Proceeding
of the

2026 International Interdisciplinary Conference on Artificial Intelligence: Engineering, Health, Finance and Humanities (IICAI)

Shahid Beheshti University, Tehran, Iran

31 MAY – 1 JUNE 2026

IEEE IRAN SECTION

IEEE

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

دانشگاه شهید بهشتی

Copyright Notice: © 2026 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Disclaimer: The content and views expressed in each essay are the sole responsibility of the respective authors. The editor/compiler of this booklet does not endorse or assume any responsibility for the views expressed by the authors. This booklet contains the accepted papers presented at the " 2026 International Interdisciplinary Conference on Artificial Intelligence: Engineering, Health, Finance and Humanities (IICAI)". Responsibility for the content of the papers lies directly with the authors and does not necessarily reflect the views of the conference's scientific committee.

Title of the Book: Conference Materials and Papers
Publication Date: May 2026

MESSAGE FROM THE CHAIR

Artificial intelligence is no longer merely a subject of academic debate or technological curiosity. As AI continues to expand, it is gradually becoming part of how countries plan for the future, how organizations make decisions, how industries improve their services, and how societies think about what lies ahead. For this reason, holding a conference in this field is not simply a scientific event; it is an opportunity to bring together different perspectives, experiences, and expectations regarding the future role of artificial intelligence in various aspects of life. One of the main goals of this conference is to strengthen the connection between knowledge and application. Universities can generate new ideas, train specialists, and develop the scientific foundations of this field. Policymakers can guide the responsible and effective use of these technologies. Industry can transform research achievements into practical tools, services, and solutions. When these three sides work together, artificial intelligence can move beyond isolated studies and create real impact.

The First International Conference on Artificial Intelligence, held last year, was an important step in this direction. It helped create a space in which universities, governmental institutions, and industry could come closer together and discuss the development and application of artificial intelligence in a more focused way. This year, at the Second International Conference on Artificial Intelligence, we hope to continue this path with greater maturity, broader participation, and clearer attention to practical needs.

This year's conference also pays special attention to the role of artificial intelligence in wealth creation, wealth management, and financial markets. This is an important and timely topic, as many economic and financial decisions today depend on data, prediction, risk analysis, and intelligent decision-making systems. We believe that serious academic and professional discussions in this area can open new paths for research, innovation, and practical cooperation.

At the same time, the conference continues to welcome papers across a wide range of topics in artificial intelligence, from machine learning, natural language processing, computer vision, robotics, and big data to ethics, governance, health, engineering, education, and the humanities. This diversity is essential, because artificial intelligence itself is not limited to a single discipline or sector.

I sincerely hope that this conference will provide a useful and constructive environment for researchers, students, scholars, policymakers, entrepreneurs, and industry representatives. I also hope that the discussions and collaborations formed during this event will continue beyond the days of the conference and contribute to the responsible development and effective use of artificial intelligence.

I warmly welcome all respected participants to the Second International Conference on Artificial Intelligence and wish them a valuable and successful conference.

IICAI

Second International Interdisciplinary Conference on Artificial Intelligence

Executive Committee

- **Executive Committee Chair:** Dr. Monireh Abdoos
- **Publications & Indexing:** Dr. Mahboobeh Riahi
- **Industry Connection:** Dr. Mojtaba Vahidi-Asl
- **International Connection:** Dr. Armin Salimi Badr
- **Public Relations:** Dr. Saeed Mohammadi
- **Pre-events:** Dr. Sadegh Aliakbari
- **Workshops:** Dr. Ramak Ghavamizadeh
- **Executive Affairs:** Mr. Yousef Borhani
- **Secretariat Manager:** Mrs. Somayyeh Andariani
- **Secretariat:** Mrs. Zohreh Arab

Organizing Committee

- Dr. Mohsen Ebrahimi Moghaddam (Chairman of the conference)
- Dr. Seyed Mahmoud Reza Aghamiri (Chairman of the conference)
- Dr. Mojtaba Vahidi Asl
- Dr. Mehrnoush Shamsfard
- Dr. Armin Salimi Badr
- Dr. Monireh Abdoos
- Dr. Freidoon Shams Alii
- Dr. Ali Jahanian

Scientific Committee

- **Scientific Committee Chair:** Dr. Mehrnoosh Shamsfard
- **Executive Assistant of the Scientific Committee:** Mostafa Karimimanesh
- Dr. Ahmad Akbari Azirani
- Dr. Ahmad Ali Abin
- Dr. Sadegh Aliakbari
- Dr. Bagher Ansari
- Dr. Mehrdad Ashtiani
- Dr. Mohammad Hadi Bokaei
- Dr. Maryam Dollou
- Dr. Mohsen Ebrahimi Moghaddam
- Dr. Abdolmoein Esghaei
- Dr. Heshaam Faili
- Dr. Mohammad Mehdi Homayounpour
- Dr. Mohsen Kahani
- Dr. Reza Khosrowabadi
- Dr. Hamed Malek
- Dr. Mohammad Hossein Moaiyeri

Scientific Committee

- Dr. Hamid Reza Mahdiani
- Dr. Behrooz Minaei
- Dr. Hamid Moghaddasi
- Dr. Pegah Nejat
- Dr. Ali Noorzad
- Dr. Maziar Palhang
- Dr. Dara Rahmati
- Dr. Mahboobeh Riahi
- Dr. Reza Safabakhsh
- Dr. Armin Salimi-Badr
- Dr. Neda Shahidi
- Dr. Yasser Shekofteh
- Dr. Mohsen Soryani
- Dr. Hadi Tabatabaee
- Dr. Alireza Talebpour
- Dr. Mojtaba Vahidi Asl
- Dr. Mohammadhadi Zahedi

Student Committee

- Amirabbas Ahadi
- Amirreza Yazdanpanah
- Aras Valizadeh
- Aria Shakoo
- Farzaneh Fakhri
- Fatemeh Zarrin Kolah
- Mahdi Khorasani
- Mohammadsina Rahimi
- Mohammadali Hashemi Pour
- Mohammadamin Rashidi
- Mohsen Norouzi
- Narges Gholam Nezhad

Student Committee

- Nasime Rahimi
- Osveh Hasanpour
- Parsa Hamzeii
- Parsa Harooni
- Razieh Delvari
- Sadra Yousef Abadi
- Shima Emami
- Zahra Azizi
- Zhina Fathi

Thanks to all the other students who helped us organize this conference.

Reviewers

- Dr.Fatemeh Abbasi
- Dr.Maqsoud Abbaspour
- Dr.Monireh Abdous
- Dr.Hatef Abdous
- Dr.Atena Abdi
- Dr.Hamed Abdi
- Dr.Ahmad Ali Abin
- Dr.Hossein Attarzadeh
- Dr.Mohammad Akhavan
- Dr.Maryam Akhavan
- Dr.Maryam Azimzadeh
- Dr.Masoumeh Azimzadeh
- Dr.Reza Azmi
- Dr.Narges Asadi
- Dr.Ali Akbar Afzalian
- Dr.Faezeh Aghazadeh Par
- Dr.Zeinab Aghahadi
- Dr.Atefeh Aghaei

Reviewers

- Dr.Ahmad Akbari
- Dr.Akbar Alem Tabriz
- Dr.Sadegh Ali Akbari
- Dr.Mohammad Alishahi
- Dr.Tayebeh Amirkhani
- Dr.Reza Ahsan
- Dr.Somayyeh AhmadKhan
- Dr.Parvin Ahmadi
- Dr.Azadeh Ahmadi
- Dr.Fatemeh Ahouz
- Dr.Mohsen Ebrahimi Moghadam
- Dr.Eslahchi
- Dr.Abdolmoein Esghaee
- Dr.Salman Eivazi Nejad
- Dr.Parsa Bagherzadeh
- Dr.Alireza Baghershemirani
- Dr.Golnaz Baghdadi
- Dr.Masoumeh Behrouz

Reviewers

- Dr.Hamid Beigi
- Dr.Mohammad Javad Parseh
- Dr.Nahid Paknezhad
- Dr.Aida Pak Niat
- Dr.Mazyar Palhang
- Dr.Zahra Pourbahman
- Dr.Boshra Pishgoo
- Dr.Shima Tabibian
- Dr.Alireza Talebpour
- Dr.Farah Torkamani Azar
- Dr. Zahra Saberi
- Dr. Alireza Sadeghirad
- Dr. Amirhossein Sadr
- Dr. Mohsen Saryani
- Dr. Armin Salimi Bad
- Dr. Reza Safabakhsh
- Dr. Hamid Safari
- Dr. Pegah Safari

Reviewers

- Dr. Hossein Sameti
- Dr. Arezoo Soltan Mohammadi
- Dr.Amir Jalali
- Dr.Hamed Jalali
- Dr.Shahram Jamali
- Dr.Ali Jahanian
- Dr.Zeinab Hajimohammadi
- Dr.Vesal Hekmi
- Dr.Ali Heirani Tabas
- Dr.Seyed Mohammad Hossein Hashemi
- Dr.Mohammad Hosseini
- Dr.Mohammadmehdi Homayounpour
- Dr.Mahdieh Khaliqfard
- Dr.Saeidreza Kheradpisheh
- Dr.Maryam Khosravi
- Dr.Reza Khosroabadi
- Dr.Kevin Khosroabadi
- Dr.Fatemeh Khoshnevisan

Reviewers

- Dr.Marzieh Davoudabadi
- Dr.Raheleh Davoudi
- Dr.Behrouz Dori
- Dr.Mehdi Delrabaei
- Dr.Hamidreza Zarandi
- Dr.Mohammadamin Zargarezadeh
- Dr.Mohammadhadi Zahedi
- Dr.Mohammadhadi Zahedi
- Dr.Hossein Zeinali
- Dr.Fatemeh Ziaei Tabar
- Dr.Abbas Rad
- Dr.Shiva Rajabi
- Dr.Rouzbehani
- Dr.Mahboubeh Riahi
- Dr. Alireza Sharifi
- Dr. Hadi Shakibian
- Dr. Ahmad Shalbaf
- Dr. Mehrnoosh Shamsfard
- Dr. Asadollah Shahbahrami

Reviewers

- Dr. Yaser Shekofteh
- Dr. Mojtaba Shourian
- Dr. Seyed Mohammad Shoari Shoar
- Dr.Mojgan Farhoudi
- Dr.Hesham Fili
- Dr. Keramatollah Ghaffari
- Dr. Aghil Ghaheri
- Dr.Jahanshah Kaboudian
- Dr.Ehsanollah Kabir
- Dr.Reza Kalantari
- Dr.Mostafa Karimi Manesh
- Dr.Mohsen Kahani
- Dr.Razieh Ganji
- Dr.Amir Laki Zadeh
- Dr.Hani Mahboubi
- Dr.Ahmad Mahmoudi
- Dr.Saeid Masoudnia
- Dr.Naser Mazini

Reviewers

- Dr.Hamed Malek
- Dr.Rouhollah Mansouri Sepehr
- Dr.Rahel Mahdian Taraqqi
- Dr.Hamidreza Mahdiani
- Dr.Mohammadhossein Moeiri
- Dr.Mohammadshahram Moein
- Dr.Ali Motaghi
- Dr.Mostafa Mostafaei
- Dr.Maedeh Mosharraf
- Dr.Hamid Moghaddasi
- Dr.Majid Moghadam
- Dr.Saeedeh Momtazi
- Dr.Nooh Monavari
- Dr.Mohammadreza Mohammadi
- Dr.Hamid Mohammadzadeh
- Dr.Mostafa Mirzaei
- Dr.Hamid Mirmoeini
- Dr.Behrouz Minaei
- Dr.Babak Nasersharif

Reviewers

- Dr.Ali Nasrabadi
- Dr.Mehdi Nassaji
- Dr.Ali Nazari
- Dr.Ehsan Nazerfard
- Dr.Pegah Nejat
- Dr.Mahmoud Neshati
- Dr.Samira Nouferesti
- Dr.Ahmad Nikabadi
- Dr.Mahya Nikouei
- Dr.Alireza Vafaei Nejad
- Dr.Alireza Valian
- Dr.Abolfazl Valizadeh
- Dr.Mojtaba Vahidi Asl
- Dr.Hadi Veisi
- Dr.Shahrbanoo Yadollahi
- Dr.Yadollah Yaghoubzadeh
- Dr.Kaabeh Yaghoubi

Scientific Supporters

دانشگاه شهید بهشتی

دانشکده مهندسی و علوم کامپیوتر

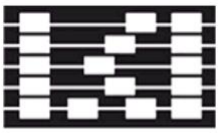
دانشکده مهندسی و علوم کامپیوتر، دانشگاه شهید بهشتی

انجمن هوش مصنوعی ایران

انجمن انفورماتیک ایران

انجمن فناوری اطلاعات و ارتباطات ایران

انجمن سیستم های هوشمند ایران



انجمن کامپیوتر ایران

انجمن ماشین بینایی و پردازش تصویر ایران

شورای عالی فضای مجازی

انجمن علوم مهندسی برق ایران

پژوهشگاه دانش‌های بنیادی (IPM)

انجمن روانشناسی ایران

انجمن صنفی کارفرمایی شرکت های فناور هوش مصنوعی و اقتصاد دیجیتال

Industrial Sponsors

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

2nd International Interdisciplinary Conference on Artificial Intelligence

2026

Engineering, Health, Finance and Humanities

31 MAY – 1 JUNE

AI4ALL

Shahid Beheshti University International Conference Center

Tracks:

- Natural Language Processing
- Image Processing and Machine Vision
- Audio and Speech Processing
- Knowledge Representation and Reasoning
- Robotics
- Theory and Applications of Machine Learning
- Cognitive and Computational Neuroscience
- Artificial Intelligence Hardware Platforms
- Data Mining and Big Data
- Bioinformatics
- AI in Games
- AI in the Humanities and Social Sciences
- AI in Health
- AI in Engineering
- AI in Governance

Special Track: Wealth Creation and Management Powered by AI

Conference Website: www.iai.sbu.ac.ir/2026

IEEE IRAN SECTION

Submission Deadline : November 28 ,2025

Notificaion of Acceptance : January 31 ,2026

Oral Papers

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

SESSION SCHEDULE
2ND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE

SUNDAY, 31 MAY

TIME	TOPIC	CHAIRPERSON	PAPER IDS
8-9	OPENING CERMONY		
9-10:40 ROOM 1	WEALTH CREATION AND MANAGEMENT USING ARTIFICIAL INTELLIGENCE, NATURAL LANGUAGE PROCESSING	DR. MINAEI , DR. SHAMSFARD	iai-1071 iai-1196 iai-1204 iai-1100 iai-1123
9-10:40 ROOM 2	DATA MINING AND BIG DATA , BIOINFORMATICS	DR. TALEBPOUR , DR. ALIAKBARY	iai-1157 iai-1184 iai-1169 iai-1142
11-12:45 ROOM 1	NATURAL LANGUAGE PROCESSING	DR. FAILI , DR. BOKAEI	iai-1156 iai-1087 iai-1232 iai-1207 iai-1126
11-12:45 ROOM 2	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN GOVERNANCE, APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN OTHER FIELDS OF ENGINEERING	DR. SALARIFARD , DR. ABBASPOUR	iai-1055 iai-1153 iai-1231 iai-1125 iai-1064
13:15-15:15 ROOM 1	IMAGE PROCESSING AND MACHINE VISION	DR. EBRAHIMI MOGHADAM , DR. MOIN	iai-1219 iai-1088 iai-1039 iai-1050 iai-1209 iai-1019
13:15-15 ROOM 2	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN OTHER FIELDS OF ENGINEERING	DR. MOAYRI, DR. MAHMOUDI	iai-1101 iai-1104 iai-1114 iai-1057 iai-1214
15:15-17 ROOM 1	THEORETICAL TOPICS IN MACHINE LEARNING, APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN THE HUMANITIES AND SOCIAL SCIENCES	DR. ABDOOS, DR. JALALY	iai-1092 iai-1218 iai-1097
15:15-17 ROOM 2	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN OTHER FIELDS OF ENGINEERING	DR. NOORZAD, DR. NICKABADI	iai-1129 iai-1065 iai-1173 iai-1201 iai-1185

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

SESSION SCHEDULE

2ND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE

MONDAY, 1 JUNE

TIME	TOPIC	CHAIRPERSON	PAPER IDS
9-10:40 ROOM 1	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN HEALTH AND HYGIENE	DR. ABIN, DR. MOGHADDASI	iai-1076 iai-1109 iai-1132 iai-1187 iai-1053
9-10:40 ROOM 2	ROBOTICS , ARTIFICIAL INTELLIGENCE IN GAMES	DR. SALIMI BADR, DR. PALHANG	iai-1192 iai-1103 iai-1002 iai-1144 iai-1188
11-12:45 ROOM 1	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN HEALTH AND HYGIENE, KNOWLEDGE ENGINEERING AND	DR. ESGHAEI, DR. KHOSROABADI	iai-1213 iai-1048 iai-1030 iai-1152 iai-1200
11-12:45 ROOM 2	AUDIO AND SIGNAL PROCESSING, AI HARDWARE PLATFORMS	DR. HOMAYOUNPOUR, DR. RAHMATI	iai-1228 iai-1078 iai-1143 iai-1154 iai-1111 iai-1227

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1002

Integrated Reactive-Proactive Multi-Agent Coordination: A SACR Communication Protocol with Strategic Planning Framework

Mostafa Mehrabi

Faculty of Computer Science and Computer Engineering

Shahid Beheshti University

Tehran, Iran

mmehrabi.mostafa@gmail.com

Monireh Abdoos*

Faculty of Computer Science and Computer Engineering

Shahid Beheshti University

Tehran, Iran

m_abdoos@sbu.ac.ir

***Abstract*—**Coordinating autonomous agents under strict constraints without extensive training is challenging. Game-theoretic approaches coordinate effectively but lack planning capabilities, while reinforcement learning requires prohibitive training data. We integrate cooperative game theory with LLM-based planning through SACR (Signal, Acknowledge, Communicate, Resolve), a lightweight protocol for priority-based conflict resolution. The framework uses forward prediction to handle saturation constraints and external interference across planning and execution phases. Dynamic coalition formation enables agents to balance formation maintenance, target acquisition, and threat avoidance simultaneously. Evaluating nine agents across formation, interference, and obstacle scenarios shows SACR eliminates all proximity violations versus 4-82 violations in baselines. With LLM planning, Claude Sonnet 4 with extended thinking achieves 89.1-100% plan adherence, outperforming Gemini 2.5 Pro's 80.9-86.2%. The framework provides interpretable, scalable coordination for constrained multi-agent systems. Implementation code is available at https://github.com/mostafamehrabii/multi-agent-sacr-simulation.

***Keywords*—**Multi-agent systems, Signal Acknowledge Communicate Resolve (SACR), cooperative game theory, large language models, planning, communication protocols.

Article Code: 1019

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Beyond Sequences: A Benchmark for Atomic Hand-Object Interaction Using a Static RNN Encoder

Yousef Azizi Movahed
Dept. of Computer Science, School of Mathematics, Statistics, and Computer Science, College of Science
University of Tehran
Tehran, Iran
azizi.movahed@ut.ac.ir

*Fatemeh Ziaeetabar**
Dept. of Computer Science, School of Mathematics, Statistics, and Computer Science, College of Science
University of Tehran
Tehran, Iran
fziaeetabar@ut.ac.ir

***Abstract*—**Reliably predicting human intent in hand-object interactions is an open challenge for computer vision. Our research concentrates on a fundamental sub-problem: the fine-grained classification of atomic interaction states, namely 'approaching','grabbing', and 'holding'. To this end, we introduce a structured data engineering process that converts raw videos from the MANIAC dataset into 27,476 statistical-kinematic feature vectors. Each vector encapsulates relational and dynamic properties from a short temporal window of motion. Our initial hypothesis posited that sequential modeling would be critical, leading us to compare static classifiers (MLPs) against temporal models (RNNs). Counter-intuitively, the key discovery occurred when we set the sequence length of a Bidirectional RNN to one (seq length=1). This modification converted the network's function, compelling it to act as a high-capacity static feature encoder. This architectural change directly led to a significant accuracy improvement, culminating in a final score of 97.60% on the benchmark split, while maintaining superior robustness on unseen subjects compared to both temporal models and strong classic baselines (SVM, Random Forest). Of particular note, our optimized model successfully overcame the most challenging transitional class, 'grabbing', by achieving a balanced F1-score of 0.90. These findings provide a new benchmark for low-level hand-object interaction recognition using structured, interpretable features and lightweight architectures.

***Keywords*—**Hand-Object Interaction, Action Recognition, Feature Engineering, Recurrent Neural Network, Benchmark, Action Anticipation

Article Code: 1030

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Simultaneous representation of decision variable and saccade direction in the parietal cortex

*Zahra Naghdabadi**
Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
znaghdabd@gmail.com

Amirreza Bahramani
Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
bahramani77@gmail.com

Abstract— Neurons in the lateral intraparietal area (LIP) maintain elevated activity during saccadic eye movements and also fire when a salient cue is in their receptive field. Most of the time, humans and animals plan to make saccades to salient targets (based on physical appearance or their goal). It is unknown whether LIP's activity is dominated by the current focus of attention or preparation for the next-saccade (attention vs. intention debate). In addition to representing the next-saccade, area LIP represents the decision variable during decision-making as some LIP neurons' responses are best explained by the accumulated evidence in favor of the chosen target. Since in some decision-making tasks, choice is reported by a saccade, it is unknown whether pre-choice activity of LIP neurons is neural representation of the decision variable or the saccade (or both). If LIP represents both signals, then it is not known whether the decision variable and the next-saccade are represented within the same neuronal population or in separate populations. We addressed these questions using population analysis on neural data recorded from macaques during a perceptual decision-making task. We found that the same population of LIP neurons represents the decision variable and the next-saccade. Simultaneous encoding of these information occurs in linearly separable subspaces of the latent space, making it easily distinguishable and communicated to distinct downstream areas. We hypothesize that by doing so, LIP can have an active role in the decision-formation process while simultaneously conveying information about the next saccadic eye movement.

Keywords— LIP, saccade, decision-making, population analysis, dimension reduction

Article Code: 1039

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Selective Cross-Feature Fusion Network (SCFF-Net) for Vehicle Re-Identification in Urban Surveillance

*Morteza Makhloughi**
Department of Computer Engineering
Zanjan University
Zanjan, Iran
m.makhloghi@znu.ac.ir

Ali Amiri
Department of Computer Engineering
Zanjan University
Zanjan, Iran
a_amiri@znu.ac.ir

Aydin Kaya
Department of Computer Engineering
Hacettepe University
Ankara, Turkey
aydinkaya@cs.hacettepe.edu.tr

Abstract— Vehicle re-identification (Re-ID) plays an increasingly important role in intelligent transportation systems (ITS) and urban surveillance by enabling vehicle tracking across multiple cameras with non-overlapping views. The task is challenging due to dramatic appearance variations caused by illumination, occlusion, background clutter, and camera viewpoints.Most deep-learning-based methods focus either on local part-level features or global semantic embeddings, leading to redundancy and loss of fine detail. To address these issues, this paper presents the Selective Cross-Feature Fusion Network (SCFF-Net) — a lightweight hybrid CNN–Transformer architecture that adaptively integrates convolutional texture cues and transformer-based global semantics. The key innovation is the Selective Cross-Feature Fusion (SCFF) module, which performs adaptive gating to filter redundant activations and emphasize complementary feature responses. Unlike previous hybrid methods with static fusion, SCFF-Net learns dynamic, input-dependent fusion weights that enhance discriminative representations. Extensive experiments on two public benchmarks, VeRi-776 and VehicleID, demonstrate that SCFF-Net achieves superior performance (85.2% mAP / 97.9% Rank-1 on VeRi-776 and 79.4% mAP on VehicleID) while maintaining real-time inference efficiency (≈ 8 ms/image). In addition, cross-domain experiments confirm its generalization capacity to unseen environments. With a balanced trade-off between accuracy and computational cost, SCFF-Net offers a practical and scalable solution for city-scale vehicle monitoring.

Keywords— Vehicle Re-Identification, Hybrid CNN– Transformer Network, Selective Feature Fusion, Intelligent Transportation Systems (ITS), Urban Surveillance.

Article Code: 1048

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

DC-GATT: Dynamic Connectivity Graph Attention Transformer for Multi-Task EEG Classification

Mohammad Ahadzadeh
Department of Biomedical Engineering
Amirkabir University of Technology
(Tehran Polytechnic)
Tehran, Iran
Mmd.Ahd@aut.ac.ir

Pouya Taghipour Langrodi†
Department of Biomedical Engineering
Amirkabir University of Technology
(Tehran Polytechnic)
Tehran, Iran
Pouya_Taghipour@aut.ac.ir

Arian Afshar†
Department of Biomedical Engineering
Amirkabir University of Technology
(Tehran Polytechnic)
Tehran, Iran
Arafsh1382@aut.ac.ir

*Golnaz Baghdadi**
Department of Biomedical Engineering
Amirkabir University of Technology
(Tehran Polytechnic)
Tehran, Iran

Abstract—Decoding cognitive states from electroencephalography (EEG) across multiple tasks and sessions is essential for developing robust Brain-Computer Interfaces (BCI). However, reaching great precision is still difficult because of the non-stationarity of neural functional interactions. Although several prior efforts have developed effective static graph representations or task-specific models, they often neglect the fine-scale temporal evolution of brain networks that discriminates more complex cognitive states. This paper presents the Dynamic Connectivity Graph Attention Transformer (DC-GATT), an end-to-end framework designed to bridge this gap by jointly modeling session-level dynamic functional connectivity with transformer-based temporal attention. Using the COG-BCI dataset, we model EEG epochs as sequences of overlapping graph snapshots to capture the temporal progression of network reconfiguration. The architecture employs a hierarchical dual-attention pooling scheme to effectively extract discriminative spatiotemporal features. DC-GATT yields an average test accuracy of 85.3% demonstrating a significant improvement over static graph baseline (70.9%). Our findings confirm that modeling the temporal dynamics of graph snapshots carries critical discriminative information, offering a scalable and robust approach for multi-task cognitive state decoding.

Keywords—EEG, Dynamic functional connectivity, Graph attention network, Transformer, Multi-task classification, Weighted phase-lag index, Brain–computer interface

Article Code: 1050

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

AdaLayer-CAM: Adaptive Multi-Layer Method for Visualization and Explanation of CNNs

*MohammadJavad GhaderiPoor**
Department of Statistics, Mathematics and Computer Science
Allameh Tabataba'i University
Tehran, Iran
mohammadjavadghaderipoor@gmail.com
(https://orcid.org/0009-0001-6340-7233)

Latifeh PourMohammadBagher
Department of Statistics, Mathematics and Computer Science
Allameh Tabataba'i University
Tehran, Iran
l_pmb@atu.ac.ir

Zahra Rahimi
Department of Statistics, Mathematics and Computer Science
Allameh Tabataba'i University
Tehran, Iran
za.rah@atu.ac.ir

Abstract—Class Activation Mapping (CAM) is one of the standard tools for visualizing the regions of focus in neural networks, showing where the model pays attention during prediction. However, many widely used approaches, includingGrad-CAM, Grad-CAM++, and LayerCAM rely only on a subset of high-level convolutional layers—usually the final network layer or the last layer within each stage—and consequently ignore many intermediate and shallow layers that contain valuable spatial and semantic information. In addition, these methods assign equal importance and weight to all selected layers, disregarding the differences in the influence of each layer. In this work, we introduce AdaLayer-CAM, a fully adaptive framework that integrates activation information from all convolutional layers and assigns each layer a weight based on the data and its activation energy. This adaptive weighting makes the layers with stronger discriminative content more prominent while reducing the impact of weaker or noisier layers. Moreover, a soft fusion mechanism replaces the traditional hard-max combination, which gradually

and thresholdedly integrates the weighted maps, resulting in more stable and accurate final heatmap localization. The resulting heatmaps preserve the structural precision of shallow layers and the semantic robustness of deeper ones—without requiring any modification to the model architecture or retraining of the network. Experimental results show that AdaLayer-CAMproduces class activation maps that are more accurate and less noisy than the original LayerCAM and other commonly used related methods, thereby improving interpretability and localization performance in weakly supervised learning approaches. The complete implementation, along with thecorresponding demonstration materials, is publicly accessible at:https://github.com/mhmdjgh/AdaLayer-CAM

Keywords—Convolutional Neural Networks (CNNs), Explainable AI, Class Activation Mapping (CAM), Adaptive LayerCAM, Multi-layer fusion, Interpretability, Weakly-supervised localization, Visualization.

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1053

Enhanced lung cancer detection through SMOTE-ENN resampling with optimized machine learning classifiers and LOOCV

*Alireza Kazempoor Choobari**

Social and Biological Network Analysis Laboratory(SBNA), Department of Computer Engineering

University of Kurdistan

Sanandaj, Iran

alireza.choobari@uok.ac.ir

Sadegh Sulaimany

Social and Biological Network Analysis Laboratory(SBNA), Department of Computer Engineering

University of Kurdistan

Sanandaj, Iran

s.sulaimany@uok.ac.ir

Abstract—Today, the use of machine learning approaches has proven to be effective in various fields of medicine and healthcare for the diagnosis and prediction of diseases, leading to accurate analysis and classification and assisting healthcare and medical professionals. Lung cancer is one of the deadliest cancers globally, and its early diagnosis is extremely important. In this research, we sought to achieve good results by performing various and methodical data preprocessing steps, such as standardization and using SMOTE-ENN for resampling and noise sample removal, utilizing powerful and optimized machine learning classifiers, and employing the most reliable validation method, namely LOOCV (Leave-One-Out Cross-Validation). We evaluated the classifiers using various evaluation metrics and ultimately achieved the best results with the Random Forest model: Accuracy = 99.38% and F1-Score = 98.85%. The output of this research indicates that appropriate preprocessing and data cleaning by ignoring and removing noisy samples, and using an optimized version of machine learning-based classifiers can lead to an improvement in lung cancer diagnosis.

Keywords—Lung cancer, Classification, Machine learning, Preprocessing.

Article Code: 1055

VerifyWise: A Modular AI Governance Platform for Data-Driven Compliance Management

*Mohammad Khalilzadeh**

Faculty of Technology and Engineering

Raja University

Qazvin, Iran

0009-0005-5540-6360

Firouzeh Razavi

Department of Information Technology Management - Electronic Business, Faculty of Humanities

Raja University

Qazvin, Iran

f.razavi@raja.ac.ir

Abstract—The advancing integration of artificial intelligence into mission-critical domains demands governance frameworks that go beyond static rules and manual auditing. This paper presents VerifyWise, a modular platform engineered for data-driven AI compliance management. Rather than relying on speculative machine learning or fuzzy logic components, VerifyWise focuses on concrete, implemented functionalities: structured risk registers, policy-to-control mapping interfaces, multi-framework compliance tracking, vendor risk assessments, evidence management, and built-in data security controls. The platform enables comprehensive risk capture and compliance monitoring across multiple governance standards, providing unified workflows for risk assessment, control implementation, third-party oversight, and audit reporting. By examining the system's architecture, data model, security design, and modules, we demonstrate how a structured, modular approach supports transparency, adaptability, scalability, and data protection in AI governance. While advanced analytics modules remain future work, the current implementation offers a robust foundation for deploying practical compliance systems aligned with international standards.

Keywords—AI Governance, Compliance Management, Risk Assessment, Policy Mapping, Data-Driven Analytics, Data Security, Intelligent Systems.

Article Code: 1057

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Adaptive Deep Learning Framework with Uncertainty-Aware Drift Detection, and Adversarial Robustness for Anomaly Classification in Power Systems State Estimation

Seyed Mohammad Shobeiry
PhD student in Electrical Engineering
Faculty of Electrical Engineering
Shahid Beheshti University
Tehran, Iran
s_shobeiry@sbu.ac.ir

*Mohammad Taghi Ameli**
Professor in Electrical Engineering
Faculty of Electrical Engineering
Shahid Beheshti University
Tehran, Iran
m_ameli@sbu.ac.ir

Mohammad Sadegh Sepasian
Associate Professor in Electrical Engineering
Faculty of Electrical Engineering
Shahid Beheshti University
Tehran, Iran
m_sepasian@sbu.ac.ir

Abstract—Power System State Estimation (PSSE) is increasingly threatened by False Data Injection Attacks (FDIAs) and Sudden Load Changes (SLC) under evolving operational conditions driven by Concept Drift (CD). To address these challenges, this study introduces an Adaptive Deep Learning (ADL) framework that integrates uncertainty-aware drift detection and adversarial robustness for anomaly classification in non-stationary power systems. The proposed framework employs multi-scale temporal attention to capture dynamic behaviors across multiple time horizons, applies Bayesian uncertainty estimation via Monte Carlo Dropout to distinguish epistemic from aleatoric uncertainty, and adopts FGSM-based adversarial training to enhance resistance to evasion attacks. A memory buffer with class-specific importance scoring ensures continual adaptation while avoiding catastrophic forgetting. An experimental evaluation on the IEEE 14-bus system shows that the ADL model maintains 83.6% accuracy under severe concept drift (only 3.4% below its baseline), while conventional deep learning models drop by 28.3%. It also achieves 97.06% accuracy under adversarial conditions, exhibiting strong generalization to various attack types. Overall, the proposed framework delivers robust, uncertainty-aware anomaly detection, enhancing cybersecurity resilience in modern power grids.

Keywords—Anomaly Classification, Adaptive Deep Learning, Power Systems State Estimation, Concept Drift.

Article Code: 1064

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Optimized Reconstructive Pruning for Trojan Detection in Neural Networks with RL-GNN

*Mohsen Askari**
Department of Computer Engineering
Malek Ashtar University of Technology
Tehran, Iran
Msn.asgary@gmail.com

Shahrooz Janbaz
Department of Computer Engineering
Malek Ashtar University of Technology
Tehran, Iran
janbaz@mut.ac.ir

Abolfazl Rahmani
Department of Computer Engineering
Malek Ashtar University of Technology
Tehran, Iran
abolfazlrahmani1@chmail.ir

Abstract—Trojan attacks, embedding malicious modifications in deep neural networks (DNNs), pose significant threats to their reliability in safety-critical applications, such as autonomous driving and medical diagnostics. These stealthy backdoors trigger erroneous outputs on specific inputs while maintaining normal behavior otherwise, challenging traditional validation methods. This study proposes an O`ptimized Reconstructive Pruning (ORP) framework to enhance Trojan detection across diverse DNN architectures, including convolutional networks and transformers. ORP integrates reconstructive pruning, which mitigates accuracy degradation post-pruning, with a novel hybrid Reinforcement Learning-Graph Neural Network (RL-GNN) optimizer. This optimizer dynamically navigates complex pruning parameter spaces, improving scalability and efficiency. Evaluated on TrojAI benchmarks comprising over 2,500 models, ORP achieves a classification accuracy of 92.3%, surpassing baseline pruning methods by 16%, while reducing execution time by 36%.Additionally, ORP reports an AUC-ROC of 0.95, F1-score of 0.92, and Precision-Recall curve area of 0.94, demonstrating superior performance across metrics. The framework demonstrates robustness in low-resource settings and noisy environments, effectively amplifying Trojan-induced asymmetries. These advancements enhance the detection of sophisticated backdoors, ensuring reliable DNN deployment. Future research may explore ORP's integration into federated learning and its applicability to emerging adaptive threats, strengthening AI security in real-world systems.

Keywords— Trojan detection, deep neural networks, reconstructive pruning, reinforcement learning, graph neural networks, AI security

Article Code: 1065

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Interpretable Ensemble Learning for Predicting the Non-linear Moment Capacity of Bolted Extended Endplate Moment Connections

Matin Alizadeh
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
0009-0009-6989-910X

S. Mohammad Hosseini
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
0009-0001-5822-7529

*Mahmoud. R. Shiravand**
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
0000-0002-8740-2242

Abstract— The Bolted Extended Endplate Moment-resisting connection (BEEP) is one of the prequalified moment-resisting steel connections specified in AISC 358 and European codes. Its reliable seismic performance is supported by extensive experimental and analytical studies. However, conventional analytical models often fail to capture the complex interplay of component yielding, bolt fracture, and stiffness degradation. In this paper, an interpretable machine learning (ML) framework is proposed to predict the nonlinear behavior of BEEP connections. A wide range of experimental studies was reviewed, analyzed, and systematically assembled into a comprehensive database. An ensemble learning framework, integrating Random Forest, XGBoost, HISTGM, KNN and LightGBM models, was trained to predict the connection's peak and ultimate strength. Trainded models have 0.95 and higher as **RR22** score for the test data. High predictive accuracy was achieved on unseen test data. Crucially, the interpretability of these models is addressed using SHapley Additive exPlanations (SHAP), providing clear, physically consistent interpretations that link geometric and material parameters to the connection performance. The result is a transparent, accurate, and rapid tool for preliminary assessment and design check of this connection.

Keywords—Bolted Extended Endplate Connection, Ensemble Learning, Explainable AI, Seismic Performance Prediction

Article Code: 1071

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Enhanced Diffusion Policy for Robust Multimodal Stock Trading through the Integration of Technical Indicators and Fundamental News

Amirali Vakili
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
am.vakili@mail.sbu.ac.ir

Mahdi Shahbazi Khojasteh
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
m.shahbazikhojasteh@alumni.sbu.ac.ir

Mahan Veisi
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
mah.veisi@mail.sbu.ac.ir

*Armin Salimi-Badr**
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a_salimibadr@sbu.ac.ir

Abstract—The pursuit of automated trading strategies that can consistently outperform market benchmarks remains a significant challenge in financial markets. Although Deep Reinforcement Learning (DRL) has shown promise for stock trading, conventional DRL approaches often rely on restrictive policy representations that fail to capture the multimodal nature of optimal trading decisions. To address this limitation, this paper applies the Enhanced Diffusion-QL (EDiffusion-QL) framework to automated stock trading on U.S. equity markets, integrating it with a Bidirectional Long Short-Term Memory (BiLSTM) state encoder that processes both technical indicators and financial news embeddings. With EDiffusion-QL, trading actions are generated by using a diffusion-based policy and twin Q-learning critics for robust value estimation. We evaluate our proposed method on a comprehensive dataset comprising 30 U.S. stocks spanning from 2009 to 2020, comparing performance against state-of-the-art DRL baselines. The results demonstrate that our model achieves superior performance, attaining the highest cumulative return, Sharpe ratio, and average profitability per trade. By modeling multi-modal action distributions, EDiffusion-QL establishes a more robust and profitable approach to portfolio optimization and automated trading.

Keywords—Automated Trading, Deep Reinforcement Learning, Diffusion Policy, Stock Market, Q-Learning.

Article Code: 1076

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Fibroglandular Tissue Classification in Breast MRI: A Comparative Study of Automated Decision Strategies

Meysam Khalaj
CIPCE, School of Electrical and Computer Engineering
College of Eng., University of Tehran
Tehran, Iran
meysamkhalaj@ut.ac.ir

Arvin Arian
Advanced Diagnostic and Interventional Radiology Research Center (ADIR), Cancer Institute,
Tehran University of Medical Sciences (TUMS)
Tehran, Iran
arvinaryan@yahoo.com

Ala Torabi
Department of Radiology Shariati Hospital
Tehran University of Medical Sciences
Tehran, Iran
alatorabi2477@gmail.com

Nasrin Ahmadinejad
Advanced Diagnostic and Interventional Radiology Research Center (ADIR), Cancer Institute,
Tehran University of Medical Sciences (TUMS), Tehran, Iran
n_ahmadinejad@yahoo.com

Masoumeh Gity
Advanced Diagnostic and Interventional Radiology Research Center (ADIR), Cancer Institute,
Tehran University of Medical Sciences (TUMS)
Tehran, Iran
p_gity@yahoo.com

Seyedeh Nooshin Miratashi Yazdi
Cancer Research Institute
Tehran University of Medical Sciences
Tehran, Iran
n_miratashi@yahoo.com

Mohammad Pooya Afshari
CIPCE, School of Electrical and Computer Engineering
College of Eng., University of Tehran
Tehran, Iran
pooya1380@gmail.com

Melika Sadeghi Tabrizi
CIPCE, School of Electrical and Computer Engineering
College of Eng., University of Tehran
Tehran, Iran
melikasadeghi16@ut.ac.ir

*Hamid Soltanian-Zadeh**
CIPCE, School of Electrical and Computer Engineering
College of Eng., University of Tehran
Tehran, Iran &
School of Cognitive Sciences, IPM
Tehran, Iran

Abstract—Fibroglandular tissue (FGT) assessment in breast magnetic resonance imaging (MRI) is a key factor in breast cancer risk evaluation and follows the BI-RADS lexicon standard. Most automated methods have focused on segmentation, while classification-based approaches remain limited. Previous studies have often analyzed both breasts together, overlooking BI-RADS recommendations for sidespecific evaluation and alternative decision strategies. This study investigates three assessment methods for automated FGT classification: the conventional BI-RADS Maximum Rule, a Probability Averaging Rule for bilateral integration, and a Lower-Uncertainty Rule that uses Shannon entropy to prioritize more confident predictions. These strategies were tested using three deep learning architectures, including MobileNetV2, ResNeXt-26, and a hybrid ViT-ResNet model, on 654 pre-contrast T1-weighted breast MRI scans. Across ten independent runs, the ViT-ResNet model with the Probability Averaging Rule achieved the highest performance, with an accuracy of 0.85, an F1 score of 0.84, and a Cohen's kappa of 0.78. Both proposed strategies surpassed the conventional rule, and the expert-annotated dataset is publicly released to enable reproducible research.

Keywords—Fibroglandular Tissue Classification, Breast MRI, BI-RADS Assessment, Deep Learning, Shannon Entropy

Article Code: 1078

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Mamba Meets Sleep: Do State Space Models Outperform CNNs for EEG Classification?

Mostafa Mehrabi

Faculty of Computer Science and Computer Engineering

Shahid Beheshti University

Tehran, Iran

mosta.mehrabi@Mail.sbu.ac.ir

*Hamed Malek**

Faculty of Computer Science and Computer Engineering

Shahid Beheshti University

Tehran, Iran

h_malek@sbu.ac.ir

Abstract—Automatic sleep stage classification from EEG signals is critical for diagnosing sleep disorders, yet most studies evaluate architectures in isolation. We present a systematic evaluation of 15 models across five families: traditional machine learning, residual CNNs, hybrid ResNet-BiLSTM, efficient CNNs, and state space models. Using Sleep-EDF with leave-one-subject-out cross-validation on 20 subjects, we assessed subject-independent generalization. Mamba Base achieved the highest accuracy (93.32% ± 3.01%) and recall (87.11%), while ResNet8 provided comparable accuracy (93.20%) with lowest variance (2.47%). Surprisingly, adding recurrent layers or increasing depth degraded performance, suggesting simpler architectures better capture temporal dependencies in 30-second EEG epochs. State space models combine transformer expressiveness with linear complexity, making them suitable for clinical deployment. Traditional ML methods achieved 90-91% accuracy with minimal computational cost, offering interpretable alternatives. Our comprehensive comparison reveals clear trade-offs between accuracy, stability, efficiency, and deployment constraints, providing practical guidance for architecture selection based on clinical requirements and resource availability. Implementation code is available at: https://github.com/mostafamehrabii/sleep-stage-classification.

Keywords— Sleep stage classification, EEG signals, state space models, Mamba, deep learning, convolutional neural networks, subject-independent validation

Article Code: 1087

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Speech-Based Conversational Business Intelligence System Leveraging Large Language Models for Accounting Analytics

*Fatemeh Mohsennia**
Department of Computer Engineering
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
fatemeh.mohsennia@mail.um.ac.ir

Mohsen Kahani
Department of Computer Engineering
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
kahani@um.ac.ir

Morteza Fardin
Zharf Pouyan Baran
Information Technology Company
Mashhad, Iran
fardin.morteza271@gmail.com

Abstract—In this study, a speech-based Conversational Business Intelligence (C-BI) system was designed and implemented to facilitate natural interaction between users and accounting software. The proposed system provides a seamless communication channel that enables non-technical users to obtain financial reports and analytical insights through Persian speech or natural language queries, without navigating complex interfaces. The architecture integrates two advanced components: an Automatic Speech Recognition (ASR) model based on Whisper, fine-tuned with domain-specific Persian accounting data, and a Large Language Model (LLM) responsible for rewriting, interpreting, and extracting key concepts from user queries. In addition, the BGE-M3 embedding model was employed to generate semantic representations of queries and perform similarity-based retrieval within a vector database. Experimental results demonstrated that the fine-tuned Whisper model (v2) achieved a Word Error Rate (WER) of 0.11 and a BLEU score of 0.817, while the Gemma3– 12B LLM reached an overall accuracy of approximately 90% in intent classification and entity extraction. These findings confirm that combining speech and language models within the Conversational BI framework provides an effective and localized solution for improving user experience in Persian accounting systems and enhancing the accessibility of financial data analytics.

Keywords—Conversational Business Intelligence, Automatic Speech Recognition, Natural Language Processing, Large Language Models, Accounting Analytic

Article Code: 1088

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Hybrid CNN-Based Deepfake Image Detection with Enhanced Cross-Forgery Generalization

Alireza Maleki
RIV Lab., Department of Computer Engineering
Bu-Ali Sina University
Hamedan, Iran
alireza.maleki@eng.basu.ac.ir

Hanie Zomorrodi
RIV Lab., Department of ComputerEngineering
Bu-Ali Sina University
Hamedan, Iran
h.zomorrodi@eng.basu.ac.ir

Hassan khotanlou
RIV Lab., Department of Computer Engineering
Bu-Ali Sina University
Hamedan, Iran
khotanlou@basu.ac.ir

Atefeh Mosafer
RIV Lab., Department of Computer Engineering
Bu-Ali Sina University
Hamedan, Iran
atefeh.mosafer2@gmail.com

Abstract—In recent years, the rapid advancement of deep learning-based technologies for generating fake content has heightened the need for precise, reliable methods to detect fake images. In this study, an efficient method for distinguishing fake images from real ones using video frames is proposed. During preprocessing, video frames are isolated using a face-detection algorithm to better concentrate on the face instead of other unrelated objects. The ResNet-101 and EfficientNet-V2-s models are used in a concurrent dual-branch architecture to extract complementary characteristics from the images at various degrees of detail. The final classification between the real and fake image classes is achieved by concatenating the outputs of these two networks and feeding them into a four-layer Multi-Layer Perceptron (MLP). Two modules—Inter-Class Data Interaction and Dual-Domain Feature Interaction—are integrated to enhance the model's generalization and its ability to accurately detect diverse types of forgery. Results from experiments on a large dataset show that the proposed model performs admirably, achieving 98.26% accuracy and an AUC of 99.84%. Additionally, an analysis of the model's performance among different datasets shows that the suggested approach outperforms earlier approaches and exhibits strong generalization.

Keywords—Deepfake Detection, Feature Interaction Learning, Cross-Forgery Generalization, Hybrid Feature Extraction

Article Code: 1092

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

SB-Softmax: A Parametric and Robust Replacement for Softmax in Deep Learning Models

*Morteza Taheri**

Department of Computer Engineering

University of Zanjan

Zanjan, Iran

taheri.morteza674@gmail.com

Abstract—The conventional Softmax often causes numerical instability and overconfident predictions because of its unbounded exponential function, especially in mixed-precision training. This paper introduces SB-Softmax (Sigmoid-Bound Softmax), a bounded and parametric normalization layer with learnable parameters α and β, initialized using Particle Swarm Optimization (PSO). SB-Softmax keeps the same normalization structure as Softmax but uses bounded derivatives, which leads to more stable gradients and lower runtime. Experiments on MNIST, CIFAR-10, and CIFAR-100 show better accuracy and calibration compared to standard Softmax, along with reduced training time. Accuracy increases from 0.7736 to 0.7882 on CIFAR-10 and from 0.4576 to 0.4887 on CIFAR-100, while Expected Calibration Error (ECE) is reduced by up to 15%. These results show that SB-Softmax can be a simple and practical replacement for Softmax in deep learning models.

Keywords — Softmax; SB-Softmax; normalization layer; mixed precision training; calibration; deep neural modeling; particle-swarm-driven optimization (PSO)

Article Code: 1097

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Development and Validation of an Explainable Machine Learning Framework for Decision Support in Student Admission Management

*Hananeh Teshnehlab**

Artificial Intelligence Specialist,

Rahyab Payam Gostaran Company,

Tehran, Iran.

h.teshnehlab@rahyabpg.co.ir

Mohammad Reza Ayatollahzadeh Shirazi

Chief Technology and Business Development Officer,

Rahyab Payam Gostaran Company,

Tehran, Iran.

mr.shirazi@rahyabvas.com

Abstract— Artificial Intelligence (AI) has become an essential component in decision-making systems across various domains. In the field of education, explainable machine learning (XML) frameworks enhance transparency and improve data-driven decision-making. An explainable multi-layer architecture is proposed for improving tenth-grade student enrollment decisions using academic and behavioral data from 195 students at Mofid Girls School, Tehran, Iran. The model integrates Decision Tree and Random Forest algorithms for feature selection and applies Logistic Regression, XGBoost, and Multilayer Perceptron (MLP) to predict enrollment outcomes. Among the models, Logistic Regression demonstrated the highest interpretability and accuracy, achieving an accuracy of 0.7599 and an AUC of 0.84. The model was embedded in an interactive chatbot-based decision support system for school administrators. The study is based on a single-school case study, and future multi-institutional validation is required to confirm generalizability.

Keywords— Machine Learning, Student Enrollment, Explainable AI, Decision Support System, Logistic Regression, XGBoost, Multilayer Perceptron.

Article Code: 1100

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Calibrating Large Language Models for High-Quality Automated Writing Feedback

Mohammad Sajjad Aghadadi

NLP Lab, School of ECE, College of Engineering University of Tehran

Tehran, Iran

sa.aghadadi@ut.ac.ir

Heshaam Faili

NLP Lab, School of ECE, College of Engineering University of Tehran

Tehran, Iran

hfaili@ut.ac.ir

Abstract—Automated assessment of IELTS Writing Task 2 remains a challenging problem due to the multidimensional nature of the scoring rubric and the difficulty of generating criterion-aligned feedback. This study proposes a PMI-based calibration framework for enhancing LLM-driven evaluation and feedback generation. Using a large dataset of 38,205 annotated IELTS responses, we compute feature-level PMI weights to adjust model outputs in two configurations: (1) a Static PMI model with a global calibration vector, and (2) a Contextual PMI model that dynamically selects calibration weights based on semantic clustering of input essays. Experimental results on a 1,000-sample test set show that PMI calibration consistently improves performance across all evaluation metrics, including Cosine similarity, BERTScore, rubric consistency, and LLM-as-a-judge ratings. Notably, the Contextual PMI model yields the strongest gains—for example, improving cosine similarity in the Coherence and Cohesion criterion from 0.462 (zero-shot) to 0.780. These findings highlight the effectiveness of PMI-driven calibration, particularly in context-adaptive settings, and demonstrate its potential as a lightweight, interpretable, and scalable approach for improving automated writing feedback.

Keywords—Automatic Writing Feedback, Calibration, Pointwise Mutual Information, LLM

Article Code: 1101

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Attention-Driven Deep Generative Modeling of Temporal Networks

*Maryam Kani**
Department of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
ma.lotfalikani@mail.sbu.ac.ir

Sadegh Aliakbary
Department of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
s_aliakbary@sbu.ac.ir

Hamed Malek
Department of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
h_malek@sbu.ac.ir

Abstract—Deep generative models have been widely applied to complex network generation, enabling the synthesis of structures that closely resemble real-world systems. Among these models, temporal network models capture time-varying interactions by allowing nodes and edges to appear or change over time. In this work, we enhance LSTM-based temporal network generation by incorporating self-attention mechanisms in two separate models, namely pre-LSTM and post-LSTM models, to better capture long-range temporal dependencies and improve model accuracy. The proposed model is evaluated on six real-world temporal networks— Bitcoin, Wikipedia, Calls, Movies, MathOverflow, and Wikipedia Edits—using structural metrics such as mean degree, wedge count, power-law exponent, clustering coefficient, and centrality measures. Experimental results show that the attention-augmented model significantly reduces generation error for networks with long temporal spans, while networks with shorter temporal ranges may exhibit marginal performance degradation due to limited temporal signal strength.

Keywords— complex networks, temporal networks, network models, deep generative models, attention mechanism

Article Code: 1103

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Energy-Elastic Proximal Policy Optimization for Autonomous Soft Tissue Cutting in Robotic Surgery

Seyed Mohammad Mahdi Pazhouh
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
s.pajoh@mail.sbu.ac.ir

*Ahmad Ali Abin**
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a_abin@sbu.ac.ir

Armin Salimi-Badr
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a_salimibadr@sbu.ac.ir

Ali Eslami
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
ali.eslami@mail.sbu.ac.ir

Hasan Ali Ahmadi
Department of Anesthesiology, School of Medicine, Shahid Beheshti University of Medical Sciences
Tehran, Iran
ha.ahmadi@sbmu.ac.ir

Abstract— With the growing use of robotic systems in surgery, the need for autonomous and efficient soft tissue manipulation is increasing. This paper proposes an energy-efficient deep reinforcement learning method for autonomous soft tissue cutting. To narrow the simulation-to-reality gap, we vary spring stiffness coefficients to better capture the mechanical behavior of different tissues. We also introduce a pinch-point sampling strategy inspired by k-means++ to improve the spatial distribution of tensioning poins, and a reward shaping term that penalizes unnecessary energy use to encourage smoother, more stable tensioning. Evaluated on 16 standard cutting patterns, our approach maintains cutting accuracy comparable to state-of-the-art methods while achieving higher stability and consistency. These results suggest that the proposed Energy-Elastic Proximal Policy Optimization framework can improve both the efficiency and reliability of robotic soft tissue manipulation.

Keywords— Robotic surgery, deep reinforcement learning, soft tissue cutting, energy-efficient tensioning, autonomous surgery

Article Code: 1104

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

DiTA-RUL: Diffusion–Transformer Augmentation for CMAPSS

Yasamin Tafakor
Computer Science Department
Amirkabir University of Technology
Tehran, Iran
yasi.tafakor@aut.ac.ir

Ali Jahan
Computer Engineering Department
Amirkabir University of Technology
Tehran, Iran
alijahan@aut.ac.ir

Reza Tavakoli
Mathematical Sciences Department
Sharif University of Technology
Tehran, Iran
seyedreza.tavakoli@sharif.edu

*Siavash Ahmadi**
Electronics Research Institute
Sharif University of Technology
Tehran, Iran
s.ahmadi@sharif.edu

Babak Khalaj
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
khalaj@sharif.edu

Abstract— Accurately estimating the Remaining Useful Life (RUL) is critical for improving the reliability and safety of industrial systems. The CMAPSS dataset, which provides run-to-failure engine trajectories, is widely used as a benchmark for RUL prediction. However, limited training samples and the challenge of capturing complex degradation patterns make effective data augmentation essential in this field. In this work, we propose a diffusion-based data augmentation framework that employs a latent structure incorporating wavelet decomposition, attention mechanisms, and sequence modeling. We then apply a diffusion model within this latent space to generate augmented sensor signals. In the backward phase of the diffusion model, a transformer-based network reconstructs realistic engine degradation signals using the learned latent representations. The generated data not only capture the statistical and temporal characteristics of true sensor signals but also improve RUL prediction compared to many existing approaches. We demonstrate the quality of the augmented data both by visual comparison with actual signals and by showing their effectiveness in boosting RUL prediction performance. Code and data: https://github.com/yasitafakor/Diffusionmodel-using-latent-space-on-CMAPSS

Keywords—Remaining Useful Life, Prognostics and Health Management, Diffusion Model, Transformer, Data Augmentation, C-MAPSS

Article Code: 1109

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Ensemble Machine Learning for Predicting Stroke Patient Survival

Seyedeh Maryam Mousavi
Department of Statistics and Epidemiology, Faculty of Medicine, Zanjan University of Medical Sciences, Zanjan, Iran.
m.mousavi6024@gmail.com

Samira Ahmadi
Social Determinants of Health Research Center, Health and Metabolic Diseases Research Institute, Zanjan University of Medical Sciences, Zanjan, Iran.
samira.ahmadizn@gmail.com

Solmaz Norouzi
Zanjan Applied Pharmacology Research Center, Health and Metabolic Diseases Research Institute, Zanjan University of Medical Sciences, Zanjan, Iran.
Research Intelligence Unit, Zanjan University of Medical Sciences, Postal Code 45156-13112, Zanjan, Iran.
snorouzibiostatistics@gmail.com

Abstract—Background: Stroke is a leading global cause of death and disability, making accurate survival prediction crucial for clinical decision-making. Recent advances in machine learning provide robust tools for survival analysis, effectively handling complex, non-linear, and high-dimensional data. This study evaluated and compared the predictive performance of Survival Support Vector Machine (SSVM), Random Survival Forest (RSF), and their ensemble models in identifying factors influencing stroke patient survival. Methods: We evaluated the performance of Survival Support Vector Machine (SSVM), Random Survival Forest (RSF), and their ensemble models in predicting survival outcomes in a cohort of 332 stroke patients from Imam Khomeini Hospital in Ardabil, who were monitored for up to 15 years following their stroke (2007-2022). Collected variables encompassed demographic, clinical, and lifestyle factors, including age, sex,

Education, comorbidities (e.g., hypertension, heart disease), stroke type (ischemic or hemorrhagic), and smoking history. Alongside the Cox model, we implemented SSVM and RSF, then constructed four ensemble models (Voting, Stacking, Bagging, and Boosting) by integrating their outputs to capitalize on the strengths of both algorithms. Performance was assessed using accuracy, sensitivity, specificity, area under the curve (AUC), and survival function estimation error. Results: Ensemble models outperformed the individual RSF, SSVM, and Cox models in predicting long-term mortality. The Voting ensemble achieved the highest performance, with an AUC of 0.855, overall accuracy of 0.833, 95% confidence interval of 0.747–0.962, sensitivity of 0.867, and specificity of 0.762. It demonstrated clear superiority over SSVM (AUC = 0.711), RSF (AUC = 0.828), and Cox (AUC = 0.768). Variable importance analysis in the optimal Voting model ranked age, hypertension, heart disease history, and hemorrhagic stroke type as the top predictors of long-term mortality. Conclusion: Integrating SSVM and RSF within ensemble frameworks enhances both predictive accuracy and interpretability in stroke survival analysis. The Voting ensemble, as the most accurate model, offers a valuable clinical tool for identifying high-risk patients and guiding targeted interventions.

Keywords— Survival Support Vector Machine, Random Survival Forest, Survival Support Vector Machine and Random Survival Forest Ensemble, Survival Analysis, Stroke.

Article Code: 1111

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Photonic Quantum Hardware for AI Optimization Tasks

*Sarah Daneshzad**

Department of Physics

Faculty of Physics, Shahid Beheshti University

Tehran, Iran

sarah.daneshzad@gmail.com

Gholam-Mohammad Parsanasab

Department of Electrical Engineering

Faculty of Electrical Engineering, Shahid Beheshti University

Tehran, Iran

gm_parsanasab@sbu.ac.ir

Abstract— Recent advances in photonic quantum computing offer a promising route for accelerating combinatorial optimization problems central to artificial intelligence (AI). This paper presents a hardware-aware simulation framework for implementing the Quantum Approximate Optimization Algorithm (QAOA) on photonic hardware platforms, extending prior photonic QAOA implementations for the Max-Cut problem [1] to more general AI-relevant graph optimization tasks. Unweighted, weighted, and balanced graph-partitioning problems are modeled as representative benchmarks, and their cost and mixer unitaries are compiled into integrated-photonics primitives based on Mach–Zehnder interferometers, programmable phase shifters, and cross-Kerr-type interactions. Depth- $\boldsymbol{p}$ photonic QAOA circuits ($\boldsymbol{p} = \mathbf{1}$-$\mathbf{3}$) are simulated using the Strawberry Fields photonic backend with realistic models of propagation loss, phase noise, and detector inefficiency. The results show monotonic improvement in approximation ratio with increasing depth, achieving up to $\boldsymbol{\rho \approx 0.87}$ for graphs with $\boldsymbol{n \leq 18}$ and less than 10% degradation under nominal noise conditions. Warm-start initialization reduces the number of optimization iterations by approximately 30% compared with random initialization. Benchmark comparisons with Goemans–Williamson and simulated annealing demonstrate that photonic QAOA attains near-classical performance for small- to medium-sized graphs, indicating that integrated photonics is a feasible platform for near-term quantum optimization.

Keywords— photonic quantum computing, Quantum Approximate Optimization Algorithm (QAOA), integrated photonics, AI optimization, and quantum hardware platforms.

Article Code: 1114

Integrating Physicochemical Analysis and Deep Learning for SGD Detection: A Case Study of the northern part of the Persian Gulf

Ali Lotfi
Dept. of Mineral and Groundwater Resources Faculty of Earth Sciences,
Shahid Beheshti University
Tehran, Iran
alilotfi256@gmail.com

Razieh Lak
Dept. Mineral and Groundwater Resources Research Institute for Earth Sciences Geological Survey of Iran
lak_ir@yahoo.com

Mohsen Ehteshami
Dept. of Mineral and Groundwater Resources Faculty of Earth Sciences,
Shahid Beheshti University
Tehran, Iran

SaeedReza Kheradpisheh
Dept. Data and Computer Sciences
Shahid Beheshti University
Tehran, Iran
kheradpisheh@ut.ac.ir

*Yaser Nikpeyman**
Dept. of Mineral and Groundwater Resources Faculty of Earth Sciences,
Shahid Beheshti University
Tehran, Iran

Abstract— Submarine Groundwater Discharge (SGD) is a hidden yet critical component of the global water cycle, facilitating the transport of nutrients and chemical compounds into coastal ecosystems. This study aims to identify potential SGD zones in the northern part of the Persian Gulf using a hybrid approach that integrates expert hydrogeological analysis with machine learning algorithms. By utilizing physicochemical seawater data—including temperature, salinity, dissolved oxygen, and chlorophyll—collected from the region, a framework was developed to automatically detect hydrochemical anomalies. The methodology involved a multi-phase analysis using supervised (Gradient Boosting) and unsupervised (SOM and K-Means) models, culminating in a deep learning model (LSTM) trained on synthetic data to address data scarcity. The results demonstrate that the developed models can identify high-probability SGD hotspots with high accuracy, particularly in the Strait of Hormuz and along the Bushehr coast. Furthermore, a significant spatial correlation was observed between these identified zones and major tectonic structures, such as the Qatar–Kazerun fault. These findings provide a scalable and intelligent pathway for monitoring coastal water resources in the Persian Gulf.

Keywords—Submarine Groundwater Discharge (SGD); Persian Gulf; Machine Learning; Deep Learning; Physicochemical Data

Article Code: 1123

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

A Multi-Layer Comparative Analysis of Human and Machine Reviewing Using Large Language Models

*Mostafa Karimi Manesh**

Faculty of computer science and engineering, Shahid Beheshti University

Tehran, Iran

m_karimimanesh@sbu.ac.ir

Mehrnoush Shamsfard

Faculty of computer science and engineering, Shahid Beheshti University

Tehran, Iran

m-shams@sbu.ac.ir

Abstract— This study conducts a comparative analysis of the behavior of human reviewers and large language models in the peer-reviewing of scientific articles, aiming to systematically examine the cognitive, analytical, and linguistic dimensions of these two approaches. In the first step, a collection of 83 papers in the field of natural language processing and 2,379 human review comments was gathered, and corresponding machinegenerated reviews were produced using the GPT-4 model. Subsequently, a four-layer annotation framework—covering article sections, evaluation aspects, reviewer functions, and comment severity—was designed to investigate behavioral and content-level differences between the two reviewing systems. The results indicate that the language model places greater emphasis on structural and technical components of the paper, such as methodology and data, and demonstrates strong performance in quantitative analysis; whereas human reviewers focus more on conceptual aspects such as novelty, comparison with prior work, and scientific impact. From a linguistic perspective, the language model tends to rewrite criticisms as constructive suggestions, while human reviewers generally adopt a more direct and critique-oriented style. The conceptual overlap between the two types of reviewing reaches up to 79% in technical sections and approximately 71% overall, reflecting a notable convergence between human and machine judgment.

Key contributions of this study include the development of a “section-wise specialized reviewing” approach and the creation of a Persian-language scientific peer-review dataset. The findings suggest that an intelligent integration of human and machine reviewing can enhance the accuracy, fairness, and transparency of scientific evaluation and pave the way for more advanced automated reviewing systems.

Keywords— Peer Review Analysis, Automated Peer Review, Large Language Models, Prompt Engineering, Scientific Paper Evaluation

Article Code: 1125

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Lightweight Hybrid Attention Recurrent Neural Network (HARN) for IDS in IoT Network Layer

*Mohammad Hossein Esfahani**
Department of Electrical Engineering
Iran University of Science and Technology (IUST)
Tehran, Iran
esfahani_m@elec.iust.ac.ir

Ali Sadr
Department of Electrical Engineering
Iran University of Science and Technology (IUST)
Tehran, Iran
sadr@iust.ac.ir

Abstract— The rapid expansion of the Internet of Things (IoT) has enhanced convenience and intelligence to many real-world applications, but it has also opened the door to increasingly sophisticated cyberattacks. Among them, zero-day attacks remain especially difficult to detect because they do not follow recognizable patterns and often bypass conventional intrusion detection systems (IDS). Addressing this issue, we introduce a lightweight Hybrid Attention Recurrent Neural Network (HARN) that combines LSTM, GRU, and attention layers to better capture the most informative patterns in network traffic. A key strength of the proposed approach is its carefully designed preprocessing stage: statistical feature extraction, Yeo–Johnson transformation, SMOTE rebalancing, and RFE-based feature selection work together to form a compact and clean feature set that noticeably improves the model's learning quality and metric outcomes. By setting the recurrent time-step to 1, the model avoids the accumulation of long-term memory and becomes more responsive to previously unseen attack behaviors. Evaluation on the TON-IoT dataset demonstrates that the proposed IDS achieves 99.67% accuracy, 99.55% precision, 99.75% recall, a 99.71% F1-score, and an AUC of 0.9984, outperforming prior methods on all metrics. These results demonstrate that the combination of the hybrid architecture and the strengthened preprocessing pipeline offers an effective and practical approach for improving security in modern IoT networks.

Keywords— Deep Learning, HARN, Intrusion Detection System, IoT, TON-IoT, Zero-Day Attack

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1126

Divide and Conquer: A Cascaded Architecture for Relation Extraction on Long-Tailed Datasets

*Parham Rahimi**

School of Computer Engineering

Iran University of Science and Technology, NAJM

Tehran, Iran

rahimi.parhaam@gmail.com

Behrouz Minaei-Bidgoli

School of Computer Engineering

Iran University of Science and Technology

Tehran, Iran

b_minaei@iust.ac.ir

Abstract— Relation Extraction (RE) from unstructured text is a critical task for populating Knowledge Graphs (KGs). The Distant Supervision (DS) paradigm, while effective for data annotation, suffers from two critical challenges: the wrong label problem (noise) and severe class imbalance in benchmark datasets. While recent generative models have shown promise, they struggle with these real-world data imperfections. This paper proposes a novel framework to address these challenges. First, we introduce a prompt-engineering methodology to convert traditional DS datasets (e.g., NYT) into a format suitable for fine-tuning generative models. Second, we demonstrate that a T5 model fine-tuned with this framework achieves state-of-the-art results on the NYT24 joint RE task, surpassing prior baselines. Finally, to combat extreme data imbalance, we propose a novel cascaded T5 architecture that improves classification F1-score by over 6% compared to a single-stage model. Our work provides a robust, generative solution for RE in noisy and imbalanced settings.

Keywords— Relation Extraction, Distant Supervision, Long-Tailed Datasets, Prompt Engineering Framework, Small Language Models

Article Code: 1129

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Interpretable Machine Learning for Rocking-Induced Settlement Prediction Using SHAP Analysis

Seyed Emad Miri

Faculty of Civil, Water, and Environmental Engineering

Shahid Beheshti University

Tehran, Iran

mirimiremad@gmail.com

*Hamid Mohammadnezhad**

Faculty of Civil, Water, and Environmental Engineering

Shahid Beheshti University

Tehran, Iran

h_mohammadnezhad@sbu.ac.ir

Abstract—Rocking shallow foundations have been increasingly recognized as an effective mechanism for seismic energy dissipation; however, their permanent settlement remains difficult to predict due to the complex soil-foundation interaction. This paper presents the development and evaluation of two supervised machine learning models, namely Extra Trees and LightGBM, using a dataset of 140 experimental cases. Based on learning curve analysis, the most optimized predictive behavior was achieved when 80% of the data was utilized during model training. Meanwhile, Extra Trees provided the best predictive performance on the test set, with a coefficient of determination of 0.945. To ensure interpretability rather than relying on black box behavior, the contribution of each input variable to the model output is investigated using Shapley Additive Explanations, describing how changes in each variable affect the predicted settlement. Accordingly, the slenderness ratio appears to be the most influential factor on rocking-induced settlement, followed by the rocking coefficient, maximum ground acceleration, and Arias intensity. On the other hand, the foundation area ratio and soil type have a relatively minor effect. The combination of ensemble learning models and Shapley Additive Explanations provides both strong predictive capability and transparent insights, therefore supporting more reliable and interpretable seismic design practices.

Keywords—Machine Learning, Rocking Shallow Foundations, SHAP, Interpretability, Supervised Learning

Article Code: 1132

NeuroPC-KGNN: A Causality-Driven Graph Neural Network Framework for Early Diagnosis of Parkinson's Disease

Fatemeh Ahouz
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
f_ahouz@sbu.ac.ir

*Ahmad Ali Abin**
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a_abin@sbu.ac.ir

Amin Golabpour
School of Allied Medical Sciences
Shahroud University of Medical Sciences
Sharoud, Iran
a.golabpour@shmu.ac.ir

Abstract—Early diagnosis of Parkinson's disease is essential for effective intervention, and voice signals provide a simple non-invasive tool for this purpose. However, many acoustic features are correlated rather than causal, limiting the robustness and generalizability of conventional machinelearning models. This study introduces NeuroPC-KGNN, a framework that integrates the Peter–Clark (PC) causal discovery algorithm with a Graph Convolutional Network (GCN) operating on a graph constructed from samples connected to their k-nearest neighbors. The PC algorithm identifies features that exhibit direct or indirect causal relationships with disease status, ensuring that non-causal and noisy variables are excluded. A GCN trained on a kNN-constructed graph then models the structural dependencies among samples. Using the UCI Parkinson dataset, NeuroPC-KGNN achieved 90% accuracy and 91% sensitivity on the test set. Permutation-based sensitivity analysis revealed that spread1, PPE, and MDVP:Flo(Hz) were the most influential features, showing the greatest impact on model performance when perturbed. The strong alignment between the causal graph and feature-importance results highlights the reliability and interpretability of the proposed framework. Our results demonstrate that NeuroPC-KGNN provides a robust, interpretable, and clinically meaningful approach for early PD diagnosis from voice data.

Keywords— Parkinson's disease, voice analysis, causal feature selection, PC algorithm, graph neural networks

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1142

Generating Pharmaceutical Molecules Using Multi-Path Deep Learning

Sanaz Hashemipour
dept. of Computer Science, Faculty of Mathematical Sciences
University of Tabriz
Tabriz, Iran
s.hashemi1400@ms.tabrizu.ac.ir

*Habib Izadkhah**
dept. of Computer Science, Faculty of Mathematical Sciences
University of Tabriz
Tabriz, Iran
izadkhah@tabrizu.ac.ir

Abolfazl Barzegar
Research Center for Biosciences and Biotechnology (RCBB)
University of Tabriz
Tabriz, Iran
barzegar@tabrizu.ac.ir

Abstract—The growing complexity of molecular data and the vast chemical space, encompassing an extraordinarily large number of potential compounds, pose significant challenges in drug discovery. In this work, the Multi-Path Long Short-Term Memory (MP-LSTM) model is proposed, a hybrid deep learning architecture for generating pharmaceutical molecules. The model employs multiple learning pathways that integrate LSTM and one-dimensional convolutional layers, enabling the extraction of diverse features from input sequences. Experimental results demonstrate that the MP-LSTM outperforms conventional LSTM and Bi-LSTM models in producing novel and chemically valid drug-like molecules. The generated compounds were further evaluated using the SwissADME platform and the RDKit library, confirming favorable pharmacokinetic properties. These results highlight the efficacy of the MP-LSTM model in exploring chemical space and demonstrate its potential to accelerate AI-driven drug discovery.

Keywords—multi-path LSTM (MP-LSTM), de novo drug design, deep learning, bioinformatics, artificial intelligence in medicinal chemistry

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1143

Persian Speech Emotion Recognition Using Multi-Scale SincNet and cGAN-based Ensemble Learning

*Ali Hekmat**

Faculty of Mathematical Sciences,

Shahid Beheshti University,

Tehran, Iran

al.hekmat@mail.sbu.ac.ir

Yasser Shekofteh

Faculty of Computer Science and Engineering,

Shahid Beheshti University,

Tehran, Iran

y_shekofteh@sbu.ac.ir

Abstract—Speech Emotion Recognition (SER), particularly for the Persian language, presents significant challenges due to data scarcity and emotional variability among speakers. This research proposes a novel multi-stage framework to enhance the accuracy, interpretability, and generalizability of Persian SER models. The proposed architecture consists of three primary stages: (1) feature extraction using a multi-scale interpretable SincNet, (2) representation learning and feature map reconstruction via conditional Generative Adversarial Networks (cGANs) to address class imbalance, and (3) a final classification decision based on a One-vs-All architecture with ensemble learning using hard and soft voting. To ensure a fair and robust evaluation, experiments were conducted on a refined version of the ShEMO dataset (with ASR-based corrected annotations) using a strict speaker-independent configuration, which is critical for real-world applicability. The results demonstrate a significant improvement over baseline methods across all key evaluation metrics. The proposed method achieved a Weighted Accuracy of 86.35%, a 1.56% increase over the best-performing baseline, HuBERT, which reached 84.79%. Furthermore, our approach achieved a Weighted F1- score of 86.36%, representing a 2.32% improvement over the HuBERT baseline's score of 84.04%.

Keywords—Speech Emotion Recognition, SincNet, cGAN, Ensemble Learning, Interpretability, ShEMO.

Article Code: 1144

Enhancing Imitation Learning for Humanoid Robots Using Vision Transformers and Time Contrastive Networks

Amirmohsen Sharifi
Dept. of Electrical and Computer Engineering
Isfahan University of Technology
Isfahan, Iran
amirmohsensois3@gmail.com

Maziar Palhang
Dept. of Electrical and Computer Engineering
Isfahan University of Technology
Isfahan, Iran
palhang@iut.ac.ir

Abstract— We propose a new method for skill imitation learning for humanoid robots using Vision Transformers and Time Contrastive Networks. For learning temporal representations of gestures, we compare two leading visual encoders, ViT-B16 and EfficientNetV2S. A triplet loss function with hard mining is utilized to generate discriminative embeddings from consecutive video frames, which provide dense reward signals for deep reinforcement learning. This enables a NAO humanoid robot to learn gesture imitation in the Webots platform. Our experimental results, strengthened by GroupKFold cross-validation, indicate that the ViT-B16 architecture consistently surpasses the fine-tuned EfficientNetV2S model. The embeddings from ViT-B16 form a more well-separated and structured latent space, as evidenced by dimensionality reduction visualizations, demonstrating a superior capture of temporal structure. Finally, the efficacy of our approach is demonstrated by integrating the ViT-B16 encoder into a reinforcement learning framework. We present quantitative evidence of policy convergence and qualitative results of the NAO robot successfully imitating a complex gesture, validating our end-to-end method in a practical robotic task.

Keywords—Imitation Learning, Vision Transformers, Time Contrastive Networks, Humanoid Robotics, Deep Reinforcement Learning, NAO Robot

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1152

Multi-Output Convolutional Neural Network for Predicting MCMI-IV Personality Profiles from Resting-State EEG

Marjan Nosouhi

School of Electrical and Computer Engineering

University of Tehran

Tehran, Iran

mnosouhi@ut.ac.ir

*Moein Esghaei**

Faculty of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

esghaei@gmail.com

Abstract— The integration of neuroscience and artificial intelligence has opened new avenues for understanding the relationship between brain activity and psychological characteristics. In this study, a multi-output Convolutional Neural Network (CNN) framework is employed to predict personality and clinical profile parameters derived from the Millon Clinical Multiaxial Inventory (MCMI) using resting-state electroencephalography (EEG) data. EEG recordings were obtained from 30 participants using a standard 19-channel montage under eyes-closed resting conditions. Following preprocessing and segmentation into fixed-length epochs, the CNN was trained using subject-wise cross-validation to prevent data leakage. Predictions were generated at the epoch level and aggregated to produce subject-level scores. Results show statistically significant trait-level correlations for selected dimensions, particularly Turbulent Personality (4B) ($r = 0.511$, $p < 0.05$, FDR-corrected), and high structural similarity (~0.95) between true and predicted inter-trait correlation matrices. These findings highlight the potential of deep learning approaches in computational psychiatry and personality neuroscience and support the feasibility of non-invasive EEG-based psychological assessment.

Keywords—electroencephalography (EEG); Millon Clinical Multiaxial Inventory (MCMI); deep learning; Convolutional Neural Networks (CNN), Personality Neuroscience, Computational Psychiatry

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1153

Vision-based Static Detection of Windows Malware with Rule-Augmented Feature Maps

Mahdi Seyfipoor
PhD at Computer Engineering
University of Tehran
Tehran, Iran
mahdiseyfipoor@ut.ac.ir

*Mohammad Mahdi Eskandari**
Undergraduate Student in Computer Engineering
University of Tehran
Tehran, Iran
m.mahdi.eskandari@ut.ac.i

Siamak Mohammadi
Associate Professor of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
smohamadi@ut.ac.ir

Abstract—Static malware detection has become a major challenge in recent years due to the increasing use of obfuscation, packaging, and multi-stage shellcode-based payloads in modern Windows threats. In this paper, we introduce a vision-based rule-augmented multi-channel framework that transforms Portable Executable (PE) files into visual feature maps composed of byte images, entropy gradients, opcode-density predictions, and a rule channel encoding highlevel malicious indicators such as shellcode regions, hand-mapping patterns, and syscall-like instruction sequences. A Vision Transformer (ViT) backbone processes these mergedrepresentations to learn global structural relationships that traditional opcode-based or feature-engineered static approaches fail to capture. Experiments were conducted on the Malimg dataset, MalwareBazaar, and a verified set of 3000 benign Windows executables. The proposed framework achieved 97.4% accuracy on Malimg and 95.1% accuracy on MalwareBazaar, yielding an overall average accuracy of 96.6% across both datasets. The corresponding F1-scores were 97.3% and 95.2% respectively, reflecting strong resilience against compressed binaries, modified UPX artifacts, and encrypted multi-stage shellcode payloads. Further experiments show that the method successfully detects samples that evade commercial engines such as Kaspersky, CrowdStrike, and ESET.

Keywords—Windows Malware Detection, Static Analysis, Vision Transformer (ViT), Rule-Augmented Feature Maps.

Article Code: 1154

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Reducing Speaker Leakage in Voice Conversion Using Speaker-Invariant Representation

*Alireza Sahragard Beirami**
Computer Engineering Department
Amirkabir University of Technology
Tehran, Iran
arsb@aut.ac.ir

Mohammad Mehdi Homayounpour
Computer Engineering Department
Amirkabir University of Technology
Tehran, Iran
homayoun@aut.ac.ir

Abstract— Voice conversion is a major research area in speech synthesis that focuses on transforming the timbre of a source speaker to match that of a target speaker. One of the most important challenges in performing voice conversion is the lack of parallel data. To handle this issue, most of the models tend to use two different encoders. One for encoding the speech content and one for encoding the speaker identity from the reference speech. However, these models suffer from the speaker leakage problem where the generated speech still contains the timbre of the source speaker. To handle this problem, we used speaker-invariant representations. To prove this, we trained TriAAN-VC, a Voice Conversion system with WavLM and WavLM-SPIN features. To assess speaker leakage, we used average Cosine similarity between the source and converted samples speaker embeddings. Objective evaluations show that using this method helps to reduce the timbre of the source speaker information in the synthesized speech.

Keywords—voice conversion, speech synthesis, speaker-invariant representation

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1156

PerSHOP: A Multi-Domain Persian Dataset for Shopping Dialogue Systems

Keyvan Mahmoudi

School of Electrical and Computer Engineering

University of Tehran

Tehran, Iran

kmahmoudi@ut.ac.ir

Heshaam Faili

School of Electrical and Computer Engineering

University of Tehran

Tehran, Iran

hfaili@ut.ac.ir

Abstract— Task-oriented dialogue systems require high-quality datasets that reflect realistic user–agent interactions. Persian, however, lacks resources in this area, especially for commercial and retail applications. We introduce PerSHOP, the first large-scale, multi-domain Persian dataset for shopping-oriented task-based dialogues. PerSHOP contains 1,061 multi-turn buyer–seller conversations grounded in a structured knowledge base of 750 products, 84 sub-domains, and 33 semantic slots. Dialogues were collected through a custom crowdsourcing platform and annotated using a hybrid rule-based and human verification pipeline. We provide benchmark results for intent classification and slot filling using zero-shot large language models and supervised transformerbased models. PerSHOP fills a critical gap in Persian dialogue resources and supports future research in domain-specific conversational AI. The dataset is publicly released for research use and is available at: https://github.com/keyvanmahmoudi/PerSHOP

Keywords—task-oriented dialogue, dialogue datasets, shopping dialogues, intent classification, slot filling, persian language

Article Code: 1157

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

PE-iForest: A Perturbation-based Framework for Explainable Anomaly Detection

*Aria Shakoo**

Department of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

a.shakoo@mail.sbu.ac.ir

Mahboobeh Riahi-Madvar

Department of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

m riahi@sbu.ac.ir

Abstract—As anomaly detection is increasingly deployed in high-stakes domains such as healthcare, finance, and cybersecurity, there is a critical need for methods that are simultaneously interpretable, scalable, and accurate. Explainable anomaly detection methods often suffer from high computational complexity in model-agnostic approaches or are limited by reliance on a single model in model-specific methods. This paper introduces a perturbation-based framework for anomaly detection that leverages vectorized perturbations to deliver high-speed explanations while remaining agnostic to the underlying detection algorithm. We instantiate the framework on Isolation Forest (iForest), a widely used scalable detector, to address its inherently "blackbox" nature and enhance user trust through interpretable outputs. We empirically evaluate PE-iForest against state-of-the-art explainers DIFFI, TreeSHAP and traditional detection baselines (LOF, HBOS, PCA) using four public benchmark datasets. Our optimized implementation achieves superior efficiency, generating explanations for 100 anomalies in 0.89 seconds on highdimensional Arrhythmia data, compared to 2.87 seconds for DIFFI. PE-iForest also yields explanations that are qualitatively and quantitatively aligned with more complex methods. Overall, the results of PE-iForest demonstrate that our simple, vectorized perturbation framework can outperform complex methods in speed while maintaining high trustworthiness.

Keywords—anomaly detection, explainable AI (XAI), Isolation Forest, perturbation explanation, SHAP, DIFFI, scalability

Article Code: 1169

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Community-Aware Influence Maximization on Heterogeneous Information Networks using Graph Neural Networks

*Amirreza Vafaei**

Department of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

ami.vafaei@mail.sbu.ac.ir

Sadegh Aliakbary

Department of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

s_aliakbary@sbu.ac.ir

Abstract—Influence Maximization (IM) is a pivotal task in network science, aiming to find a minimum set of seed nodes that optimizes the diffusion of information. While modern neural-driven frameworks, such as MAHE-IM, have shown success in Heterogeneous Information Networks (HINs), they often rely on less powerful embedding techniques like Skip-gram (e.g., metapath2vec) and lack the direct integration of the network's intrinsic modular organization. To overcome these constraints, we propose GNN-IM-CD (Graph Neural Network for Influence Maximization based on Community Detection). This Algorithm introduces two main enhancements: 1) replacing the Skip-gram component with a powerful Heterogeneous Graph Neural Network (H-GNN) to extract more expressive node features; and 2) integrating an Enhanced Community Detection (CD) layer using an aggressive Louvain method to ensure variety among seeds and leverage sub-structure knowledge. Experimental results on real-world HIN datasets demonstrate that GNN-IM-CD notably exceeds the performance of the leading MAHE-IM framework under both the Independent Cascade (IC) and Linear Threshold (LT) models, validating the power of combining GNNs with community structure awareness for effective influence spread.

Keywords—Influence Maximization, Heterogeneous Networks, Community Detection, Graph Neural Networks

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 11 73

Cross-domain prediction of ultimate strength in corroded reinforced concrete columns using domain adaptation techniques

Melina Sadeghi
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
meli.sadeghi@mail.sbu.ac.ir

Pooria Poorahad
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
p_poorahadanzabi@sbu.ac.ir

*Mahmoud R. Shiravand**
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
m_shiravand@sbu.ac.ir

Abstract—Reinforced concrete (RC) columns in corrosive media undergo complex nonlinear degradations, making accurate seismic performance assessment challenging. While machine learning (ML) offers predictive capabilities, conventional models often fail to generalize across datasets from different laboratories due to domain shifts in material properties and loading protocols. This study addresses this limitation by employing two domain adaptation techniques: correlation alignment (CORAL) and domain-adversarial neural networks (DANN). Using a comprehensive database of experimental tests on corroded RC piers, the methods are evaluated across three distinct domain split scenarios: failure mode, axial load ratio, and corrosion ratio. The study focuses on predicting the ultimate strength of the structures. Results indicate that both techniques significantly outperform baseline ML models in cross-domain scenarios, with DANN showing superior robustness. The findings demonstrate that domain adaptation effectively mitigates distribution discrepancies, offering a reliable framework for assessing aging infrastructure even when training and target data diverge.

Keywords—artificial intelligence, machine learning, domain adaptation, corrosion, RC columns

Article Code: 1184

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Adaptive Data Analysis for Density Estimation: A Pólya-Tree Approach

*Amir Hossein Hadavi**
Information Systems and Security Lab (ISSL)
Sharif University of Technology
Tehran, Iran
ah_hadavi@ee.sharif.edu

Mohammad Reza Aref
Information Systems and Security Lab (ISSL)
Sharif University of Technology
Tehran, Iran
aref@sharif.edu

Mohammad M. Mojahedian
Information Systems and Security Lab (ISSL)
Sharif University of Technology
Tehran, Iran
mojahedian@sharif.edu

Abstract—This paper revises the Adaptive Data Analysis (ADA) framework by modeling more constructive analyst–data interactions aimed at improving inference accuracy. Focusing on distribution estimation within a nonparametric Bayesian setting, we employ Pólya trees (PTs) as priors to enable adaptive selection of nonparametric queries that reduce estimation error without increasing query numbers. The hierarchical, interpretable, and conjugate nature of the framework allows subjective beliefs to be intuitively encoded as priors and efficiently updated to posteriors with tractable computations. Simulations validate the superior efficiency and robustness of the model compared with the nonadaptive approach. This structured ADA modeling offers a more realistic framework for research in human-in-the-loop systems and cognitive modeling of belief updating.

Keywords—Adaptive data analysis, Bayesian density estimation, Non-parametric priors, Pólya trees

Article Code: 1185

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Soil Shear Strength Prediction Using Genetic-Optimized Boosting Machine Learning Models

Mohammadreza Ghadami
Faculty of Civil, Water and Environmental Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
mo.ghadami@mail.sbu.ac.ir
ORCID ID: 0009-0002-2039-5460

Ali Noorzad
Faculty of Civil, Water and Environmental Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
a_noorzad@sbu.ac.ir
ORCID ID: 0000-0002-3785-7679

*Hamid Mohammadnezhad**
Faculty of Civil, Water and Environmental Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
h_mohammadnezhad@sbu.ac.ir
ORCID ID: 0000-0001-8697-2430

Abstract—Shear strength represents a vital characteristic of soils, viewed as their inherent ability to withstand failure under the influence of forces applied to the soil structure. However, it is worthy to note that conventional laboratory-based methods for the determination of soil shear strength parameter are soil sampling, time-consuming, and limited in scalability**.** In this context, empirical methods such as Skempton (1957), Karlsson and Viberg (1967), and Wroth and Wood (1985) were employed, including coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and mean square error (MSE). Among these empirical approaches, the Skempton approach (1957) performed better than the other empirical methods, achieving superior results across themetrics 0.6684, 0.0496, 0.0353, 0.1074, and 0.0022. Therefore, in this study, it has been tried to utilize a novel hybrid data-driven approach developed with the integration of machine learning boosting algorithms and a genetic models (which have been used by other

researchers) in testing phase (with a $R^2 = 0.7976$, RMSE = 0.0273, MAE = 0.0363, MAPE = 0.0837, and MSE = 0.0013). Also, in order to improve the estimation, a clear predictive equation has been developed by the hybrid CatBoost-GA model, which provides proper and practical formulation for engineering applications. Moreover, SHapley Additive exPlanations made it clear that the liquidity index is the important variable controlling shear strength, followed by the dry density index. Surprisingly, lower-importance variables, that is, the plasticity index and loam content, exhibited weak associations with small median SHAP analysis, while wet density, which has been emphasized in prior literature, was also found to be less effective. In the light of this subject, the proposed hybrid ML–GA framework presents a reliable, interpretable, and computationally efficient alternative to conventional laboratory testing for the soil shear strength prediction.

Keywords— Shear Soil Strength, Machine Learning, Genetic

algorithm- SHAP Analysis

Article Code: 1187

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Enhancing Drug Synergy Prediction through Heterogeneous Graph Transformers and Relation-Aware Attention Networks

Dorsa Abdi
Faculty of Computer Science and Mathematics
Iran University of Science and Technology
Tehran, Iran
dorsa.abdi14@gmail.com

Abdussalam Nik Kerdar
Faculty of Computer Science and Mathematics
Iran University of Science and Technology
Tehran, Iran
salamnik42@gmail.com

*Maedeh Sadat Tahaei **
Faculty of Computer Engineering
Iran University of Science and Technology
Tehran, Iran
mstahaei@aut.ac.ir

Amirali Maroufi
Faculty of Materials science and Engineering & Computer Science
Iran University of Science & Technology
Tehran, Iran
Amiralimaroufi82@gmail.com

Parinaz Kanan
Faculty of Computer Science and Mathematics
Iran University of Science and Technology
Tehran, Iran
parinaz.kanan@gmail.com

Abstract— In the context of ever-more-complex biomedical data, drug discovery has had to evolve to be a data-driven science. A complex disease like cancer often cannot be treated successfully with a single agent due to the possibility of resistance or failure at the network level of interactions between agents, as well as the biological patient. Thus, it is a computationally difficult problem to predict drug synergy—the added therapeutic effect of a pair of drugs. Recently, KGANSynergy illustrated the ability of knowledge graph attention networks to learn multi-relational biomedical dependencies. However, it is a general baseline model and is also limited in heterogeneous-oriented entity and relation encoding. In the present work, we build upon KGANSynergy in a com parative and heterogeneous-aware manner by implementing two alternative graph learning paradigms: (1) Heterogeneous Graph Transformer (HGT) for type-subtyped message passing, and (2) Relation-Aware Graph Attention (RAGAT) for fine-tuned relational modeling. Using publicly available benchmark datasets of drug-cell line interactions, we demonstrate how the

use of heterogeneity-aware approaches improves predictive accuracy and interpretability to inform preclinical drug screening. Ultimately, our work shows how multirelational learning is critical as we move towards explainable and scalable drug combinations. In our study, we focused on the binary classification task of drug synergy prediction and achieved state-of-the-art performance on the DrugCombDB dataset, where the proposed variants outperformed the baseline, and the HGT-enhanced architecture delivered the strongest results with an AUC of 0.8951 and an F1 score of 0.8214.

Keywords—drug synergy, knowledge graph, graph neural network, heterogenous graph transformer, attention mechanism

Article Code: 1188

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

CAMQL-NPC: Enhancing Survival Behaviors in Dynamic Game Environments via Internal State Augmentation

*Nima Salami**
Faculty of Computer Science and Engineering Shahid Beheshti University of Technology
Tehran/Iran
nimasalami115@gmail.com

Hassan Haghighi
Faculty of Computer Science and Engineering Shahid Beheshti University of Technology
Tehran/Iran
h_haghighi@sbu.ac.ir

Abstract—Designing intelligent Non-Player Characters (NPCs) in survival games requires agents to balance objective completion with self-preservation. A common limitation in conventional Reinforcement Learning (RL) frameworks, including standard Q-Learning, is the construction of a state space based exclusively on external environmental signals (e.g., target proximity), thereby overlooking the agent's internal physiological condition. Such a constraint often results in suboptimal behaviors, where agents pursue rewards despite critical health conditions, resulting in premature termination. This paper proposes a Context-Aware Modified Q-Learning (CAMQL) framework that integrates Internal State Augmentation (ISA). By incorporating the agent's health status into the state representation and employing a dynamic survival-weighted reward function, the NPC learns to prioritize evasion over looting during critical scenarios. Experimental results in a Unity-based simulation demonstrate that CAMQL achieves a significantly higher convergence rate and survival probability compared to baseline MQL approaches, effectively bridging the gap between aggressive exploration and defensive preservation.

Keywords— Non-Player Character (NPC), Context-Aware Q-Learning, Internal State Augmentation, Survival Games, Dynamic Reward Shaping

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1192

Neural-Network Surrogate Modeling for Fast Optimization of Compliant Flapping Mechanisms

*Hassan Sayyaadi**
Department of Mechanical Engineering
Sharif University of Technology
Tehran, Iran
sayyaadi@sharif.edu

Ali Zouelm
Department of Mechanical Engineering
Sharif University of Technology
Tehran, Iran
alizouelm@gmail.com

Abstract— Recent research on bioinspired flapping-wing robots has shown that compliant flapping mechanisms play a crucial role in achieving high aerodynamic efficiency. However, optimally designing compliant mechanisms that can reproduce the complex flapping patterns of birds using only a single actuator remains a significant challenge and constitutes a difficult optimization problem. This paper addresses the issues of computational speed and cost in optimizing such mechanisms, and aims to leverage deep learning to design a compliant mechanism capable of autonomously adjusting its flapping pattern under varying flight conditions. A planar two-segment flapping-wing mechanism with two configurations is introduced and modeled: one with two degrees of freedom and a single linear spring, and another with three degrees of freedom and four springs (including compliant hinges). Both configurations are driven by a single rotational motor. Using multibody simulations, a reference MATLAB/SimMechanics model was employed to generate a dataset of roughly 1,300 samples per configuration. A multilayer perceptron (MLP) neural network was then trained as a fast surrogate model to map design parameters to performance indicators for the target of mimicking birds. Subsequently, surrogate-based optimization was performed using a multi-objective genetic algorithm to identify design parameters capable of reproducing target bird-inspired wingbeat patterns for medium birds operating at 3 and 6 Hz. Validation

simulations demonstrate that the proposed framework can learn the dynamic behavior of the mechanisms under aerodynamic loading across different frequencies and reproduce frequency-dependent patterns unattainable with purely rigid mechanisms. Moreover, the optimization time is reduced from tens of hours to less than a minute. A comparative analysis of the two configurations further shows that increasing mechanical compliance (through additional and varied springs) improves the fidelity of pattern reproduction.

Keywords—Surrogate Model, Multilayer Perceptron (MLP), Flapping Mechanisms, Bio-Inspired Robotics, Data-Driven Modeling, Optimization

Article Code: 1196

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Technical Analyst Attention Network (TAAN): An Interpretable Deep Learning Model for Algorithmic Trading in Crypto and Forex Markets

Ali Tavassolian
Faculty of Computer Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
ali.tavassolian.7@gmail.com

Fatemeh Yousefloei
Faculty of Computer Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
f.yosefloee@mail.sbu.ac.ir

Mojtaba Vahidi Asl
Faculty of Computer Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
mo vahidi@sbu.ac.ir

Monireh Abdoos
Faculty of Computer Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
m abdoos@sbu.ac.i

Abstract—Designing profitable algorithmic trading systems for highly volatile and complex markets, such as Forex and cryptocurrency, remains a significant challenge. While powerful predictors, many contemporary deep learning models lack the intrinsic ability to replicate the analytical logic of an expert trader. A human analyst operates hierarchically: they first evaluate signals across different time periods within a single indicator, and then synthesize insights from the most important indicators for a final decision—a process for which common deep learning architectures are not designed. To address this, we introduce the Technical Analyst Attention Network (TAAN), a novel architecture specifically engineered to emulate this analytical workflow. At its core, TAAN features a custom-designed hierarchical attention mechanism. The first attention layer identifies the most salient signals by assessing multiple time periods within each indicator. The second layer then finalizes the analysis by weighing the importance of the indicators themselves (e.g., market structure vs. momentum). We evaluated this architecture on high-volume Forex (EUR/USD, GBP/USD) and cryptocurrency (BTC/USD, ETH/USD) data. TAAN significantly outperformed benchmark models, achieving an accuracy of 95.43% and a

Matthews Correlation Coefficient (MCC) of 0.8107. Furthermore, a backtest on EUR/USD yielded a 192.16% cumulative return and a 3.6 Sharpe ratio. This superior performance, achieved with a remarkably smaller parameter count, validates our design approach.

Keywords— Algorithmic Trading, Interpretable Deep Learning, Attention Mechanism, Human-Inspired AI,Automated Feature Engineering, Technical Analysis, Forex, CryptocurrencyTrading

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1200

A Spatio-Temporal Deep Reinforcement Learning Framework for Algorithmic Trading Using Genetically Evolved Dynamic Knowledge Graphs

Fatemeh Yousefloie
dept. of Computer Engineering
Shahid Beheshti University
Tehran, Iran
f.yosefloee@mail.sbu.ac.ir

Ali Tavassolian
dept. of Computer Engineering
Shahid Beheshti University
Tehran, Iran
a.tavassolian.7@gmail.com

Mehrnoosh Shamsfard
dept. of Computer Engineering
Shahid Beheshti University
Tehran, Iran
m-shams@sbu.ac.ir

Monireh Abdoos
dept. of Computer Engineering
Shahid Beheshti University
Tehran, Iran
m abdoos@sbu.ac.ir

Abstract—Financial markets dynamic and non-stationary behavior challenges traditional, static algorithmic trading systems. While modern deep learning models offer high predictive power, their Black-box nature creates a critical gap in trust and explainability. To address these challenges, this paper introduces a novel spatio-temporal deep reinforcement learning framework for algorithmic trading. At each time step, our system employs a genetic algorithm to construct a dynamic knowledge graph. This graph serves as a real-time, optimized map of effective technical analysis strategies for the current market state. A sequence of these evolving graphs is then fed into a hybrid neural architecture. This architecture leverages a Graph Attention Network (GAT) to capture the complex relationships within each strategic map and an LSTM to model their temporal dynamics. The resulting rich spatio-temporal representation serves as the state representation for a deep reinforcement learning agent. This agent is trained using the PPO algorithm to learn an optimal policy for executing trades and selecting risk management levels (TP and SL). This approach represents a paradigm shift from static prediction models to an adaptive agent that learns to make decisions based on the evolution of the market's strategic structure.

Experimental results demonstrate that our framework not only achieves superior performance in generating profitable signals but also offers a novel form of hybrid explainability, enabling the analysis of the strategic knowledge graphs underlying the agent's decisions.

Keywords—Algorithmic Trading, Reinforcement Learning, Dynamic Knowledge Graphs, Genetic Algorithms, Explainable AI (XAI)

Article Code: 1201

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Finite element model updating using computational intelligence - based methods: A Case study of theSan Roque Valley bridge, California

Fatemeh Z. Dokhanian
Faculty of Civil, Water and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
f_dokhanian@sbu.ac.ir

Negin G.sabagh
Faculty of Civil, Water and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
n_ghorbansabagh@sbu.ac.ir

*Mahmoud R. Shiravand**
Faculty of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
m_shiravand@sbu.ac.ir

Abstract— In recent years, updating finite element (FE) models has become an effective tool for reducing uncertainties and improving the accuracy of structural numerical models. In this study, a reinforced concrete bridge was selected as a case study, and its FE model was analyzed using OpenSees software and field-measured vibration data. To perform the updating process, computational intelligence (CI) methods and genetic metaheuristic algorithms were used to minimize the difference between the numerical responses and the measured natural frequencies. The results indicate that the genetic algorithm (GA), through its intelligent search operators, can modify essential model parameters, such as the elastic modulus of the deck and piers, as well as the abutment stiffness, enabling the numerical model to accurately replicate the bridge's dynamic behavior. The updated model can then be used to predict the structure's response under different loading conditions and for future seismic analysis, both of which play an important role in the management and maintenance of bridges.

Keywords— Model Updating, Bridge, Genetic Algorithm, Numerical Simulation, metaheuristic algorithm, Computational Intelligence

Article Code: 1204

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Robust Algorithmic Trading in Volatile Markets Using Pessimistic Transformer-Enhanced TD3 in Continuous Action Spaces

Amirhossein Ghozati
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
am.ghozati@mail.sbu.ac.ir

*Armin Salimi-Badr**
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a_salimibadr@sbu.ac.ir

Abstract—Contemporary financial markets, marked by their rapid evolution and inherent volatility, pose substantial challenges to algorithmic trading due to pervasive uncertainty. Traditional approaches frequently prove insufficient, and even advanced deep reinforcement learning (DRL) models may succumb to overestimation biases. In response, we present the Pessimistic Transformer-based TD3 (PT-TD3), an innovative extension of the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm, specifically adapted for continuous action spaces. Our methodology integrates a Transformer encoder to discern long-term patterns and interdependencies among multiple assets, alongside a pessimism mechanism that exploits variances between twin critics to conservatively modulate Q-values, thereby curtailing unwarranted overconfidence in risky decisions. Additionally, we enhance the state representation with technical indicators such as RSI and MACD, while refining the reward function through volatility penalties to promote greater stability. Evaluated on a diversified portfolio of 10 U.S. stocks over the period from 2015 to 2023, PT-TD3 yields a total return of 52.14% (exceeding the benchmark's 46.03%), a Sharpe ratio of 3.63 (versus 2.76), and a maximum drawdown of -5.06% (compared to -7.95%), surpassing competing techniques including LSTM-PPO hybrids and CNN-GRU-Attention frameworks. An ablation study reveals optimal performance at $\alpha = 0.4$, underscoring how this calibration effectively manages uncertainty without compromising potential gains. Ultimately, PT-TD3 furnishes a more resilient framework for traders operating in unpredictable market conditions.

Keywords— Deep Reinforcement Learning, Algorithmic Trading, TD3, Pessimism, Transformer Attention, Multi-Asset Portfolio, Continuous Actions

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1207

Benchmarking Persian Proofreading Tools: A Fine-grained, Error-Typed Evaluation

Zahra Saaberi
Faculty of computer science and engineering
Shahid Beheshti University
Tehran, Iran
z_saaberi@sbu.ac.ir

*Mehrnoush Shamsfard**
Faculty of computer science and engineering
Shahid Beheshti University
Tehran, Iran
m-shams@sbu.ac.ir

Abstract— We present a dedicated benchmark, ParsProof, for systematic evaluation of Persian proofreading systems, featuring fine-grained annotations across 52 error types tailored to the specific linguistic properties of Persian language. Unlike most of the existing datasets, our benchmark enables detailed error-level analysis. We evaluated seven widely used Persian proofreading tools—Virastyar, Negar, Virastman, Farsiyar, Paknevis, Viravirast, and Neviise—as well as GPT- 4.1. Among the specialized proofreading systems, Paknevis achieved the highest F1 score. GPT-4.1, however, generated the largest number of correct edits overall, though it also introduced more unnecessary or erroneous changes. The benchmark highlights the strengths and limitations of current tools at a granular level, providing a robust foundation for developing more accurate and linguistically informed Persian proofreading systems. A publicly available subset of the dataset is provided for research purposes.

Keywords—Proofreading tools, fundamental NLP tools, Persian, benchmarking

Article Code: 1209

Early Prediction of Inhibitory Control Lapses from Unconstrained Gaze Dynamics

Rezvane Mirzapour
Faculty of Electrical Engineering
K. N. Toosi University of Technology
Tehran, Iran
rezvane.mirzapour@email.kntu.ac.ir

Farhane Mirzapour
Faculty of Electrical Engineering
K. N. Toosi University of Technology
Tehran, Iran
farhane.mirzapour@email.kntu.ac.ir

*Mehdi Delrobaei**
Faculty of Electrical Engineering
K. N. Toosi University of Technology
Tehran, Iran
delrobaei@kntu.ac.ir

Abstract—Predicting human cognitive errors from early behavioral signals is vital for responsive human-computer interaction but often depends on specialized eye-trackers, limiting scalability. This study explores whether it is possible to predict inhibitory control lapses early using a standard laptop webcam. We have created a computer vision pipeline that analyzes webcam video to extract gaze coordinates, which are refined and classified into directions (left, center, right). Gaze dynamics features from the first 10-25% of a trial (150-375 ms post-stimulus) were used in an emotional Simon task. A Gradient Boosting classifier, trained via Leave-One-Participant-Out cross-validation, distinguished correct from error trials. The model predicted subsequent response errors with remarkable performance, achieving 97% accuracy and an F1-score of 71%, significantly outperforming baseline models. Findings show that early gaze data reflect inhibitory processing, providing a framework for cognitive monitoring. This supports practical applications such as personalized training and driver safety.

Keywords—Biomechatronic Systems, Affective Computing, Cognitive Errors, Predictive Modeling, Gradient Boosting

Article Code: 1213

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Reliability-Aware Prompting for Diabetic Retinopathy Classification using Vision-Language Models

*Danial Zohourian**
dept. of Computer and
Data Science, Faculty of Mathematical Science
Shahid Beheshti University
Tehran, Iran
danizohourian@gmail.com

Dorsa Asgari
dept. of Computer and
Data Science, Faculty of Mathematical Science
Shahid Beheshti University
Tehran, Iran
d.asgari@mail.sbu.ac.ir

Sadegh Madadi
dept. of Biomedical Engineering
Amirkabir University of Technology
Tehran, Iran
smadadi@aut.ac.ir

Hadi Farahani
dept. of Computer and
Data Science, Faculty of Mathematical Science
Shahid Beheshti University
Tehran, Iran
h_farahani@sbu.ac.ir

Abstract—Diabetic Retinopathy (DR) remains one of the leading causes of preventable blindness, and automated grading of fundus images plays a crucial role in early diagnosis and treatment. Vision–language models (VLMs), such as CLIP and its medical variant FLAIR, have recently shown promise for zero-shot medical image classification without task-specific training. However, their performance is highly sensitive to prompt phrasing, and optimal label formulations vary across disease stages. In this study, we evaluate the FLAIR model on the APTOS dataset for both binary and five-class DR classification. After establishing a strong baseline, we introduce Reliability-Aware Prompting (RAP), a lightweight method that identifies prompt sets demonstrating high precision for specific DR stages on a training split and selectively applies them at inference time. Without modifying model parameters, RAP improves classification reliability by leveraging class-specific prompt trustworthiness. The results highlight the importance of prompt selection in medical VLMs and suggest a simple, training-free path to more dependable diagnostic performance.

Keywords—diabetic retinopathy, artificial intelligence, visionlanguage models, prompt ensemble

Article Code: 1214

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Comparative Assessment of Process-Based and Deep Learning Models for Runoff Simulation: A Case Study of the Zayandehrood River Basin

Mohamad Saeed Zarkhan
Phd student, Department of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
m_zarkhan@sbu.ac.ir

*Azadeh Ahamdi**
Associate Professor, Department of Civil, Water, and Environmental Engineering
Shahid Beheshti University
Tehran, Iran
az_ahmadi@sbu.ac.ir

Abstract—Efficient water resource governance necessitates high-precision hydrological simulations. This research evaluates and compares the runoff modeling capabilities of the process-based Soil and Water Assessment Tool (SWAT) alongside two deep learning (DL) frameworks: Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN). The predictive performance of these models was analyzed at the Eskandari and Ghalehshahrokh gauging stations, employing five statistical indices: Nash-Sutcliffe Efficiency (NSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Correlation Coefficient (CC), and Percent Bias (PBIAS). The training phase results indicated a superior fit for the DL models, which achieved NSE values between 0.79 and 0.91, outperforming SWAT's range of 0.71 to 0.78. However, the DL models experienced a significant reduction in accuracy during the testing interval, with NSE values dropping to 0.49–0.57. Specifically, the CNN architecture yielded a PBIAS of -21.23, signaling a notable underestimation of runoff. In contrast, the SWAT model demonstrated superior consistency and robustness during the validation stage, sustaining an NSE of 0.71. This suggests that SWAT possesses better generalizability and remains a more dependable tool for runoff simulation in this basin.

Keywords—Runoff Simulation, SWAT, LSTM, CNN, Upstream Zayandeh-Rood Basin.

Article Code: 1218

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

I²KAN: Interpretable Interaction-aware Kolmogorov-Arnold Networks for Time-Series Anomaly Detection

Amirhossein Sadr
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a.sadr@mail.sbu.ac.ir

*Mahboobeh Riahi-Madvar**
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
m_riahi@sbu.ac.ir

Abstract— Time Series Anomaly Detection (TSAD) is critical for the reliability of modern IT infrastructure, cloud services, and industrial systems. While deep learning models have advanced forecasting-based TSAD, they often overfit to local fluctuations and lack interpretability, limiting their trustworthiness in safety-critical applications. This paper introduces I²KAN, a novel framework that leverages Kolmogorov-Arnold Networks (KAN) to address these dual challenges. I²KAN enhances KAN's architecture with two key strategies: an interaction-aware feature weighting mechanism that dynamically quantifies and incorporates the multivariate dependencies within time series data, and an interpretability pipeline that provides transparent rationales for its predictions. By moving beyond the black-box nature of traditional deep models, I²KAN not only achieves high detection accuracy but also offers crucial insights into the underlying anomaly patterns. Extensive experiments on the ECG5000 and UCR benchmarks demonstrate that I²KAN achieves state-of-the-art performance, outperforming MLP, KAN, and KAN-AD models in Accuracy, ROC-AUC, and F1 score. Furthermore, we employ SHAP analysis to validate and visualize the model's decision-making process, reinforcing its credibility for real-world deployment.

Keywords— Time Series Anomaly Detection, KolmogorovArnold Networks, Interpretability, Multivariate Interaction, Deep Learning

Article Code: 1219

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

CLIP-Joint-Detect: End-to-End Joint Training of Object Detectors with Contrastive Vision-Language Supervision

*Behnam Raoufi**
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
behnam.raoufi93@sharif.edu

Hossein Sharify
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
sharifir.hossein@ee.sharif.edu

Mohamad Mahdee Ramezanee
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
ramezani.mm@ee.sharif.edu

Khosrow Hajsadeghi
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
ksadeghi@sharif.edu

Saeed Bagheri Shouraki
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
bagheri-s@sharif.edu

***Abstract*—**Conventional object detectors rely on cross-entropy classification, which can be vulnerable to class imbalance and label noise. We propose CLIP-Joint-Detect, a simple and detector-agnostic framework that integrates CLIP-style contrastive vision-language supervision through end-to-end joint training. A lightweight parallel head projects region or grid features into the CLIP embedding space and aligns them with learnable class-specific text embeddings via InfoNCE contrastive loss and an auxiliary loss term, while all standard detection losses are optimized simultaneously. The approach applies seamlessly to both two-stage and one-stage architectures. We validate it on Pascal VOC 2007+2012 using Faster R-CNN and on the large-scale MS COCO 2017 benchmark using modern YOLO detectors (YOLOv11), achieving consistent and substantial improvements while preserving real-time inference speed. Extensive experiments and ablations demonstrate that joint optimization with learnable text embeddings markedly enhances closed-set detection performance across diverse architectures and datasets.

***Keywords*—** Object Detection, CLIP, Faster R-CNN, YOLO

Article Code: 1227

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Towards Automated Self-Aware System Design: Process-Based Models to MAPE-K Architecture

Maryam Samsam Sokhiravi
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
m.samsamsokhiravi@sbu.ac.ir

Seyed-Hosein Attarzadeh-Niaki
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
h attarzadeh@sbu.ac.ir

Abstract—We present a systematic design flow to connect abstract behavioral models of self-aware cyber-physical systems to their executable realization. We use an Abstract Process-based Self-aware Model (APSA) and concretely capture it as a Scenario-Aware Dataflow (SADF) model and guarantee timing and resource bounds. The functional components of APSA are then systematically mapped to the roles of the MAPE-K reference architecture and implemented using the MAPE Specification Language (MSL), which automatically generates executable Abstract State Machines (ASMs). A smart-home comfort-control case study demonstrates that the resulting self-adaptive system remains fully traceable to its high-level behavioral model.

Keywords—Self-Awareness, Cyber-Physical Systems, MAPE–K, Scenario-Aware Dataflow (SADF), Formal System Design, Design Flow, Smart Home.

Article Code: 1228

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Ensemble-Guided Distillation for Compact and Robust Acoustic Scene Classification on Edge Devices

*Hossein Sharify**
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
sharifir.hossein@ee.sharif.edu

Behnam Raoufi
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
behnam.raoufi93@sharif.edu

Mahdy Ramezani
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
ramezani.mm@ee.sharif.edu

Khosrow Hajsadeghi
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
ksadeghi@sharif.edu

Saeed Bagheri Shouraki
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
bagheri-s@sharif.edu

Abstract— We present a compact, quantization-ready acoustic scene classification (ASC) framework that couples an efficient student network with a learned teacher ensemble and knowledge distillation. The student backbone uses stacked depthwise-separable "expand–depthwise–project" blocks with global response normalization to stabilize training and improve robustness to device and noise variability, while a global pooling head yields class logits for efficient edge inference. To inject richer inductive bias, we assemble a diverse set of teacher models and learn two complementary fusion heads: z1, which predicts perteacher mixture weights using a student-style backbone, and z2, a lightweight MLP that performs per-class logit fusion. The student is distilled from the ensemble via temperature-scaled soft targets combined with hard labels, enabling it to approximate the ensemble's decision geometry with a single compact model. Evaluated on the TAU Urban Acoustic Scenes 2022 Mobile benchmark, our approach achieves state-of-the-art (SOTA) results on the TAU dataset under matched edge-deployment constraints, demonstrating strong performance and practicality for mobile ASC.

Keywords—acoustic scene classification,knowledge distillation, teacher ensemble, model compression

Article Code: 1231

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

A Methodological Framework for Assessing Language Models for Security Applications

Mostafa Rahimi
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
mos_rahimi@sbu.ac.ir

Maghsoud Abbaspour
Faculty of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
maghsoud@sbu.ac.ir

Abstract—The absence of standardized and reliable benchmarks poses a major challenge in selecting the most suitable large language model (LLM) for various security-related tasks. This limitation arises from the diversity and specificity of such tasks, the mismatch between benchmark metrics and real-world infrastructure requirements, and the limited maturity and adoption of newly proposed benchmarks. Moreover, discovering newly vulnerabilities in source code remains a slow process, as it demands up-to-date security expertise, deep programming knowledge, and extensive code analysis across multiple execution paths. To address these constraints, large language models can be leveraged as intelligent assistants capable of accelerating vulnerability analysis and detection. These models can aid security professionals by (1) understanding vulnerabilities found in source code and mitigate it through retrieval-augmented architectures, (2) performing intelligent static code reviews to identify all possible data and control flows without execution, and (3) generating complex test cases to facilitate inter-function transitions. Specifically, we propose a process-oriented, scenario-driven framework to assess the capabilities of LLMs in security-related tasks and systematically identify the most effective models for real-world software security.

Keywords—security task, secure code generation, large language models, evaluation framework

Article Code: 1232

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Stacked Interaction Module for Refining Subject-Object Dynamics in Video Captioning

Mahdi Khouri-shandiz

School of Automotive Engineering, Iran University of Science and Technology

Tehran 16846-13114, Iran

mahdi_khoorishandiz@auto.iust.ac.ir

Abdollah Amirkhani

School of Automotive Engineering, Iran University of Science and Technology

School of Computer Engineering, Iran University of Science and Technology

Tehran 16846-13114, Iran

amirkhani@iust.ac.ir

Abstract—This Current video captioning models fail to grasp complex Subject-Object interactions, yielding 'entangled' representations where roles are blurred. We introduce SIMCap, a novel hierarchical architecture that disentangles these interactions by explicitly separating roles. Instead of relying on complex external graphs, SIMCap leverages an innovative Stacked Interaction Module (SIM). This module empowers Subjects to actively query a rich memory bank of Objects and video context. This process is stacked iteratively, allowing the model to progressively refine its understanding of the interaction. By intrinsically modeling relationships, SIMCap achieves State-of-the-Art results on the standard MSRVTT and MSVD benchmarks, demonstrating decisive superiority, particularly in the CIDEr semantic accuracy metric.

Keywords—Video captioning, Interaction, hierarchical, Subject-Object.

Poster Papers

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

POSTER PAPER PRESENTATION PROGRAM

SECOND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE

	TIME	TOPIC	PAPER IDS
SUNDAY, 31 MAY	9–10	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN OTHER FIELDS OF ENGINEERING	iai-1166 iai-1182 iai-1221 iai-1191 iai-1189 iai-1198
	10–11	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN HEALTH AND HYGIENE	iai-1203 iai-1186 iai-1215 iai-1056 iai-1195 iai-1151
	11–12	APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN GOVERNANCE AND THE HUMANITIES AND SOCIAL SCIENCES, WEALTH CREATION AND MANAGEMENT USING ARTIFICIAL INTELLIGENCE	iai-1008 iai-1174 iai-1003 iai-1081 iai-1168
	13–14	DATA MINING AND BIG DATA	iai-1043 iai-1194 iai-1028 iai-1217 iai-1080 iai-1197
	14–15	NATURAL LANGUAGE PROCESSING	iai-1128 iai-1089 iai-1054 iai-1211 iai-1225 iai-1160
MONDAY, 1 JUNE	9–10	BIOINFORMATICS, KNOWLEDGE ENGINEERING AND KNOWLEDGE-BASED SYSTEMS, COGNITIVE AND COMPUTATIONAL NEUROSCIENCE	iai-1046 iai-1108 iai-1038 iai-1139
	10–11	IMAGE PROCESSING AND MACHINE VISION	iai-1124 iai-1212 iai-1180 iai-1134 iai-1205
	11–12	IMAGE PROCESSING AND MACHINE VISION, THEORETICAL TOPICS IN MACHINE LEARNING	iai-1233 iai-1224 iai-1138 iai-1193

Article Code: 1003

WealthAI: An End-to-End Framework for Explainable and Personalized AI-Driven Wealth Management

*Mohammadali Soltanshahi**
AI & Smart Governance Unit
SKAI
Tehran, Iran
ali.soltanshahi@gmail.com

Amirhossein Kafi
Smart Banking Research Unit
SKAI
Tehran, Iran
amirhkafi@gmail.com

Mojtaba Raeessafari
Smart Banking Policy Unit
SKAI
Tehran, Iran
mgrd81@gmail.com

***Abstract*—**The convergence of deep reinforcement learning, large language models, and federated learning is redefining the landscape of wealth management. This paper introduces WealthAI, a novel, end-to-end framework that integrates these technologies to deliver personalized, explainable, and privacy-preserving investment strategies. Our architecture features a multi-agent deep reinforcement learning (DRL) core for dynamic portfolio optimization, an explainable AI (XAI) module that generates human-readable rationales for every decision, and a federated learning protocol that ensures client data privacy. We evaluate WealthAI on a comprehensive dataset spanning 14 years (2010–2024) across equities, cryptocurrencies, and commodities. Results demonstrate a 22.7% annualized return with a Sharpe ratio of 1.68, significantly outperforming traditional benchmarks. Crucially, our XAI module increases user trust by 67% in user studies, while federated learning reduces data leakage risk to near zero. This work provides a practical, ethical, and high-performance blueprint for the future of AI in finance.

***Keywords*—**Artificial Intelligence, Wealth Management, Deep Reinforcement Learning, Explainable AI, Large Language Models, Federated Learning, Portfolio Optimization, Algorithmic Trading, Robo-Advisors, Financial Ethics.

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1008

Artificial Intelligence-Driven Wealth Creation and Management

*Farzad Baneshi**

Department of Management,

Shiraz University

Shiraz, Iran

farzadanesh99@gmail.com

Mahsa Nadimpour

Department of Basic Sciences,

Shahid Chamran University

Ahvaz, Iran

nadimpour.mahsa@gmail.com

Abstract—The integration of Artificial Intelligence (AI) into wealth creation and management has redefined modern financial ecosystems by enhancing decision-making, automation, and predictive accuracy. AI-driven models, such as Machine Learning (ML) and Natural Language Processing (NLP), enable investors and financial institutions to analyze large-scale market data, identify emerging trends, and optimize asset allocation strategies. Through algorithmic trading, robo-advisors, and intelligent portfolio management, AI reduces human error and improves risk assessment. Moreover, AI-based systems facilitate personalized investment solutions, increasing efficiency and accessibility in financial services. However, challenges related to data privacy, transparency, and regulatory frameworks remain critical for sustainable adoption. This paper explores the transformative potential of AI in wealth generation and management, emphasizing both opportunities and ethical considerations in the evolving digital economy.

Keywords—Artificial Intelligence, Machine Learning, Wealth Management, Algorithmic Trading, Financial Technology, Predictive Analytics

Article Code: 1028

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Analyzing Stack Overflow Question Types and Answer Characteristics: Implications for Designing and Benchmarking Coding Assistant LLMs

*Omid Mohammadi Kia**

Shahid Beheshti Faculty of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

o_mohammadikia@sbu.ac.ir

Mahmood Neshati

Shahid Beheshti Faculty of Computer Science and Engineering

Shahid Beheshti University

Tehran, Iran

m_neshati@sbu.ac.ir

Abstract—Large language models (LLMs) like GitHub Copilot excel as coding assistants, yet struggle with diverse developer queries; Stack Overflow (SO), a key Q&A hub, features conceptual, debugging, and best practice questions, but optimal answer formats—length, code, images, references— remain underexplored, hindering LLM design and benchmarking. We analyze 50,000 high-engagement SO questions from 2020 to 2024, classifying them via hybrid annotation and fine-tuned RoBERTa with high accuracy and 92 percent inter-annotator agreement. Results show: debugging questions, 42 percent of the sample, favor short code-rich answers, 78 percent under 150 words with 91 percent including code and strong acceptance correlation; conceptual queries, 31 percent, prefer long explanations with references, 65 percent over 400 words and 82 percent referenced; best practices, 27 percent, balance multimodal responses with 72 percent code, 56 percent images, and 71 percent references, with associations statistically significant. We propose adaptive LLM prompting and stratified benchmarks, We release an annotated dataset on GitHub to enhance context-aware coding aids.

Keywords—Stack Overflow analysis, Question classification, Answer characteristics, Coding assistants, Large Language Models (LLMs), Benchmarking

Article Code: 1038

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

ChatGPT 4 and personality prediction

Zahra Eslami
Department of Biomedical engineering
Shahid Beheshti University
Tehran, I.R. Iran
zaahraaeslami2@gmail.com

Marcus Cheetham
Department of Internal Medicine
University Hospital Zurich
Zurich, Switzerland
Marcus.Cheetham@usz.ch

Seyed Abolfazl Valizadeh
Department of Biomedical engineering
Shahid Beheshti University
Tehran, I.R. Iran
s_valizadeh@sbu.ac.ir

***Abstract*—**Large Language Models (LLMs) such as ChatGPT-4 have been successfully applied to infer the Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism) from general text. However, their use for personality inference from LLM user interaction text (e.g., queries and retrieved information) remains underexplored. We hypothesize that conscientiousness acts as a moderator, increasing the predictive accuracy of LLM-inferred traits due to the linguistic alignment between characteristic conscientious text patterns and LLM training data. Participants (N=36) completed a validated self-report of the Big Five and used ChatGPT-4 for over 12 weeks. ChatGPT-4 was used to generate personality inferences from their text inputs. Correlational and moderation analyses revealed that the correlation between self-reported and LLM-inferred traits was significantly stronger for individuals with higher self-reported conscientiousness. This effect was most notable for the trait of conscientiousness itself, with modest, consistent moderation effects observed across the other Big Five traits. Predictive accuracy remained moderate overall, aligning with existing computational personality research. These findings suggest that the linguistic consistency of conscientious individuals facilitates more accurate LLM-based personality inference, emphasizing the need for improved model calibration and input assessment.

***Keywords*—**GPT4, Personality prediction, prompts styling

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1043

Customer Segmentation in Online Tire Sales Using RFM and Quantity: A Comparison of K-Means, MiniBatchKMeans, and DBSCAN with Quality Improvement via Density-Based Refinement

*Seyed Mohammadreza Jalalian Shahri**
M.Sc.in Telecommunication Eng.
Smart Systems Department
Kavir Tire Company
Birjand, Iran
mr.jalalian.ieee@gmail.com

Seyed Hashem Mohtashami
M.Sc. in Artificial Intelligence
Smart Systems Department
Kavir Tire Company
Birjand, Iran
mohtashami.ai@gmail.com

Abstract—Customer segmentation remains a critical challenge in e-commerce, where diverse purchasing behaviors demand robust analytical frameworks. This study introduces a practical approach based on five years of real online sales data from a tire manufacturing company. Behavioral features were derived from the RFM–Qty model (recency, frequency, monetary value, average order amount, and purchase quantity). Three clustering algorithms—K-Means, MiniBatchK-Means, and DBSCAN—were systematically compared. Internal evaluation metrics indicated k = 8 as the optimal cluster number. DBSCAN was subsequently applied as a refinement stage to filter out approximately 3.28% of noisy data points, after which reapplying K-Means on the cleaned dataset led to significant improvements (the silhouette coefficient increased from 0.3876 to 0.3967, and the Davies–Bouldin index decreased from 0.9182 to 0.9022). The final solution yielded eight actionable customer segments with clear behavioral profiles and managerial implications, such as loyalty programs for "Champions," reactivation campaigns for "At-risk" customers, and strategies to convert "One-time buyers" into repeat purchasers. A two-dimensional PCA visualization capturing about 82% of the variance was employed to illustrate cluster separation. The findings highlight that incorporating density-based noise filtering prior to clustering enhances segmentation effectiveness and supports more targeted marketing decisions in real-world ecommerce environments. This work contributes to both academic research on customer analytics and the practical design of data-driven marketing strategies

Keywords— Online sales, customer clustering, RFM, K-Means, MiniBatchK-Means, DBSCAN, HDBSCAN, PCA, cluster evaluation

Article Code: 1046

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Predicting Plasma Protein Binding (Fraction Unbound) with Machine Learning Using Molecular Descriptors

Arash Maghsoudlou

Department of Pharmaceutics and Pharmaceutical Nanotechnology

School of Pharmacy, Shahid Beheshti University of Medical Sciences

Tehran, Iran

arash_maghsoudlou@sbmu.ac.ir

*Fatemeh Ghorbani-Bidkorpeh**

Department of Pharmaceutics and Pharmaceutical Nanotechnology

School of Pharmacy, Shahid Beheshti University of Medical Sciences

Tehran, Iran

f.ghorbani@sbmu.ac.ir

M. Soltani

Department of Electrical & Computer Engineering

University of Waterloo

Waterloo, Ontario, Canada

msoltani@uwaterloo.ca

Abstract—The unbound fraction in plasma (fu) is an important pharmacokinetic parameter that directly affects drug distribution, clearance, and efficacy in vivo. While experimentally determining fu is costly and labour-intensive, developing reliable methods for its prediction is highly desirable. In this paper, we propose the use of machine learning techniques to estimate fu, using both physicochemical properties of compounds and in vitro-derived molecular fingerprints from the Krumpholz et al. dataset. The Gradient Boosting model achieved the highest predictive performance, resulting in an R^2 score of about 0.61, a root mean squared error of about 0.19, and a mean absolute error of about 0.15. By comparison, the Random Forest model achieved an R2 value of about 0.53. Lipophilicity/polarity-related descriptors were identified as the most important predictors of fu, indicating that fu is driven by these properties.

Keywords—plasma protein binding, fraction unbound, machine learning, molecular descriptors

Article Code: 1054

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Claim Verification in Persian: A Hybrid Pipeline with LLMs and Rationale-Quality Classification

Sara Bourbour *
Information Technology department
Tarbiat Modares University
Tehran, Iran
s.bourbour@modares.ac.ir

MohammadAli SeifKashani
Computer engineering department
Sharif University of Technology
Tehran, Iran
ma.seifkashani@ce.sharif.edu

Fateme Aali
Information Technology department
Tarbiat Modares University
Tehran, Iran
f.alee@modares.ac.ir

Abstract—We present a Persian claim-verification pipeline that combines a fast deep stance classifier with an LLM-based verifier and a rationale-quality filter. Our corpus mixes private and public news articles paired with social-media posts. Annotation proceeds in two passes: (i) human stance labels (support, refute, unrelated) on post–article pairs; (ii) for pairs initially labeled as negative (refute/unrelated), an instructionfollowing LLM produces short Persian rationales which are then audited by humans for validity. Invalid rationales expose systematic failure modes and are used to augment training via class-conditional oversampling and rationale-aware multi-task learning. We train LaBSE- and ParsBERT-based classifiers and calibrate probabilities by temperature scaling. On a confidential test set, oversampling the negative class yields the strongest evidence classification for LaBSE (Accuracy 0.71, F1 0.80, Precision 0.68, Recall 0.97), while GPT-4o provides higher recall but lower precision among LLM baselines. The rationale-quality head improves the reliability of LLM-driven decisions by gating low-quality explanations and stabilizes paragraph aggregation. We discuss deployment-oriented choices— probability calibration, paragraph-level scoring, and human-inthe-loop validation—that increase trustworthiness without disclosing sensitive dataset details.

Keywords—claim verification, fakenews, rationale quality, stance detection.

Article Code: 1056

An interpretable framework based on deep learning and the Internet of Things for predicting risk dynamics in chronic patients

Behnaz Pouriayevali

PhD Student, Department of Management and Health Information Technology, School of Management and Medical Information Sciences

Isfahan University of Medical Sciences

Isfahan, Iran.

Behnaz.pouriayevali@ gmail.com

Dr Asghar Ehteshami

Associate Professor, Department of Management and Health Information Technology, Health Information Technology Research Center

Isfahan University of Medical Sciences

Isfahan, Iran

ehteshami@mng.mui.ac.ir

Abstract— Introduction: Effective management of chronic conditions necessitates analytical approaches capable of capturing temporal dynamics within physiological data streams. Conventional machine learning methodologies frequently encounter limitations in this domain, functioning as opaque "black boxes" that obscure decision-making processes. Objective: This investigation introduces an interpretable framework integrating Internet of Things (IoT) infrastructure with deep learning architectures to forecast health deterioration risks among individuals with chronic conditions. Methods: Temporal data were obtained from the established PhysioNet repository. A long short-term memory (LSTM) network was developed and optimized to anticipate critical health events. Model performance was evaluated against four baseline approaches (artificial neural network and random forest). Furthermore, the SHAP (SHapley Additive exPlanations) methodology was employed to elucidate model predictions and identify predominant risk determinants. Results: The LSTM architecture demonstrated superior predictive capability relative to baseline models, attaining 95% accuracy with an area under the ROC curve (AUC) of 0.97. SHAP-based analysis identified

diminished sleep quality and abrupt alterations in physical activity patterns as the most significant predictors of patient risk. Conclusion: The developed framework not only enhances predictive precision but also fosters confidence in and adoption of intelligent health systems through clinically relevant explanatory capabilities.

Keywords: Internet of Things, deep learning, LSTM, E-health, Explainable Artificial Intelligence (XAI), chronic disease management, predictive analytics.

Article Code: 1080

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

A Comprehensive Review of Big Data Analytics Approaches in Social Networks Using Graph Neural Networks: Methods, Challenges, and Future Directions

Mohammad Javad Khedmati Yavand

Faculty and Research Institute of Artificial Intelligence and Cognitive Sciences

Imam Hussein (AS) Comprehensive University

Tehran, Iran

mjkhedmati@ihu.ac.ir

Hussein Saberi

Faculty and Research Institute of Artificial Intelligence and Cognitive Sciences

Imam Hussein (AS) Comprehensive University

Tehran, Iran

hsaberi@ihu.ac.ir

Mohammad Ali javadzade

Faculty and Research Institute of Artificial Intelligence and Cognitive Sciences

Imam Hussein (AS) Comprehensive University

Tehran, Iran

javadzade @ihu.ac.ir

Abstract— The rapid growth of social media platforms generates massive amounts of complex and interdependent data that traditional machine learning methods cannot analyze effectively. Because this type of data is non-Euclidean and relational in nature, it requires more advanced models that can directly exploit graph structures. GNNs have emerged as strong tools in extracting meaningful structural and behavioral patterns within large-scale social networks by leveraging message-passing and neighborhood aggregation mechanisms. This study presents a systematic and comprehensive review of foundational concepts, major GNN architectures, and their principal applications in social network big data analytics. It also presents a detailed, quantitative comparison of state-of-the-art methods for fundamental tasks such as node classification, link prediction, community detection, sentiment analysis, and fake news detection. In addition, the study investigates some critical challenges faced, such as scalability, noise and incomplete data, privacy issues, and limited generalizability, and discusses their algorithmic implications. Finally, the work points out some emerging research directions: the integration of GNNs with Large Language Models, dynamic and temporal graph learning, federated graph learning, and improved model explainability. This review

aims to present researchers with a consolidated updated view of the advances, limitations, and promising future directions of GNN-based analysis in large-scale social networks.

Keywords— Graph Neural Networks, Big Data, Social Networks, Graph Analytics, Deep Learning, Natural Language Processing, Model Comparison, Algorithmic Challenges.

Article Code: 1081

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Beyond the Victors: Correcting Survivorship Bias in Organizational Analytics with Explainable AI

*Seyed Mohsen Tofighi**

PhD. Candidate Department of Computer Engineering

Islamic Azad University, North Tehran Branch

Tehran, Iran

seyedmohsen.tofighi@iau.ac.ir

Hasan Pouladi

PhD. Candidate Department of Management and Economic

Islamic Azad University, Science and Research Brancch

Tehran, Iran

h.pouladi@iau.ac.ir

Abstract— Organizational analytics frequently falls prey to survivorship bias, a cognitive pitfall where analysis focuses on successful entities ("survivors") while neglecting unsuccessful ones. This practice yields distorted insights and flawed managerial strategies, particularly in high-turnover sectors such as the insurance industry. This paper makes a conceptual and methodological contribution by introducing the Analytical Debiasing Framework (ADF), which reframes Explainable Artificial Intelligence (XAI) as a cognitive debiasing mechanism rather than a post-hoc interpretability tool.Using a dataset of 1,000 insurance agents (2020–2024), encompassing both retained and departed employees, we employ an XGBoost model enhanced with SHAP (SHapley Additive exPlanations). Results demonstrate that models trained exclusively on "survivors" systematically mask critical performance drivers and exhibit reduced predictive power. In contrast, a holistic model incorporating all agents achieves superior accuracy (AUC = 0.91) and reveals the central role of organizational support mechanisms—factors largely invisible in survivor-only datasets.By making the reasons behind predictions actionable and cognitively meaningful, the ADF enables double-loop organizational learning, challenging entrenched managerial mental models and advancing fairness-aware, explanationdriven decision-making.

Keywords—Survivorship Bias, Explainable AI (XAI), Organizational Learning, Algorithmic Fairness, Insurance Industry.

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1089

Implementation of a Retrieval-Augmented Generation-Based Question Answering System Using Large Language Models

*Mohammad Mahdi HajiRamezanAli**
Shahab Danesh University
Qom, Pardisan, Iran
hajiramezanali@shdu.ac.ir

Mohammad Mehdi Homayounpour
Shahab Danesh University
Qom, Pardisan, Iran
homayoun@shdu.ac.ir

Abstract— Question-answering (QA) systems, as a major subfield of natural language processing, play a crucial role in the development of search engines, virtual assistants, customer service platforms, and educational tools. This requirement is especially important for Persian, a language that faces limitations in training data and computational resources; In this study, we designed and implemented a hybrid framework based on large language models (LLMs) and retrieval-augmented methods. First, the architecture of the QA system was examined using several LLMs and their performances were compared. Next, the system's ability to respond multi-step and complex questions was evaluated to assess its effectiveness under more realistic conditions. To increase system reliability, we developed a mechanism for identifying irrelevant (out-of-domain) queries that leverages model-generated data and a lightweight neural network. Experimental results indicate that "Qwen2-7BInstruct", when combined with retrieval augmentation and ranker models, achieved the best performance among the models considered. Furthermore, the proposed out-of-domain detection mechanism substantially improved the system's performance outperforming humans by 9.51% in accuracy metric and prevented generation of irrelevant responses.

Keywords—Retrieval augmented generation, RAG, chatbot, large language models, out of domain detection, out-of-scope detection

Article Code: 1108

Enhancing Single-Cell Transcriptomic Quality in Gastric Cancer: An Intersective Approach Combining Adaptive and Manual Quality Control

Fatemeh Khalaji-Pirbalouti

MSc student of Genetics, Department of Genetics,

Faculty of Basic Sciences

Shahrekord University

Shahrekord, Iran

fkhalaji078@gmail.com

*Modjtaba Emadi-Baygi**

Associate Professor of Genetics, Department of Genetics,

Faculty of Basic Sciences

Shahrekord University

Shahrekord, Iran

emadi-m@sku.ac.ir

Abstract— Single-cell RNA sequencing (scRNA-seq) has emerged as an indispensable tool for dissecting cellular heterogeneity in complex malignancies such as gastric cancer (GC). The presence of technical artifacts, including ambient RNA and stressed or dying cells, necessitate stringent quality control (QC). We present a robust QC framework, implemented in the Python environment using Scanpy package, and applied to publicly available GC scRNA-seq dataset from the Gene Expression Omnibus (GEO) database. Our approach intersects adaptive QC, based on the median absolute deviation (MAD) for objective, outlier-based filtering, with manual QC, where thresholds are visually determined via histogram and violin plot inspection to remove low-quality cells. Specifically, we targeted three critical metrics: the number of detected genes, total unique molecular identifier counts (UMI), and the fraction of counts from mitochondrial genes per cell. Here, we profiled the transcriptomes of 134,367 cells from 28 patients with gastric cancer. By intersecting the set of cells removed by both methods, we derive a refined, high-quality cell population of 116,666 cells. This intersective methodology

provides a flexible yet robust approach to mitigate the impact of technical noise, resulting in a cleaner count matrix optimized for downstream analysis of the GC tumor microenvironment. In conclusion, our intersective QC framework delivers a reproducible, high-quality gastric cancer scRNA-seq dataset by comprehensively mitigating technical artifacts. This dual-filtering strategy not only enhances the fidelity of cellular heterogeneity analyses but also establishes a robust foundation for uncovering molecular mechanisms within complex tumor microenvironments.

Keywords—scRNA-seq, Quality control, Gastric cancer, Bioinformatics, Expression matrix

Article Code: 1124

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

DRL-Guided Neural Batch Sampling for Semi-Supervised Pixel-Level Anomaly Detection

Amirhossein Khadivi Noghredeh
School of Mathemathics, Statitstics and Computer Science
College of Science, University of Tehran
Tehran, Iran
khadivi.ahn@gmail.com

Abdollah Safari
School of Mathematics, Statistics and Computer Science
College of Science, University of Tehran
Tehran, Iran
a.safari@ut.ac.ir

*Fatemeh Ziaeetabar**
School of Mathematics, Statistics and Computer Science
College of Science, University of Tehran
Tehran, Iran
fziaeetabar@ut.ac.ir

Firoozeh Haghighi
School of Mathematics, Statistics and Computer Science
College of Science, University of Tehran
Tehran, Iran
fhaghighi@ut.ac.ir

Abstract— Anomaly detection in industrial visual inspection is challenging due to the scarcity of defective samples. Most existing methods rely on unsupervised reconstruction using only normal data, often resulting in overfitting and poor detection of subtle defects. We propose a semi-supervised deep reinforcement learning framework that integrates a neural batch sampler, an autoencoder, and a predictor. The RL-based sampler adaptively selects informative patches by balancing exploration and exploitation through a composite reward. The autoencoder generates loss profiles highlighting abnormal regions, while the predictor performs segmentation in the lossprofile space. This interaction enables the system to effectively learn both normal and defective patterns with limited labeled data. Experiments on the MVTec AD dataset demonstrate that our method achieves higher accuracy and better localization of subtle anomalies than recent state-of-the-art approaches while maintaining low complexity, yielding an average improvement of 0.18 in $\boldsymbol{F1max}$ and 0.08 in AUC, with a maximum gain of 0.28 in $\boldsymbol{F1max}$ in the best case.

Keywords—Anomaly detection, semi-supervised learning, REINFORCE, neural batch sampler, MVTec AD

Article Code: 1128

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

From RLHF to DPO and Beyond: A Review of Preference Optimization Methods for Aligning Large Language Models

Poorya Saneei
Department of Computer Engineering
Iran University of Science and Technology
Tehran, Iran
poorya_saneei@iust.ac.ir

Parham Rahimi
Department of Computer Engineering
Iran University of Science and Technology,
NAJM
Tehran, Iran
pa_rahimi@alumni.iust.ac.ir

*Behrouz Minaei-Bidgoli**
Department of Computer Engineering
Iran University of Science and Technology
Tehran, Iran
b_minaei@iust.ac.ir

Abstract— Large Language Models (LLMs) are not aligned with human values by default, and this challenge has fostered the development of various alignment methods to make models helpful, honest, and harmless. One of the first practical methods for solving this problem was Reinforcement Learning from Human Feedback (RLHF), but its complexity, instability, and high computational cost have prevented it from large-scale application. Since then, many alignment techniques have been suggested, but a consistent review of their evolution is missing. This paper traces the evolution of the alignment methods, focusing on how it evolved from RLHF to the family of algorithms called Direct Preference Optimization, or simply DPO. It elegantly reparametrizes the alignment problem into a supervised learning problem. Further, we consider how the new generation of algorithms overcame the limitations of DPO along three major axes: better computational efficiency by removing the reference model, flexibility of formats and objectives, and optimization for complex goals such as creativity. This paper provides an overview of how alignment research has moved from complexity to simple, efficient, and modular algorithms and shows what the future of alignment will be like.

Keywords—Large Language Models, LLM Alignment, Preference Optimization, Reinforcement Learning from Human Feedback (RLHF), Direct Preference Optimization (DPO), Reward Model, Reference-Free Algorithms.

Article Code: 1134

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Development of 3D Neural Cellular Automata for Learning Spatio-Temporal Patterns

*Yashar Rezazadeh Shahir**
Computer Engineering
Shahid Rajaee Teacher Training University
Tehran, Iran
yasharrezazadehshahir@gmail.com

Akram Beigi
Computer Engineering
Shahid Rajaee Teacher Training University
Tehran, Iran
akrambeigi@sru.ac.ir

Abstract— Neural Cellular Automata (NCA) have emerged as a powerful framework that integrates machine learning with mechanistic modeling to learn and reproduce complex spatiotemporal dynamics. Despite their promise, existing studies have largely focused on two-dimensional formulations, limiting their applicability to inherently volumetric phenomena. In this work, we introduce a three-dimensional extension of Neural cellular automata capable of learning, generating, and sustaining complex structures in 3D environments. Building upon the architecture of Richardson et al. (2024) in Learning Spatio-Temporal Patterns with Neural Cellular Automata, our model enables efficient training and robust generalization across a variety of volumetric patterns. Experimental results demonstrate that 3D NCA can synthesize intricate structures while maintaining computational efficiency. These findings broaden the potential of NCA-based systems and offer new opportunities in domains such as computational biology, medical image analysis, materials science, and artificial life.

Keywords—Neural Cellular Automata, 3D Modeling, Evolutionary Computation, Machine Learning, Spatio-Temporal Pattern Processing

Article Code: 1138

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

AdaKD-LoRA: Adaptive Knowledge Distillation with Low-Rank Adaptation for Rehearsal Free Continual Learning

*Mohammad Ghabel Rahmat**
Department of Computer Engineering,
Ka.C. Islamic Azad University
Karaj, Iran
M.ghabelrahmat@iau.ac.ir

Samira Vejdanparast
Department of Computer Engineering,
Ka.C. Islamic Azad University
Karaj, Iran
S.vejdanparast@iau.ac.ir

Azam Bastan Fard
Department of Computer Engineering,
Ka.C. Islamic Azad University
Karaj, Iran
Bastanfard@kiau.ac.ir

Abbas Jalilvand
Department of Computer Engineering,
Ka.C. Islamic Azad University
Karaj, Iran
Jalilvand@iau.ac.ir

Abstract — Rehearsal free continual learning requires models to acquire new task-specific knowledge without erasing previously learned information, despite the absence of stored exemplars. However, many distillation-based strategies employ a static knowledge-distillation weight, applying a fixed regularization strength across all tasks. This static formulation is unable to reflect the varying degrees of model drift that occur during incremental learning, often resulting in either excessive rigidity or insufficient retention. To address this limitation, we introduce AdaKD-LoRA, a parameter efficient framework that combines low-rank adaptation with dynamically regulated distillation. In this approach, a pretrained backbone remains frozen to preserve general representations, while lightweight LoRA modules encode task-specific updates. The strength of the distillation term is adaptively adjusted based on the divergence between consecutive task models, allowing the method to increase emphasis on stability when drift is high and to maintain flexibility when new classes require stronger plasticity. Experiments on a ten-task Class-Incremental Learning benchmark show that this adaptive formulation substantially

reduces forgetting and achieves more stable knowledge preservation compared with Fine-Tuning, LwF, and LoRA-only baselines. AdaKD-LoRA delivers higher accuracy on earlier tasks, more favorable backward transfer, and competitive overall performance, demonstrating that replacing static distillation with an adaptive mechanism provides a more effective and scalable solution for rehearsal-free continual learning.

Keywords— Continual Learning, Rehearsal-Free, Adaptive Knowledge Distillation, LoRA, Class-Incremental Learning, Catastrophic Forgetting.

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1139

A novel idea for Abductive Planning on Temporal Knowledge Graphs

*Amir Hossein Sharafi**
Najm
Department of Mathematics, Tafresh University
Tafresh, Iran
Am_sharafi@tafreshu.ac.ir

Alireza Shahbazi
Najm
Tehran, Iran
shahbazi@najm.ac

Behrouz Minaei Bidgoli
Najm
School of Computer Engineering, Iran University of Science and Technology
Tehran, Iran
b_minaei@iust.ac.ir

Abstract—Reasoning and planning over Dynamic Knowledge Graphs (DKGs) are critical for intelligent systems, but real-world applications are often hindered by incomplete information. While theoretical frameworks exist for integrating temporal and action formalisms, they often fall short in providing robust planning strategies for sparse-knowledge environments. This paper introduces a comprehensive framework grounded in standard Description Logics (ALC) that operates on a "Generate-and-Test" paradigm. The core contribution is a Recursive Hierarchical Strategy that integrates a Goal-Oriented A* Planner with a Level-by-Level Abductive Reasoner. The process begins with an A* search to find action sequences (defined in OWL) that satisfy a set of goals. When action preconditions are not met, they are treated as abductive assumptions. The system employs an AAA-style Abductive Reasoner to find simpler explanations for these assumptions. The planner is then recursively invoked to find actions that satisfy these simpler assumptions. This cycle continues up to a fixed depth. Finally, the generated plans are validated using a Temporal Projection Engine that assigns dummy timestamps to actions to verify goal satisfaction and logical consistency. We also demonstrate how this framework can proactively discover "Contradictory Stories" by planning for logically conflicting goals

Keywords— AI Planning, Dynamic Knowledge Graph, Temporal Reasoning, Abductive Reasoning, AAA Solver, A* Search, Inconsistency Detection.

Article Code: 1151

AI-CADx in Retinal Disease: A Systematic Review of Explainability, Privacy, and Scalability for Next-Generation Telemedicine

*Mohammad Shojaeinia**
PhD in Medical Informatics
Department of Health Information Technology and Management, School of Allied Medical Sciences
Shahid Beheshti University of Medical Science
Tehran, Iran
Shojaeinia.m@sbmu.ac.ir

Hamid Moghaddasi
Professor in Health information management and Medical Informatics
Department of Health Information Technology and Management, School of Allied Medical Sciences
Shahid Beheshti University of Medical Sciences
Tehran, Iran
Moghaddasi@sbmu.ac.ir

Abstract— The evolution of Artificial Intelligence–based Computer-Aided Diagnostic (AI-CADx) systems is exerting a profound influence on contemporary medical diagnostics, shifting their role from auxiliary support tools for clinical specialists to scalable, population-level screening solutions. This transition has been particularly consequential in telemedicine environments and resource-constrained healthcare settings. The present systematic review analyzes the developmental trajectory of AI-CADx systems, with emphasis on advances in model architectures, implementation methodologies, and real-world clinical integration, using retinal disease assessment as a representative application domain. A comprehensive literature search of PubMed, IEEE Xplore, arXiv, and Google Scholar (2015–2025) yielded 847 initial records, from which 78 studies met the inclusion criteria for final synthesis. Recent research demonstrates that key innovations are progressively overcoming long-standing barriers to clinical adoption. The fusion of multimodal imaging data with Explainable AI (XAI) methodologies, such as GradCAM++, substantially enhances model interpretability by making decision pathways more transparent, thereby strengthening clinician confidence in

automated outputs. In parallel, emerging privacy-preserving computational approaches, including federated learning and differential privacy, support secure processing of sensitive patient data while maintaining compliance with confidentiality requirements. The effectiveness of these synergistic strategies is illustrated by systems such as MadhuNetri and KIDROP, which report high sensitivity and specificity in identifying referable diabetic retinopathy and retinopathy of prematurity, respectively, exemplifying the potential of AI-driven platforms to alleviate workforce shortages and broaden access to high-quality ophthalmic care. This review offers three novel contributions: (1) a synthesized framework integrating technical performance, privacy preservation, and explainability dimensions specific to retinal AI-CADx; (2) a critical analysis of bias and generalizability challenges across diverse populations; and (3) a human-centered governance model for responsible clinical deployment. Despite these advances, the routine clinical deployment of AI-CADx systems continues to face significant challenges. Variability in regulatory requirements, unresolved issues related to algorithmic bias, and the need for comprehensive, humancentered governance frameworks remain critical considerations. Addressing these factors is essential to ensuring the long-term transparency, fairness, and reliability of AIenabled diagnostic technologies.

Keywords—Artificial intelligence, computer-aided diagnosis, ophthalmology, deep learning, explainable AI (XAI), multimodal fusion, privacy-preserving computation.

Article Code: 1160

Beyond Semantics: A Perception-Based CNN Model for Quantifying Phonetic Rhythm in Speech

*Mohammad Mahdi Peyravi**
Research Institute for Interdisciplinary Quran Studies Shahid Beheshti University
Tehran, Iran
mohammadmahdi.peyravi@iau.ac.ir

Dr. Alireza Talebpour
Research Institute for Interdisciplinary Quran Studies Shahid Beheshti University
Tehran, Iran
talebpour@sbu.ac.ir

Dr. Zeynab Hajimohammadi
Research Institute for Interdisciplinary Quran Studies Shahid Beheshti University
Tehran, Iran
z_hajimohammadi@sbu.ac.ir

Abstract—Quantifying subjective, prosodic features of speech, such as rhythm, remains a significant challenge in Natural Language Processing. A primary obstacle is "semantic bias," where a listener's perception of sound is heavily influenced by the meaning of the word. This paper proposes a novel methodology to overcome this challenge and develop a robust computational model for objectively quantifying phonetic rhythm. Our method involves three key steps: 1) Operationally defining rhythm within a continuous two-dimensional space (Percussiveness/Softness and Tempo/Slowness) based on phonetics. 2) Collecting perceptual ground-truth data from native speakers via an online survey. 3) Critically, to eliminate semantic bias, we designed and utilized a corpus of phonotactically valid meaningless sentences. These sentences were generated using a heuristic function to ensure full coverage of the rhythmic spectrum. Finally, a Convolutional Neural Network (CNN) was trained on this perception data, learning to predict rhythmic scores directly from phoneme sequences. Experimental results demonstrate that the CNN architecture significantly outperforms a baseline LSTM model, confirming our hypothesis that phonetic rhythm is a local phenomenon best captured by CNNs. This work provides a foundational tool for computational stylistics and advanced Text-to-Speech (TTS) systems.

Keywords—Computational Phonetics, Speech Rhythm, Convolutional Neural Network (CNN), Perception-Based Learning, Semantic Bias, Natural Language Processing

Article Code: 1166

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Federated LLM-Based SIEM with Meta-Model Aggregation: Enabling Collaborative, Scalable, and Privacy-Preserving Threat Detection

Masoud GanjKhani
Faculty of Computer Science and Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
m_ganjkhani@sbu.ac.ir

Alireza Shameli-Sendi
Faculty of Computer Science and Engineering
Shahid Beheshti University (SBU)
Tehran, Iran
a_shameli@sbu.ac.ir

Abstract— The escalating sophistication of cyber threats necessitates a paradigm shift from traditional, isolated Security Information and Event Management (SIEM) systems toward intelligent, collaborative defense frameworks. This paper presents a novel Federated Large Language Model (LLM)- Based SIEM architecture that combines privacy-preserving federated learning principles with the advanced reasoning capabilities of LLMs for distributed network intrusion detection. Our approach enables multiple security nodes to collaboratively learn and improve detection accuracy by sharing abstracted attack patterns rather than sensitive raw traffic data, thereby ensuring strict privacy preservation. To provide structured threat intelligence and standardized attack taxonomy, all network intrusions are classified according to the MITRE ATT&CK framework, enabling alignment with industry-standard threat modeling and facilitating interoperability with existing security infrastructure. Experimental validation demonstrates that the proposed system achieves 79% overall accuracy with a 16.1% improvement across five federated learning cycles, successfully detecting 11 distinct attack categories mapped to MITRE ATT&CK techniques, including Reconnaissance (Port Scan - 100%), Resource Hijacking (DoS/DDoS), Credential Access (Brute Force – 87.5%), and Initial Access (SQL Injection, Phishing). The distributed architecture comprises

three autonomous nodes sharing a centralized knowledge base that accumulates 196 validated detection rules through collaborative learning, with Port Scan attacks achieving perfect 100% detection accuracy. This work demonstrates that LLM-based federated learning can deliver competitive detection performance while maintaining data sovereignty and providing interpretable, MITRE-aligned threat intelligence for modern SIEM deployments.

Keywords— Federated SIEM, LLM, MITRE ATT&CK, Privacy-Preserving Classification, Knowledge Sharing

Article Code: 1168

Generative AI and Individual Entrepreneurial Orientation

*Saeed Mohammadi**
Faculty of Management and Accounting
Shahid Beheshti University (SBU)
Tehran, Iran
Email: sae_mohammadi@sbu.ac.ir

Mahshid Mirzaee
Faculty of Entrepreneurship
University of Tehran
Tehran, Iran
Email: mahshidmirzaee3@gmail.com

Abstract— This research conducts a contemporary empirical analysis of generative artificial intelligence (GenAI)'s effects on the five facets of individual entrepreneurial orientation (Ind.EO): innovativeness, proactiveness, risk-taking, autonomy, and competitive aggressiveness. Grounded in resource-based view (RBV) theory, it posits positive associations and evaluates them through logistic regression using survey responses from 428 Iranian entrepreneurs. Findings reveal notable boosts in four facets among GenAI adopters—innovativeness (OR=1.96), proactiveness (OR=2.10), autonomy (OR=1.93), and competitive aggressiveness (OR=2.22)—yet no influence on risk-taking (OR=1.26, p=0.312). Such outcomes emphasize GenAI's capacity to elevate specific entrepreneurial traits while delineating limits, offering insights for theory, training, and strategy in AI-enhanced entrepreneurship.

Keywords— Generative AI, individual entrepreneurial orientation, innovativeness, proactiveness, risk-taking, autonomy, competitive aggressiveness.

Article Code: 11 74

Resilience or Recession? A Longitudinal Analysis of NFT Market Dynamics and AI Generated Value During the Crypto Downturn

*Aria Shakoo**
Department of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
a.shakoo@mail.sbu.ac.ir

Amirhossein Ashrafian
Department of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
s.ashrafian@mail.sbu.ac.ir

Maedeh Mosharraf
Department of Computer Science and Engineering
Shahid Beheshti University
Tehran, Iran
m mosharraf@sbu.ac.ir

***Abstract*—**The Non-Fungible Token (NFT) market has transitioned from the speculative hyper-growth of 2021-2023 to a period of stabilization and correction in 2024-2025, often termed the "crypto winter." While prior literature has extensively documented value drivers in bull markets, the asset pricing dynamics during prolonged market contractions remain underexplored. This paper presents a comprehensive longitudinal analysis of this downturn, focusing on two critical dimensions: the valuation reversal of Artificial Intelligence (AI)-generated assets versus traditional hand-drawn collections, and the comparative robustness of machine learning models in predicting prices under high-volatility conditions. We curated a dataset of over 113,000 transactions from 2024-2025 and benchmarked findings against a 2023 baseline dataset of near 7 million transactions. Our results reveal a stark inversion of market sentiment: while AI assets commanded a premium in 2023, they traded at a near 76% discount compared to traditional collections in 2024 ($\mu \approx 0.92$ Ether (ETH) vs 4.16 ETH), signaling a "flight to quality." Furthermore, we evaluated two pricing models: a Deep Learning (DL) Multi-Layer Perceptron (MLP) and an ensemble Random Forest (RF). Contrary to the hypothesis that deep neural networks would better capture non-linear market volatility, the RF model ($R^2 = 0.74$) significantly

outperformed the MLP (R 2 = 0.71) on the winter data. This comparative study suggests that in sparse, high-noise financial environments, ensemble decision trees remain more robust than deep architectures.

Keywords—NFT, Crypto Winter, Generative AI, Price Prediction, Deep Learning, Random Forest, Financial Modeling

Article Code: 1180

ORB-SVM: An Innovative Hybrid Framework for Efficient Brain Tumor Detection from MRI Scans

*Amirhosein Azarpour**

Department of Computer Science

Shahid Beheshti University

Tehran, Iran

amirazarpour07@gmail.com

Abstract—Brain cancer remains one of the most significant challenges in modern medicine, where the accuracy of early stage diagnosis is a decisive factor in patient survival and treatment efficacy. Although Magnetic Resonance Imaging (MRI) is the established gold standard for visualizing neurological structures, the interpretation of these high dimensional scans is often complicated by subjective variability among practitioners and the inherent noise present in complex medical images. While contemporary approaches frequently rely on high parameter deep learning architectures, such models often involve significant computational costs and require extensive data for effective training. This study introduces a hybrid framework that utilizes the Oriented FAST and Rotated BRIEF (ORB) algorithm for precise feature extraction and a Support Vector Machine (SVM) for classification [15] [11]. The proposed approach achieves a substantial data reduction of approximately 99.5%, which effectively minimizes the influence of non informative background data while preserving critical diagnostic patterns essential for tumor identification. By balancing feature sparsity with a robust kernel based classifier, this methodology addresses the limitations of over parameterized systems while maintaining high diagnostic integrity. Experimental evaluations conducted on the Br35H dataset demonstrate that the framework attains

a classification accuracy of 97.5%. The findings suggest that the integration of localized feature representation and optimized classification provides a reliable and resource efficient alternative for medical image analysis, offering a structured solution that maintains performance without the need for extensive computational overhead.

Keywords—ORB, Support Vector Machine, Brain Tumor, Data Compression, Medical image processing, MRI images

Article Code: 1182

An Automated Modular Framework for MITRE ATT&CK Mapping

*Alireza Katani**

Department of Computer Engineering

Isfahan University of Technology

Isfahan, Iran

a.katani@ec.iut.ac.ir

Abstract— The significant surge in cyberattacks in recent years has placed a heavy burden on both governmental and private organizations. Security analysts are often overwhelmed with a multitude of repetitive, time-consuming tasks, many of which could be effectively automated. One of the most crucial tasks is the manual and labor-intensive process of mapping diverse security events, Security Information and Event Management (SIEM) rules, and intrusion-detection alerts to the techniques defined in the MITRE ATT&CK framework. This not only hampers analyst productivity but also compels organizations to increase staffing and absorb higher operational costs. In this study, we present an automated approach to alleviate this burden for both analysts and organizations. Our method utilizes a curated dataset consisting of attack descriptions and alerts generated by various intrusion detection tools and SIEM platforms, all of which have been labeled according to MITRE ATT&CK techniques. This dataset was used to fine-tune a lightweight large language model (LLM), enabling the development of an intelligent system capable of automatically mapping diverse cybersecurity text inputs to MITRE ATT&CK techniques.

Keywords—Cybersecurity, MITRE ATT&CK, LLM, SIEM

Article Code: 1186

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Zero-Shot, Standard Fine-Tuning, and Curriculum Learning Approaches for VQA in GI Endoscopy

Mahdi Azmoodeh-Kalati
Institute of Medical Science and Technology
Shahid Beheshti University
Tehran, Iran
m.azmoodehkalati@alumni.sbu.ac.ir

Mohammad Sadegh Maghareh
Computer Science Department
Amirkabir University of Technology
Tehran, Iran
maghareh@aut.ac.ir

*Reza Lashgari**
Institute of Medical Science and Technology
Shahid Beheshti University
Tehran, Iran
r_lashgari@sbu.ac.ir

Abstract— Visual Question Answering (VQA) on gastrointestinal (GI) endoscopy images is a challenging multimodal task due to complex visual content and specialized clinical language. This paper presents a comparative study of three inference strategies for GI endoscopy VQA using a large vision–language model: (1) zero-shot inference without domain training, (2) conventional fine-tuning on the target dataset, and (3) curriculum-guided finetuning with progressively harder training stages. Building on our prior work, we adapt a state-of-the-art 2-billion-parameter vision–language model to the Kvasir-VQA-x1 dataset (6,500 images, 159k QA pairs) using Low-Rank Adaptation (LoRA). Experimental results on a held-out evaluation set demonstrate that domain-specific fine-tuning yields substantial performance gains over the zero-shot baseline (e.g., BLEU improves from 0.0700 to 0.2513). Moreover, curriculum learning further boosts accuracy across metrics (e.g., BLEU 0.2667), indicating better generalization to diverse question complexities. We describe the three-stage curriculum setup in detail, where each “easy” question is seen three times and “moderate” questions twice during training, potentially reinforcing core knowledge. The proposed curriculum-guided strategy achieves more balanced VQA performance across simple and complex questions than standard fine-tuning. We discuss the impact of repeated

exposure to easier samples on learning and outline future steps to vary the curriculum schedule. Overall, our work highlights curriculum learning as an effective method to improve GI endoscopy VQA, outperforming both zero-shot and standard fine-tuning approaches in this domain.

Keywords— Visual Question Answering, Gastrointestinal Endoscopy, Curriculum Learning, Fine-Tuning, Vision-Language Models

Article Code: 1189

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Hybrid Fuzzy-Q Learning Approach for Load Balancing in IoT Networks

*Fereshteh Taghizadeh**
School of Electrical and Computer Engineering, Shiraz University, Shiraz,Iran
ftaghizadeh@shirazu.ac.ir

Mohsen Raji
School of Electrical and Computer Engineering, Shiraz University, Shiraz,Iran
mraji@shirazu.ac.ir

Morteza Keshtkaran
School of Electrical and Computer Engineering, Shiraz University, Shiraz,Iran
mkeshtkaran@cse.shirazu.ac.ir

Abstract—Artificial Intelligence is redefining the optimization of complex IoT routing systems, enabling adaptive, data-driven decision-making in scenarios where traditional methods fall short. In traditional IoT networks, maintaining reliable and energy-efficient routing under dynamic traffic patterns and topology changes remains a critical challenge. To address this, we propose Fuzzy-QRPL, a hybrid load-balancing framework that integrates fuzzy logic with Q-learning to enhance routing efficiency in RPL-based IoT networks. Two key metrics— Residual Energy Potential (EP) and Congestion Index (CI)—are computed from queue occupancy, end-to-end delay, number of child nodes, and residual energy. Nine fuzzy rules are defined, each initialized with a Q-value and updated according to a reward based on the packet delivery ratio (PDR). This AI-enhanced mechanism enables dynamic and intelligent parent selection under varying network conditions. Cooja simulation results across diverse network sizes demonstrate that Fuzzy-QRPL significantly improves PDR, reduces delay, and increases the number of successfully delivered packets compared to conventional RPL objective functions.

Keywords— RPL, Load Balancing, Fuzzy Logic, Q-learning, IoT Networks, Parent Selection, Reinforcement Learning

Article Code: 1191

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Applications of Artificial Intelligence in Upstream Oil and Gas Industries

Parisa Salimi
Department of Research & Development
Oil Design and Construction Company
Tehran, Iran
p.salimi@odcc.ir

Zahra Nargessi
Department of Research & Development
Oil Design and Construction Company
Tehran, Iran
z.nargessi@odcc.ir

Fatemeh Alamolhoda
Department of Research & Development
Oil Design and Construction Company
Tehran, Iran
f.alamolhoda@odcc.ir

Abstract—Artificial intelligence has emerged in recent years as a transformative foundational technology in the oil and gas industry, with its greatest impact observed in the upstream sector where data complexity, high operational costs, and technical uncertainties necessitate rapid and precise decision-making. Advances in machine learning algorithms, deep neural networks, and large-scale data analytics have enabled significant improvements across exploration, drilling, and production operations. In exploration, AI-based models enhance seismic data processing, geological feature detection, uncertainty analysis, and subsurface structure prediction, thereby increasing the accuracy of identifying new reservoirs and reducing investment risks. In drilling operations, predictive algorithms support real-time monitoring of well parameters, decreasing the likelihood of events such as stuck pipe, fluid loss, and formation instability, while optimizing well trajectory design. Furthermore, in reservoir engineering and production management, data-driven models can more accurately predict reservoir behavior, production rates, pressure dynamics, fracture networks, and saturation changes, offering an effective alternative to classical simulation approaches. The application of reinforcement learning and intelligent production control systems further contributes to improved recovery factors, reduced operating costs, and optimized well performance.

Given the rapid evolution of digital technologies and the shifting landscape of the global energy market, the adoption of artificial intelligence is no longer a technological option but a strategic necessity for enhancing efficiency, operational agility, and mitigating technical and economic risks in the upstream domain. Investment in data-centric infrastructure, integration of information systems, and development of skilled human resources is essential for harnessing the full potential of AI and advancing toward intelligent, resilient, and sustainable production in the oil and gas industry.

Keywords—*Artificial Intelligence, Big Data, Upstream, Oil and Gas Industries, Digitalization*

Article Code: 1193

Performance Analysis of Variational Quantum Classifiers for Classification Tasks in the NISQ Era

Saghar Kiani
Dept. of Computer Engineering
Ka.c., Islamic Azad University Karaj
Karaj, Iran
Saghar.Kiani@iau.ir,
ORCID: 0009-0002-7284-9268,

Paria Kamalpour
Dept. of Computer Engineering
Ka.c., Islamic Azad University Karaj
Karaj, Iran
Paria.Kamalpour@iau.ir,
ORCID: 0009-0002-5242-1192,

Sarina Ghafouri
Dept. of Computer Engineering
Ka.c., Islamic Azad University Karaj
Karaj, Iran
Sarina.Ghafouri@iau.ir,
ORCID: 0009-0007-7564-1530

*Niloofar Mirzaei Chahardeh**
Dept. of Computer Engineering
SR.C., Islamic Azad University Tehran
Tehran, Iran
Niloofar.Mirzaeichahardeh@iau.ir,
ORCID: 0000-0002-2025-6949

***Abstract*—**Quantum Machine Learning (QML) leverages superposition, entanglement, and high-dimensional Hilbert spaces to model complex data structures that can be difficult for classical algorithms to approximate efficiency. In the Noisy Intermediate-Scale Quantum (NISQ) era, limited qubit counts, shallow circuits, and hardware noise constrain practically realizable quantum circuits and motivate the study of variational and hybrid approaches.This work presents a small-scale empirical benchmark of a two-qubit Variational Quantum Classifier (VQC) on the Iris dataset compared with a classical RBF-SVM baseline under noiseless and simulated depolarizing noise conditions. The VQC, implemented in Qiskit Aer with angle encoding and COBYLA optimization, achieves an accuracy of 0.93 ± 0.02 and an F1-score of 0.92 ± 0.03, close to the SVM's 0.94 ± 0.01 and 0.93 ± 0.02, with a small but statistically detectable advantage for the classical model ($p = 0.04$, Cohen's $d = 0.48$). Under a 2% depolarizing noise channel, the VQC exhibits a smaller relative drop in accuracy than the SVM, suggesting favorable behavior under mild simulated noise. The study serves as an educational, reproducible benchmark rather than a demonstration of quantum advantage.

***Keywords*—**quantum machine learning, variational quantum classifier, nisq, depolarizing noise, feature encoding.

Article Code: 1194

LLU-460: A Dynamic Dataset of Large Language Model Usage Among Students at Islamic Azad University, Karaj

Alireza Jafari
Dept. of Computer Engineering, KA.C.,
Islamic Azad University
Karaj, Iran
alireza.jafarii@iau.ir
ORCID: 0009-0004-0198-8513

Seyed Mohammad Amin Alemohammad
Dept. of Computer Engineering and Sciences, CT. C., Islamic Azad University Tehran, Iran
mohammadamin.alemohammad@iau.ir
ORCID: 0009-0002-3714-2910

*Niloofar Mirzaei Chahardeh**
Dept. of Computer Engineering, SR.C.,
Islamic Azad University
Tehran, Iran
Niloofar.Mirzaeichahardeh@iau.ir
ORCID: 0000-0002-2025-6949

Abstract— In recent years, the use of large language models and AI tools has become increasingly common among university students, yet their actual usage patterns remain poorly understood. To investigate this, we developed an interactive web application that collected responses from students at Islamic Azad University, Karaj. Unlike traditional surveys, the system adapted questions based on each student's prior answers, enabling more precise and context-aware data collection. Using this platform, 460 complete responses were gathered. Although the dataset is moderate in size, it reveals meaningful patterns, including the frequency of AI tool usage, preferred tools, and students' perceptions of their academic utility. A notable aspect is the slightly higher proportion of male participants, introducing a minor imbalance. Despite this, the dataset reflects real student behavior and can serve as a valuable resource for future research. For modeling and prediction, we applied two supervised learning algorithms: Random Forest and XGBoost. XGBoost performed consistently on this limited dataset, achieving approximately 0.80 accuracy, whereas Random Forest showed signs of overfitting with unrealistically perfect results. Overall, our findings demonstrate that real student data can model AI tool usage patterns and suggest that expanding the dataset will enhance the reliability of future analyses. The LLU- 460 dataset and Streamlit web application are publicly available.

Keywords— llms, web application, llu-460, xgboost, random forest.

Second International Interdisciplinary Conference on Artificial Intelligence

Article Code: 1195

Personalized Serious Games for ADHD: A Systematic Review and Evidence Synthesis

Mohammad Cheraghi
dept. of Computer Engineering and Sciences, Islamic Azad University, Central Tehran Branch
Tehran, Iran
Mohammad.Cheraghi@iau.ir
ORCID:0009 0005 4076 7980

Milad Hassani
dept. of Information and Technology Managemennt, College of Farabi, University of Tehran
Qom, Iran
Miladhassani@ut.ac.ir

*Niloofar Mirzaei Chahardeh**
dept. of Computer Engineering, SR.C.,
Islamic Azad University
Tehran, Iran
Niloofar.Mirzaeichahardeh@iau.ir
ORCID: 0000-0002-2025-6949

Abstract—Approximately 5-7% of children globally experience ADHD, a neurodevelopmental disorder manifesting through difficulties in sustained attention, excessive motor activity, and impaired impulse control. Game based digital interventions have emerged as engaging therapeutic tools, with adaptive personalization proposed to amplify clinical benefits by tailoring task parameters to each user's cognitive profile. We conducted a systematic synthesis of 20 investigations examining adaptive game based interventions for pediatric and adolescent ADHD populations, spanning publications from January 2011 through August 2024. Our multi database search strategy (five databases) yielded 72 distinct records, with 20 ultimately satisfying eligibility requirements following full text screening. Adaptive features predominantly implemented through artificial intelligence and machine learning algorithms encompassed dynamic difficulty adjustment, individualized performance feedback, and customized reward timing. Relative to fixed difficulty comparison conditions, adaptive interventions produced modest to moderate gains in attentional capacity and working memory performance, alongside variable but predominantly favorable impacts on symptom ratings. User engagement remained uniformly elevated across studies, while treatment emergent adverse effects occurred infrequently. Nevertheless, evidence certainty was constrained by methodological limitations including modest sample sizes, insufficient assessor blinding, and heterogeneous outcome

.

instrumentation resulting in moderate to high bias risk across the corpus. Adaptive game based platforms thus constitute a provisionally encouraging supplementary approach within multimodal ADHD treatment, particularly for augmenting cognitive capacities. Subsequent investigations should emphasize adoption of uniform assessment instruments, extended monitoring periods, explicit algorithmic documentation, and methodologically robust designs conducive to independent replication and quantitative pooling.

Keywords—adhd, serious games, personalization, adaptive algorithms, machine learning

Article Code: 1197

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

A Multi-Dimensional Taxonomy and Practical Framework for MLOps Monitoring

Tina Hafezi
dept. of Computer Engineering,
National University of Skills
Tehran, Iran
Email: tinahafezii@gmail.com
ORCID: 0009-0003-4011-8943

Sara Pour Faraj
dept. of Computer Engineering,
National University of Skills
Tehran, Iran
Email: sarahpourfaraj@gmail.com
ORCID: 0009-0006-1044-2728

Seyed Mohammad Amin Alemohammad
dept. of Islamic Azad University
Central branch
Tehran, Iran
Email:
mohammadamin.alemohammad@iau.ir

*Niloofar Mirzaei Chahardeh **
dept. of Computer Engineering, Islamic Azad University, Science and Research Branch
Tehran, Iran
Email:
Niloofar.Mirzaeichahardeh@iau.ir
ORCID: 0000-0002-2025-6949

Abstract— Unlike conventional software systems, machine learning (ML) systems operate in dynamic environments that can contain non-stationary data, delayed labels, and continuously changing behavior. These differences make monitoring a fundamental competency within MLOps, providing the observability needed to maintain the reliability, stability, and long-term business value of ML systems. This paper proposes a three-dimensional taxonomy of ML monitoring metrics encompassing data quality constraints, statistical drift metrics, classical performance indicators, behavioral proxies, system health measures, and business performance metrics, including cost and value dimensions, to address the most common practical scenarios. Building upon this taxonomy, we introduce a practitioner-oriented framework for selecting, integrating, and operationalizing monitoring metrics across the ML lifecycle. The framework incorporates metrics aligned with stakeholder objectives and links technical and business indicators. It also integrates early-warning diagnostics, automated responses, and both label-dependent and label-independent monitoring. Furthermore, for maintaining resilience against data drift and concept drift, interpretability, human-in-the-loop evaluation, and

automated retraining mechanisms are emphasized as essential elements. We also outline observability modules, model lineage tracking, automation triggers, and retraining logic that operationalize the proposed framework. The presented approach provides a unified foundation for designing robust monitoring systems that preserve model performance over extended periods, reduce operational costs, and strengthen the reliability and business alignment of production ML systems.

Keywords— mlops, machine learning monitoring, data drift, concept drift, model metrics.

Article Code: 1198

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Hybrid Machine Learning and Deep Learning Approach for Minimum Airfare Prediction

Mohammad Safaiyan
Department of Electrical and Computer Engineering, Faculty of Computer Engineering, Isfahan University of Technology
Isfahan, Iran
Email: m.safaiyan77@gmail.com

Ali Mahvash Mohammadi
Department of Software Engineering, Faculty of Computer Engineering, University of Isfahan
Isfahan, Iran
Email: alimahvashm4@gmail.com

Abstract— Airfare prediction is a challenging task due to dynamic pricing mechanisms, limited real-time visibility into fare updates, and nonlinear demand–supply interactions. This paper focuses on Minimum Fare Prediction: forecasting the lowest airfare a route will reach within the next three days so that users can decide whether to buy now or wait. We propose a hybrid framework that integrates gradient-boosted tree ensembles (XGBoost and Histogram Gradient Boosting Regressor) with temporal neural architectures (1D-CNN, LSTM, and a CNN→LSTM hybrid) trained on a unified feature space. The models are trained and evaluated on a proprietary Pateh.com dataset of approximately 49.7 million flight search records collected over 480 days, one of the largest corpora analyzed in airfare prediction research. Each instance corresponds to the minimum observed daily fare for a specific origin–destination route, enriched with 20 engineered features including calendar descriptors, lagged fares, and rolling 3-day minimum prices. To evaluate practical usefulness, we introduce Regression Weighted Accuracy (RWA), a tolerance-based metric that rewards predictions within a ±10% band of the true price and penalizes larger deviations. Using a time-series cross-validation scheme with 12 folds, our hybrid CNN→LSTM model achieves higher RWA than existing baselines such as k-nearest neighbors, gradient boosting, and bagging regression trees on a key route (IFN–DXB), and reaches up to 0.60 RWA on selected high-traffic routes. The approach is currently being A/B tested within Pateh's production system to provide "buy now vs. wait" recommendations and display the predicted minimum fare for the next three days.

Keywords—Minimum airfare prediction; Hybrid ensemble learning; XGBoost; LSTM; 1D-CNN; Dynamic pricing systems; Regression Weighted Accuracy (RWA); Large-scale time series

Article Code: 1203

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

From Code to Clinic: Bridging the Implementation Gap in AI-Based Clinical Decision Support

*Ehsan Ghazanfari Savadkoohi**
Department of Health Information Technology and Management
School of Allied Medical Sciences,
Shahid Beheshti University of Medical Sciences
Tehran, Iran
ghazanfarisavadkoohi@gmail.com

Farkhondeh Asadi
Department of Health Information Technology and Management
School of Allied Medical Sciences,
Shahid Beheshti University of Medical Sciences
Tehran, Iran
asadifar@sbmu.ac.ir

Hassan Emami
Department of Health Information Technology and Management
School of Allied Medical Sciences,
Shahid Beheshti University of Medical Sciences
Tehran, Iran
haemami@sbmu.ac.ir

Abstract— Artificial intelligence-based clinical decision support systems (AI-CDSS) are increasingly proposed to improve diagnosis, risk prediction, and clinical workflow. However, a persistent implementation gap remains: many models report strong retrospective performance, yet few are sustainably integrated into routine care. This narrative review examines the technical, human, organizational, regulatory, and economic factors contributing to this gap and proposes a lifecycle-oriented roadmap for sustainable AI-CDSS implementation. We conducted a structured narrative synthesis of English-language literature published between January 2015 and September 2025, including empirical deployment studies, systematic and narrative reviews, implementation science frameworks, reporting guidelines, and regulatory or policy documents. The analysis focused on technical infrastructure, workflow integration, human factors, governance, equity, and post-deployment monitoring. The review shows that implementation failure is rarely explained by algorithmic performance alone. Instead, failure often emerges from socio-technical misalignment, including fragmented data infrastructure, dataset shift, limited prospective validation, unclear clinical ownership, workflow

disruption, alert fatigue, liability uncertainty, and weak reimbursement incentives. Compared with traditional rulebased CDSS, AI-CDSS introduce additional challenges related to model drift, opacity, local calibration, adaptive regulation, and long-term surveillance. To address these challenges, we propose a three-phase roadmap: local readiness assessment and silent-mode calibration; workflow-integrated and graded deployment; and sustained algorithmic stewardship supported by MLOps, governance mechanisms, and equity monitoring. Sustainable implementation requires a shift from modelcentric evaluation to system-centric translation, treating AICDSS as adaptive clinical services that require continuous validation, human-centered design, institutional accountability, and stakeholder alignment.

Keywords— Artificial Intelligence; Clinical Decision Support Systems; Implementation Science; Algorithmic Stewardship; MLOps; Translational Gap; Socio-technicalSystems

Article Code: 1205

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Comparative Evaluation of YOLO-Based and CNN Models for Wheat-Head Detection in Precision Agriculture

Shahin Rostami
dept. of Computer Engineering,
KA.C., Islamic Azad University
Karaj, Iran
shahin.rostami2310@iau.ir
ORCID:0009-0002-9581-3655

Shervin Danesh
dept. of Computer Engineering,
KA.C., Islamic Azad University
Karaj, Iran
shervin.danesh7902@iau.ir
ORCID:0009-0001-9159-8786

*Niloofar Mirzaei Chahardeh**
dept. of Computer Engineering,
SR.C., Islamic Azad University
Tehran, Iran
Niloofar.Mirzaeichahardeh@iau.ir
ORCID:0000-0002-2025-6949

Abstract— Precise identification of wheat heads is essential for crop yield estimation, plant condition monitoring, and automated support systems in precision agriculture. This study presents a standardized comparative evaluation of four deep learning detectors YOLOv8, YOLOv11, Faster R-CNN, and MobileNetV2 for wheat head localization on the Global Wheat 2020 dataset. All models were trained and tested under identical preprocessing and experimental settings to ensure fairness. The comparison focuses on detection accuracy, inference speed, and deployment feasibility on resource-constrained edge devices for real-time field applications. Evaluation metrics include precision, recall, F1-score, mAP@0.5, and per-image inference time.Among the evaluated models, YOLOv8 achieved the best overall balance between accuracy and speed, reaching a benchmark mAP@0.5 of 0.96 while maintaining real-time performance. YOLOv11 demonstrated a competitive trade-off between detection robustness and computational efficiency, making it suitable for wheat head detection tasks. Faster R-CNN obtained high precision but suffered from longer inference times, limiting practical on-site deployment. MobileNetV2 provided faster execution but with lower detection reliability.

Additional ablation experiments indicate that transformer-based feature integration and adaptive convolution significantly improve YOLOv11's stability in densely populated field scenes. Overall, this work provides a fully standardized, direct comparison of the four models on the Global Wheat 2020 benchmark and establishes a new performance reference for wheat head detection.

Keywords— wheat-head detection, deep learning, precision agriculture, YOLOv11, real-time object detection.

Article Code: 1211

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

A Systematic Review of Tokenizer Adaption Techniques for Large Language Models

Niloofar Mortazavi
Faculty of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
NiloofarMortazavi@ut.ac.ir

*Mostafa Amiri**
Faculty of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
Mostafa.Amiri@ut.ac.ir

Hesham Faili
Faculty of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
hfaili@ut.ac.ir

Abstract—This systematic review synthesizes keyresearch contributions in adapting Large Language Models(LLMs) to diverse linguistic contexts, focusing specifically onnovel methods for transferring or replacing tokenizers andinitializing associated vocabulary embeddings. The field ofLarge Language Model (LLM) adaptation has rapidlyevolved beyond simple vocabulary expansion towardssophisticated strategies that aim for high-fidelity transferwith minimal post-hoc fine-tuning. Recent methodsprimarily focus on how to initialize the new embeddinglayer—the lexical interface—to maximize preservation of theLLM's knowledge encoded in the frozen transformer core.

Keywords—cross-lingual adaptation, tokenizer transfer,tokenizer replacement, large language models

Article Code: 1212

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Intelligent Leaf Disease Diagnosis Using EfficientNetB0 Combined with CBAM in a MultiTask Framework

*Nasrin Nikootadbir**
Faculty of Engineering
University of Birjand
Birjand, Iran
nasrin_nikootadbir@birjand.ac.ir

Mohammad Hasan Majidi
Faculty of Engineering
University of Birjand
Birjand, Iran
m.majidi@birjand.ac.ir

Abstract— Automatic plant leaf disease detection plays a crucial role in smart agriculture by reducing crop losses and enabling timely treatment. In this study, a multi-task deep learning model based on EfficientNetB0 is proposed to simultaneously predict three outputs: leaf health status, plant species, and disease type. To enhance the feature representation of discriminative regions, the Convolutional Block Attention Module (CBAM) is integrated into the network. Furthermore, a specialized Gated Disease Head is designed to exploit the semantic dependency between plant species and disease type, while a Consistency Supervision mechanism is introduced to prevent contradictory predictions among the outputs. The model is trained on a dataset containing approximately 4,000 leaf images from barberry, jujube, and pomegranate plants, using a two-stage optimization strategy including warm-up and fine-tuning phases. Experimental results demonstrate that the proposed approach achieves 97% accuracy in health classification, 93% in species identification, and 93% in disease recognition, while maintaining a significantly lower parameter size compared to state-of-the-art methods. Overall, the proposed framework provides an efficient, lightweight, and reliable solution for real-world intelligent plant disease monitoring systems.

Keywords—Deep learning, CBAM, EfficientNetB0, multi-task learning, plant disease detection, consistency

Article Code: 1215

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Artificial Intelligence in ADHD Diagnosis and Treatment: A PRISMA-Aligned Systematic Review

Fatemeh Ranjbar
dept. of Computer Engineering, KA.C.,
Islamic Azad University
Karaj, Alborz, Iran
F.Ranjbar0241@iau.ir
ORCID: 0009-0004-4069-2600

Rashin Gholijani Farahani
dept. of Computer Engineering, KA.C.,
Islamic Azad University
Karaj, Iran
Rashin.Gholijanifarahani@iau.ir
ORCID: 0009-0001-5092-3052

Mohammad Cheraghi
dept. of Computer Engineering and Sciences,
Islamic Azad University, Central Tehran Branch
Tehran, Iran
Mohammad.Cheraghi@iau.ir
ORCID: 0009-0005-4076-7980

Seyed Sepehr Mirhoseyni Nayeri
dept.of Computer Engineering KA.C.,
Islamic Azad University
Karaj, Iran
S.Mirhoseyni3252@iau.ir
ORCID: 0009-0008-1554-8374

*Niloofar mirzaei Chahardeh**
dept. of Computer Engineering, SR.C.
Islamic Azad University
Tehran, Iran
Niloofar.Mirzaeichahardeh@iau.ir
ORCID: 0000-0002-2025-6949

Abstract— ADHD is one of the most prevalent neurodevelopmental disorders worldwide, affecting both children and adults. Conventional diagnostic approaches remain limited by subjectivity, inter-rater variability, and cultural or clinical biases. Artificial Intelligence has emerged as a promising avenue for developing more objective and data-driven diagnostic methods. This systematic review, conducted in accordance with PRISMA 2020 guidelines, synthesizes findings from 64 empirical studies published between 2020 and 2025, covering a broad spectrum of AI-enabled approaches including functional neuroimaging (fMRI), electrophysiological signals (EEG), physiological sensing, behavioral monitoring, speech and NLP analysis, and digital therapeutic tools such as Virtual Reality and robotics. To the best of the authors' knowledge, this is the first PRISMA-compliant review to comprehensively address all six of these modalities within a single synthesis. Machine learning and deep learning models performed remarkably well in controlled research settings, with internal validation accuracy for fMRI- and EEG-based approaches frequently

reported between 94 and 99%. However, these high accuracy values are almost exclusively derived from internal or single-site validation, raising substantial concerns regarding overfitting and real-world generalizability. In the therapeutic domain, VR- and robotics-based interventions demonstrated high engagement and feasibility, yet consistent evidence for the transfer of such benefits to everyday functioning remains scarce. This review identifies a substantial gap between experimental AI models and clinically deployable systems, and proposes that bridging this gap will require cross-site datasets, rigorous external validation, transparent reporting standards, and robust ethical frameworks ensuring fairness, safety, and privacy in real-world ADHD assessment and care.

Keywords—adhd, artificial intelligence, machine learning, diagnosis, digital health

Article Code: 1217

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Cross-Dataset Empirical Evaluation of NSGA-II for Multi-Objective Feature Selection in Intrusion Detection Systems

*Mahdis Rahmani**
Department of Mathematics and Computer Science
Allameh Tabataba'i University
Tehran, Iran
mahdis_rahmani@atu.ac.ir

Fereshteh-Azadi Parand
Department of Mathematics and Computer Science
Allameh Tabataba'i University
Tehran, Iran
parand@atu.ac.ir

***Abstract*—**Intrusion Detection Systems (IDS) play a crucial role in network security, as they monitor and detect cybersecurity attacks. However, the increasing complexity and dimensionality of IDS datasets pose significant challenges for training intrusion detection models. Redundant or irrelevant features within these datasets can decrease classifier performance and increase computational overhead. As a result, feature selection becomes an essential step in the preprocessing phase. Many existing feature selection methods used in IDS datasets rely on single-objective optimization, which often struggles to effectively balance the trade-off between minimizing feature count and maximizing accuracy. This study conducts a cross-dataset empirical evaluation of the Non-dominated Sorting Genetic Algorithm II (NSGA-II) for multi-objective feature selection in IDS datasets. We evaluate the proposed framework using nine benchmark datasets namely NSL-KDD, UNSW-NB15, CIC-IDS2017, CSE-CIC-IDS2018, UKM-IDS20, UNR-IDD, N-BaIoT, BoT-IoT, and X-IIoTID, which encompass traditional networks, modern networks, and IoT environments.Experimental results demonstrate that NSGA-II reduces the feature space by an average of 72.71% while maintaining detection accuracy. In many cases, feature reduction results in improved performance, with the best instance achieving a 1.28% increase in accuracy. Although, four configurations experienced minor accuracy degradation (average -0.56%),

highlighting the importance of careful dataset-classifier pairing. These findings highlight NSGA-II as a generalizable and effective method for feature selection in high-dimensional IDS applications.

Keywords—Non-dominated Sorting Genetic Algorithm II (NSGA-II), Feature Selection, Intrusion Detection Systems (IDS), Multi-objective Optimization

Article Code: 1221

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

AI-Driven Materials Genome: Accelerated Discovery of Multi-Functional Sensing Materials

Farzane Hasheminia
School of Advanced Thecnologies
Iran University of Science and Technology
Tehran, IRAN
Farzane_Hasheminia@nt.iust.ac.ir

*Sadegh Sadeghzadeh**
School of Advanced Thecnologies
Iran University of Science and Technology
Tehran, IRAN
Sadeghzadeh@iust.ac.ir

Behrouz Minaei-Bidgoli
School of Computer Engineering
Iran University of Science and Technology
Tehran, IRAN
b_minaei@iust.ac.ir

Abstract— Multi-functional sensing materials play a critical role in advanced technologies, including medical devices, environmental monitoring, and smart infrastructure. The discovery and development of these materials using traditional methods is often time-consuming and costly, relying predominantly on experimental testing. The materials genome approach, which integrates computational data, simulations, and predictive analyses, provides a pathway to accelerate the materials discovery process. In this article, a roadmap for the design of multi-functional sensing materials is presented, encompassing data-driven strategies, predictive algorithms, and experimental feedback loops. Limitations such as data scarcity, model uncertainty, and challenges in result interpretability are also discussed, along with perspectives on developing faster and more efficient methods for materials discovery and optimization.

Keywords—materials genome, sensor, materials design, aidriven, dataset.

Article Code: 1224

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

An Extended Table Detection Framework for Enhanced Recognition of Persian Tables in OCR Systems

Aras Valizadeh
Computer science and engineering
Shahid Beheshti University
Tehran, Iran
ara.valizadeh@mail.sbu.ac.ir

Bita Baroutian
Computer science and engineering
Shahid Beheshti University
Tehran, Iran
bi.baroutian@mail.sbu.ac.ir

Fateme Taraghi
Computer science and engineering
Shahid Beheshti University
Tehran, Iran
f.taraghi@mail.sbu.ac.ir

Zahra Mohammadi
Computer science and engineering
Shahid Beheshti University
Tehran, Iran
zahra.mohammadi4@sbu.ac.ir

Atefe Aghaei
Computer science and engineering
Shahid Beheshti University
Tehran, Iran
a_aghaei@sbu.ac.ir

*Mohsen Ebrahimi Moghaddam**
Computer science and engineering
Shahid Beheshti University
Tehran, Iran
m_moghadam@sbu.ac.ir

Abstract—Currently, with the widespread usage of automated document processing and retrieval systems, the importance of accurate OCR has increased significantly. However, robust OCR for Persian documents remains a challenging task, particularly in the presence of complex table layouts that exacerbate the intrinsic difficulties of the Persian script. Existing OCR and VLM-based methods often face limitations in reliably identifying and interpreting tables. These approaches are further hindered by structural hallucinations that compromise downstream reasoning and extraction. In response to these challenges, a modular, OCR-agnostic framework has been proposed. This has been specifically designed to improve table recognition in Persian documents. The framework is designed to integrate with existing OCR systems, enhancing the robustness of table detection across a wide range of real-world layouts. A dedicated dataset of Persian tables was gathered from different sources to facilitate systematic evaluation. The experimental results demonstrate that the proposed approach achieves superior table recognition performance in comparison to state-of-the-art OCR and VLM methods. This establishes a strong baseline for future research in Persian table analysis.

Keywords—Persian OCR, table detection, document layout analysis, vision–language models (VLMs), structural recognition

Article Code: 1225

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Active Learning Strategies for Imbalanced Clinical Note Categories in ICU Mortality Prediction

*Mina Afzali**
Department of Computer Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
m.afzali@mail.um.ac.ir

Mohsen Kahani
Department of Computer Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
kahani@um.ac.ir

Maryam Karrabi
Department of Computer Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
maryam.karrabi@mail.um.ac.ir

Nadieh Armin
Department of Computer Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
n.armin@mail.um.ac.ir

Abstract— Accurate mortality prediction in intensive care units (ICUs) is critical for improving patient outcomes and optimizing clinical decision-making. Clinical notes contain rich unstructured information that can complement structured data in predictive models. However, these datasets are often highly imbalanced, with far fewer mortality cases than survivals, limiting the performance of traditional machine learning methods. In this study, we perform a comparative analysis of different clinical note categories for ICU mortality prediction and introduce an active learning framework to address data imbalance. By iteratively selecting the most informative samples for model training, active learning improves predictive accuracy while reducing annotation costs, resulting in faster model training and effective management of data imbalance. Our experiments demonstrate that the application of Active Learning led to a significant reduction in both model execution time and the number of labeled samples required for training. Crucially, this enhancement in efficiency was concurrent with an improvement in predictive performance, as quantified by elevated AUC scores for several categories.

Keywords— mortality prediction, unstructured clinical notes, data imbalance, active learning, Note embedding, MIMIC III

Article Code: 1233

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

Enhancing Facial Emotion Recognition Using YOLO11 Classification on the AffectNet Dataset

*Reza Nasiri**
Dept. of Computer Engineering
Faculty of Engineering, Shahid Chamran University of Ahvaz
Ahvaz, Iran
Reza.Nasiri@stu.scu.ac.ir

Seyed Enayat Alavi
Dept. of Computer Engineering
Faculty of Engineering, Shahid Chamran University of Ahvaz
Ahvaz, Iran
se.alavi@scu.ac.ir

Mohammad Javad Rashti
Dept. of Computer Engineering
Faculty of Engineering, Shahid Chamran University of Ahvaz
Ahvaz, Iran
mohammad.rashti@scu.ac.ir

Abstract— Facial Expression Recognition (FER) is a core component of affective computing, enabling intelligent systems to interpret human emotional states in real-world environments. Despite considerable progress, FER performance on large-scale in-the-wild datasets such as AffectNet remains limited due to illumination variations, pose changes, occlusions, intra-class variability, and annotation noise. Existing architectures including attention-enhanced CNNs, landmark-guided models, and transformer-based approaches have achieved incremental improvements but generally remain below 70% accuracy on the AffectNet-7 benchmark. In this study, we introduce a comprehensive FER framework based on YOLO11- Classification, a high-capacity classification architecture with 28.3 million parameters, optimized feature extraction, and enhanced receptive-field modeling. Without relying on handcrafted FER-specific modules such as facial alignment, landmark extraction, or region-attention engineering, the proposed system achieves 79.50% Top-1 accuracy on AffectNet-7, surpassing the previous state-of-the-art (NorFace, 68.69%) by an absolute margin of 10.81%. Extensive experiments demonstrate the robustness of YOLO11-Classification under real-world challenges, including occlusions, extreme head poses, and class imbalance, while confusion matrix analysis confirms consistent improvements across all emotional categories, particularly those with limited representation. These findings establish a new state-of-the-art for AffectNet-7 and highlight the potential of high-capacity general-purpose architectures as powerful alternatives to FER-specific handcrafted designs.

Keywords— Facial Expression Recognition (FER), YOLO11, AffectNet, Deep Learning, Affective Computing, Convolutional Neural Networks.

Venue: Chamran Highway – Daneshjoo Boulevard – Shahid Beheshti University

IICAI
Second International Interdisciplinary Conference on Artificial Intelligence

IEEE IRAN SECTION

IEEE

دانشگاه شهید بهشتی