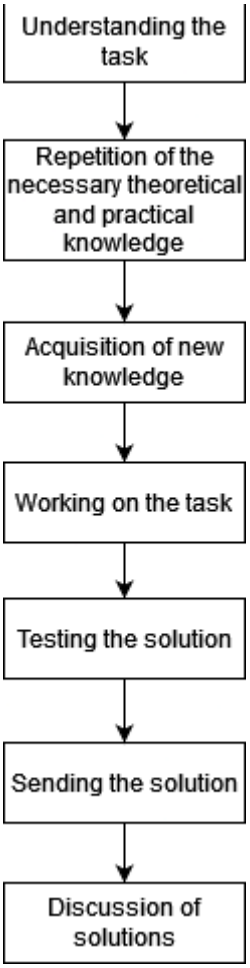
4.1 Task structure

After acquiring theoretical and practical knowledge in areas related to web programming and web security, students were given challenge-based tasks that they had to solve individually during laboratory exercises.

The tasks were designed to be based on challenges and to be related to the security of web applications. Each task was scheduled for a specific time. After that time, the students sent their material to the server that contained solution in the form of program code and textual file with information about their experiences during the work on task and, if necessary, detailed explanation of the solution. Also, in some tasks it is mandatory to write down what could be the potential security issues and vulnerabilities and how they can be avoided and what are the most suitable techniques or methods for that. A discussion followed where different approaches to solving tasks and encountered problems were discussed.

These are practical tasks that always required submitting created code or code snippets, so it can be said that these tasks have characteristics of code-entry challenges.

In the schema below, the general structure of flow of every task is presented.



1.figure: Structure of flow of every task

During each part of the task, students can ask the subject professor questions if something is unclear or if they need additional explanations. In the description of each task, web links to resources containing material for reviewing acquired knowledge and acquiring new knowledge have been added.

As can be seen in the schema, there is a section for testing the solution. This part is crucial because, in some tasks, additional information necessary for problem-solving cannot be obtained without testing.

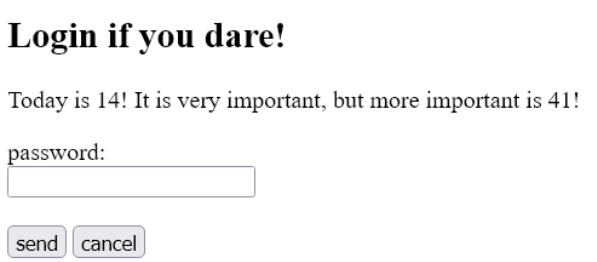
4.2 Examples of challenge-based tasks

In this section, some challenge-based tasks and their possible solutions will be presented. Some of them are easier and require less time to solve, but some require more testing attempts and more knowledge to be properly solved.

4.2.1 Finding the correct password

Task text: *Analyze html code of form on given url. Try to find the correct password. For getting password, you should write php code that will get important data. Write an explanation of your solution.*

Students were required to analyze the HTML code of a web page that contains text messages and a login form with a password input field and two buttons: one for submitting data and one for resetting the form.



2. figure: Web form of task 1

Below is the HTML code of the first task's webpage.

```
<!DOCTYPE html>
<html lang="en">
<body>
<h2>Login if you dare!</h2>
<p>Today is 14! It is very important, but more
important is 41!</p>

<form action="check.php" method="post">
<label for="password">password:</label><br>
<input type="password" id="password"
name="password"><br>
<input type="hidden" name="l1"
value="41rkPe3lA2fhgutijgkdkfjiti" >
<input type="hidden" name="l2"
value="ab0526ad3fc8bd9909482126d72deb432a0efbe54
a29ba0da3b64b" ><br>
<input type="submit" value="send">
<input type="reset" value="cancel">
</form>
</body>
</html>
```

When examining the source code of the HTML page, two hidden fields named l1 and l2 can be seen within the form. Each of these fields has an initial value assigned, consisting of a string of random numbers and letters. The length of the first field is 26 characters, and the second is 54 characters.

If the user refreshes the page multiple times or sends a new request to the server, they will notice that the value of the hidden field l2 changes, but only the first 14 characters. The remaining 40 characters stay the same.

Since the content is a random string of characters, a student might think it is a type of hash value. Knowing some hash algorithms, they might conclude that the SHA-1 algorithm generates a hash of 40 characters in length. By copying the 40 characters that do not change, they can use free online tools to try to find the plain text of the hash value. These tools contain many hash values and their corresponding plain text values. Some of these values are automatically generated, while others are manually added by volunteers. This type of service does not decrypt data in real-time but searches its own database based on the input provided. From an ethical standpoint, the purpose of these services is not to steal someone's password but to highlight the weakness of a password. Of course, some users of these services may have malicious intentions.

Entering the found hash value from the hidden field into an online tool will not result in finding the plain text version of the text. The student should continue analyzing the information on the web page. Above the web form, there is a message: *Today is 14! It is very important, but more important is 41!*

From this message, the student might deduce that the important information is that 41 is the reverse of 14. By using a programming language, such as PHP, to apply a built-in function or method to reverse a string, they can attempt to find the plain text value for the reversed string using an online service. In PHP, the function used for this purpose is *strrev*. For the reversed value, the student will easily find the plain text version of the password.

4.2.2 Uploading larger jpg and png files

Task text: *Analyze html code and try to upload jpg and png files that are bigger than 1MB. Don't forget to insert your index number from Subotica Tech. Write the explanation of your solution.*

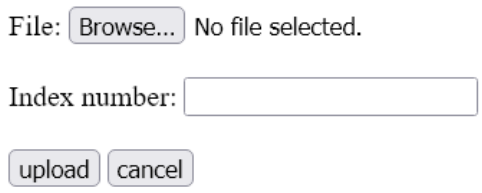
The HTML code of a web page contains file selection fields, a text field for entering an index number, and two buttons, one for submitting data and one for resetting the form.

Below is the HTML code of the second task's webpage.

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <title>File upload</title>
</head>
<body>
<form method="post" name="upload"
action="upload.php" enctype="multipart/form-
data">

    <input type="hidden" name="MAX_FILE_SIZE"
value="102400" >
    <label for="if">File:</label> <input
type="file" name="file" id="if"
accept="image/jpeg"><br><br>
    <label for="index">Index number:</label>
<input type="text" name="index"
id="index"><br><br>
    <input type="submit" name="sb" id="sb"
value="upload">
    <input type="reset" name="rb" id="rb"
value="cancel">
</form>

</body>
</html>
```



2. figure: Web form of task 2

By analyzing the HTML code, the following information can be obtained:

- the name of the page receiving submitted data through the form (action attribute),
- the value of the hidden field MAX_FILE_SIZE,
- and the value of the accept attribute.

In web forms, MAX_FILE_SIZE is a hidden input field that defines the maximum file size (in bytes) that the server will allow users to upload through that specific form. This ensures server-side restrictions on file size limits for uploads. The value of this field can be easily changed by a potential attacker.

The accept attribute in a file input type specifies the types of files that the file input element can accept for file selection. It assists browsers in filtering the files displayed in the file picker dialog to show only those files that match the specified types.

Based on the obtained values, the file size is limited to 100KB and selecting files of type jpg is preferred. The student can try changing the file type in the file picker to all types and select a png file instead. There are several solutions for uploading a png file larger than 1MB.

The first solution involves using the "Inspect element" option within the web browser to delete the hidden field MAX_FILE_SIZE. This action does not delete the original file on the server but changes the HTML code that is fetched and currently active in the web browser. Without this hidden field, there would be no size limit for the file. The student will only need to enter their index number and submit a png file larger than 1MB.

The second solution requires copying the entire HTML code into another HTML file and deleting the hidden field and the accept attribute. The newly created HTML file should be run locally in a web browser to test whether uploading a png file larger than 1MB will succeed. If the upload is successful, the student will receive confirmation of the file being sent.

It's notable here that there are several security vulnerabilities on the server side in program code. There is no verification of the size and type of the uploaded file or the origin domain from which the data is sent.

4.2.3 Try to find your own password

Text of task: *Try to get your password from hash. You can't remember the correct password, but you know that it contains 4 digits. Also, you find another hash or random string. Maybe it is connected to something with hash. Write php code and try to find the correct password.*

When looking at the length of the found hash, the result is 32 characters. A student familiar with hash algorithms can immediately conclude that the MD5 algorithm was used for hashing. Also, based on the string with random values, they might think that the SALT technique was used to obtain the hash value. SALT technique involves adding a unique, random value to a password before hashing it to ensure that identical passwords have different hash values and to protect against precomputed hash attacks. Since SALT can be added to the beginning, end,

or both ends of the password, the student must take all of this into account.

The password consists of 4 digits, but it is not yet known where the SALT is located within the password. The student can write program code that will hash values from 0000 to 9999 and check the obtained hash against the given hash. If they go through all combinations and do not find the password, the next step is to add the SALT value to the beginning of the potential password and perform the hashing with this data. If the result is still negative, the procedure is repeated by adding the SALT value to both the beginning and end, or only to the end of the potential password.

Some of the students tried to enter the hash into the search option available on some online services to obtain the plaintext based on that value. None of them succeeded. The reason is that these online services did not use SALT, which was generated randomly in this task, and therefore, even such a simple data of 4 digits cannot be found

5 Future plans

After the initial test implementation of tasks in the educational process and based on the feedback from students who worked on these tasks, future steps can be defined for the next implementation of these challenge-based tasks.

Solving certain tasks required more time and effort from some students. There were instances where some tasks were not successfully completed. One of the suggestions was to divide the tasks by difficulty level.

For this purpose, a control test (both theoretical and practical) would be conducted to determine the students' knowledge levels. Based on this level, tasks of the corresponding difficulty or tasks of the next level of difficulty could be assigned to the students.

Another suggestion was to categorize the tasks and have them completed after covering a larger teaching unit within the curriculum. This way, these types of tasks would be done multiple times throughout the semester.

Some students suggested that the tasks could be done in teams, because that think it could help them to solve problems more effectively.

Since many tasks require the input of certain data that were verified on the backend, the plan is to log all these attempts in a database. Based on this data, the students' thought processes and approaches to problem-solving can be determined.

A potential improvement in defining tasks would involve surveying industry experts to identify the areas where they see the most security vulnerabilities among developers in their companies, as well as the challenges that even they sometimes find problematic.

6 Conclusions

In educating software and cyber engineers, a gap often exists between theoretical and practical knowledge. Therefore, it is crucial to develop and apply innovative educational methods that provide hands-on experience with real-world challenges.

This paper has explored the integration of challenge-based tasks in the education of cybersecurity and software engineers. These tasks were implemented in two curricula: Security in E-Business Systems and Web Programming. The challenges were defined to follow the OWASP Top 10 standard for developers and web application security.

Section 2 provided a brief overview of the education of security aspects and fundamental information about web application security. The importance of challenge-based tasks was highlighted in Section 3 through key arguments. The implementation of these tasks in the curricula was detailed in Section 4, with specific tasks and their possible solutions presented. Based on feedback from students, future plans for the continued use and improvement of challenge-based tasks were discussed in the subsequent section. The key points from each section were summarized to emphasize the critical role of innovative, hands-on educational methods in bridging the gap between theoretical knowledge and practical skills in the field of cybersecurity and software engineering.

References

[1] Z. Čović, Z. Papp, H. Manojlović and J. Simon, "Hackathon-based Teaching Method in the Training of Software Engineers", Proceedings of the 12th International Conference on Applied Internet and Information Technologies AIIT 2022, Zrenjanin, Serbia, 2022, pp. 108-116

[2] Cheung, R. S., Cohen, J. P., Lo, H. Z., & Elia, F. (2011). *Challenge based learning in cybersecurity education*. Athens: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

[3] G. Bombaerts, D. Martin and K. Doulougeri, "Structured and open Challenge-Based Learning in Engineering Ethics Education," 2022 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 2022, pp. 1-8, doi: 10.1109/FIE56618.2022.9962652.

[4] Gasiba, T., Lechner, U., Pinto-Albuquerque, M., Zouitni, A. (2020). Design of Secure Coding Challenges for Cybersecurity Education in the Industry. In: Shepperd, M., Brito e Abreu, F., Rodrigues da Silva, A., Pérez-Castillo, R. (eds) Quality of Information and Communications Technology. QUATIC 2020. Communications in Computer and Information Science, vol 1266. Springer, Cham. https://doi.org/10.1007/978-3-030-58793-2_18

[5] Čović, Z. (2024). Threats and Vulnerabilities in Web Applications and How to Avoid Them. In: Kovács, T.A., Nyikes, Z., Berek, T., Daruka, N., Tóth, L. (eds) Critical Infrastructure Protection in the Light of the Armed Conflicts. HCC 2022. Advanced Sciences and Technologies for Security Applications. Springer, Cham. https://doi.org/10.1007/978-3-031-47990-8_9

[6] A. J. A. Wang, "Security testing in software engineering courses," 34th Annual Frontiers in Education, 2004. FIE 2004., Savannah, GA, USA, 2004, pp. F1C-13, doi: 10.1109/FIE.2004.1408561.

BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Customization of the Physiological Parameter Assessment Using Fuzzy Logic

Felisberto David Wandi Chivela*, Zoltán Papp **, Edit Laufer ***
*Óbuda University, Doctoral School of Applied Informatics and Applied Mathematics, Budapest, Hungary
** University of Novi Sad, Hungarian Language Teacher Training Faculty, Subotica, Serbia
*** Óbuda University, Bánki Donát Faculty of Mechanical and Safety Engineering, Budapest, Hungary
pfyreu@stud.uni-obuda.hu, zoltan.papp@magister.uns.ac.rs, laufer.edit@bgk.uni-obuda.hu

***Abstract* — Effective health monitoring is very important for individuals engaged in sports and physical activities due to the diverse physiological responses exhibited by each participant. Traditional methods often fail to deal with the complexity of individual health profiles, highlighting the necessity for personalized assessment methods. In this paper, a hierarchical fuzzy model is presented, which is intended to assess the risk level of the current physical activity. In order to personalize the evaluation statistics-based approach was used to tune the membership functions. The model presented provides both numerical and linguistic assessments of risk, demonstrating consistent trends between improved membership functions and medical recommendations. Extensions for future work are also included.**

***Keywords*:** fuzzy logic, risk assessment, sports activity, patient monitoring, membership functions, statistical evaluation.

1. INTRODUCTION

In today's world, the positive outcomes of regular exercise in preventing illness, aiding in recovery, and promoting an active lifestyle in general are universally recognized. Engaging in sports promotes an active lifestyle and enhances the overall quality of life [1].

Regular physical exercise offers numerous benefits for overall health and well-being. It improves cardiovascular health by strengthening the heart and reducing the risk of heart disease and high blood pressure, while also aiding in weight management through calorie burning and muscle building. Exercise boosts mood by releasing endorphins, reduces stress, anxiety, and depression, and enhances energy levels by improving circulation and nutrient delivery. Additionally, it strengthens muscles and bones, promotes better sleep, and enhances brain health and cognitive function. Exercise also boosts the immune system, reduces the risk of chronic diseases, and increases longevity, ultimately leading to a higher quality of life through improved mobility, reduced pain, and increased independence [2].

However, these benefits depend on individual capabilities and medical advice. Failure to consider personal fitness levels and medical guidance can result in potential dangers, such as overexertion or injury. Engaging in activities beyond our current capabilities and exercising with incorrect duration, frequency, and intensity levels can be counterproductive and fail to yield beneficial results. Factors like chronic illnesses, age, and other relevant sub-factors must be carefully evaluated to ensure that participating in sports remains a safe and beneficial activity, rather than exacerbating existing health issues.

It is noticeable that in contemporary times, patient monitoring devices have become indispensable in our daily routines [3]. The widespread adoption of Internet of Things (IoT) technology has led to the development of increasingly sophisticated systems with broader functionality. Consequently, the continuous monitoring and recording of physiological data have become accessible to a wider audience. As a result, research focus has intensified on evaluating physiological parameters, aiming to enhance safety in everyday life by enabling prompt recognition of any health deterioration. This research field's significance has been particularly highlighted by the COVID-19 pandemic, emphasizing the critical need for remote diagnosis. Utilizing such applications has played a pivotal role in curbing the spread of the virus by minimizing visits to medical facilities for less severe illnesses [4]. IoT devices have the capacity to facilitate remote health monitoring and emergency notification systems. From basic blood pressure and heart rate monitors to sophisticated gadgets capable of overseeing specialized implants like pacemakers, Fitbit electronic wristbands, or advanced hearing aids, the spectrum of health monitoring devices is vast [5]. Specialized sensors establish a network of intelligent devices capable of gathering, processing, transmitting, and analysing crucial data across various environments. This includes linking in-home monitoring devices with hospital-based systems, enhancing connectivity and data utilization [6]. Nonetheless, health monitoring systems prove beneficial not solely for the elderly with chronic conditions but also for individuals coping with cardiac conditions [7][8]. Moreover, such systems can prove advantageous for healthy patients as well, aiding in monitoring their physical activity and assessing the risk or their performance level [9]. The key features of health monitoring systems include utilization of wireless communication, portability, non-invasiveness, ease of use, compactness, and minimization of device count [10]. Overall, the integration of IoT in healthcare plays a crucial role in managing chronic illnesses and in disease prevention and control. Remote monitoring becomes feasible through robust wireless solutions. This connectivity empowers healthcare professionals to capture patient data and employ sophisticated algorithms in health data analysis [11].

A significant challenge of creating patient monitoring systems lies in medical applications, where numerous factors defy simple quantification and the boundaries delineating normal, increased, and abnormal values remain ambiguous and vague. Medicine frequently applies linguistic descriptions. Soft computing techniques prove highly valuable in addressing the challenges encountered in medical applications. The fuzzy approach demonstrates significant utility and efficiency in these domains and their counterparts, such as risk management. Its compatibility with human language and ability to manage uncertainty, imprecision, and subjectivity in both data and evaluation processes make it particularly valuable. Fuzzy-based methods typically yield more realistic results presented in a user-friendly format [12] [13].

Another challenge in constructing patient-specific models stems from the system's behaviour is that it is influenced by numerous factors, some of which may be unidentified, with complex and often unknown interactions among them [14]. While the membership functions of the fuzzy model can be tailored in a patient-specific manner by accounting for the maximum number of relevant factors, it remains challenging to fully consider the combined effect of all significant factors unique to individual patients.

Thus, various approaches aim to minimize the number of inputs while considering a wide range of influential factors. Patient specificity can be ensured through different methods which include utilizing personal medical recommendations [15] as well as establishing thresholds via equations or tables derived from statistical data on personal characteristics like age, sex, and fitness level, aggregating results from patients with similar traits [16][17]; finally, membership functions (MFs) can be tuned based on the input-output pairs with fuzzy-neural system [18] [19].

The main goal of this study is to develop a risk assessment model in which the evaluation is customized according to the patient's characteristics. For this reason, statistics-based approaches are studied and built in the evaluation process. To handle this issue, authors focus on a specific subsystem, namely, the "Current physical status", because it is where user-specific tuned membership functions are most crucially required.

This paper is organized as follows, in section 2, the overall hierarchical model structure is presented. Section 3 shows the investigated subsystem structure – 'current physical status'. Section 4 presents the proposed statistics-based evaluation followed by a case study in section 5. Section 6 discusses the results of the use of the personal statistics. Section 7 draws the conclusion.

2. The Overall Model Structure

The overall model has a hierarchical multilevel clustered structure, which facilitates both model expansion and simplification of the evaluation process. The evaluation uses a Mamdani-type fuzzy inference system. The model's structure aligns closely with the logic of the evaluation process. The classification of input parameters relies on the logical connections between them. Three primary groups have been delineated, indicating whether they pertain to the patient's medical condition, characterize their sport activity behaviour, or describe the environmental conditions.

Within the primary groups, further classification is possible based on the permanence of the parameters. These include permanent parameters (such as sex), quasi-permanent or infrequently changing factors (such as chronic diseases and occupation), and real-time variables (for example, blood pressure and heart rate). The structure is derived from the model outlined in [20], depicted in Figure 1.

On the left-hand side of the diagram are the identified risk factors influencing the calculated temporal risk level. The middle section of the figure delineates which parameters belong to each risk factor group through three blocks. The highest level of the hierarchy is situated on the right-hand side, responsible for computing the actual risk level based on input from the problem groups. The main groups constitute the subsystems of the model, with their contributions to the overall risk level computed separately during processing. The subsystems, along with their varying parameters, undergo real-time evaluation, while the remainder of the model is evaluated offline before real-time assessment begins.

The medical condition of the patient is characterized by the Medical Condition group, which represents the most crucial and intricate subsystem. Personal conditions primarily determine the patient's load capacity, and most interactions among input factors occur within this group. The first input factor in this group is Disease Condition, encompassing chronic diseases such as hypertension, diabetes, and cardiac diseases, among others. While these diseases are quasi-permanent factors, their severity may vary over time. The second input is the Current Physical Status subsystem, which offers information about the patient's current condition. This assessment is based on measured parameters such as heart rate, systolic and diastolic blood pressure. Additionally, associated metrics and factors influencing these parameters are utilized as input to construct patient-specific membership functions.

The input factor, Basic Physical Information, serves to characterize the fundamental attributes and living circumstances of the patient. Mental stress holds particular significance within this subsystem. Despite exerting a weaker influence on physiological parameters compared to physical activity, it can notably elevate heart rate. To delineate the sport activity habits of the patient, the Activity Load subsystem is employed. Its subfactors delineate the intensity, duration, and frequency of the patient's activity, specifying how vigorously (Intensity), how long per occasion (Duration), and how often per week (Frequency) the activity is performed. Finally, the third main subsystem Environmental condition uses combined subfactors to characterize the environment. The temperature is combined with the actual humidity (TH) and wind (TW) because of their influence on thermal sensation. This group has more importance in the case of outdoor sports, but humidity and temperature together can influence the risk level indoors too [21].

The parameters of the group "Current physical status" subsystem, which is the main topic of this paper, changes in real-time. Monitoring these values ensures continuous control. Therefore, there is a need to customize it for each patient separately, i.e., a flexible risk assessment framework is required.

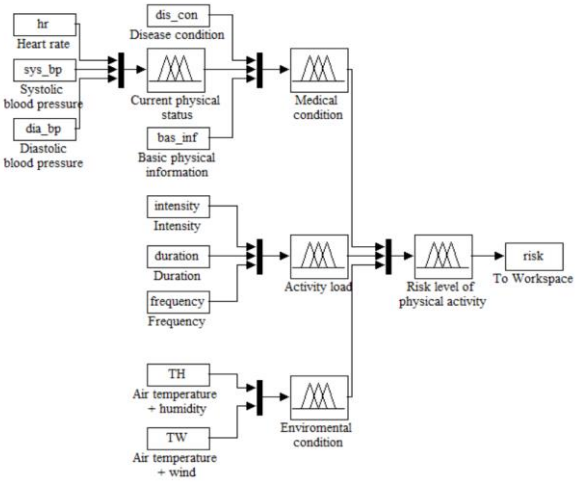


Figure 1. The fuzzy model structure

3. The Current Physical Status Subsystem

The initial parameter that was considered is the heart rate, which represents the number of myocardial contractions per minute. It is influenced by various external and internal factors, with numerous interactions among them. Patient-specific membership functions related to this parameter can be defined based on the maximum heart rate, representing the highest attainable value under physical exertion. Approximately 30 factors can affect this threshold, including age, sex, weight, time of day, environmental conditions, physical fitness, activity intensity, mood, medications, and certain medical conditions [22]. Instead of relying solely on the maximum heart rate, it is customary to utilize a predictive maximum heart rate estimated by the OwnZone function of the Polar heart rate monitoring device. This prediction is derived from personal parameters such as age, sex, and the patient's resting heart rate [23]. The OwnZone function leverages heart rate variability, which can be assessed using the Polar device or an oscillometric blood pressure monitor [24]. Obtaining the predicted maximum heart rate requires only a brief 5-minute test before monitoring, unlike traditional methods such as progressive exercise testing or VO2max measurement, which are typically conducted in laboratory settings and may not be feasible for all individuals, especially those with cardiovascular conditions or elderly individuals on medications affecting heart function and circulation. Additionally, this method offers the advantage of daily or multiple daily assessments, allowing for consideration of fluctuations in the maximum heart rate during evaluation.

Understanding the training target is crucial, as it determines the optimal heart rate and associated risk level. The target heart rate zone, expressed as a percentage of the maximum heart rate, is detailed in figure 2. This represents the original Polar zone where Vlow (<50%), light (50-60%), medium (60-70%), mhigh (70-85%), high (85-95%), vhigh (>95%).

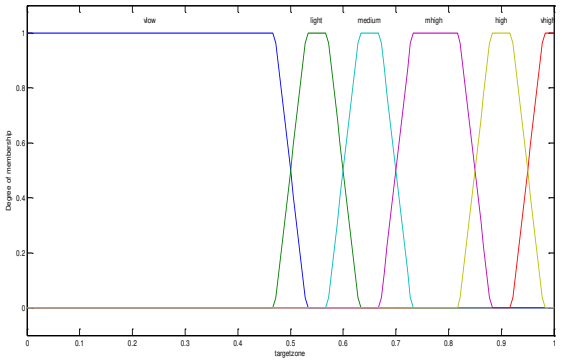


Figure 2. Target zones

This percentage indicates the permissible intensity level for the individual. It varies based on the activity's objective (rehabilitation or prevention), the individual's athletic background (regular athlete or beginner). Consequently, personalized zone limits can be computed, and membership functions can be adjusted to accommodate the specific characteristics of the individual patient.

Nevertheless, limits can be specified depending on the patient's condition. Figure 3 shows a case for a patient under medical treatment, debilitated, cardiovascular disease, respiratory disease or rehabilitation; here, from the original target zones, vlow would be the appropriate one based on the table, that's why it became the target (<50%), mhigh is the original light zone (50-60%), vhigh is the merge of the rest (medium, mhigh, high, vhigh) (>60%)).

Figure 4 shows a case for a beginner level sports individual; here, from the original target zones, light or medium would be the appropriate one based on the table, that's why target zone is created merging the vlow, light and medium zones (<70%), mhigh is the original mhigh zone (70-85%), vhigh is the merge of the rest (high, vhigh) (>85%). Figure 5 shows a case in which a person exercises regularly; here, from the original target zones, mhigh would be the appropriate one based on the table, that's why the target zone defined as a merge of the vlow, light, medium and mhigh zones (<85%), mhigh is the original high zone (85-95%), vhigh is the original vhigh (>95%).

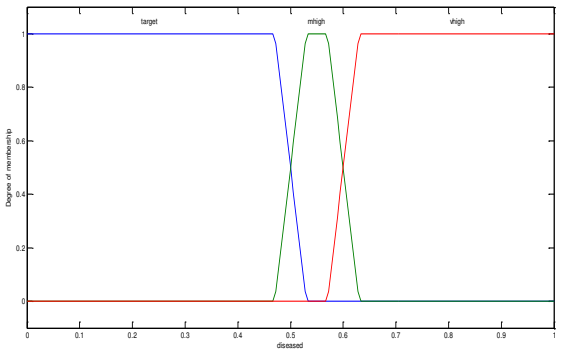


Figure 3. Illness, rehabilitation

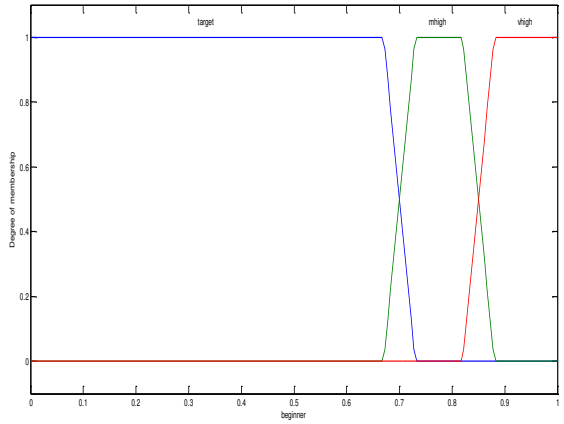


Figure 4. Beginner

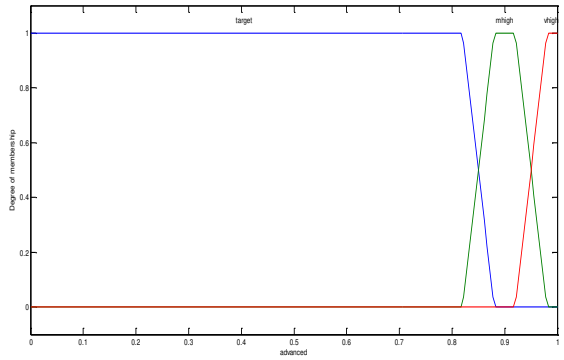


Figure 5. Regularly do sport

The second parameter under consideration is blood pressure, recognized as the most critical cardiovascular risk factor. Blood pressure signifies the force exerted by blood against vessel walls, notably in arteries. This force fluctuates rhythmically due to the heart's cyclic contractions. Systolic pressure denotes the peak pressure generated by the contracting left ventricle, while diastolic pressure refers to its lowest point during relaxation. To comprehensively evaluate blood pressure, both systolic and diastolic readings are essential, thus constituting the two additional input parameters in the Current Physical Status subsystem. Typically, systolic pressure rises during progressive exercise, while diastolic pressure remains steady or experiences a slight decline. The response at maximal or submaximal effort levels varies based on factors such as age, gender, and physical fitness. Older patients generally exhibit higher blood pressure readings, but this relationship inversely correlates with physical fitness, with better fitness levels associated with lower measured values. Additionally, men tend to have higher maximum systolic blood pressure compared to women. The blood pressure thresholds referred to in this paper stem from maximum values linked to age and gender, as detailed in [25], notwithstanding adjusted according to target zones, which also influence optimal blood pressure. Figure 6 shows the systolic blood pressure with the following parameters: low (<187), normal (187-204), increased (204-220), abnormal (>220). Figure 7 illustrates the diastolic blood pressure with the following constraints: Low (<80), normal (80-84), increased (85-90), abnormal (>90).

Table 1 illustrates the maximum systolic and diastolic blood pressure values categorized by age and sex [25].

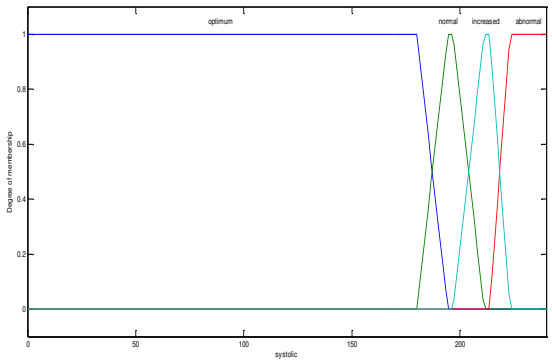


Figure 6. Systolic blood pressure

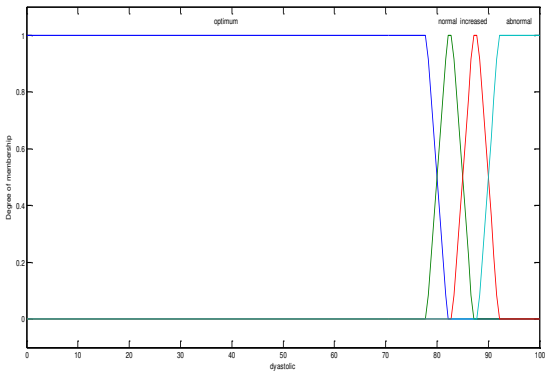


Figure 7. Diastolic blood pressure

1. TABLE: NORMAL BLOOD PRESSURE RESPONSE

Age	Men		Women	
	Systolic	Diastolic	Systolic	Diastolic
20-29	161-203	59-83	136-176	58-82
30-39	164-204	64-88	138-182	63-85
40-49	167-209	68-92	144-190	67-89
50-59	170-216	71-95	153-201	69-93
60-69	173-221	72-96	162-210	68-84
70-79	169-223	71-97	160-210	73-93

The output membership functions (risk levels) are: vsafe (<0,2), msafe (0,2-0,4) medium (0,4-0,6), mdangerous (0,6-0,8), vdangerous (0,8-1) as illustrasted in figure 8. (vsafe represents the smallest, while vdangerous the highest risk).

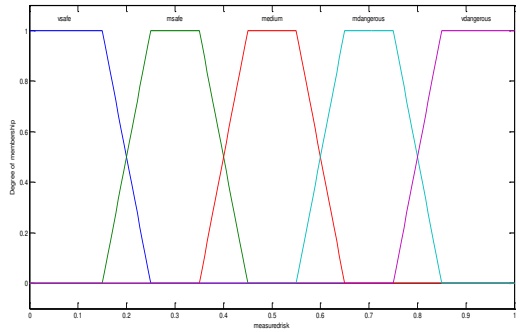


Figure 8. Risk level

The structure of the Current Physical Status subsystem is depicted in Figure 9, with input factors positioned on the left side and influential factors, crucial for refining membership functions, situated at the top of the figure. Input factors encompass heart rate (HR), systolic blood pressure (SBP), and diastolic blood pressure (DBP), while influential factors include disease condition, which transcribes whether the individual is diseased or someone who is a beginner or advanced in terms of sports, basic physical information which includes age and sex, and finally, the training target. This subsystem elaborates on the "Current Physical Status" input within the overarching model, as illustrated in Figure 1.

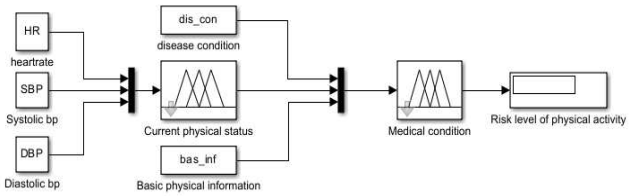


Figure 9. Current physical status subsystem

4. STATISTICS-BASED EVALUATION

4.1. Membership functions construction – an overview

Deriving membership functions from training data is a core challenge in fuzzy set theory. There are no definitive guidelines for selecting the appropriate method for generating these functions. Additionally, the task is complicated by a lack of consensus on how to define and interpret membership functions. For instance, Dubois and Prade [26] discuss the complexities and differing interpretations involved, underscoring the subjective nature of defining membership functions in fuzzy set theory. Therefore, various methods can be employed to generate membership values based on the desired interpretation.

Extensive literature focuses on creating membership functions to reflect subjective perceptions of vague concepts. However, these methods often cannot be directly applied to practical problems like fuzzy logic applications, which require modelling uncertainty in input data. There are no standard measures to evaluate the accuracy of generated membership functions, particularly for abstract concepts. Therefore, models must be flexible and easily adjustable to optimize algorithm performance. Given the importance of membership functions, multiple methods may be necessary, tailored to specific problems and data types.

An overview of some of these methods can be found in [27]. The authors provide a solid background on the various techniques available for generating membership functions. The discussed key techniques encompass Heuristic Methods, rooted in expert knowledge and intuitive grasp of problem domains, where the construction of membership functions relies on rules derived from human expertise and experience. Probability-Possibility Transformations involve converting probabilistic data into fuzzy membership functions, utilizing the interplay between statistical data and fuzzy sets to manage uncertainty. Cluster Analysis, exemplified by methods like fuzzy c-means (FCM), identifies natural data groupings to construct membership functions, assigning data points to clusters with varying membership degrees for smoother function creation. Neural Networks employ artificial neural networks to learn membership functions dynamically, particularly adept for complex pattern recognition tasks. Genetic Algorithms optimize membership functions by emulating natural selection, searching for optimal parameters to enhance fuzzy system performance. Histograms and Density Estimation employ statistical methods to estimate data density distributions, aiding in precise membership function creation through visual representation of data distribution. More details on the histogram method will be outlined in the next section, as this is the method used in this paper.

Although these methods were originally proposed for pattern recognition purposes, they can still be relevant to patient monitoring, and there is potential to combine them to create robust and accurate membership functions tailored to medical data.

A further contribution was introduced by Medaglia [28], who proposed an innovative method for constructing membership functions in convex normal fuzzy sets using Bézier curves. This technique offers significant flexibility and efficiency. The Bézier curve-based mechanism allows users to intuitively manipulate the shapes of membership functions to fit given data sets with minimal discrepancy. The paper includes several numerical experiments comparing this method to conventional approaches, demonstrating its superiority in producing accurate and reliable membership functions. One key advantage of this method is its ability to handle various data shapes intuitively. Traditional methods often require complex calculations and are less adaptable to different data distributions. In contrast, the Bézier curve-based approach simplifies the process, making it accessible even for those with limited technical expertise in fuzzy set theory. By harnessing these techniques, researchers and practitioners can develop more sophisticated monitoring systems that improve diagnostic accuracy, adapt to dynamic patient conditions, and support personalized patient care, ultimately leading to better health outcomes and enhanced quality of care for patients.

In the literature, fuzzy sets are frequently represented using triangular, trapezoidal, and bell-shaped membership functions [29][30].

A trapezoid shape input membership function is given by:

$$\mu_v(x) = \begin{cases} 0, & x < v_1 \\ \frac{x-v_1}{v_2-v_1}, & v_1 \le x \le v_2 \\ 1, & v_2 \le x \le v_3 \\ \frac{v_4-x}{v_4-v_3}, & v_3 \le x \le v_4 \\ 0, & x > v_4 \end{cases} \quad (1)$$

where, v_1, v_2, v_3, v_4 are the parameters of the membership function.

In the method proposed by Devi and Sarma [31], a parametric representation of the histogram is utilized to estimate fuzzy membership functions. This is achieved through the rational function approximation, where the parameters of the function are derived by applying least squares fit to the histogram values. Once obtained, these parameters are normalized to ensure that the function's maximum value is one. To determine the membership value for any given sample, these normalized parameters are substituted back into the rational function. This approach is particularly useful for classifying unlabelled samples. For each feature within each class, histograms are constructed, and the parameters representing the membership function are determined accordingly.

To obtain the desired results, the input MFs are tuned according to the personal characteristics of the patient. For simplicity, these values are represented by trapezoidal MFs.

4.2 Membership function fitting to the Histogram

Due to the complex interactions between input factors, it is difficult to precisely evaluate their effects on the measured values.

The data collected and recorded during monitoring can be used to assess the patient's current condition. Furthermore, these data can be recorded in a personal profile to personalize the evaluation in the future. According to the current state of the patient, the previous measurements performed under the approximately same conditions can be considered. Statistics, such as histograms, can be created based on these values to provide further insights. This histogram represents the normal reaction of the patient under the given conditions.

When a histogram is available, a membership function (MF) can be defined based on it. This function is piecewise linear whose highest point corresponds to the domain with the maximum value in the histogram and the rest of the function is created based on the remaining histogram values. Further details on this method can be found in [32][33].

This is how the membership functions of the inputs HR, SBP and DBP are created. After the histogram-based functions are available, original membership functions can be tuned accordingly. The functions tuned in this way are more reliable since the patient's normal reactions and the medical recommendation are taken into account together. These functions are the 'improved MFs' that can be obtained by simply calculating the mean of the correspondent parameters of the histogram-based functions and the tuned original membership functions.

5. CASE STUDY

In this section, the membership functions representing the medical recommendation for an 'Advanced Male 20-29' are presented, as well as the measurement-based statistics for a specific patient. The Mamdani-type inference system was implemented in MatLab Fuzzy Logic Designer alongside the rule base.

In order to define the membership functions for heart rate, consideration should be given to the predicted maximum heart rate and the recommended intensity (as outlined in Figure (2). Three antecedent fuzzy sets are applied: "target," representing the ideal heart rate zone for the patient; "mhigh," indicative of an elevated heart rate; and "vhigh," denoting a very high heart rate that is not advisable for the patient. These zones and the actual input value are delineated as a percentage of the individual's maximum heart rate. Figure 10 illustrates the heart rate antecedent sets for the group 'Advanced Male 20_29' which features males aged between 20-29 years who regularly do sports at an advanced level.

The membership functions for systolic blood pressure are also depicted in Figure 11. The thresholds establish the antecedent fuzzy sets: "low," representing hypotonic values; "normal," indicating the desired SBP value; "increased," signifying a somewhat higher but still acceptable value; and "abnormal" which is not recommended for the patient due to increased risk. Similarly, the antecedent sets for diastolic blood pressure are presented in Figure 12.

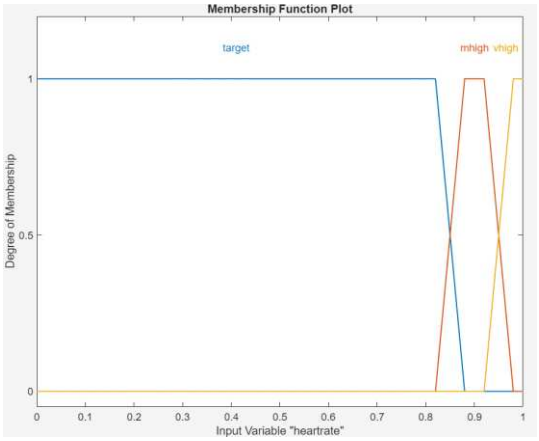


Figure 10. Heart rate zones for Advanced Male 20-29

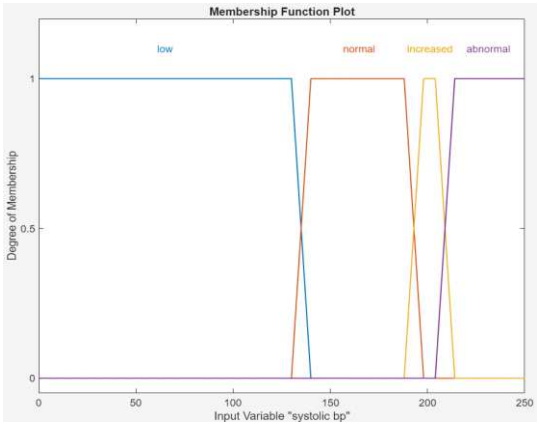


Figure 11. Systolic blood pressure for Advanced Male 20-29

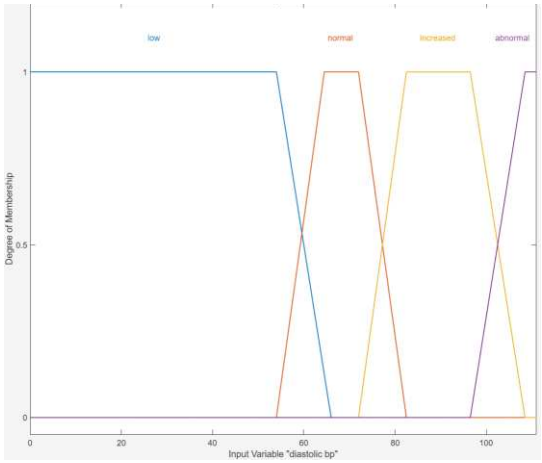


Figure 12. Diastolic blood pressure for Advanced Male 20-29

Utilizing the influential parameters described above, the membership functions can be adjusted to align with patient characteristics. Figures 13-15 show the statistics-based membership functions of a 21-Year-Old elite badminton athlete. His personal parameters are as follows: resting heart rate, HRrest=60bpm, maximum heart rate, HRmax=195bpm, systolic blood pressure, SBP=120mmHg, maximum systolic blood pressure, SBPmax=235mmHg, diastolic blood pressure, DBP=55mmHg, maximum diastolic blood pressure, DBPmax=82mmHg, weight=175kg. The athlete performed an incremental treadmill running test for the evaluation of maximal oxygen consumption (VO2max), anaerobic threshold, and time to exhaustion. He started exercising at a treadmill speed of 2.7 km/h and an inclination of 10% gradient for 3 min, and the speed and inclination were gradually increased every 3 min until he was exhausted or fatigued volitionally. Heart rate variability was examined using the Polar heart rate monitor over a period of 5 min at rest in the supine position [34]. In all cases, the examined HR, SBP and DBP were incrementally generated from the resting values to close to the maximum parameters. The graphics presented from this point forward in this chapter are dedicated to the profile of the 21-year-old elite badminton athlete.

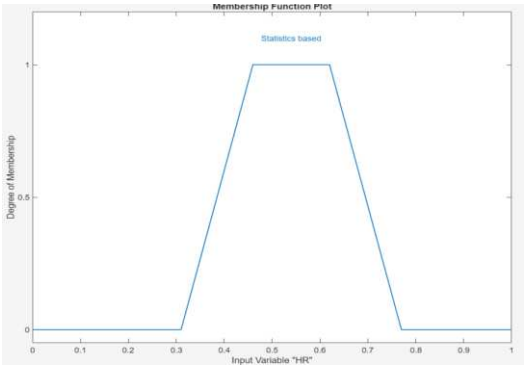


Figure 13. Stats-based fuzzy set representing the personal statistics for HR

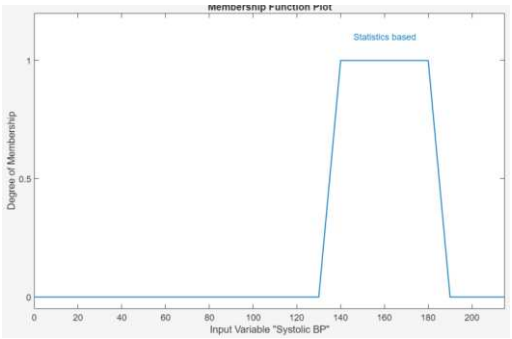


Figure 14. Stats-based fuzzy set representing the personal statistics for SBP

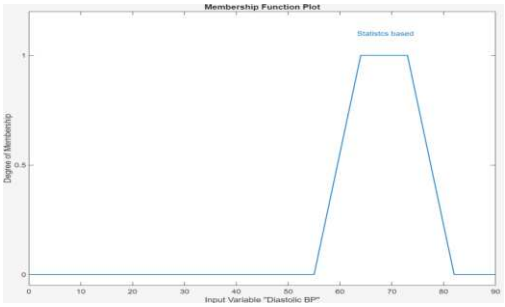


Figure 15. Stats-based fuzzy set representing the personal statistics for DBP

Next, the statistics-based MFs were compared to the medical recommendations and then aggregated to the 'improved MFs' as shown in figures 16-18. These new membership functions (figures 19-21) were then used to evaluate the risk level of the patient (see table 2).

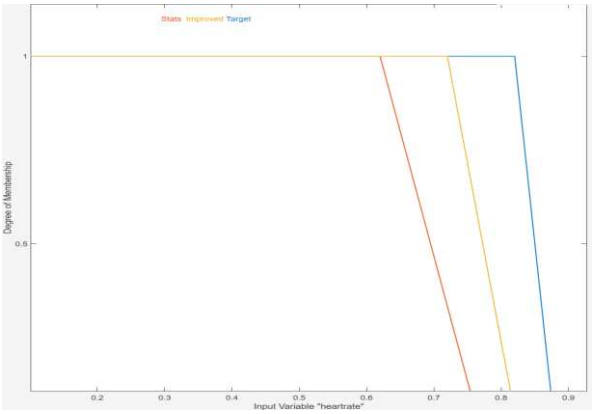


Figure 16. Comparison between the statistics-based (Input Values: [0, 0, 0.62, 0.77]) and the medical recommendation MFs for HR (Input Values: [0, 0, 0.82, 0.88])

The HR input domain for the statistics-based function ranges from 0 to 0.77, while the medical recommendation function ranges from 0 to 0.88. The significant difference occurs in the mid-to-high range values, specifically from 0.62 to 0.88. In the range [0.62, 0.77], the statistics-based function provides a more detailed classification of HR values. This range is crucial as it indicates moderate risk levels where the athlete's HR is elevated but not yet in the high-risk category. The medical recommendation function does not start this categorization until a higher range [0.82,

0.88], suggesting that the athlete-specific model is more sensitive to increases in HR, thus providing earlier warnings and potentially better risk management during physical activity. The personalized statistics-based function allows for a broader range of HR values to be classified as "moderately safe" before reaching "moderately dangerous," reflecting the athlete's higher tolerance for elevated heart rates during intense physical activity. The parameters of the 'improved' MFs for HR are [0 0 0.74 0.825] which were obtained by taking the mean of the corresponding parameters of the statistics-based function and the medical recommendation function.

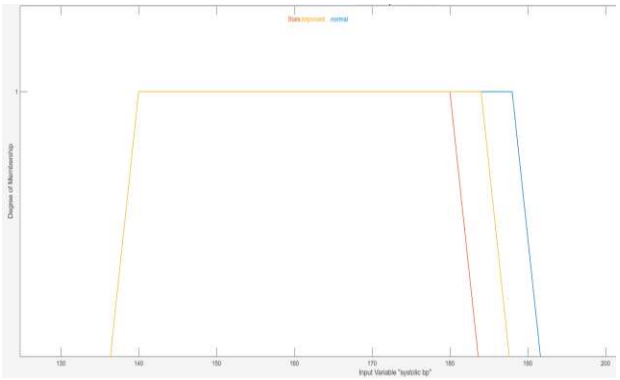


Figure 17 . Comparison between the statistics-based (Input Values: [130, 140, 180, 190]) and the medical recommendation ([130, 140, 188, 198]) MFs for SBP

The range of interest here is [180,190] for the statistics-based function versus [188,198] for the medical recommendation. In the range [180, 190], the statistics-based function identifies SBP values within this interval as moderate to high risk. This is significant because it indicates that the athlete-specific function flags elevated SBP values earlier than the medical recommendation function, which only starts this categorization at higher SBP values [188, 198]. By focusing on this range, it is evident that the personalized model is tailored to detect potential cardiovascular strain at lower thresholds. This early detection capability enables better prevention and management strategies during high-intensity exercise. Similar to the previous analysis, the parameters of the 'improved' MFs for SBP are [0, 0, 184, 194].

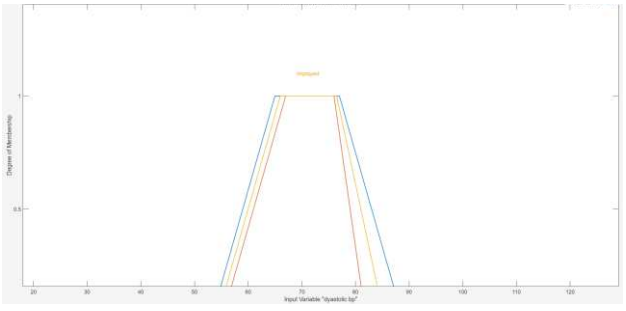


Figure 18. Comparison between the statistics-based (Input Values: [55, 67, 76, 82]) and the medical recommendation (Input Values: [53, 65, 77, 89]) MFs for DBP

The input domain includes values from resting diastolic pressure to the peak DBP observed during the athlete's maximal activity levels. In the range [76 82], the statistics-based function assigns higher risk levels compared to the medical recommendation, which considers values up to 89 before assigning similar risk levels. This indicates that the athlete-specific function is more conservative and sensitive to increases in DBP. The range [76 82] is critical as it represents values where the athlete's diastolic pressure is elevated but still below the extreme high-risk category. This sensitivity helps in the early detection and management of cardiovascular risks specific to the athlete's physiology. Analogously, the parameters of the 'improved' MFs for DBP are [54, 66, 76.5, 85.5].

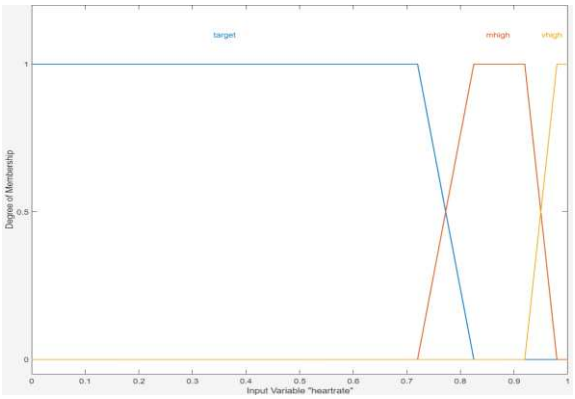


Figure 19. Improved MFs for HR

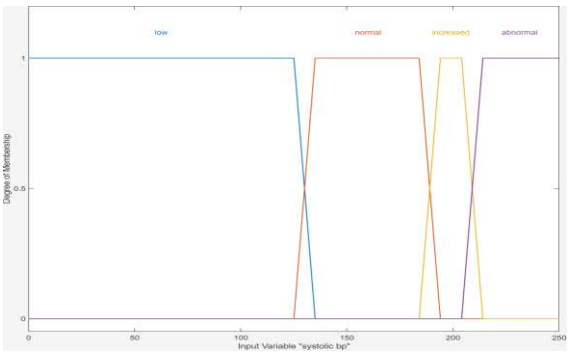


Figure 20. Improved MFs for SBP

				Improved MFs		Medical recommendation	
HR (bpm)	% of HRmax (bpm)	SBP (mmHg)	DBP (mmHg)	Risk Level (numerical)	Linguistic terms	Risk Level (numerical)	Linguistic terms
60	0.31	120	55	**0.108**	**Very safe**	**0.111**	**Very safe**
70	0.36	130	58	**0.226**	**Moderately safe**	**0.121**	**Very safe**
80	0.41	140	61	**0.21**	**Moderately safe**	**0.193**	**Moderately safe**
90	0.46	150	64	**0.153**	**Very safe**	**0.131**	**Very safe**
100	0.51	160	67	**0.106**	**Very safe**	**0.106**	**Very safe**
110	0.56	170	70	**0.106**	**Very safe**	**0.106**	**Very safe**
120	0.62	180	73	**0.106**	**Very safe**	**0.106**	**Very safe**
130	0.67	190	76	**0.361**	**Moderately safe**	**0.224**	**Moderately safe**
140	0.72	200	79	**0.561**	**Medium**	**0.539**	**Medium**
150	0.77	210	82	**0.692**	**Moderately dangerous**	**0.698**	**Moderately dangerous**

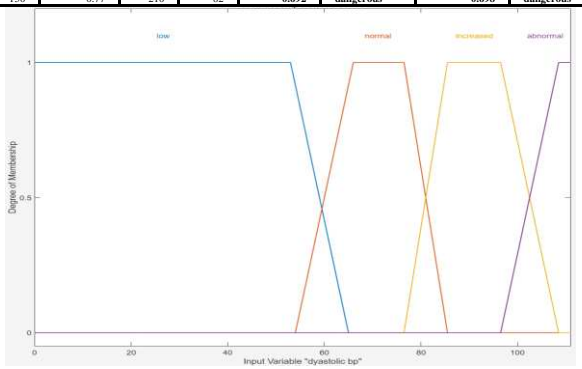


Figure 21. Improved MFs for DBP

Our analysis of the input domains where statistics-based functions and medical recommendation functions differ highlights the significant impact of personalized modelling. The specific ranges of [0.62, 0.77] for HR, [180, 190] for SBP, and [76, 82] for DBP illustrate areas where the statistics-based functions offer more detailed and early risk categorization. These distinctions underscore the critical importance of personalized health monitoring, enabling timely and precise interventions tailored to individual physiological responses, especially for athletes.

Moreover, incorporating the improved MFs into our analysis demonstrates the value of a balanced approach. The improved MFs, which are calculated as the means of the adjacent statistics-based and medical recommendation functions, provide a smoother and more adaptive risk assessment. These functions offer early warnings similar to the personalized statistics-based functions while gradually aligning with the thresholds set by medical recommendations. This comprehensive and timely health monitoring approach is particularly beneficial for managing the cardiovascular health of athletes, ensuring their safety, and optimizing performance.

6. RESULTS

The risk evaluation of personal statistics is illustrated in Table 2. The table presents a comprehensive risk evaluation of a 21-year-old male elite badminton athlete, focusing on heart rate (HR), systolic blood pressure (SBP), diastolic blood pressure (DBP), and their corresponding risk levels both numerically and in linguistic terms. The comparison of risk levels derived from improved membership functions (MFs) and medical recommendations offers valuable insights into the athlete's cardiovascular status under varying physiological conditions.

2. TABLE: RISK EVALUATION OF A 21-YEAR-OLD MALE BADMINTON ATHLETE

At lower heart rates and blood pressure values the risk level is predominantly categorized as ‘very safe’. The medical recommendations are very similar with the improved MFs, suggesting a low risk for cardiovascular events. Which means that in this zone the statistics-based approach is reliable as it presents similar results as the medical recommendation. For moderated heart rates and blood pressure values the risk level varies from ‘very safe’ to 'moderately safe’ with a slight variation between the improved MFs and the medical recommendation. For higher heart rates and blood pressure values we notice an increase in cardiovascular risk in both the improved MFs and the medical recommendation.

The consistency between improved MFs and medical recommendations validates the reliability of the model, particularly at higher heart rates where risk levels are more pronounced. This approach emphasizes the importance of personalized risk assessment models, ensuring the patient’s safety and optimal performance management.

7. CONCLUSION

In sports and physical activity, ensuring health monitoring is crucial for participants across all levels. Each individual engages in sports with varying physiological responses and health profiles, making personalized assessment methods essential. Traditional approaches to health monitoring do not always take that into account.

Integrating fuzzy logic into health monitoring systems can be advantageous in dealing with such challenges as it can handle uncertainty, imprecision, or subjectivity of the input data. It can offer more accurate and personalized results, leading to safer practice of physical activity.

In this paper, the authors analysed the risk levels of the current activity using a hierarchical fuzzy model structure with a focus on the current physical status subsystem. To assess the patient's current condition data (which includes but is not limited to HR, SBP, DPB values as well as sampling frequency, duration, sex, and activity type) is collected and recorded during monitoring. This data can be recorded in a personal profile to personalize the evaluation in the future and the previous measurements performed under the approximately same conditions can be considered. The statistics-based approach was used. Histograms were created representing personal statistics, i.e., the normal reaction of the patient under the given conditions. When the histogram is available, a membership function (MF) can be defined based on it. This function is piecewise linear whose highest point corresponds to the domain with the maximum value in the histogram and the rest of the function is created based on the remaining histogram values. This way, the membership functions of the inputs HR, SBP and DBP are created. After the histogram-based functions are available, original membership functions can be tuned accordingly. The functions tuned in this way are more reliable, since the

patient's normal reactions and the medical recommendation are taken into account together.

This study shows that there is consistency between the improved MFs and the medical recommendations, which validates the reliability of the model, particularly at higher heart rates where risk levels are more pronounced. This approach emphasizes the importance of personalized risk assessment models, ensuring the patient's safety and optimal performance management.

In the future, authors aim to develop different mathematical methods that can be used to represent the patient's statistics and for fitting the membership functions.

ACKNOWLEDGEMENT

This work was supported by the Fuzzy Systems Scientific Group at the Bánki Donát Faculty of Mechanical and Safety Engineering of Óbuda University.

REFERENCES

[1] Michelini, E. (2014). *The role of sport in health-related promotion of physical activity: the perspective of the health system*. Springer.

[2] Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2013). Long-term health benefits of physical activity–a systematic review of longitudinal studies. *BMC public health*, *13*, 1-9.

[3] Akkaş, M. A., Sokullu, R., & Çetin, H. E. (2020). Healthcare and patient monitoring using IoT. *Internet of Things*, *11*, 100173.

[4] Kotorov, R., Chi, L., & Shen, M. (2020). Personalized monitoring model for electrocardiogram signals: diagnostic accuracy study. *JMIR Biomedical Engineering*, *5*(1), e24388.

[5] Ersue, M., Romascanu, D., Schoenwaelder, J., & Sehgal, A. (May 2015). Management of Networks with Constrained Devices: Use Cases. *IETF Internet Draft*.

[6] Dey, N., Hassanien, A. E., Bhatt, C., Ashour, A. S., & Satapathy, S. C. (2018).

[7] C. De Capua, A. Meduri, and R. Morello, "A smart ECG measurement systems based on web-service-oriented architecture for telemedicine applications," *IEEE Trans. Instrum*. Meas., vol. 59, no. 10, pp. 2530–2538, Oct. 2010.

[8] A. Alamri, C. Jongeun, and A. El Saddik, "AR-REHAB: An augmented reality framework for poststroke-patient rehabilitation," *IEEE Trans. Instrum. Meas.,* vol. 59, no. 10, pp. 2554–2563, Oct. 2010.

[9] F. Rahnman, A. Kumar, G. Nagendra, and G. Sen Gupta, "Network approach for physiological parameters measurement*," IEEE Trans. Instrum. Meas.*, vol. 54, no. 1, pp. 337–346, Feb. 2005

[10] L. Fanucci, S. Saponara, T. Bacchillone, M. Donati, P. Barba, I. Sánchez-Tato, *et al*., "Sensing devices and sensor signal processing for remote monitoring of vital signs in CHF patients," *IEEE Trans. Instrum. Meas*., vol. 65, no. 3, pp. 553–569, Mar. 2013.

[11] Vermesan, O., & Friess, P. (Eds.). (2013). Internet of things: Converging technologies for smart environments and integrated ecosystems. River Publisher.

[12] M. Takács, "Extended fuzzy methods in risk management," in *Proc. 14th WSEAS Int. Conf. Appl. Math*., Puerto De La Cruz, Spain, Dec. 2009, pp. 300–304.

[13] A. Karime, M. Eid, J. M. Alja'am, A. El Saddik, and W. Gueaieb, "A fuzzy-based adaptive rehabilitation framework for home-based wrist training," *IEEE Trans. Instrum. Meas*., vol. 63, no. 1, pp. 135–144, Jan. 2013.

[14] E. Tóth-Laufer, M. Takács, I.J. Rudas, "Interactions Handling Between the Input Factors in Risk Level Calculation, "in *Proc. of the 11th IEEE Int. Symposium on Applied Machine Intelligence and Informatics*, Herl'any, Slovakia, January 31-February 2, 2013, pp. 71-76, doi: 10.1109/SAMI.2013.648097

[15] R.-E. Precup, R.-C. David, E. M. Petriu, M.-B. Radac, S. Preitl, and J. Fodor, "Evolutionary optimization-based tuning of low-cost fuzzy controllers for servo systems," Knowl. -Based Syst., vol. 38, pp. 74-84, Jan. 2013.

[16] S. Szénási, Z. Vámossy, M. Kozlovszky, "Preparing Initial Population of Genetic Algorithm for Region Growing Parameter Optimization", 4th IEEE International Symposium on Logistics and Industrial Informatics (LINDI), 5-7. Sept 2012, pp. 47-54.

[17] A. Gegov, Fuzzy Networks for Complex Systems - A Modular Rule Base Approach. Berlin, Heidelberg: Springer-Verlag, 2010.

[18] E. Tóth-Laufer, M. Takács, I.J. Rudas, "Conjunction and Disjunction Operators in Neuro-Fuzzy Risk Calculation Model Simplification" in 13th IEEE Internation Symposium on Computational Intelligence and Informatics (CINTI 2012), Budapest, Hungary, November 20-22, 2012, pp. 195-200, ISBN: 978-1-4673-5204-8, IEEE Catalog Number: CFP1224M-PRT, DOI: 10.1109/CINTI.2012.6496759.

[19] A. Gegov, "Advances in fuzzy systems and networks," in Proc. IEEE Conf. of Intelligent Systems (IS 2012), Sofia, Bulgaria, 2012, pp. 33-40.

[20] Y. Wu, Y. Ding, and H. Xu, "Comprehensive fuzzy evaluation model for body physical exercise," in Risk Life System Modeling and Simulation (Lecture Notes in Computer Science). New York, NY, USA: SpringerVerlag, 2007, pp. 227–235.

[21] J. Ogorevc, A. Podlesek, G. Gersak, and J. Drnovsek, "The effect of mental stress on psychophysiological parameters," in Proc. IEEE Int. Workshop Med. Meas. Appl., May 2011, pp. 294–299.

[22] K. Hottenrott, Training with the Heart Rate Monitor. Heidelberg, Germany: Quelle & Meyer, 2007.

[23] (2014, Jan. 21). Polar RS800CX User Manual—Polar USA [Online]. Available:http://www.polar.com/e_manuals/RS800CX/Polar_RS800CX_ user_manual_English/manual.pdf

[24] S. Ahmad, M. Bolic, H. Dajani, V. Groza, I. Batkin, and S. Rajan, "Measurement of heart rate variability using an oscillometric blood pressure monitor," IEEE Trans. Instrum. Meas., vol. 59, no. 10, pp. 2575–2590, Oct. 2010.

[25] M. C. Sieira, A. O. Ricart, and R. S. Estrani, "Blood pressure response to exercise testing," in Apunts Med Esport. Amsterdam, The Netherlands: Elsevier, 2010, pp. 191–200

[26] D. Dubois, H. Prade, Fuzzy sets - a convenient fiction for modeling vagueness and possibility, IEEE Trans. on Fuzzy Systems 2 (1) (1994) 16-21.

[27] Medasani, S., Kim, J., & Krishnapuram, R. (1998). An overview of membership function generation techniques for pattern recognition. *International Journal of approximate reasoning*, *19*(3-4), 391-417.

[28] Medaglia, A. L., Fang, S. C., Nuttle, H. L., & Wilson, J. R. (2002). An efficient and flexible mechanism for constructing membership functions. *European Journal of Operational Research*, *139*(1), 84-95.

[29] J. Dombi, Membership function as an evaluation, Fuzzy Sets and Systems 35 (1990) 1–21.

[30] Corrente, S., Greco, S., & Słowiński, R. (2017). Handling imprecise evaluations in multiple criteria decision aiding and robust ordinal regression by n-point intervals. *Fuzzy Optimization and Decision Making*, *16*, 127-157.

[31] Devi, B., & Sarma, V. V. S. (Year). Estimation of Fuzzy Memberships from Histograms. School of Automation, Indian Institute of Science, Bangalore, India.

[32] Tóth-Laufer, E., & Várkonyi-Kóczy, A. R. (2014, June). A personal profile-based patient-specific anytime risk calculation model. In *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (pp. 1-6). IEEE.

[33] Tóth-Laufer, E., & Várkonyi-Kóczy, A. R. (2014). Personal-statistics-based heart rate evaluation in anytime risk calculation model. *IEEE Transactions on Instrumentation and Measurement*, *64*(8), 2127-2135.

[34] Tai, C. C., Chen, Y. L., Kalfirt, L., Masodsai, K., Su, C. T., & Yang, A. L. (2022). Differences between Elite Male and Female Badminton Athletes Regarding Heart Rate Variability, Arterial Stiffness, and Aerobic Capacity. *International journal of environmental research and public health*, *19*(6), 3206.

BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Origami's Mathematical Precision: Transforming Medicine Through Folded Geometry

Enisa Trubljanin
University of Donja Gorica (Faculty of Applied Science, Mathematics), Podgorica, Montenegro
enisa.trubljanin@udg.edu.me

Abstract

Origami constructions, inspired by the ancient Japanese art of folding paper, are gaining importance in mathematics and science. The paper explores the mathematical aspects of origami, with a focus on their geometric constructions and applications. Applications of origami in the field of medicine were investigated. The paper also discusses and addresses advanced topics such as modular origami and origami computing. Through all these topics, the richness and depth of mathematical ideas hidden behind folded paper are shown, thus emphasizing their importance in the modern world. An understanding of origami mathematics was provided, thus providing an incentive for further research and applications in medicine. Through this work, the goal is to understand the potential of origami as a creative and interdisciplinary tool in the medical field, highlighting opportunities for innovation and progress.

Keywords: Origami, Origami geometry, Retina

1. INTRODUCTION

Origami, the Japanese art of folding paper, is not just an aesthetic form of entertainment or decoration; it represents a deep mathematical and geometric discipline that has given birth to some of the most innovative solutions in modern science [1]. While origami is traditionally associated with the art of harmonious folding of paper, its application is increasingly becoming a subject of interest and research. This paper introduces the world of the use of origami constructions in medicine, highlighting the historical context, theoretical foundations, mathematical problems and contemporary applications of this fascinating discipline.

Pioneering works such as Akira Yoshizawa and Robert J. Lang contributed to the development of origami theory and opened the door for mathematical analysis in this area [1]. Today, origami is not only an art, but also a tool for solving complex mathematical problems and engineering challenges.

The geometric basis of origami constructions lies in simple axioms and rules of paper folding. On the other hand, behind that simplicity are hidden deep mathematical theories that reach into the areas of algebra, topology and geometry. Mathematical concepts such as the construction of angles, lengths and surfaces are reinterpreted through the prism of origami, opening up new areas of research and application.

Contemporary applications of origami extend beyond the boundaries of mathematics into fields such as architecture, robotics and medicine. In medicine, origami is used to design implants and surgical instruments that can be folded inside the patient's body. This paper explores the breadth of origami mathematics, highlighting its importance for medical purposes. The paper deepens this application in medicine, providing a comprehensive insight into the world of origami constructions and their vital role in various areas of human activity in medicine.

One area that is gaining increasing attention is medicine, where origami is proving to be an extremely useful technique with the potential for revolutionary applications. Through the precise folding and shaping of, origami provides unique opportunities for the development of innovative medical devices, biomaterials and therapeutic interventions [2].

This paper explores the diverse applications of origami in medicine, exploring how this ancient art can contribute to the advancement of modern medical practice. Through a review of available

literature, analysis of current research and case studies, it is shown how origami can be creatively and effectively used in diagnostics, therapy, tissue reconstruction, as well as in the development of innovative medical devices and biomaterials. It has been studied how features of origami, such as precision, flexibility and personalization, open the door to advanced medical applications that can improve the quality of life of patients and facilitate the work of medical staff. In order to intuitively understand the topic of origami constructions, some examples of models from everyday life are shown *(Figure 1)*.

Figure 1. Examples of some origami models

2. ORIGAMI CONSTRUCTIONS

Origami mathematics is an interdisciplinary field that combines geometry, algebra, number theory, and other mathematical disciplines to explore the basic principles and theorems underlying paper folding. Through the study of theoretical foundations, mathematicians are able to explore new ideas, theorems and algorithms that have wide application in practice.

Over the centuries, origami constructions have evolved from simple shapes to complex geometric structures. Although origami have traditionally been associated with aesthetic and artistic value, their mathematical potential has only recently become the subject of systematic study. One of the most significant moments in the development of origami mathematics was the work of the American mathematician Robert J. Lang. Lang is known for his discovery of Lang angles, a key concept in origami theory [3]. His works established the mathematical foundations of origami geometry and provided tools for solving complex problems in this field.

The development of origami constructions in mathematics did not stop at basic theorems. Through development, origami constructions have gone from simple artistic forms to complex mathematical structures with a wide range of applications even in the fields of medicine.

3. DEVELOPMENT OF ORIGAMI IN MEDICAL APPLICATION

Origami goes back a long way, where it originally served as a means of entertainment and artistic expression. However, in the medical community, its application began to be investigated with increasing attention as its potentials for solving medical problems and improving therapeutic approaches were discovered.

Throughout history, key events have marked the evolution of origami in medicine. Pioneering work by researchers and medical professionals has resulted in numerous innovations, from simple origami models used in education and simulation, to complex origami-inspired medical devices and biomaterials [4]. Origami models began to be used as an educational tool for medical students, allowing them to better understand anatomy and medical procedures, and have evolved into an advanced application in medicine today. The origami technique is the inspiration for many structures in medicine today *(Figure 2)*.

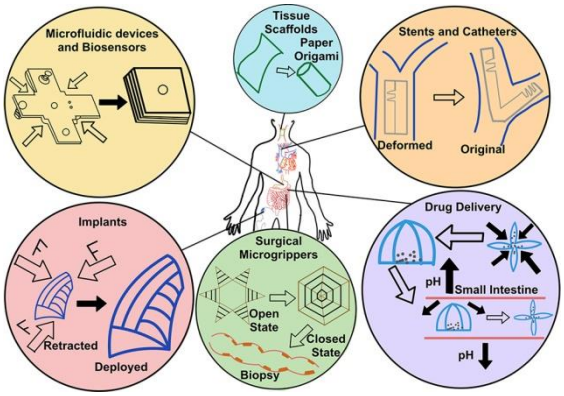


Figure 2. Some of the inspirations for the application of origami in medicine [4]

3.1 Benefits of origami in medicine

Origami requires a high level of precision in folding and shaping. This feature allows medical professionals to create accurate models of anatomical structures or medical devices needed for diagnosis, therapy or training. Origami allows customization of shape and size according to the specific needs and requirements of medical procedures or patients. This flexibility makes origami particularly useful in fields such as

reconstructive surgery and the creation of personalized medical devices [5].

Using origami can reduce the need for invasive medical procedures. For example, origami-inspired medical devices could be smaller and more flexible, allowing for less traumatic procedures and faster patient recovery. Combining origami principles with medical technology can lead to innovative solutions and advances in medicine. Origami-inspired medical devices or biomaterials may open up new areas of research and therapeutic possibilities. The origami technique uses basic materials that can reduce the cost of manufacturing medical models or devices compared to traditional methods that require more expensive materials or technologies [5].

One of the key advantages of origami medical devices is their ability to adapt to the patient's anatomy. Inspired by origami constructions, these devices can be folded into a small form for easy insertion into the patient's body, and then unfolded in the appropriate place to perform their function *(Figure 3)* [6]. Origami endoscopy tools enable precise and minimally invasive exploration of the inside of the body, which reduces trauma and patient recovery time. Origami principles can be applied in reconstructive surgery to shape and reconstruct tissues that have been damaged or lost through trauma, disease or surgery.

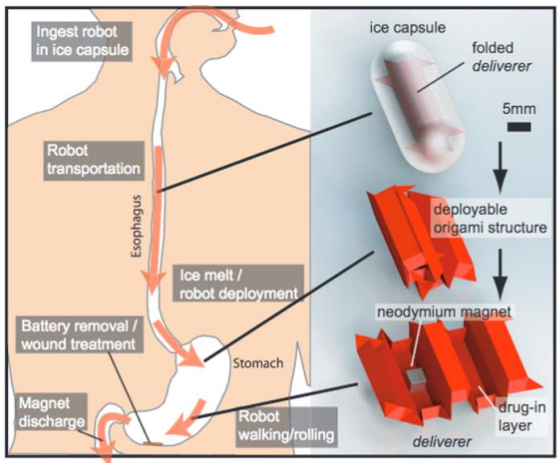


Figure 3. An example of a capsule containing an origami structure
[6]

The origami technique can be used to design and create three-dimensional templates that can be used as guides during reconstructive surgeries. These templates allow surgeons to precisely shape transplanted or reconstructed tissue sections, leading to better functional and aesthetic outcomes for patients [7].

4. ORIGAMI BIOMEDICAL IMPLANTS

Origami biomedical implants represent an innovative technique that combines the principles of origami with biomedical technology to develop advanced medical devices and implants [8]. This technique is based on the idea of reshaping materials, such as metals or polymers, into complex three-dimensional structures through folding and folding, thus enabling the creation of devices that are adaptable, compact and functional.

Biomedical origami implants represent a field of constant innovation in medicine. New technologies and materials are constantly being developed to improve the characteristics and performance of these implants, providing new opportunities for the therapy and treatment of various medical conditions.

Origami is used in the design of medical implants and prostheses. Folding origami-inspired implants can be easily transported and implanted at the site of intervention, then unfolded to provide the required support or functionality [9]. These implants are particularly useful in areas such as tissue reconstruction, regenerative medicine and orthopedics.

Foldable sensors are used to monitor a patient's vital signs or detect specific biomarkers in the blood. Sew sensors are also used for the purpose of folding origami at a specific location in the body. These devices enable fast, accurate and non-invasive diagnostics, which improves the quality and efficiency of medical care.

The concept of Origami implant design represents an innovative approach to 3D integration, aiming to tackle the challenges associated with size and cost limitations in biomedical implants [9]. By breaking down large systems into smaller chips and employing advanced 3D integration techniques, these components can be seamlessly folded to achieve compactness for implantation and later unfolded within the body. This strategy allows for the partitioning of electronics into functional blocks, facilitating mass production, while also enabling the assembly of customized implants from these cost-effective modules.

The application of this method can be particularly advantageous in the context of retinal prostheses. The retina is a layer of photosensitive nerve cells located at the back of the eye, which plays a key role in the vision process. Its basic function is to convert light signals received by the eye into electrical impulses that the brain can interpret as visual information [10]. For the purposes

of forming and installing a retinal prosthesis in the eye, origami principles have proven to be extremely useful. Instead of utilizing a single large chip, the implementation of multiple smaller chips dispersed across a flexible substrate offers a promising alternative. Prior to implantation, this substrate can be folded, allowing for ease of insertion, and subsequently unfolded within the eye to conform to its curved shape *(Figure 3)* [11].

By doing so, the implant achieves a customized fit with the retina, optimizing electrode contact and enhancing the effectiveness of stimulation. Consequently, our innovative origami-based approach introduces the potential to develop an elongated planar system that can be seamlessly folded into a compact structure for minimally invasive surgical procedures, before transitioning into its functional configuration.

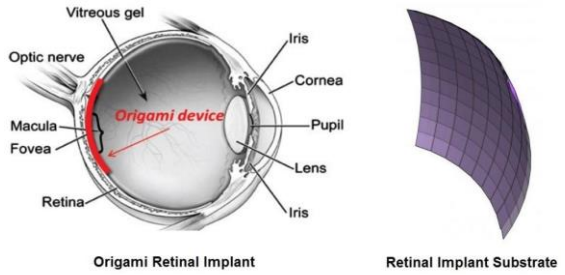


Figure 3. Appearance of the retinal implant substrate [12]

Furthermore, the adaptability of such a system to the eye's curvature allows for the optimization of chip and electrode placement through the strategic design of the origami structure [11].

4.1 Sensors in an origami retinal implant

Sensors made using the origami technique for the retina of the eye are sophisticated medical devices designed to monitor the health of the eye and diagnose various diseases. This technology uses the principles of origami geometry to create flexible and adaptable sensors that can be precisely placed on the surface of the retina.

Typically, retinal sensors are made of thin, flexible materials such as polymer or silicone that allow for comfortable placement on the surface of the eye without causing discomfort or damage [12]. Origami techniques are used to shape these materials into complex three-dimensional structures that can be folded and shipped in a relatively small space, then unfolded and placed on the surface of the retina.

The process of making these sensors involves several steps. First, the material is carefully selected and prepared to be compatible with human tissue and the environment of the eye. Folding and shaping techniques inspired by origami geometry are then used to transform the material into the desired shape. This may involve folding, folding or removing parts of the material to achieve the desired structure.

Once the sensors are shaped, they can be mounted on a thin foil or support that is then carefully placed on the surface of the retina. This application requires high precision and careful manipulation to ensure that the sensors are properly positioned and securely attached to the retina.

Once installed, the sensors can monitor various parameters of the eye, such as intraocular pressure, temperature, pH value or the concentration of certain molecules in tears [12]. That data can be vital for monitoring eye health and diagnosing conditions like glaucoma, diabetic retinopathy or dry eye.

By using the origami technique to create sensors in the retina of the eye, high precision, flexibility and comfort for patients is achieved. This innovative technology promises to improve the diagnosis and monitoring of eye diseases, paving the way for personalized treatment approaches and better management of ophthalmic conditions.

Given that the Origami implant is positioned within the body, the precision of chip alignment poses a challenge. Additionally, this alignment may shift over time due to patient mobility and tissue dynamics. Consequently, the proximity communication system must intermittently assess its alignment and adjust accordingly to optimize power usage while maintaining a desired data transmission rate. Apart from optimizing power efficiency, alignment sensing serves as a valuable tool for monitoring the implant's deployment progress [12]. Considering the stringent power limitations inherent in implants situated within delicate organs like the eye, the alignment adjustment process must prioritize energy conservation and computational simplicity.

5. ENDOSCOPIC INSTRUMENTS

Flexible endoscopic instruments represent a key aspect in the diagnosis and therapy of internal organs, enabling precise visualization and intervention without the need for invasive surgery. Origami principles are increasingly applied in the

design of these instruments in order to achieve greater flexibility, precision and adaptability.

Internal anatomical structures are often complex and diverse. Flexible endoscopic instruments must be able to adapt to different shapes and curvatures of organs in order to efficiently perform diagnostic or therapeutic procedures.

The origami approach allows engineers to design endoscopic instruments that can be bent and shaped in real time. This allows surgeons to precisely maneuver instruments through complex anatomies, such as bowel bends or gastric curvatures, with minimal trauma or discomfort to the patient [13].

By using flexible endoscopic instruments inspired by origami principles, it is possible to reduce the need for invasive surgical procedures. These instruments allow access to internal organs through natural openings or small incisions, which reduces the risk of complications and speeds up the patient's recovery process.

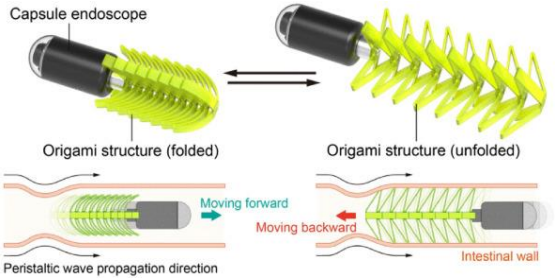


Figure 4. Endoscopic capsule for visualization of internal organs [14]

Using origami principles, endoscopic instruments with improved folding and folding mechanisms can be developed, which can improve their compactness and maneuverability. For example, inspired by origami techniques, instruments can be designed that can be folded compactly during transport and then expand to the appropriate shape and size during use. This facilitates instrument handling and may reduce the need for additional auxiliary tools during procedures. Also, the application of the origami principle can enable the development of endoscopic instruments with multiple functions, which can further improve the efficiency and practicality of these medical devices [15].

6. ORIGAMI COMPUTING

Origami algorithms and computer geometry represent a field of research that is oriented towards the application of mathematical principles and techniques of computer geometry in the analysis, simulation and generation of origami constructions. This area enables the development of sophisticated algorithms for the automatic generation of origami models with different characteristics and properties and even those that are used for medical purposes.

One of the main properties of origami algorithms is geometric transformation. These algorithms use various geometric transformations, such as translation, rotation and scaling, to manipulate and transform paper modules or parts of an origami model. These transformations allow the creation of different shapes and structures of origami models with the desired characteristics and properties which are used as such and to create models that have practical application in medicine..

In addition, origami algorithms also use various optimization techniques to find optimal configurations of origami models in terms of efficiency, stability or aesthetics [16]. These algorithms use mathematical methods such as linear programming, genetic algorithms or simulated annealing to find the best solution according to the given goals and constraints. They have a special role in the context of creating a model that will take up as little space as possible.

Computational geometry is important in the development of origami algorithms, allowing efficient manipulation of geometric data and structures. These algorithms use different algorithms and techniques to analyze and manipulate origami models, such as algorithms for cutting, overlapping checks or solving spatial problems. Examples of origami algorithms include algorithms for the automatic generation of origami models, algorithms for analyzing the stability of origami structures or algorithms for optimizing the shape and dimensions of origami models [16].

7. CONCLUSION

Researching origami mathematics and its application in medicine is a fascinating journey through the history, theory and application of this unique combination of art and science. Through consideration of various aspects of origami mathematics, its connections with other sciences are revealed, thus revealing the complex connections between paper folding and fundamental mathematical concepts. Through solving problems in medicine using the origami technique, new ways of approaching classical mathematical challenges are discovered.

Origami, as the art of translating a flat surface into complex three-dimensional structures, exhibits astonishing mathematical precision that brings a potential revolution in the medical field. Through the combination of geometry, engineering and

medical needs, origami offers innovative approaches in diagnosis, therapy and surgery. This synergy between art and science enables the development of sophisticated medical devices such as stents, microrobots and affordable diagnostic tools.

Advances in medical origami design enable more personalized therapies, minimize invasiveness and optimize treatment outcomes. The application of origami geometry in the design of retinal implants provides flexibility and adaptability, enabling precise placement and reducing the risk of complications. Furthermore, origami-inspired microrobots represent the possibility of precisely delivering drugs or performing microsurgical procedures in places inaccessible by traditional methods. These advanced devices promise to revolutionize the way diseases are diagnosed and treated, particularly in areas such as neurosurgery, cardiology and oncology.

Despite the promising potential, there are challenges that must be addressed. This includes developing reliable materials that are compatible with the human body, ensuring precision in the manufacturing process, as well as adapting regulatory frameworks to enable rapid integration of these technologies into clinical practice. However, with growing interest from the research community and industry, the difficulties will be overcome, and origami will remain a key technology in transforming medicine. In each fold lies an untold story of imagination becoming reality, of complex mathematical ideas taking shape through simple folding processes. Origami constructions are a symbol of harmony between art, mathematics and other sciences.

References

[1] Hatori, K. (2011). *History of Origami in the East and the West before Interfusion. Origami*, *5*, 3-11.

[2] Lang, R. J. (2007). *The science of origami*. Physics world, *20*(2), 30.

[3] Debnath, S., & Fei, L. J. (2013). *Origami theory and its applications: a literature review*. World academy of science, engineering and technology, 1131-1135.

[4] Ahmed, A. R., Gauntlett, O. C., & Camci-Unal, G. (2021). *Origami-inspired approaches for biomedical applications*. ACS omega, *6*(1), 46-54.

[5] Sargent, B., Butler, J., Seymour, K., Bailey, D., Jensen, B., Magleby, S., & Howell, L. (2020). *An origami-based medical support system to mitigate flexible shaft buckling*. Journal of Mechanisms and Robotics, *12*(4), 041005.

[6] Rus, D., & Tolley, M. T. (2018). *Design, fabrication and control of origami robots.* Nature Reviews Materials, *3*(6), 101-112.

[7] Bolaños Quiñones, V. A., Zhu, H., Solovev, A. A., Mei, Y., & Gracias, D. H. (2018). *Origami biosystems: 3D assembly methods for biomedical applications*. Advanced Biosystems, *2*(12), 1800230.

[8] Green, D. W., Watson, G. S., Watson, J., & Abraham, S. J. (2012). *New biomimetic directions in regenerative ophthalmology.* Advanced Healthcare Materials, *1*(2), 140-148.

[9] Loh, M., & Emami-Neyestanak, A. (2015). *Capacitive proximity communication with distributed alignment sensing for origami biomedical implants*. IEEE Journal of Solid-State Circuits, *50*(5), 1275-1286.

[10] Cheng, D. L., Greenberg, P. B., & Borton, D. A. (2017). *Advances in retinal prosthetic research: a systematic review of engineering and clinical characteristics of current prosthetic initiatives.* Current Eye Research, *42*(3), 334-347.

[11] Bareket, L., Barriga-Rivera, A., Zapf, M. P., Lovell, N. H., & Suaning, G. J. (2017). *Progress in artificial vision through suprachoroidal retinal implants*. Journal of Neural Engineering, *14*(4), 045002.

[12] Liu, Y., Park, J., Lang, R. J., Emami-Neyestanak, A., Pellegrino, S., Humayun, M. S., & Tai, Y. C. *(2013, June). Parylene* origami structure for intraocular implantation. In 2013 Transducers & Eurosensors XXVII: The 17th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS & EUROSENSORS XXVII) (pp. 1549-1552). IEEE.

[13] Ge, Y., Lalitharatne, T. D., & Nanayakkara, T. (2022). *Origami inspired design for capsule endoscope to retrograde using intestinal peristalsis.* IEEE Robotics and Automation Letters, *7*(2), 5429-5435.

[14] Alian, A., Zari, E., Wang, Z., Franco, E., Avery, J. P., Runciman, M., ... & Mylonas, G. (2023). *Current engineering developments for robotic systems in flexible endoscopy.* Techniques and Innovations in Gastrointestinal Endoscopy, *25*(1), 67-81.

[15] Ranzani, T., Russo, S., Schwab, F., Walsh, C. J., & Wood, R. J. (2017, May). Deployable stabilization mechanisms for endoscopic procedures. In *2017 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1125-1131). IEEE.

[16] Lang, R. J. (1996, May). *A computational algorithm for origami design*. In : Proceedings of the twelfth annual symposium on Computational geometry (pp. 98-105)

BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Prediction of wine quality using machine learning techniques

Erestina Gjeka,
University of Donja Gorica, Podgorica, Montenegro
erestina.gjeka@udg.edu.me

***Abstract* — Climate change has affected every sector of nature, especially healthcare in recent years. These changes have affected the vineyards but also the characteristics of the wine. In this research project, two natural factors were taken into account, temperature and annual precipitation. At times when machine learning had not yet been discovered, each process was very complicated and time-consuming. Therefore, machine learning is a very smart move to get fast and accurate results. Pearson correlation coefficient was used to come to a conclusion.**

***Keywords*:** Temperature · Precipitation · Wine · **Pearson correlation coefficient**.

1 INTRODUCTION

Problems due to climate change have led to debates on many topics, research, and even this research paper. Because of all this, there was a need to do various studies to determine the impact of temperature and precipitation on wine quality. This is very important because it gives us information on how to use temperature and precipitation to get the type of wine we want. This research paper will deal specifically with the analysis of temperature and precipitation as the main factors influencing the quality of wine. Two types of wine were considered: white and red, the two most commonly used types. Italy is known for its wine and the variety of grapes it produces; that is why the information for the research paper is based on this country; however, the same method of analysis can be used for any other country. In order to study the data on each wine in more detail and more precision, three hypotheses were made. Depending on these hypotheses, what truly affects the quality of wine can be seen. A detailed analysis was made for each of them.

– Hypothesis 1: The amount of alcohol in the wine depends on the temperature.

– Hypothesis 2: Precipitation affects the acidity of wine.

– Hypothesis 3: Precipitation affects the pH of the wine.

2 PEARSON CORRELATION COEFFICIENT

The Pearson correlation coefficient is often referred to only as the correlation coefficient. Below are its characteristics and how it can be used. The correlation coefficient allows clear monitoring of the relationships that the two variables have with each other. This coefficient gives results that show how strong the relationship between two variables is [1].

The formula used to calculate the Pearson correlation coefficient is:

$$r_{XY} = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} \quad (1)$$

r_{XY} = Pearson correlation coefficient

n = number of samples

x = first variable

y = second variable.

In statistics, the correlation coefficient is very important because, in addition to strength, it also gives the direction in which the two variables move. The direction can take two forms: positive or negative. Positive means that if one variable grows, so does another, and if one decreases, so does the other. If a negative result is obtained, it means that one variable grows while the other simultaneously decreases. It is necessary to be careful when choosing variables because it is not good to analyze two variables that have little to do with each other. That's because, even if we get results that show a high impact between two variables, we cannot say exactly that this is a completely correct solution if we have selected variables incorrectly. The choices made for the analysis to be further studied used logic that shows that temperatures and precipitation really affect wine production. However, this should be checked to determine the accuracy or eventual inaccuracy of the claim [2].

As noted earlier, the correlation coefficient gives us two pieces of information about the variables: the direction and strength of their connection. The direction has already been processed; we are left with the strength of the connection between the two variables. If:

– $r_{XY} = +1$; this is the perfect positive link

– $r_{XY} \geq 0.75$; this is a pretty strong positive link.

– $0.25 \leq r_{XY} \leq 0.75$, this is a moderate positive link.

If r_{XY} has the value greater than 0 and less than 0.25, then that connection is positive but weak. If $r_{XY} = 0$, it means that if one variable increases or decreases, the other variable

does not change. If r_{XY} has the value greater than −0.25 and less than 0, then that connection is negative but weak.

– r_{XY} = -1; this is the perfect negative link

– $r_{XY} \leq$ -0.75; this is a pretty strong negative link.

– -0.75 ≤ $r_{XY} \leq$ -0.25, this is a moderate negative link [1].

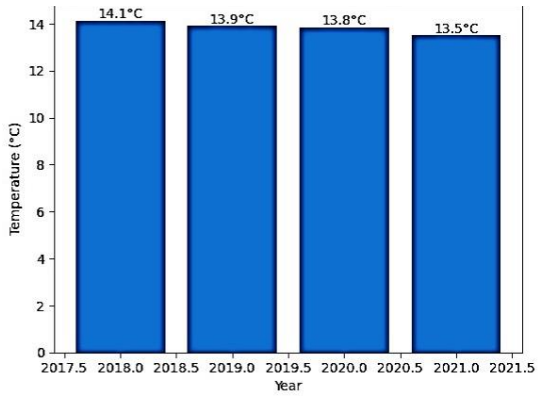
The correlation between the variables is displayed using graphics known as heat maps. Heat maps allow data to be visualized via matrices. It is very useful, especially in cases where links between multiple variables are sought, not just two. The matrix is created in such a way that it allows us to notice key data with one glance and to study the strength of the connections that the variables have with each other. Variables are displayed through columns and rows, while the numbers we get in the matrix are indicators of the strength of the relationship between each two variables [3].

3 ANALYSIS OF PRECIPITATION AND TEMPERATURE IMPACTS ON WINE CHARACTERISTICS

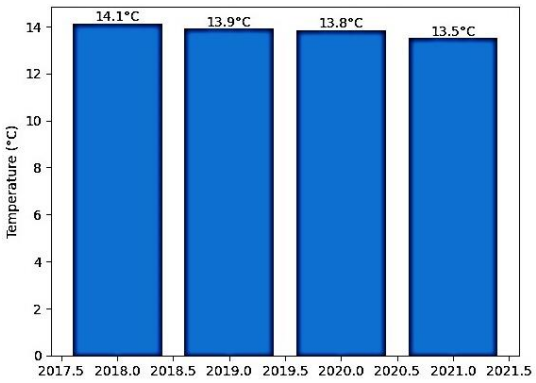
Data analysis is a process that must be carefully considered, both in data selection and in further work. During this analysis, we will not only get a link between the characteristics of wine and temperature or rain, but we will also get much more than that thanks to heat maps.

The country to which all the data relates is Italy, because it is one of the largest wine producers in the world. There are many wine-producing cities, but among the highest quality are cities like Sicily, Tuscany, Piedmont, and many others. Italy is a country where climatic conditions differ, which means that each zone in which grapes can be grown differs from another zone. Even within the same zone, there may be small variations in temperature and precipitation. Since Italy is not a small country, the average temperature and precipitation values for the whole country have been taken into account. The average wine characteristics in some famous wine-producing cities are also determined [4]. The aim of the analysis is to link the effects of climate characteristics on wine production in order to act on its production in the best possible way. Data were collected through online websites, not only for temperatures and precipitation but also for wine. All data were analyzed using the Pearson correlation coefficient and presented via different graphs. This analysis aims to conclude how much effect temperature and precipitation have on wine and what characteristics of wine are affected by them.

The initial step is to find the data. The rain data in Italy used in the project is from this website [5], and the temperature data were taken from the website [6]. Data collected and grouped is always desirable to be displayed using charts. The main goal is to compare the annual level of data from 2018 to 2021. Initially, precipitation data will be analyzed, and years will be identified with the largest and smallest amounts of precipitation. Once these results are obtained, the same process will be repeated for temperatures. The second step will be to present the characteristics of the wine. Two types of wine, red and white, were considered in this analysis. Graphs will allow the characteristics of white wine and red wine to be compared. The three main characteristics to be considered during this analysis are alcohol, acidity, and pH.



1. figure: Average precipitation value in Italy for the years 2018-2021.

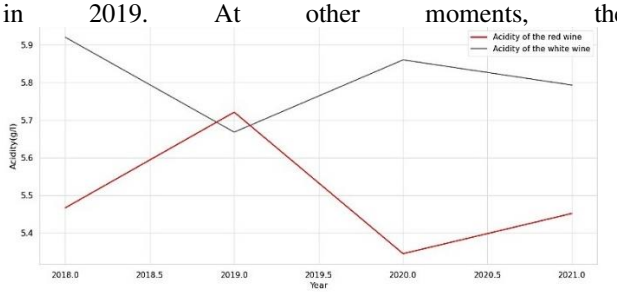


2. figure: Average temperature value in Italy for the years 2018-2021

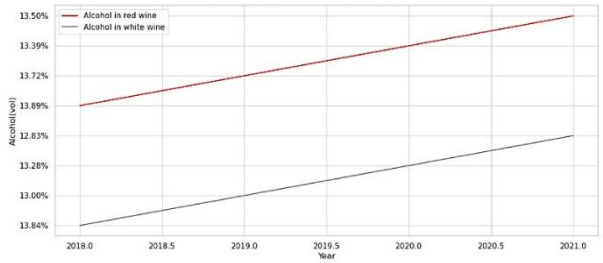
Figure 1 shows a precipitation graph for Italy over the past 4 years. Displaying data in this way allows information to be processed faster. The largest precipitation was recorded in 2019, with 890.8mm. This shows the average amount of precipitation in 2019, expressed in millimeters. On the other hand, it was noted that the precipitation in 2021 was lower compared to previous years, at 730.6mm. In Figure 2, the temperature table for Italy over the past four years is shown.

It can be observed that the highest temperature was recorded in 2018, at 14.1 °C. The lowest temperature compared to others was that of 2021, with 13.5 °C. Some graphs related only to the characteristics of the wine will be displayed. The three charts below will compare the properties of white and red wine.

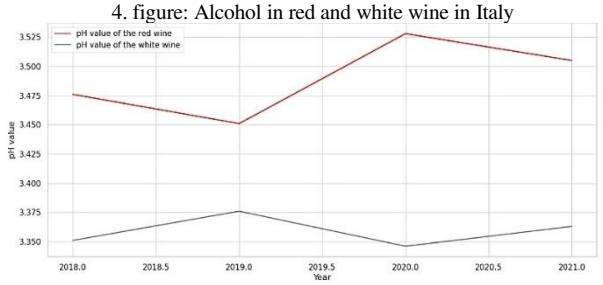
Looking at Figure 3, it can be seen that at one point the acidity of white and red wine almost coincided, and that was in 2019. At other moments, the



3. figure: Acidity in red and white wine in Italy



4. figure: Alcohol in red and white wine in Italy



5. figure: pH value of red and white wine in Italy

acidity values of red and white wine do not approach each other; on the contrary, the acidity of red wine is almost always lower than the acidity of white wine.

It can be said that, in general, for these years, the acidity of red wine is lower than that of white wine. The most interesting part of this graph is the difference between white and red wine, as it will be further used for analysis.

In Figure 4, which is shown below, looks as if there is a linear correlation between variables; when one grows or falls, the other does the same. However, a pattern isn't really shown, so no general statement can be drawn.

From Figure 5, it can be concluded that the pH value of red and white wine is never the same. Red wine always has a higher pH value than white wine.

The main graphics are shown with basic data for climate and wine characteristics. All the elements needed to form a Pearson correlation coefficient were found.

3.1 Pearson correlation coefficient calculated for precipitation and properties of wine

Table 1 shows the assessment using the Pearson correlation coefficient that uses formula (1). In the Figure 6, the code that has been used to form the Table 1 below is shown.

```
whitewine = pd.read_csv('/kaggle/input/white-wine-and-precipitation-in-italy-
whitewine.head()

redwine = pd.read_csv('/kaggle/input/red-wine-and-precipitation-in-italy-2021
redwine.head()

data_merged = pd.merge(whitewine, redwine, on="mm.")
data_merged.head()

corr_matrix = data_merged.corr(method="pearson")
corr_matrix.head(10)
```

6. figure: The code for calculating the Pearson correlation coefficient for precipitation and wine characteristics

	mm.	alcohol_x	acidity_x	pH_x	alcohol_y	acidity_y	pH_y
mm.	1.000	-0.316	-0.827	0.720	0.168	0.842	-0.654
alcohol_x	-0.316	1.000	0.760	-0.621	0.600	-0.281	-0.082
acidity_x	-0.827	0.760	1.000	-0.924	0.083	-0.812	0.518
pH_x	0.720	-0.621	-0.924	1.000	0.228	0.913	-0.730
alcohol_y	0.168	0.600	0.083	0.228	1.000	0.511	-0.808
acidity_y	0.842	-0.281	-0.812	0.913	0.511	1.000	-0.919
pH_y	-0.654	-0.082	0.518	-0.730	-0.808	-0.919	1.000

1. table: The results shown were calculated using the Pearson correlation coefficient for precipitation and wine characteristics.

Table 1 gives the results of the Pearson correlation coefficient for precipitation. Each variable that has a suffix x belongs to white wine, while those that have a suffix y belong to red wine. This table shows the result calculated by the Pearson correlation coefficient, which establishes a connection between every two elements. Not only by analyzing the relationship between temperature and precipitation with the characteristics of wine but also by analyzing the link between the properties of red and white wine.

Through the Pearson correlation coefficient, we can confirm or refute the hypotheses formed at the beginning of this analysis. As mentioned earlier, all results obtained are between -1 and 1, with -1 showing a perfectly negative link where one variable decreases while the other grows and 1 showing a perfect positive bond where one variable grows or decreases while the other also increases or decreases.

Data with values between -0.25 and 0.25 cannot provide reliable information that can be confirmed with certainty because the link between them is not strong enough, which means that their connection is weak. On the other hand, results above 0.25 and below -0.25, but especially those above 0.75 and below -0.75, can provide information with greater certainty as their connection is stronger.

The obtained data via the Pearson correlation coefficient will be displayed in the form of heat maps to make it easier to read. The heat map will display the results of the Pearson correlation coefficient for each type of wine separately. As seen in Table 1, there is also an analysis of the relationship between the characteristics of both wines, but this link is not significant for further analysis. Therefore, the results will be displayed separately in the white and red wine heat map.

"Seaborn" and "Matplotlib" have been two packages that have helped to earn the desired results (heat maps); using them, the coding has become easier and better visualized. On the other side, "font_scale" and "fig_size" can be used to arrange the photo how we desire it to be. In all of the heat maps, "font_scale" has been equal to 1.1, and "fig_size" has been equal to (8, 8).

```
import seaborn as sns
import matplotlib.pyplot as plt
corr_matrix1 = whitewine.corr(method="pearson")
print(corr_matrix1)

sns.set(font_scale=1.1)
plt.figure(figsize=(8, 8))
sns.heatmap(corr_matrix1, annot=True, cmap="coolwarm", square=True, vmin=-1, vmax=1)

plt.show()
```

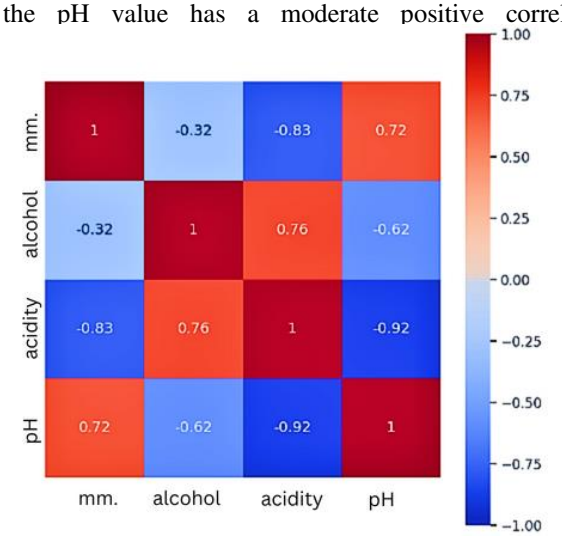
7. figure: The code that has been used to gain the heat map for precipitation and characteristics of the white wine

The first heat map is generated using the code shown in Figure 7. This heat map depicts the properties of white wine with precipitation.

The code shown in Figure 8, was used to generate a heat map of red wine properties and precipitation.

In Figure 8, a graphic display represents the Pearson correlation coefficient calculated only for white wine and precipitation. In the same way, a heat map is obtained for red wine which is shown in Figure 10. The analysis will be performed in a way that takes into account only data that has a strong link.

The heat map shown in Figure 8 shows that the following factors must be taken into account for the impact of precipitation on white wine properties: acidity and pH value. The acidity has a strong negative correlation, while the pH value has a moderate positive correlation.

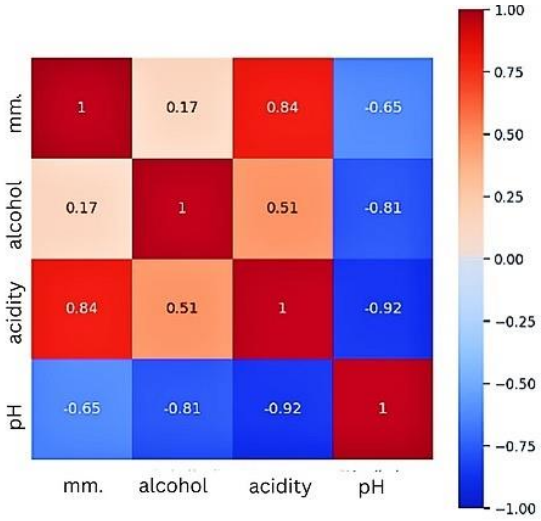


8. figure: Heat map for white wine characteristics and precipitation

```
import seaborn as sns
import matplotlib.pyplot as plt
corr_matrix2 = redwine.corr(method="pearson")
print(corr_matrix2)

sns.set(font_scale=1.1)
plt.figure(figsize=(8, 8))
sns.heatmap(corr_matrix2, annot=True, cmap="coolwarm", square=True, vmin=-1, vmax=1)
plt.show()
```

9. figure: The code that has been used to gain the heat map for precipitation and characteristics of the red wine



10. figure: Heat map for red wine characteristics and precipitation

3.1.1 Analysis of the impact of precipitation on the acidity of wine.

White wine has an acidity value of -0.83. This suggests that acidity and precipitation in wine are inversely proportionate. When one number rises, the other falls. When there is less rain, the acidity is higher; when there is more rain, the acidity is lower.

According to the book [7] the relationship between wine and rainfall (precipitation) may be reinforced. The quantity of acid in wine is frequently responsible for its sourness. Wine, as is known, is produced from grapes, and if the grapes are in their best condition, then the wine will also taste better. Grapes ripen at their own pace; nevertheless, a perfectly ripe grape equivalent to grapes picked the previous season can never be achieved, even if they are cut at the same time and location but in consecutive years. The amount of sour wine is influenced by the age of the grapes. According to the same book [7] "too much rain during flowering time in late spring/early summer can lead to poor fruit set and, later on, increased fungal diseases and low yields."

According to this book, apple acid levels rise in colder temperatures while falling in warmer ones. It is possible to define chilly weather situations as those that are inherently distinguished by rain. Of course, it should be kept in mind that not all grape varieties are affected negatively by this association. Because of the positive correlation between red wine acidity and precipitation, which means that when precipitation grows, acidity also grows, and when precipitation decreases, acidity also decreases, it is important to delve into greater depth. White wine, on the other hand, has the opposite impact.

Red wine and white wine are manufactured using distinct techniques, and typically, white wine is made from grapes that are stronger than those used to make red wine. Additionally, it uses grapes that are less developed than those used to make red wine. The graph in Figure 3 supports the assertion that white wine has higher acidity than red wine. The book [8] says: "Red grapes tend to have a longer growing season and thicker skins, which make them more resilient to changes in weather conditions... For white grapes, excessive rainfall during the ripening period can lead to a reduction in acidity and a loss of aromatic intensity. White grapes tend to ripen earlier than red grapes and have a shorter growing season, so they are more vulnerable to changes in weather conditions.", which supports the idea that has been validated. While the heat map in Figure 8 indicates that the ratio between acidity in white wine and precipitation is strong and negative, the heat map in the other Figure 10 indicates that the relationship between acidity in red wine and precipitation is strong and positive (value of +0.84). We may infer that ***hypothesis 2*** is verified as true based on these two manuals, the books [3] and [4], and based on the study done using Pearson's correlation coefficient.

3.1.2 Analysis of the impact of precipitation on the pH value of wine.

In Figures 8 and 10, it is shown that there is a moderate relationship between pH and precipitation. Figure 8 displays the findings of the Pearson's correlation coefficient, which demonstrates a positive relationship between precipitation and the pH of white wine. Precipitation and red wine's pH have a negative connection, just like previously. This implies that while one number

increases, the other decreases. A more thorough investigation is also necessary in this final instance, when white wine and red wine should be separated.

According to Article [9], it is stated that "Usually, if a wine has a high acid level, it will have a low pH.". Indirect analysis can be used to determine the link between pH and rainfall levels. It has been shown that acidity has an impact on pH value, and since pH is a phrase that indicates how much acidity is present in a substance, this enables an understanding of the relationship between these two terms. Additionally, precipitation in this instance is related to both acidity and pH. It was said that white wine's acidity and rainfall had an inverse relationship, meaning that as rainfall rises, acidity falls and vice versa.

The acidity of red wine, on the other hand, has a positive relationship with rainfall, which implies that these two values rise or fall simultaneously. As was already mentioned, there is a positive correlation between rainfall and white wine's pH whereas a negative correlation exists between red wine's pH and rainfall. Analyzing the relationship between pH and acidity for both wines is sufficient.

The acidity of white wine has a negative relationship with pH, as seen by the heat map in Figure 8. This means that when the pH value falls, acidity increases, which makes sense because the lower the pH value, the higher the acidity. In a prior analysis, it was determined that as the acidity of white wine increases, the levels of precipitation decrease since there is a negative correlation between the two. This leads to the conclusion that, when calculated based on acidity, pH value decreases when rainfall is reduced. This demonstrates once more that precipitation and pH value are positively correlated in white wine.

A red wine analysis will now be addressed. The acidity of red wine has a positive link with rainfall, as previously demonstrated, but this can also be observed in Figure 10 where the heat map is shown. This means that acidity and rainfall both rise and reduce at the same time. As a result, there is a direct link. Red wine, on the other hand, has a negative association with acidity. This demonstrates that as the pH value decreases, acidity increases in the same manner that it does in white wine, as this is a general concept that always applies. The examination of the Pearson correlation coefficient calculation in the heat map in Figure 10 leads to the conclusion that as precipitation falls, the pH value lowers. Because it has been indirectly proven that precipitation influences the pH of both wines, ***hypothesis 3*** is accepted.

3.2 Pirson correlation coefficient calculated for temperature and wine properties

The data in the figures were analyzed using the Pearson correlation coefficient, and the table is the result of that analysis with the aid of the following code:

```
whitewine = pd.read_csv('/kaggle/input/white-wine-and-temperature-in-italy-in-2
whitewine.head()

redwine = pd.read_csv('/kaggle/input/red-wine-and-temperature-in-italy-in-2021-
redwine.head()

data_merged = pd.merge(whitewine, redwine, on="celcius")
data_merged.head(15)

corr_matrix = data_merged.corr(method="pearson")
corr_matrix.head(15)
```

11. figure: The code for calculating the Pearson correlation coefficient for temperature and wine characteristics

2. table: The results shown were calculated using the Pearson correlation coefficient for temperature and wine characteristics.

	celcius	alcohol_x	acidity_x	pH_x	alcohol_y	acidity_y	pH_y
celcius	1.000	0.867	0.351	-0.267	0.702	0.149	-0.424
alcohol_x	0.867	1.000	0.760	-0.621	0.600	-0.281	-0.082
acidity_x	0.351	0.760	1.000	-0.924	0.083	-0.812	0.518
pH_x	-0.267	-0.621	-0.924	1.000	0.228	0.913	-0.730
alcohol_y	0.702	0.600	0.083	0.228	1.000	0.511	-0.808
acidity_y	0.149	-0.281	-0.812	0.913	0.511	1.000	-0.919
pH_y	-0.424	-0.082	0.518	-0.730	-0.808	-0.919	1.000

Through Table 2 , we can see which of the data has a positive, negative, strong, or weak relative. Data that have a continuation of x refer to white wine, while data that have a continuation of y refer to red wine.

The identical procedure that was used for precipitation will be used for temperature. Formula (1) has been used to evaluate data relationships. Table 2 shows Pearson's correlation coefficient for the temperature and properties of white and red wine in the same table. Through Table 2, it is possible to analyze in detail what correlation exists between temperature and white or red wine regarding their acidity, alcohol, or pH values.

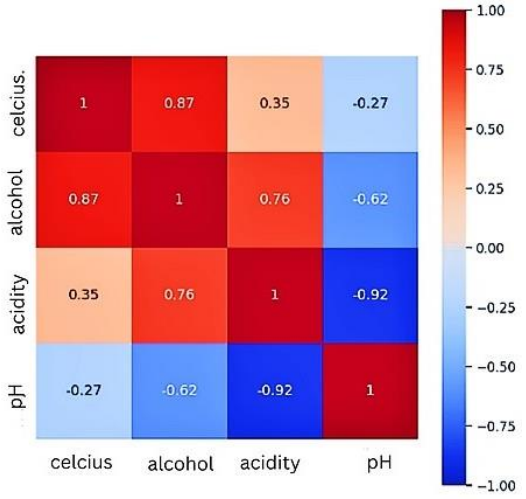
As highlighted in Table 1, we will divide this information into two correlation maps, where the Pearson correlation coefficient will be calculated, especially for red and white wine. As previously done when analyzing data on wine in relation to precipitation, only strong correlations are important. As has already been said, the data is easier to analyze using a heat map. The following are two heat maps where the published results were taken from Table 2. Using the coloring of the heat map, it is easily noticed which are strong and weaker distances.

The first heat map is generated using the code shown in Figure 12.

```
import seaborn as sns
import matplotlib.pyplot as plt
corr_matrix1 = whitewine.corr(method="pearson")
print(corr_matrix1)

sns.set(font_scale=1.1)
plt.figure(figsize=(8, 8))
sns.heatmap(corr_matrix1, annot=True, cmap="coolwarm", square=True, vmin=-1, vmax=1)
plt.show()
```

12. figure: The code that has been used to gain the heat map for temperature and characteristics of the white wine

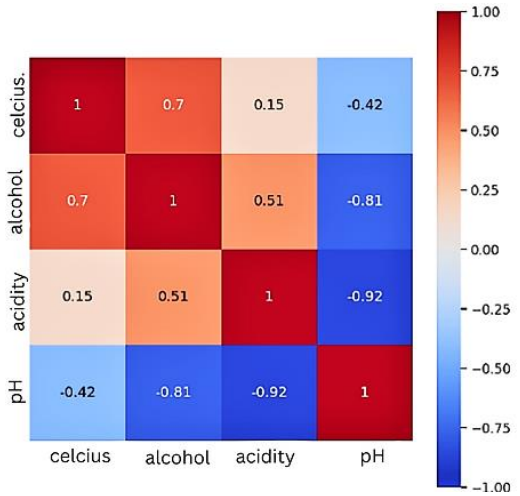


13. figure: Heat map for white wine characteristics and temperature

```
import seaborn as sns
import matplotlib.pyplot as plt
corr_matrix2 = redwine.corr(method="pearson")
print(corr_matrix2)

sns.set(font_scale=1.1)
plt.figure(figsize=(8, 8))
sns.heatmap(corr_matrix2, annot=True, cmap="coolwarm", square=True, vmin=-1, vmax=1)
plt.show()
```

14. figure: The code that has been used to gain the heat map for temperature and characteristics of the red wine



15. figure: Heat map for red wine characteristics and temperature

The same way as before, first the code will be shown, and then the output of that code, which will be the second heat map. The code in Figure 14 was used to generate a heat map of red wine properties and temperature.

In the first heat map shown in Figure 13, which represents a correlation map for Pearson's correlation coefficient between white wine characteristics and temperature, we see that only alcohol has a strong correlation with temperature. This means that for white wine, we will only analyze the relationship between temperature and alcohol. As for the second heat map in Figure 15, showing the correlation map for Pearson's correlation coefficient calculated for the relationship between the red wine and temperature properties, we can notice that only alcohol has a strong correlation with temperature. This means that we will only analyze that correlation here.

3.2.1 Analysis of the impact of temperature on alcohol in wine

The first heat map in Figure 13, displays a correlation map using Pearson's correlation coefficient between alcohol and temperature in white wine, which is 0.87. This suggests that in white wine, there is a positive relationship between alcohol and temperature. The quantity of alcohol in wine increases as the temperature rises. However, no mention is made of temperatures during fermentation, but rather of climatic temperatures during grape growing. However, because alcohol is temperature sensitive, further evidence is required to prove this association. This information is derived from a link [10], which states that temperature and climatic change affect wine and the level of alcohol in wine. This link says: "Warmer temperatures mean lower acidity and higher sugars in the grapes. Sugar converts to alcohol, so the end effect is a powerful acidity in the wine and higher alcohol levels." As a conclusion, wine is obtained with a higher amount of alcohol and a lower amount of acid. This proves that the temperature has an impact on the amount of alcohol in white wine.

Similarly, it has been established that temperature influences the alcohol content in red wine. The second heat map in Figure 15 displays the findings of the Pearson correlation coefficient between red wine qualities and temperature. Alcohol and temperature have a high reciprocal association, according to Pearson's correlation coefficient. This connection has a value of 0.7, indicating that alcohol and temperature have a significantly strong relationship.

Also, red wine as well as white wine have the same relationship in terms of alcohol and temperature. In the same way as with white wine, data were collected regarding this link [10], and based on Pearson's correlation coefficient, it can be concluded that temperature affects alcohol for red wine as well. As a conclusion, hypothesis 1 is confirmed as correct.

4 Conclusion

Many pieces of information from various books and websites were gathered during this analysis. In addition to the literature, the Pearson correlation coefficient and correlation maps were used, which showed the results of the coefficient. Temperature data were collected from the [5] and [6] websites, while wine data were collected from the website [11].

Initially, it was considered that the general notion that precipitation and temperature impact both white and red wine qualities had a foundation. The broad theory is then subdivided into three prosthetic hypothesis. According to this study, temperature and precipitation should be employed in favor of growing vines and creating the desired type of wine. Certain sites and methods of grape cultivation are essential.

The influence of climate change on vineyards should be considered every year and more, as climate change is becoming a severe issue. As a result, their influence on vineyards cannot be overlooked.

References

[1] Turney, Sh. (2020). Pearson Correlation Coefficient (r) | Guide and Examples.

[2] Frost, J. Interpreting Correlation Coefficients, Statistics by Jim.

[3] Kumar, A. (2022). Correlation Concepts, Matrix & Heatmap using Seaborn .

[4] MacNeil, K. (2015). The Wine Bible. 2nd edn. Workman Publishing, New York, pages: 34, 601-609.

[5] https://tradingeconomics.com/italy/precipitation (Accessed: 21.03.2023)

[6] https://tradingeconomics.com/italy/temperature (Accessed: 21.03.2023)

[7] Buglas J., A. (2022). An Introduction to Viticulture, Winemaking and Wine: From Vineyard to Wine Glass. Cambridge Scholars Publishing, UK, page 66.

[8] Robinson, J. (2015). The Oxford Companion to Wine, Oxford University Press.

[9] Hale, N. (2022). What is acidity in wine?, Wine Enthusiast. https://www.winemag.com/2019/06/19/what-is-acidity-in-wine/

[10] https://www.thirtyfifty.co.uk/spotlight-climate-change.asp#:~:text=Warmer\%20temperatures\%20mean\%20lower\%20acidity,feel\%20lighter\ %20in\%20the\%20mouth. (Accessed: 01.04.2023)

[11] https://www.superiore.de/ (Accessed: 20.03.2023)

BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Bevezetés az IoT világába: azon belül testek internete és kiberbiztonság

Introduction to the world of IoT: including the internet of Bodies and cyber security

Nagy Attila*, Horváth-Kiss Anikó**, Rajnai Zoltán***

Óbudai Egyetem, Budapest, Magyarország

*attila.nagy@uni-obuda.hu; **kiss.aniko@bgk.uni-obuda.hu; ***rajnai.zoltan@bgk.uni-obuda.hu

Összefoglalás **— A cikk fő célja a Dolgok Internetének bemutatása, különös tekintettel a Testek Internetére. Célja, hogy ráirányítsa a figyelmet a jövőben rejlő lehetőségekre ezen a területen, miközben felhívja a figyelmet a kiberbiztonsággal kapcsolatos kihívásokra. Az emberi testek internetkapcsolatának előretörése a társadalmak egészségesebbé válásához vezethet, ami miatt kiemelten fontos lesz a megfelelő védelem biztosítása.**

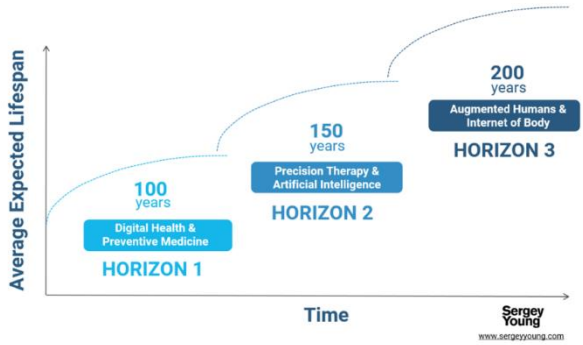
Kulcsszavak**:** IoT, IoB, sérülékenységek

Abstract — The main purpose of this article is to present the Internet of Things, with special regard to the Internet of Bodies. It aims to focus attention on future opportunities in this field while raising awareness of the challenges associated with cyber security. Advances in the Internet connection of human bodies can lead to healthier societies, which makes it extremely important to ensure adequate protection.

Keywords**:** IoT, IoB, vulnerabilities

1 BEVEZETÉS

A longevitás (a hosszú élet tudománya) egy viszonylag új kutatási terület. A longevitás területe forradalmi változásokat hozhat az emberi társadalmakban. Akár 100, 150 vagy 200 évig is élhet majd az ember. Persze ez még egy hosszú folyamat lesz az emberiségnek, de a lehetőségek kezdenek megnyílni. [1] A fejlődési szakaszok három horizontra vannak bontva. Az első horizont a digitális egészségügy és a megelőző gyógyászat technológiai fejlődésére épül, amelyek segítségével átlagosan 100 évig élhetünk. A második horizont a mesterséges intelligenciát (AI) és a precíziós terápiákat foglalja magában, amelyek potenciálisan 150 évre növelik élettartamunkat. A harmadik horizont pedig a kiterjesztett emberek és a testek internete (IoB) innovációinak és fejlesztéseinek köszönhető, amelyek potenciálisan 200 éves (vagy több!) átlagos élettartamot eredményezhetnek. Az (1. ábrán) lehet megnézni a különböző horizontokat. [2]



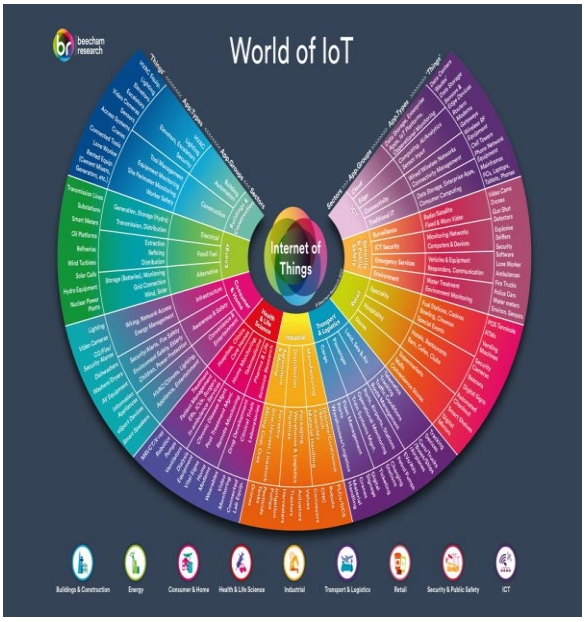
1. ábra: A hosszú élettartam innovációjának horizontja

Az ember életének meghosszabbítása és minőségének javítása több tényezőn múlik. Az egyik ilyen tényező a testek internete (IoB). Ebben a részben először is bemutatjuk a dolgok internetét (IoT), majd áttérünk a testek internetére, ismertetjük a hozzá tartozó eszközöket, végül pedig megvizsgáljuk a potenciális biztonsági kockázatokat. Ha az emberek élettartamának növelésében a Testek Internete kulcsszerepet játszik, akkor a biztonság kiemelt fontosságú lesz ezen a területen. Természetesen a kiberbiztonság minden szektorban létfontosságú.

2 A DOLGOK INTERNETE

Az „Internet of things” fogalmat, ha magyarra szeretnék fordítani akkor az a dolgok internete. Eme logikát követve,

akkor az „Internet of bodies” magyar fordításban a testek internete lesz. Az olvasóknak szeretnénk bemutatni, hogy mi a testek internete, és hogy milyen kiberbiztonsági veszélyek lehetségesek ezen a területen. Mielőtt bele vágnánk szeretnénk a dolgok internete területeit bemutatni, és hogy mi tartozik hozzá nagyvonalakban, a következő ábrán látható az IoT világának területitérképe, melyet 2008-ban vezettek be (2. ábra).



2. ábra: Az IoT területei

Kilenc kulcsfontosságú üzleti területet tartalmazz, melyek a következők: építés és építőipar, energia, fogyasztói és otthoni, egészség és élet tudomány, ipari, szállítás és logisztika, kiskereskedelem, biztonság és közbiztonság, információs és kommunikációs technológia (ICT). [3]

Építés és építőipar: Az IoT alkalmazása az építészetben és építőiparban lehetővé teszi az intelligens épületek és infrastruktúrák létrehozását, amelyek hatékonyabb energiafelhasználást, jobb fenntarthatóságot és magasabb biztonsági szintet biztosítanak. [4]

Energia: Az IoT technológiák segítségével az energiaipar hatékonyabbá válhat, lehetővé téve az okos mérők, energiahatékony rendszerek és a távoli energiakezelés bevezetését, ami csökkentheti az energiafelhasználást és a költségeket. [5]

Fogyasztói és otthoni: Az IoT megoldásokkal a fogyasztók számára intelligens otthonok és okos eszközök nyújtanak kényelmet, energiahatékonyságot és biztonságot, miközben lehetővé teszik az eszközök távoli vezérlését és felügyeletét. [6]

Egészség és élet tudomány: Az IoT az egészségügyben és élet tudomány területén lehetővé teszi az okos egészségügyi eszközök, távmonitorozás és egészségügyi adatelemzés fejlesztését, amelyek segítenek a betegségek korai felismerésében és a betegellátás javításában. [7]

Ipari: Az IoT alkalmazása az ipari területeken lehetővé teszi a gyártási folyamatok automatizálását, az üzemhatékonyság javítását és a gyártási adatok elemzését, ami növeli a termelékenységet és csökkenti a költségeket. [8]

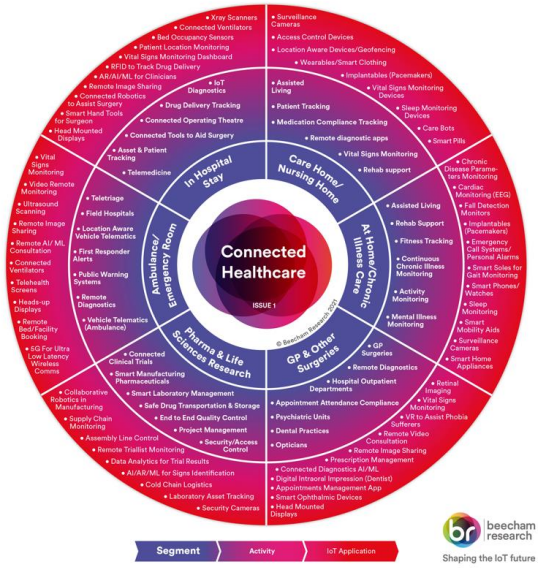
Szállítás és logisztika: Az IoT technológiák alkalmazása a szállításban és logisztikában lehetővé teszi az okos járművek, flottakezelés és raktározás fejlesztését, ami javítja az áruk követését, a szállítási hatékonyságot és a logisztikai folyamatokat. [9]

Kiskereskedelem: Az IoT megoldások alkalmazása a kiskereskedelmi területen lehetővé teszi az okos üzletek, az élmény alapú kiskereskedelem és a vásárlói analitika fejlesztését, ami növeli a vásárlói elköteleződést és javítja az élményt. [10]

Biztonság és közbiztonság: Az IoT technológiák alkalmazása a biztonság és közbiztonság területén lehetővé teszi a térfigyelő rendszerek, okos városok és vészhelyzeti felügyelet fejlesztését, ami javítja a bűn- és balesetmegelőzést, valamint az incidenskezelést. [11]

Információs és kommunikációs technológia (ICT): Az IoT integrálása az információs és kommunikációs technológiák területén lehetővé teszi az okos hálózatok, a távoli elérés és a felhőalapú szolgáltatások kialakítását, ami elősegíti az adatok hatékonyabb kezelését és az információk könnyebb elérését. Ez a fejlesztés lehetővé teszi az üzleti folyamatok automatizálását, az adatelemzési és döntéstámogató rendszerek kialakítását, valamint az ügyfélkapcsolatok és szolgáltatások optimalizálását. Az IoT és az ICT összekapcsolása a digitális átalakulás egyik kulcsfontosságú tényezője, ami új üzleti lehetőségeket teremt és javítja a vállalatok hatékonyságát és versenyképességét. [12]

Az összekapcsolt egészségügyhöz is nagyon sok eszköz tartozik és terület melyet a következő ábrán lehet megnézni (ábra 2.)



3. ábra: A hat kulcsfontosságú területet

Az 3. ábrán hat kulcsfontosságú területet emelnek ki, melyek a következők: kórházi tartózkodás, gondozás otthon és idősek otthona, otthon és krónikus betegség ellátás, háziorvosi rendelők és más műtétek, gyógyszergyártás és élettudományok kutatás, mentők és sürgősségi szoba. [13] A testek internete a krónikus betegség ellátás területéhez tartozik. Ez a terület olyan eszközöket és technológiákat foglal magában, amelyek lehetővé teszik az egészségügyi adatok gyűjtését és elemzését az otthoni környezetben élő emberek számára, például okos eszközökkel és hordozható egészségügyi eszközökkel.

Minden egyes terület megérdemelne egy saját fejezetet, de mi most a testek internetéről fogunk foglalkozni.

3 TESTEK INTERNETE

A "Testek Internete" kifejezést 2016-ban alkották meg, és ez egy viszonylag új területet jelöl. Ezek az eszközök az emberi testet figyelik biometrikus, fiziológiai vagy viselkedési adatok gyűjtése révén. Az információkat vezeték nélküli vagy hibrid hálózatokon keresztül továbbítják más eszközök felé. Ezután egy központi számítógép elemzi és értékeli ki ezeket az adatokat. [14] Az adatok kiértékelése az adatbányászat segítségével történik így állapítható meg az ember egészségi állapota.

3.1 Eszközök

Okos órák és fitneszkarkötők (Smart watches and fitness trackers)

A jelenleg legnépszerűbb és széles körben elterjedt IoB (Testek Internete) eszközök közé tartoznak az okos órák és fitneszkarkötők. Ezek az eszközök mozgás-, szívritmus- (beleértve az eltérések detektálását), alvás- és az újabb készülékekben a véroxigénszint-monitorozásra képes érzékelőket használnak. Emellett széleskörű egészségügyi és fitnesz támogatást nyújtanak, például képesek figyelemmel kísérni az esetleges eleséseket vagy a szokatlan mozgásmintákat, amelyek görcsrohamot jelezhetnek, és ennek megfelelően értesíthetik a családtagokat és a gondozókat. [15]

"Okos" vagy "digitális" pirulák (’Smart’ or ’digital’ pills)

Az úgynevezett okos pirulák olyan gyógyszerformák, amelyek egy lenyelhető érzékelőt kombinálnak a hagyományos gyógyszerhatóanyaggal. Ezek az érzékelők képesek rögzíteni a bevétel időpontját, az adagot és a gyógyszer típusát, továbbá információt szolgáltatnak a páciens aktivitási szintjéről. Jelenleg is használatosak többek között kemoterápia, valamint bipoláris zavar és skizofrénia kezelésére. A jövőbeni fejlesztések lehetővé tehetik, hogy ezek a pirulák nyomon kövessék a belső állapotokat és a gyógyszerekre adott reakciókat, sőt, automatikusan szabályozzák az adagolást a beteg válaszreakciói alapján. [15]

Okos kontaktlencsék (Smart contact lenses)

Ezek a lencsék lehetővé teszik a vér biomarkereinek, például a glükóz, a koleszterin, a nátrium- és káliumionok monitorozását és értékelését egy nem invazív módon, csupán a szemfolyadék felhasználásával. A jövőben a kiterjesztett valóság alkalmazásával készült kontaktlencsék is megjelenhetnek, amelyek további információkat jelenítenek meg a felhasználó látóterében, vagy akár rögzíthetik is azt, amit az emberek látnak. [15]

Agy-gép interfészek (Brain-computer interfaces)

Az agyi implantátumokat már használják arra, hogy a súlyosan bénult emberek képesek legyenek robotkarokat irányítani a saját táplálkozásukhoz, és közvetlenül szöveget generáljanak (90 karakter per perc sebességgel) az agyi jeleik felhasználásával, nem pedig a neuromuszkuláris útvonalakon keresztül. A nem invazív fejfedők is alkalmazhatók a robotkarok egyre összetettebb irányítására, habár itt a technológiai fejlődés lassabb. Ezen technológiák még nem kerültek kereskedelmi forgalomba. [15]

Mesterséges hasnyálmirigyek (Artificial pancreases)

A mesterséges hasnyálmirigy a folyamatos glükózmonitorozást (CGM) és az inzulinpumpát ötvözi, mesterséges intelligenciát alkalmazva az inzulinadagolás automatizálására a CGM által mért adatok alapján. [15]

Szívritmus-szabályozók (Cardiac pacemakers)

Az összekapcsolt szívritmus-szabályozók képesek valós idejű információkat biztosítani a páciens szívritmus-változásairól, és lehetővé teszik a távoli kezelést, mint például a szívritmus-érzékelési vagy szabályozási küszöbök beállítását. Az újonnan fejlesztett szívritmus-szabályozók képesek a testben történő biológiai lebomlásra, amikor már nincs rájuk szükség, és együttműködve más érzékelőkkel hatékonyabban érzékelik az eltéréseket. [15]

Összekapcsolt ruházat (Connected clothing)

A ruházatba beágyazott érzékelőkkel rendelkező ruhadarabok képesek a szívritmus és a mozgás figyelésére. Egyes darabok monitorozzák a testhőmérsékletet is, és képesek alkalmazkodni, hogy a viselőjük kényelmesen érezze magát. Az okospelenkák újítása, hogy képesek érzékelni és jelenteni az újszülöttek bélműködését. [15]

Érzékelőkkel felszerelt kórházi ágyak (Sensor-equipped hospital beds)

Az érzékelőkkel felszerelt kórházi ágyak olyan beépített szenzorokkal rendelkeznek, amelyek képesek a testhőmérséklet, szívverés, véroxigén-szint, vérnyomás és egyéb élettani adatok folyamatos monitorozására. Ez lehetővé teszi az egészségügyi személyzet számára, hogy valós időben nyomon kövessék a betegek alapvető élettani paramétereit. [15]

Figyelő monitorok (Attention monitors)

A szemkövetés olyan technológia, amely meghatározza, hogy egy személy hova néz, legyen szó számítógéphez rögzített eszközről irodai környezetben vagy szemüvegről vezetés közben és mindennapi helyzetekben. Néhány fejlettebb prototípus képes az agyi tevékenység észlelésére és a szemmozgások elemzésére is, így valós időben képesek különböző kognitív folyamatokat, mint a kognitív terhelés, a fáradtság, az elkötelezettség és a figyelem monitorozására. Ezek az eszközök hang- vagy tapintási visszajelzést is adhatnak, például, ha egy felhasználó figyelmetlenné válik. [15]

Testbe ültetett érzékelők (Body-implanted sensors)

A bőr alá ültethető bioszenzorok lehetővé teszik a biológiai folyamatok pontosabb és részletesebb nyomon követését a hagyományos viselhető eszközöknél. Ezek az eszközök továbbfejlesztett funkciókkal is rendelkeznek, mint például egy bőrbe integrált interfész, amely lehetővé teszi a felhasználó számára, hogy távolról irányítsa más eszközöket. További fejlesztés alatt álló érzékelők, amelyeket a fogakhoz rögzítenek, képesek monitorozni a felhasználó által bevitt glükóz, só és alkohol mennyiségét. [15]

Női technológiai termékek (Female technology products)

Egyre több technológiai terméket fejlesztenek ki kifejezetten a nők egészségének és jóllétének elősegítésére; ezeket gyakran "femtech" termékeknek nevezik. A Testek Internete kategóriájába tartozó eszközök között megtalálhatók a hordozható mellszívók, a medencefenék-erősítő eszközök, hűsítő hatású karkötők, amelyek enyhítik a hőhullámokat, valamint olyan csatlakoztatott készülékek, amelyek képesek a méhnyak nyálka monitorozására a termékenység nyomon követése érdekében.

Beültethető mikrochipek (Implantable microchips)

Az RFID és NFC (Near field communication) mikrochipek az emberi testbe ültethetők információk tárolására, mint például a személyes adatok vagy az ajtók nyitására és vásárlások lebonyolítására.

Érzelmi szenzorok (Emotion sensors)

Az érzelmeket érzékelő eszközök, amelyek még fejlesztés alatt állnak, képesek azonosítani a felhasználó érzelmi állapotát az arc kifejezése, mikromozgások, testtartás, gesztusok, agyi és szívtevékenység, bőrvezetőképesség és egyéb jelek alapján. Ezek az adatok felhasználhatók olyan környezeti változtatások elérésére, amelyek javítják a hangulatot.

Látás- és hallássegítő eszközök (Vision and hearing aids)

Számos viselhető eszköz és implantátum áll rendelkezésre, beleértve a beépített kamerával rendelkező mesterséges lencséket és cochleáris eszközöket, amelyeket az érzékelés helyreállítására vagy fokozására használhatnak. Ezek az eszközök lehetővé teszik videó- és hangfelvételek készítését, valamint képesek azonosítani, ha a felhasználó elesett, együtt más viselkedési jelzőkkel.

Wellness-szkennelő alkalmazások (Wellness scanning apps)

Egyes vállalatok olyan technológiákat fejlesztenek, amelyek lehetővé teszik az egészségi állapot mérését invazív beavatkozások és különleges eszközök nélkül. Egy ilyen alkalmazás képes a hagyományos okostelefon-kamerával készített, 30 másodperces arcvideó elemzésével becsülni a pulzusszámot, a stressz szintet és további egészségügyi jellemzőket, majd ezek alapján egy átfogó "wellness" pontszámot adni.

Olfaktorikus érzékelők (Olfactory sensors)

Ezeket az érzékelőket okosfogkefékbe építhetik be vagy önálló modulként alkalmazhatják, hogy gyűjtsék a felhasználó leheletéből származó apró, a biológiai aktivitáshoz vagy betegségekhez kapcsolódó anyagmennyiségeket.

Bőrre felhelyezhető érzékelők („ESkin") (Sking applied sensors)

Rugalmas fóliaérzékelők, amelyeket tapasz formájában helyezhetünk a bőrre, és amelyek képesek jeleket rögzíteni, mint például a szívverés — ami a fóliát rezgésbe hozza — és az izzadságszint, amelyre a fólia sóval érintkezve reagál.

Hordozható agyi szenzorok (Wearable brain sensors)

Olyan fejhallgatók, amelyek elektromos agyi aktivitást mérnek a fejbőrön elhelyezett elektródák segítségével. Ezáltal képesek például a koncentráció szintjének és a fáradtságnak a megállapítására. [15]

A könnyebb átláthatóság végett egy táblázatba szedtük a különböző IoB eszközöket.

1. táblázat: IoB eszközök

IoB eszközök
Okos órák és fitneszkarkötők
"Okos" vagy "digitális" pirulák
Okos kontaktlencsék
Agy-gép interfészek
Mesterséges hasnyálmirigyek
Szívritmus-szabályozók
Kapcsolt ruházat
Érzékelőkkel felszerelt kórházi ágyak
Figyelő monitorok
Testbe ültetett érzékelők
Női technológiai termékek
Beültethető mikrochipek
Érzelmi szenzorok
Látás- és hallássegítő eszközök
Wellness-szkennelő alkalmazások
Olfaktorikus érzékelők
Bőrre felhelyezhető érzékelők
Hordozható agyi szenzorok

4 SÉRÜLÉKENYSÉGEK

4.1 *OWASP top 10 IoT*

Gyenge, kitalálható vagy beépített jelszavak (Weak guessable, or hardcoded passwords)

Az IoT eszközök gyakran rendelkeznek webalapú felületekkel, amelyeket konfigurációra és kezelésre használnak, ezek mellett hitelesítési mechanizmusok is találhatók az eszközökben, mint például soros konzolok, hálózati szolgáltatások stb. Ha ezeket a felületeket nem megfelelően állítják be, a támadók hozzáférhetnek érzékeny információkhoz és engedély nélkül módosíthatják az eszköz beállításait. A SISA IoT biztonsági tesztelése során kiderült, hogy a tesztelt IoT eszközök többségénél kitalálható jelszavak és felhasználónév lista volt használatban. Egy másik kritikus hiba a rögzített jelszavak beépítése, ahol a fejlesztők beprogramozott hitelesítő adatokat helyeznek az IoT eszközök komponenseibe, például a firmware-be.

A támadás enyhítése:

A gyártóknak megfelelő hitelesítési és jelszókezelési kontrollokat kellene bevezetniük annak biztosítása érdekében, hogy a jelszavak biztonságosak és nehezen kitalálhatóak legyenek. Továbbá a felhasználókat arra kellene ösztönözni, hogy változtassák meg az alapértelmezett jelszavakat az eszközeiken, és az eszközök beállításakor használjanak erős, egyedi és összetett jelszavakat. [16],[17]

Nem biztonságos hálózati szolgáltatások (Insecure network services)

A nem biztonságos hálózati szolgáltatások a hálózati protokollok, szolgáltatások vagy konfigurációk sebezhetőségeire utalnak, és általában magukban foglalják a nem titkosított kommunikációs protokollokat, a gyenge hálózati biztonsági beállításokat, valamint az elavult vagy sebezhető szoftverek használatát. A támadók ezeket a sebezhetőségeket kihasználva lophatnak érzékeny adatokat, indíthatnak támadásokat más rendszerek ellen, vagy jogosulatlanul férhetnek hozzá az eszközhöz.

A támadás enyhítése:

A biztonságos hálózati protokollok, például a Transport Layer Security (TLS) alkalmazása, valamint a hálózati szolgáltatások rendszeres frissítése segíthet enyhíteni ezt a sebezhetőséget. A SISA azt is javasolja, hogy rendszeresen végezzenek hálózati sebezhetőségi értékelést és vörös csapat gyakorlatokat az IoT hálózatokban lévő kritikus biztonsági hibák azonosítása érdekében. [16],[17]

Nem biztonságos ökoszisztéma interfészek (Insecure ecosystem interfaces)

Ez a sebezhetőség az IoT ökoszisztéma különböző komponensei közötti nem biztonságos interfészekből ered. Sok IoT eszköz gyengén védett interfészekkel rendelkezik (web, API, mobil interfészek) külső rendszerekkel, mint például felhőszolgáltatások, más IoT eszközök és hagyományos IT rendszerek. A támadók ezeket az interfészeket használhatják érzékeny adatok elérésére, támadások indítására más rendszerek ellen, vagy az eszköz és annak funkcióinak irányítására. A SISA IoT biztonsági tesztelő csapata egy piacvezető IoT eszköz biztonságának értékelése során találkozott egy olyan API-val, amelynek segítségével minden felhasználó UUID-ját generálhattuk, és felhasználhattuk annak élő helyzetének, jelszavaknak, az alkalmazáshoz csatlakoztatott egyéb eszközeinek, e-mail címének stb. megszerzésére. Annak ellenére, hogy az eszközalkalmazásnak 300 ezer+ letöltése van iOS-en és Androidon, és a gyártónak több mint 100 éves piaci múltja

van, ez a sebezhetőség a hatékony biztonsági intézkedések hiányára utal.

A támadás enyhítése:

Az API-k gyakori frissítése, szigorú hozzáférés-szabályozás alkalmazása a bizalmas API-khoz és interfészekhez való hozzáférés korlátozására, biztonságos kommunikációs csatornák megvalósítása az IoT ökoszisztéma különböző komponensei között, valamint titkosítás alkalmazása ajánlott intézkedések ennek a sebezhetőségnek a mérséklésére. [16],[17]

Biztonságos frissítési mechanizmus hiánya (Lack of secure update mechanism)

Az IoT eszközök gyakran úgy vannak tervezve, hogy alacsony költségűek, kis energiafogyasztásúak és könnyen használhatóak legyenek, ami azt eredményezheti, hogy a biztonsági szempontokat figyelmen kívül hagyják a tervezési folyamat során. Különösen a biztonságos frissítési mechanizmus hiánya teszi az IoT eszközöket sebezhetővé az ismert biztonsági résekkel és kihasználási lehetőségekkel szemben. A támadók kihasználhatják az elavult firmware-t vagy szoftvert az eszköz biztonságának kompromittálására. Az IoT fizetési rendszerekben ez a sebezhetőség komoly következményekkel járhat, beleértve a pénzügyi veszteségeket, a bizalmas információkhoz való jogosulatlan hozzáférését, és a kritikus rendszerek működésének zavarását.

A támadás enyhítése:

Az ilyen funkciók bevezetése, mint a digitális aláírások, a visszaállítás elleni mechanizmusok, a biztonságos szállítás (a frissítés titkosított formában történő küldése, a frissítés aláírása stb.), valamint a firmware hitelesítése az eszközön segíthet a gyártóknak kezelni ezt a sebezhetőséget. [16],[17]

Nem biztonságos vagy elavult komponensek használata (Use of insecure or outdated components)

Az IoT eszközökben használt nem biztonságos vagy elavult komponensek alkalmazása egyre növekvő aggodalomra ad okot a technológia világában. Sok IoT eszköz harmadik féltől származó komponensek felhasználásával készül, amelyek tartalmazhatnak sebezhetőségeket, amiket a támadók kihasználhatnak az eszköz biztonságának kompromittálására.

A támadás enyhítése:

Az IoT eszközökben használt összes szoftver és komponens (beleértve a firmware-t, könyvtárakat és keretrendszereket) rendszeres frissítése és javítása, valamint egy folyamat létrehozása a komponensekben lévő biztonsági sebezhetőségekről szóló értesítések figyelemmel kísérésére és fogadására az IoT ökoszisztémában, a sebezhetőség mérséklésének néhány legjobb gyakorlata. [16],[17]

Elégtelen adatvédelem (Insufficient privacy protection)

Sok IoT eszköz gyűjt és tárol érzékeny személyes adatokat, azonban gyakran hiányoznak a megfelelő adatvédelmi és adatbiztonsági intézkedések. Ez magában foglalhatja az adatgyűjtést a felhasználó beleegyezése nélkül, az adatok biztonsági ellenőrzések nélküli tárolását, valamint az adatok megfelelő engedélyek nélküli harmadik felekkel történő megosztását.

A támadás enyhítése:

Az adatvédelmi elvek tervezésbe történő beépítése, a titkosítás alkalmazása az érzékeny adatok továbbítása és tárolása során, valamint a felhasználók beleegyezésének megszerzése az adatgyűjtéshez és felhasználáshoz azok az effektív mérséklő intézkedések, amelyek alkalmazhatók. [16],[17]

Nem biztonságos adatátvitel és tárolás (Insecure data transfer and storage)

Az IoT eszközök esetében komoly aggodalomra ad okot az adatok titkosítás nélküli, nyílt szöveges átvitele és tárolása. Az IoT eszközök nagy mennyiségű személyes és érzékeny információt gyűjtenek és tárolnak, és a támadók közbeékelődhetnek az adatátvitelbe vagy manipulálhatják az adatokat azok továbbítása során, illetve kihasználhatják a gyenge tárolási mechanizmusokat.

A támadás enyhítése:

Biztonságos protokollok, mint például az HTTPS használata az adatátvitelhez, az érzékeny adatok nyugalmi állapotban történő titkosítása, szilárd hozzáférés-szabályozási rendszerek kialakítása, valamint az adattárolási gyakorlatok rendszeres auditálása hatékony intézkedések az IoT eszközök adatátvitelének és tárolásának biztosítására. [16],[17]

Eszközkezelés hiánya (Lack of device management)

Az IoT eszközök hatékony kezelésének hiánya veszélyeztetheti az egész hálózatot. Az effektív eszközkezelés hiánya lehetővé teszi a támadók számára, hogy távolról manipulálják vagy irányítsák az IoT eszközöket. A nem megfelelő kezelés jogosulatlan hozzáférést, firmware manipulációt vagy az eszközök módosítását eredményezheti. A SISA IoT eszköztesztelési értékelései során több esetben is kiderült, hogy az eszközök lejárt SSL tanúsítványokkal rendelkeztek, így a webes kommunikáció HTTP-n keresztül történt. Mivel az eszköz nem biztosított frissítéseket, az SSL tanúsítványok nem lettek megújítva, ami sebezhetővé tette az eszközt.

A támadás enyhítése:

Erős hitelesítési mechanizmusok, mint például egyedi eszközhitelesítő adatok bevezetése, és a hozzáférés-szabályozások érvényesítése, hogy az eszközkezelési funkciók csak a jogosult személyzet számára legyenek elérhetőek, csökkentheti ezt a kockázatot. [16],[17]

Nem biztonságos alapértelmezett beállítások (Insecure default settings)

Az IoT eszközökön azok a konfigurációk, amelyeket a gyártó változatlanul hagy, potenciális biztonsági kockázatoknak tehetik ki az eszközt. Ezek a beállítások magukban foglalhatják az alapértelmezett felhasználóneveket és jelszavakat, nyitott portokat és a nem titkosított kommunikációt. Gyakran az alapértelmezett beállítások a "minimális" megközelítést képviselik, vagy akár bevezethetnek IoT biztonsági sebezhetőségeket, például beépített jelszavakat vagy root jogosultságokkal futó kitett szolgáltatásokat.

A támadás enyhítése:

Az alapértelmezett felhasználónevek, jelszavak és konfigurációk megváltoztatása az eszközök első beállítása során, valamint a szükségtelen szolgáltatások és portok letiltása a támadási felület csökkentése érdekében olyan intézkedések, amelyek mérsékelhetik ezt a sebezhetőséget. [16],[17]

Fizikai védelem hiánya az IoT rendszerekben (Lack of physical hardening)

Az IoT rendszerek fizikai védelem hiányára utal, ha nem valósítanak meg fizikai biztonsági intézkedéseket. Ez teszi a beágyazott eszközöket sebezhetővé különféle hardveres támadásokkal és firmware manipulációval szemben, így engedélyezve a hackerek számára az olyan jogosulatlan hozzáféréseket, mint a root soros bejelentkezés, érzékeny információk kinyerése stb., amelyek távoli támadások végrehajtására vagy az eszköz feletti irányítás megszerzésére használhatók.

A támadás enyhítése:

Néhány intézkedés, amelyet meg lehet tenni az eszköz fizikai megerősítésére, beleértve a hibakereső portok letiltását vagy elszigetelését, a biztonságos indítás használatát a firmware érvényesítésére, a manipuláció-észlelési mechanizmusok alkalmazását, valamint az érzékeny információk eltávolítható memóriakártyán való tárolásának elkerülését. [16],[17]

5 ÖSSZEGZÉS

A kutatási terület, amit "longevitás" néven ismerünk, a hosszú élet tudománya, és viszonylag újnak számít. Ez a tudományág jelentős változásokat hozhat az emberi társadalmakban, például elősegítheti, hogy az emberek akár 100, 150, vagy akár 200 éves korukig is éljenek. Ez egy hosszú folyamat lesz az emberiség számára, de a lehetőségek már kezdenek körvonalazódni. A longevitás fejlesztési szakaszait három "horizont" jellemzi: az első a digitális egészségügy és a megelőző gyógyászat technológiai fejlődésére épül, a második horizont a mesterséges intelligenciát (AI) és a precíziós terápiákat foglalja magában, míg a harmadik a kiterjesztett emberek és a testek internete (IoB) innovációira támaszkodik. Ezek a fejlesztések potenciálisan lehetővé teszik az emberek számára, hogy jelentősen meghosszabbított életet éljenek.

A "testek internete" (IoB) olyan technológiát jelent, amely az emberi testen viselhető vagy beültetett eszközökkel gyűjt adatokat, mint például biometrikus, fiziológiai vagy viselkedési információkat. Ezek az adatok vezeték nélküli vagy hibrid hálózatokon keresztül továbbíthatók más eszközökre, ahol központi számítógépek elemezhetik és értékelhetik ki őket. Az IoB eszközök jelentőségét nem csak a hosszabb élettartam elősegítésében, hanem a mindennapi egészségügyi és életminőségi javulásokban is kiemelkedőnek ígérkezik.

A "dolgok internete" (IoT) szintén fontos szerepet játszik, különböző üzleti területeken belül, mint például az építőipar, az energiaipar, a fogyasztói és otthoni eszközök, az egészségügy és élettudományok, az ipar, a szállítás és logisztika, a kiskereskedelem, a biztonság és közbiztonság, valamint az információs és kommunikációs technológiák. Ezek a területek hozzájárulnak az okos infrastruktúrák létrehozásához, az energiafelhasználás hatékonyságának javításához, az intelligens egészségügyi eszközök fejlesztéséhez, és számos más előnnyel szolgálnak mind a magánszemélyek, mind a társadalom számára.

A longevitás és az IoB további kutatása és fejlesztése kulcsfontosságú lesz az emberiség jövője szempontjából, hiszen ezek a technológiák nem csupán az élet hosszát képesek növelni, hanem a minőségét is jelentősen javítani. Azonban fontos, hogy ezen technológiák fejlődése mellett a kiberbiztonsági kihívásokra is nagy figyelmet fordítsunk, hiszen az adatvédelem és a biztonságos használat garantálása nélkülözhetetlen a felhasználók bizalmának megőrzéséhez.

IRODALOMJEGYZÉK

[1] S. Young, *A fiatalodás tudománya*. 2023.
[2] S. Young, "Three Horizons of Longevity Innovation." [Online]. Available: https://sergeyyoung.com/three-horizons-of-longevity-innovation, 2020
[3] "World of IoT sector map." [Online]. Available: https://www.beechamresearch.com/download-details/world-of-iot-sector-map/, 2024
[4] "IoT in construction: Top benefits, Use-case and application." [Online]. Available: https://toolsense.io/equipment-management/iot-in-construction-top-benefits-use-cases-application/#:~:text=IoT%20has%20diverse%20applications%20in,on%20time%20and%20within%20budget.
[5] M. M. Alenazi, "IoT and Energy," in *Internet of Things - New Insights*, M. K. Habib, Ed., IntechOpen, 2024. doi: 10.5772/intechopen.113173.
[6] "Everything you need to know about consumer IoT (CIoT)." [Online]. Available: https://www.hitechnectar.com/blogs/consumer-iot-ciot/, 2023
[7] "How Internet of things (IoT) is impacting life sciences and healthcare industry." [Online]. Available: https://www.42gears.com/white-papers/how-internet-of-things-is-impacting-life-sciences-healthcare-industry/, 2023
[8] "What is industrial IoT (IIoT)?" [Online]. Available: https://www.cisco.com/c/en/us/solutions/internet-of-things/what-is-industrial-iot.html
[9] "IoT in transportation and logistics - The ultimate guide." [Online]. Available: https://www.teletracnavman.com/fleet-management-software/telematics/resources/iot-in-transportation-logistics-the-ultimate-guide, 2024
[10] "How IoT in retail is changing the global retail industry." [Online]. Available: https://onomondo.com/blog/how-iot-in-retail-is-changing-the-global-retail-industry/, 2023
[11] "Understanding the IoT for public safety." [Online]. Available: https://iothink-solutions.com/all_resources/understanding-the-iot-for-public-safety/, 2023
[12] "Information and communication technology (ICT)." [Online]. Available: https://www.techopedia.com/definition/24152/information-and-communications-technology-ict, 2023
[13] "Connected healthcare sector chart." [Online]. Available: https://www.beechamresearch.com/download-details/connected-healthcare-sector-chart/, 2024
[14] "What is the Internet of Bodies (IoB), and why should you care?" [Online]. Available: https://itrexgroup.com/blog/internet-of-bodies-iob-definition-benefits-examples/, 2022
[15] "The future of the Internet of Bodies," 2023. [Online]. Available: https://files.microcms-assets.io/assets/8ba880c1ada24b3286662c41b2822851/b70814cee4424407819ae201cca24153/Future%20of%20IoB%20Full%20report%20FINAL%20SOIF%2005.31.pdf, 2023
[16] "Internet of things (IoT) Top 10 2018." [Online]. Available: https://wiki.owasp.org/index.php/OWASP_Internet_of_Things_Project#tab=IoT_Top_10
[17] "The OWASP IoT top 10 vulnerabilities and how to mitigate them." [Online]. Available: https://www.sisainfosec.com/blogs/the-owasp-iot-top-10-vulnerabilities-and-how-to-mitigate-them/, 2023

BÁNKI KÖZLEMÉNYEK --- BÁNKI REPORTS
VI. ÉVFOLYAM 2. SZÁM --- VOLUME VI NO. 2

Optimizing Neural Network Hyperparameters Using Genetic Algorithms for Predicting Student Adaptability in Online Education

Safarpour Motealegh Mahalegi Homayoun*(1),
István Nagy**(2)
*, ** Óbuda University / Institute of Mechatronics and Vehicle Engineering, Budapest, Hungary
homayoun.safarpour@stud.uni-obuda.hu (1)
nagy.istvan@bgk.uni-obuda.hu (2)

***Abstract* — Predicting student adaption is a crucial component of studying online learning material. Machine learning algorithms are crucial in this situation. Deep learning is a fundamental concept in machine learning algorithms. This work used Python in the Jupyter Notebook environment to implement the deep learning approach for forecasting students' adaptation to online learning. The Keras and Tensorflow libraries were used to construct a neural network model using the Kaggle dataset. The data is divided into testing data and training sets and utilize the Keras plot_model utility method to visualize the neural network model. Construct the deep learning model with two hidden layers, each employing randomly picked activation functions from relu, sigmoid, tanh, elu, and selu. Additionally, include one output layer with the softmax activation function. After undergoing a fine-tuning procedure until the alterations stabilized, this model achieved an accuracy of 89.63%.**

***Keywords*:** Evolutionary Algorithms, Neural Network Optimization, Adaptive Learning Systems, Educational Data Mining, Hyperparameter Tuning, Predictive Analytics, Automated Machine Learning, Student Adaptability.

Summary*— *In this paper, we used genetic algorithms (GA) for the optimization of neural network hyperparameters for predicting student adaptability in online learning. Additionally, the findings of our study demonstrate that this strategy improves the neural network design and enhances our comprehension of the aspects that impact student adaptation with the use of modern machine learning techniques [4]. However, this method reduces the result of the loss function and improves the accuracy of the model, which shows that it is crucial to adjust the number of layers and neurons as well as select the desired activation function when exploring hyperparameter spaces, resulting in improved accuracy and reduced error rates [5].

1 INTRODUCTION

especially considering the constant accumulation of data in the student's academic records in higher education. The educational management methods are not genuinely set up to help educational administrators identify which pupils have been under threat of leaving their education. There is plenty of clear data on the topic of education, this data is classified into five categories: gender, age, education level, load-shedding finance, and quantity statistics. The three levels of their adaptivity—low, moderate, and high—are represented in the data. to find instances of prediction for student adaptation to online training. Classical learning environments and online learning systems are different types of educational frameworks; The goal of higher education institutions is to improve the quality of training by optimizing neural network hyperparameters. Using genetic—algorithms inspired by principles of natural selection [4] — is the topic of a later study. Online learning is fantastic for its flexibility, but it can take some adjustment for students. To figure out what helps them succeed, we looked at a Kaggle dataset [1] focused on student adaptability. We are using a neural network; think of it as a powerful analytical tool built with TensorFlow and Keras to dig into that data. This network has multiple layers for complex learning to prevent overfitting, is tailored for multi-class classification [3], and can sort information to give us insights into how students adapt.

1.1 Genetic Algorithms in Hyperparameter Optimization

Using Genetic Algorithms (GAs) to build our neural network is a game-changer, it helps the system automatically find the best settings for itself, leading to more powerful and efficient learning. Since figuring out how students adapt to online learning is complex, picking the right analysis tools (activation functions) within the network is super important [5], the GA determined ReLU as the optimal hidden layer activation function and Softmax for the output layer.

1.2 Activation Functions

figuring out how students adapt to online learning is complex, picking the right analysis tools—activation functions—within the network is super important [5], the GA determined ReLU as the optimal hidden layer activation function and Softmax for the output layer.

1.2.1 The Role ReLU

The Rectified Linear Unit (ReLU), used in the hidden layers, is notable for its simplicity and efficacy, efficiently passing positive inputs while nullifying negative ones [6].

This function is represented by Figure 1, which illustrates its operation of passing positive inputs unchanged while negating negative inputs.

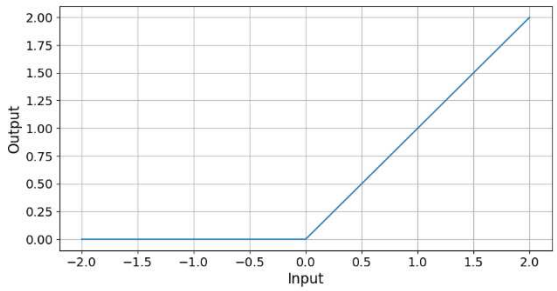


Figure 1: ReLU Activation Function

1.2.2 The Application of SoftMax

In the output layer, the SoftMax function excels at multi-class classification tasks by normalizing the network's outputs, thus providing clear probabilistic insights into student adaptability [7]. As shown in Figure 2, This graphical representation underscores the function's capacity to normalize the network's output, facilitating the derivation of clear, probabilistic insights regarding student adaptability.

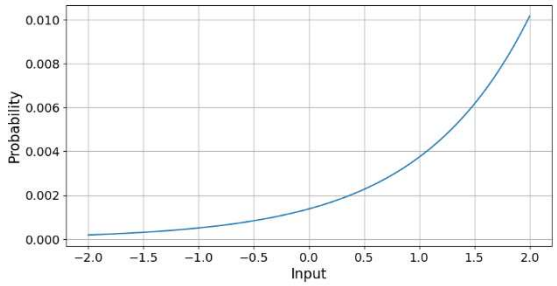


Figure 2: SoftMax Probability Distribution.

1.3 Empirical and Theoretical Foundations

The inclusion of ReLU and SoftMax is supported by extensive research highlighting their effectiveness in deep learning applications, enhancing network performance and computational efficiency, and interpreting outputs as probabilities [8,9].

The strategic selection of ReLU and SoftMax activation functions is instrumental in the development of our neural network model for predicting student adaptability in online education. Their incorporation is grounded in both theoretical and empirical evidence, highlighting their respective roles in ensuring the model's effectiveness and interpretability.

1.4 TensorFlow as a Development Framework

TensorFlow—Google's powerful framework—is the backbone of our research on optimizing neural networks for adaptability prediction. It goes beyond a set of tools – TensorFlow lets us build truly tailored networks. Key here are its dynamic computation graphs (ideal for complex student adaptability relationships) and the option for GPU acceleration. Early on, the dataset's size slowed down training, but switching to GPUs made all the difference [10].

1.5 Leveraging TensorFlow for Educational Data Mining

We chose TensorFlow for a reason, we need something capable of handling real-world messiness in online learning data, and flexible enough to let us zero in on those subtle adaptability patterns. TensorBoard's visualizations are lifesavers for spotting training errors...especially for dimensionality reduction [11].

This is not just about tech, Using TensorFlow underscores our commitment to a truly data-driven approach in educational research. It is about applying the latest machine -learning techniques to understand how students adapt and using that knowledge to improve online learning for everyone [12].

2 METHODOLOGY

2.1 Data Preparation and Dataset Description

he first step is to import the necessary libraries and then load the dataset using the Pandas library. To convert categorical variables to numerical representations, we use the Sklearn preprocessing library's LabelEncoder function. The next step is data cleaning to ensure that the data is "clean" enough for analytical work, which means it appropriately represents the information you plan to study without distortions caused by poor data quality. Eliminate noise and errors, handle missing data, and ensure data consistency. The dataset [1], contains demographic, educational, and infrastructure data for 1200 students, which we divided into two groups: 20% for validation and 80% for evaluating our models. The preprocessing phase is critical for improving data quality, precision, and dependability in predictive performance [2][16].

```
data_path = "path_to_dataset.csv"
df = pd.read_csv(data_path)
categorical_columns = ['Gender', 'Education Level', ...]
for col in categorical_columns:
    le = LabelEncoder()
    df[col] = le.fit_transform(df[col])
X_train, X_val, y_train, y_val = train_test_split(
    df.drop('target_column', axis=1),df['target_column'],
test_size=0.2, random_state=42)
```

Figure 3: Data Handling Step.

2.2 Genetic Algorithm for Hyperparameter Optimization

Genetic Algorithm was used to optimize the structure of a feedforward multiple-layer neural network. Optimized variables in chromosomes involve the number of artificial neurons and hidden layers and training parameters such as population size, maximal learning step size, percentage of the fittest chromosomes for crossing-over, number of random mutations per chromosome, and crossing-over intensity per chromosome [6,7]. An appropriate artificial neural network was developed and trained for each of the parameters listed above. A population of ten produced models with varying training mistakes. A single artificial neural network was trained for a maximum of fifty

iterations (epochs). the same number of individuals as in the initial population is chosen to continue to the next generation for the fittest were selected to create next-generation chromosomes, which underwent mutation and crossing-over procedures. The goal of structure optimization was to find the parameter combinations that resulted in the artificial neural network with the lowest training error and highest accuracy [8], and to find the best activation function.

2.3 Model Training and Evaluation

For the training of our model, which employed a genetic algorithm, we constructed the chromosomes using most of the listed weights. The structure optimization application provided the learning parameters and neural network topology. The training process concluded once the rate-loss function of the artificial neural network reached its minimum value. After that, the models were used to make output values based on the input data sets that were used for training and validation [9].

This search is vital for identifying optimal model configurations that might not be accessible through traditional optimization techniques [10]. These enhancements underscore the efficacy of GAs in refining the NN model, making it a more reliable tool for understanding and predicting student adaptability [11].

2.4 TABLES, FIGURES, AND CODE SNIPPETS

In this section, we present essential visual aids and code excerpts that substantiate our methodology and findings, clarifying the model's performance and the effectiveness of the GA optimization process.

2.4.1 Neural Network Performance Before GA Optimization

Before the application of Genetic Algorithms (GA) for hyperparameter tuning, we assessed the performance of our neural network model to establish a baseline for subsequent optimization [5]. This evaluation is crucial, as it highlights the initial capabilities of the model and identifies potential areas for enhancement through the sophisticated search techniques that GA provides.

The neural network, designed with a multi-layered architecture and initiated with heuristic hyperparameters, was subjected to extensive training and validation processes [6]. Initial training spanned a considerable number of epochs, allowing the model substantial time to learn from the training data. We recorded key performance metrics during this phase, including accuracy and loss on both training and validation datasets [7]. These metrics offered insights into the model's learning progression and its ability to generalize.

The accuracy metric, indicative of the model's predictive correctness, and the loss metric, reflective of the model's error magnitude, were monitored at each epoch [8]. These metrics served not only to assess the efficacy of the model but also to detect early signs of overfitting or underfitting—conditions that could compromise the model's performance on new, unseen data.

Examining these initial results critically is vital, as they set the groundwork for the subsequent application of GA optimization. Improvements in the model's performance post-optimization can be directly attributed to the GA's more effective navigation of the hyperparameter space compared to the initial heuristic approach [9].

By analysing the neural network's behaviours before GA optimization, we aim to draw meaningful comparisons between the pre- and post-optimization phases. This comparison will not only highlight the impact of GA but will also affirm the robustness of the optimization process itself [10].

Figure 4 shows neural network training and validation accuracy over 50 epochs [11]. Training accuracy initially rises sharply, indicating that the model is quickly assimilating the data. This sharp increase fades around the 10th epoch, suggesting the model stabilizes around an optimal training data set of parameters. The figure shows that, since the validation set does not affect the model's weight adjustments and indicates how well the model generalizes to unseen data, validation accuracy increases more slowly [12]. The model performs better on training data than validation data, indicating overfitting. At the end of the training period, both accuracies stabilize, with the validation accuracy fluctuating but rising. A model that adapts may need more training or tuning [13].

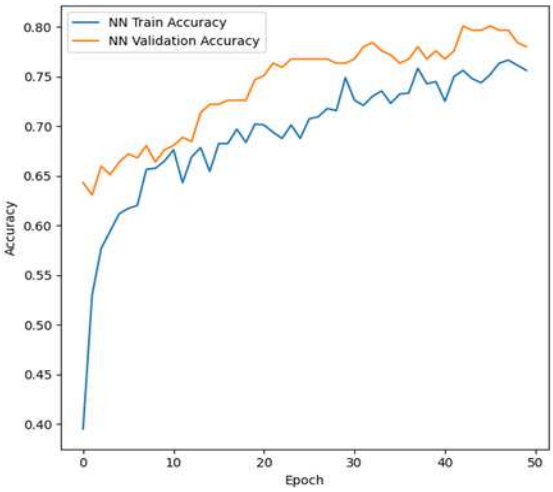


Figure 4: Model accuracy comparison.

Figure 5 shows the value of the loss function in 50 epochs of training. Loss, a key metric, measures the difference between predicted and actual values. Lower values indicate better performance. As training epochs progress, loss curves flatten, indicating diminishing returns. Training loss drops sharply from 1.1 to below 0.6, while validation loss follows, indicating effective learning and generalization. Limited overfitting is indicated by the narrow training-validation loss gap.

Overall, the data from both figures suggest that the model is learning effectively; however, there might be opportunities for enhancement, potentially through the implementation of techniques aimed at reducing overfitting, such as adding dropout layers, employing regularization, or expanding the variety and volume of training data. Further experiments to fine-tune the model's

hyperparameters could also help achieve a more optimal balance between bias and variance.

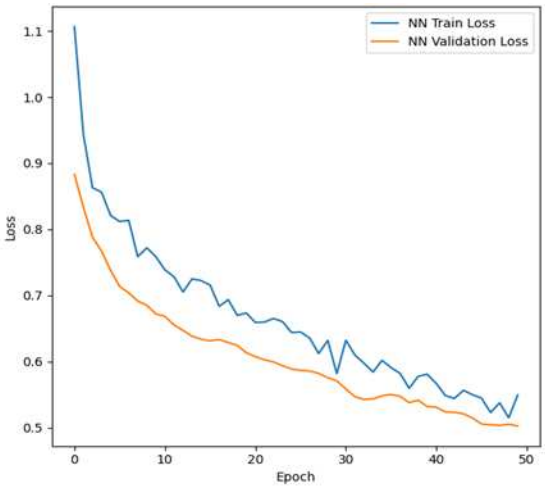


Figure 5: Model loss comparison.

2.5 Comparative Analysis Post Genetic Algorithm Optimization

The application of Genetic Algorithm (GA) optimization exemplifies the evolutionary capabilities inherent in machine learning methodologies. Figures 6 and 7 illustrate the quantifiable improvements in neural network accuracy and efficiency as direct outcomes of GA optimization. These figures are crucial as they not only document the progression of the model's performance metrics over iterative epochs but also highlight the substantial enhancements brought about by GA intervention.

Figure 6 presents a comparative analysis of the accuracy rates achieved by the neural network before and after the application of GA optimization over 50 epochs. The GA Train Accuracy exhibits a higher trajectory compared to the pre-optimization NN Train Accuracy, suggesting more robust learning from the training data due to GA optimization. The GA Validation Accuracy also shows an improvement, consistently maintaining a higher level than the NN Validation Accuracy. This enhancement indicates that the GA has effectively improved the model's ability to generalize. Notably, after an initial period of volatility, the validation accuracy stabilizes, demonstrating gradual improvement and suggesting that GA may have contributed to mitigating overfitting to the training data—a common challenge in machine learning models.

These sections have been refined to maintain academic rigor, offering detailed insights into the performance metrics and the impact of GA optimization. The technical descriptions are precise, and the narrative is structured to guide the reader through the progression and outcomes of the research effectively.

In **Figure 7**, we can see how using GA boosted our model's learning. Both the GA Train Loss and Validation Loss lines dropped dramatically, indicating that our model performed much better at its task. See how the GA model's train loss starts high around 1.0 and quickly decreases to about 0.4; The validation loss decreasing along with it shows it's not just memorizing training data but learning to generalize. Compare that to the standard neural network line – the loss decreases, but more slowly. This shows that our GA optimization made a real difference – by the end, the GA model simply performs better.

Figure 6: Model accuracy comparison after applying GA.

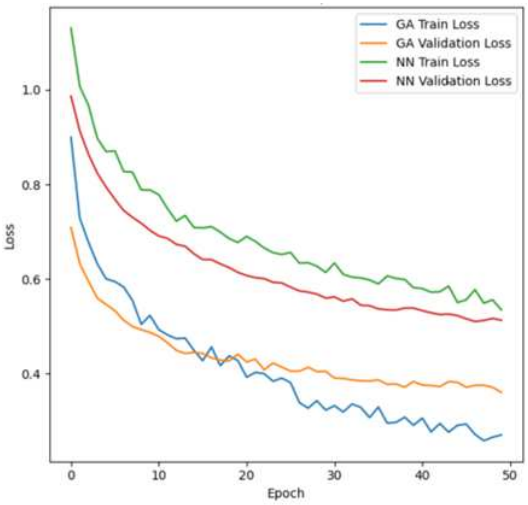


Figure 7: Comparison of Training and Validation Loss for GA and NN Models.

These figures collectively indicate a successful optimization process. The GA's strategic exploration of hyperparameter space appears to have endowed the neural network with an enhanced capacity to learn and predict more accurately, as reflected in the higher accuracy and lower loss observed post-optimization.

2.6 Analysing Adaptability Influences: A Methodological Approach

The goal is to create a model that can analyse these indicators both alone and in combination so that we can distinguish the level of student adaptation to improve learning performance or which student is on leave from university or college. In this way, the dataset serves as both empirical evidence and a foundation for methods to modify instructional technology and includes variables such as 'Adaptivity Level,' 'Gender,' 'Age,' 'Education Level,'

'Load-shedding,' 'Financial Status,' and more. When combined, these factors offer a comprehensive understanding of the flexibility of online learning. We aim to build more inclusive and successful online educational frameworks by analysing data thoroughly to understand the complex interplay of variables impacting student adaptation. Several factors may influence the adaptability of online learning; the following table illustrates the dataset's important characteristics.

Table 1:Dataset Overview

Gender	Age	Education Level	Load-shedding	Financial	Adaptivity Level
Boy	21-25	University	Low	Mid	Moderate
Girl	21-25	University	High	Mid	Moderate
Girl	16-20	College	Low	Mid	Moderate
Girl	15-Nov	School	Low	Mid	Moderate
Girl	16-20	School	Low	Poor	Low
Boy	15-Nov	School	Low	Poor	Low
Boy	15-Nov	School	Low	Mid	Low
Boy	15-Nov	School	Low	Mid	Moderate
Boy	16-20	College	Low	Mid	Low
Boy	15-Nov	School	Low	Mid	Moderate
Girl	16-20	University	Low	Mid	Low
Girl	16-20	College	Low	Mid	Low
Boy	15-Nov	School	Low	Mid	Moderate
Girl	16-20	College	Low	Mid	Low

2.7 Fine-Tuning for Top Performance

We put our model through its paces over 50 training sessions (epochs) with the crucial goal of ensuring it could learn effectively from the data it was given and then apply that knowledge to completely new situations. After training and tuning the model with validation data, we tested its performance with unseen data. For this goal, Table 2 provides a snapshot of model accuracy and error rates—*loss function output*—at various stages of training. This is where the human element comes in; we analyze these metrics to see if any adjustments are needed that have not been done by GA or not.

Table 2: Model Training Parameters and Results

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	57.88%	68.46%	0.8837	0.6911
2	65.77%	73.86%	0.7218	0.617
3	72.20%	74.69%	0.664	0.5805
4	71.47%	78.01%	0.6328	0.5583
5	73.03%	75.93%	0.6319	0.5435
6	75.00%	78.01%	0.5827	0.5095
...	...	...	...	...
50	88.90%	89.63%	0.2684	0.3408

Adding more layers to the neural network design, from three to four, implies that the educational data may be more complexly abstracted. At the same time, the network's enhanced representational power is shown by the growth of neuron counts, from 100 to 150 in the first layer and from 50 to 120 in the second and also we reduced tenfold to 0.001, enabling finer-grained modifications when training the model and also The dropout rate was also adjusted from 0.5 to 0.3 to complement these structural improvements; this should help reduce the likelihood of overfitting and improve the model's ability to generalize to other types of datasets. Better performance metrics quantify the results of this optimization process. The model demonstrated exceptional skill in learning from the provided data and in generalizing to new data subsets, as seen by an increase in training accuracy to 82.47% and validation accuracy to 80.32%.

Concurrently, the model's loss measures decreased, with training loss dropping to 0.35 and validation loss to 0.33, highlighting the enhanced predictive accuracy after optimization.

The empirical statistics demonstrating the effect of the GA are summarized in Table 3, which follows this narrative. By comparing the parameters of the neural network before and after optimization, it shows how the method improved performance.

Table 3: Hyperparameter Optimization Results

Parameter	Before Optimization	After Optimization
Number of Layers	3	4
Neurons in Layer 1	100	150
Neurons in Layer 2	50	120
Learning Rate	0.01	0.001
Dropout Rate	0.5	0.3
Activation Function	ReLU	ReLU
Training Accuracy	75.62%	82.47%
Validation Accuracy	78.01%	80.32%
Training Loss	0.5	0.35
Validation Loss	0.48	0.33

This table not only illustrates the GA's role in optimizing our neural network model but also serves as a prelude to the ensuing results and discussion section. It prepares the reader for a deeper analysis of the performance improvements observed, setting a solid empirical foundation for the subsequent interpretative commentary on the model's enhanced ability to predict student adaptability in online education settings.

Figure 8 presents the initialization code for the GA population [15]. This snippet provides insight into how we generated an initial population of potential solutions (NN configurations) for the optimization process.

```
def create_individual():
    individual = []
    for _ in range(n_layers - 1):
        neurons = random.randint(neuron_min, neuron_max)
        activation = random.choice(list(activation_function_map.keys()))
        individual.append(neurons)
        individual.append(activation_function_map[activation])
    individual.append(num_classes)
    individual.append(activation_function_map['softmax'])
    return individual
```

Figure 8:GA Population Initialization

And in Figure 9 demonstrates the function used to evaluate each individual NN configuration's performance within the GA.

```
def evaluate_model(individual):
    model = Sequential()
    for i in range(0, len(individual) - 2, 2):
        neurons = individual[i]
        activation = list(activation_function_map.keys())[individual[i + 1]]
        model.add(Dense(neurons, activation=activation))
    model.add(Dropout(0.5))
    model.add(Dense(num_classes, activation='softmax'))
    model.compile(optimizer='Adam', learning_rate=0.01),
    loss='categorical_crossentropy', metrics=['accuracy'])
    model.fit(x_train, y_train_categorical, epochs=10, batch_size=32,
              verbose=0, validation_data=(x_val, y_val_categorical))
    _, accuracy = model.evaluate(x_val, y_val_categorical, verbose=0)
    return accuracy,
```

Figure 9: Model Evaluation Function

2.8 Research Insights

our results demonstrate that Genetic Algorithms (GA) may enhance the prediction of students' adaptation to online courses by optimizing the hyperparameters of neural networks and promising results across various domains [12]. This optimization results in reduced loss and increased accuracy leading to less complex models with better performance on time series prediction problems [13] for improving educational data mining and aiding in the creation of adaptive learning systems.

3 RESULTS AND DISCUSSION

In this study, we enhanced the architecture of a Neural Network (NN) by using Genetic Algorithms (GA) to figure out the adaptability of online students. The hyperparameters, such as the number of layers, number of neurons per layer, choice of activation functions, and dropout rates, were fine-tuned for optimization.

3.1 Results

GA improved NN model performance significantly. Initial validation accuracy for the NN model was 78.01% with a loss of 0.48. GA optimization increased model validation accuracy to 89.63% and decreased loss to 0.3408. These findings show that GA can navigate hyperparameter space to get the best NN setup.

3.2 Discussion

volving techniques to develop artificial neural network weights and structure can represent complicated connections from raw process data. Genetic algorithms can work on chromosomes with many parameters if the crossing-over and mutation processes are set properly. The approaches may be used for many optimization and model prediction problems.

4 REFERENCES

[1] Hasan, M., & Suzan, M. D. (2021). "Students Adaptability Level in Online Education Kaggle Dataset." DOI: 10.1109/ICCCNT51525.2021.9579741.

[2] Hesami, M., & Jones, A. M. P. (2020). "Application of machine learning models in plant cell and tissue culture." Applied Microbiology and Biotechnology, 104(20), 8549–8564. DOI: 10.1007/s00253-020-10851-5.

[3] Lee, J. (2018). "Deep Learning for Multi-Class Prediction of Student Performance in Educational Data." International Journal of Recent Technology and Engineering (IJRTE), 7(6), 2155–8219.

[4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[5] Bengio, Y. (2012). "Practical recommendations for gradient-based training of deep architectures." In Neural Networks: Tricks of the Trade (pp. 437-478). Springer, Berlin, Heidelberg.

[6] Asadi, B., & Jiang, H. (2020). "On Approximation Capabilities of ReLU Activation and Softmax Output Layer in Neural Networks."

[7] Stanojevic, A., Wozniak, S., Bellec, G., Cherubini, G., Pantazi, A., & Gerstner, W. (2022). "An Exact Mapping From ReLU Networks to Spiking Neural Networks."

[8] Yang, Y., Wu, Y., Yang, H., & Xiang, Y. (2023). "Nearly Optimal Approximation Rates for Deep Super ReLU Networks on Sobolev Spaces."

[9] Schmidt-Hieber, J. (2019). "Lecture 4: Mathematics for Deep Neural Networks: Statistical theory for deep ReLU networks."

[10] Peng, N. (2021). "Research on the effectiveness of English online learning based on neural network." Neural Computing and Applications.

[11] Zhou, Y., Niu, K., Lv, H., Lu, G., & Pan, Y. (2023). "CGDC-LSTM: A novel hybrid neural network model for MOOC dropout prediction." International Joint Conference on Neural Networks (IJCNN).

[12] Cui, Y., Surpur, C., Ahmad, S., & Hawkins, J. (2016). "A comparative study of HTM and other neural network models for online sequence learning with streaming data." International Joint Conference on Neural Networks (IJCNN).

[13] Holland, J. H. (1992). "Genetic algorithms." Scientific American, 267(1), 66-73.

[14] Goldberg, D. E., & Holland, J. H. (1988). "Genetic algorithms and machine learning." Machine Learning, 3(2), 95-99.

[15] Whitley, D. (1994). "A genetic algorithm tutorial." Statistics and Computing, 4(2), 65-85.

[16] Hark, C., Okumuş, H., & Uçkan, T. (2022). "Adaptation to Online Education: An Educational Data Mining Application." Journal of Educational Data Mining, DOI: 10.53070/bbd.1199055.