

8. Резюмета на публикациите за участие в конкурса на български и английски език

Александър Илиев Илиев

конкурс за професор в област на висше образование 4. Природни науки, математика и информатика, професионално направление 4.6. Информатика и компютърни науки, научна специалност: Информатика (Конгнитивни модели и изкуствен интелект за мултимодални данни), обявен в ДВ, бр. 54/12.06.2026 г.

Монографии:

- A. **Alexander I. Iliev**, "AI Popular Systems. Architecture, Capabilities, and Comparative Analysis", LAP Lambert Academic Publishing, ISBN: 978-66-30-21444-4, 2026, pp. 292
- B. **Alexander I. Iliev**, "Practical Benchmarking and Hands-On Evaluation of AI Popular Systems", LAP Lambert Academic Publishing, ISBN: 978-66-30-22514-3, 2026, pp. 164

Учебници:

- C. **Alexander I. Iliev** (2023), "Artificial Intelligence Smart Systems Design Applied Use Cases Part 1", LAP Lambert Academic Publishing, ISBN: 978-620-6-78830-0, pp. 136
- D. **Alexander I. Iliev** (2023), "Artificial Intelligence in Smart Systems Applied Use Cases Part 2", LAP Lambert Academic Publishing, ISBN: 978-620-7-45055-8, 2023, pp. 168
- E. **Alexander I. Iliev**, "Smart Python for Machine Learning and Intelligent Modeling, Foundations and Classical Machine Learning | Part 1", LAP Lambert Academic Publishing, ISBN 978-66-30-10726-5, 2026, pp. 316
- F. **Alexander I. Iliev**, "Smart Python for Machine Learning and Intelligent Systems, Deep Learning, Transfer Learning, and AI Engineering | Part 2", LAP Lambert Academic Publishing, ISBN 978-66-30-21092-7, 2026, pp. 324

Глава от книга:

- G. Alexander I. Iliev, "Perspective chapter: Emotion detection using speech analysis and Deep Learning", Publisher: IntechOpen, Book title: "Emotion Recognition - Recent Advances, New Perspectives and Applications", Editor: Seyyed Abed Hosseini, 2023, DOI: 10.5772/intechopen.110730, Link

Полезен модел:

- H. "Big Data Processing System", Type: Utility model, Authors: A.I. Shikalanov, L.M. Kirilov, E.P. Kovacheva, R.V. Nikolov, E.D. Shoikova-Stoyanova, A.I. Iliev, L.I. Gotsev, Patent Office of the Republic of Bulgaria, Patent No. 5104 U1 / 29.08.2025

Публикация №1 – в том на конференция:

- 1. **Iliev, A.I.**. Content Discovery Using Perceptual Automation. MEDES 2018 - 10th International Conference on Management of Digital EcoSystems, Association for Computing Machinery, 2018, ISBN:978-1-4503-5622-0, DOI:10.1145/3281375.3281399, 233-238 **Без JCR или SJR – индексирани в WoS или Scopus** [Линк](#)

Резюме:	
	В тази работа се предлага иновативна система за откриване и избор на медийно съдържание на базата на човешкото поведение. Проектът имаше две части: първо, във фазата на възприятие, речевият сигнал се използва за анализ с цел откриване и извличане на емоционални подсказки; и второ, във фазата на автоматизация, текстова информация беше използвана за

	определяне на настроението чрез методология за обработка на естествен език. Речевите подсказки първо се улавят от сигнала и след това се класифицират в три емоционални категории: възходящо, позитивно или щастливо; беземоционално, равномерно или неутрално; и спадащо, негативно или тъжно. След това, за да съответства събраните емоции от речта с конкретна контекстуална информация, извлечена от реален източник, чрез техники за обработка на естествен език бяха обработени пет книги от "Игра на тронове". Резултатите бяха обобщени и направено беше съпоставяне между двете системи, така че книгите да могат да бъдат избирани на базата на степента на всяка изразена емоция от говорещите. Това може да се извършва по ненаатрапчив начин без директно и явно вмешателство между човек и машина.
Abstract:	In this work, an innovative media content discovery and selection system is proposed based on human behavior. There were two parts to this project: first, in the perception phase, a speech signal is used for analysis with the intention to detect and extract emotional cues; and second, in the automation phase, textual information was used to determine the sentiment using natural language processing methodology. The speech cues are first captured from the signal and then classified in three emotional categories: upbeat, positive, or happy; emotionless, flat, or neutral; and downbeat, negative or sad. Then in order to match the collected emotions from speech to specific contextual information captured from a real source, using natural language processing techniques five books from The Game of Thrones were processed. The results were summarized and mapping between the two systems was drawn so that books can be chosen based on the degree of each emotion portrayed from speakers. This can be done in a non-intrusive fashion without direct and explicit human-machine intervention.

Публикация №2 – в том на конференция:

2. **Iliev, A.I., Stanchev, P..** Smart multifunctional digital content ecosystem using emotion analysis of voice. 18th International Conference on Computer Systems and Technologies CompSysTech'17, 1369, ACM, 2017, ISBN:978-1-4503-5234-5, pp.58-64, DOI: 10.1145/3134302.3134342, **SJR: 0,159 (2017) – индексирани в WoS или Scopus (Scopus)** [Линк](#)

Резюме:	В опит да се установи подобрена архитектура, ориентирана към услугите (SOA) за съвместим и персонализиран достъп до дигитални културни ресурси, автоматична детерминистична техника може потенциално да доведе до подобрение на търсенето, препоръчването и персонализирането на съдържанието. Такава техника може да бъде разработена по много начини, използвайки различни средства за търсене и анализ на данни. Тази статия се фокусира върху използването на гласово и емоционално разпознаване в речта като основно средство за предоставяне на алтернативен начин за разработване на новаторски решения за интеграция на слабо свързаните компоненти, които обменят информация на базата на общ модел на данни. Параметрите, използвани за изграждане на векторите на характеристиките за анализ, носеха информация за височина, време и продължителност. Те бяха сравнени с глобалната симетрия, извлечена от говорения източник чрез инверсно филтриране. Сравнение с техните първи производни също беше предмет на изследване в тази статия. Говореният източник беше 100-минутна театрална пиеса, съдържаща четири мъжки говорещи, и беше записана при 8kHz с 16-битова резолюция на пробите. Бяха целени четири емоционални състояния: щастлив, ядосан, страх и неутрален. Класификацията беше извършена с метода на най-близките съседи (k-Nearest Neighbor). Тренировъчни и тестови експерименти бяха извършени в три сценария: 60/40, 70/30 и 80/20 минути съответно. Близко сравнение на всяка характеристика и тяхната скорост на промяна показва, че характеристиките във времевия домейн постигат по-добри резултати с по-малко изчислително напрежение от техните производни. Освен това беше постигната верна степен на разпознаване до 95% с избраните характеристики.
---------	--

Abstract:	In an attempt to establish an improved service-oriented architecture (SOA) for interoperable and customizable access of digital cultural resources an automatic deterministic technique can potentially lead to the improvement of searching, recommending and personalizing of content. Such technique can be developed in many ways using different means for data search and analysis. This paper focuses on the use of voice and emotion recognition in speech as a main vehicle for delivering an alternative way to develop novel solutions for integrating the loosely connected components that exchange information based on a common data model. The parameters used to construct the feature vectors for analysis carried pitch, temporal, and duration information. They were compared to the glottal symmetry extracted from the speech source using inverse filtering. A comparison to their first derivatives was also a subject of investigation in this paper. The speech source was a 100-minute-long theatrical play containing four male speakers and was recorded at 8kHz with 16-bit sample resolution. Four emotional states were targeted namely: happy, angry, fear, and neutral. Classification was performed using k-Nearest Neighbor method. Training and testing experiments were performed in three scenarios: 60/40, 70/30 and 80/20 minutes respectively. A close comparison of each feature and its rate of change show that the time-domain features perform better while using lesser computational strain than their first derivative counterparts. Furthermore, a correct recognition rate was achieved of up to 95% using the chosen features.
------------------	--

Публикация №3 – в том на конференция:

3. **Iliev Al., P. Stanchev.** Glottal Attributes Extracted from Speech with Application to Emotion Driven Smart Systems. 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2018) - Volume 1: KDIR, 2018, ISBN:978-989-758-330-8, pp. 297-302, DOI: 10.5220/0006951002970302, **Без JCR или SJR – индексирани в WoS или Scopus (Scopus)**

Резюме:	Всяко новаторско разработване на умен систем зависи от взаимодействието човек-компютър и също така е зависимо директно или индиректно от емоцията на потребителя. В тази статия предлагаме идея за разработване на умен систем, използвайки извличане на настроение от речта с възможно приложение в различни области от нашия ежедневен живот. За крос-валидация бяха използвани два различни корпуса на реч, с обучение и тестване върху всяко множество. Системата е независима от текст, съдържание и пол. Емоции бяха извлечени от женски и мъжки говорители. Системата е устойчива на външен шум и може да бъде приложена в области като забавление, персонализация, автоматизация на системи, обслужващи индустрии, сигурност, наблюдение и много други.
Abstract:	Any novel smart system development depends on human-computer interaction and is also dependent either directly or indirectly on the emotion of the user. In this paper we propose an idea for the development of a smart system using sentiment extraction from speech with possible application in various areas in our everyday life. Two different speech corpora were used for cross-validation with training and testing on each set. The system is text, content, and gender independent. Emotions were extracted from both female and male speakers. The system is robust to external noise and can be implemented in areas such as entertainment, personalization, system automation, service industries, security, surveillance, and many more.

Публикация №4 – в том на конференция:

4. **Iliev, A. I., Mote, A., Manoharan, A..** Cross-Cultural Emotion Recognition and Comparison Using Convolutional Neural Networks. Digital Presentation and Preservation of Cultural and Scientific Heritage Conference Proceedings, Vol. 10, Institute of Mathematics and Informatics –

BAS, 2020, ISSN:1314-4006, pp. 089-102 **Без JCR или SJR – индексирани в WoS или Scopus (Web of Science)** [Линк](#)

Резюме:	Статията има за цел да определи сравнение на емоциите в три различни култури, а именно канадски френски, италиански и северноамерикански. Това беше постигнато чрез използване на проби от реч за всяка от трите езика, които бяха обект на нашето изследване. Използваните характеристики бяха MFCCs и бяха пропуснати през конволюционна невронна мрежа, за да се провери тяхната значимост за задачата за разпознаване на емоции чрез реч. Бяха обучени и тествани три различни системи, по една за всеки език. Точността беше 71,10%, 79,07% и 73,89% съответно за всеки от тях. Целта беше да се докаже, че векторите на характеристиките, които използвахме, представяха добре всяка емоция. В края беше направено сравнение между всяка емоция, пол и език, и се наблюдаваше, че освен емоцията неутрална, всяка друга емоция се изразяваше по малко различен начин от всяка култура. Речта е един от основните начини за разпознаване на емоции и представлява привлекателна област за изучаване с приложение към представянето и запазването на различни културни и научни наследства.
Abstract:	The paper sets to define a comparison of emotions across 3 different cultures namely Canadian French, Italian, and North American. This was achieved using speech samples for each of the three languages subject to our study. The features used were MFCCs and were passed through convolutional neural network in order to verify their significance for the task of emotion recognition through speech. Three different systems were trained and tested, one for each language. The accuracy came to 71.10%, 79.07%, and 73.89% for each of them respectively. The aim was to prove that the feature vectors we used were representing each emotion well. A comparison across each emotion, gender and language was drawn at the end and it was observed that apart from the emotion <i>neutral</i> , every other emotion was expressed somewhat differently by each culture. Speech is one of the main vehicles to recognize emotions and is an attractive area to be studied with application to presenting and preserving different cultural and scientific heritage.

Публикация №5 – в том на конференция:

5. **Iliev, A. I., Stanchev, P. L.** Smart Ecosystems through Voice and Images. CATA 2020, San Francisco: EPiC Series in Computing, Vol. 69, Proceedings of 35th International Conference on Computers and Their Applications, 2020, ISSN:2398-7340, pp. 256-263, doi:10.29007/ztmt, **SJR (2020): 0,21, индексирани в WoS или Scopus (Scopus)**

Резюме:	В тази статия обобщаваме на високо ниво някои от популярните умни технологии, които могат да формират много екосистеми на умни градове. По-конкретно ще акцентираме върху автоматизацията на различни процеси на базата на извличане и анализ на дигитални медии чрез речеви сигнали и изображения. В момента съществуват много стандартизирани системи за персонализация и препоръчване на дигитално медийно съдържание, както и различни услуги в разнообразни области. Повечето от тях са разработени с предвид взаимодействието между човек и машина. Обикновено това се постига чрез традиционна употреба на мишка и клавиатура. Потребителят ръчно въвежда своя отговор, който след това се записва от системата за по-нататъшен анализ.
Abstract:	In this article we summarize at a high-level some of the popular smart technologies that may contrive many smart city ecosystems. More specifically we will emphasize the automation of various processes based on the extraction and analysis of digital media, through speech signals and images. Currently, there are many productized systems for personalization and recommendation of digital media content as well as various services in different areas. Most of them are

	developed with human-machine interaction in mind. Usually, this is done through a conventional use of a mouse and a keyboard. The user types their response manually, which is then recorded by the system for further analysis.
--	--

Публикация №6 – в книга:

6. **Alexander I. Iliev, Peter L. Stanchev.** Smart Services Using Voice and Images. Special Issue on "Digital Ecosystems and Social Networks" DESN, TLDKS: LNCS, Volume XLVII, 12630, ISBN 978-3-662-62918-5, <https://doi.org/10.1007/978-3-662-62919-2>, pp.137-154, Springer, 2021, 20, **SJR (Scopus):0.407 Q2 (Scopus)** [Линк](#)

Резюме:	В тази глава ще акцентираме върху някои от най-забележителните напредъци в умните технологии, които формират екосистемата на умния град. Освен това ще подчертаем автоматизацията на множество разработки на базата на извличане и анализ на дигитални медии, използвайки реч и изображения. В момента съществува множество практични системи, използвани за персонализация и препоръчване на различни медии. От друга страна, има разнообразни типове услуги в различни области, които пряко извличат полза от тези напредъци. Повечето от тях са създадени с предвид методологията за взаимодействие между човек и машина, където хората трябвало да взаимодействат с машините по различни начини. В миналото този тип взаимодействие е бил извършван чрез използването на традиционни интерфейси като мишка и клавиатура, където потребителят трябвало ръчно да въвежда отговор, който след това бивал записван от машината за последващ анализ. Ето защо, за да опростим тези видове взаимодействия и да допринесем за подобряване на услугите, нови методологии трябва да бъдат изучавани, откривани и разработвани, за да подобрят услуги като препоръчване и персонализация.
Abstract:	In this chapter, we will emphasize on some of the most prominent advances in smart technologies that formulate the smart city ecosystem. Furthermore, we will be highlighting the automation of numerous developments based on the extraction and analysis of digital media, using speech and images. At present, there is a multitude of practical systems used for personalization and recommendation of different media. On the other hand, there are assorted types of services in different areas that are directly benefiting from these advancements. Most of them were created with human-machine interaction methodology in mind, where people had to interact with the machines in various ways. In the past this type of interaction has been completed through the use of conventional interfaces such as a mouse and a keyboard, where the user had to type a response manually, which was in turn recorded by the machine for subsequent analysis. Therefore, in order to simplify these types of interactions and lead to improvement of services, new methodologies must be studied, discovered and developed so as to improve services such as recommendation and personalization services.

Публикация №7 – в том на конференция:

7. **Iliev, A.I., Dewli M., Kalkan M., Prakash P.K., Turkar R..** Acoustic Event Detection and Sound Separation for security systems and IoT devices. CompSysTech '21: International Conference on Computer Systems and Technologies, ACM International Conference Proceeding Series, 2021, ISBN: 978-145038982-2, DOI:<https://doi.org/10.1145/3472410.3472441>, pp. 34-39. **SJR (Scopus):0.232, индексан в WoS или Scopus (ACM Digital Library)** [Линк](#)

Резюме:	Когато мислим за аудио данни, мислим за музика и реч. Обаче, множеството от различни видове аудио данни съдържа огромно разнообразие от различни звуци. Човешкият мозък може да идентифицира звуци като два автомобила, които се блъскат един в друг, някой, който плаче, или експлозия на бомба. Когато чуем такива звуци, можем да разпознаем източника на звука и събитието, което ги е причинило. Можем
---------	---

	да изградим изкуствени системи, които да разпознават акустични събития, точно както хората. Разпознаването на акустични събития (AED) е технология, която се използва за откриване на акустични събития. Не само че можем да открием акустичното събитие, но също така да определим времеви интервал и точния момент на настъпване на всяко събитие. Тази статия има за цел да използва конволюционни невронни мрежи при класифицирането на околна среда, свързани с определени акустични събития. Класификацията и откриването на акустични събития имат множество приложения в реалния свят, като откриване на аномалии в индустриални инструменти и машини, умни домашни системи, приложения за сигурност, маркиране на аудио данни и създаване на системи за помощ на хора с увреден слух. Въпреки че звуците от околната среда могат да обхващат голямо разнообразие от звуци, ще се съсредоточим специално върху определени градски звуци в нашето изследване и ще използваме конволюционни невронни мрежи (CNNs), които традиционно се използват за класификация на изображения, за нашия анализ на аудио данни. Моделът, когато му бъде предоставен пробен аудио файл, трябва да може да зададе класификационен етикет и съответен точен резултат.
Abstract:	When we think of audio data, we think of music and speech. However, the set of various kinds of audio data, contains a vast multitude of different sounds. Human brain can identify sounds such as two vehicles crashing against each other, someone crying or a bomb explosion. When we hear such sounds, we can identify the source of the sound and the event that caused them. We can build artificial systems which can detect acoustic events just like humans. Acoustic event detection (AED) is a technology which is used to detect acoustic events. Not only can we detect the acoustic event but also, determine the time duration and the exact time of occurrence of any event. This paper aims to make use of convolutional neural networks in classifying environmental sounds which are linked to certain acoustic events. Classification and detection of acoustic events has numerous real-world applications such as anomaly detection in industrial instruments and machinery, smart home systems, security applications, tagging audio data and in creating systems to aid the hearing-impaired individuals. While environmental sounds can encompass a large variety of sounds, we will focus specifically on certain urban sounds in our study and make use of convolutional neural networks (CNNs) which have traditionally been used to classify image data, for our analysis on audio data. The model, when given a sample audio file must be able to assign a classification label and a corresponding accuracy score.

Публикация №8 – в том на конференция:

- Iliev, A.I., Manoharan, A., Mote, A.. Multi-Lingual Emotion Classification using Convolutional Neural Networks.** Lecture Notes in Computer Science, Proc. 13th International Conference on Large-Scale Scientific Computations, 13127, Springer, 2022, ISBN:978-3-030-97548-7, ISSN:03029743, DOI:10.1007/978-3-030-97549-4_52, pp. 456-463. **SJR (Scopus):0.320 Q3 (Scopus)** [Линк](#)

Резюме:	Емоциите играят централна роля в човешката взаимодействие. Интересното е, че различните култури имат различни начини да изразят същата емоция, и това мотивира нуждата от изучаване на начина, по който емоциите се изразяват в различни части на света, за да разберем тези разлики по-добре. Тази статия има за цел да сравни 4 емоции, а именно гняв, щастие, тъга и неутралност, както са изразени от говорители на 4 различни езика (канадски френски, италиански, северноамерикански английски и германски - Берлински диалект) чрез съвременни методи за цифрова обработка на сигнали и конволюционни невронни мрежи.
Abstract:	Emotions play a central role in human interaction. Interestingly, different cultures have different ways to express the same emotion, and this motivates the need to study the way emotions are expressed in different parts of the world to understand these differences better. This paper aims to compare 4 emotions

	namely, anger, happiness, sadness, and neutral as expressed by speakers from 4 different languages (Canadian French, Italian, North American English and German - Berlin Deutsche) using modern digital signal processing methods and convolutional neural networks.
--	--

Публикация №9 – в том на конференция:

9. **Iliev, A.I., Anand A..** Huber Loss and Neural Networks Application in Property Price Prediction. Future of Information and Communication Conference (FICC), Springer series "Lecture Notes in Networks and Systems", 2023, ISSN:23673370, 23673389, **SJR (Scopus):0.151 Q4 (Web of Science)** [Линк](#)

Резюме:	В тази статия целим да изследваме пазара на недвижими имоти в Германия, и конкретно сме взели данни за Берлин и приложили различни напредничави методи на невронни мрежи и оптимизация. Винаги е трудно за хората да оценят на каква цена е най-добре за имот и в него участват различни категорийни и числови характеристики. И основното предизвикателство е изборът на модел, функция на загубата и персонализирането на невронната мрежа, за да се адаптира най-добре към данните за пазара. Разработихме проект, който може да се използва за прогнозиране на цената на имота в Берлин. Първо работихме върху извличането на текущата пазарна цена на имота от интернет чрез уеб скрепинг. След това направихме интензивен изследователски анализ на данните, подготвихме най-добрите данни за експерименти. След това изградихме 4 различни модела и работихме върху най-добрите функции на загубата, които могат да се приложат към нашия модел, и табулирахме квадратните и абсолютните грешки за същото. Тествахме нашия модел с текущите имоти на пазара, и резултатите от примерите са показани. Тази методология може да бъде приложена ефективно, и резултатите могат да бъдат използвани от хората, които се интересуват от инвестиции в недвижими имоти в Берлин, Германия.
Abstract:	In this paper we aim to explore the Real Estate Market in Germany, and particularly we have taken a dataset of Berlin and applied various advanced neural network and optimization techniques. It's always difficult for people to estimate what price is best for a property and there are various categorical and numerical features involved in it. And the main challenge is to choose the model, loss function and customize the neural network to best fit for the marketplace data. We have developed a project that can be used to predict the property price in Berlin. Firstly, we have worked on procuring the online current market price of the property by web scraping. Then we did intensive exploratory Data Analysis on it, prepared the best data for experiments. Then we build 4 different models and worked on the best loss functions which can suite our model and tabulated the mean squared and mean absolute errors for the same. We have tested our model with the current on the market properties, and the sample results are plotted. This methodology can be applied efficiently, and the results can be used by the people who are interested in investing in real estate in Berlin, Germany.

Публикация №10 – в том на конференция:

10. **Iliev, A.I., Rajasekaran R.K..** Hybrid Recommendation System. Future of Information and Communication Conference (FICC), Springer series "Lecture Notes in Networks and Systems", 2023, ISSN:23673370, 23673389, https://doi.org/10.1007/978-3-031-28076-4_1, **SJR (Scopus):0.151 Q4 (Web of Science)** [Линк](#)

Резюме:	В тази статия целим да изследваме и приложим хибридна система за препоръки, използвайки съвместно базирана и базирана на съдържание система за препоръки, която може да бъде приложена в пазара за ремонти. Разработихме няколко варианта на предложената архитектура на хибридната система за препоръки и сравнихме тяхното представяне с метрика, наречена степен на попадение. Освен това обсъдихме подготовката на данните и математическото обяснение на моделите за машинно обучение, които използвахме за хибридната система за препоръки. За да постигнем това, използвахме различни python рамки, използвани за обработка на естествен език и системи за препоръки.
Abstract:	In this paper we aim to explore and implement hybrid recommendation system using collaborative and content based recommendation system that can be applied to repairs marketplace. We developed multiple variants of proposed hybrid recommendation system architecture and compared their performance with a metric called hit rate. Furthermore, we discussed about data preparation and mathematical explanation of the machine learning models we used for hybrid recommendation system. To achieve this we used different python frameworks used for natural language processing and recommendation systems.

Публикация №11 – в том на конференция:

11. Iliev, A.I., Raksha A.. Text Regression Analysis for Predictive Intervals using Gradient Boosting. Future of Information and Communication Conference (FICC), Springer series "Lecture Notes in Networks and Systems", vol. 652, 2023, ISSN:23673370, 23673389, https://doi.org/10.1007/978-3-031-28073-3_18, SJR (Scopus):0.151 Q4 (Web of Science) [Линк](#)

Резюме:	В тази статия целим да изследваме анализа на текстови данни с помощта на различни методи за векторизация на данните и прилагане на методи за регресия. В момента имаме много техники за категоризация на текст, но съществуват много малко алгоритми, посветени на текстова регресия. Но много регресионни модели предсказват само една оценена стойност. Въпреки това е по-ефективно да се предсказва числов интервал, отколкото една единствена оценка, тъй като можем да бъдем по-сигурни, че истинската стойност ще бъде в рамките на интервала, отколкото да считаме оценената стойност за истинска. Целим да комбинираме и двата тези метода в тази статия. Целта на тази статия е да се съберат текстови данни, да се почистят и да се създаде текстов модел за регресия, за да намерим най-подходящия алгоритъм, който може да се използва във всяка ситуация, чрез използване на квантилна регресия.
Abstract:	In this paper, we aim to explore text data analysis with the help of different methods to vectorize the data and carry out regression methods. At present, we have a lot of text categorization techniques, but very few algorithms are dedicated to text regression. But many regression models predict only a single estimated value. However, it's more efficient to predict a numerical range than a single estimate, since we can be more sure that the true value will be within the range rather than considering the estimated value as the true value. We aim to combine both these techniques in this paper. The goal of this paper is to collect text data, clean the data, and create a text regression model to find the best-suited algorithm that could be used in any situation, through the use of quantile regression.

Публикация №12 – в том на конференция:

12. Iliev, A.I., Nimmala A., Rahiman R.A., Raju S., Chilakalanerpu S.. Fake Review Recognition using an SVM Model. Computing Conference (CC), Springer series "Lecture Notes in Networks and Systems", vol 652, 2023, ISSN:23673370, 23673389, https://doi.org/10.1007/978-3-031-37963-5_61, SJR (Scopus):0.151 Q4 (Web of Science) [Линк](#)

Резюме:	Тази статия има за цел да разработи система за разпознаване на фалшиви рецензии за продукти на Amazon чрез код на Python. Електронните търговски платформи формират все по-голяма част от световния бизнес пазар, а рецензиите играят ключова роля за определяне на автентичността и качеството на продукт. Рецензиите и коментарите играят жизненоважна роля в електронната търговия, допринасяйки за добрата или лошата репутация на продукта и съответно на продавача. Често на тези платформи се публикуват фалшиви рецензии, които в крайна сметка водят до повишаване или намаляване на рейтингите на тези продукти. За да насърчават продуктите си и да дискредитират конкурентите, безнравни участници наемат бот ферми, купуват легитимни акаунти и използват Sybil акаунти. Много фирми наемат хора да публикуват измислени положителни рецензии за техните стоки или услуги или несправедливо критични рецензии за стоките или услугите на техните конкуренти. Тази процедура предоставя неточна информация на потенциалните купувачи на тези стоки; затова ни трябва начин да разпознаем и премахнем лъжливите рецензии. Освен това тя пряко влияе върху надеждността както на продавача, така и на платформата. Чрез този проект ние се стремим да използваме техники SVM в Python скрипт, за да разграничим фалшивите рецензии от автентичните, които в крайна сметка определят качеството на продукта и надеждността на продавача.
Abstract:	This paper aims to develop a fake product review recognition system on Amazon using Python code. E-commerce websites form an ever-expanding portion of the global business market, and reviews play a vital role in deciding the authenticity and quality of a product. Reviews and comments play a vital role in the ecommerce domain contributing to a product's and in turn a seller's good or bad reputation. There are often fake reviews posted on these platforms which ultimately result in either raising or dropping the ratings of these products. To promote products and disparage rivals, immoral actors hire bot farms, purchase legitimate accounts, and use Sybil accounts. Many businesses employ people to post fictitious favourable reviews about their goods or services or unfairly critical reviews about the goods or services of their rivals. This procedure provides inaccurate information to potential buyers of these goods; hence we need a way to identify and eliminate bogus reviews. Moreover, It directly affects the reliability of both the seller as well as the platform. Through this project, we aim to utilize SVM techniques in Python script to differentiate fake reviews from authentic ones, which ultimately decide the quality of the product and reliability of the seller.

Публикация №13 – в том на конференция:

13. Iliev, A.I., Umapathy S.G.. Style Accessory Occlusion using CGAN with Paired Data. Computing Conference (CC), Springer series "Lecture Notes in Networks and Systems", vol 739, 2023, ISSN:23673370, 23673389, https://doi.org/10.1007/978-3-031-37963-5_70, **SJR (Scopus):0.151 Q4 (Web of Science)** [Линк](#)

Резюме:	Виртуалното премерване (Virtual try-on - VTO) е вълнуваща и завладяваща концепция, при която крайните потребители или клиентите могат да получат почти реалистичен опит от продукта или артикула, който ги интересува. В областта на VTO, Допълнена реалност (AR) и Виртуална реалност (VR) допринасят активно. Въпреки това тези технологии изискват сложно оборудване като VR очила или изчислително мощни смартфони, за да предоставят добър опит с AR. За да решим този проблем с изискването на скъпо оборудване, Машинното обучение (ML) и Изкуствен интелект (AI) допринасят за изследвания и приложения в VTO. В тази връзка може да се разгледа Генеративното моделиране, особено за VTO на електронна търговска платформа. Генеративните модели са нестабилни. За да решим този проблем с нестабилността и да създадем
---------	---

	уникален опит с генеративни модели в областта на VTO, се използват данни в двойки, за да се оклудира стилът аксесоар като шапка или кеп на портретни изображения.
Abstract:	Virtual try-on (VTO) is an exciting and captivating concept where end users or customers can gain a near-realistic experience of the product/ item they are interested in. In the field of VTO, Augmented Reality (AR) and Virtual Reality (VR) actively contribute. However, these technologies require sophisticated hardware equipment like a VR Glass /VR Headgear or computationally capable smartphones to provide a good experience with AR. To solve this issue of requiring expensive hardware, Machine Learning (ML) and Artificial Intelligence (AI) have contributed to research and applications in VTO. To this end, Generative Modelling can be considered, especially for VTO on an e-commerce platform. Generative models are volatile. To solve this instability issue and to create a unique experience with generative models for the VTO domain, paired-wise data is used to occlude a style accessory like Hat/ Cap on portrait images.

Публикация №14 – в научно списание:

14. Iliev A.I., D. Singh, S. Chittiyala. Bias Detection and Mitigation in Large Language Models for Code Generation. IEEE Internet of Things, IEEE, 2025, DOI:10.1109/IJOT.2025.3636829, **SJR (Scopus):2.483 Q1 - оглавява ранглистата (Web of Science) [Линк](#)**

Резюме:	Големите езикови модели (LLM) са основни инструменти в съвременното разработване на софтуер, като значително ускоряват кодирането, рационализират отстраняването на грешки и осигуряват по-широк достъп до усъвършенствани алгоритмични функции. Тяхното въздействие се простира в нововъзникващи области като Интернет на нещата (IoT), където генерираните от изкуствен интелект код все повече управлява поведението на периферните устройства и интеграцията на интелигентни системи. LLM обаче често наследяват и разпространяват пристрастия, присъстващи в техните обучителни данни. Тези пристрастия могат да се появят в генерираните от изкуствен интелект код, което води до етично проблематични, алгоритмично изкривени или дори опасни резултати – особено тревожни в критични за безопасността и социално чувствителни IoT среди. Това изследване изследва пристрастията в генерираните от LLM код и предлага многостранен подход, комбиниращ контекстуален анализ на код, контрафактуално инженерство на бързи отговори и обучение с подсилване с човешка обратна връзка (RLHF), за да открие и смекчи такива пристрастия. Показваме как тези техники разкриват скрити пристрастия в конвенциите за именуване, логиката на решенията и моделите на кодиране и демонстрираме как RLHF може да намали пристрастията в мащаб, като същевременно запазва функционалността – постигайки 49% намаление на пристрастията в оценяваните бенчмаркове. Нашите открития подчертават необходимостта от строги рамки за оценка на пристрастията, внимателно куриране на данните и прозрачни работни процеси за разработка. Разглежда се компромисът между намаляване на пристрастията и прецизност на кода, с последици за етичното използване на изкуствен интелект в IoT системи и други софтуерни контексти с високи залози. Насърчават се по-нататъшни изследвания в областта на хибридните техники за премахване на пристрастията, идентифицирането на междусекторни пристрастия, енергийно ефективното моделиране и разработването на стандартизирани етични практики за кодиране. Ключови термини - IoT, Модели с големи езици, Откриване на пристрастия, Генериране на код, Контрафактуални подкани, Обучение с подсилване, RLHF (Reinforcement Learning - ограничено високочестотно обучение), Етичен изкуствен интелект, Алгоритмична справедливост, Контекстуален анализ на кода.
Abstract:	Large Language Models (LLMs) are essential tools in modern software development, significantly accelerating coding, streamlining debugging, and

	enabling broader access to advanced algorithmic features. Their impact extends into emerging domains like the Internet of Things (IoT), where AI-generated code increasingly drives edge device behavior and smart system integration. However, LLMs often inherit and propagate biases present in their training data. These biases can surface in AI-generated code, leading to ethically problematic, algorithmically skewed, or even unsafe outcomes—particularly concerning in safety-critical and socially sensitive IoT environments. This research investigates bias in LLM-generated code and proposes a multi-faceted approach combining Contextual Code Analysis, Counterfactual Prompt Engineering, and Reinforcement Learning with Human Feedback (RLHF) to detect and mitigate such biases. We show how these techniques reveal hidden biases in naming conventions, decision logic, and coding patterns, and demonstrate how RLHF can reduce bias at scale while preserving functionality – achieving a 49% reduction in bias across evaluated benchmarks. Our findings emphasize the need for rigorous bias evaluation frameworks, careful data curation, and transparent development workflows. The tradeoff between bias reduction and code precision is explored, with implications for the ethical use of AI in IoT systems and other high-stakes software contexts. Further research is encouraged in hybrid debiasing techniques, intersectional bias identification, energy-efficient modeling, and the development of standardized ethical coding practices. Index Terms— IoT, Large Language Models, Bias Detection, Code Generation, Counterfactual Prompts, Reinforcement Learning, RLHF, Ethical AI, Algorithmic Fairness, Contextual Code Analysis
--	--

Публикация №15 – в том на конференция:

15. **Iliev A.I.**, Nithin Jayagovindan. Intelligent Autonomous Vehicles (IAV) using Artificial Intelligence Focusing on Perception. IS&T International Symposium on Electronic Imaging Science and Technology, 37, 3, 2025, DOI:<https://doi.org/10.2352/ei.2025.37.3.mobmu-324,324-327>. **SJR 2024 (Scopus):0.222 Международно академично издателство**
[Линк](#)

Резюме:	Изкуственият интелект (ИИ) допринася значително за разработването на автономни превозни средства по несравним начин. Тази статия очертава техники и алгоритми за внедряване на Интелигентни автономни превозни средства (ИАВ), използващи ИИ алгоритми за възприемане на трафика, вземане на решения и контрол в автономни превозни средства чрез обединяване на откриване на пътни сценарии, откриване на пътни ленти, семантична сегментация, откриване на пешеходци и класификация и откриване на пътни знаци. Съвременните компютърно зрение и алгоритми, базирани на дълбоки невронни мрежи, позволяват анализ в реално време на различни данни за превозните средства чрез изкуствен интелект. Динамиката на превозното средство се конституира чрез ИИ в системите за управление на превозните средства за повишена безопасност и ефективност, за да се гарантира, че те са оптимизирани с времето. Освен това, статията ще обсъди и предизвикателствата и възможните бъдещи насоки, ще подчертае как ИИ има потенциала да управлява автономните превозни средства към по-безопасни и по-надеждни, както и интелигентни транспортни системи. Това е надеждата на бъдещето, при което мобилността е интелигентна, устойчива и достъпна с комбинацията от ИИ с автономни превозни средства.
Abstract:	Artificial Intelligence (AI) contributes significantly to the development of autonomous vehicles in an unmatched way. This paper outlines techniques and algorithms for the implementation of Intelligent Autonomous vehicles (IAV) leveraging AI algorithms for traffic perception, decision-making and control in autonomous vehicles through merging traffic scenario detection, traffic lane detection, semantic segmentation, pedestrian detection, and traffic sign classification and detection. The modern computer vision and deep neural networks-based algorithms enable the real-time analysis of different vehicle data

	through artificial intelligence. The vehicle dynamics are constituted through AI in vehicle control systems for increased safety and efficiency to ensure that they are optimized with time. In addition, the paper will also discuss challenges and possible future directions, underscore how AI has the potential of driving autonomous vehicles towards safer and more reliable as well as intelligent transportation systems. This is the hope of the future whereby mobility is intelligent, sustainable, and accessible with the combination of AI with autonomous vehicles.
--	---

Публикация №16 – в научно списание:

16. Chikalanov, A., Kirilov, L., Kovatcheva, E., Nikolov, R., Shoikova, E., **Iliev, A.I.**, Gotsev, L.. A Model of Big Data Architecture on the Base of FIWARE Components. Proceedings of the Bulgarian Academy of Sciences, 9, 76, Bulgarian Academy of Sciences, 2023, ISSN:1310–1331, DOI:<https://doi.org/10.7546/CRABS.2023.09.10>, 1393-1401. **SJR (Scopus):0.16, JCR-IF (Web of Science):0.329 Q3 (Scopus)** [Линк](#)

Резюме:	Интензивното развитие на хардуерни и софтуерни инструменти предоставя много възможности за решаване на различни реални задачи, като например работа с големи масиви от данни. Това от своя страна е мощен стимул и възможност за разработването на различни подходи и инструменти за тяхното решаване. След обсъждане на няколко добре познати архитектури на BigData е проектиран модел на архитектура на BigData. Архитектурата е организирана в три раздела: периферен раздел, раздел за обработка и съхранение и раздел за приложения. Моделът ще бъде реализиран с помощта на FIWARE компоненти.
Abstract:	The intensive development of hardware and software tools provides many opportunities for solving various real-world tasks, such as working with large data sets. This, in turn, is a powerful incentive and opportunity for the development of different approaches and tools for solving them. After discussing several well-known BigData architectures a model of a BigData architecture is designed. The architecture is organized in three sections: edge section, processing-and-storage section and application section. The model will be implemented using FIWARE components.

Публикация №17 – в том на конференция:

17. **Iliev A.I.**, A.. Discovery of Deepfakes in Art. Digital Presentation and Preservation of Cultural and Scientific Heritage, 15, BAN, 2025, DOI:<https://doi.org/10.55630/dipp.2025.15.5>, 55-64. **SJR (Scopus):0.243 SJR, непопадащ в Q категория (Web of Science)** [Линк](#)

Резюме:	Разпространението на deepfakes (дълбочинни фалшиви изображения) – генерирани или манипулирани от изкуствен интелект медии – трансформира пейзажа на съвременното изкуство. Дълбоко генеративните модели, включително GAN (глобални асоциативни мрежи), VAE (видео-ориентирани функционалности), дифузионни модели и трансформатори, позволиха на творците да изследват нови творчески сфери, като едновременно с това повдигат критични въпроси относно автентичността, етиката и откриването. Тази статия представя цялостен анализ на deepfake технологиите в пет ключови медийни модалности: изображение, видео, текст, реч и музика. Разглеждаме архитектурите, които позволяват създаването на съдържание, и най-съвременните техники, използвани за откриване. Освен това, оценяваме точността, устойчивостта и практическото приложение на откриването, като включваме диаграми, сравнителни таблици и формули за ефективност. Тази работа има за цел да предостави балансирана перспектива върху възможностите и предизвикателствата, породени от синтетичните медии в художествената област.
---------	---

Abstract:	The proliferation of deepfakes AI-generated or manipulated media has transformed the landscape of contemporary art. Deep generative models, including GANs, VAEs, diffusion models, and Transformers, have enabled artists to explore new creative realms while simultaneously raising critical questions around authenticity, ethics, and detection. This paper presents a comprehensive analysis of deepfake technologies across five key media modalities: image, video, text, speech, and music. We examine the architectures that enable content creation, and the state-of-the-art techniques used for detection. Further, we evaluate detection accuracy, robustness, and practical implementation, incorporating diagrams, comparative tables, and performance formulas. This work aims to provide a balanced perspective on the opportunities and challenges posed by synthetic media in the artistic domain.
-----------	--

Публикация №18 – в том на конференция:

18. Mansoor, N., **Iliev A.I.**, (2024). DeepFake Detection Using Deep Learning. In: Arai, K. (eds) Intelligent Computing. SAI 2024. Lecture Notes in Networks and Systems, vol 1018. Springer, Cham. , Print ISBN978-3-031-62272-4, Online ISBN978-3-031-62273-1, https://doi.org/10.1007/978-3-031-62269-4_14, SJR (Scopus):0.166, Q4 (Web of Science), [Линк](#)

Резюме:	Произхождащ от 2017 г., Deepfakes използва алгоритми за изкуствен интелект, по-специално генеративни състезателни мрежи (GAN), за генериране на подвеждащи видеоклипове и изображения. Въпреки широкообхватните приложения на GAN, възникват етични опасения поради потенциалната им злоупотреба. Бързото развитие на технологиите в мултимедията доведе до появата на злонамерени дейности, като например манипулация с DeepFake в политиката и развлеченията. Тъй като манипулираните визуализации стават все по-реалистични, необходимостта от подход за откриване на DeepFake става от съществено значение. Изследването използва две различни архитектури на конволюционни невронни мрежи за класификация, подчертавайки нарастващото значение на откриването на deepfake в съвременното общество. В това изследване са разработени техники за откриване на DeepFake чрез изграждане на два CNN модела, базирани на архитектури ResNet50 и DenseNet121. Направен е сравнителен анализ за двата модела въз основа на параметрите за точност и ефективност на F1-оценка. Експерименталните резултати показват, че всички CNN модели, използвани в изследването, постигат добри резултати с точност от 86% и F1-оценка от 0,87 съответно.
Abstract:	Originating in 2017, Deepfakes utilizes artificial intelligence algorithms, particularly Generative Adversarial Networks (GANs), to generate deceptive videos and images. Despite the wide-ranging applications of GANs, ethical concerns emerge due to their potential misuse. The rapid advancement of technology in multimedia has led to the emergence of malicious activities, such as DeepFake manipulation in politics and entertainment. As manipulated visuals become increasingly realistic, the need for a DeepFake detection approach becomes essential. The research employs two distinct Convolutional Neural Network architectures for classification, emphasizing the escalating importance of deepfake detection in contemporary society. In this research, DeepFake detection techniques are developed by building two CNN models based on ResNet50 and DenseNet121 architectures. A comparative analysis is done for the two models based on accuracy and F1-score efficiency parameters. Experimental results indicate that all CNN models employed in the study achieve good performance with an accuracy of 86% and F1-score of 0.87 respectively.

19. **Iliev A.I.**, Mansoor, N.. Explainable AI for Deepfake Detection. Applied Sciences, 15(2), 725, MDPI, 2025, DOI:<https://doi.org/10.3390/app15020725>, SJR (Scopus):0.555, JCR-IF (Web of Science):2.7 Q2 (Web of Science), [Линк](#)

Резюме:	Нарастващият технологичен напредък доведе до опасения относно злоупотребата с него в политиката и развлеченията, което направи надеждните методи за откриване от съществено значение. Това проучване въвежда техника за откриване на deepfake, която подобрява интерпретируемостта, използвайки алгоритъма за мрежова дисекция. Изследването се състои от два етапа: (1) откриване на фалшиви изображения с помощта на усъвършенствани конволюционни невронни мрежи като ResNet-50, Inception V3 и VGG-16 и (2) прилагане на алгоритъма за мрежова дисекция, за да се разберат вътрешните процеси на вземане на решения в моделите. Производителността на конволюционните невронни мрежи се оценява чрез F1-оценки, вариращи от 0,8 до 0,9, демонстрирайки тяхната ефективност. Чрез анализ на чертите на лицето, научени от моделите, това проучване предоставя обясними резултати за класифициране на изображенията като реални или фалшиви. Тази интерпретируемост е от решаващо значение за разбирането как работят моделите за откриване на deepfake. Въпреки че съществуват многобройни модели за откриване, те често нямат прозрачност в процесите си на вземане на решения. Това изследване запълва тази празнина, като предлага прозрения за това как тези модели разграничават реални от манипулирани изображения. Констатациите подчертават важността на интерпретируемостта в дълбоките невронни мрежи, осигурявайки по-добро разбиране на техните йерархични структури и процеси на вземане на решения.
Abstract:	The surge in technological advancements has resulted in concerns over its misuse in politics and entertainment, making reliable detection methods essential. This study introduces a deepfake detection technique that enhances interpretability using the network dissection algorithm. This research consists of two stages: (1) detection of forged images using advanced convolutional neural networks such as ResNet-50, Inception V3, and VGG-16, and (2) applying the network dissection algorithm to understand the models' internal decision-making processes. The CNNs' performance is evaluated through F1-scores ranging from 0.8 to 0.9, demonstrating their effectiveness. By analyzing the facial features learned by the models, this study provides explainable results for classifying images as real or fake. This interpretability is crucial in understanding how deepfake detection models operate. Although numerous detection models exist, they often lack transparency in their decision-making processes. This research fills that gap by offering insights into how these models distinguish real from manipulated images. The findings highlight the importance of interpretability in deep neural networks, providing a better understanding of their hierarchical structures and decision processes.

20. Iliiev, A.I., Malvika Panwar M. AI in Cryptocurrency. Future of Information and Communication Conference (FICC 2023), Springer series "Lecture Notes in Networks and Systems", Vol. 652, 2023, ISBN: 978-3-031-28072-6, Online ISBN: 978-3-031-28073-3, ISSN:23673370, 23673389, https://doi.org/10.1007/978-3-031-28073-3_14, pp. 205-217, **SJR (Scopus): 0.171 Q4 (Scopus), [Линк](#)**

Резюме:	Това проучване изследва предсказуемостта на шест значими криптовалути за предстоящите два дни, използвайки техники с изкуствен интелект, т.е. алгоритми за машинно обучение, като „случайна гора“ и „градиент“, за прогнозиране на цената на тези шест криптовалути. Изследването ни показва, че машинното обучение може да се разглежда като средство за прогнозиране на цените на криптовалутите. Системата за машинно обучение се учи от минали данни, изгражда модели за прогнозиране и прогнозира резултата за нови данни, когато ги получи. Точността на прогнозирания резултат се влияе от количеството данни, тъй като колкото повече данни има, толкова по-добре моделът може да предвиди резултата. Резултатите показват, че с метриката за ефективност „accuracy score“,
---------	---

	която използвахме за това изследване, успяхме да изчислим точността на алгоритмите и да установим, че и двата алгоритъма – съответно „random forest“ и „gradient boosting“ – се представят добре за криптовалути като Solana (98,07%, 98,14%), Binance (96,56%, 96,85%) и Ethereum (96,61%, 96,60%), с изключение на Tether (0,38%, 12,35%) и USD coin (–0,59%, 1,48%). Резултатите показват, че и двата алгоритъма работят ефективно с повечето криптовалути, което може да бъде допълнително подобро чрез използване на алгоритми за дълбоко обучение като ANN, RNN или LSTM.
Abstract:	This study investigates the predictability of six significant cryptocurrencies for the upcoming two days using AI techniques i.e., machine learning algorithms like random forest and gradient for predicting the price of these six cryptocurrencies. The study presents to us that machine learning can be seen as a medium to predict the prices of cryptocurrencies. A machine learning system learns from past data, constructs prediction models, and predicts the output for new data whenever it gets it. Predicted output's accuracy is influenced by the quantity of data since the more data there is, the better the model can predict the output. The results show that with the accuracy score performance metric, which we employed for this study, we were able to calculate the accuracy of the algorithms and find that both algorithms random forest and gradient boosting respectively performed well for the cryptocurrencies such as Solana (98.07%,98.14%), Binance (96.56%, 96.85%), and Ethereum (96.61%, 96.60%)), except for Tether (0.38%, 12.35%) and USD coin (–0.59%, 1.48%), the results demonstrate that both algorithms work effectively with most cryptocurrencies which can be further increased by using deep learning algorithms like ANN, RNN or LSTM.

21. Mansoor N., **Iliev, A.I.**, Artificial Intelligence in Forensic Science. Future of Information and Communication Conference (FICC 2023), Springer series "Lecture Notes in Networks and Systems", Vol. 652, 2023, ISBN: 978-3-031-28072-6, Online ISBN: 978-3-031-28073-3, ISSN:23673370, 23673389, https://doi.org/10.1007/978-3-031-28073-3_11, **SJR (Scopus):0.171 Q4 (Scopus)**, [Линк](#)

Резюме:	Изкуственият интелект е бързо развиваща се технология, която се използва в различни индустрии. Този доклад предоставя общ преглед на някои приложения на изкуствения AQ1 интелект, използвани в криминалистиката. Той също така описва последните изследвания в различни области на криминалистиката и ние внедрихме модел за случай на употреба в дигиталната криминалистика. За разлика от традиционната криминалистична идентификация, случаите на употреба на изкуствен интелект в областта на криминалистиката са многобройни. Този доклад описва списък с приложения на изкуствен интелект, които се използват в криминалистиката, някои от последните проучвания, проведени в тази област, и внедряването на модела за случай на употреба на изкуствен интелект в дигиталната криминалистика.
Abstract:	Artificial intelligence is a rapidly evolving technology that is being used in a variety of industries. This report provides an overview of some artificial AQ1 intelligence applications used in forensic science. It also describes recent research in various fields of forensics, and we implemented a model for a use case in digital forensics. Unlike traditional forensic identification, the use cases of Artificial intelligence are numerous in the field of forensic science. This report describes a list of Artificial intelligence applications which are used in forensic science, some of the recent studies conducted in this field, and the implementation of the model for an AI use case in digital forensics.

22. Balaji, K., **Iliev A.I.**, Comprehensive Feature Selection Methods for Predicting Diabetes and Stroke Risk. In: Arai, K. (eds) Intelligent Computing (SAI 2024), Lecture Notes in Networks and Systems, Vol. 1019. Springer, Cham, ISBN: 978-3-031-62272-4, Online ISBN: 978-3-031-62273-1, https://doi.org/10.1007/978-3-031-62273-1_9, pp. 128–145, **SJR (Scopus):0.166, Q4 (Scopus)**, [Линк](#)

Резюме:	Изследователската задача, озаглавена „Всеобхватни методи за подбор на характеристики за прогнозиране на риска от диабет и инсулт сред женени мъже/жени“, изследва връзката между различните рискови фактори и превъзходството на AQ1 диабета и инсулта сред женените хора. Основната
----------------	--

	цел на проекта е да се открият важните рискови фактори, свързани с диабета и инсулта AQ2 сред хора, които са женени, разведени или овдовели. За да се постигне това, се използва стабилен процес за избор на характеристики, за да се идентифицират най-влиятелните фактори сред широк набор от променливи. След задълбочен преглед на литературата, проектът продължава с подготовката на данните. Това включва избора на 21 релевантни характеристики въз основа на познания в областта и анализ на данните. Освен това, наборът от данни преминава през цялостен процес на почистване, за да се гарантира качеството и точността на данните. За да се изберат най-подходящите характеристики за диабет и инсулт, почиственият набор от данни е подложен на различни методи за избор на характеристики, включително някои нови и иновативни методи, както и някои добре познати методи. Тези характеристики са обучени с помощта на различни алгоритми за машинно обучение и класификатори, а най-подходящите характеристики, влияещи върху диабета и инсулта, са оценени въз основа на показателите за ефективност. След това е изграден модел за прогнозиране както за диабет, така и за инсулт, използвайки най-подходящите характеристики, които го засягат. Иновативният характер на методологията, интегриращ различни техники за подбор на характеристики, допринася за развитието на областта и повишава точността и надеждността на предсказуемите модели. Резултатите от изследването отварят врати за по-нататъшни изследвания и осигуряват значителен принос в областта на здравеопазването.
Abstract:	The studies task titled “Comprehensive Feature SelectionMethods for Predicting Diabetes and Stroke Risk Among Married Male/Female Individuals” explores the relationship among diverse danger elements and the superiority of AQ1 diabetes and stroke among married people. The undertaking’s primary aim is to discover the important thing threat elements related to diabetes and stroke AQ2 amongst folks who are married, divorced, or widowed. To reap this, a robust feature choice process is hired to identify the most influential factors among an extensive set of variables. After a thorough literature review, the project proceeds with data preparation. This involves the selection of 21 relevant features based on domain knowledge and data analysis. Additionally, the dataset undergoes a comprehensive cleaning process to ensure data quality and accuracy. To select the most relevant features for diabetes and stroke the cleaned dataset was made to undergo different feature selection methods including some new and innovative methods along with some well-known methods. These features were trained using different machine learning algorithms and classifiers and the most relevant features affecting diabetes and stroke were evaluated based on the performance metrics. Then a prediction model for both diabetes and stroke were built using the most relevant features affecting it. The methodology’s innovative nature, integrating various feature selection techniques, contributes to the advancement of the field and enhances the accuracy and robustness of predictive models. The research findings open doors for further investigations and provide a significant contribution to the healthcare domain.

23. Murthy, H. M. K., **Iliev A. I.** Secure Curation and Reuse of Digital Scientific Data Using Retrieval-Augmented Generation Systems. Digital Presentation and Preservation of Cultural and Scientific Heritage, Vol. 15, Institute of Mathematics and Informatics - BAS, 2025, ISSN: 1314-4006, Online ISSN: 2535-0366, DOI:<https://doi.org/10.55630/dipp.2025.15.17>, pp. 187-196. **SJR (Scopus):0.217, индексиран в Scopus, Web of Science, Линк**

Резюме:	Това проучване изследва сигурното и интелигентно куриране на цифрови научни данни, използвайки системи за генериране на добавени данни (RAG). Фокусирайки се върху подходите за запазване на поверителността и повторното използване на структурирани набори от данни в здравеопазването, ние предлагаме модели за откриване на измами, подобряване на семантичната интерпретация, персонализация и сигурен достъп до цифрови знания в критични сектори.
---------	---

	В тази статия изследваме как безопасно да се пренасочват структурирани набори от данни за здравеопазването, базирани на RAG архитектурата. Също така се отчитат проблемите, свързани с регулирането на семантиката, интелигентното търсене на информация и приемането на подходящи етични подходи към изкуствения интелект. В този смисъл, тази работа обогатява темата на конференцията, като прилага надежден изкуствен интелект в културното наследство и областите с интензивно използване на данни, за да се улесни устойчивото и сигурно куриране на цифрови научни активи.
Abstract:	This study explores secure and intelligent curation of digital scientific data using Retrieval-Augmented Generation (RAG) systems. Focusing on privacy-preserving approaches and reuse of structured healthcare datasets, we propose models for fraud detection, enhancing semantic interpretation, personalization, and secure access to digital knowledge assets in critical sectors. In this paper, we explore how to safely repurpose structured healthcare datasets based on the RAG architecture. It also acknowledges issues of semantics regulation, smart search information, and the adoption of proper ethical approaches to artificial intelligence. On that note, this work enriches the conference's theme by applying trustworthy artificial intelligence into heritage and data-intensive domains to facilitate sustainable and secure curation of digital scientific assets.