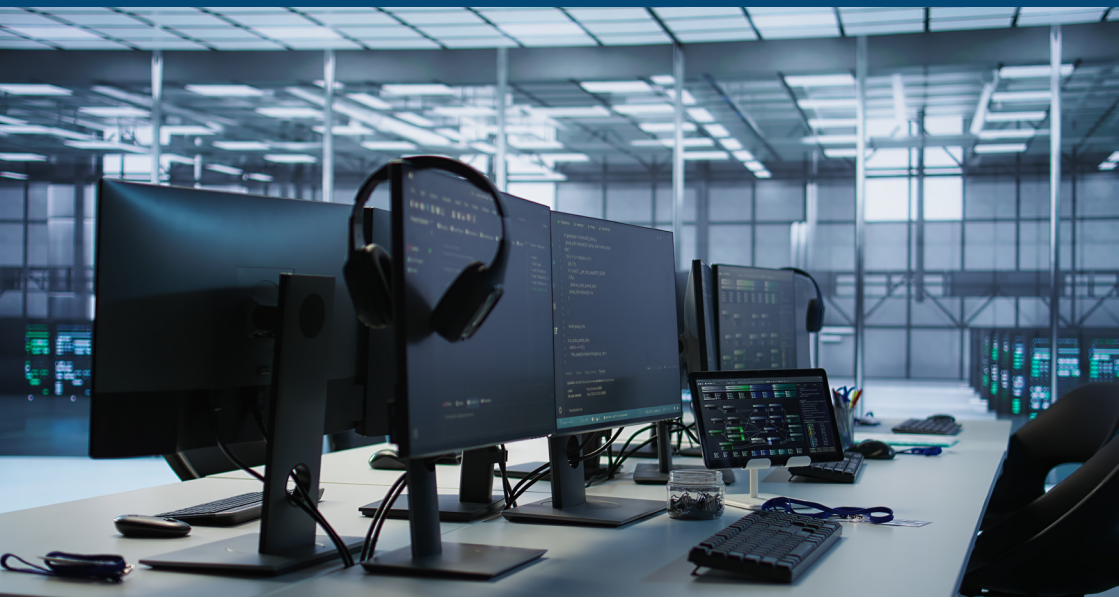
DAMSS
DATA ANALYSIS METHODS FOR SOFTWARE SYSTEMS



16th Conference on

DATA ANALYSIS METHODS for Software Systems

November 27–29, 2025

Druskininkai, Lithuania,
Hotel "Europa Royale"

https://www.mii.lt/DAMSS

LITHUANIAN COMPUTER SOCIETY
VILNIUS UNIVERSITY, INSTITUTE OF DATA SCIENCE AND DIGITAL TECHNOLOGIES
LITHUANIAN ACADEMY OF SCIENCES

16th Conference on

DATA ANALYSIS METHODS
for Software Systems

November 27–29, 2025

Druskininkai, Lithuania, Hotel "Europa Royale"
https://www.mii.lt/DAMSS

VILNIUS UNIVERSITY PRESS
Vilnius, 2025

Contacts:

Dr. Jolita Bernatavičienė
jolita.bernataviciene@mif.vu.lt
Prof. Olga Kurasova
olga.kurasova@mif.vu.lt
Tel. (+370 5) 2109 315

https://doi.org/10.15388/DAMSS.16.2025
ISBN 978-609-07-1200-9 (digital PDF)

Preface

The 16th International Conference on Data Analysis Methods for Software Systems (DAMSS-2025) is taking place in Druskininkai, Lithuania. As in previous years, it is organized at the same venue and time, fostering active scientific exchange among participants.

First organized in 2009 with 16 papers, the event began as a workshop under an international project and has evolved into a prominent conference. Founded by the Institute of Mathematics and Informatics, now the Institute of Data Science and Digital Technologies at Vilnius University, and supported by the Lithuanian Academy of Sciences and the Lithuanian Computer Society, the conference maintains high academic and organizational standards through long-term collaboration. The conference has grown into one of Lithuania's leading events in computer science and computer engineering, promoting interdisciplinary research and fostering genuine innovation.

This year, the conference features 90 presentations representing nine countries, with a total of 132 registered participants from Armenia, France, Italy, Lithuania, Poland, Romania, Slovenia, Spain, and Sweden. With participation from seven Lithuanian universities, science and higher education institutions, the conference serves as the main annual meeting for the country's computer science community. The conference's international character fosters project collaboration, facilitates the exchange of knowledge, and supports the development of innovative ideas on a global scale.

The annual organization of DAMSS facilitates the rapid exchange of new ideas within the scientific community and is unique in providing opportunities for practical collaboration. Researchers from Lithuania and abroad, along with representatives from the industry and public sectors, are encouraged to develop joint projects, apply research to practical applications, and address both business and societal challenges, thereby aligning research with market needs and industry competencies.

A key highlight of the conference is the participation of young scientists. Doctoral and Master's students from Lithuania and other countries

can present their research, engage in discussions, and gain valuable experience. This year, we have a special student presentation session.

Traditionally, most presentations are delivered in poster format, providing an interactive platform for discussion and feedback. Oral sessions are mainly for keynote speakers. This year, we are pleased to invite six keynote speakers.

Five IT companies and the Research Council of Lithuania support DAMSS-2025. This also demonstrates the relevance of the conference topics to the business sector.

Topics covered by the conference include applied mathematics, artificial intelligence, big data, bioinformatics, blockchain technologies, business rules management, cybersecurity, data science, deep learning, high-performance computing, data visualization, machine learning, medical informatics, modelling educational data, ontological engineering, optimization, quantum computing, signal and image processing.

This book compiles the presentations delivered at the DAMSS-2025 conference.

PARTNER

International Federation for Information Processing
ifip.org

DAMSS 2025 SUPPORTED BY:

General Sponsors

Neurotechnology
neurotechnology.com

Novian
novian.lt

Research Council of Lithuania
lmt.lt

3RTechnology
3rt.lt

Main Sponsor

Visoriai Information Technology Park (VITP)
vitp.lt

Sponsor

Asseco Lithuania
asseco.lt

Yield Estimation in Hydroponic Sprouts from High-Frequency RGB Imaging

Danylo Abramov[1,3], Eimantas Zaranka[1,3], Rūta Juozaitienė[1,3], Andrius Grigas[2], Tomas Krilavičius[1,3]

[1] Vytautas Magnus University
[2] Vytautas Magnus University Agriculture Academy
[3] Centre for Applied Research and Development

danylo.abramov@vdu.lt

Hydroponic fodder production requires reliable and cost-effective yield monitoring to optimize irrigation and lighting. This study investigates whether high-resolution RGB imaging, without additional multispectral sensors, can provide accurate yield estimation for short-cycle wheat sprouts. Trays were continuously weighed and photographed every five minutes in a controlled environment over multiple 7-day cycles. Images were preprocessed in HSV color space with white suppression, segmented using the Segment Anything Model, and transformed into interpretable features describing greenness, coverage, texture, and color balance. Several machine learning models were compared, including Linear Regression, Ridge, Lasso, Support Vector Regression (SVR), Random Forest (RF), tuned Random Forest, Extreme Gradient Boosting (XGB), LightGBM (LGBM), and Multilayer Perceptron (MLP). Results demonstrated that tree-based and boosting models markedly outperformed linear baselines. Random Forest achieved the best mean absolute error (MAE = 0.0060) and symmetric mean absolute percentage error (sMAPE = 0.208), while LightGBM produced the lowest root mean square error (RMSE = 0.0143). XGB also showed competitive accuracy, whereas SVR and MLP performed moderately, and linear models lagged behind (MAE > 0.08). Feature-importance analysis revealed that greenness and colour indices, such as AGI, mean_H, Proxy_NDVI, and mean_R, were the dominant predictors across all models. The findings confirm that accurate yield estimation is feasible using RGB-only imaging, providing a practical solution for real-time greenhouse monitoring. This approach can be further extended to other crops and integrated into smart greenhouse control systems.

Quantum Machine Learning for Image Classification in Healthcare: Algorithms, Applications, and Future Prospects

Sasan Ansarian, Remigijus Paulavičius, Ernestas Filatovas

Institute of Data Science and Digital Technologies
Vilnius University

sasan.ansarian@mif.vu.lt

One of the incentives for quantum machine learning (QML) is its potential to achieve significant computational advantages, such as exponential speedups and improved accuracy, over classical machine learning (ML) by using quantum phenomena. QML goals to solve problems too complex for classical computers, enabling richer data representations, more efficient algorithms, and the ability to tackle large, high-dimensional datasets. This could revolutionise in fields of medical image analysis, enabling through advanced simulations, improving the accuracy and speed of analysis, and creating more precise predictive models for disease spread and patient outcomes. In this work, we systematically review studies published between 2020 and 2025 that apply QML, identified through database searches and screened using PRISMA guidelines. The review analyses the evolution of QML datasets and evaluation strategies used in this field, highlighting their strengths and limitations. Preliminary findings illustrate that most applications remain at the proof-of-concept step and are constrained by limited quantum hardware and dataset availability. To advance the field toward clinical relevance, this work identifies key research priorities, including the development of standardised, quantum-ready image datasets, unified evaluation protocols, hybrid quantum–classical model optimisation, and the co-design of algorithms and hardware. Focusing on these areas will make it clear that QML provides valuable benefits for diagnostics, personalised care, and healthcare operations.

Acknowledgements. This research was funded by the Research Council of Lithuania under Grant Agreement No. S-ITP-25-5.

Neural Dynamics Revealed by Power-Law Analysis of sEEG Across Brain Regions

Karolina Armonaite[1], Livio Conti[2], Camillo Porcaro[3], Tomas Ruzgas[1], Franca Tecchio[4]

[1] Department of Applied Mathematics
Kaunas University of Technology
[2] National Institute of Nuclear Physics (INFN)
Section of Tor Vergata, Rome, Italy
[3] Department of Neuroscience & Padova Neuroscience Center
University of Padova, Padova, Italy
[4] Laboratory of Electrophysiology for Translational Neuroscience
Institute of Cognitive Sciences and Technologies
National Research Council, Rome, Italy

karolina.armonaite@ktu.lt

The brain is a complex system that exhibits self-organisation, emergent dynamics, and signatures of scale-invariance. Evidence suggests that neuronal activity displays scale-free properties, which can be quantitatively captured by analysing spectral power distributions. In this study, we investigate multivariate time series obtained from stereotactic EEG (sEEG) recordings across 36 anatomically defined brain regions under resting conditions. Our analysis focused on scale-free dynamics, quantified by estimating power spectral densities (PSD) using the Fast Fourier Transform (FFT) on 10-second segments tapered with a Hamming window and modelling them in log–log space. Prior to transformation, all signals were filtered using low- and high-pass Butterworth filters to reduce prominent oscillatory components. Specifically, a low-pass filter with an 8 Hz cutoff was used to isolate the low-frequency range (0.5–4 Hz), and a high-pass filter with a 33 Hz cutoff was applied to isolate the high-frequency range (34–80 Hz). We then tested for power-law scaling of the form P(f)∝f^(-β) where P(f) denotes spectral power at frequency f, and β is the scaling exponent. Power-law behavior was evaluated across 36 brain regions in low- and high-frequency ranges, by log-transforming frequencies and mean PSD values. Linear trends (β) were estimated via the weighted least squares method, with weights

defined as w_i=1/(Δy_i^2), where Δy_i represents the uncertainty in y_i=log⁡(PSD(f_i)). Model fitting revealed two different frequency regimes characterized by power-law scaling: a low-frequency interval and a high-frequency interval. Statistical comparison of fitted exponents β indicated significant differences between cortical and subcortical regions in the high-frequency band.

High-Performance Artificial Intelligence: Bridging Advanced Computing and AI

Hrachya Astsatryan, Yuri Shoukourian, Vladimir Sahakyan, Levon Aslanyan, Hasmik Sahakyan

Institute for Informatics and Automation Problems
National Academy of Sciences of the Republic of Armenia

hrach@sci.am

Where high-performance computing (HPC) meets artificial intelligence (AI), new scientific discoveries are being reshaped. This paradigm, however, exposes a critical impedance mismatch: traditional HPC infrastructures struggle with the unique demands of AI workloads, while the success of data-intensive AI methods demands robust, scalable, and intelligent computational resources. In collaboration with other teams in HPC and AI-related research, our research addresses this dual challenge through two pillars. AI4HPC-a strategic approach that infuses the HPC stack with AI and novel methods to create an intelligent infrastructure-and HPC4AI, which is using this optimized foundation to accelerate large-scale scientific breakthroughs.

The AI4HPC pillar aims to develop intelligent optimization methods targeting performance, energy efficiency, and scalability across five core areas: hardware-software co-design, multi-parametric system optimization, scalable data analytics, advanced simulation frameworks, and wide-scope algorithm optimization. Moving beyond synthetic benchmarks, we validate these methods through real-world applications. All methods are empirically evaluated on national, testbed and international infrastructures, including the Aznavour supercomputing system.

The HPC4AI pillar demonstrates the tangible impact of this infrastructure by powering discovery across diverse domains. This includes a cloud-based ML service that classifies over 40 million astronomical spectra using custom CNNs and Mamba-based architectures; scalable platforms for environmental monitoring using satellite data; and high-fidelity quantum simulations using hybrid classical and physics-informed neural network models. Altogether, these applications validate a full, production-grade AI-HPC ecosystem that directly translates infrastructural innovations into transformative scientific insights.

Zero-Knowledge Proofs for Digital Image Authenticity

Laura Atmanavičiūtė

Kaunas Faculty
Vilnius University

laura.atmanaviciute@knf.vu.lt

Ensuring the authenticity of digital images has become a major challenge in an era where editing tools and generative models can easily alter visual content. It is often difficult to tell whether an image is genuine or the result of manipulation, especially when metadata or watermarks can be removed or forged. Zero-knowledge proofs (ZKPs) provide a cryptographic approach to this problem by allowing one party to prove that an image has undergone only permitted transformations, without revealing any information about the original version. This approach combines verifiability and privacy, offering a possible foundation for privacy-preserving image authentication systems. This study investigates how modern zero-knowledge proof systems can be applied to verify image transformations. It focuses on recent advances in zk-SNARKs, zk-STARKs, and recursive or folding-based proof schemes such as Halo2 and Nova, which make it possible to generate compact proofs that can be quickly verified even for complex computations. The research begins with a theoretical analysis of these systems, comparing their setup assumptions, proof sizes, verification speeds, and scalability. By representing standard image operations, such as cropping, resizing, brightness and contrast adjustment, or filtering as arithmetic circuits, the study shows how visual transformations can be expressed in a form suitable for zero-knowledge verification. This allows the verifier to check that a given output image truly results from a valid transformation of some original image, without ever revealing that original. The experimental part of the research builds on an open-source folding-based zk-SNARK framework designed for verifiable image transformations. The proof-of-concept experiment tests how efficiently such proofs can be generated and verified for high-resolution images, using common transformation scenarios. Several performance indicators are analyzed, including proof

generation time, peak memory use, proof size, and verification speed. The results demonstrate that proofs remain compact and that verification consistently takes less than a second, confirming the practical potential of zero-knowledge verification. However, the proving phase still requires considerable computational effort, particularly for larger images, where time and memory demands grow rapidly. These findings highlight the main trade-off in current zero-knowledge systems: while verification is efficient and independent of input size, the cost of proof generation remains a key limitation. Despite this, the research confirms that zero-knowledge proofs can serve as a promising basis for privacy-preserving verification of digital images and other multimedia content. The approach aligns well with broader trends in decentralized authenticity frameworks and digital provenance systems. Future improvements in circuit optimization, hardware acceleration, and hybrid post-quantum schemes could make zero-knowledge verification practical for real-world use, bridging the gap between cryptographic research and applied media authenticity.

FNDTN Protocol: A Bitcoin-Native Decentralized Scientific Publishing System with Cryptographically Verifiable Peer Review

Kazimieras Bagdonas

Kaunas University of Technology

kazimieras.bagdonas@ktu.lt

We present FNDTN Protocol, a decentralized scientific publishing system built on Bitcoin's Layer 1 and Layer 2 infrastructure. The protocol addresses fundamental inefficiencies in traditional academic publishing: centralized control, unpaid peer review, slow publication cycles, and lack of transparency. By leveraging Bitcoin's Taproot smart contracts for trustless escrow, Lightning Network for micropayments, Discreet Log Contracts (DLCs) for conditional payment distribution, and Nostr for censorship-resistant communication, FNDTN creates an economically sustainable ecosystem where reviewers are compensated proportionally to their contribution quality. The system employs threshold encryption to maintain reviewer anonymity during evaluation while enabling post-publication identity revelation for reputation building. We demonstrate how game-theoretic incentive alignment, combined with cryptographic guarantees, produces a Nash equilibrium favouring high-quality peer review. The protocol achieves censorship resistance through distributed storage sharding (256 shards with 3x replication), deterministic reviewer selection via TF-IDF semantic matching, and Bitcoin timestamping for immutable publication records. Our economic model shows reviewers can earn 30,000-45,000 satoshis per review, creating the first viable market for professional peer review services. Early simulations indicate the system can process 100,000,000+ papers annually with sub-7-day review cycles while maintaining Byzantine fault tolerance for up to 33% malicious participants.

Machine Learning Approaches to Student Dropout Prediction: A Systematic Review (2020–2024)

Amanda Balnionyte

Vilnius Gediminas Technical University

amanda.balnionyte@gmail.com

Student dropout remains a persistent issue in higher education that affects not only individuals but also institutions and national education systems. For universities, every student who withdraws represents a loss of potential, resources, and reputation. Over the last few years, artificial intelligence, and particularly machine learning, has been increasingly applied to understand and predict this phenomenon. By analysing digital traces of student activity, such as Moodle logs, these models can uncover behavioural patterns that often precede academic disengagement.

This study presents a systematic literature review of research published between 2020 and 2024, focusing on machine learning approaches to student dropout prediction. The aim was to identify which algorithms perform best, which data types and features are most informative, and what methodological challenges remain in this growing research field. Across the reviewed work, algorithms such as Random Forest, Gradient Boosting Machine, XGBoost, and LightGBM, consistently achieved the strongest results, with accuracy and F1 scores frequently exceeding 0.90. Simpler methods such as Naïve Bayes or single Decision Trees performed less reliably, especially with imbalanced datasets.

The most influential predictors of dropout were found to be academic performance indicators and behavioural engagement variables drawn from Moodle activity, such as frequency and recency of logins. In contrast, demographic and financial data offered limited predictive value on their own. Several recurring gaps were also identified. Many studies relied on small, single-institution datasets, and evaluation metrics were not always consistent. Few papers addressed model transparency or ethical concerns such as bias and privacy.

These limitations underline the need for explainable and ethically responsible AI models that can be scaled across institutions. Future research should focus on improving model scalability, feature engineering, and real-time data integration. Progress in these areas could significantly improve the accuracy, robustness, and adaptability of dropout prediction systems in dynamic learning environments.

Computational Analysis of Mechanisms Governing the Sensitivity and Efficiency of Enzyme-Based Biosensors and Bioreactors

Romas Baronas

Institute of Computer Science
Vilnius University

romas.baronas@mif.vu.lt

Enzymes play a crucial role in analytical biosensing systems due to their ability to specifically recognize analytes (substrates) and catalyse their conversion into products that can be readily detected using conventional analytical methods, such as electrochemical, optical, and other techniques [1]. In such systems, enzymes are primarily used in immobilized forms as biosensors or bioreactors. Enzyme-based biosensors and bioreactors are widely used in various fields, including medical and clinical diagnostics, environmental monitoring, as well as industrial and biotechnological processes [2, 3]. Computational modelling of enzyme-based biosensors and bioreactors enables the simulation of biosensor responses and bioreactor yields under both steady-state and transient conditions. The simulations consider biosensors and bioreactors with complex geometries and kinetic schemes that describe the action of biocatalysts. Mathematical and computational tools are widely used to optimize existing biochemical systems and to develop novel ones [4]. The aim of this work was to investigate the influence of the partitioning and diffusion limitations on the efficiency of enzyme-based bioreactors and biosensors using a three-layer model involving different schemes of the enzyme kinetics [5-7]. Exact steady state analytical solutions for the substrate and reaction product concentrations and the bioreactor effectiveness, as well as biosensor response, were obtained for the first and zero-order reaction rates. The transient action of enzyme-based bioreactors and amperometric as well as potentiometric biosensors was numerically investigated using the finite difference technique. The numerical simulator has been programmed in Java. Mathematical modelling of the diffusion-limited membrane and the conditions under which

the same values of the steady state characteristics are obtained when simulating the treated system by purposefully changing the values of the diffusion and distribution coefficients have been investigated [5, 6]. The application of different specific types of interface conditions, perfect contact and partition conditions, at which the steady state biosensor response is the same for both types of interface conditions, has been studied. To simplify the analysis, effective diffusion coefficients for the overall layer, comprising the diffusion-limiting membrane and the outer diffusion (Nernst) layer, have been identified to reduce the three-layer model to an equivalent two-layer model [5, 6]. In particular, it was determined that at relatively high substrate concentrations and in the presence of external diffusion limitation, the transient response of an amperometric biosensor exhibiting uncompetitive substrate inhibition may follow a five-phase pattern, depending on the model parameter values. Specifically, the response starts from zero, reaches a global or local maximum, decreases to a local minimum, increases again, and finally decreases to a steady intermediate value [7]. Managing such oscillations in the transient biosensor response is crucial for accurately predicting the analyte concentration, e.g., glucose in human blood, using a glucose biosensor.

References

[1] H. Bisswanger, Enzyme Kinetics: Principles and Methods, 2nd ed., Wiley-Blackwell, 2008.

[2] B.D. Malhotra, C.M. Pandey, Biosensors: Fundamentals and Applications, Smithers Rapra, 2017.

[3] P.M. Doran, Bioprocess Engineering Principles, 2nd edn., Academic Press, 2013.

[4] R. Baronas, F. Ivanauskas, J. Kulys. Mathematical Modeling of Biosensors, 2nd ed. Springer, 2021.

[5] R. Baronas. Nonlinear effects of partitioning and diffusion-limiting phenomena on the response and sensitivity of three-layer amperometric biosensors. Electrochim. Acta, 2024, 478, 143830.

[6] R. Baronas. Nonlinear effects of partitioning and diffusion limitation on the efficiency of three-layer enzyme bioreactors and potentiometric biosensors. J. Electroanal. Chem., 2024, 974, 118698.

[7] R. Baronas. Non-monotonic effect of substrate inhibition in conjunction with diffusion limitation on the response of amperometric biosensors. Biosensors, 2025, 15, 441.

Feature Stability Index (FSI): A Multi-Axis Metric for Assessing Robustness of Features in Imbalanced Fraud Detection

Dalia Breskuvienė, Gintautas Dzemyda

Institute of Data Science and Digital Technologies
Vilnius University

dalia.breskuviene@mif.vu.lt

In highly imbalanced domains such as credit card fraud detection, model explanations are often dominated by a few seemingly influential features. However, the importance of these features can vary considerably when data distribution, model architecture, or random initialization changes, raising concerns about reproducibility and trustworthiness. To address this, we introduce the Feature Stability Index (FSI), a unified metric that quantifies the robustness of feature importance patterns across different experimental conditions.

FSI is not a feature selection method but a diagnostic measure that evaluates how consistently a feature set maintains its relevance under three axes of variation: model choice, random seed, and temporal data window. The metric aggregates axis-specific stability components S_{model}, S_{window}, and S_{seed} into a single interpretable score, weighted by parameters α, β, and γ. Two complementary formulations are proposed: FSI-CV, which measures numerical consistency of feature importance magnitudes, and FSI-IE, which assesses the stability of feature set inclusion frequencies through entropy analysis.

Experiments using credit card transaction data demonstrate that FSI effectively distinguishes stable from unstable importance signals. Features with high FSI values exhibit consistent relevance across models and resampling, while unstable features fluctuate significantly, indicating potential sensitivity to model design or data drift.

Unlike traditional overlap-based stability indices such as Jaccard or Kuncheva, FSI captures both numeric and inclusion stability and attri-

butes instability to specific axes, providing a deeper understanding of model behavior. The proposed FSI framework supports robustness auditing of feature importance results and contributes to more reliable model interpretation in dynamic, high-risk environments such as financial fraud detection.

AMBER C2: Enhancing Cyber Defence with Ethical Adversarial Machine Learning

Arnoldas Budžys, Juozas Dautartas, Haroldas Jomantas, Viktor Medvedev, Olga Kurasova

Institute of Data Science and Digital Technologies
Vilnius University

arnoldas.budzys@mif.vu.lt

The cybersecurity world is divided into adversaries seeking to disrupt operations and defenders protecting sensitive infrastructure. The blue team protects and monitors systems while the red team attempts to breach them. The white team designs, implements, and manages the exercise infrastructure. The penetration testers use red team methods to find vulnerabilities before hostile actors do.

We are working on the AmberC2 project, which focuses on a secure Command and Control (C2) framework that integrates Adversarial Machine Learning (AML) to support realistic but controlled cyber exercises. We can manage ethically disguised malicious software in an isolated laboratory, applying strict security measures and comprehensive auditing. Our goal is training, evaluation, and research that strengthen defence. AmberC2 supports payload generation, delivery, and control channels that can be restricted, redirected, or terminated as needed. The framework explores evasion and concealment techniques inspired by AML methods, testing them only for instrumented purposes. These techniques help simulate advanced persistent threats while maintaining security. Approach-wise, AmberC2 investigates malware obfuscation and evasion techniques through AML methods. These methods correspond to the advanced persistent threats we are trying to replicate, thus offering protection measures against such attacks. The project presents the system architecture, management framework, and responsible operating procedures. The goal is to increase the realism of exercises, reveal gaps in modern defences, and accelerate the development of resilient security solutions. In the future, the variety of scenarios in

networks, operating systems, and cloud platforms will be expanded, and the threat generation policy will be refined to reflect changing methods.

Acknowledgments. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-MIP-24-116.

Explainable Multi-Label Chest X-Ray Analysis for Severity-Aware Triage

Paulius Bundza, Justas Trinkūnas

Dept. of Information Systems
Vilnius Gediminas Technical University

justas.trinkunas@vilniustech.lt

Chest X-rays are a high-volume, time-critical modality, yet interpretation remains resource-intensive and error-prone in the context of radiologist shortages, motivating AI "second reader" systems to reduce turnaround time and inter-reader variability. This work presents MCADS, a multi-label deep-learning system that detects 18 chest radiographic abnormalities from a single image and explains predictions using Grad-CAM, extending common 14-label setups (e.g., ChestX-ray14, CheXpert) by combining broader label coverage with interactive explainability and an end-to-end web application. MCADS implements embedded case-level prioritization based on mean predicted abnormality probability, with three triage bands-insignificant (≤19%), moderate (20-30%), and significant (≥31%)-to surface potentially urgent cases for earlier review. The system is evaluated across eight public datasets (NIH ChestX-ray14, CheXpert, MIMIC-CXR with CheXpert labels, Google DS1, RSNA Pneumonia, SIIM-ACR Pneumothorax, PadChest, and VinDr-CXR), using AUC-ROC per pathology as the primary performance metric.

The model is based on DenseNet-121 from the TorchXRayVision library with sigmoid output heads for 18 labels, pretrained on a multi-dataset chest X-ray corpus and followed by post-hoc temperature scaling and per-label decision thresholds to improve calibration. For each predicted abnormality, Grad-CAM heatmaps are generated and stored alongside probabilities, enabling transparent visual review and supporting quality-assurance workflows. MCADS is deployed as a Django (ASGI) web application with Celery/Redis asynchronous workers, Nginx/Gunicorn serving, and PostgreSQL plus media storage for persistence. The clinical workflow, implemented as a single-page interface, follows a BPMN-modeled pathway: the technologist acquires the chest X-ray,

MCADS performs preprocessing, multi-label inference, and Grad-CAM generation, and the radiologist reviews, confirms, or edits the findings before forwarding the report to a lung disease specialist for treatment planning; history and administrative views are included.

MCADS achieves strong cross-dataset discrimination for many labels, for example AUC-ROC up to 0.93 for cardiomegaly and 0.95 for effusion on PadChest, 0.93 for pneumothorax on VinDr-CXR, and 0.84 for pneumonia on CheXpert, while maintaining state-of-the-art DenseNet-121 performance when served through the web application without observable degradation due to integration. The system generalizes across all eight datasets without retraining, and on a 2-core CPU the inference latency after warm-up is approximately 5-30 seconds per image, with model caching reducing subsequent latency; Grad-CAM heatmaps typically align with relevant anatomical regions (e.g., consolidation regions, pleural spaces), supporting clinician trust in the explanations. Current limitations center on the reliability and coverage of the system: Grad-CAM heatmaps can produce misleading or "hallucinated" saliency patterns that visually emphasize non-pathologic regions, and there is no dedicated edge-case detection mechanism, so rare or atypical studies may not be identified as such and can be misclassified or under-triaged. Future work is planned on prospective clinical validation, improved robustness across institutions, and multimodal fusion with additional clinical data.

References

[1] Cohen, J. P., Viviano, J. D., Bertin, P., Morrison, P., Torabian, P., Guarrera, M. G., Lungren, M. P., Chaudhari, A., Brooks, R., Hashir, M., & Bertrand, H. (2022). TorchXRayVision: A library of chest X-ray datasets and models. Proceedings of the 5th International Conference on Medical Imaging with Deep Learning (MIDL), 172, 1-19. https://doi.org/10.48550/arXiv.2111.00595

Improving Malware Detection by Analyzing Similarities of Multi-Category Benign Software

Juozapas Rokas Čypas, Juozas Dautartas, Olga Kurasova, Viktor Medvedev

Institute of Data Science and Digital Technologies
Vilnius University

juozapas.cypas@mif.vu.lt

In today's digital world, the importance of cybersecurity is increasing rapidly. The evolution of technology and AI enables various threat actors to evolve malware as well as the methods of evading malware detection. In this study, we aim to analyze modern malware evasion methods presently used in the wild. It's important to identify such methods so that they can be analyzed and studied by security researchers for the purpose of improving the defense infrastructure of software systems. Traditionally, machine learning based analysis of Windows Portable Executable (PE) file static features uses datasets that have either two classes (benign or malicious), or multiple malware classes (e.g., worms, Trojans, ransomware, spyware), and one benign software class. Our proposed method and dataset (DOI:10.18279/MIDAS.265677) focus on the analysis of multiple categories of benign software (office tools, security, media, etc.) and just one class of malware. These categories are used to train a classifier that can distinguish benign software based on its static features. Our concept is that for a given malware sample, it is possible to identify a benign category corresponding to the lowest expected detection rate. This approach analyzes the similarity between each malware sample and multiple categories of benign software. For each instance of malware, we identify the closest benign category and then inject the most characteristic static features of that benign cluster into the malware sample to trick the classifier and maximize the evasion rate. Preliminary results indicate that, after injecting features from the category of benign files selected based on similarity, the initial classifier demonstrates a decrease in the detection rate of such concealed malware in comparison to the original malware file.

Acknowledgments. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-MIP-24-116.

Mimer – Sweden's AI Factory Transforming Access to Trustworthy AI

Petra Dalunde

Director Strategic Partnerships
AI Factory Mimer
RISE Research Institutes of Sweden

petra.dalunde@ri.se

Mimer, Sweden's newly established AI Factory, represents a transformative initiative aimed at democratizing access to trustworthy AI across the Swedish and European innovation ecosystem by supporting the AI innovators. As one of 19 AI Factories and 13 Antennas launched across Europe in 2025, Mimer is part of the European Union's strategic effort to realize the goals of the Digital Decade 2030. This initiative complements broader EU actions to create a unified internal digital market and to elevate AI competencies across sectors. Mimer provides free access to high-performance compute resources, AI expertise, curated datasets, and tailored support to startups, SMEs, and public sector organizations. By removing barriers to entry, Mimer enables smaller actors to experiment, develop, and deploy AI solutions that are both innovative and ethically grounded. In this presentation, Petra Dalunde, Director Strategic Partnerships at Mimer, will briefly outline the EU policy context and then delve into the operational model of Mimer, detailing the guiding functions to other EU Innovation Infrastructure initiatives like TEFs, EDIHs, AIoD etc, as well as services offered to different target groups and how these services are designed to foster responsible AI development. The talk will highlight how Mimer's infrastructure and collaborative framework empower organizations to scale AI projects, accelerate time-to-impact, and contribute to a more inclusive and competitive European AI landscape. Through real-world examples and strategic insights, the session will demonstrate how Mimer is not only a national resource but also a key node in Europe's vision for a resilient, sovereign, and human-centric digital future.

Prompt-Based Bias Detection Using Linguistic Metrics

Edgaras Dambrauskas[1,2], Eglė Rimkienė[1,2], Justina Mandravickaitė[1,2], Dmytro Klepachevkyi[1,2], Milita Songailaitė[1,2], Tomas Krilavičius[1,2]

[1] Vytautas Magnus University
[2] Center for Applied Research and Development

edgaras.dambrauskas@vdu.lt

The rise of disinformation in the media presents significant challenges to public understanding as well as democratic processes. With rapid online dissemination, distinguishing between reliable journalism and deliberately misleading content has become extremely important. Therefore, this study examined narrative construction patterns in disinformation and trustworthy news through analysis of English news articles covering selected international events (2015-2023).

We evaluated knowledge graph (KG) grounding for narrative extraction and compared grounded vs. ungrounded variants across events, causal links and frames with role-specialised two-model ensembles per structure type (extractors + formatter). We used Mistral-Small-3.1-24B as a formatter across tasks. The following ensembles of smaller LLMs were used as extractors: entity linking (Qwen-2.5-72b + Llama-3.1-8b), framing (Gemma-3-27b-it + Qwen-2.5-72b), whole narrative extraction (Gemma-3-27b-it + Mistral-Small-3.1-24b). We also assessed a 'critic' component for missing links, contradictions and inconsistencies and applied Mistral-Small-3.1-24b + Gemma-3-27b-it for this task. On a human-annotated set, event micro-precision reached 0.87 for disinformation and 0.75 for trustworthy news. Furthermore, a grounded Mistral-Small-3.1-24b variant resulted in denser event/relation graphs than its ungrounded counterpart and larger ungrounded Llama-3.3-70B-Instruct, also aligned better to the KG schema, which integrated Abstract Meaning Representation (AMR) parses and FRED (an automatic system that derives RDF/OWL ontologies and linked data from natural-language text). Causal link detection was recall-limited (ranking-insensitive) with modest precision (micro

0.50/0.45 for disinformation/trustworthy news, accordingly), while 'critic' performance was ranking-limited with large oracle headroom, achieving moderate precision (0.54–0.57) across classes. Also, KG grounding improved gap-finding, raising missing-link coverage from 2-4% (no KG) to 54-57% and up to 100% on an annotated sample. An entity-linking ablation with a simple two-way consensus reached full coverage without disagreement on the targeted sample.

Taken together, these results support a practical recipe in which KG grounding and ensembles of smaller LLMs deliver competitive quality at a manageable cost.

Modelling Pattern Formations of Bacteria: Influence of Gravity

Boleslovas Dapkūnas, Romas Baronas, Remigijus Šimkus

Vilnius University

boleslovas.dapkunas@mif.vu.lt

In microcontainers, growing bacterial colonies of various species, such as *Escherichia coli* and *Bacillus subtilis*, self-organise and form patterns. One of the phenomena that can be observed in patterns during physical experiments is the formation of plumes: vertical structures descending from the larger aggregate of bacteria near the top of the microcontainer. The mechanism of plume formation is still poorly understood. The use of mathematical modelling can help fill the gaps in knowledge.

Studies of mathematical models for bacterial pattern formation have intensified since the introduction of Keller–Segel partial differential equations model for chemotaxis in 1971 [1]. When modelling *Escherichia coli*, experiments have shown that the dynamics of oxygen have to be taken into account [2]. Hillesdon et al. [3] have shown that by coupling the Keller–Segel model with the fluid flow equation, plume formation can be modelled in colonies of *Bacillus subtilis*.

By coupling Keller–Segel model involving oxygen dynamics with Navier–Stokes incompressible fluid flow equations, we investigate the effects of gravity on the modelled *Escherichia coli* plume formation. The model consists of partial differential equations describing the dynamics of bacteria density, self-excreted chemoattractant concentration, oxygen concentration, and fluid dynamics. We also investigate the influence of these dimensionless model parameters: Schmidt number, Rayleigh number, and oxygen cut-off threshold. The numerical simulation was carried out using the finite difference technique.

References

[1] E.F. Keller, L.A. Segel, Model for chemotaxis, *J. Theor. Biol.*, 30(2):225–234, 1971

[2] R. Šimkus, R. Baronas, Ž. Ledas, A multi-cellular network of metabolically active E. coli as a weak gel of living Janus particles, *Soft Matter*, 9(17):4489–4500, 2013.

[3] A.J. Hillesdon, T.J. Pedley, O. Kessler, The development of concentration gradients in a suspension of chemotactic bacteria, *Bull. Math. Biol.*, 57:299–344, 1995.

A Longitudinal Analysis of Temporal Anomaly Detection in Telegram Cybersecurity Channels

Andrius Daranda, Lina Kankevičienė, Julija Daranda

Alytus Faculty
Kauno kolegija Higher Education Institution

andrius.daranda@go.kauko.lt

The exponential growth of cybersecurity information on social media platforms presents challenges for assessing information or detecting threats. This study develops a comprehensive framework for automated cyber threat prioritization by analyzing temporal patterns, content quality indicators, and community engagement. Our methodology employed content similarity, multi-method burst detection, and credibility scoring based on technical metrics. The analysis identified meaningful temporal patterns and revealed anomalous activity days. Moreover, it has been established that specific content characteristics correlate with higher engagement. We found a quality-reach trade-off, where highly technical content exhibits lower immediate virality but greater long-term value. Also developed a credibility scoring system to prioritize quality. This research presents a systematic and quantifiable approach for early-warning cyber threat systems, providing insights for strategies to combat cybersecurity information overload.

CardioCopilot: An AI-Powered Virtual Consultant for Cardiovascular Prevention

Gabrielė Dargė, Deividas Valiuška, Eimantas Zaranka, Ričardas Krikštolaitis, Tomas Krilavičius

Faculty of Informatics
Vytautas Magnus University

gabriele.darge@vdu.lt

CardioCopilot is an AI-powered virtual consultant designed for cardiovascular disease prevention and health promotion. It combines large language models (LLMs) with retrieval-augmented generation (RAG) methods to provide accurate, personalised recommendations. The system integrates a FAISS vector database, LangChain framework, and OpenAI's GPT-4o model to ensure reliability and data security. Its analysis is based on the latest European Society of Cardiology guidelines, enabling assessment of arterial hypertension, dyslipidemia, diabetes, and lifestyle factors such as nutrition and physical activity. User data, stored in a MySQL relational database, is integrated in a way that allows individualised blood pressure analysis and recommendations while ensuring data protection. The system maintains strict conversational boundaries and topic limitations, focusing exclusively on cardiovascular disease-related queries. Testing demonstrated high medical accuracy, reduced risk of hallucinations, and strong multilingual support. CardioCopilot acts as a trusted intermediary between individuals and healthcare professionals, helping to identify risks early and promote healthier lifestyle decisions.

Feature Level Deception or When Malware Wears a Mask

Juozas Dautartas, Juozapas Rokas Čypas, Viktor Medvedev, Olga Kurasova

Institute of Data Science and Digital Technologies
Vilnius University

juozas.dautartas@mif.stud.vu.lt

Today's digital landscape shows an unsettling race between cyber defense and offense fields. The rise in popularity of machine learning (ML) has made this race even more intense as these technologies have become an integral part of our everyday security tools and products. These tools integrate various ML algorithms that have been trained on large datasets of static and dynamic malware features or patterns of malicious network traffic.

Therefore, it comes as no surprise that adversaries are implementing various attacks against these classifiers used by security products. That's why testing and validating current defenses is a critical part of a cybersecurity professional's job. In this research, we will analyze a targeted adversarial attack against classical ML malware classifiers. We will focus on Windows API calls from various benign classes as well as malware. These data will be used to impersonate a specific benign class using feature injection techniques. The adversarial samples will be applied to test trained ML classifiers as well as real products.

This research is conducted for ethical and research purposes with an aim to make cybersecurity defenses more robust and reliable. As these realistic and malicious functionality preserving samples can be used to train more accurate malware classifiers in the future.

Acknowledgments. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-MIP-24-116.

Automatic Propaganda Technique Classification in Lithuanian News Articles Using Pretrained Language Models

Greta Dovidaitytė, Gražina Korvel

Institute of Data Science and Digital Technologies
Vilnius University

greta.dovidaityte@mif.vu.lt

Propaganda is a powerful tool used to influence the opinions or actions of the audience, and propaganda techniques are the methods used to achieve it. For centuries, propaganda had a neutral connotation; however, it has recently become associated with foul intentions, manipulation, and deception. The development and increased availability of communication technologies created a favorable environment for the rapid dissemination of digital propaganda. One of the ways to spread it became the news media. Fortunately, fast advancement in machine learning technologies has led to the development of systems that can automatically detect and classify propaganda in news articles. The recent propaganda technique classification studies took advantage of the creation of pretrained language models (PLMs). The rich, context-aware text representations created by PLMs helped to capture subtle propaganda rhetorical cues and achieve state-of-the-art results. However, most studies focus on the English language, creating a research gap for various low-resource languages, including Lithuanian. This study focuses on the classification of propaganda techniques in Lithuanian news articles using three pre-trained language models – multilingual BERT, XLM-RoBERTa, and LitLat. The experiments were performed on a new Lithuanian propaganda dataset created by the Vilnius University propaganda and disinformation research project ATSPARA. In addition to the model performance comparison, this research focuses on model interpretability using Explainable AI frameworks, extensive data analysis, and linguistic differences between propaganda techniques.

Acknowledgements: The authors are thankful for the data provided by the ATSPARA project, supported by the Lithuanian Government Priority Research Program "Building Societal Resilience and Crisis Management in the Context of Contemporary Geopolitical Developments" (implemented through the Lithuanian Research Council) under Grant No. S-VIS-23-8. The authors are also thankful for the high-performance computing resources provided by the Information Technology Open Access Center of Vilnius University.

Predicting Students' Achievements Using Machine Learning Methods

Hande Safiye Erdal[1,2], Monika Katilienė[1,2], Tomas Krilavičius[1,2], Tomas Urbonas[3]

[1] Vytautas Magnus University
[2] Center for Applied Research and Development
[3] Sonaro

hande.safiye.erdal@vdu.lt

Schools increasingly seek decision support systems that rely on data to improve learning accomplishment and support underachieving pupils. Despite the growing demand, limited research has been conducted on incorporating real-time educational data to predict student achievement. This study uses e-diary data collected from 12 schools between 2017 and 2023, covering 11,577 student records. The dataset includes information on students' academic background, teachers' experience, and school-level characteristics. For prediction, the focus was placed on subjects with the most frequently recorded grades since these subjects provide enough data to allow reliable forecasting. The goal of prediction was to estimate the final average grade of a subject at the end of the current semester. Several experimental setups were tested: forecasting 20, 30, 40, 50, and 60 days before the end of the semester; forecasting the next semester based on the previous one; forecasting the upcoming year based on the past year; and making long-term forecasts up to three years ahead. Different modeling strategies were compared, including semester-specific, subject-specific, grade-specific, and universal approaches. The universal strategy was ultimately selected for its scalability and stronger generalization across different cases. A variety of regression models were tested, including linear models and ensemble methods such as random forest, gradient boosting, and CatBoost. Model performance was evaluated using standard metrics: mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). Across all experiments and prediction horizons, CatBoost consistently outperformed the other models. For

instance, in short-term forecasts made 20 days before the end of the semester, the best-performing models achieved R^2 values higher than 0.95. Longer-term predictions remained reasonably accurate. These results indicate that ensemble learning methods, particularly gradient boosting approaches such as CatBoost, are effective for forecasting student performance based on e-diary data. The model achieves high accuracy across multiple time horizons, providing valuable support for data-driven educational decision-making.

Quantum Computing for Vehicle Routing Problems: State of the Art and Challenges

Ernestas Filatovas[1], Remigijus Paulavičius[1], Dmitry Podkopaev[1,2]

[1] Institute of Data Science and Digital Technologies
Vilnius University
[2] Systems Research Institute
Polish Academy of Sciences

podkop@ibspan.waw.pl

Numerous variants of the vehicle routing problem (VRP) are important in operations research and practice. Last-mile delivery problems attract particular attention because of the large numbers of customers, products, and vehicles involved. These problems are usually formulated as mixed-integer linear programming (MILP) or other combinatorial optimization models. Real-life problems of large scale, as they appear in business and logistics, are often not solvable with traditional MILP or heuristic algorithms due to their high computational complexity, especially when realistic constraints such as time windows, vehicle capacities, service priorities, or multi-depot structures must be taken into account. The rapid progress of quantum computing in recent years has enabled the first experimental applications of quantum methods to routing and logistics. Quantum annealing, in particular, has shown promise for handling larger problem instances and for delivering high-quality heuristic solutions within short runtimes. This makes it suitable for hybrid approaches, where many smaller subproblems are decomposed from a large initial model and solved on a quantum device, while the overall coordination relies on classical optimization. At the same time, important challenges remain, such as limited solution accuracy, overhead of mapping logical variables to physical qubits, and restricted problem sizes. These challenges continue to constrain applicability, while the adequate representation of rich real-world constraints in quantum formulations is still an open research issue. This talk reviews the state of the art in applying quantum computing to

solving large-scale VRPs. We describe first examples of realistic problem solving, outline the potential for building real-world applications with the emphasis on last-mile delivery, and highlight current hardware and algorithmic approaches. The relative advantages of quantum annealing devices are discussed, along with key challenges that must be addressed before practical adoption becomes possible.

Evaluation of Multimodal AI Models for Eyeglasses Recognition

Henrikas Giedra, Dalius Matuzevičius

Vilnius Gediminas Technical University

henrikas.giedra@vilniustech.lt

The automatic annotation of large-scale datasets is an important step for developing reliable computer vision systems. Eyeglasses detection plays a significant role in biometric preprocessing, demographic analysis, and downstream facial recognition applications. Manual labeling of eyeglasses in large datasets is labor-intensive and prone to inconsistencies, motivating the use of multimodal AI models to automate this process. This research investigates the effectiveness of state-of-the-art vision–language models (VLMs) in automatically labeling eyeglasses within facial image datasets. We benchmark five representative VLMs – Gemini 2.0 Flash, Gemini 2.5 Flash, Gemini 2.5 Pro, Mistral, and Gemma 3 – on two widely used face datasets: CelebAMask-HQ and FFHQ. The evaluation framework examines detection accuracy and labeling consistency across varied demographic attributes and image conditions. Robustness is further assessed under challenging factors such as partial occlusions, reflections, and stylistic variations of eyewear. Beyond quantitative performance, we analyze model outputs for systematic biases and misclassification patterns, offering insights into the reliability of VLMs in automated annotation pipelines. The findings demonstrate the potential of modern VLMs to serve as effective, scalable, automated tools for dataset labelling, reducing human effort while improving annotation consistency. This study provides comparative benchmarks and methodological insights to support the integration of multimodal AI into computer vision research and application pipelines.

Acknowledgements. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-24-12.

Investigation of Heterogeneous Service Times in Outpatient Clinics Using Clustering and Discrete-Event Simulation

Emilija Jakuškaitė, Dovydas Kaunietis, Lukas Velička, Aušra Gadeikytė

Department of Applied Informatics
Faculty of Informatics
Kaunas University of Technology

ausra.gadeikyte@ktu.lt

In recent years, long waiting times for healthcare services have been a major policy concern in most OECD countries, according to a 2020 OECD report. Healthcare service providers often face challenges in managing patient queues, leading to longer waiting times, negative patient experience, and inefficient resource utilization. Traditional queue management approaches are often insufficient within clinical environments due to dynamic patient flow, varying service durations, and limited medical resources. This study examines outpatient appointment scheduling and queue modelling strategies in order to reduce patient waiting times and evaluate service efficiency, using the open-access Hangu clinic dataset. This dataset contains service time records with heterogeneous features such as patient demographics, medical problems, and previous visit information. The analytical part of the investigation involves exploratory data analysis and clustering of patient data to identify groups with different service durations. The optimal number of clusters was selected based on the silhouette coefficient and the elbow method, using k-means and k-median methods. Three patient clusters with associated service durations and inter-arrival times were used as input for a discrete-event simulation to evaluate queue dynamics. The proposed model allows the prediction of different scenarios of patient flow intensities and the assessment of physician workload through the resource utilization coefficient. Furthermore, the prototype of the queue management system was developed using

the Flask framework and SimPy library for discrete-event simulation. The simulation was enhanced with the scikit-fuzzy library for wait time predictions. Also, real-time queue visualization and QR code-based tracking were implemented. The simulation results demonstrated that scheduling appointments according to patient cluster characteristics can reduce average waiting times.

Evaluation Metrics for Explainable AI

Rokas Gipiškis, Olga Kurasova

Institute of Data Science and Digital Technologies
Vilnius University

rokas.gipiskis@mif.vu.lt

Explainable Artificial Intelligence (XAI) has become one of the key components in ensuring transparency and accountability in AI systems. However, the evaluation of XAI methods remains inconsistent and fragmented. Here, we analyse the categorisations, methodologies, and properties of various evaluation metrics to highlight key differences in evaluation approaches and the lack of consensus across studies. Our findings emphasise challenges such as inconsistent definitions, limited implementation, and the absence of standardised evaluation frameworks. By comparing existing taxonomies and metric classifications, we identify gaps and provide insights to support the development of more robust and comparable XAI evaluation methodologies. Finally, we highlight underexplored application areas beyond classification, such as semantic segmentation, as promising directions for evaluation research.

Single-View Depth Map Integration for Robust Eyeglass Temple Extraction

Ervinas Gisleris, Artūras Serackis, Dalius Matuzevičius

Vilnius Gediminas Technical University

ervinas.gisleris@vilniustech.lt

Precise extraction of eyeglass temples from images is challenging due to their slender geometry, reflective surfaces and tendency to be occluded. Conventional segmentation methods often fail to capture these subtle spatial cues, resulting in incomplete or inaccurate delineation. This research explores the potential of using single-view depth estimation as an additional method to improve the segmentation of eyeglass temples. We benchmark state-of-the-art monocular depth estimation models to evaluate their generated depth maps in the context of the fine-grained geometry of eyeglass temples. Our analysis involves quantitative and qualitative evaluations across various datasets with different lighting, pose, and occlusion conditions. Based on the benchmarking outcomes, we propose novel depth map integration strategies that exploit geometric information derived from single-view depth to reinforce segmentation robustness. These strategies include depth-guided attention mechanisms and adaptive fusing of depth and RGB features. Experimental results demonstrate that incorporating single-view depth maps yields improvements in segmentation accuracy. We establish a standardized evaluation protocol for assessing the contribution of depth features to segmentation performance, providing reproducible benchmarks for future research. The findings highlight the significance of leveraging monocular depth cues for enhancing robustness in eyeglass temple extraction tasks.

Acknowledgements. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-24-12.

Recommender Methods Based on Collaborative Filtering in E-Learning

Eligius M.T. Hendrix, Alejandro Fuster-López, Pablo Guerrero-García, Jesús Martínez-Cruz

Universidad de Málaga

eligius@uma.es

E-learning has a long tradition. In an e-learning environment, a computer system aids students to learn online, providing exercises and learning material. Recommender systems also have a long tradition. They are known in platforms providing entertainment like Netflix and products like Amazon. Clients obtain suggestions for the next products based on their past behaviour.

Recommender systems are relatively new in e-learning. On one hand, we have the assessment material on question and study material and on the other hand, a database of user profiles which record the background of the student and the history of how questions were answered and how long it took them.

In a European research project called iMath, we investigated various ways to create a recommender system for e-learning on mathematical subjects based on several methodologies. The aim was to facilitate learning and understanding for a student.

This provided a discussion of how to measure performance in learning indicators of the system; when is this successful? In the discussions about the effectiveness, we decided that methods should aim to predict which next question or which new material (video, description) may help the student further in obtaining insight, where success is measured as getting the next question well answered. This means we measure whether the next question is answered correctly.

Moreover, we had a discussion about concept maps; how do we know that students cover all the concepts available in the topic to be learned, and is there a logical order?

Like in other recommender systems, given past information of users in a database, machine learning methods like Random Forest have been

suggested and implemented. Another way is to simply classify questions on their difficulty level, given the success rate of the answers. Machine learning methods like random forest are data and energy-hungry. Instead, we implemented methods for collaborative filtering. The recommendations are based on other users of the system who followed a similar path, and for which some questions have been successful. The computation does not require learning of systems, but is based on matrix factorization methodologies, which require far less computation and energy.

Our investigation goes in the direction of comparing the machine learning ideas with matrix factorization methods, measuring effectiveness and efficiency.

Identification of Target Genetic Variations via the Mathematical Properties of the Genetic Code

Kamilija Jablonskaitė[1], Ali Aldujeli[2], Ieva Čiapienė[2], Ugnė Meškauskaitė[2], Indrė Matulevičiūtė[3], Vacis Tatarūnas[2], Mantas Landauskas[1]

[1] Kaunas University of Technology
[2] Lithuanian University of Health Sciences
[3] Hospital of Lithuanian University of Health Sciences

kamilija.jablonskaite@ktu.edu

One of the most common types of genetic variation is the single nucleotide polymorphism (SNP). It refers to the presence of a particular nucleotide at a specific position in the genetic code. Such genetic variations are often associated with various phenotypes and are commonly investigated in genome-wide association studies (GWAS) to uncover genotype–phenotype relationships [1]. Genetic variations might span regions of genetic code. This study aims at identifying target genetic variations with respect to mathematical properties of genetic code. Genetic sequence of interest is encoded as integer number sequence comprising of the first 4 natural numbers, each assigned to a particular nucleotide. In this way a gene, chromosome or any other genetic sequence could be further analyzed mathematically. Patterns, variability, statistical distribution, etc. could be assessed using a number of methods. Here Shannon entropy and generalized permutation entropy are computed for moving windows of the encoded sequence. Entropy based features are a proven class of methods for genetic data analysis [2]. For example, differences in entropy had been used to detect mutations in the virus DNA [3]. The aforementioned numerical features here are analyzed as a new sequence. Local extreme values (peaks) are identified and corresponding SNPs are selected. These are the target genetic positions to be further analyzed by experts in order to determine possible common aspects.

Acknowledgements. This research was supported by the Research and Innovation Funds of Kaunas University of Technology and Lithuanian University of Health Sciences (MEGRAC, Grant No. INP2025/7).

References

[1] Ding, X., & Guo, X. (2018). A survey of SNP data analysis. Big Data Mining and Analytics, 1(3), 173-190.
[2] Orlov, Y. L., & Orlova, N. G. (2023). Bioinformatics tools for the sequence complexity estimates. Biophysical reviews, 15(5), 1367-1378.
[3] Vopson, M. M., & Robson, S. C. (2021). A new method to study genome mutations using the information entropy. Physica A: Statistical Mechanics and its Applications, 584, 126383.

Detecting Maritime Anomalies Using LSTM and XGBoost Quantile Regression-Powered Prediction Interval Models and Dynamic Thresholding

Lukas Janušauskas, Robertas Jurkus, Julius Venskus

Institute of Applied Mathematics
Vilnius University

lukas.janusauskas@mif.stud.vu.lt

Since October of 2023 Baltic undersea cables have been damaged numerous times, some of these incidents are suspected to be an act of sabotage [1, 3]. Such was the case of the 2024 November 17-18 submarine cut incident, when a bulk carrier, Yi Peng 3, lowered its anchor, dragged it on the sea floor for approximately 160 kilometers and cut a telecommunications cable between Sweden and Lithuania [4]. This case, along with other incidents of cable disruption, illustrates the importance of maritime anomaly detection systems, which would allow for swift and reliable detection and response to irregular vessel behavior. We analyzed the Automatic Identification System (AIS) data of the Yi Peng 3 vessel at the time of the incident. This analysis heavily influenced the proposed approach to maritime anomaly detection. The main finding of the analysis was that the difference between heading and course was much larger than what would be expected normally, under meteorological conditions at the time and place of the incident. Thus, the proposed anomaly detection method revolves around the difference between heading and course. Using both AIS and meteorological data, we trained machine learning models to estimate prediction intervals of this measurement in a future window of 25 minutes. We trained eXtreme Gradient Boosted decision trees (XGBoost) and Long-Short Term Memory (LSTM) neural networks with a quantile loss function, which enables models to produce estimates of the conditional quantiles. This allowed us to obtain prediction intervals easily and subsequently have an anomaly detection mechanism. In this study, the effectiveness of XGBoost and LSTM models at maritime anomaly detection was compared using metrics specific to prediction intervals, such as prediction interval

coverage probability (PICP) and prediction interval normalized average width (PINAW). It was determined that LSTM models have a PICP of 97.77 % and PINAW of 0.022, while XGBoost achieved an inferior performance, having a smaller PICP of 94.09 % and an approximately equal PINAW of 0.019. Regarding the actual anomaly detection mechanism, simple prediction interval-based anomaly flagging was deemed as an insufficient approach due to the numerous false positives it produced. To mitigate this problem, we used the Dynamic Thresholding method [2], which enables us to determine the anomaly threshold without previous knowledge of the percentage of anomalous cases within the dataset and can be used for anomaly pruning, which will be expanded upon in further research. Most importantly, the model evaluated Yi Peng 3 AIS and meteorological data, and the anomaly was successfully identified using model output and threshold, obtained from Dynamic Thresholding. In conclusion, we propose an approach to maritime anomaly detection that can flag anomalies with a low false positive rate. The method is shown to be effective at detecting anomalous behavior in the Yi Peng 3 Baltic Sea cable disruption case. However, the set of known anomalies, which could be used for the evaluation of this system, is incredibly small, therefore, reliable ways for generating synthetic anomalous vessel movements will be investigated in further research.

Acknowledgements. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-MIP-24-117.

References

[1] BNS. (2024, 11 20). *Chinese ship under surveillance following Baltic Sea cables damage*. Retrieved from https://www.lrt.lt/en/news-in-english/19/2418936/chinese-ship-under-surveillance-following-baltic-sea-cables-damage

[2] Hundman, K. a. (2018). Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.* ACM. doi:10.1145/3219819.3219845

[3] Lehto, E. (2025, 8 11). *Finland charges Eagle S tanker captain, officers over cable cuts*. Retrieved from Reuters: https://www.reuters.com/markets/commodities/finland-charges-eagle-s-tanker-captain-officers-over-cable-cuts-2025-08-11/

[4] LRT.lt. (2024, 11 28). *Chinese ship suspected of deliberately cutting cables in Baltic Sea – media.* Retrieved from https://www.lrt.lt/en/news-in-english/19/2426033/chinese-ship-suspected-of-deliberately-cutting-cables-in-baltic-sea-media

Agent-Based Modelling in Sustainable Finance: A Systematic Literature Review

Megang N. Junile, Audrius Kabasinskas

Kaunas University of Technology

junmeg@ktu.lt

Sustainable finance requires simultaneous optimization of financial returns and sustainability outcomes, creating modelling challenges that traditional financial approaches cannot adequately address. This systematic review examines agent-based modelling (ABM) applications in sustainable finance, analyzing 45 studies published between 2005 and 2025 following PRISMA guidelines. Our analysis reveals significant gaps. Environmental dimensions dominate (47% of studies), while social and governance aspects remain underrepresented. Nearly half of all studies (49%) lack clear alignment with EU sustainability frameworks, 93.3% rely on single validation methods, undermining policy credibility. Geographically, research concentrates in developed countries, with Africa, South America, and Australia absent. Methodologically, 67% of possible agent-interaction-network combinations remain unexplored. These findings highlight urgent priorities for future research: implementing multi-validation approaches, expanding geographic coverage to underrepresented regions, strengthening alignment with regulatory frameworks, and developing integrated ESG models. Addressing these gaps is essential for realizing ABM's potential to inform sustainable finance policy and support global sustainability transitions.

Toward a Multi-Dimensional Assessment of Blockchain Decentralization: Empirical Evidence and MCDM-Based Integrative Framework

Mindaugas Juodis, Ernestas Filatovas, Remigijus Paulavičius

Institute of Data Science and Digital Technologies
Vilnius University

mindaugas.juodis@mif.vu.lt

Blockchain decentralization remains one of the most debated and multidimensional topics in distributed systems research. Although decentralization is often referred to as a core blockchain property, its quantitative assessment across different system perspectives (wealth, network, consensus, transactions, etc.) is still fragmented. This work contributes to the ongoing academic effort to establish a measurable and comparable framework for evaluating decentralization from multiple perspectives. We already examined wealth decentralization by analyzing cryptocurrency ownership concentration across major blockchains, including Bitcoin, Ethereum, and Layer-2 networks such as Arbitrum, Optimism, and Polygon. The results revealed structural concentration patterns in asset distribution when analyzed with classical inequality measures (e.g., Gini coefficient, Shannon entropy, Herfindahl–Hirschman Index). We also introduced a randomized group-sampling estimation approach that provides more robust concentration estimates across heterogeneous blockchain ecosystems. We also investigated transaction-level concentration patterns across UTXO- and account-based models using various metrics and compared Bitcoin, Ethereum, and multiple Layer-2 solutions, uncovering systemic dependencies between transaction-flow structure and the underlying ledger architecture. Building on these findings, our current research agenda expands to three additional aspects of decentralization: consensus, governance, and network. Each aspect is approached through distinct empirical

analyses – validator and staking concentration in consensus protocols, voting power and proposal diversity in governance systems, and node distribution across geographies, ISPs, and client implementations in network topology. Finally, we outline an ongoing effort to integrate all five aspects – wealth, transactions, consensus, governance, and network – into a unified Composite Decentralization Index (CDI). The CDI will be developed using multi-criteria decision-making (MCDM) methods, enabling weighted aggregation of diverse indicators and comparative evaluation across blockchain ecosystems. This approach is expected to yield an empirically grounded and methodologically coherent measure of holistic decentralization, bridging the gap between theoretical definitions and observable blockchain behavior.

A Dynamical Network Framework for Relational State Analysis

Rūta Juozaitienė[1], Ernst C Wit[2]

[1] Vytautas Magnus University, Kaunas, Lithuania
[2] Università della Svizzera italiana, Lugano, Switzerland

ruta.juozaitiene@vdu.lt

Dynamic networks are often analysed as binary structures, where ties are either present or absent. However, many real-world systems evolve through richer trajectories of relational states —such as acquaintance, friendship, and close friendship— whose dynamics contain substantive information. To address this, we introduce a continuous-time modelling framework for relational state networks, in which each edge occupies one of several substantively distinct states and evolves by transitioning between them over time. Transition intensities are driven by state-dependent covariates that might be decomposed into anchoring effects reflecting the influence of the current state and pulling effects associated with the attractiveness of the target state. Both linear and smooth non-linear dependencies are accommodated through a generalised additive model (GAM) representation. We consider two common observation schemes. When complete event histories are available, we employ a Cox-type partial likelihood with nested case–control sampling, yielding efficient estimation of both parametric and smooth covariate effects. When only repeated cross-sectional snapshots are observed, we derive a binary-case likelihood that has the potential to be generalised to a multi-state setting. Simulation studies demonstrate that the proposed estimators recover parameters with high accuracy across diverse scenarios. Furthermore, an empirical application to the Teenage Friends and Lifestyle Study illustrates that our approach reproduces substantive findings of established modelling techniques, while offering substantial gains in computational efficiency. Overall, the proposed framework provides a scalable approach to analysing persistent relational states. It retains the interpretability of classical network effects, generalises them to multi-state ties, and supports estimation under both ideal (full-history) and more realistic (panel) observation designs, thereby broadening the empirical scope of dynamic network analysis.

Clustering Dynamic Proximity Graphs to Profile Vessel Traffic Density over Time in a Baltic Sea Case Study

Robertas Jurkus[1], Povilas Treigys[2], Julius Venskus[1]

[1] Institute of Applied Mathematics
Vilnius University
[2] Institute of Data Science and Digital Technologies
Vilnius University

robertas.jurkus@mif.vu.lt

Maritime traffic in dynamically evolving coastal regions such as the Baltic Sea is characterised by continuously shifting vessel interactions that reflect operational intent, navigational constraints, and seasonal or traffic-driven changes in behaviour. Understanding these interactions not at the level of individual trajectories, but as evolving maritime interaction networks, is increasingly recognised as a critical foundation for situational awareness, traffic monitoring, and future autonomous navigation support systems. In this study, we model vessel interactions as time-indexed proximity graphs, where nodes represent vessels and edges are formed whenever two vessels are within a defined nautical-mile encounter threshold. Instead of analysing individual vessels in isolation, we extract connected graph components at each time step, representing localised, interaction-driven traffic formations. For each such component, we compute a rich set of graph-level structural descriptors, including node density, degree distribution statistics, clustering properties, connectivity and path efficiency measures, and centrality-based interaction intensity indicators. These feature vectors, one per traffic formation per time step, are then processed using unsupervised clustering methods, allowing us to discover and categorise recurring traffic regimes without imposing prior assumptions regarding vessel types, traffic rules, or temporal segmentation. The resulting clusters capture distinct and interpretable maritime traffic states, ranging from sparse and fragmented motion to dense, hub-like and coordinated interaction structures. This provides a data-driven characterisation of macroscopic traffic behaviour over time

and enables the identification of stable or recurrent traffic patterns that could aid in traffic monitoring, seasonal analysis, and maritime traffic complexity assessment. A Baltic Sea AIS (Automatic Identification System) case study confirms the practical viability and scalability of the proposed framework and highlights its potential as a foundation for higher-level maritime intelligence, including situational classification, strategic planning, or predictive traffic state modelling in future work.

Acknowledgements. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-MIP-24-117.

Reinforcement of Investment Decision Making Using Sentiment Analysis

Audrius Kabašinskas[1], Jurgita Černevičienė[1], Hamza Wasim[2]

[1] Kaunas University of Technology
[2] Sabanci University, Turkey

audkaba@ktu.lt

This study investigates whether sentiment analysis of financial news, retrieved through the Yahoo Finance API, can enhance the prediction of price movement directions across diversified asset classes including stocks, cryptocurrencies, and commodities. A sentiment indicator ranging from –1 to 1 is derived from news headlines and integrated with traditional financial metrics. Four machine learning models – XGBoost, Random Forest, Support Vector Machine (SVM), and Logistic Regression – are employed to predict directional price movements. To ensure transparency and interpretability, Explainable AI techniques (SHAP and LIME) are applied to decompose model predictions, revealing the contribution of sentiment versus financial features in driving investment signals. Using a multicriteria decision-making framework reinforced by machine learning predictions, the study evaluates the effectiveness of sentiment-driven indicators in guiding portfolio decisions. The results demonstrate the potential of combining sentiment analysis with interpretable machine learning to enhance decision-making frameworks in volatile and complex markets, while XAI ensures that model recommendations remain actionable and trustworthy for investors.

Acknowledgements. The presentation is related to project "Intelligent liquidity strategies" 02-020-K-0064 that is funded by the European Union Funds for the period 2021-2027 under the Measure No. 05-001-01-05-07 "Establishing a coherent system for the promotion of innovative activities".

Artificial Intelligence and Multimodal Data Fusion System for Assessing and Detecting Fraud in Applicants' Videos

Diana Kalibatiene, Algirdas Laukaitis, Kęstutis Normantas, Julius Jancevičius, Mindaugas Jankauskas, Dovilė Jodenytė

Vilnius Gediminas Technical University

julius.jancevicius@vilniustech.lt

When applying to international universities, an applicant must go through the admission process, which includes submitting documents about the school previously attended, documents confirming language skills, an application for admission, and a recording of an interview. The applicant's admission data is collected in different formats: paper format, scanned images, text documents (word, pdf), and video recordings. Processing such multimodal data is a time-consuming and human resource-intensive process, often carried out by one person, which is why selection decisions may be subjective and do not reveal the real level of preparation of the applicant. The aim of this research is to mitigate biases and enhance the precision of applicant assessments, ensuring that suitable candidates are not overlooked due to limitations in conventional manual evaluation methods by integrating multimodal data and applying advanced deep learning models for image analysis and natural language processing.

Automated video interviews (AVIs), which use machine learning (ML) algorithms to assess applicants, are becoming increasingly popular [1] because they improve efficiency and reduce the influence of human bias [1-4] and speed up the assessment process [5]. Such algorithms automatically infer applicant knowledge, skills, abilities, and other characteristics [6], such as personality traits [7]. While ML promises to efficiently and accurately infer applicant interviews, there is little empirical evidence to support the validity of algorithmic methods for applicant evaluation and personnel assessment [1]. Moreover, fairness problems in automatic interview assessment systems, especially video-based automated interview assessments, have less been addressed despite their prevalence in the recruiting field [3].

The data collected by VILNIUS TECH, such as certificates submitted for admission, copies of diplomas, and students' achievements in individual

study subjects, provide a unique opportunity to automate and unify the objective assessment of all applicants. By applying the newest deep learning models, we have the opportunity to search for connections between assessments in certificates, the nature of which varies in different countries and universities. Moreover, application of deep learning allows us to search for connections between the achievements of graduates from various countries and assess the possibilities of effectively studying one or another study program in Lithuania, and recommend a study program or field to students, depending on their achievements at school or a university in another country.

Acknowledgements: These results are part of the project "Artificial intelligence and multimodal data fusion system for assessing and detecting fraud in applicants` videos" (FAIR-VID). This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-25-14.

References

[1] Hickman, L., Tay, L., & Woo, S. E. (2024). Are automated video interviews smart enough? Behavioral modes, reliability, validity, and bias of machine learning cognitive ability assessments. Journal of Applied Psychology. https://doi.org/10.1037/apl0001236 Kim, C., Choi, J., Yoon, J., Yoo, D., & Lee, W. (2023). Fairness-aware multimodal learning in automatic video interview assessment. IEEE Access, 11, 122677-122693. https://doi.org/10.1109/ACCESS.2023.3325891.

[2] Oswald, F. L., Behrend, T. S., Putka, D. J., & Sinar, E. (2020). Big data in industrial-organizational psychology and human resource management: Forward progress for organizational research and practice. Annual Review ofOrganizational Psychology and Organizational Behavior, 7(1). https://doi.org/10.1146/annurev-orgpsych-032117-104553.

[3] Kim, C., Choi, J., Yoon, J., Yoo, D., & Lee, W. (2023). Fairness-aware multimodal learning in automatic video interview assessment. IEEE Access, 11, 122677-122693. https://doi.org/10.1109/ACCESS.2023.3325891

[4] Putra, B., Azizah, K., Mawalim, C. O., Hanif, I. A., Sakti, S., Leong, C. W., & Okada, S. (2024). MAG-BERT-ARL for Fair Automated Video Interview Assessment. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3473314

[5] İlhan, Ü. D., Güler, B. K., Turgut, D., & Duran, C. (2025). Opportunities and challenges of asynchronous video interviews: Perceptions of human resources professionals from Türkiye. PLoS One, 20(6), e0325932.

[6] Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: Why HR is set to fail the big data challenge. Human Resource Management Journal, 26, 1–11. https://doi.org/10.1111/1748-8583.12090.

[7] Rotolo, C. T., Church, A. H., Adler, S., Smither, J. W., Colquitt, A. L., Shull, A. C., Foster, K. B., & Foster, G. (2018). Putting an end to bad talent management: A call to action for the fi eld of industrial and organizational psychology. Industrial and Organizational Psychology, 11(2), 176–219. https://doi.org/10.1017/iop.2018.6.

Anonymization Corpus for Automated Anonymization/Encoding of GDPR-Relevant Data

Danguolė Kalinauskaitė[1,2], Milita Songailaitė[1,2], Justina Mandravickaitė[1,2], Justinas Dainauskas[1,2], Agnė Paulauskaitė-Tarasevičienė[3], Tomas Krilavičius[1,2]

[1] Vytautas Magnus University
[2] Centre for Applied Research and Development
[3] Kaunas University of Technology

tomas.krilavicius@vdu.lt

The The enforcement of the General Data Protection Regulation (GDPR) has significantly reshaped personal data processing practices. Its implementation has required data controllers to reassess and adapt the methodologies and tools previously used in handling personal data. These changes have led to a growing need for effective data anonymization, making the automation of this process increasingly important, particularly for data processing entities managing large volumes of data. However, the implementation of this task presents considerable challenges, as effective automation of data anonymization requires access to large datasets to efficiently train machine learning or deep learning models to identify and anonymize different categories of personal information. Furthermore, the training process often has to start entirely from scratch, since annotated corpora suitable for the intended training scenarios are scarce or, in some cases, unavailable.

To address the aforementioned issues, we present our project aimed at developing a specialized anonymization corpus of at least 10 million words, annotated with GDPR-relevant data categories that reflect general personal information pertaining to real-world individuals. The final annotated corpus will be used for automated data anonymization/encoding in accordance with GDPR requirements, as well as for training machine learning or deep learning models. The corpus annotations will encompass two types of data relevant under the GDPR framework: (1) data of general categories, comprising personal identifiers such as

names, surnames, usernames, identification numbers, phone numbers, demographic characteristics, locations, and other related attributes; and (2) data of special categories, including information concerning religious or philosophical beliefs, political views or memberships, sexual orientation, and other sensitive aspects. Data of the first category will represent up to 98% of all annotated data in the corpus, while those of the second category will constitute 2-5% of the total annotations. The final corpus will be composed of texts of different genres: 80% will consist of administrative documents from various fields, 10% - media articles, and 10% - scientific articles. The annotated corpus will be validated. For this task, tools will be developed that enable training machine learning or deep learning solutions (e.g., language technology solutions) using the anonymization corpus and demonstrating their performance with a certain level of accuracy.

The prepared anonymization corpus will be publicly available on dedicated open-access platforms. For this purpose, a methodology will be developed to hide personal information annotated in texts and, at the same time, ensure the integrity and coherence of the texts.

Acknowledgements: These results are part of the project "Developing an Anonymization Corpus" implemented from the funds of the Economic Recovery and Resilience Plan "New Generation Lithuania", project No. 02-100-K-0001. We thank Eglė Rimkienė for helping to prepare the presentation of the results.

Lithuanian Summaries Corpora for Automatic Text Summarization

Danguolė Kalinauskaitė[1,2], Milita Songailaitė[1,2], Justinas Dainauskas[1,2], Justina Mandravickaitė[1,2], Virginijus Marcinkevičius[3], Tomas Krilavičius[1,2]

[1] Vytautas Magnus University
[2] Centre for Applied Research and Development
[3] Vilnius University

danguole.kalinauskaite@vdu.lt

Every text conveys certain information that substantially defines the idea and essence of that text. Creating a concise representation of a source text while maintaining the essential information and the meaning of the content is known as text summarization. Manual text summarizing is a time-consuming and overwhelming task, especially when working with large volumes of text. Therefore, the ability to automatically obtain the most important information from texts is of particular importance for many different practical applications, such as text summarization for research purposes (e.g., to process and analyze large amounts of content), text classification, question answering, finding key points in legal documents or big paragraphs, analyzing messages conveyed by texts to detect disinformation in media or other sources, creating headlines, etc.

There are two primary types of text summarization techniques: (1) extractive summarization and (2) abstractive summarization. The extractive method uses existing sentences from a text and combines them to create a summary, while the abstractive technique generates new sentences to convey key ideas from the original content, i.e., rephrases information in a more concise and coherent manner. We present our work in progress on creating and validating Lithuanian summaries corpora, which are intended to be used as resources for training deep learning-based automatic summarization systems to produce high-quality summarized textual content. For our corpora, we use texts from 4 different areas, namely, media, law, health care, and

information technologies. Texts in these areas differ in various aspects, which in turn allows us to provide the necessary diversity of texts in terms of their content, structure, length and topics covered. To ensure a consistent corpora development process and measurable final results, a list of predefined requirements for corpora creation and validation was prepared. By presenting the outcomes of the corpora development process, we discuss different stages of preparing representative and well-balanced corpora which are based on clear text selection principles. We give a particular focus on the text annotation process, which determines the further successful use of the prepared new resources for training artificial intelligence models. Finally, we discuss the main challenges we have already encountered in collecting, processing, and annotating the texts needed for our corpora.

To assess the suitability of the created corpora for building automatic summarization systems, tools will be developed that allow deep learning solutions to be trained using the new resources and demonstrate their operation. We address this validation process as well.

Acknowledgements: These results are part of the project "Summaries Corpora for the Needs of Artificial Intelligence" implemented from the funds of the Economic Recovery and Resilience Plan "New Generation Lithuania", project No. 02-101-K-0001. We thank Eglė Rimkienė for helping to prepare the presentation of the results.

Few-Shot Isolated Sign Language Recognition with SlowFast and Prototypical Networks

Michał Kalinowski, Bożena Kostek

Faculty of Electronics, Telecommunications and Informatics
Gdańsk University of Technology

bozena.kostek@pg.edu.pl

Sign language is the primary means of communication for deaf and hard-of-hearing individuals worldwide. It is estimated that over 70 million people have significant hearing loss. Unlike spoken languages, sign languages are not standardized globally and exhibit significant variation across regions. Even in countries sharing the same spoken language, such as the United States and the United Kingdom, where English predominates, American Sign Language (ASL) and British Sign Language (BSL) differ markedly and are not mutually intelligible. Currently, around 300 distinct sign languages are in use globally. This diversity in sign languages creates a persistent communication barrier not only between sign language users and individuals without hearing impairments but also among sign language users from different regions within the same country. Moreover, sign language recognition systems often struggle with data scarcity, as available corpora are typically too small to support robust model training and generalization.

The present study addresses this limitation by framing sign recognition as a few-shot learning problem under the prototypical network paradigm. A modified SlowFast convolutional neural network is employed to extract spatiotemporal embeddings from video samples, facilitating metric-based comparison between support-set exemplars and query clips. Training is conducted on a subset of classes, while evaluation targets entirely unseen classes drawn from the LSA64 dataset, thereby rigorously testing generalization capability. On the held-out test split, this approach achieves 88% accuracy, demonstrating that prototype-based few-shot learning can effectively extract features from unseen data and thus successfully recognize novel sign classes with minimal examples.

These findings underscore the potential of combining powerful video feature extractors with metric learning for data-efficient sign language recognition. While the model excels at clustering most classes, it struggles to distinguish overlapping classes.

Future work could leverage margin-based loss functions and optimized hyperparameters to enhance performance, paving the way for more robust recognition systems.

Peculiarities of CT Volumetric Imaging Towards the Optimal Image-Guided Radiotherapy

Greta Karpavičienė, Algimantas Kriščiukaitis, Reda Čerapaitė-Trušinskienė, Robertas Petrolis, Diana Meilutytė-Lukauskienė, Renata Paukštaitienė

Dept. Physics, Mathematics and Biophysics
Lithuanian University of Health Sciences

algimantas.krisciukaitis@lsmu.lt

Ensuring selective irradiation of target tissues is one of the biggest challenges in radiotherapy. Methods and devices of Image-Guided Radiotherapy (IGRT) are elaborated with the aim of ensuring that the prescribed radiation dose is delivered accurately to the tumour while minimising exposure to surrounding healthy tissues. Technical solutions ensure a few-millimetre, or even sub-millimetre precision of the irradiation beam, while with currently used mechanical means of patient positioning, we can expect much bigger positioning deviations, reaching even centimetre range. Patient positioning deviation to a certain extent is related to changes in soft tissue density and volume, which change during the period of treatment. Therefore, the discovery of reliable reference structures in routinely performed daily Cone-Beam Computed Tomograms (CBCT) was one of the aims of this study. Having the reliable reference structures, we carried out the retrospective estimation of patient position deviation during the whole treatment cycle and evaluated possible dynamics of unwanted irradiation of tumour-surrounding critical organs. The study was conducted in patients with head and neck cancer treated in the Lithuanian University of Health Sciences Kaunas Clinics Affiliated Hospital of Oncology, Department of Radiotherapy. Patients' positioning was evaluated using volumetric images obtained by the CBCT machine integrated into the Halcyon V3.1 linear accelerator (Varian Medical Systems, Palo Alto, CA, USA). Custom-made algorithms of hard tissue segmentation and actual patient position estimation were elaborated in MATLAB (MathWorks, USA) environment. The hard

tissue structures in volumetric images, in particular the mandible and part of the skull, were segmented and adjusted using mathematical morphology algorithms. We found these structures as reliable reference landmarks for patient position estimation. We found the deviation of actual patient position ranging from 1 to 3,5 mm, which resulted in changes in irradiation ranging from 0,016 to 0,057 Gy/fraction in the planned target volume and in critical surrounding organs (e.g. larynx, parotid, etc.) as well. The values indicate that it can cause significant damage to the surrounding organs. In conclusion, we state that specially selected hard tissue structures can serve as reliable landmarks of patient position, while soft tissues eventually change. The development of more precise image-guided radiotherapy methods can significantly reduce the damage to tumour-surrounding organs.

AI and Smartwatches for Eating Behavior Detection: Supporting Dementia Caregivers

Gabrielė Kasputytė[1,2], Danylo Abramov[1,3], Paulius Savickas[1,2], Anton Volčok[1,3]

[1] Vytautas Magnus University
[2] Baltic Intelligence Solutions
[3] Centre for Applied Research and Development

gabriele.kasputyte@vdu.lt

The healthcare sector is facing major challenges due to population ageing, rising prevalence of chronic conditions, and the demand for personalised care. One critical issue is irregular nutrition, which is often linked to neurodegenerative diseases such as Alzheimer's and Parkinson's. Traditional dietary monitoring approaches that rely on self-reporting are impractical for older adults or individuals with cognitive decline, highlighting the need for automated and unobtrusive solutions. This study introduces an artificial intelligence–based eating behaviour recognition system that leverages smartwatch inertial sensor data (accelerometer and gyroscope) to detect eating episodes in real time with minimal user involvement. A prototype system was developed and evaluated using real-world data collected via a custom Garmin smartwatch application. Multiple machine learning methods (logistic regression, KNN, SVM, RF, GBM) and deep learning architectures (LSTM, GRU, 1-D CNN, CNN–LSTM) were compared. Feature importance analysis using SHAP further highlighted the most informative motion patterns for eating detection. Results demonstrated that a Random Forest (RF) classifier using 90-second windows achieved the best performance (F1 = 0.907). Deep learning models, especially CNN–LSTM hybrids, achieved competitive results when data were properly preprocessed, but did not surpass RF. The findings confirm that automated eating behavior detection from wearable inertial data is both feasible and practically applicable. Such systems hold strong potential for dementia care and rehabilitation, enabling continuous, non-invasive monitoring that supports caregivers while promoting early identification of nutrition-related health risks.

Acknowledgements. This project is funded by the European Union's NextGenerationEU program, No. 02-018-K-0499.

NO-GAP: AI-Based Analytics for Student Achievement Trends in Lithuanian Schools

Monika Katilienė[1,2], Anton Volčok[1,2], Rasa Erentaitė[3], Rimantas Vosylis[3]

[1] Vytautas Magnus University
[2] Centre for Applied Research and Development
[3] Kaunas University of Technology

monika.zdanaviciute@card-ai.eu

School managers in Lithuania face challenges in translating abundant educational data into actionable insights. Traditional e-diary reports provide only basic summaries, while national datasets require advanced analytic skills, limiting evidence-based decision-making. We present NO-GAP, an AI-powered analytics tool developed to analyze multi-year, population-wide student and school data. Integrating longitudinal datasets covering approximately 50,000 students from around 1,000 schools, NO-GAP applies AI-based clustering methods to detect patterns in student achievement across cohorts, schools, and subgroups. The tool identifies groups of students with similar learning trajectories, uncovers differences in performance between schools, and tracks achievement trends over time. By comparing results to national and municipal benchmarks, NO-GAP generates actionable insights and targeted recommendations, helping school managers understand which areas require intervention and which strategies are most effective. Using these capabilities, NO-GAP supports in-depth exploration of trends and profiles, enabling school managers to monitor performance, identify disparities, and implement evidence-based strategies to improve educational outcomes. The tool's visualisations and data summaries make complex information accessible and actionable for decision-making across multiple levels of school management.

Augmentative and Alternative Communication Personalization with PECS Cards: Evaluation of Reinforcement Learning Algorithms

Oskaras Klimašauskas, Asta Slotkienė

Vilnius Gediminas Technical University

oskaras.klimasauskas@stud.vilniustech.lt

Augmentative and alternative communication (AAC) systems, such as the Picture Exchange communication System (PECS), are widely used to support children with autism spectrum disorder (ASD) in real communication with adults. To improve their usability and learnability, the systems are digitalised. Artificial intelligence capabilities are applied to increase communication effectiveness. Most existing research works, based on AI models, try to build classifiers for the child's reaction states, where the data to train the models is fully labelled. However, data labelling is prone to subjective interpretations by the specialist. For each child with ASD, their individuality makes it impossible to determine the rule-based data and their decisions. The paper proposes the research and evaluation of reinforcement learning algorithms, where the agent learns an optimal strategy for presenting an appropriate PECS card to facilitate effective communication between a child and an adult. The goal is to investigate the performance of reinforcement learning algorithms (SARSA, SARSA(λ), Actor–Critic and Expected SARSA) for selecting a suitable PECS during augmentative and alternative communication. The investigation results demonstrate that the policy-gradient Actor–Critic algorithm outperforms other SARSA algorithms. When the agent started from position (1,1), Actor–Critic required an average of seven steps - 9.1% fewer than SARSA and Expected SARSA, and 5.4% fewer than SARSA(λ). In the second strategy, when the agent started from the matrix centre, Actor–Critic, SARSA(λ), and Expected SARSA each completed the task in three steps, while SARSA required approximately 33.3% steps. For AAC development, the Actor-Critic algorithm is the most effective approach, as it enables rapid convergence to the user's behaviour, minimising the

"cold start" problem and ensuring that the system effectively addresses the child's communication needs. Meanwhile, the Sarsa(λ) results confirm that in the PECS cards selection process, it is important to evaluate sequences of actions (choices of PECS cards), and not just individual steps, as this allows for faster learning of more complex communication chains.

Human Stress Detection from Body Movement Using Real-Time Video Feed

Marius Klumbys, Gintautas Šedys, Eglė Butkevičiūtė, Tomas Blažauskas

Faculty of Informatics
Kaunas University of Technology

gintautas.sedys@ktu.lt

Stress affects productivity, health, and safety, so practical tools for monitoring it are valuable. This study explores stress prediction possibilities on humans using real-time videos (web camera). Accurate video-only stress prediction is difficult because the model lacks direct physiological cues (e.g., HRV, EDA) that are among the most informative indicators of stress. Relying solely on skeletal motion and facial landmarks typically underperforms multimodal (physiology-augmented) approaches.

The aim of this study is to find whether it is possible to accurately predict stress using only video feed data. The project uses SWELL-KW dataset for training and validating model accuracy, it is a multimodal collection of posture, facial expression and physiology recordings captured during rest, neutral writing and stress tasks (time pressure, interruptions). The dataset features have been reconstructed, cleaned and combined so that every minute had consistent features and labels - the data was prepared to be reproduced with MediaPipe (body pose and face landmarks library). Gradient boosted tree with a minute long window was used for the model - before training, each person was normalized to cancel out their natural posture and facial differences, and simple history features, particularly rolling averages, were added. Currently, with leave-one-out cross validation, the model reached about 57% balanced accuracy and 61% ROC-AUC, which basically sums up how well the model separates stress vs. non-stress across all possible probability thresholds. An AUC of 100% would be perfect separation, while 50% is just a random guessing, so at 61% there is a lot of room to grow.

This work delivers a practical template for stress monitoring by relying on standard webcams and on-device processing. It offers an alternative to wearable sensors for tracking workload strain, giving organizations or individual users a non-intrusive way to monitor their well-being. These early results are encouraging and further investigations should focus into privacy-preserving on-device models, longitudinal evaluation, and robustness to occlusion and lighting.

Simulation of Cyber Incident Response Using Artificial Intelligence

Dmitrij Kolodynskij

Vilnius Gediminas Technical University

d.kolodynskij@gmail.com

Modern cybersecurity operations face increasing complexity, where static playbooks are often insufficient to handle dynamic and evolving threats. This research explores the simulation of cyber incident response using Artificial Intelligence to create adaptive, data-driven response frameworks. The proposed approach employs Large Language Models to automatically generate incident scenarios and response plans, simulating real-world decision-making under controlled conditions. By integrating AI-driven reasoning with mock operational environments, the simulation enables experimentation with various response strategies – ranging from automated execution to human-in-the-loop decision paths.

From an organizational perspective, cyberattacks are inevitable for a variety of reasons – organizations cannot identify every possible vulnerability in their systems and cannot fully eliminate the human factor, often regarded as the weakest link in cybersecurity. Incident response capabilities enable affected companies to detect, contain, and recover from security incidents efficiently, while well-structured response scenarios also support the wider community by helping to prevent similar attacks in the future. In this context, the generation of incident response plans through AI-based simulation becomes a practical necessity: it allows organizations to proactively develop, test, and refine their response strategies in a controlled, risk-free environment. This capability is essential for reducing response times, enhancing coordination among security operations teams, and improving overall cyber resilience.

Results demonstrate that AI-based simulation can enhance incident response readiness, support the testing of security playbooks, and facilitate continuous improvement in cyber defense strategies. The study contributes to the ongoing development of intelligent, scalable, and explainable cyber response systems.

Modelling of Cyber Security Attacks in Large Asynchronous Network Flows

Virgilijus Krinickij, Linas Bukauskas

Cyber Security Laboratory
Institute of Computer Science
Faculty of Mathematics and Informatics
Vilnius University

virgilijus.krinickij@mif.vu.lt

Cyber security is one of the most versatile subfields today. For modelling possible attacks on organisations network, professionals need to develop a system that would be capable of generating asynchronous network flows. These network flows would be used for the assessment of possible cyber events in the future to prevent attackers from breaching the organization. This work presents a practical, production-minded architecture for distributed network-flow collection on Linux hosts. The created system couples a message-driven control plane and enables on-demand capture, dynamic filter updates, and resilient return-path streaming of packet artifacts for storage and analysis. On endpoints, a Python client built on Scapy performs interface-level sniffing and applies BPF filters supplied asynchronously via RabbitMQ. Filters are created using predefined records in a database structure that are pushed as queue messages from the RabbitMQ broker server to the client machine. The client machine can take multiple instances of the network filter. The agent in the client machine, based on the filter, produces captured packets, which in turn are serialized and returned over callback queues, where the RabbitMQ server component persists them to PCAP files and logs event metadata. Also, time-based anomaly injection and cyberattack templates are supported for testing and experiments. The asynchronous aspect of the system comes from a predefined network configuration, which is a part of the overall system design. We assume that a predefined network configuration is a set of networks in a laboratory setup where we simulate real-life networks with different configurations and at different times. For this, we use Proxmox, a powerful open-source virtualization

platform. Operationally, Proxmox hosts provides virtual machine and Linux container (LXC) placement, while an Ansible playbook codifies provisioning for system automation. System automation is needed so that there would be less downtime in different environments where the system would be needed. For future cyberattack assessment in produced PCAP files we will use different algorithms like Dynamic Time Warping (DTW), windowed DTW, Needleman-Wunch, Smith Waterman. A possible machine learning functionality can also be added for high-throughput live data flow analysis in the RabbitMQ server machine. We also discuss web-exposed upload vulnerabilities, emphasizing why executable payloads must never be writeable in web-served paths and how defense-in-depth (MIME/extension whitelists, server-side execution blocks) prevents remote code execution. Altogether, the solution demonstrates a portable, automatable pipeline for remote capture and adaptive telemetry at scale, balancing performance, operability, and security.

Advancing Cybersecurity Competence Assessment Frameworks Integrating Human Factor Dimensions

Karina Kulikauskaitė, Karolis Baranauskis

Vilnius Gediminas Technical University

karina.kulikauskaite@vilniustech.lt

Cybersecurity has become a priority field in digital infrastructures because the number of sophisticated threats is increasing constantly. While significant progress has been made in developing technical defenses, the human factor persists as the most significant source of vulnerability. Cybersecurity competence assessment frameworks have traditionally emphasized technical knowledge and procedural skills while underrepresenting the human factor, such as emotional states, stress, fatigue, teamworking, etc. This research introduces an improved framework for evaluating cybersecurity competence with explicit attention to human factors. The framework integrates cognitive, behavioral, and situational dimensions into a multidimensional evaluation model. The framework was tested through controlled experiments involving participants from the Security Operations Center of an international enterprise, encompassing both technical and non-technical roles as well as diverse levels of professional experience. Cybersecurity competence was evaluated through quizzes, real-world scenarios, and simulations, while participants' physiological parameters, including heart rate variability (HRV), respiratory rate, and peripheral capillary oxygen saturation (SpO_2), were simultaneously monitored throughout the experiment. The quantitative analysis of participants' responses, emotional states, and physiological measurements was made, enabling the identification of the levels of stress-inducing questions and cases. The findings indicate that HRV and SpO_2 can serve as predictive indicators of human factors and performance outcomes in competence assessment tests. Experiment confirmed that integrating human factor dimensions increases the validity and predictability of cybersecurity competence assessment models, while an improved framework provides a holistic understanding and contributes to the development of adaptive training methodologies.

Efficiency of Vision-Language Models on Image Caption Generation: Case of Lithuanian Language

Tautvydas Kvietkauskas[1], Pavel Stefanovič[1], Vytas Mulevičius[2], Artūras Nakvosas[2,3]

[1] Faculty of Fundamental Science
Vilnius Gediminas Technical University
[2] Neurotechnology
[3] Institute of Data Science and Digital Technologies
Vilnius University

tautvydas.kvietkauskas@vilniustech.lt

Vision-language models have gained significant popularity in recent years because they can simultaneously address problems related to image and text analysis. They combine computer vision and natural language processing techniques to perform tasks such as image description, picture-based question-answering systems, and multi-modal search. Recently, these models have become increasingly important in developing advanced applications, such as autonomous vehicles, medical diagnostics, and content management. Many Vision-language models are adapted to the most popular languages, such as English, Spanish, and Chinese, but lack integration with less popular languages, like Lithuanian. This study analysed the effectiveness of various Vision-language models, such as BLIP, Gemma3, Qwen, and others, using pre-prepared data collected from Lithuanian news portals. Thus, to expand the research data, the Flickr8k dataset was selected, and its captions were translated into Lithuanian. The research dataset consists of photos associated with news articles and their corresponding captions below each image. Given that many models cannot generate captions in Lithuanian, a study was conducted to translate captions from Lithuanian to English and vice versa. Traditional evaluation metrics, such as BLEU, METEOR, ROUGE, and BERTscore, were used to evaluate the research results. The results of the experimental investigation show that models trained with languages of smaller countries, such as Lithuania, can be sufficiently accurate.

Graph Topology-Based District Heating Network Preparation for Probabilistic Resilience Assessment

Karolis Lašas, Sigitas Rimkevičius

Lithuanian Energy Institute

karolis.lasas@lei.lt

Systems of various kinds must be able to adapt to the changing nature of technologies. New inventions replace old ones and introduce new systems. This is especially true for systems of high importance - those that directly affect people's health and well-being, such as district heating systems (DHS). For example, in northern countries, heating is not just a convenience but also a necessity. The only commonality between new and old systems is that they ultimately fail due to different factors. Some factors can be predicted in advance (such as aging infrastructure) and addressed beforehand. In contrast, others are unpredictable (like extreme temperature changes, war, cyber-attacks, etc.) - these are known as High Influence Low Probability (HILP) events. To assess how well a system is prepared to handle a HILP event, a metric called resilience can be used. Resilience is a system's ability to absorb the impact of a disruption, survive its consequences, and recover by restoring the system to its previous or improved state. However, there is no single, unified definition of resilience, nor is there a single way to measure it. Some methods involve full thermal-hydraulic simulations of DHS and digital twins, while others use various system attributes, such as flexibility or robustness, to create a single resilience metric. This research uses a graph representation of the district heating network (DHN) and employs graph topology metrics to evaluate pipeline sections, which will later be applied for probabilistic resilience assessment. Cluster analysis and partitioning are also used to assign nearest consumer nodes to the corresponding pipeline sections, as direct connections are not available in the data. The dataset used in this research is open-source GIS (Geographical Information System) data, which includes various details about the district heating network of Utena city in Lithuania. A graph of the DHN was created with 1518

nodes and 1504 edges, featuring over 1100 unique pipeline sections serving customers in Utena city. Physical properties of each section, such as length, pipe thickness, and diameter, along with graph topology metrics like centrality and connectivity, are used to prepare a dataset describing the network for resilience assessment. While resilience evaluation is essential for ensuring secure and reliable systems, it should not be the final step. To build a truly resilient system, one must assess its current preparedness for HILP events and use that information to make improvements. Therefore, future research will focus on system optimization based on resilience assessment results.

Leveraging LLM Ensembles for KG-Grounded Narrative Extraction: The Case of Disinformation and Trustworthy News

Justina Mandravickaitė

Vytautas Magnus University

justina.mandravickaite@vdu.lt

The rise of disinformation in the media presents significant challenges to public understanding as well as democratic processes. With rapid online dissemination, distinguishing between reliable journalism and deliberately misleading content has become extremely important. Therefore, this study examined narrative construction patterns in disinformation and trustworthy news through analysis of English news articles covering selected international events (2015-2023). We evaluated knowledge graph (KG) grounding for narrative extraction and compared grounded vs. ungrounded variants across events, causal links and frames with role-specialised two-model ensembles per structure type (extractors + formatter). We used Mistral-Small-3.1-24B as a formatter across tasks. The following ensembles of smaller LLMs were used as extractors: entity linking (Qwen-2.5-72b + Llama-3.1-8b), framing (Gemma-3-27b-it + Qwen-2.5-72b), whole narrative extraction (Gemma-3-27b-it + Mistral-Small-3.1-24b). We also assessed a 'critic' component for missing links, contradictions and inconsistencies and applied Mistral-Small-3.1-24b + Gemma-3-27b-it for this task. On a human-annotated set, event micro-precision reached 0.87 for disinformation and 0.75 for trustworthy news. Furthermore, a grounded Mistral-Small-3.1-24b variant resulted in denser event/relation graphs than its ungrounded counterpart and larger ungrounded Llama-3.3-70B-Instruct, also aligned better to the KG schema, which integrated Abstract Meaning Representation (AMR) parses and FRED (an automatic system that derives RDF/OWL ontologies and linked data from natural-language text). Causal link detection was recall-limited (ranking-insensitive) with

modest precision (micro 0.50/0.45 for disinformation/trustworthy news, accordingly), while 'critic' performance was ranking-limited with large oracle headroom, achieving moderate precision (0.54–0.57) across classes. Also, KG grounding improved gap-finding, raising missing-link coverage from 2-4% (no KG) to 54-57% and up to 100% on an annotated sample. An entity-linking ablation with a simple two-way consensus reached full coverage without disagreement on the targeted sample. Taken together, these results support a practical recipe in which KG grounding and ensembles of smaller LLMs deliver competitive quality at a manageable cost.

Introducing Quantum Machine Learning: Where AI Meets Quantum Computing

Marco Marcozzi, Ernestas Filatovas, Remigijus Paulavičius

Institute of Data Science and Digital Technologies
Vilnius University

marco.marcozzi@mif.vu.lt

The already revolutionary capabilities of Artificial Intelligence (AI) are getting a massive upgrade by incorporating the potential of quantum computing. This merging of two powerful fields creates Quantum Machine Learning (QML), a new approach that can dramatically enhance problem-solving. This talk explores the intersection of these two technologies, reviewing where machine learning techniques are being adapted for use on quantum systems and what advantages these hybrid techniques bring to solve challenges in many fields. We categorize and analyze current research to identify the most common applications and the specific algorithms that researchers are using to achieve these advances across different areas.

Acknowledgements. This research was funded by the Research Council of Lithuania under Grant Agreement No. S-ITP-25-5. The authors are also thankful for the HPC resources provided by the IT Research Center of Vilnius University.

Forecasting the Student Pipeline for Lithuanian Universities

Jurgita Markevičiūtė, Alfredas Račkauskas

Institute of Applied Mathematics
Vilnius University

jurgita.markeviciute@mif.vu.lt

The research analyses demographic and educational trends from 2000 to 2025 and projects future student flows to 2036 and 2043. Using national and regional data, it finds a long-term decline in the number of schoolchildren and graduates due to low birth rates, migration, and ageing populations, though slight stabilisation has appeared in recent years—mainly in Vilnius, Kaunas, and Klaipėda districts. The share of graduates choosing Lithuanian universities dropped sharply after 2009 but is gradually recovering. Forecasts suggest around 9,000–11,000 graduates will enrol in Lithuanian universities annually, with Vilnius, Kaunas, and Klaipėda maintaining dominance and smaller regions facing ongoing decline. The study concludes that universities must adapt to demographic shifts, strengthen regional access, and expand international recruitment to sustain enrollment stability.

Moving Forward Sustainable National Forest Inventory in Lithuania with Cutting-Edge Technologies: From Prototype to Wide-Scale Adoption

Arnas Matusevičius, Nerijus Šakinis, Linas Urbonas, Gabrielė Kasputytė, Tomas Krilavičius

Centre for Applied Research and Development
Vytautas Magnus University

arnas.matusevicius@vdu.lt

The Lithuanian National Forest Inventory (NFI) plays a pivotal role in advancing sustainable forest management, evidence-based policymaking, and climate action through systematic forest data collection and analysis. The NFI Information System is an advanced digital infrastructure built on a robust MS SQL Server database and client–server architecture, enabling efficient data processing, spatial analysis, and strategic decision-making.

In order to ensure long-term national forest strategic goals and to meet evolving national (Forest Law, data lake and archival data orders) and EU legal requirements (LULUCF Regulation and the forthcoming EU Forest Monitoring Regulation), two key modernization directions have been initiated. The first focuses on transitioning from the currently used file-based system to a centralized NMIIS platform, ensuring data consistency, accessibility, and interoperability across all levels of forest monitoring. The second involves modernizing the NFI Information System itself – upgrading the existing prototype into a fully scalable, centralized, and intelligent platform designed to meet the future demands of sustainable forestry, data integration, and real-time analytics.

The new architecture will enable seamless integration with emerging technologies, including AI-driven analysis, machine learning models, and remote sensing data sources. Field crews will benefit from mobile-enabled tools for immediate data validation, automated calculations, and assisted measurements. Beyond technological advancement, this

modernization underscores Lithuania's commitment to innovation, efficiency, and digital resilience in sustainable forest governance.

Acknowledgements. This research is funded by the Horizon Europe Framework Programme (HORIZON) under the call Teaming for Excellence (HORIZON-WIDERA-2022-ACCESS-01-two-stage) – Creation of the Centre of Excellence in Smart Forestry "Forest 4.0", No. 101059985. It is co-funded by the European Union under the project FOREST 4.0 – "Ekscelencijos centras tvariai miško bioekonomikai vystyti", No. 10-042-P-0002.

Correlation-Based Log Data Analysis for Cybersecurity Incident Detection

Dalius Mažeika, Dominykas Volodka

Vilnius Gediminas Technical University

dalius.mazeika@vilniustech.lt

Sophisticated cybersecurity incidents are constantly increasing. Traditional detection methods frequently fail to identify multi-stage or distributed attacks, primarily due to the high volume, heterogeneity, and decentralized nature of log data, which remains a critical source of evidence for detecting anomalous activities within networked systems.

This research investigates correlation-based techniques for analyzing distributed log data to enhance cybersecurity incident detection. Two datasets were employed in the investigation: a publicly available dataset containing Windows-based authentication events, and a custom dataset generated from log data collected within a virtual IT infrastructure subjected to controlled attack scenarios. Four log aggregation methods were applied to consolidate the data into a unified dataset. Depending on the characteristics of the data, statistical, time-based, model-based, and behavior-based aggregation techniques were employed. The aggregated data were normalized using either the Min-Max scaling approach or the StandardScaler method to ensure comparability across features. For the correlation analysis, Pearson correlation and Kendall's rank correlation were applied. In addition, Euclidean distance and cosine similarity measures were applied to further examine relational patterns within the data.

A modular testbed was developed to conduct the experiments, and a series of tests was carried out to examine the correlation between different types of log data. The analysis focused on the relationships between network traffic logs, DNS logs, process logs, and authentication logs. Various cyberattacks were simulated within the virtual environment to assess the method's effectiveness under realistic threat scenarios. The correlation between log data from key system components – including the IIS server, firewall, VSFTPD, SSHD service, and system audit logs—was thoroughly examined. The experimental results indicate that, with an appropriate configuration, the proposed method can effectively detect anomalies and identify malicious activities.

X-Ray Lung Diseases Classification Using Deep Neural Network

Edita Mažonienė, Dmitrij Šešok

Vilnius Gediminas Technical University

edita.mazoniene@vilniustech.lt

Numerous lung illnesses' prevalence has significantly increased over the past three decades due to population growth, harmful environmental influences, and an increase in life expectancy, according to researchers who have examined health and patient indicators as well as associated risk factors for various chronic respiratory diseases. As a result, hundreds of millions of people globally, across all age groups, suffer from a variety of respiratory illnesses, and it is also noted that patient ages are generally decreasing. Long-term clinical outcomes, including early mortality, depend on early detection and treatment of chronic respiratory disorders. Appropriate diagnostics are necessary to promptly diagnose respiratory diseases; in this instance, the most common (regular) radiological examination is recommended, and an initial diagnostic X-ray is performed. Radiologists' interpretation of radiography images remains a significant concern due to human ability to detect subtle visual features in the images, even though radiography can be completed quickly and widely due to the commonality of chest radiology imaging systems in hospitals. Numerous advancements in the field of image analysis and categorization using deep neural networks have made progress. DNN-based systems surpass many traditional learning models because they are getting more accurate while working with large amounts of data.

The findings of this article's application of DenseNet to categorise 14 different diseases from chest X-ray pictures from the NIH (National Institutes of Health) ChestX-ray14 dataset were obtained. The main goal was to enhance the accuracy of lung disease differentiation while using a pretrained DenseNet model with different Data Preprocessing and Data Augmentation techniques. We tested the suggested model using evaluation measures like recall, precision, and F1-score. The suggested deep learning model has shown that three out of fourteen diseases: Cardiomegaly (AUC 0,90195), Emphysema (AUC 0.91859) and Effusion (0.900281) increased in classification accuracy.

Not All Listeners Are the Same: Acoustic Cue Use in Synthetic Speech Quality Judgments

Gerda Ana Melnik-Leroy, Gediminas Navickas

Institute of Data Science and Digital Technologies
Vilnius University

gerda.ana.melnik@gmail.com

Standard TTS (text-to-speech) evaluation pipelines often assume a homogeneous "typical listener," risking misleading conclusions about perceived quality. We test this assumption by comparing congenitally blind and sighted listeners on a ternary AX discrimination task for assessing synthetic speech quality (using Lithuanian neural TTS).

On each trial, two renditions of the same word were presented and listeners indicated whether A was more distorted, X was more distorted, or both sounded the same. Using three synthesis quality levels (LOW, MEDIUM, HIGH), we defined two conditions: LOW–HIGH (easier as larger quality gap) and LOW–MEDIUM (harder as small quality gap). Both groups performed better in the easier condition, yet they diverged as the quality gap narrowed: sighted listeners showed reduced accuracy, whereas blind listeners were comparatively stable. Crucially, difficulty was stimulus- and group-specific: the same lexical items shifted between "easy" and "hard" across groups, rather than increasing uniformly with nominal task complexity.

To probe the perceptual mechanisms used, we conducted item-level acoustic analyses contrasting cue differences between "easy" and "hard" items for each group. We performed segment-level (phoneme) annotation for each token, extracted acoustic measures per segment under each quality condition, and computed per-cue quality deltas (LOW–HIGH or LOW–MEDIUM) that were z-normalized. For each group, we ranked words by accuracy, took the top-10 "easy" and top-10 "hard," and plotted, for each listener group and cue, the difference between the mean z-normalized quality deltas of the top-10 easy and top-10 hard words. The results showed that the same acoustic cues that were

informative for one group were uninformative, or even confusing, for the other, indicating distinct perceptual strategies. This tells us that synthetic speech isn't perceived the same way by all listeners. Listener background shapes how speech is judged, even when intelligibility is high. Accordingly, TTS evaluation should not collapse across heterogeneous listeners or across items. Protocols that report group-wise and item-level results, and inspect cue separability can expose asymmetric confusions that standard aggregate scores miss. Treating listener diversity as a first-order factor is necessary to avoid mischaracterizing model quality and to ensure that improvements generalize across populations.

How Evolutionary Algorithms Can Efficiently Explore and Exploit the Search Space

Marjan Mernik

University of Maribor, Slovenia

marjan.mernik@um.si

Evolutionary Algorithms mimic nature with mechanisms such as selection, crossover and mutation to solve various optimisation problems. To properly apply Evolutionary Algorithms, a deep understanding of various selection, crossover and mutation operators is required. However, exploration and exploitation are crucial steps and even more essential concepts for any search algorithm. On the other hand, these fundamental concepts are not very well understood among practitioners using evolutionary algorithms. Furthermore, how to measure exploration and exploitation directly is an open problem in Evolutionary Computation. In this talk, I will first introduce the basic ingredients of every evolutionary algorithm and point out many problems and mistakes inexperienced users face, as well as different applications of evolutionary algorithms. In the second part of my talk, our novel direct measure of exploration and exploitation will be explained as based on attraction basins — parts of a search space where each part has its own point called an attractor, to which neighbouring points tend to evolve. Each search point can be associated with a particular attraction basin. If a newly generated search point belongs to the same attraction basin as its parent, the search process is identified as exploitation; otherwise, as exploration. In the last part, I will mention some open problems regarding computing attraction basins for continuous problems.

Evolution of Artificial Intelligence in Radiological Detection of Lumbar Disc Hernias: Emphasising Trustworthiness and Explainability

Jolanta Miliauskaitė[1], Asta Slotkienė[1], Susana Moleirinho[2], António Fernandes[2], Luís Mendes Gomes[3,4], José Machado[3]

[1] Institute of Data Science and Digital Technologies
Vilnius University
[2] SliceD Group Research Center, Portugal
[3] Informatics Department, ALGORITMI, LASI
University of Minho, Braga, Portugal
[4] Informatics Department, Faculty of Sciences and Technology
University of the Azores

jolanta.miliauskaite@mif.vu.lt

It is widely recognised that artificial intelligence (AI) is profoundly transforming medical imaging, particularly in the radiological detection of lumbar disc hernias [1-2]. This study presents a comprehensive bibliometric analysis based on global research articles extracted from the Web of Science database, aiming to map out the advances made in radiology concerning lumbar disc hernia detection, with an emphasis on trustworthiness and explainability [3-4]. The main contributions and findings have been identified, systematised, and visualised through keyword mapping of relevant AI techniques applied to this domain. These methods primarily facilitate diagnostic accuracy enhancement, automated segmentation, and classification of lumbar spine structures, thereby addressing clinical challenges such as subjective image interpretation and inter-observer variability. A critical focus is placed on the evolution of frameworks ensuring AI trustworthiness, including robustness, fairness, privacy compliance, and clinical reliability, alongside approaches to explainable AI (XAI) that promote transparency by visualising model decision-making processes, such as heatmaps highlighting key anatomical regions on MRI scans. These features are vital for promoting clinician confidence and ensuring ethical, safe AI deployment in clinical workflows.

The present study reveals main trends indicating progressive integration of multimodal data, including clinical, imaging, and genomic information, to enhance diagnostic precision and patient stratification. Additionally, it underscores the challenges faced, such as heterogeneous imaging protocols, limited availability of high-quality annotated datasets, and the need for standardised validation practices. The results support researchers and clinicians by providing valuable insights into AI applications in lumbar disc hernia radiology, guiding future investigations focused on developing robust, interpretable, and clinically relevant AI systems. This work underlines the indispensable role of explainability and trustworthiness as complementary pillars underpinning the responsible adoption of AI technologies, ultimately advancing patient care and resource efficiency in modern radiology.

References

[1] Wang, H., Jin, Q., Li, S., Liu, S., Wang, M., & Song, Z. (2024). A comprehensive survey on deep active learning in medical image analysis. Medical Image Analysis, 95, 103201.

[2] Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. Zeitschrift fuer medizinische Physik, 29(2), 102-127.

[3] Donthu N, et al. How to conduct a bibliometric analysis: an overview and guidelines. *J Bus Res*. 2021;133:285–296.

[4] Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of informetrics, 11(4), 959-975.

Exploring the Behaviour of the Descent-Ascent Heuristic Optimization Algorithm

Alfonsas Misevičius, Gintaras Palubeckis, Dovilė Verenė

Department of Multimedia Engineering
Kaunas University of Technology

alfonsas.misevicius@ktu.lt

In this work, the main goal is to present some results of the computational experiments with a heuristic optimization algorithm based on the descent-ascent (DA) principle. The descent-ascent principle-based algorithm was recently proposed in [1]. The core idea of this algorithm is that the greedy improving operations (moves between solutions) are interchangeably combined with the random non-improving moves (perturbations) during the optimization process. It should be noted that our context is the minimization problems, so the descent procedure means the improvement process, while the ascent procedure should be linked to the random perturbation operations. As a case study, we investigate the well-known, NP-hard combinatorial optimization problem – the quadratic assignment problem (QAP) [2]. This problem has a long history; it has quite many important practical applications and still continues to attract the attention of many researchers. The experiments are on the basis of the component-based paradigm [3], where the important algorithmic features (parameters) of the DA algorithm are chosen and examined. In particular, the following features/parameters were investigated:

- the total number of iterations of the DA algorithm;
- the number of iterations of the tabu search (TS) procedure within the iterated tabu search algorithm;
- the number of iterations of the iterated tabu search (ITS) algorithm within the DA algorithm;
- the strength of the perturbations (mutations).

In the experiments, we have used challenging benchmark data instances from [4] and [5]. As the main algorithm quantitative performance

criterion, we adopt the average percentage deviation (θ^{-}) of the objective function, which is calculated by the following formula: $\theta^{-}=(z^{-}\text{-BKV})/\text{BKV}\times 100[\%]$, where z^{-} is the average objective function value and BKV denotes the best known ((pseudo-)optimal) value of the objective function. The obtained results of the conducted experiments demonstrate how the methodical reconfiguration of the selected particular components/parameters – guided by the designer's expertise – well improves the overall performance of the DA algorithm. The findings indicate that, among the examined components, the total number of iterations of the DA algorithm and the number of iterations of the iterated tabu search algorithm within it play the most crucial role in achieving high efficiency. The main conclusion is that – by extracting different particular components and options of the algorithm – we can study their impact and influence on the behaviour of the algorithm. Also, by combining them in a proper way, we can reconfigure the given algorithm; we can find the most promising redesigned algorithm configuration/architectures and build new powerful algorithm variants.

References

[1] Misevičius, A., Palubeckis, G., Verenė, D., & Žekienė, G. (2025). A descent-ascent principle based algorithm for the quadratic assignment problem. In Information and Software Technologies, 31st International Conference, ICIST 2025, Proceedings, Communications in Computer and Information Science (CCIS), Springer, Cham, in press.

[2] Çela, E. (1998). The Quadratic Assignment Problem: Theory and Algorithms, Kluwer, Dordrecht.

[3] Hoos, H.H. (2012). Programming by optimization. Communications of the ACM, 55(2), 70–80. https://doi.org/10.1145/2076450.2076469.

[4] De Carvalho, Jr., S.A., & Rahmann, S. (2006). Microarray layout as a quadratic assignment problem. In German Conference on Bioinformatics, GCB 2006, Lecture Notes in Informatics – Proceedings, Volume P-83, Huson, D., Kohlbacher, O., Lupas, A., Nieselt, K., & Zell, A., Eds., Gesellschaft für Informatik, Bonn, pp. 11–20.

[5] Taillard, E.D. (1991). Robust taboo search for the QAP. Parallel Computing, 17, 443–455. https://doi.org/10.1016/S0167-8191(05)80147-4.

A Fly in the Ointment. Multiword Expressions and Their Challenges for NLP Tools and Tasks

Verginica Barbu Mititelu

Romanian Academy Research Institute for Artificial Intelligence "Mihai Drăgănescu", Romania

vergi@racai.ro

Commonly referred to as expressions, locutions, idioms, or constructions, multiword combinations that exhibit idiosyncratic behaviour across linguistic levels constitute a pervasive and well-recognised challenge in natural language. Within computational linguistics, such constructions have long been identified as a significant 'bottleneck' for Natural Language Processing. Despite the substantial progress brought about by large-scale language models, the robust and systematic handling of these non-compositional units remains an open problem. In this talk, I will discuss the computational challenges posed by these phenomena and present the broader efforts of the international research community dedicated to their linguistic and computational modelling.

A Hybrid CNN and Reinforcement Learning Approach for Parameter Estimation from Magnetic Tweezer Images

Tautvydas Naudžiūnas, Guoda Perminaitė, Povilas Treigys

Institute of Data Science and Digital Technologies
Vilnius University

tautvydas.naudziunas@mif.vu.lt

Extracting algorithm parameters based on visual outputs presents a significant challenge, as the relationship between the appearance of an image and the underlying parameters that generated it is often non-linear and lacks clear statistical correlation. Traditional supervised learning approaches, such as Convolutional Neural Networks, struggle with this task due to their reliance on direct feature-to-parameter mapping and can be harmed by noisy inputs which is detrimental especially in magnetic tweezer experiment data, where the images are both noisy and low resolution. While powerful, transformer-based models are often impractical for this domain due to their demand for large datasets.

To address these limitations, we propose a hybrid framework that combines Convolutional Neural Networks with Reinforcement Learning for precise parameter regression. The method follows a multi-stage pipeline: first, a DnCNN-based denoising block is used to enhance the quality of the input image. By adaptively adjusting denoising parameters based on validation loss, this block achieved better performance than traditional denoising techniques in the context of analyzing diffraction ring images from magnetic tweezers experiments. Next, a feature extraction CNN then processes this cleaned image to encode its characteristics into a compact representation. It serves as the initial state for a reinforcement learning model, specifically a policy gradient reinforcement learning approach, which predicts and refines the parameter vector. The agent is fine-tuned through a feedback loop where its predictions are executed

by the external algorithm, and the resulting performance metric is used as a reward signal to guide further learning.

Experimental results demonstrate that our framework achieves superior regression accuracy compared to conventional CNN models, showing consistently lower error and greater prediction stability, while avoiding the data requirements of transformer models, making it suitable for a wide range of applications with limited training data. Due to the denoising stage the proposed method is especially effective in domains dealing with low-resolution or noisy visual data such as microscopy.

Acknowledgements. The authors are thankful for the high-performance computing resources provided by the ITAPC of Vilnius University.

Initial Study of Lithuanian Emotional Speech Synthesis

Gediminas Navickas, Gražina Korvel, Gintautas Tamulevičius

Institute of Data Science and Digital Technologies
Vilnius University

gediminas.navickas@mif.vu.lt

Emotional speech synthesis is a complex area of research that aims to generate speech that sounds natural and conveys human emotions. Despite the rapid progress of neural text-to-speech (TTS) methods, synthesis of emotional expression poses significant challenges in all languages, even high-resourced ones. The main challenges are related to still clearly undefined acoustic features of emotions, different levels and types of emotions (e.g., cold anger and hot anger), mixed emotions, limited interpretability and control of the emotional speech synthesis process. Lastly, the absence of an emotional speech corpus also restricts the capabilities of modern models. For low-resource languages (such as Lithuanian), these challenges and tasks become even more complex.

Recent literature has identified a shift from traditional rule-based and statistical parametric methods to deep generative approaches. Emotional TTS systems are based on deep neural networks (DNNs), variational autoencoders (VAEs), generative adversarial networks (GANs), transformers, and diffusion models. These models lead to the following emotional speech synthesis strategies:

- Explicit training of models using an emotionally labelled speech corpus.
- Transfer learning of emotional speech, thus avoiding the need for large amounts of data.
- Semi-supervised training methods, based on learning from both unlabelled and labelled data.

To achieve synthesis control and interpretability, another paradigm should not be dismissed: modifying neutral synthesized speech to provide the desired emotional content. This paradigm would require

a detailed analysis of emotional speech, a large corpus of emotional speech data, and additional models for the speech transformation.

Initiating the study of emotional Lithuanian speech synthesis, we began by assessing the State-of-the-Art methods in speech synthesis, the availability of emotional speech corpora, and model transferability. This report summarizes the main trends, methods, and challenges identified in recent studies, outlining how these insights can be used in the development of emotional speech synthesis for the Lithuanian language.

Forensics of Corpus Messages in Popular Android Device Applications

Matas Niewulis, Virgilijus Krinickij, Linas Bukauskas

Cybersecurity Laboratory
Institute of Computer Science
Faculty of Mathematics and Informatics
Vilnius University

matas.niewulis@mif.stud.vu.lt

With the rapid evolution of mobile technologies, smartphones have become an essential part of modern communication and personal data storage. Increasing functionality, accessibility, and integration of mobile devices with cloud-based ecosystems have significantly expanded the scope and sensitivity of the data they handle. The Android operating system dominates the global market, serving as the foundation for a wide variety of devices with differing architectures, firmware, and storage structures. As devices used by a user accumulate data containing personal communications, browsing history, application interactions, and metadata, they have become a critical focus for digital forensics and data-driven research. Due to frequent system updates, encryption mechanisms, and diverse manufacturer implementations, retrieving and analyzing Android data remains a complex and often inconsistent process. Existing tools are typically tailored for specific device types or Android versions, resulting in compatibility limitations and only partial data extraction. Such limitation creates a strong need for a universal and automated solution that can reliably extract and process data across different Android environments.

To address these challenges, this research aims to model and develop a Linux-based solution for more automated data acquisition and analysis from Android devices. The provided solution integrates various technological components to establish communication with the target device through Android Debug Bridge, enabling full access to its file system and corpus. The proposed solution would capture a complete physical bit image of the device's data partition and organize

extracted information into structured directories for further processing. Key data sources include a corpus which in turn undergoes analytical processing using a Naive Bayes classifier to identify correlations in an aim to visualize relational structures within the corpus. Unlike manual examination methods, an automated workflow would reduce human error and significantly improve the efficiency and repeatability of mobile data analysis. The classifier's output and analytical summaries would be presented to the analyst in a human readable format, allowing clear interpretation and standardized reporting. The research outcome would result in a more stable and automated data extraction and evaluation for forensics analysts.

Cluster-Wise Modelling of Migrations at Municipality Level in Poland: Some Observations and Conclusions

Jan W. Owsiński[1], Jarosław Stańczak[1], Przemysław Śleszyński[2], Rafał Wiśniewski[2]

[1] Systems Research Institute
Polish Academy of Sciences
[2] Institute of Geography and Spatial Organization
Polish Academy of Sciences

Jan.Owsinski@ibspan.waw.pl

We present some results from a project, devoted to modelling and analysis of internal migration flows in Poland at the municipality level (some 2500 units). The basic model analysed makes migration linearly dependent upon unemployment. The dependence is positively verified, and the associated maps, corresponding to the consecutive years over two decades (2003-2022) provide a very pungent spatial image.

Despite the positive verification of this simple model, its statistical characteristics are rather feeble. Hence, a study was undertaken with analogous models identified for subsets (clusters) of municipalities, employing a simile of classical k-means. Given the known dependence of the outcome from the k-means-like procedure upon the starting point, various initial configurations were considered. The results are exemplified and some broader conclusions drawn therefrom, related to a wide scope of applications of clustering, in, for instance, artificial intelligence and machine learning.

Radiomic Texture-Based Clustering of Myocardial Damage Severity from Cardiac MRI

Arina Peržu[1], Žygimantas Abramikas[2], Nojus Burdaitis[1], Giedrė Balčiūnaitė[2], Sigita Glaveckaitė[2], Darius Palionis[3], Nomeda Rima Valevičienė[3], Jolita Bernatavičienė[1]

[1] Faculty of Mathematics and Informatics, Vilnius University
[2] Clinic of Heart and Vascular Diseases, Institute of Clinical Medicine, Faculty of Medicine, Vilnius University
[3] Vilnius University Hospital Santaros Klinikos

arina.tichonovskaja@mif.stud.vu.lt

Cardiovascular diseases remain the leading cause of death worldwide, creating a pressing demand for early, objective, and AI-driven diagnostic tools. Cardiac magnetic resonance imaging (MRI), when combined with radiomics, enables the extraction of quantitative texture features that capture microstructural myocardial changes associated with fibrosis and tissue remodeling – alterations that are not always detectable otherwise. However, a critical challenge persists: can these texture-derived biomarkers autonomously cluster patients according to myocardial damage severity *without* predefined labels, and do such clusters reflect meaningful clinical differences? In this study, we propose an unsupervised machine learning framework that groups patients based on radiomic texture features extracted from multiphase cardiac MRI. A total of 46 features were computed using GLCM, GLRLM, Fourier and wavelet transforms, Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and fractal analysis, followed by feature selection using statistical testing, random forest importance, LASSO regression, and correlation analysis. Clustering was performed using the k-means algorithm combined with UMAP for nonlinear dimensionality reduction. The resulting clusters were evaluated using established clinical biomarkers, including native T1 relaxation time and extracellular volume fraction (ECV) as imaging-based indicators of myocardial fibrosis, collagen volume fraction (CVF) derived from histology as a

direct measure of fibrotic tissue change, and global longitudinal strain (GLS) as a functional marker of myocardial contractility. A clear and consistent trend was observed: clusters characterized by increased texture heterogeneity and gray-level irregularity were associated with elevated T1, ECV, and CVF values alongside reduced GLS, indicating advanced myocardial impairment, while clusters with homogeneous texture profiles corresponded to preserved myocardial structure and function. These findings demonstrate that radiomics-based clustering can autonomously uncover clinically relevant stages of myocardial damage, highlighting the potential of AI-enabled texture analysis as a powerful approach for early detection, risk profiling, and personalized management in cardiology.

Acknowledgements. The study has received approval for a biomedical research permit (No. 2025/5-1665-1119, 2025-05-12).

Constrained Branch and Bound Algorithm for Modularity Maximisation

Aurimas Petrėtis, Julius Žilinskas

Institute of Data Science and Digital Technologies
Vilnius University

aurimaspetretis@gmail.com

Modularity is one of the most popular metrics to measure the community clustering quality of a graph. Modularity maximisation is a technique which utilises this metric to perform community detection. However, modularity maximisation is an NP-hard problem and cannot be used for big graphs. We present a branch and bound algorithm to perform an exact search for specified number of communities on this problem. Various optimisation techniques like vertex ordering, high initial lower bound evaluation and real time upper bound evaluation are used to speed up the search. Algorithm performance gain compared to full search is evaluated on Zachary Karate Club graph and on LFR benchmark graphs.

Evaluating Recommendation Approaches for Employee Benefit Personalisation

Eglė Rimkienė[1,2], Marius Marmakas[1,2], Rimantė Povilavičienė[1,2], Justinas Dainauskas[1,2], Tomas Krilavičius[1,2]

[1] Vytautas Magnus University
[2] Center for Applied Research and Development

rimante.povilaviciene@vdu.lt

Employee benefit platforms are designed to engage and support the well-being of employees by giving them budgets to spend across a wide variety of categories. In practice, however, the wide range of options can quickly become overwhelming. Combined with irrelevant suggestions, the system may be perceived as ineffective. As a result, the system risks failing to increase employee satisfaction, benefit providers will be overlooked, and allocated budgets will be left unused. This raises a key research question: which recommendation approaches can best balance efficiency, personalisation, and fairness in such environments?

Two collaborative filtering approaches are investigated and compared under identical data and serving constraints to address this issue. The first is implicit-feedback matrix factorisation, with alternating least squares (ALS), which captures usage patterns effectively when historical data is abundant. The second is a two-tower neural retrieval model, which generates user embeddings from contextual and behavioural signals, directly addressing cold-start situations and enabling controllable personalisation (e.g., campaign promotion) without retraining.

The study integrates two data sources. The relational database supplies the live catalogue (benefit metadata, providers, categories, region hierarchy, purchases) and enforces availability by geography per company. Mixpanel captures user behaviour events – views, selections, purchases - treated as implicit feedback with calibrated weights, so stronger intent (e.g., checkout) counts more than casual browsing. Before ranking, the candidate list is filtered by regions to match real availability, and results are organised into categories for presentation.

Preliminary results show that ALS is most effective in environments with stable histories and identities, offering fast and efficient inference. In contrast, the two-tower model demonstrates stronger cold-start performance, flexibility across new markets, and more precise personalisation. Evaluation metrics such as Hit Rate, Recall, and NDCG illustrate how the two approaches perform differently depending on the evaluation criteria. These findings suggest that while ALS is efficient with dense histories, neural retrieval provides broader applicability by solving cold start scenarios and enabling adaptive boosting and overrides.

Comparative Analysis of ECG Data Augmentation Methods in Arrhythmia Classification

Jonas Mindaugas Rimšelis, Povilas Treigys, Zigmantas Kęstutis Juškevičius, Jolita Bernatavičienė

Institute of Data Science and Digital Technologies
Vilnius University

jonas.rimselis@mif.vu.lt

An electrocardiogram (ECG) measures electrical signals from the heart to capture various cardiovascular conditions. Distinct patterns arise in ECG during abnormal heartbeats, which facilitate the recognition of cardiovascular diseases through non-invasive ECG. Single-lead Holter devices allow uninterrupted, continuous monitoring of heart performance during everyday tasks and the identification of cardiovascular diseases. Deep learning methods are utilized for classifying heartbeats and raising awareness of deteriorating health [1]. Since abnormal heartbeats occur rarely, even in patients diagnosed with arrhythmia, data used for training models are imbalanced, leading to poor generalization and robustness [2]. Data augmentation is utilized to mitigate label balancing issues. Data augmentation techniques can be divided into traditional augmentation and generative deep learning methods. While traditional augmentation utilizes transformations of existing data to synthesize training data, generative methods utilize Generative Adversarial Networks (GANs) [3], Variational Autoencoders (VAEs) [4], and Diffusion Discrete Probabilistic Models (DDPMs) to create artificial signals [5]. As a traditional augmentation technique, SMOTE has been applied to ECG datasets, but some practitioners have raised concerns that it may implicitly distort morphological or temporal properties of ECG signals due to its interpolation mechanism. In contrast, generative methods tend to synthesize signals that mimic real-world data but tend to simplify signal morphology [6]. Furthermore, there is a lack of research on synergies between preprocessing and data augmentation techniques. In this study, a literature review is performed to capture the most prominent

and efficient data augmentation methods for ECG considering heartbeat classification in arrhythmia cases. Furthermore, synergies between preprocessing and data augmentation methods are analyzed. The review is followed by a comparative analysis of leading augmentation approaches, focusing particularly on ECG signals generated using DDPMs for the MIT-BIH Arrhythmia Database classification task. It is hypothesized that a 1Ds Convolutional Neural Network (CNN) classifier will show better performance in abnormal beat classification when trained on data augmented by DDPM than by other methods.

Acknowledgements. Funding was provided by the Research Council of Lithuania (LMTLT), agreement Nr. S-ITP-25-9.

References:

[1] O. A. Oke and N. Cavus, "A systematic review on the impact of artificial intelligence on electrocardiograms in cardiology," Int. J. Med. Inform., vol. 195, p. 105753, 2025, doi: https://doi.org/10.1016/j.ijmedinf.2024.105753.

[2] Z. Chen et al., "A novel imbalanced dataset mitigation method and ECG classification model based on combined 1D_CBAM-autoencoder and lightweight CNN model," Biomed. Signal Process. Control, vol. 87, no. PB, p. 105437, 2024, doi: 10.1016/j.bspc.2023.105437.

[3] M. Kuntalp and O. Düzyel, "A new method for GAN-based data augmentation for classes with distinct clusters," Expert Syst. Appl., vol. 235, no. March 2023, p. 121199, 2024, doi: 10.1016/j.eswa.2023.121199.

[4] Y. Xia, W. Wang, and K. Wang, "ECG signal generation based on conditional generative models," Biomed. Signal Process. Control, vol. 82, no. May 2022, p. 104587, 2023, doi: 10.1016/j.bspc.2023.104587.

[5] E. Adib, A. S. Fernandez, F. Afghah, and J. J. Prevost, "Synthetic ECG Signal Generation Using Probabilistic Diffusion Models," IEEE Access, vol. 11, no. May, pp. 75818–75828, 2023, doi: 10.1109/ACCESS.2023.3296542.

[6] L. Simone, D. Bacciu, and V. Gervasi, "ECG synthesis for cardiac arrhythmias: Integrating self-supervised learning and generative adversarial networks," Artif. Intell. Med., vol. 167, no. May 2024, p. 103162, 2025, doi: 10.1016/j.artmed.2025.103162.

Photoplethysmography-Based Obstructive Sleep Apnea Detection Using ANN: Out of Distribution Study

Mantas Rinkevičius[1], Oskar Pfeffer[2], Nando Hegemann[2], Vaidotas Marozas[1]

[1] Kaunas University of Technology
[2] Physikalisch-Technische Bundesanstalt

vaimaro@ktu.lt

Obstructive Sleep Apnea (OSA) is a prevalent sleep disorder characterized by recurrent episodes of upper airway obstruction during sleep, resulting in intermittent hypoxia and sleep fragmentation. Early detection of OSA is crucial to prevent associated cardiovascular and metabolic complications, however, the gold-standard diagnostic method, polysomnography (PSG), is expensive, complex, and limited in accessibility. As OSA induces alterations in autonomic nervous system activity, it can potentially be screened using low-cost wearable devices equipped with unobtrusive photoplethysmography (PPG) sensors, which are well-suited for monitoring pulse rate and wave-related physiological features. Objective: This study aims to compare signal feature-based artificial neural network (ANN) models for detecting apnea segments and estimating OSA burden, using a robust evaluation framework based on out-of-distribution testing across two datasets. Methods: Two whole-night long open-source -PSG datasets, MESA and OSASUD, were utilized in this study. The MESA dataset was used for training (100 subjects) and validation (30 subjects), while the OSASUD dataset served as an out-of-distribution test set (30 subjects). Five physiological features were extracted: pulse wave interval (Tp), peak-to-peak amplitude (App), maximum slope of the PPG waveform (Smax), systolic time (ST), diastolic time (DT), and peripheral oxygen saturation (SpO_2). Several ANN architectures were compared, including CNN-GRU and Inception1D models. Performance was evaluated using segment-level classification metrics (balanced accuracy and F1-score) and the mean absolute error of apnea burden estimation. To assess robustness, the signal sampling rate was

progressively reduced to simulate lower signal quality conditions. Results: Out-of-distribution testing consistently yielded lower performance across all deep learning models compared to in-distribution evaluation (balanced accuracy: 74.8% vs. 68.4%; F1-score: 71.5% vs. 58.4%). Models using either PPG features or SpO_2 alone achieved inferior OSA segment classification performance compared to those incorporating a fusion of all available features (pulse wave–derived and SpO_2). The CNN-GRU architecture consistently outperformed Inception1D,. Notably, a reduction in PPG sampling rate from 256 Hz to 20 Hz resulted in only a minor – change (~1%) in balanced accuracy, indicating robustness to lower signal -sampling. Conclusions: The findings indicate that PPG-based pulse wave analysis features provide complementary information to neural network classifiers, enhancing OSA classification performance compared to using peripheral oxygen saturation (SpO_2) alone, a conventional feature in OSA research. Model performance varied across architectures, particularly under out-of-distribution evaluation, highlighting the importance of assessing generalization. Future work should further investigate model calibration and classification uncertainty to improve the reliability of deep learning–based OSA detection.

Acknowledgements. The project (22HLT01 QUMPHY) has received funding from the European Partnership on Metrology, co-financed by the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

Graph-Theory Based Identification of Company Groups for Tax Evasion Risk

Tomas Ruzgas, Karolina Armonaite

Department of Applied Mathematics
Kaunas University of Technology

tomas.ruzgas@ktu.lt

The identification of economically related groups of companies is an important element in the assessment of tax evasion risks. This study proposes and applies a graph-theoretic approach to identify such groups using data from tax registers and company ownership networks. Unlike traditional classifications based on legal or administrative criteria, the proposed method defines company groups as structural units derived from interconnected ownership links between legal entities.

The methodology is based on the construction of a directed graph, where each node represents a company, and the edges represent ownership relationships with assigned weights corresponding to ownership percentages. Three graph-based grouping algorithms are explored and evaluated. The first identifies weakly connected components (WCCs) in the graph, which ensures maximal inclusion of all indirectly related companies. The second and third approaches apply different graph filtering techniques: one uses eigenvector centrality-based filtering, while the other removes edges below a specified ownership threshold, helping to eliminate insignificant or formal links. The empirical analysis is based on a dataset of legal entities with shareholder information from Lithuania's tax registry. The study analyses the resulting group structures formed by each method, comparing the number of groups, their sizes, and structural characteristics. The WCC-based approach yields fewer, larger groups with high node connectivity, while threshold-based methods produce more granular clusters with stronger internal ownership links. Groups are further analysed through the centrality of nodes to identify key companies within each group.

The proposed methodology enables visual representation of ownership networks and supports the detection of business groups that may

operate in a coordinated manner for tax planning or evasion purposes. Visual analysis of selected groups illustrates the effectiveness of centrality filtering in revealing economically meaningful substructures, such as hubs and branching ownership paths.

Overall, the study demonstrates the feasibility and interpretability of graph-theoretic techniques for identifying company groups. These approaches provide useful insights for auditors, analysts, and policymakers when assessing the complexity of inter-company relationships and detecting structures that may increase the risk of non-compliance or tax base erosion.

Enhancing Requirements Engineering QA with Retrieval-Augmented Generation and Knowledge Graphs

Darius Sabaliauskas, Jolanta Miliauskaitė

Institute of Data Science and Digital Technologies
Vilnius University

darius.sabaliauskas@mif.stud.vu.lt

Automated question answering (QA) systems in requirements engineering (RE) can greatly speed up the processes of specification analysis, validation, and decision-making. While traditional QA models, such as those based on BERT architectures, excel at extracting specific spans of text, they often encounter difficulties with long, fragmented, and domain-specific requirement texts [1-2]. Retrieval-Augmented Generation (RAG) improves the precision of answers by using external knowledge during the inference stage [3]. In this research, we assess five RAG strategies for RE QA like sparse lexical retrieval with BM25 [4], dense vector-based retrieval using semantic embeddings [2], a hybrid semantic reranking process utilizing cross-encoders [1], and graph-enhanced retrieval leveraging concept-based knowledge expansion [5]. Furthermore, we implement a multi-hop retrieval extension that incorporates entity-level reasoning to uncover contextual evidence spanning multiple segments [6]. These methods are designed to alleviate information overload by pinpointing potentially relevant parts of requirement documents, potentially enhancing answer precision and interpretability in subsequent reasoning activities. We assess the models using a domain-specific RE dataset with four standard QA metrics: Exact Match (EM) and F1 score for span-level accuracy, ROUGE-L for lexical similarity, and BERTScore for semantic alignment. The experimental outcomes indicate that retrieval-based approaches do not consistently surpass the non-retrieval BERT+LSTM baseline in all scenarios. Although some configurations, especially dense and multi-hop retrieval, show localized improvements in semantic alignment and contextual relevance, other RAG variants perform similarly or worse.

These results imply that while retrieval augmentation can boost performance in particular cases, its efficacy is largely contingent on the retrieval strategy and the quality of context selection, underscoring the necessity for meticulous method design in RE-focused QA systems.

References

[1] Fu, H., Yao, Y., Lin, J., & Liu, Z. (2023). S-BERT: A semantic information selecting approach for open-domain question answering. arXiv:2305.17413.

[2] Karpukhin, V., et al. (2020). Dense Passage Retrieval for Open-Domain Question Answering. ACL.

[3] Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS.

[4] Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. Foundations and Trends in IR, 3(4), 333–389.

[5] Xu, X., et al. (2024). GraphTrace: A modular retrieval framework combining knowledge graphs and LLMs for multi-hop question answering. NeurIPS.

[6] Zhao, K., et al. (2024). Advancing Large Language Models with Enhanced RAG: Evidence from Biological UAV Swarm Control. Aerospace Science and Technology.

.

Visualizing and Controlling the Optimization Process of Geometric Multidimensional Scaling

Martynas Sabaliauskas

Institute of Data Science and Digital Technologies
Vilnius University

martynas.sabaliauskas@mif.vu.lt

Multidimensional Scaling (MDS) is a special technique for dimensionality reduction and data visualization, aiming to represent high-dimensional data in a low-dimensional Euclidean space. The quality of this representation is typically improved by minimizing a stress function.

However, the stress function is well-known for being non-convex, characterized by a complex surface with numerous local minima. Traditional optimization algorithms, most notably SMACOF (Scaling by Majorizing a Complicated Function), often become trapped in these inferior solutions. Furthermore, these algorithms operate as "black boxes", providing researchers with little insight into the optimization process and no tools to identify or improve an inferior result. The consistency of the outcome heavily depends on the random initial configuration, making MDS analysis often frustratingly unclear.

This research introduces a paradigm shift by leveraging the Geometric Multidimensional Scaling (GMDS) method to transform this unclear process into one that is transparent, interpretable, and controllable. GMDS is derived from a clear geometric interpretation of the optimization problem, where the iterative step is defined as the centroid of intermediate points representing ideal locations based on individual dissimilarities. Crucially, this geometric step has been proven to correspond exactly to the anti-gradient direction of the local stress function. We utilize this analytical foundation to deconstruct the optimization process.

By employing dynamic visualization tools (Desmos, GeoGebra) and conducting a thorough asymptotic analysis of the GMDS iteration formula, we map the optimization process. The asymptotic analysis reveals that as a point approaches singularities (coinciding points) or diverges

to infinity, the subsequent GMDS steps converge to well-defined hyperspheres. These hyperspheres form a “topographical map” of the stress structure, visualizing the basins of attraction.

While this visualization provides profound insight, we found that visual inspection alone is insufficient to consistently differentiate global minima from local ones. This motivated the development of a novel, analytically driven assessment and iterative improvement cycle. This enhanced GMDS framework allows us to analytically determine whether any point in the configuration is trapped in a local minimum. If so, the algorithm systematically relocates the point to a globally superior position and re-optimizes the configuration.

Empirical studies demonstrate that this enhanced GMDS consistently outperforms SMACOF, achieving lower stress values. Additionally, the intrinsic structure of GMDS is highly suitable for parallel computing, allowing for the efficient processing of large datasets. By integrating geometric intuition, visual analysis, and analytical precision, this work not only presents a superior algorithm but also fundamentally changes how researchers can interact with and control dimensionality reduction optimization problems.

Ranks of Hankel Matrices in Estimation of Remaining Useful Life of Bearings

Paulina Sabanskytė, Mantas Landauskas

Department of Mathematical Modelling
Faculty of Mathematics and Natural Sciences
Kaunas University of Technology

paulina.sabanskyte@ktu.edu

Estimation of the remaining useful life (RUL) of bearings is usually performed using various methods, including machine learning, entropy-based, or other more or less classical methods. This study aims to use the ranks of the associated Hankel matrix. Algebraic techniques in predictive diagnostics are still less common, although some applications of Hankel matrices, for example, have received more attention recently [2, 3, 4]. The study further adds to the development of Hankel matrix-based applications in predictive diagnostics. Bearing acceleration measurements with respect to the vertical axis (obtained from the scientific experimentation platform PRONOSTIA [1] are analysed. A fixed sliding window is used, and the list of associated Hankel matrices is formed. As in our previous research, the computation of the singular values of the Hankel matrices is carried out by using the SVD decomposition. The optimal parameters for matrix size, threshold for singular value importance are determined previously. It must be noted that the pseudo-rank of the Hankel matrix is considered, since almost all real-world sequences are contaminated by noise. To reduce said noise, the fixed window of the moving average is used. In this way, pseudo-rank is employed as the property of a vibrational signal and a predictor of RUL. A new step in the study is the development of the prediction model to receive multidimensional input: a list of pseudo-ranks. Analysis of real-world diagnostic data shows that the correlation between pseudo-rank and remaining useful life could be seen when there was about 5 to 10 \% RUL left. Thus, a one-dimensional predictor is only suitable for analysing short RUL. We have constructed and trained an artificial neural network (ANN) with a single and then with multidimensional input, pseudo-

rank(s) of a Hankel matrix. A number of computational experiments are performed to compare the two approaches.

Acknowledgements. Part of the presented research is supported by the Research Council of Lithuania (LMT grant no. S-ST-25-197).

References

[1] Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N. (2012). PRONOSTIA: An experimental platform for bearings accelerated degradation tests, 1-8.

[2] Ojha, S., & Shelke, A. (2025). Estimation of the Remaining Useful Life of Axially Loaded Members Through Nested Least-Square Support Vector Regression. Journal of Earthquake Engineering, 29 (1), 243-264.

[3] Fan, W., Jiang, F., Li, Y., & Peng, Z. (2024). A Hankel matrix-based multivariate control chart with shrinkage estimator for condition monitoring of rolling bearings. IEEE Transactions on Automation Science and Engineering.

[4] Bisoyi, S., Rathi, A. K., & Mahato, S. (2024, November). Improvements in Prediction of Remaining Useful Life of Roller Bearings Using Deep Learning Models. In 2024 8th International Conference on System Reliability and Safety (ICSRS) (pp. 688-694). IEEE.

Causal Insights into Stroke Mortality Risk Reduction

Virgilijus Sakalauskas, Dalia Kriksciuniene

Kauno Kolegija

virgilijus.sakalauskas@go.kauko.lt

Stroke Stroke is a disease with severe and lasting consequences for patients' health, which necessitates the use of a broad variety of methods for health recovery. The research for identifying the most effective treatment methods is crucial for reducing the impact of stroke outcomes. Predictive machine learning models are widely applied for exploring indicators of patient survival; however, these models are limited in their explanatory power and ability to identify the most effective treatment. In this study, we aim to apply causal inference and its explanatory power to clinical and demographic data of 944 stroke patients treated at the Clinical Centre of Montenegro.

Based on the research literature, we have selected eight variables for analysis, including treatment methods, age, health status, and stroke type. Our goal was to assess the impact of treatment methods on stroke mortality rates. To reduce bias and create fair comparisons, we applied Propensity Score Matching (PSM), which enabled us to form groups of patients with similar baseline characteristics. To understand the different effects of treatments on various patients, we applied the causal forest algorithm. Causal forests extend the random forest algorithm to estimate Conditional Average Treatment Effects (CATEs) at the individual level. This not only quantifies the expected benefit of a treatment for a given patient profile but also helps to identify which subgroups benefit most—an essential step toward personalised medicine.

We evaluated the causal effect of three treatment strategies on stroke-related mortality using matched datasets and Causal Forests analysis. The treatments compared were:

- Treatment 2: Antiplatelet therapy (ANTIAGREGAC)
- Treatment 1: Anticoagulant therapy (ANTIKOAG)
- Treatment 0: Other medications or no targeted treatment

The Antiplatelet therapy (treatment 2) consistently showed the most positive and reliable impact on survival after stroke, approximately 8.6% survival chance is higher compared to Anticoagulant therapy (treatment 1), and 15% higher than No targeted treatment (treatment 0). On the other hand, Anticoagulant therapy showed no consistent benefit, with CATE at -6.5%, possibly resulting in some harm, depending on the patient's profile.

In Causal Forests, feature importance doesn't indicate which variables predict outcomes (as in standard models) - it indicates which variables help explain the differences in treatment effects between patients.

Our calculations identified the Age and Health Status as the most influential features. These findings highlight the importance of personalised treatment decisions, as patient characteristics - especially Age - strongly influence treatment response.

This research demonstrates the power of data science in making causal healthcare insights and supporting the design of targeted care pathways.

Context-Aware Modeling of Temporal Semantics for Deferred Reasoning

Juan C. Santana Santana, José J. Hernández Cabrera, José J. Hernández Gálvez, José Évora Gómez

Universidad de Las Palmas de Gran Canaria, Spain

juan.santanasantana@ulpgc.es

Temporal expressions in natural language are intrinsically uncertain: they carry ambiguity, vagueness, and cultural variability that cannot be reduced to deterministic timestamps without loss of meaning. Traditional approaches to temporal processing in NLP have treated this uncertainty as a problem to be eliminated through immediate normalization into ISO 8601 formats.

Orchilla challenges this assumption by proposing a formal language that treats uncertainty as a representable and manipulable semantic entity – a first-class component of temporal meaning rather than a computational obstacle. The framework introduces a compositional and operator-based modeling language that encodes temporal expressions as functional transformations over primitive temporal anchors. Instead of enforcing premature disambiguation, it allows expressions such as "early spring," "three days after Easter," or "before Christmas" to retain their internal semantic structure and context-dependence through a set of formal operators. This enables explicit reasoning about what is known, what remains underspecified, and how meaning evolves as contextual information accumulates.

The architecture integrates neural sequence labeling for temporal detection, symbolic decomposition into atomic components, and a deferred inference engine that resolves formal representations only when sufficient context is available. This complete architecture has been fully implemented as an experimental system to empirically validate the theoretical proposal, demonstrating its practical feasibility and the operational soundness of the deferred-inference paradigm. Beyond quantitative result Orchilla offers an explanatory contribution: it redefines uncertainty as an operational dimension of meaning. By converting ambiguity

into a structured object of computation, the framework bridges linguistic theory, knowledge representation, and temporal reasoning.

This approach suggests a broader paradigm for temporal natural language understanding, one in which interpretive indeterminacy is not a source of error but a structured space of possible meanings that can be formally modeled, deferred, and resolved. In doing so, Orchilla advances both the theoretical and computational foundations of temporal semantics in language.

Data Analysis Method for Understanding Human States Through Conversations: A Real-World Case in Diabetes Monitoring

Juan C. Santana, Jose J. Hernandez Cabrera, Jose J. Hernández Gálvez, José Evora, Juan A. Montiel

Universidad de Las Palmas de Gran Canaria, Spain

josejuan.hernandez@ulpgc.es

Monitoring daily lifestyle and emotional states in chronic disease management is essential, but conventional methods (manual diaries, structured forms) impose burdens on users and often yield sparse, low-fidelity data. In collaboration with a health-tech partner, we deployed a system for diabetic patients to submit voice notes describing their meals, physical activity, mood, and – for female participants – menstrual status. For instance, a patient might record: "Today I had toast with tomato and olive oil for breakfast, walked for 20 minutes, and I feel a bit stressed." This single utterance encapsulates multiple overlapping pieces of information: nutrition, exercise, emotional state. The challenge is to convert such noisy, unstructured, multimodal conversational data into structured, actionable records.

In this work, we present Semantic Sensoring, a data analysis framework that transforms free-form conversational input into structured human state annotations via multi-label classification. Rather than forcing a single label per message, our approach allows multiple tags simultaneously thus capturing the multifaceted nature of daily narratives.

The pipeline begins with automatic speech-to-text transcription optimized for healthcare vocabulary, followed by advanced tokenization designed for informal conversational syntax. A multi-label classification model assigns multiple semantic tags to each utterance, relying on low-resource learning strategies (few-shot, transfer learning, semantic augmentation) to confront the scarcity of labeled data. Finally, the extracted tags are structured into machine-readable formats.

Applied to a pilot cohort of diabetic users, Semantic Sensoring successfully identified combinations of nutrition events, activity, emotional

states, and menstrual indications with promising accuracy, even using limited training samples. The resulting structured dataset can feed downstream models – such as personalized health monitors, alert systems, or behavioral analytics engines. This case demonstrates how conversational inputs, even in voice form, can be leveraged as rich phenotypic sensors of human lifestyle and state. By bridging the gap between casual human communication and structured health data, Semantic Sensoring offers a new paradigm for data-driven health systems oriented around human narratives.

Multi-Agent System for Location of Park-and-Ride Hubs

Sathuta Piripun Sellapperuma, Algirdas Lančinskas

Institute of Data Science and Digital Technologies
Vilnius University

Sathuta.Sellapperuma@mif.vu.lt

Park-and-Ride (P&R) hubs are important elements of modern urban infrastructure. They help to reduce congestion and pollution by allowing travellers to leave their cars in easily accessible locations in the city and commute inside the city by bicycle, scooter, public transit, or other ecological vehicles.

Choosing locations for new P&R hubs can be a challenging task because it must consider competition between existing hubs, travel demand from many regions, customer (or travellers in our case) needs, available locations for the hubs, and other properties. Customer behaviour adds additional uncertainty when locating P&R hubs. Some travellers may always choose the most attractive hub, while others may distribute their choice among several hubs depending on their daily needs and the attractiveness of the hubs. These different behaviour models lead to different estimates of hub utility.

To ensure that new hub locations perform well under different customer behaviours, it is necessary to find solutions that remain effective independent of the customer behaviour. To address this need, it is not necessary to find the best solution for a single customer behaviour model, but rather a robust solution or set of them that are simultaneously good for different customer behaviour models. This solution balances the trade-offs between different customer behaviours and provides a practical recommendation for city planners.

We propose a Multi-Agent System (MAS) to identify robust locations for a given set of new facilities from a given set of location candidates. In this system, Customer Behaviour Agents represent different customer behaviour models and are able to propose solutions considering that customer behaviour. Each customer behaviour agent uses reinforcement

learning to explore candidate solutions, learn from its experience in evaluating the utility of the solutions and propose solutions taking into account the negotiation status. A Mediator Agent coordinates interaction between customer behaviour agents and their negotiation process. Agents follow negotiation strategies to identify optimal candidate locations in the context of different customer behaviour models and highlight the robust solution or a set of them.

Acknowledgements. This research has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-24-8.

Numerical Infinities and Infinitesimals in Optimization

Yaroslav D. Sergeyev

University of Calabria, Rende, Italy

yaro@dimes.unical.it

In this talk, a recent computational methodology is described (see [1,2]). It has been introduced with the intention to allow one to work with infinities and infinitesimals numerically in a unique computational framework. It is based on the principle "The part is less than the whole" applied to all quantities (finite, infinite, and infinitesimal) and to all sets and processes (finite and infinite). The methodology uses as a computational device the Infinity Computer (a new kind of supercomputer patented in several countries) working numerically with infinite and infinitesimal numbers that can be written in a positional system with an infinite radix. On a number of examples (numerical differentiation, divergent series, ordinary differential equations, etc.) it is shown that the new approach can be useful from both theoretical and computational points of view. The main attention is dedicated to applications in optimization (local, global, and multi-objective) (see [1-7]). The accuracy of the obtained results is continuously compared with results obtained by traditional tools used to work with mathematical objects involving infinity.

For more information see the dedicated web page https://www.theinfinitycomputer.com. The web page developed at the University of East Anglia, UK is dedicated to teaching the methodology: https://www.numericalinfinities.com

References

[1] Sergeyev Ya.D. and De Leone R., eds, Numerical Infinities and Infinitesimals in Optimization. Springer, Cham, 2022.

[2] Sergeyev Ya.D. Numerical infinities and infinitesimals: Methodology, applications, and repercussions on two Hilbert problems, EMS Surveys in Mathematical Sciences, 2017, 4(2), 219–320.

[3] De Leone R., Egidi N., Fatone L. The use of grossone in elastic net regularization and sparse support vector machines, Soft Computing, 24, 17669-17677, 2020.

[4] Astorino A., Fuduli A. Spherical separation with infinitely far center, Soft Computing, 24, 17751-17759, 2020.
[5] De Leone R., Fasano G., Roma M., Sergeyev Y.D. Iterative grossone-based computation of negative curvature directions in large-scale optimization, Journal of Optimization Theory and Applications, 186(2), 554-589, 2020.
[6] De Leone R., Fasano G., Sergeyev Ya.D. Planar methods and grossone for the Conjugate Gradient breakdown in nonlinear programming, Computational Optimization and Applications, 71(1), 73-93, 2018.
[7] Cococcioni M., Pappalardo M., Sergeyev Ya.D. Lexicographic multiobjective linear programming using grossone methodology: Theory and algorithm, Applied Mathematics and Computation, 318, 298–311, 2018.

Comparative Evaluation of Diffusion-Based ECG Denoising and Hierarchical Kalman Filtering with Online Learned Evolution Priors

Simonas Šimėnas, Dmytro Klepachevskyi, Povilas Treigys, Jurgita Markevičiūtė, Jolita Bernatavičienė

Institute of Data Science and Digital Technologies
Vilnius University

simonas.simenas@mif.stud.vu.lt

ECG signal denoising is a crucial preprocessing task for accurate downstream cardiac analysis, as ECGs are susceptible to various noise interferences (e.g., baseline wander from respiration, muscle artifacts and electrode motion) that hide meaningful morphological features. This challenge is further amplified in wearable device conditions, where the environment, motion and respiration introduce frequent and diverse noise sources compared to stationary clinical equipment. Approaches like the Hierarchical Kalman Filtering With Online Learned Evolution Priors (HKF) [1] leverage adaptive state-space modelling of heartbeat dynamics to enhance signal quality, outperforming deep-learning methods, such as fully convolutional denoising Auto-Encoders, in tasks involving the noise removal of Additive Gaussian Noise (AGN). Recent diffusion-based generative models like: DMAM-ECG [2], BeatDiff [3] and DeScoD-ECG [4], have demonstrated state-of-the-art (SOTA) performance in denoising ECG signals corrupted by aforementioned noise interfaces. While HKF authors mention the model as an alternative to current deep-learning methods, no detailed comparison under real noise conditions between HKF and SOTA diffusion-based models currently exists. This work presents a comparative analysis of diffusion-based models against HKF under identical training conditions. All models were trained under the same time constraints, tracking the performance of each model every 10 epochs to ensure fair evaluation. Standard error metrics like: Sum of Squared Distances (SSD), Percentage RMS Difference (PRD), Maximum

Absolute Distance (MAD), and Cosine Similarity (Cos Sim), were used to quantify the performance of each model. We evaluated denoising performance on the ECG signals from the MIT-BIH Arrhythmia Database [5], which were synthetically corrupted with three different noise types from the MIT-BIH Noise Stress Test Database [6]: baseline wander (BW), muscle artifact (MA), electrode motion (EM), that appear in clinical settings. Noise power is set randomly at signal-to-noise ratios (SNR) varying from 20% to 200%. We conducted the analysis of performance by stratifying by noise level to quantify the robustness of each model with the increase of noise. In addition, we enforced inter-patient splits, ensuring no subject overlap in the training and testing sets, and the use of different noise segments, to prevent memorization and to evaluate the true generalization of models. Experiments were conducted in matched-noise scenarios, where the model was trained and tested on the same noise type using an unbalanced dataset for training with the test set retaining the global class ratio, while hyperparameters for each of the methods were set based on the provided information in each of their corresponding papers. Our results highlight the superior denoising results achieved by diffusion-based models in comparison to HKF under matched training (epoch/time) constraints, while at the cost of higher inference latency.

Acknowledgements. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-25-9.

References

[1] T. Locher, G. Revach, N. Shlezinger, R. J. G. van Sloun and R. Vullings, "HKF: Hierarchical Kalman Filtering with Online Learned Evolution Priors for ECG Denoising," in ICASSP 2023 – IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023.

[2] Z.-D. Hu, Y. Hong, J.-Y. Huang, K.-H. Chen, W.-Q. Zhao, A. Grau, E. Guerra, C.-S. Wang ir F.-Q. Zhang, „DMAM-ECG: A Diffusion Model with Self-Attention Module for ECG Signal Denoising," Journal of Network Intelligence, t. 9, nr. 3, p. 1279–1292, 2024.

[3] L. Bedin, G. Cardoso, J. Duchateau, R. Dubois and E. Moulines, "Leveraging an ECG Beat Diffusion Model for Morphological Reconstruction from Indirect Signals," in Advances in Neural Information Processing Systems 37 (NeurIPS 2024) – Main Conference Track, 2024.

[4] H. Li, G. Ditzler, J. Roveda ir A. Li, „DeScoD-ECG: Deep Score-Based Diffusion Model for ECG Baseline Wander and Noise Removal," IEEE Journal of Biomedical and Health Informatics, t. 28, nr. 9, p. 5081–5091, 2024.
[5] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, p. 45–50, 2001.
[6] G. B. Moody, W. K. Muldrow and R. G. Mark, "A noise stress test for arrhythmia detectors," in *Computers in Cardiology*, 1984.

Cyber Incident Detection in Network Flow Utilising Machine Learning and Apache Kafka

Adrija Sipavičiūtė, Virgilijus Krinickij

Cybersecurity Laboratory
Institute of Computer Science
Faculty of Mathematics and Informatics
Vilnius University

adrija.sipaviciute@mif.stud.vu.lt

With cybercrime showing no signs of slowing down, passive network flow analysis is not enough, and implementing near real-time automated network traffic analysis is crucial to detect as many threats as possible. Various algorithms have been used for intrusion or anomaly detection, including Smith-Waterman, Needleman-Wunsch, variations of Dynamic Time Warping, and negative selection algorithm V-detector, with integration of analytics engine for large-scale data processing. These algorithms typically concentrate on specific attributes. Smith-Waterman and Needleman-Wunsch focus on sequence similarity, Dynamic Time Warping captures temporal alignment, and the V-detector extends the analysis to multiple attributes. Unfortunately, these algorithms, even with improvements, tend to be slow and require high computational resources to analyse large datasets, making them unsuitable for real-time network traffic analysis. To overcome these obstacles, alternatives are sought.

One of the possible options is applying supervised machine learning algorithms. After the initial training phase, machine learning models require far less computational power for inference compared to the resource-intensive traditional algorithms. Furthermore, machine learning algorithms can incorporate a wider range of features such as packet size, protocol types, and even payload. This enables richer analysis compared to sequence alignment methods, which focus mainly on sequence similarity of two sequences.

This research addresses the fundamental problem that current algorithms and methods for network flow alignment are not capable of

effectively managing the scale and speed of real-time network flows. In this work, we are working on analysing near real-time network traffic gathered in a safe laboratory environment. The environment is built on a Proxmox hypervisor hosting multiple virtual machines and a network infrastructure. The gathered datasets adhere to already existing synthetic data. Network flow data is captured in real-time and stored directly in comma-separated (CSV) format, rather than the traditional packet capture (PCAP) format. This approach ensures immediate compatibility with machine learning algorithms, reducing the need for time-consuming data format conversions. Moreover, only relevant features are recorded at the point of collection, eliminating the overhead of post-processing feature engineering.

To handle the large-scale data efficiently, big data technologies are employed – Apache Kafka providing storage and real-time streaming, and Apache Spark enabling seamless integration with machine learning models. Supervised classifiers, such as Random Forest, Naïve Bayes, and Gradient Boosting, are applied, selected for their ability to distinguish between normal and anomalous network traffic. Model performance is evaluated using standard machine learning evaluation metrics and by measuring the time analysis takes in given computational resources.

Survey of NLP Techniques for Detecting Alport Syndrome in Electronic Health Records

Gabrielė Skirmantaitė, Gražina Korvel

Institute of Data Science and Digital Technologies
Vilnius University

gabriele.skirmantaite@mif.vu.lt

Rare diseases present significant challenges for healthcare systems worldwide due to their low prevalence, diverse clinical presentations, and considerable impact on patients. Diagnosing these diseases is often delayed because symptoms commonly overlap with more prevalent conditions, complicating timely and accurate recognition. Alport syndrome, a hereditary disorder affecting renal, auditory, and ocular systems, exemplifies the critical need for early and precise diagnosis to improve clinical outcomes.

The digitization of healthcare through electronic health records (EHRs) has resulted in extensive datasets combining structured patient information with voluminous unstructured clinical narratives such as progress notes and laboratory reports. While structured data allows straightforward computational analysis, much clinically relevant information remains embedded in free-text formats, necessitating advanced analytical methods. Natural language processing (NLP) offers a powerful framework for leveraging this unstructured data by automating tasks including named entity recognition, text classification, medical outcome prediction, and information retrieval. These tasks facilitate the identification of relevant clinical phenotypes, patient stratification, and support evidence-based decision-making.

This work presents a systematic literature review surveying the application of NLP techniques for rare disease diagnosis using EHR data, with a focus on identifying challenges such as linguistic complexity, domain-specific terminology, scarcity of annotated datasets, and data privacy concerns. Owing to the limited availability of research specifically addressing NLP for Alport syndrome, this rare disease is employed as a

motivating exemplar rather than the sole subject of analysis. The review critically evaluates prevailing methodologies, highlights existing gaps, and outlines promising future directions, including data-efficient learning, cross-lingual model adaptation, and privacy-preserving collaborative frameworks. By consolidating insights from the literature, this study aims to inform the development of NLP-based tools to accelerate diagnosis and enhance personalized care for patients with rare diseases.

Beyond Forgetting: How Elastic Weight Consolidation Enhances New Language Acquisition in LLMs

Vytenis Šliogeris

Neurotechnology

vytenis@neurotechnology.com

Large Language Models (LLMs) perform impressively in high-resource languages but falter in low-resource settings, such as Lithuanian, where they score worse on translated benchmarks compared to their English counterparts. This suggests that the underlying domain knowledge is present but obscured by the model's limited proficiency in the low-resource language. However, directly attempting to enhance it through naïve continual pretraining (CP) on new data typically causes catastrophic forgetting. We address these challenges by CP Gemma-2 (2B) on a 10% subset of the Lithuanian component of the CulturaX corpus, using Elastic Weight Consolidation (EWC) to preserve the model's prior capabilities while aiming to unlock latent domain knowledge through enhanced Lithuanian fluency. EWC is a regularization framework that constrains updates to parameters deemed important for past tasks, thereby reducing their plasticity during CP. Fisher information was estimated from the Massive Multitask Language Understanding (MMLU) benchmark to identify domain-knowledge-critical parameters.

We evaluated both linguistic fluency and domain knowledge across English and Lithuanian. Fluency was measured via perplexity on the TruthfulQA and Lithuanian Q/A datasets, while domain knowledge was assessed on 7 benchmarks: ARC-Easy, Belebele, GSM8K, HellaSwag, MMLU, TruthfulQA, and Winogrande.

The results show two principal outcomes. First, EWC effectively mitigates catastrophic forgetting: the model retains its English reasoning and factual knowledge, including mathematical ability, that otherwise degrade under standard CP. Second, by improving Lithuanian fluency, EWC enhances access to the model's existing domain knowledge in Lithuanian, improving performance on 5 of 7 benchmarks and achieving

lower perplexity. Overall, EWC acts as both a protective regularizer and a cross-lingual facilitator, enabling low-resource language adaptation without sacrificing original performance. This study demonstrates a practical and reproducible pathway toward continually improving multilingual LLMs.

Evaluating Open Weight Large Language Models for Disinformation Detection as Assistants

Milita Songailaitė[1,2], Veronika Bryskina[1,2], Tomas Krilavičius[1,2]

[1] Vytautas Magnus university
[2] Centre for Applied Research and Development

milita.songailaite@vdu.lt

Disinformation remains one of the biggest problems in today's information environment. Stories originating from Russia pose a significant threat to democracy in Eastern Europe. As more and more people rely on digital media, there is a growing demand for tools that can help people critically evaluate the reliability of information that are easy to use and can be used by many people at once. Large language models (LLMs) have shown great promise in this area, but most research so far has focused on closed, proprietary systems that are not available to everyone.

This research investigates the effectiveness of open weight LLMs, including Qwen 3, Gemma 3, Deepseek, LLaMA 3, and others, as practical tools for detecting deception, even on a smaller scale suitable for everyday applications. We analyse three architectures that help assistants: basic one-shot prompting, contextually enriched few-shot prompting, and retrieval-augmented generation (RAG). Each is evaluated against media texts containing Russian disinformation narratives and neutral reporting, with a particular focus on Lithuanian-language sources, which pose special challenges due to low-resource linguistic settings. Our goal is to show how open weight systems can be useful when reading the news by looking at not only how well they detect things but also how consistent, adaptable, and clear they are.

Preliminary results suggest that smaller open-source models, while less powerful than their commercial counterparts, can still provide meaningful guidance when carefully prompted or augmented. Few-shot prompting improves contextual awareness of manipulative patterns, while RAG architectures show promise in bridging knowledge

gaps and grounding judgments in external evidence. However, the research highlights important limitations: sensitivity to prompt design, inconsistencies across different narrative framings, and notable difficulties in handling underrepresented languages. Overall, the results indicate that open-source LLMs, when used correctly, could help improve media literacy and make democracies more resilient in regions where disinformation efforts are most prevalent.

Analysing Bulk- and Microviscosity Effects from Multi-Level Computation Results of the Properties of BODIPY-Based Molecular Sensors

Stepas Toliautas, Domantas Narkevičius

Institute of Chemical Physics
Faculty of Physics
Vilnius University

stepas.toliautas@ff.vu.lt

Molecular compounds based on boron-dipyrromethene (BODIPY) have been shown to be promising candidates for microscopic, single-molecule scale sensing of environment properties, such as viscosity or temperature [1]. It is also possible to anchor the sensors to a specific type of microscopic environment, e.g., a lipid cell membrane, where the restricted molecular drift and its fluctuations result in a measurable estimate of the bulk viscosity and temperature, respectively [2]. This work builds upon an existing quantum-chemical model of microviscosity sensitivity applied to the snapshots of molecular dynamics simulations of a BODIPY sensor anchored in a bilayer lipid membrane [3]. Intensity and timescales of the dynamic changes in expected microviscosity sensitivity are evaluated by analysis of the results of the computations of the molecular properties with the aim to determine how much the bulk drift (spanning 2-12 ns) influences the fluorescence lifetime-based viscosity measurements of the sensors (0,1-5 ns). Quantum-chemical computations and subsequent data analysis were performed using resources at the supercomputer "VU HPC" of Vilnius University in the Faculty of Physics location.

References

[1] K. Maleckaitė et al., Molecules 27, 23 (2022).
[2] D. Narkevičius, master thesis, Vilnius University (2024).
[3] S. Toliautas et al., in ECAMP15, Innsbruck, Austria (2025).

Time Series-Based Analysis of Ion Channel Gating in Response to Diverse Activation Stimuli

Paulina Trybek[1], Beata Dworakowska[2], Aleksander Bies[1], Agata Wawrzkiewicz-Jałowiecka[3]

[1] Institute of Physics, Faculty of Science and Technology
University of Silesia;
[2] Institute of Biology, Department of Physics and Biophysics
Warsaw University of Life Sciences;
[3] Department of Physical Chemistry and Technology of Polymers
Silesian University of Technology

paulina.trybek@us.edu.pl

The mechanisms of ion transmission across the cell membrane have been extensively investigated over the decades through theoretical studies, computational modeling, and a wide range of experimental approaches aimed at elucidating ion channel function. Among them, large-conductance calcium- and voltage-activated potassium (BK) channels play a crucial role in regulating membrane excitability, calcium signaling, and cellular homeostasis. Due to their dual activation by voltage and intracellular calcium, BK channels exhibit highly complex gating behavior, which makes them an excellent model system for studying nonlinear dynamics in ion channel activity.

In this work, we analyze patch-clamp time series data recorded from BK channels activated by various modulators that increase the probability of channel opening, under different experimental conditions such as varying membrane potentials. We identify time series patterns that differentiate between distinct activation states and experimental setups, providing insight into how modulatory factors influence channel kinetics as revealed by sequence dynamics. Our approach integrates classical signal analysis techniques, feature extraction, frequency-domain characterization, and machine learning to provide a comprehensive framework for studying ion channel gating. To enhance signal interpretability, the nonlinear mode decomposition was employed to isolate significant com-

ponents from noisy data, enabling a more detailed signal characterization and revealing subtle temporal dependencies.

The main objective of this study is to assess whether standard or advanced time series methodologies can effectively recognize distinct channel signatures that are characteristic of specific experimental conditions. The comparative analysis of different approaches demonstrates which types of signal processing and analytical methods are most suitable for interpreting complex ion channel activity. In particular, the results indicate that although machine learning techniques can efficiently classify signals from different group, their limited interpretability constrains their utility in explaining the underlying physiological mechanisms. Therefore, complementing data-driven models with interpretable features – such as power spectrum analysis, signal complexity measures, and temporal pattern evaluation – is essential for obtaining a more comprehensive understanding of channel behavior.

This integrative strategy not only enhances the interpretative value of the results but also provides a robust framework for characterizing the multifaceted gating dynamics of ion channels, allowing for insights into various aspects of channel gating, such as sensitivity to modulatory factors and dependence on experimental conditions.

Gamification of Large Language Models

Michal Valko

Stealth AI Startup / Inria / ENS

michal.valko@inria.fr

Reinforcement learning from human feedback (RLHF) is a go-to solution for aligning large language models (LLMs) with human preferences; it passes through learning a reward model that subsequently optimizes the LLM's policy. However, an inherent limitation of current reward models is their inability to fully represent the richness of human preferences and their dependency on the sampling distribution. In the first part we turn to an alternative pipeline for the fine-tuning of LLMs using pairwise human feedback. Our approach entails the initial learning of a preference model, which is conditioned on two inputs given a prompt, followed by the pursuit of a policy that consistently generates responses preferred over those generated by any competing policy, thus defining the Nash equilibrium of this preference model. We term this approach Nash learning from human feedback (NLHF) and give a new algorithmic solution, Nash-MD, founded on the principles of mirror descent. NLHF is compelling for preference learning and policy optimization with the potential of advancing the field of aligning LLMs with human preferences. In the second part of the talk, we delve into a deeper theoretical understanding of fine-tuning approaches as RLHF with PPO and offline fine-tuning with DPO (direct preference optimization) based on the Bradley-Terry model and come up with a new class of LLM alignment algorithms with better both practical and theoretical properties. We finish with the newest work showing links between and building on top of them.

[arXiv:2312.00886, arXiv:2310.12036, arXiv:2402.05749, arXiv:2402.02992, arXiv:2310.17303, arXiv:2405.08448, arXiv:2410.17055, arXiv:2503.19612, arXiv:2505.19731]

Fine-Grained Object Detection for Precision Beekeeping

Gabriela Vdoviak, Tomyslav Sledevič

Vilnius Gediminas Technical University

gabriela.vdoviak@vilniustech.lt

Real-time and non-invasive bee monitoring at the hive entrance is crucial for the advancement of precision beekeeping practices and the maintenance of colony health. Conventional monitoring methods are characterized by their labour-intensiveness and disruptive nature, thus requiring the development of autonomous vision-based systems. However, robust detection models must effectively handle challenging real-world scenarios, including lighting variations, frequent object occlusion, motion blur, and the complexity of classifying multiple insect groups and distinguishing small objects like pollen grains.

In this study, we investigated the performance of the YOLO model on a dataset consisting of honeybee and wasp classes. The model was refined to enhance computational efficiency while preserving sensitivity to fine-grained details. A comparative analysis of two annotation approaches for pollen carrying-bees was also conducted.

The findings showed that the architecturally modified model achieved higher detection precision in comparison to the baseline model, while maintaining comparable inference speed. This demonstrates model's suitability for efficient deployment on embedded platforms. The analysis of annotation strategies provided valuable insights into labeling choices. The separate pollen class posed significant challenges due to the size of the target, but targeted architectural modifications successfully improved its detection rate by a substantial margin. In contrast, the combined pollen-bee class demonstrated a higher overall detection stability, a critical factor for maintaining tracking continuity in subsequent behavioral analysis.

This work demonstrates that fine-grained object detection, achieved through strategic architectural modifications and data labeling, is essential for robust, multi-class surveillance, indicating the potential for efficient, autonomous hive monitoring in diverse field environments.

A Methodological Review of Anomaly Detection Methods for Maritime Hybrid Warfare Analysis

Julius Venskus[1,2], Robertas Jurkus[1,2], Povilas Treigys[2], Darius Drungilas[1]

[1] Klaipeda University
[2] Vilnius University

julius.venskus@ku.lt

The seabed is host to the critical arteries of the global economy and digital communication, including a vast network of gas pipes, electrical interconnectors, and submarine fiber-optic cables. In the contemporary landscape of hybrid warfare, these vital assets have become a prime target for state-sponsored actors seeking to exert pressure and disrupt Western nations. Adversaries are increasingly leveraging commercially-flagged vessels – often part of a "shadow fleet" – to conduct covert operations, creating a significant challenge for maritime security. Recent incidents, where vessels have allegedly damaged subsea infrastructure by dropping anchor, highlight a critical vulnerability: the difficulty of distinguishing between legitimate maritime activity and preparations for sabotage.

This paper conducts a methodological review of anomaly detection methods to assess their efficacy in identifying ship behaviors potentially indicative of threats to subsea infrastructure. The central analytical challenge lies in the ambiguity of the raw data; actions such as slowing down, loitering, or anchoring can appear benign in isolation. This review therefore moves beyond simplistic kinematic analysis to synthesize and evaluate methods that achieve contextual awareness by fusing data from multiple scientific domains. The reviewed approaches include: (1) Geospatial and Oceanographic models, which provide context on seabed topography and the precise location of critical infrastructure; (2) Behavioral and Economic baselines, which define normative vessel activity (e.g., designated anchorage zones, typical fishing patterns) for specific areas; and (3) Advanced Kinematic analysis, for detecting subtle deviations in a vessel's dynamic signature.

The principal finding of this review is that standard anomaly detection algorithms are insufficient and generate an unacceptably high rate of false positives when applied in isolation. The most effective methodologies are those that fuse a vessel's dynamic behavior with its geospatial context relative to known infrastructure. The review concludes that the ability to detect threats of this nature is critically dependent on an analytical paradigm that can answer not just "what is the ship doing?" but "is what the ship doing normal for this specific location?"

This paper provides a structured overview of the methods capable of meeting this complex challenge, offering a roadmap for developing the next generation of surveillance tools required to protect our most critical undersea assets.

Acknowledgements. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-MIP-24-117.

Enhancing Financial Insight Through Machine Learning-Driven KPI Correlation Analysis

Ilona Veitaitė[1], Audrius Lopata[2], Saulius Gudas[3]

[1] Institute of Social Sciences and Applied Informatics
Vilnius University
[2] Faculty of Informatics
Kaunas University of Technology
[3] Institute of Data Science and Digital Technologies
Vilnius University

ilona.veitaite@knf.vu.lt

The research explores how financial auditing is being reshaped by the rapid expansion of digital data and the integration of artificial intelligence (AI) and machine learning (ML). Traditional methods that rely on manual checks and limited data samples are no longer sufficient to ensure reliability in today's complex financial environment. By leveraging KPI ratio correlation analysis, predictive algorithms, and process mining, auditing can move beyond retrospective reviews and provide continuous, data-driven assurance. This shift, sometimes referred to as "Auditing 2.0," enhances the detection of anomalies and strengthens financial oversight. A central theme is the role of KPI correlation analysis in uncovering subtle links between financial indicators. These relationships are valuable for identifying anomalies, assessing financial health, and predicting potential risks before they escalate.

The study emphasizes that while simple ML models such as decision trees offer transparency, they often fail to capture the complexity of financial data. More advanced techniques, including neural networks and Long Short-Term Memory (LSTM)-based systems, excel in predictive tasks but present challenges in terms of interpretability. Hybrid approaches that combine both simplicity and sophistication appear to be the most effective, offering a balance between accuracy and explainability.

The research also highlights the importance of continuous auditing, where monitoring and testing occur in real-time rather than at periodic

intervals. This proactive approach allows errors and fraudulent activities to be detected and addressed immediately, reducing the possibility of them becoming embedded in financial statements. The integration of AI-driven process mining adds further depth, enabling auditors to model and analyze business processes more effectively, compare expected workflows with actual performance, and provide near-instant assurance after critical business events.

Finally, an experimental analysis of KPI datasets revealed that most indicators are not strongly correlated, suggesting that a broad range of them can provide unique insights for ML applications. However, correlation analysis alone is insufficient for identifying the most valuable features; additional statistical and feature-selection methods are needed to refine the inputs and improve model reliability.

The findings point to a future where AI-enhanced auditing combines advanced analytics, transparency, and regulatory alignment to improve trust and effectiveness, while also raising important questions about fairness, security, and ethical use of AI in financial decision-making.

DOGO: Pet Avatar Personalization Utilizing Latent Diffusion Models

Eimantas Zaranka[1,2], Monika Katilienė[1,2], Tomas Krilavičius[1,2], Tadas Žiemys[3]

[1] Vytautas Magnus University
[2] Centre for Applied Research and Development
[3] Dogo App

eimantas.zaranka@card-ai.eu

The recent rapid advancements in generative image models have introduced a possibility for the personalisation and customisation of images while preserving subject-specific features across various generated scenarios. Traditional fine-tuning approaches often require large datasets or manual artistic interventions, therefore limiting the accessibility and needing a constant model retraining for new subject integration.

This research investigates the application of latent diffusion-based generative models for personalised dog avatar generation, which enables innovative interactions in digital environments such as social media platforms, pet owner communities, and other related services. Personalised avatars hold a significant emotional value for owners, strengthen client-provider relationships, and boost engagement with products or services. The experiments conducted in this research consisted of three fine-tuning/inference techniques: DreamBooth with Low-Rank Adaptation (LoRA), Textual Inversion, and Image Prompt (IP) Adapters, which were applied to different base models, including Stable Diffusion 1.5 (SD1.5), Stable Diffusion XL (SDXL), and Stable Diffusion 3 (SD3), along with their fine-tuned variants for specific realistic or pet focused generations. By fine-tuning DreamBooth LoRA and Textual Inversions on a limited set of user-provided images, or running an inference with a pre-trained IP adapter, the models generate avatars that accurately replicate the pet's distinctive traits while supporting stylistic modifications from text and/or image prompts. Furthermore, experiments were conducted using commercial solutions, including DALL-E 3, Grok Image Banana Nano, SeeDream, Adobe Firefly and others. After the experiments, the

pivotal role of model variation selection is noticed, which drastically impacts personalisation efficacy, as it remarkably influences output quality. Among variants, SDXL-based models such as Juggernaut and Ratatoskr demonstrated superior performance. Notably, default base Stable Diffusion models proved inadequate, often producing artefacts during inference; thus, additionally fine-tuned variants were essential for artefact mitigation. Evaluating different fine-tuning strategies, DreamBooth LoRA personalisation method showed the most promise in visual feature preservation, while generating dog avatars that closely resembled original images and allowing for a wide range of stylistic variations. Meanwhile, Textual Inversion unsatisfactorily retained specific subject features, frequently reverting to generic representations. IP Adapters were not fine-tuned, and during the experiments, only generalised open-source versions were used in order to validate the feasibility of usage. While IP Adapters showed promising results, the results were suboptimal, where newly generated images showed feature preservation capabilities but fell short in image customisation. The main limitation of DreamBooth LoRA usage in production is the requirement for per-subject adapter training, which increases computational demand and processing time, complicating deployment in scalable pipelines.

Future research will focus on developing a novel, first of its kind, IP Adapter specialized for dog feature extraction, enabling a robust zero-shot personalization pipeline that will be efficient and preserves subject identity without a need of fine-tuning the adapter. This study will advance the theoretical framework of generative models for personalisation and will support practical applications, allowing pet owners to digitally commemorate and share their companions' achievements and unique features with the world.

Preventing Memory Based Data Leaks Through Cryptographic Remote Attestation Mechanisms

Damian Olgierd Zykovič, Virgilijus Krinickij, Linas Bukauskas

Cybersecurity Laboratory
Institute of Computer Science
Faculty of Mathematics and Informatics
Vilnius University

olgierd.zykovic@mif.stud.vu.lt

Modern commerce platforms are frequent targets for cybercriminals. Attackers usually try to exfiltrate customers' private data, which in most cases is an essential resource in organisational data stores. Most breaches lead to monetisation through the resale of data on various platforms. Today, organisations rely on strong, unbreakable encryption without a key, but encryption alone cannot guarantee complete protection. In the majority of conventional system configurations, decrypted data resides in system memory in plaintext once accessed for legitimate use. If the machine is compromised by malware, the attackers can extract the data directly from memory, thus bypassing encryption entirely. Modern malware is fully capable of residing stealthily in memory and stealing decrypted data in real time.

To address this challenge our research investigates the possibility for integration of remote attestation mechanisms with Trusted Platform Modules (TPMs) to cryptographically prove to a remote verifier that the firmware is in an unaltered state and is trusted before any sensitive operation is allowed. Remote attestation enables the system to cryptographically prove that the current state of software and hardware configuration matches a previously known, safe state. To achieve this, we will explore the use of Platform Configuration Registers (PCRs), a special register within the TPM module that holds the cryptographic fingerprints of system components. With special registers, we can bind the use of a TPM-based key to a certain state of the device; the key can be sealed

to an expected set of PCR values. In the event that the machine fails remote attestation, due to malicious activity or undocumented changes, access to cryptographic keys will be denied. We expect to reduce the risk of memory-based data leaks by enforcing decryption only on verified with remote attestation systems. This approach focuses on establishing and keeping Zero Trust Architecture (ZTA) from the system boot to data access. The success metrics will rely on the rate of successful attestations under controlled versus compromised conditions, system performance overhead that was introduced by attestation checks and the mean time to detection and denial in tampered environments.

16th Conference

DATA ANALYSIS METHODS FOR SOFTWARE SYSTEMS

Compiler **Jolita Bernatavičienė**
Prepared for press and published by
Vilnius University
Institute of Data Science and Digital Technologies
4 Akademijos St., LT-08412 Vilnius

Vilnius University Press
9 Saulėtekio Av., III Building, LT-10222 Vilnius
info@leidykla.vu.lt, *www.leidykla.vu.lt*
Books online *bookshop.vu.lt*
Scholarly journals *journals.vu.lt*