
EDITOR'S NOTE

It is a pleasure to introduce this special issue of Computer Research and Modeling. In this volume, we are collecting a number of contributions in the broad area of algorithms, mathematical modeling, and software engineering. The breadth of such contributions is ample due to the current trends in Computer Science, which nowadays are more and more interdisciplinary. We believe the best way to catch up with this rapid development is to follow its multifaceted aspects and enclose their understanding within our horizon of investigation. Certainly, Artificial Intelligence in its various flavors, and mainly Large Language Models, is not something that has been left behind.

To facilitate the reading of this special issue, we summarize here the content with a brief description (articles are ranked alphabetically by first author's surname).

Regarding the paper “Computational treatment of natural language text for intent detection” (A. S. Adekotujo, T. Enikuomehin, B. Aribisala, M. Mazzara, A. F. Zubair). The process of algorithmically determining user intent from a given statement is known as intent detection and it plays a crucial role in task-oriented conversational systems. Lack of data and the efficacy of intent detection systems have been hindered by the fact that the user's intent text is typically characterized by short, general sentences and colloquial expressions. In this paper, an intent detection model is presented which accurately classifies and detects user intent. The performance of the model has been compared with existing models showing improvements to them. It is hoped that the novel model will aid in ensuring information security and social media intelligence.

The paper “Automating high-quality concept banks: leveraging LLMs and multimodal evaluation metrics” (U. Ahmad, V. Ivanov) explores the potential of large language models (LLMs) for generating high-quality concept banks and proposes a multimodal evaluation metric to assess the quality of generated concepts. The authors investigate three key research questions: the ability of LLMs to generate concept banks comparable to existing knowledge bases like ConceptNet, the sufficiency of unimodal text-based semantic similarity for evaluating concept-class label associations, and the effectiveness of multimodal information in quantifying concept generation quality compared to unimodal concept-label semantic similarity. Multimodal models outperform unimodal approaches in capturing concept-class label similarity. Furthermore, the generated concepts for the CIFAR-10 and CIFAR-100 datasets surpass those obtained from ConceptNet and the baseline comparison, demonstrating the standalone capability of LLMs in generating high-quality concepts. Being able to automatically generate and evaluate high-quality concepts will enable researchers to quickly adapt and iterate to a newer dataset with little to no effort before they can feed that into concept bottleneck models.

The paper “A study of traditional and AI-based models for second-order intermodulation product suppression” (A. Degtyarev, N. Bakholdin, A. Maslovskiy, S. Bakhurin) investigates neural network models and polynomial models based on Chebyshev polynomials for interference compensation. It is shown that the neural network model provides compensation for parasitic interference without the need for parameter tuning, unlike the polynomial model, which requires the selection of optimal delays. The L-BFGS method is applied to both architectures, achieving a compensation level comparable to the LS solution for the polynomial model, with an NMSE result of -23.59 dB and requiring fewer than 2000 iterations, confirming its high efficiency. Additionally, due to the strong generalization ability of neural network architectures, the first-order method for neural networks demonstrates faster convergence compared to the polynomial model. In 20 000 iterations, the neural network model achieves a 0.44 dB improvement in compensation level compared to the polynomial model. In contrast,

the polynomial model can only achieve high compensation levels with optimal first-order method parameter tuning, highlighting one of the key advantages of neural network models.

The current paper provides a performance and required resources comparison of the traditional memory polynomial-based scheme and NN-based model for IMD2 cancellation.

The paper “Non-linear self-interference cancellation on base of mixed Newton method” (A. Degtyarev, S. Bakhurin) investigates a potential solution to the problem of Self-Interference Cancellation (SIC) encountered in the design of In-Band Full-Duplex (IBFD) communication systems. The suppression of self-interference is implemented in the digital domain using multilayer nonlinear models adapted via the gradient descent method. The presence of local optima and saddle points in the adaptation of multilayer models prevents the use of second-order methods due to the indefinite nature of the Hessian matrix. This work proposes the use of the Mixed Newton Method (MNM), which incorporates information about the second-order mixed partial derivatives of the loss function, thereby enabling a faster convergence rate compared to traditional first-order methods. By constructing the Hessian matrix solely with mixed second-order partial derivatives, this approach mitigates the issue of “getting stuck” at saddle points when applying the Mixed Newton Method for adapting multilayer nonlinear self-interference compensators in full-duplex system design. The Hammerstein model with complex parameters has been selected to represent nonlinear self-interference. This choice is motivated by the model’s ability to accurately describe the underlying physical properties of self-interference formation. Due to the holomorphic property of the model output, the Mixed Newton Method provides a “repulsion” effect from saddle points in the loss landscape. The paper presents convergence curves for the adaptation of the Hammerstein model using both the Mixed Newton Method and conventional gradient descent-based approaches. Additionally, it provides a derivation of the proposed method along with an assessment of its computational complexity.

In the paper “Extraction of characters and events from narratives” (A. V. Kochergin, Z. Sh. Kholmatova), the authors explore two prominent techniques for event extraction from narratives: statistical parsing of syntactic trees and semantic role labeling. While these techniques were investigated by different researchers in isolation, this study directly compares the performance of the two approaches using the custom dataset. The analysis shows that statistical parsing of syntactic trees outperforms semantic role labeling in event and character extraction, especially in identifying specific details. Nevertheless, semantic role labeling demonstrates good performance in correct actor identification.

The paper “Review of algorithmic solutions for deployment of neural networks on lite devices” (S. A. Khan, S. Shulepina, D. Shulepin, R. A. Lukmanov) reviews techniques aimed at improving the efficiency and sustainability of machine learning models. These techniques have evolved over time to adapt to modern hardware and increasingly larger neural networks with growing numbers of parameters.

Upcoming challenges, such as outliers in large attention models, highlight the importance of techniques such as adaptive clipping to improve optimizations. Similarly, quantization outperforming pruning offers valuable insights into where future research efforts should be directed. Additionally, there is a growing need to develop more advanced knowledge distillation techniques for upcoming neural network architectures, calling for continuing innovation in this area.

The paper “Analysis of the physics-informed neural network approach to solving ordinary differential equations” (I. V. Konyukhov, V. M. Konyukhov, A. A. Chernitsa, A. Dyussenova) focuses on using physics-informed neural networks, particularly multi-layer perceptrons, to tackle Cauchy initial value problems with various types of right-hand-side functions. It examines how elements like neural network architecture, optimization algorithms, and software implementation impact on the performance of the learning process and solution accuracy. The efficiency of popular machine learning frameworks in Python and C# is analyzed, revealing that C# can reduce neural network training time by 20–40%. The choice of activation functions, particularly sigmoid and hyperbolic tangent, significantly influences

both the learning process and the accuracy of the solution. Key findings include that the minimum of the loss function is achieved with a specific number of neurons in a single-layer neural network, and increasing the number of neurons does not necessarily enhance training results. For single-layer networks, the Adam method proves the most effective optimization, while the LBFGS algorithm shows better results for two- and three-layer networks, often requiring shorter training times for similar accuracy. Additionally, for Cauchy problems producing oscillating solutions with diminishing amplitudes, it is shown that a neural network with a variable weight coefficient enhances accuracy near the solution interval's endpoint.

Although human skills are heavily involved in the Requirements Engineering process, methodology and formalism play a determining role in providing clarity and enabling analysis. The paper "Deriving specifications of dependable systems" (M. Mazzara) proposes a method for deriving formal specifications, which apply to dependable software systems. The proposed approach is pragmatic to its target audience: techniques must scale and be usable by nonexperts if they are to make it into an industrial setting.

The paper "Efficient diagnosis of cardiovascular disease using composite deep learning and explainable AI technique" (S.N. Qaisrani, A. Khattak, M. Zubair Asghar, R. Kuleev, G. Imbugwa) integrates Bidirectional Long Short-Term Memory (Bi-LSTM) and Convolutional Neural Networks (CNN) into a hybrid deep-learning model, employing Explainable Artificial Intelligence (XAI) techniques to achieve accurate and interpretable diagnoses of cardiovascular diseases (CVDs). A balanced dataset containing 14 attributes related to heart health was utilized from the UCI repository. Techniques for data preprocessing, such as oversampling and standardization, were employed to mitigate class imbalance and facilitate scale-free analysis. Ten configurations of the Bi-LSTM + CNN hybrid were evaluated, focusing on the optimization of parameters such as filter size, Bi-LSTM unit size, and activation functions. The top-performing model attained remarkable metrics, with an accuracy of 99.05 %, a precision of 99 %, a recall of 99 %, and an F1-score of 99 %. It surpassed conventional machine learning classifiers such as Support Vector Machines (87 %) and Random Forest (82 %), in addition to independent deep learning models like CNN and LSTM (83–87 %). Incorporating SHAP enhanced interpretability, revealing the most significant features, including thalassemia and chest pain type. This improvement fosters increased trust and usability within medical environments. Limitations included dependence on a singular dataset and insufficient investigation of pre-trained embeddings or alternative deep-learning architectures. Future research must integrate diverse datasets, investigate advanced feature selection techniques, and employ pre-trained models such as Word2Vec or GloVe. The study offers a solid and clear solution for the early and accurate prediction of cardiovascular disease, facilitating notable progress in AI-based healthcare.

The preliminary study "NLP-based automated compliance checking of data processing agreements against General Data Protection Regulation" (O. Okonicha, A. A. Sadovykh) explores the effectiveness of NLP models including SBERT, BERT, and GPT2 for GDPR compliance checks of data privacy policies of various organizations. In testing using the OPP-115 and ACL Coling datasets, the following models are assessed at both the sentence and policy level. The findings show that SBERT has the best performance at policy levels having an accuracy of 0.57, a precision of 0.78, a recall of 0.83, and an F1-score of 0.80. Whereas BERT shows an accuracy of 0.63, a precision of 0.70, a recall of 0.50, and an F1-score of 0.55. Lastly, GPT2 has fairly good results, especially in terms of precision both at the sentence level with 0.72 and at the policy level with 0.82. The study concludes that, while all models contribute to automating GDPR compliance checks, SBERT outshines on the policy level, while BERT is best at the sentence level. Despite these results, further research is needed to enhance model performance and address existing limitations.

The paper "Enhancing DevSecOps with continuous security requirements analysis and testing" (A. A. Sadovykh, V. V. Ivanov) introduces an automated approach to security in DevSecOps, integrating

security requirements analysis and mapping to Security Technology Implementation Guides directly into CI/CD pipelines. Using the ARQAN tool for semantic mapping and RQCODE to formalize requirements as enforceable code, the method ensures real-time compliance and automated testing. Evaluation in industrial automation case shows alignment with standards like IEC 62443, automating 66 % of relevant STIG guidelines.

The paper “Generating database schema from requirement specification based on natural language processing and large language model” (N. Salem, K. Al-Tarawneh, A. Hudaib, H. Salem, A. Tareef, H. Salloum, M. Mazzara) presents an innovative solution that leverages Large Language Models (LLMs) and Natural Language Processing (NLP) to automate the generation of database schemes from system requirements. The authors introduce the ERD bot, a tool that can interpret both textual specifications and entity-relationship diagrams to produce accurate SQL commands for database creation. This approach significantly streamlines the database design process, reducing the time and effort traditionally required in minimizing human error. The integration of LLMs enhances the tool's ability to understand complex requirements and generate precise database structures. The comparative analysis with existing databases demonstrates the tool's effectiveness and potential to improve productivity for systems analysts and software engineers. Overall, the paper makes a valuable contribution to the field by showcasing how advanced AI techniques can be applied to practical software engineering tasks, and it opens avenues for future enhancements like UML diagram generation.

The paper “A survey on the application of large language models in software engineering” (N. Salem, A. Hudaib, K. Al-Tarawneh, H. Salem, A. Tareef, H. Salloum, M. Mazzara) emphasizes that, while LLMs have been widely applied to software engineering tasks like testing and development, their use in software requirements remains relatively unexplored. This presents a promising area for future research. The sources note that, in total, researchers have applied LLMs to 43 different areas in software engineering, including traditional tasks such as program repair and more complex tasks like fuzzing and GUI testing. For example, researchers have explored using LLMs for generating software specifications, repairing software specifications, and classifying both functional and nonfunctional software requirements. However, more research is needed to address the challenges that are unique to the requirements phase of the software development lifecycle.

The paper “On some mirror descent methods for strongly convex programming problems with Lipschitz functional constraints” (O. S. Savchik, M. S. Alkousa, F. S. Stonyakin) describes an approach to constructing subgradient methods for strongly convex programming problems with several strongly convex (inequality-type) functional constraints. The key idea of the proposed methods is to combine two schemes: the first is with switching on productive and nonproductive steps and the second is a recently proposed modification of mirror descent for convex programming problems, allowing one to ignore some of the functional constraints on the nonproductive steps of the algorithms. Thus, the paper describes a subgradient scheme with switching between productive and nonproductive steps for strongly convex programming problems in the case where the objective function and functional constraints satisfy the Lipschitz condition. Mirror descent-type methods for strongly convex minimization problems with relative Lipschitz-continuous and relatively strongly convex objective functions and constraints are also proposed. The generalization of the described approach to the situation with the assumption of the availability at iterations of the algorithm δ -subgradients of the objective function or functional constraints is investigated, which can potentially save computational costs of the method by eliminating the requirement of the availability of the exact value of the subgradient at the current point. For all the methods considered, theoretical estimates of the quality of the solution indicate the optimality of these methods from the point of view of lower oracle estimates.

The paper “Tree species detection using hyperspectral and Lidar data: A novel self-supervised learning approach” (L. Shaheen, B. Rasheed, M. Mazzara) demonstrates the effectiveness of Self-

Supervised Learning (SSL) for tree species detection with hyperspectral and LiDAR data. The model achieves the highest accuracy of 97.5%, and outperforms other methods, such as Semi-Supervised Learning (95.56%), Supervised Learning with CNN (95.03%), Random Forest (95%), and Support Vector Machines (68%). The SSL model is able to extract robust and transferable features from unlabeled data significantly contributing to its superior performance. A data subsampling experiment demonstrates the SSL model's robustness with limited labeled data, achieving 87.58% accuracy with 10% labeled data and 92.55% accuracy with 50% labeled data. The results emphasize the potential of SSL as a scalable and robust solution for ecological monitoring and biodiversity assessment, particularly in data-scarce environments.

The paper "Subgradient methods for weakly convex problems with a sharp minimum in the case of inexact information about the function or subgradient" (F. S. Stonyakin, E. A. Lushko, I. D. Tretyak, S. S. Ablyaev) is devoted to subgradient methods for minimizing Lipschitz μ -weakly convex functions, which are not necessarily smooth. We investigate situations where inexact information about the objective function value or subgradient is used in the iterations of the subgradient method with Polyak step-size. We prove that with a specific choice of a starting point, the subgradient method with an analogue of Polyak step-size converges with a geometric rate on the class of μ -weakly convex functions with a sharp minimum in the case of additive inexactness in subgradient values. When both the value of the function and its subgradient at the current point are known with an error, the convergence to a certain neighborhood of the set of exact solutions is shown. Estimates of the solution quality produced by the subgradient method with the corresponding analogue of the Polyak step are obtained. Furthermore, the paper proposes a subgradient method with a clipped step size and provides an estimate of the quality of its solution on the class of μ -weakly convex functions with a sharp minimum. Numerical experiments are conducted to solve the problem of low-rank matrix recovery.

The paper "Fast and accurate x86 disassembly using a graph convolutional network model" (N. Strygin, N. Kudasov) introduces an improvement over the state-of-the-art in x86 disassembly using machine learning techniques, in particular, combining ideas of DeepDi and Probabilistic Disassembly. The main contribution is a model that can make fast and accurate predictions for superset disassembly as well as an extended dataset, which combines an existing ByteWeight dataset with a number of Debian packages. The model and dataset provided are open source, which is not the case for some other solutions in the area, which contributes positively to the area as a whole. Further research may focus on tuning the hyperparameters and handling large binary files.

The paper "Reinforcement learning in optimisation of financial market trading strategy parameters" (R. Vetrin, K. Koberg) investigates the application of Reinforcement Learning (RL) to high-frequency algorithmic trading in cryptocurrency derivatives and explores how RL can help traditional trading approaches for better accuracy. Specifically, the work proposes parameter estimation of statistical models for trading using RL, such that these statistical models are more accurate than in the case where the parameters in these statistical models are obtained by traditional means. The authors explore various popular RL techniques such as Deep Q-Network, Twin Delayed Deterministic Policy Gradient and Proximal Policy Optimization in combination with classic market-making algorithms, such as the Ornstein–Uhlenbeck model. Numerical results are shown that provide some confirmation that the proposed approach could lead to a more profitable trading strategy and that the computational speed is within accepted in high-frequency time frames.

In the paper "Communication-efficient solution of distributed variational inequalities using biased compression, data similarity and local updates" (R. E. Voronov, E. M. Maslennikov, A. N. Beznosikov), the authors present a novel algorithm for solving distributed variational inequalities. Variational inequalities (VIs) constitute a broad class of problems with applications in a number of fields, including game theory, economics, and machine learning. Today's practical applications of VIs

are becoming increasingly computationally demanding. It is thus necessary to employ distributed computations to solve such problems in a reasonable time. In this context, workers have to exchange data with each other, which creates a communication bottleneck. The algorithm presented in this paper employs the following techniques to reduce the cost and the number of communications: the similarity of local operators, the compression of messages and the use of local steps on devices. Furthermore, the authors derive the theoretical convergence of the algorithm and perform some experiments to demonstrate the applicability of the method.

The paper “Regularization and acceleration of Gauss–Newton method” (N. E. Yudin, A. V. Gasnikov) proposes a family of Gauss-Newton methods for solving optimization problems and systems of nonlinear equations based on the ideas of using the upper estimate of the norm of the residual of the system of nonlinear equations and quadratic regularization. The paper presents a development of the “Three Squares Method” scheme with the addition of a momentum term to the update rule of the sought parameters in the problem to be solved. The resulting scheme has several remarkable properties. First, the paper algorithmically describes a whole parametric family of methods that minimize functionals of a special kind: compositions of the residual of a nonlinear equation and a unimodal functional. Such a functional, entirely consistent with the “gray box” paradigm in the problem description, combines a large number of solvable problems related to applications in machine learning, with the regression problems. Secondly, the obtained family of methods is described as a generalization of several forms of the Levenberg–Marquardt algorithm, allowing implementation in non-Euclidean spaces as well. The algorithm describing the parametric family of Gauss-Newton methods uses an iterative procedure that performs an inexact parameterized proximal mapping and shift using a momentum term. The paper contains a detailed analysis of the efficiency of the proposed family of Gauss–Newton methods; the derived estimates take into account the number of external iterations of the algorithm for solving the main problem, the accuracy and computational complexity of the local model representation and oracle computation. Sublinear and linear convergence conditions based on the Polak–Lojasiewicz inequality are derived for the family of methods. In both observed convergence regimes, the Lipschitz property of the residual of the nonlinear system of equations is locally assumed. In addition to the theoretical analysis of the scheme, the paper studies the issues of its practical implementation. In particular, in the experiments conducted for the suboptimal step, the schemes of effective calculation of the approximation of the best step are given, which makes it possible to improve the convergence of the method in practice in comparison with the original “Three Square Method”. The proposed scheme combines several existing modifications of the Gauss–Newton method which are frequently used in practice. In addition, the paper proposes a monotone momentum modification of the family of developed methods, which does not slow down the search for a solution in the worst case and demonstrates in practice an improvement in the convergence of the method.

In such an intricate scenario of contributions, the reader may initially find some difficulties in the navigation of content. We encourage each one to scratch below the surface to realize the deep network or relationships that all the articles can offer as a whole. We hope that this introduction can represent an effective compass to identify different paths of interest and different perspectives to look at the overall effort.

Cooperating with scientists from a broad spectrum of disciplines has been an enriching experience for us, and the final result represents a considerable step forward for the development of computing as a science. With every step of scientific development, possibly more questions remain open than those that have been closed.

*Manuel Mazzara
Innopolis University*