ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY

DOCTORAL (PhD) DISSERTATION
(Thesis book)

Mohammadreza Azodinia

Artificial Intelligence for Modeling Ground-Penetrating Radar Data

Supervisor:
Dr. Amir Mosavi

Doctoral School of Applied Informatics and Applied Mathematics

Budapest,
May 2026

Abstract

This thesis is focused on artificial intelligence-based modeling for non-invasive structural health monitoring of pavements using ground penetrating radar data. The main objective is defined as support for informed and predictive decision making in road maintenance and infrastructure management. Pavement structural condition is estimated from subsurface sensing data, asphalt temperature, and pavement layer thickness, which are obtained from radiotechnical and sensor-based measurements. Nondestructive and high-resolution data acquisition is combined with artificial intelligence methods, which enables scalable and cost-efficient monitoring. Several machine learning and deep learning approaches are examined, including Support Vector Machine (SVM), Random Forests (RF), Artificial Neural Network (ANN), and long short-term memory models (LSTM). Model performance is evaluated by root mean square error (RMSE) and correlation coefficient (CC) during training, validation, and testing phases. Strong accuracy is achieved by SVM during training, while robust generalization is achieved by RF during testing. In addition, superior predictive accuracy and stability are demonstrated by LSTM across unseen datasets, hence this model is identified as the most effective approach for sequential pavement data. The results confirm that artificial intelligence enables automated, real time, and continuous assessment of pavement structural condition. As a result, maintenance actions are scheduled in a timely manner, which leads to extended pavement service life and reduced lifecycle costs. Furthermore, integration of the proposed models into smart city and smart mobility platforms is supported, which enhances road safety and resource allocation. Therefore, this thesis establishes a unified foundation for artificial intelligence based, non-invasive pavement monitoring and sustainable urban transportation management.

Keywords: artificial intelligence, ground penetrating radar, non-invasive assessment, pavement structural health monitoring, machine learning, deep learning, predictive maintenance, smart city infrastructure, pavement condition evaluation, sustainable transportation

Introduction

T1: Machine Learning for Pavement Performance and Service Life Prediction

Research problem: Existing ML reviews in pavement engineering lack a unified reproducible methodology for gap identification.

Methodology: A systematic PRISMA-guided review of 328 sources is conducted, covering supervised, unsupervised, and deep learning methods for pavement performance and RSL prediction.

Key result: A structured synthesis identifies the main limitations of current data-driven pavement models, including data quality, model transparency, and limited adoption of deep learning. These gaps motivate the experimental contributions in T3, T4, and T5.

Supporting publication: Azodinia, M., Shayakhmetov, A., & Mosavi, A. (2025). Machine learning for pavement performance and service life prediction. Eurasian Journal of Mathematical and Computer Applications, 13(4), 41-52. **[T1]**

Structural health monitoring and predictive pavement analysis are essential for road safety, durability, and operational efficiency within modern transport networks. These activities support planning, design, construction, mobility assessment, transport modelling, and maintenance phases. Conventional pavement evaluation methods are constrained by high cost, logistical demands, and traffic disruption, which limits suitability for large scale network management. For this reason, machine learning approaches are increasingly applied to address these limitations and close existing research gaps. Variables such as asphalt temperature and pavement layer depth are commonly adopted, because they influence structural response and pavement behaviour. As a result, predictive models support proactive maintenance planning, improved resource allocation, and enhanced road safety. Furthermore, integration with smart cities and smart mobility initiatives enables adaptive and data driven infrastructure governance. Accordingly, stress evolution modeling with AI has been also discussed in detail **[T2]**.

Before the adoption of artificial intelligence, SHM relied mainly on manual inspection and direct testing techniques. These methods include visual damage assessment, load testing, vibration analysis, acoustic emission inspection, ultrasonic testing, radiographic evaluation, GPR assessment, infrared thermography, and falling weight deflectometer measurements. Although valuable diagnostic information is provided, predictive capability remains limited. Early stage or subsurface defects are often undetected, while extensive operational interruptions are required. Artificial intelligence alters SHM because rapid and continuous analysis of large datasets is enabled in real time. ML and DL support forward looking assessments through sensor data, imagery, and historical records, which improves prediction of deterioration and maintenance demand.

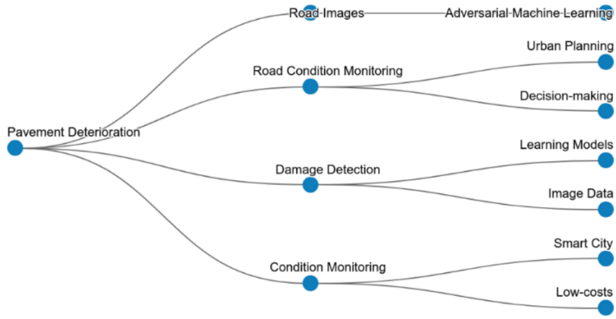

Fig.1: Taxonomy of ML Applications for Pavement Modelling

Numerous investigations confirm the advantages of AI for pavement SHM. ML methods such as decision trees, RF, SVM, and ANN are applied to forecast pavement degradation, RSL, and maintenance demand by use of traffic, climate, and material data. DL, particularly CNNs, is applied to GPR data for subsurface defect detection, which improves deterioration prediction and maintenance planning. Real time monitoring approaches based on embedded sensors capture pavement response to traffic and weather, hence earlier fault detection is achieved with minimal disruption. In addition, large scale data platforms integrate traffic patterns, weather histories, and performance records, which improves maintenance prioritization and resource allocation. These applications demonstrate that AI improves performance forecasting, continuous monitoring, and sustainable pavement management.

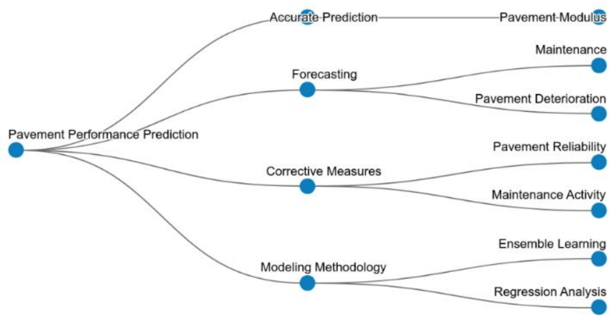

Fig.2: Taxonomy of DL Applications in Pavement Modelling

Despite substantial progress, several challenges remain. Model generalization is often limited because calibration is performed for specific climatic or operational conditions. Predictive robustness is weakened by incomplete data integration and limited representation of uncertainty related to traffic and environmental variability. Moreover, AI models often operate as black boxes, which reduces transparency and confidence among practitioners. Weak linkage with PMS also restricts practical application. Recent advances highlight the need for hybrid modelling, uncertainty quantification, explainable AI, and continuous learning strategies. Developments such as big data analytics, IoT integration, GAN based data generation, and geospatial mapping show strong potential. However, operational cost, data acquisition frequency, and implementation complexity require careful evaluation. Therefore, a multidisciplinary approach is required to ensure resilient, interpretable, and sustainable AI based SHM solutions **[T2]**.

T2: Machine Learning for Stress Evolution Modeling

Research problem: No unified taxonomy of ML methods for stress evolution modeling exists, which limits systematic method selection across scientific domains.

Methodology: Supervised, unsupervised, and deep learning methods are classified according to their suitability for stress evolution problems using a systematic PRISMA-based review.

Key result: A new cross-domain taxonomy is produced that maps ML method characteristics to stress evolution modeling requirements and defines hybrid modeling directions.

Supporting publication: Wang, P., Mosavi, A., Felde, I., Azodinia, M., et al. (2025). Machine learning for modeling stress evolution. Acta Polytechnica Hungarica, 22(12), 95-114. **[T2]**

Materials and Methods

In this thesis as described in **[T3]**, modelling is conducted to predict pavement SHM based on data driven approaches. AC surface temperature, AC thickness, and BS thickness are treated as independent variables, while RSL is defined as the dependent variable. Seven ML and DL techniques are applied, namely SGD, XGBoost, ANN, ANN combined with PSO, ANN combined with GWO, AdaBoost, and LSTM. Model development is performed using Python 3.6.5 with TensorFlow, Scikit learn, and XGBoost libraries, as well as MATLAB 2018, on identical hardware configurations. The dataset is divided into training and testing subsets using a 75:25 ratio. Hyperparameters are adjusted through a guided trial and error procedure, while MAE, RMSE, and R^2 are used to support model selection. Basic models such as SGD and

ANN are adopted as reference approaches, whereas ensemble and hybrid models are applied to capture nonlinear relationships. LSTM is included because sequential pavement behaviour is represented over time **[T3]**.

The dataset is obtained from a radiotechnical device based on GPR **[T4]**, which allows nondestructive estimation of pavement layer thickness. Radargrams are collected using air coupled and ground coupled antennas to capture subsurface electromagnetic responses. Data preparation includes normalization, treatment of missing values, and correction of signal noise and antenna height effects. Calibration of dielectric properties is applied to improve depth estimation accuracy. This procedure enables continuous data acquisition across large areas and supports both automated and data driven analysis.

Feature selection is applied to improve predictive accuracy and reduce computational burden. The Relief score is used to identify the most informative features based on their ability to distinguish between different data instances. This method evaluates local differences between neighbouring samples and assigns higher weights to features with stronger discriminative power. As a result, dimensionality is reduced while relevant information is preserved. The selected features are then used as inputs to the ML and DL models, which enhances robustness and reliability of pavement SHM predictions.

Fig 3. Feature selection with the Relief score

The Relief score has shown strong potential in SHM, especially for managing high-dimensional data from numerous sensors. It helps reduce dimensionality by identifying and keeping only the most informative features, which is essential when dealing with large datasets common in SHM applications.

$$W[A] = W[A] + \sum_{i=1}^{m} \frac{1}{m} (diff\left(A, x_i, Miss(x_i)\right)^2 - (diff\left(A, x_i, Hit(x_i)\right)^2) \tag{1}$$

$$diff\left(A, x_i, x_j\right) = \frac{\left|x_j^A - x_i^A\right|}{Max(A) - Min(A)} \tag{2}$$

Results

T3: Deep Learning for Structural Health Monitoring of Pavements

Research problem: Temporal pavement degradation is not adequately captured by static ML models, which limits predictive maintenance accuracy.

Methodology: Seven ML and DL models are compared, including LSTM, which models sequential pavement behaviour using asphalt temperature and layer thickness as inputs.

Key result: LSTM achieves the lowest RMSE and highest CC across all phases, which confirms its suitability for time-series pavement SHM.

Supporting publication: Azodinia, M. et al. (2025). Decision Making: Applications in Management and Engineering, Vol. 8, Issue 1. **[T3]**

Thesis 3: The state-of-the-art and the research gap had been investigated and published in **[T1&T2].** The modeling novelty and contributions are presented in the publication titled: Deep Learning for Structural Health Monitoring of Pavements for Improving Road Maintenance and Management **[T3]** and the model had be even further improved in the article titled: AI-based Structural Health Monitoring of Pavements: An Informed Decision-Making Approach **[T4]**. Feature selection results are illustrated, which presents Relief scores for five features, namely A303, A302, A301, A118, and A117. All features record positive scores, which confirms their contribution to classification performance. The highest relevance is observed for A303, A302, and A301, which indicates strong discriminative capability. Features A118 and A117 show slightly lower scores, however meaningful relevance is still demonstrated. This outcome supports dimensionality reduction while key information is retained for SHM related tasks. Based on the selected features, model training is conducted and the results are reported in

Table 2. The training outcomes reveal distinct learning behaviour across SVM, RF, and ANN. Among these models, SVM demonstrates the most consistent performance. The lowest RMSE value of 0.841 is achieved alongside a high CC value of 0.991, which indicates strong agreement between predicted and observed values.

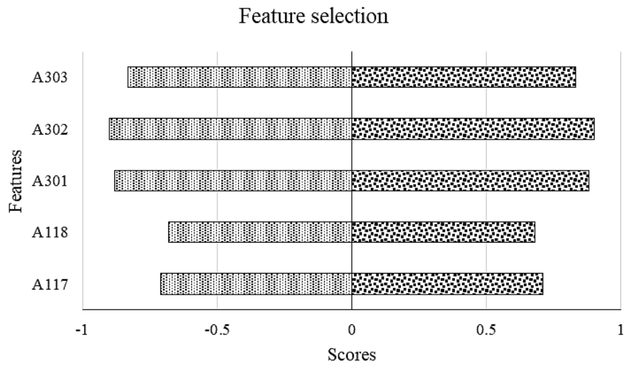

Fig 4. Results of the feature selection

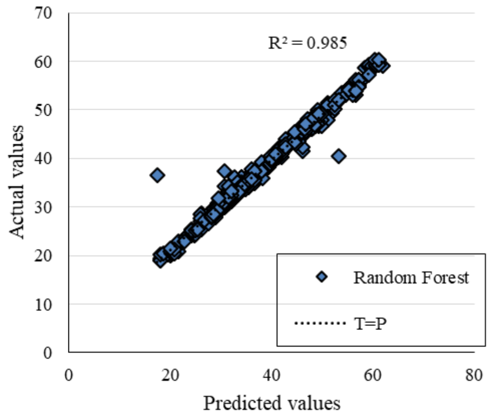

Fig 5. Results of the testing phase for RF

T4: AI-based Structural Health Monitoring of Pavements

Research problem: Non-destructive radiotechnical subsurface sensing data require a systematic feature selection and model comparison pipeline to support informed maintenance decisions.

Methodology: Relief-based feature selection is applied, followed by comparative evaluation of SVM, RF, and ANN on GPR-derived pavement condition data.

Key result: RF achieves the lowest mean percentage error (above 1%) with minimal variability. SVM remains reliable below 2%.

Supporting publication: Azodinia, M. et al. (2026).

Decision Making: Applications in Management and Engineering, Vol. 9, Issue 1. **[T4]**

Thesis 4: AI-based Structural Health Monitoring of Pavements: An Informed Decision-Making Approach [T4]. According to this thesis, the mean percentage error analysis during testing confirms the comparative model performance. ANN shows the highest mean error, around 9%, with large variability, indicating low accuracy and poor stability. SVM performs reliably, with a mean error below 2% and a narrow error bar, demonstrating consistency. RF achieves the lowest mean error, just above 1%, with minimal variability, reflecting high precision and strong generalisation to unseen data. These results reinforce that RF provides the most accurate and robust predictions, SVM remains dependable, and ANN is unsuitable without substantial tuning or redesign.

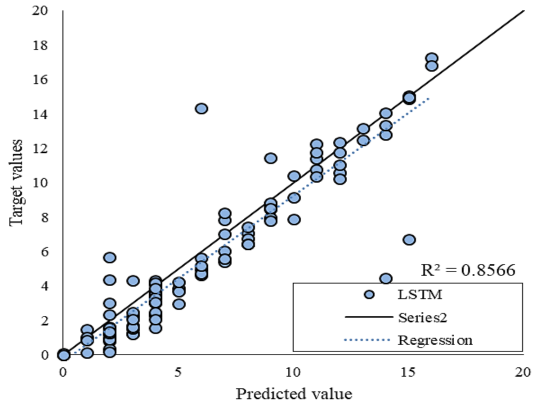

Fig 6. The Plot Diagram of the Testing Phase LSTM

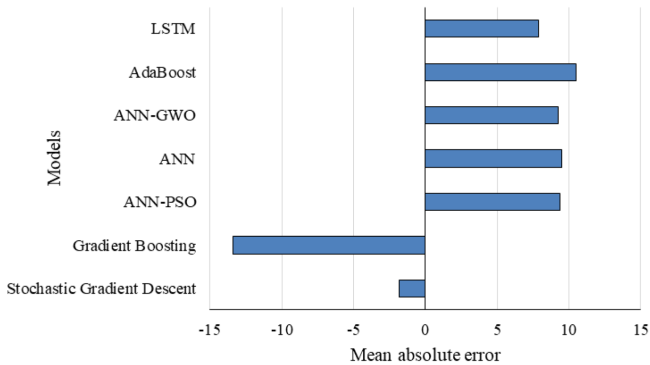

Fig 7. Square Relative Error of Testing Phase

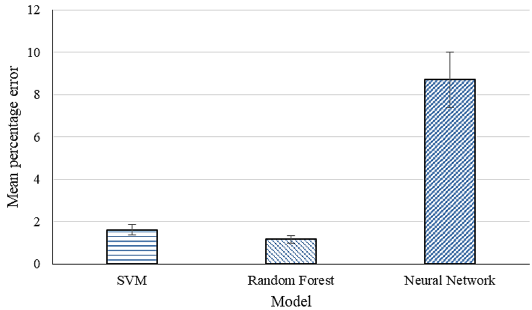

Fig 8. Mean percentage error values in the testing phase

Discussion

T5: Deep Learning for Structural Integrity Assessment of Pavements

Research problem: The trade-off between predictive accuracy and computational efficiency in DL models for pavement RSL prediction is not systematically evaluated.

Methodology: GRU and LSTM models are applied to structural field measurements and compared with respect to RMSE, CC, and computational cost.

Key result: LSTM achieves superior predictive accuracy. GRU offers a computationally efficient alternative suitable for real-time deployment.

Supporting publication: Azodinia, M., Mosavi, A., & Besheyev, D. (2026). Journal of Applied Computer

Science and Telecommunications, 1(1), 38-42. [T5]

The results of this thesis demonstrate that AI, particularly LSTM, provides highly accurate predictions for non-invasive pavement SHM using GPR data. LSTM achieves the lowest RMSE and highest CC, reflecting its capacity to model sequential and temporal patterns inherent in pavement degradation. Hybrid ANN models, such as ANN-GWO, improve predictive performance through optimisation of network parameters but remain less effective at capturing long-term dependencies compared with LSTM. Traditional ML approaches, including XGBoost and SGD, perform moderately, as they cannot fully exploit the temporal structure of GPR datasets. These findings highlight that combining non-invasive GPR sensing with advanced AI models enables precise estimation of critical pavement parameters, such as layer thickness, subsurface conditions, and surface temperature, which is essential for proactive maintenance and infrastructure planning. Furthermore, the study confirms that model selection should account for both predictive accuracy and generalisation to unseen real-world data, ensuring reliable monitoring across diverse pavement conditions. Further elaboration and discussions given in **[T3, T4, T5].**

Conclusion

This thesis establishes that non-invasive GPR data, when integrated with AI models, provides an effective and scalable approach for pavement SHM. LSTM demonstrates the highest predictive performance, followed by optimized ANN models, while other ML methods show lower reliability. The combination of AI and non-invasive sensing enables automated, real-time, and accurate monitoring without disrupting traffic, offering cost-effective and safe infrastructure assessment. These findings support proactive maintenance strategies, improve pavement longevity, and optimise resource allocation. The approach is adaptable to different urban and regional contexts, providing a foundation for large-scale deployment of AI-driven, non-invasive SHM. Future work should focus on expanding datasets, integrating traffic and environmental data, exploring hybrid and ensemble AI techniques, and implementing explainable AI to enhance interpretability and practical adoption in infrastructure management.

In the overall assessment of this dissertation, Thesis Statements **T3**, **T4**, and **T5** constitute the core independent scientific results. T3 establishes LSTM as the most effective model for time-series pavement SHM. **T4** demonstrates RF and SVM superiority for GPR-based prediction. **T5** confirms the accuracy-efficiency trade-off between LSTM and GRU for RSL prediction. Thesis Statement **T2** is accepted as a structured methodological synthesis and cross-domain taxonomy. Thesis Statement **T1** serves as a foundational literature contribution that identifies the research gaps motivating the experimental work in **T3**, **T4**, and **T5**.

References

T1: Azodinia, M., Shayakhmetov, A., & Mosavi, A. (2025). Machine learning for pavement performance and service life prediction. Eurasian Journal of Mathematical and Computer Applications, 13(4), 41–52. https://doi.org/10.32523/2306-6172-2025-13-4-41-52

T2: Wang, P., Mosavi, A., Felde, I., **Azodinia**, M., Delavar, A., Amirov, A., & Shi, W. (2025). Machine learning for modeling stress evolution. Acta Polytechnica Hungarica, 22(12), 95–114. https://doi.org/10.12700/aph.22.12.2025.12.

T3: M. **Azodinia**, K. Iskakov, O. Indira, S. Ardabili, N. Karballaeezadeh, A. Mosavi, Deep Learning for Structural Health Monitoring of Pavements for Improving Road Maintenance and Management, Decision Making: Applications in Management and Engineering, Volume 8, Issue 1 (2025).

T4: M. **Azodinia**, K. Iskakov, M. Turarova, S. Ardabili, A. Mosavi, AI-based Structural Health Monitoring of Pavements: An Informed Decision-Making Approach, Decision Making: Applications in Management and Engineering, Volume 9, Issue 1 (2026).

T5: M. **Azodinia**, A., Mosavi, Besheyev, D., Deep Learning for Structural Integrity Assessment of Pavements and Improved Road Maintenance, The Journal of Applied Computer Science & Technology (JACST), №1, 2025

Azodinia, M., Mudabbir, M., Ardabili, S., Várkonyi Kóczy, A. R., Iskakov, K., & Mosavi, A. (2025). Service life modeling of pavement with ensemble learning. In Proceedings of the IEEE 12th International Conference on Computational Cybernetics and Cyber Medical Systems (ICCC) (pp. 167–174). IEEE.

Mosavi, A., & **Azodinia**, M. (2025). Gated recurrent unit for modeling structural integrity of pavements for enhanced road maintenance. In Proceedings of the 10th International Scientific and Practical Conference on Computer Science and Applied Mathematics (CSAM 2025) (pp. 462–469). Al Farabi Kazakh National

University.

Azodinia, M., Iskakov, K., Turarova, M., Ardabili, S., & Mosavi, A. (2025). Machine learning for structural health monitoring of pavements. In Proceedings of the XVII Saginov Readings International Scientific and Practical Conference (pp. 11–13). Karaganda Technical University.

Azodinia, M., Mosavi, A., Delavar, A., Iskakov, K., Turarova, M., & Tatin, A. (2025). How machine learning is transforming pavement performance and service life prediction. In Proceedings of the IEEE 23rd International Symposium on Intelligent Systems and Informatics (SISY 2025) (pp. 331–337). IEEE.

Further publications of the candidate not used in the thesis

Mudabbir, M., Mosavi, A., Felde, I., Moniz, N., Iskakov, K., & **Azodinia**, M. (2025). Modeling remaining service life and structural health monitoring of roads with machine learning and deep learning. In Proceedings of the IEEE World Symposium on Applied Machine Intelligence and Informatics (SAMI 2025) (pp. 413–421). IEEE.

Voskanyan, M., Ghorbani, H., & **Azodinia**, R. (2024). Utilizing citizen driven scientific endeavors for freshwater pollution surveillance. In Proceedings of the IEEE World Symposium on Applied Machine Intelligence and Informatics (SAMI 2024) (pp. 505–512). IEEE.

Voskanyan, M., Ghorbani, H., & **Azodinia**, M. (2024). An investigation of the hydrochemical parameters for natural monuments. In Proceedings of the IEEE International Conference on Computational Cybernetics and Cyber Medical Systems (ICCC 2024) (pp. 327–331). IEEE.

Beheshtian, S., Roodbari, S. K., Ghorbani, H., **Azodinia**, M., Mudabbir, M., & Várkonyi Kóczy, A. R. (2024). Comparative evaluation of machine learning and Bayesian deep learning methods for estimating ultimate recovery in shale well reservoirs. In Proceedings of the IEEE International Conference on Computational Cybernetics and Cyber Medical Systems (ICCC 2024) (pp. 17–23). IEEE.

Mosavi, A., **Azodinia**, M., & Akhmetzhanov, M. (2025). Gated recurrent unit for modeling structural integrity of pavements for enhanced road maintenance. In Proceedings of the 10th Scientific Conference Computer Science and Applied Mathematics (pp. 462–469). Al Farabi Kazakh National University.

Beheshtian, S., Roodbari, S. K., Ghorbani, H., **Azodinia**, M., & Mudabbir, M. (2024). Machine learning prediction of gas hydrates phase equilibrium in porous media. In Proceedings of the IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI 2024) (pp. 417–423). IEEE.

Azodinia, M., Mudabbir, M., & Karimipour, A. (2023). Numerical investigation of two phase Al2O3 nanofluid in a microchannel equipped with bump through slip flow. Engineering Analysis with Boundary Elements, 155, 1028–1034.

Ghorbani, H., Krasnikova, A., Ghorbani, P., Ghorbani, S., Hovhannisyan, H. S., Minasyan, A., Minasian, N., Ahmadi Alvar, M., Stepanyan, H., & **Azodinia**, M. (2023). Improving the estimation of coronary artery disease by classification machine learning algorithms. In Proceedings of the IEEE International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO EPE) (pp. 159–165). IEEE.

Ghorbani, H., Krasnikova, A., Ghorbani, P., Ghorbani, S., Hovhannisyan, H. S., Minasyan, A., Minasian, N., Ahmadi Alvar, M., Stepanyan, H., & **Azodinia**, M. (2023). Prediction of heart disease based on robust artificial intelligence techniques. In Proceedings of the IEEE International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO EPE) (pp. 167–173). IEEE.

Azodinia, M., & Hajdu, A. (2016). A novel method to increase the performance of recommender systems using a parallel CBIR approach. International Journal of Computer Science and Information Security, 14(10), 205–213.

Azodinia, M., & Hajdu, A. (2016). A novel combinational relevance feedback based method for content based image retrieval. Acta Polytechnica Hungarica, 13(5), 121–134.

Azodinia, M., Farrokhi, V., & Hajdu, A. (2014). Constant time median filtering of extra large images using Hadoop. In Proceedings of the 9th International Conference on Applied Informatics (ICAI 2014) (pp. 93–101).

Azodinia, M. R., & Hajdu, A. (2015). A recommender system that deals with items having an image as well as quantitative features. In Proceedings of the IEEE International Symposium on Intelligent Signal Processing (WISP 2015). IEEE.

Azodinia, M. R., & Hajdu, A. (2015). A method for image retrieval using combination of color and frequency layers. International Journal of Computer Applications, 118(3), 10–13.

Mosavi, A., **Azodinia**, M., Hewage, K. N., Milani, A. S., & Yeheyis, M. (2011). Reconsidering multiple criteria decision making problems of construction workers with the aid of Grapheur. In Proceedings of the EnginSoft International Conference (pp. 1–5).

Abbreviation	**Full term**
AI	Artificial Intelligence
GPR	Ground Penetrating Radar
SHM	Structural Health Monitoring
ML	Machine Learning
DL	Deep Learning
SVM	Support Vector Machine
RF	Random Forest
ANN	Artificial Neural Network
LSTM	Long Short Term Memory
CNN	Convolutional Neural Network
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
CC	Correlation Coefficient

R^2	Coefficient of Determination
RSL	Remaining Service Life
SGD	Stochastic Gradient Descent
PSO	Particle Swarm Optimization
GWO	Grey Wolf Optimization
AC	Asphalt Concrete
BS	Base/Subbase (context dependent)
IoT	Internet of Things
GAN	Generative Adversarial Network
PMS	Pavement Management System
MATLAB	Matrix Laboratory (software environment)
XGBoost	Extreme Gradient Boosting