



# Contemporary Machine Learning in Smart Cities: A Review of Quality, Measurability, Explainability, Privacy, and Security

Anja DELIĆ, Jelena KOVAČ, Nevena GLIGOROV, Branislav S. RISTIĆ, Marko GORDIĆ, Radovan TUROVIĆ, Dinu DRAGAN, Dušan B. GAJIĆ, Veljko B. PETROVIĆ

**Abstract:** This review paper examines current research on machine learning applications in smart cities, explicitly focusing on quality, measurability, explainability, privacy, and security. By synthesizing findings from recent studies, we uncover new trends, methods, and ways to measure performance that are key for developing and deploying these systems. We discuss how data privacy and system security challenges intertwine with the technical requirements of quality assurance and explainability, and we propose future research directions that could foster more reliable and accountable machine-learning solutions in urban environments. This review aims to provide researchers and practitioners with an overview of the current landscape, facilitating a multidisciplinary dialogue on enhancing trust and efficacy in smart city technologies.

**Keywords:** artificial intelligence, internet of things, machine learning, smart building, smart city, smart energy, smart health, smart home, smart logistics, smart security, smart transport.

## 1 INTRODUCTION

A smart city represents a symbiosis of technology and people in an urban environment. With Inter of Things (IoT) on the rise, we see more and more applications of data analysis and informed decision-making. This is primarily enabled by a surge in number of various sensors available in the urban environments. Due to their rising numbers, we are getting a high quantity of data, which is ideal for informed decision-making. However, this high-volume influx and complexity of the data would be overwhelming for traditional pipelines to process. That is where artificial intelligence (AI), i.e., machine learning (ML), methods are introduced. Foremost, they enable automatic data analysis and learning of the trends in the data. Secondly, as computer methods, they can leverage high computing power to deliver the results quickly, even in real-time. And finally, they can learn complex trends in the data.

Smart city applications are already successfully implemented in various cities. For example, Midtown in Motion program in New York that improves traffic in real-time [1] and Virtual Singapore initiative that through virtual duplication of the city allows for various experiments and tests to be done virtually [2].

The smart city field of research is vast, and there are even comprehensive review papers on subfields. In this already well-established field, we aim to review some of the latest research of ML applications in smart cities and see how they address the following five criteria: quality, measurability, explainability, privacy, security. For quality we focus on how the researchers quantify the quality of their proposed solution. For measurability we focus on the datasets, mainly how the quality of the solutions will translate to the real-world use-case. For explainability, we concentrate on the solution's ability to explain the reason for the given result to the end-user. For privacy we focus mainly on how the research addresses the privacy of people and how this data is prevented from exposure to malicious users. Finally, for

security, we emphasize how the research addresses the security and integrity of the system itself against malicious activities. These five questions emerged as a necessary condition for the successful application of ML in smart cities. These issues do not concern only the researchers and industry professionals, but also other stakeholders, such as the policymakers, end-users, and others.

Although not exhaustive, this research aims to review current literature in terms of quality, measurability, explainability, privacy, and security. The review will yield the potential paths toward improvement of current and future initiatives in smart cities.

The structure of the rest of the paper is as follows. The second section shows a quick overview of machine learning methods and metrics presented in reviewed papers. The third section is dedicated to the overview of the research with focus on quality, measurability, explainability, privacy, and security. The last section concludes our findings on recent ML applications in smart cities.

## 2 TECHNOLOGIES AND METRICS

The rapid advancement of AI and the widespread adoption of IoT solutions have driven the development of diverse computational strategies to address increasingly complex problems. At the core of AI lies ML[3], which encompasses classical algorithms such as k-nearest neighbors (k-NN), support vector machines (SVM), decision trees (DT), and Naïve Bayes (NB). These methods rely on pattern extraction from data to perform classification and regression tasks. A specialized subset of ML, deep learning (DL), employs multilayered artificial neural networks (ANNs) to automatically learn features from raw inputs. Prominent DL architectures include convolutional neural networks (CNNs) (e.g., VGG16, VGG19, MobileNet) for image analysis and long short-term memory (LSTM) networks for sequential data processing. Hybrid models, such as CNN-

LSTM, integrate spatial and temporal learning capabilities, further expanding the applicability of DL in various domains. Additionally, graph neural networks (GNNs) have emerged as powerful tools for processing structured data, such as social networks, molecular structures, and knowledge graphs.

Concurrently, ensemble learning techniques (e.g., random forest (RF), gradient boosting, XGBoost [4], LightGBM[5]) and optimization algorithms, including genetic algorithms (GA) and particle swarm optimization (PSO), enhance predictive accuracy and facilitate hyperparameter tuning. Reinforcement learning (RL) methodologies, such as Deep Q-Networks, Multi-Agent Actor-Critic, and Proximal Policy Optimization, enable agents to make sequential decisions by continuously learning from environmental feedback. Furthermore, federated reinforcement learning (FRL) distributes the learning process across multiple agents while preserving data privacy. Neuro-symbolic AI, which combines neural learning with symbolic reasoning, is emerging as a promising approach for bridging data-driven and rule-based AI methodologies. Additionally, self-supervised learning and contrastive learning have revolutionized AI by reducing dependency on labeled data, improving feature representations across diverse domains. Beyond model-centric approaches, edge computing reduces latency and bandwidth requirements by shifting computational processes closer to data sources—an essential consideration for IoT networks. Similarly, federated learning enables decentralized model training across multiple devices, improving both privacy and scalability. Neuromorphic computing, inspired by biological neural networks, promises ultra-efficient AI implementations through event-driven processing and hardware acceleration. In specialized applications, methodologies such as two-stage demand forecasting, mixed-integer linear programming, and fuzzy logic are employed to solve intricate optimization problems. Additionally, data augmentation techniques, including Stable Diffusion for synthetic image generation, enhance model robustness. Advanced heuristics (e.g., Harris-hawk optimization [6] or the walrus optimization algorithm), Markov decision processes, and explainable AI (XAI) techniques further enrich the AI toolkit. Emerging paradigms, such as digital twin frameworks, frequently utilized in 6G network optimization, and distributed fog (DF) computing, demonstrate the seamless integration of AI-driven solutions into broader computational ecosystems. The integration of quantum machine learning (QML) introduces quantum-enhanced optimization and classification capabilities, expanding AI's reach into domains such as cryptography, materials science, and drug discovery. Simulation of urban technology (SUMO) plays a crucial role in modeling and optimizing smart city infrastructures, with platforms like CityFlow enabling large-scale traffic simulations and COSMOS testbed providing a real-world environment for testing AI-driven networking and urban computing solutions. Ensuring model explainability is a crucial challenge; techniques such as Grad-CAM [7], LIME [8], and SHAP [9] provide insights into model decision-making processes, while privacy-by-design principles and biometric-based cryptography safeguard sensitive information. These approaches collectively ensure that AI remains robust, interpretable, and efficient across a wide range of real-world applications. Evaluating AI model performance requires context-dependent metrics. In classification tasks, metrics such as accuracy, precision, recall (sensitivity), specificity, F1-score (F1), Matthews correlation coefficient (MCC), and area under the ROC curve (AUC-ROC) capture various aspects of predictive performance. Accuracy

quantifies overall correctness, while precision measures the proportion of true positives among predicted positives. Recall (sensitivity) assesses the proportion of actual positives correctly identified, and specificity evaluates the proportion of actual negatives correctly classified. The F1-score provides a harmonic mean between precision and recall, whereas AUC-ROC quantifies the trade-off between the true positive rate (TPR) and false positive rate (FPR) across different classification thresholds. Additionally, metrics such as true negative rate (TNR), miss rate, false alarm rate (FAR), detection rate (DR), and average classification accuracy (ACA) offer further insights into classification performance. Additionally, MCC is also used to better quantify the real accuracy of the model, like F1-score, mainly used on binary classification tasks. Model calibration techniques, such as Brier score and expected calibration error (ECE), ensure reliable probability estimates in probabilistic classifiers. For regression tasks, key evaluation metrics include root mean square error (RMSE), normalized RMSE (NRMSE), mean absolute error (MAE), symmetric mean absolute percentage error (SMAPE), and coefficient of determination ( $R^2$ ). RMSE emphasizes larger deviations in predictions, while MAE provides an average measure of absolute prediction error. SMAPE normalizes percentage errors, and  $R^2$  represents the proportion of variance explained by the model. Beyond classification and regression, domain-specific and system-level performance indicators—such as average waiting time, average speed, traffic conflict rate, energy efficiency, carbon emissions, computing time, stock-keeping-oriented prediction error costs, device density, packet deadline adherence, latency, buffer size, wireless communication bandwidth, power consumption, event detection time, and processing throughput (e.g., FPS, CPU/GPU utilization)—ensure a comprehensive evaluation framework. Robustness metrics, including adversarial robustness scores, model uncertainty quantification (e.g., entropy-based measures), and fairness metrics such as demographic parity and equalized odds, contribute to assessing AI reliability and ethical compliance. These metrics account for real-world constraints, resource efficiency, and the broader operational impact of AI-based systems, ensuring sustainable and equitable deployment across industries.

### 3 SMART CITY

As is already mentioned, smart city is a vast field. In this review we will analyze current research from the following subfields in order: smart mobility, smart healthcare, smart energy, smart industry, smart living, smart safety and smart security.

#### 3.1 Smart Mobility

Smart mobility is a key part of smart cities. It helps make transportation more efficient, safer, and more sustainable. Papers considered in this review include traffic signal management based on congestion, connected autonomous vehicles, and their coordination and management of public transportation schedules and routes. Most of the proposed methods use some form of Deep RL (DRL) such as Deep Q Network [10–14], Multi-Agent Actor-Critic [15] or Proximal Policy Optimization [16]. In addition to DRL, some methods incorporated fuzzy logic [11], Mixed-Integer Linear programming [12], or Markov Decision Process [16] for additional optimization

or modeling. In other papers, MLP was combined with Convolutional layers [17] or LSTM [18].

### 3.1.1 Quality

The quality aspect is well explained. Most papers include some form of simulation. Therefore, the mentioned metrics are mostly empirical and are connected to sustainability and efficiency. The most frequent metric is average waiting time [10,11,13,14,16]. In [11], the average waiting time has three forms that are considered: average waiting time, average waiting time for a left turn, and average waiting time per car. Other mentioned metrics are energy efficiency [13,15,16], carbon emission [10,15] average speed [14] and traffic conflicts [10]. Papers that included only computational simulation used F1 [17,18] and computing time [12].

Additionally, the proposed methods were compared with other competent methods that the authors chose, including other forms of machine learning.

### 3.1.2 Measurability

Proposed methods are mostly evaluated through simulations and include multiple scenarios. The simulation frameworks used for simulations are Simulation of Urban Mobility [10,11,13,16], City Flow [14] and COSMOS testbed [19]. These frameworks supported the usage of modeling scenarios from the real world. In [12], authors chose computer simulation for evaluation, whereas a simulation was not conducted at all in [18]. The setup of the simulations mostly included datasets from real-world cities [10,11,13,14,18,19]. In [17], the authors tested cyber-attack detection using a Car-Hacking dataset containing data on real cyber-attacks. In [12], authors used real-world and generated datasets to broaden their simulation scenarios.

### 3.1.3 Explainability

The papers considered have used some form of DRL. Since DRL involves neural networks, its models are seen as black boxes. Authors mainly focused on testing their proposed methods without paying much attention to explainability. However, the ones that did use SHAP to find the correlation between decisions and input data values [14,16,17]. Authors of [14] observed how the number of vehicles in different lanes influences the next state of signal control. In [16], the authors showed how modeled factors affect the braking and acceleration of connected autonomous vehicles.

### 3.1.4 Privacy and Security

Privacy and security are considered as important in smart mobility as in other fields. However, not many practical implementations were mentioned. In [19], the authors designed and tested intelligent intersection nodes using a COSMOS testbed, which includes cameras for object detection. They blurred the faces of passengers and license plates.

In [17], authors proposed an explainable framework for detecting cyber-attacks in connected autonomous vehicles. The framework is designed to monitor vehicle networks and detect

anomalies in behavior. The authors of [20] have theoretically proposed a four-layer framework for detecting abnormalities in connected autonomous vehicle networks. The detection of cyber-attacks is based on changes in the energy that IoT sensors produce. For encryption, asymmetric methods like DSA, RSA, and ECDSA are considered.

## 3.2 Smart Healthcare

Smart healthcare leverages advanced technologies, such as AI, IoT, big data analytics, and cloud computing, to enhance the quality, efficiency, and accessibility of medical services. It addresses challenges like overwhelming data from electronic health records (EHRs), medical imaging, and wearable devices by using AI and ML to uncover hidden patterns and enable early intervention for chronic conditions. Additionally, smart healthcare bridges the gap in medical service access through remote monitoring, telemedicine, and AI-driven decision support, ensuring timely care regardless of location. This approach improves individual patient outcomes and can revolutionize public health by enhancing disease tracking, outbreak prediction, and resource allocation during crises.

### 3.2.1 Quality

AI, ML, and DL are being extensively applied in the healthcare field for disease diagnosis, patient monitoring, and disease imaging [21–23]. The work in [24] tackled the challenge of limited COVID-19 X-ray datasets by aggregating multiple sources and applying transfer learning with CNN architectures (VGG16, VGG19, MobileNet). VGG19 achieved the best performance with 96.97% accuracy, 99% recall, and 100% F1-score and precision. Authors in [25] used ML to predict type 2 diabetes and prediabetes risk in middle-aged individuals, assessing model performance with AUC-ROC and precision. The research in [26] utilized a decision tree classifier model with RMSE as the evaluation metric in patient monitoring. In contrast, an IoMT-enabled system for eldercare [27] achieved 93.6% training accuracy and 91.8% validation accuracy, with sensitivity, specificity, and precision being other metrics used in its evaluation. Medical image tasks, for example, the work in [28], developed deep learning models in the detection of brain tumors through MRI, utilizing accuracy, precision, recall, FPR, and TNR for the improvement of classification reliability.

### 3.2.2 Measurability

In disease diagnosis, the work in [24] validated their model using cross-validation, reducing bias and confirming generalizability. They compared VGG16, VGG19, and MobileNet to ensure the best model was not arbitrarily chosen. Similarly, authors in [25] used cross-validation and grid search, further validating their model on real-world data from the Stockholm Diabetes Preventive Program, which included over 8,000 participants, ensuring its practical applicability.

In patient monitoring, the work in [26] used RMSE to quantify prediction accuracy and tested multiple models for robustness. They benchmarked predictions against real-world EHRs to validate their decision tree-based model. The research proposed in [27] tested multiple ML techniques (SVM, KNN, Decision Tree, ANN), with ANN performing best. Their validation

included experiments across diverse datasets and benchmarking against existing studies to confirm superior performance.

For medical imaging, the work in [28] validated their deep CNN models for brain tumor detection by comparing results with previous studies and testing across two distinct MRI datasets. This approach ensured generalizability and reduced overfitting risks.

### 3.2.3 Explainability

Explainability in AI-driven healthcare is crucial due to its direct impact on patient outcomes, as errors can lead to severe medical consequences. Initially, AI models prioritized predictive accuracy, often relying on deep learning techniques with limited transparency. However, recent advancements have emphasized explainability to ensure model interpretability, allowing healthcare professionals to validate AI-generated insights and reduce risks. This growing importance is evident in Figure 1, which illustrates the increasing number of PubMed

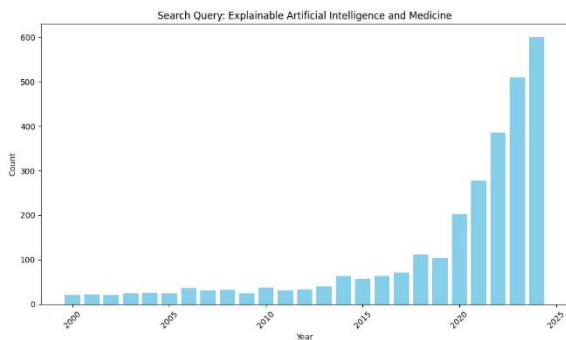


Figure 20 Number of papers on XAI in medicine over years

search results for "explainable AI and medicine" over the years, highlighting the rising interest in transparent AI models in healthcare.

The paper [29] reviewed the shift from black-box models to XAI in healthcare, highlighting techniques such as LIME and SHAP. They discussed toolkits like AI Explainability 360 and Alibi, which support practical XAI applications, and presented case studies involving classification models, deep learning explainability, and an early warning score system (XAI-EWS).

The work in [24] integrated CNNs with Grad-CAM [7] to interpret COVID-19 predictions from chest X-rays, generating heatmaps that identified critical lung regions, thereby enhancing radiologists' trust in model decisions. Similarly, in [25] employed SHAP to analyze diabetes risk based on electronic health records, identifying BMI and glucose levels as key predictive factors.

Authors in [30] proposed "HealthXAI", an XAI framework for early detection of cognitive decline using smart home sensor data. Their system combined collaborative data mining with XAI to detect behavioral and mobility anomalies associated with cognitive dysfunction. A key feature was its ability to generate natural language explanations via decision tree models, improving clinician interpretability. Validated on real-world data from 192 participants, the model demonstrated substantial predictive accuracy and scalability, offering a valuable tool for early diagnosis and intervention.

### 3.2.4 Privacy and Security

The increasing adoption of IoMT and edge computing enhances healthcare efficiency and exposes it to rising cybersecurity threats. The 2025 Cybersecurity Report by Check Point highlights a 47% increase in weekly attacks on healthcare organizations in 2024, making it the second most targeted industry. Groups like ALPHV exploit system vulnerabilities, as seen in the Change Healthcare attack, which disrupted hospital operations and forced ransom payments [31].

The authors in [32] analyzed privacy-preserving technologies, emphasizing cryptography, access control, anonymization, and blockchain. They highlighted regulatory gaps and proposed a privacy-by-design approach for integrating security. The work in [33] examined edge computing vulnerabilities, recommending blockchain and attribute-based encryption to protect patient data while advocating for federated learning to enhance security in decentralized systems.

The paper [34] explored security challenges in Healthcare IoT, particularly in remote patient monitoring. Their study stressed the risks of cyberattacks on interconnected medical devices, including infusion pumps and pacemakers, and emphasized strong encryption, blockchain, and biometric-based cryptography as key solutions. Federated learning was highlighted as a method to preserve data privacy while training AI models, ensuring compliance with HIPAA and privacy-by-design principles.

## 3.3 Smart Energy

Smart energy field aims to optimize energy management so that overall production and energy consumption lowers, possibly cutting the current cost [35]. Since energy is the foundation of urban areas, efficient and sustainable energy system is the prerequisite for future growth and progress of urban development.

### 3.3.1 Quality

Reviewed papers show models ranging from CNN-LSTM to LGBM and KNN. A key element that needs to be present in all of them is evaluating the quality of their approach.

A team [36] verified the quality of their approach by implementing their model in a number of cities worldwide, thus proving its effectiveness. However, despite the effectiveness of this approach, it is not always practical, and most researchers opt for some sort of metric.

In the case of [37], three proposed models used deep ANNS, CART DTs, and RF machine learning methods. They evaluated performance using NRMSE and SMAPE metrics to determine the best model.

In [38], researchers proved that enough energy is produced by solely using sunlight, even on days with very little to no sunlight. The usage of LGBM and KNN was proposed. By utilizing  $R^2$ , RMSE, and MAE, they validated the accuracy of their models. The metrics' values were presented clearly and organized, as well as training time and memory usage. In addition, they included the advantages and disadvantages of using suggested methods.

Spatiotemporal dependencies in energy management cannot be overlooked. Considering how they affect energy production and consumption, a model using CNN-LSTM architecture was recommended in [39]. MAE and  $R^2$  were used to evaluate the performance of the specified model.

Some models more precisely described the security aspect of energy management in smart cities. F1 score was used in most of them. Aside from it, in [40], they used FAR, precision, DR, accuracy, and recall. In [41], aside from the F1 score, accuracy, precision, recall, and AUC score were used to determine the accuracy of their model.

Authors in [35] decided to compare their results with those provided by methods already in action. It was established that energy consumption could be lowered by between 24% and 32%, while the cost of current could be lowered by 18.6% to 20.6%.

### 3.3.2 Measurability

Many papers use data collected from real-world situations to train models tested through various simulations. In the case of [36], sensors were deployed in various locations across the city to represent the energy required for the functioning of different areas. The work in [37] used data provided by the Croatian energy management information system. As mentioned, the authors of [38] considered using sunlight as their primary energy source. Thanks to the data used by the microgrid system from Thailand, simulations could be done on the proposed model. The research in [39] focused on getting data from cities with different climatic conditions, so they picked Mumbai, Delhi, Bengaluru, Chennai, and Kolkata in India, representing four different climate areas.

Different approaches were taken to this matter from the security aspect. The work in [40] relied on using an already-created physical system to run simulations in Matlab. The paper also gave an insight into what a cyber-attack would look like, with small but notable differences in voltage. It also includes the time necessary to resolve the cyber-attack.

In the case of [41], Kaggle provided the data for the model. It consisted of around 3000 instances separated into 12 classes. The model proved to measure the correct values by correctly predicting classes for all of the instances.

Sometimes, the explicit explanation of where the data is coming from or how measurability is proven is not detailed, like in the cases of [42] with a CNN model or [35], where they predicted the outdoor temperature and used PV-generated values for their models, which were later used for various simulations.

### 3.3.3 Explainability

Every resident utilizes AI systems for energy management in smart cities. The system must be explainable to ensure they understand why a specific decision was made. Unfortunately, not all models can be that easily explained, at least not without explaining the algorithm behind it. This was the case in most of the papers mentioned above. One of the exceptions was [40]. They used diagrams, similar to simulations, to explain the model's decision-making process. By detecting small oscillations in voltage or current, both the system and humans can tell when a cyber-attack is occurring. In the case of communication via a wireless connection, no explanation details were provided.

### 3.3.4 Privacy and Security

No matter how well a system performs, its security is paramount. An insecure system poses a potential threat to human lives and could result in losses of several million euros. The privacy and security aspect is marked as essential to every system in all papers. Addressing these issues strengthens people's trust. Some proposed two identification intrusion detection and encryption measures and security layers that provide guarded system communications [43].

Others proposed the usage of a DF network, based on FFT and DL, for the detection of cyber-attacks. Data shared amongst agents of production units and load is analyzed using FFT, which extracts coefficients and provides them as input indices to the DL model responsible for cyber-attack detection. Data shared via wireless communication, the amount of current and voltage that agents measured, and reference signals of controllers are the monitored parameters [40].

Hacking into a global server (GS) represents a significant problem for all local building energy management systems (LBEMS) connected to it. A selective parameter method is used in the federated reinforcement learning (FRL) training process of heating, ventilation, and air conditioning (HVAC) agents to prevent this. By applying this method, each agent chooses a random part of its local model and delivers it to GS, ensuring that hackers cannot retrieve important information, such as private energy data of buildings [35].

The team behind [41] described their model thoroughly, backing everything up with comprehensive mathematical arguments. SSAE classification is used for detecting even the smallest cyberattacks and real-time intrusion detection. WOA improves hyperparameters.

## 3.4 Smart Industry

This section covers Smart industry applications, focusing on product defect detection, worker safety compliance, and demand forecasting tasks. Research in this category relies heavily on classification models and occasionally employs regression approaches—particularly in demand forecasting [44]. While some studies provide broad comparisons of diverse methods [45], many concentrate on refining a single proposed solution [44,46].

### 3.4.1 Quality

Our literature review revealed that many studies primarily focused on classification [44–46]. This is predominantly due to the nature of the data, which is often categorical and requires classification to make sense of it. These include the classification of defective products and worker safety compliance. Neural networks dominate the classification tasks, while the regression task, particularly in demand forecasting, follows a two-stage (dual) approach. In this approach, a classification model first predicts whether the demand is zero or non-zero. If non-zero, a subsequent regression model estimates the precise demand quantity [44].

Regarding performance metrics for classification tasks, the most commonly reported metrics include accuracy, precision, recall, F1 score, specificity, and AUC [44–46]. These metrics are used to evaluate the models' performance and compare different models.

Precision and recall are both essential because their balance is vital in situations where the consequences of

misclassifications—either overlooking a risk or raising an unnecessary alert—can be substantial. In these cases, the F1 Score provides a unified measure when both metrics require equal weight. Specificity is crucial when accurately identifying negative instances is imperative, and the AUC offers broader insights into a model's overall discrimination ability under varying thresholds.

Demand forecasting [44] represents the sole regression application in the reviewed studies. The authors argue that Stock-keeping-oriented Prediction Error Costs are particularly suitable for this task because they capture the economic impact of forecasting errors on inventory management.

### 3.4.2 Measurability

Data used to train models within the smart industry consisted of labeled images and time series data, depending on the task. The labeled images were used for classification tasks, such as worker safety compliance [46], while time series data was used for demand forecasting [44]. The data was collected from real-world scenarios, such as steel manufacturing facilities and customer sales data. The image data was then preprocessed and augmented to increase the dataset size and variability. Data augmentation techniques included horizontal flipping, rotation, and image lightness and saturation modifications. In addition, synthetic data was generated via stable diffusion to diversify the training samples further[46].

The studies prevalently used k-fold cross-validation, with k ranging from 5 to 10 [44,46]. This technique was used to evaluate the model's performance and generalization ability. The data was split into training and validation sets, with the validation set used to tune the model's hyperparameters.

Authors of the [45] engaged in hyperparameter tuning to optimize the model's performance. They leveraged Genetic Algorithm (GA) to find the best hyperparameters for their model. This approach allowed them to find the optimal set of hyperparameters that maximized the model's performance.

### 3.4.3 Explainability

Concerning the explainability of the models, the authors of [44] propose the use of XAI techniques to interpret the model's predictions. They argue that the use of XAI techniques can provide insights into the model's decision-making process but must be used with caution due to the potential sensitivity of the data. The demand forecasting model explains its predictions within their proposed solution by highlighting the most influential features and their values. This lets the user understand the factors contributing to the model's predictions.

The authors of [47] propose a cognitive approach to model training to mimic human decision-making processes. This design, they argue, can offer enhanced interpretability by aligning the model's decision logic more closely with human reasoning.

### 3.4.4 Privacy and Security

To the best of our knowledge, no study explicitly addressed privacy concerns within the smart industry. Most regulations and

guidelines—such as GDPR and ePrivacy—are mentioned in the context of data collection and processing.

#### Security

Regarding security, none of the reviewed works explicitly delve into protective measures for smart industry applications. This highlights a notable gap, especially given the potential risks of cyber threats in networked industrial environments. However, the authors of [47] mention that the final judgment should be left to the human operator when dealing with decision-making models in worker safety compliance. This is due to the potential risks associated with fully automated decision-making systems, which could lead to catastrophic consequences in safety-critical environments.

## 3.5 Smart Living

This section covers Smart living applications, encompassing homes, education, and tourism. Research in this category spans from predicting customer socio-demographic characteristics[48] to assisting teachers in evaluating student performance[49]. Neural networks also dominate this area, with a few exceptions where other models are used, such as multinomial Naive Bayes [50]. In contrast to the smart industry, the smart living domain employs models such as LLMs.

### 3.5.1 Quality

In the reviewed literature, performance is predominantly quantified using classification metrics such as accuracy, precision, recall, and F1-score [50–52]. In the work focusing on customer social-demographic identification via smart meter data [48], the authors additionally report F1-macro and ACA to compare their channel attention architecture against other machine learning methods. Similarly, gesture recognition in smart homes [51] and sentiment analysis in tourism [50] report accuracy, precision, recall, and F1-score as core evaluation metrics.

The work on digital twins for 6G network optimization [53] further supplements accuracy with domain-specific measures, including device density, packet deadlines, latency, and buffer size, thereby capturing the nature of network performance.

In multi-stage problem formulations, a single metric cannot fully capture overall system performance. For instance, classification accuracy was employed as the principal performance measure in hyperparameter optimization for federated learning in online exam monitoring [52]. Meanwhile, hyperparameters, such as learning rate and number of epochs, were determined via an evolutionary approach combining PSO and GA.

The study on leveraging large language models for evaluating coloring activities [49] also measures performance by accuracy.

The study on prompt engineering for knowledge creation [54] assesses performance by comparing the depth and clarity of LLM-generated responses with original student discourse. The paper's findings emphasize improvements in response quality and depth when using Chain-of-Thought prompting for knowledge creation; however, they do not provide clear numerical measures of performance.

### 3.5.2 Measurability

The measurability of models performing tasks in smart living is primarily based on the data type and the nature of the task. For instance, in [48], the authors utilize time series data derived from smart meter readings across Irish households and businesses. The dataset is partitioned into training, validation, and testing subsets, ensuring robust performance evaluation and generalization analysis.

In the context of gesture recognition in smart homes [51], the authors accompanied classification metrics with system-level measures such as wireless communication bandwidth, power consumption, event detection time, and processing throughput (measured in frames per second as well as CPU/GPU utilization).

The sentiment analysis study [50] employs k-fold cross-validation (with a k value set to 10) to mitigate overfitting and ensure the reliability of results when analyzing YouTube comment data.

Additionally, the digital twin framework for 6G networks [53] quantifies operational parameters—such as latency and buffer size—using defined optimization formulas, bridging the gap between abstract model accuracy and real-world network performance.

Regarding the hyperparameter optimization for online exam monitoring [52], the authors conducted experiments with varying iteration counts. They examined the stability and convergence of the evolutionary search process, therefore providing quantitative evidence of the efficacy of their hybrid GA-PSO strategy.

Similarly, the study on LLM-based evaluation of coloring activities [49] introduces a custom dataset comprising 14 distinct problems that test both logical and spatial reasoning, with supplementary materials provided through a publicly accessible repository.

The study on prompt engineering for knowledge creation [54] ensures measurability by analyzing 721 discourse turns, with 272 complete question–answer pairs serving as the dataset to compare original student responses against LLM outputs.

### 3.5.3 Explainability

Explainability remains a critical concern. This is especially true in applications where model decision interpretability affects end-user trust and operational transparency. In [48], channel attention mechanisms are introduced within a fully convolutional network. This modification improves classification performance and enhances interpretability by highlighting relevant input time series data features. Although the approach is not entirely transparent, it offers insights into feature prioritization. These insights help in understanding the socio-demographic inference process.

For online exam monitoring [52], the authors note that using PSO and GA for hyperparameter tuning is less explainable due to the "black-box" nature of the evolutionary process. They partially mitigate this drawback by examining the model's sensitivity to specific hyperparameters. In addition, chain-of-thought (CoT) prompting is used in educational support [54] and logical task grading [49] to reveal the internal reasoning process of deep neural networks. Although still under development, this technique represents a promising direction for enhancing the transparency of complex model predictions.

### 3.5.4 Privacy and Security

Privacy considerations are notably sparse across the reviewed studies. The gesture recognition work [51] implicitly addresses privacy by advocating for on-device processing, thereby reducing the exposure of raw sensor data over networks. In the federated learning framework for online exam monitoring [52], data localization is designed to preserve privacy by keeping sensitive student data on local devices. Despite these measures, explicit privacy-preserving techniques are not directly addressed, suggesting a potential area for future enhancement.

The reviewed studies address security minimally. None of the contributions explicitly develop security measures, even in cases where system integrity and safe operation (such as in 6G network optimization [53]) are of great importance. In domains like smart home gesture recognition or smart meter analysis, the emphasis remains on performance and accuracy rather than defending against potential cyber threats. This highlights an opportunity for future work to integrate security frameworks into smart system applications.

### 3.6 Smart Safety

Public safety plays a crucial role in ensuring a better quality of life for citizens. As smart cities integrate IoT, AI, and big data analysis, there is room for improvement in monitoring, predicting, and mitigating risks to public safety. These technologies enable real-time surveillance and rapid response to emergencies, allowing for effective management of a wide variety of incidents, from criminal activities to natural disasters.

The tasks were mainly classification-based, and traditional methods such as K-NN, DT, and SVN were applied. Of the modern methods, CNNs and LSTMs were employed. There were also hybrid methods that combined two different methods for greater performance.

#### 3.6.1 Quality

In safety applications, the task is to identify or predict a state that threatens safety. In other words, this task is a classification task. As such, aside from simple accuracy, additional metrics are used. To better depict performances on disbalanced datasets F1 [55–62] and MCC [57,61]. Aside from them, to better understand the shortcomings of the trained model, along with the aforementioned metrics, some papers also produce various metrics from the confusion matrix, such as precision and recall (i.e. sensitivity) [55,57,59–62]. Some papers give additional metrics from the confusion matrix such as [57,62]. The best option is when the paper shows the confusion matrix [61,63].

Among the usual, there are also ROC and AUC [55,60]. Others are OOB [63], Kappa [59], paired t-test and Taylor diagram [64].

In a rare case of a regression task, paper [64] uses metrics  $R^2$ , RMSE, MSE, and MAE to display the accuracy of the air quality index prediction. In contrast, the same metrics are used by [65] to predict microbial concentration in water. Another regression task example is in [66], where authors predict pedestrian movement and use Displacement Error and Final Displacement Error metrics.

Apart from metrics that describe the accuracy of the model in various ways, sometimes the energy and time performances of the methods are constrained. These metrics are reported on a per-instance-classification basis. In papers [58,61,64], time per classification is introduced.

### 3.6.2 Measurability

For safety, there is a good trend toward using real-world data. All the papers on safety report real-world datasets; however, open access to the datasets still remains a topic that can be improved upon.

Directly openly available datasets, e.g., from Kaggle, Youtube or other sources, use papers [55–57,60,62,63].

In the paper [55], it states that the data will be made available upon request.

In the specific case of classifying several leaks in a pipe in [61], in the situation of missing real-world data, a dataset from a controlled real-life experiment was made, which is very close to a real-life scenario. Similar work was done in [65] for microbial concentration measurements.

Others do not state whether the dataset will be made available.

### 3.6.3 Explainability

For papers that use DT and derivative models, explainability is supported to a certain degree through the method itself [55,60]. In paper [55], explainability is additionally addressed through the gini index. In paper [65], SHAP is used to acquire explanations as to why the water is not drinkable.

Exceptionally, paper [56] uses a novel approach to explainability with the support of LLM to generate an explanation for the decision about the near-miss classification in traffic.

### 3.6.4 Privacy and Security

Privacy and security are usually not considered. Most information is mentioned in papers such as [57]. In some cases, this is not a problem if the data does not have privacy-sensitive data [55,58,60,61,64,65]. However, for the rest, this could be improved upon.

## 3.7 Smart Security

Securing IoT-based smart city infrastructures has emerged as a central research focus, with most work concentrating on detecting and mitigating cyberattacks, primarily DDoS. The application of various ML algorithms, including SVM, RF, KNN, NB, DT, GRB, XGBoost, ANN, CNN, LSTM, Autoencoder Classifier, QSVM, RBN, and ADA Boost, has proven to be highly effective in classifying incoming network traffic as benign or malicious. These studies report high overall detection accuracy and substantial reductions in false positives, instilling confidence in the progress of IoT security [67]. Some solutions extend beyond DDoS, targeting additional threats such as port mapping, brute-force attacks, MITM, malware, false data injection, buffer overflow, zero-day attacks, and SSH password guessing [68,69].

## 3.7.1 Quality

A recurring goal across these security solutions is achieving high detection accuracy with minimal latency. Papers commonly report performance metrics such as accuracy, recall, precision, F1, FPR, TPR, and computation time. Notably, XGBoost achieved 99.9% accuracy with an FPR of 0.06 [70], while QSVM reported accuracy rates between 99.1% and 99.4% [71].

## 3.7.2 Measurability

Most approaches are validated against established or publicly available datasets to confirm real-world applicability. Performance metrics like confusion matrices are frequently presented alongside comparative analyses with other methods. Where advanced paradigms are introduced, researchers measure reductions in computational overhead relative to classical baselines [71]. In honeypot-driven studies, comprehensive log data is collected to demonstrate the system's capacity to detect DDoS and secondary intrusion attempts (e.g., SSH password guessing) [70,72].

## 3.7.3 Explainability

While many studies provide mathematical explanations of ML algorithms, most solutions operate as black-box systems from an end user's perspective. Even when theoretical steps are detailed, non-technical users often find the outcomes opaque [67,68].

## 3.7.4 Privacy and Security

Privacy preservation is generally considered a secondary objective, with only a limited number of papers explicitly addressing it. One study anonymizes IP addresses at the data collection stage [67]. Another advocate for federated learning is ensuring that raw data remains at local sites while only aggregated parameters are transmitted to a central server, minimizing exposure risks [73].

On the question of security, several frameworks propose comprehensive strategies that go beyond detection accuracy and computational metrics. These strategies, such as specialized feature extraction pipelines that capture critical aspects of network traffic and the storage of metadata in a blockchain to ensure tamper resistance and trustworthy intrusion detection [74], instill confidence in the future of IoT security. Honeypot-driven approaches complement these strategies by analyzing large volumes of decoy traffic to uncover emerging threats [72].

Trust-based quarantining, as proposed in [75], enhances security by immediately isolating suspicious nodes, reducing the likelihood of large-scale breaches. Moreover, some research incorporates fog computing and blockchain technology, leveraging SHA-256 hashing and ECDS digital signatures to safeguard data integrity [74].

## 4 Conclusion

Current AI-based solutions in smart cities demonstrate promising results in classification and regression tasks, primarily supported by widely accepted metrics such as accuracy, precision, recall, F1, and RMSE. These quality metrics allow clear performance comparisons, but there is room for expanding multi-criteria evaluations to reflect real-world complexity better. Measurability often benefits from real-world datasets, which increases the practical relevance of findings; however, the scarcity of openly available, standardized data repositories continues to restrict reproducibility among different research groups. While emerging explainability methods like LIME, SHAP, and Grad-CAM show that explainability is gaining traction, many solutions remain opaque, indicating the importance of further research on transparent decision processes. Privacy concerns are variably addressed through anonymization and federated learning, yet comprehensive privacy-by-design approaches remain uncommon. Security, though frequently tackled in intrusion detection and related IoT-focused frameworks, would benefit from more integrated, end-to-end cybersecurity strategies. Bridging these gaps—particularly in data availability, explainable architectures, and holistic privacy-security measures—can reinforce public trust, accelerate adoption, and guide sustainable innovation in future smart cities.

## Acknowledgements

This research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-137/2025-03/200156) and the Faculty of Technical Sciences, University of Novi Sad through project “Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad 2025” (No. 01-50/295).

## 5 REFERENCES

- [1] *New York unveils “Midtown in Motion” traffic management system.* (2025, February 8). World Highways. <https://www.globalhighways.com/wh12/news/new-york-unveils-midtown-motion-traffic-management-system>
- [2] *Virtual Singapore – Building a 3D-Empowered Smart Nation—Geospatial World.* (n.d.). Retrieved February 23, 2025, from <https://geospatialworld.net/prime/case-study/national-mapping/virtual-singapore-building-a-3d-empowered-smart-nation/>
- [3] Smolyakov, V. (2024). *Machine Learning Algorithms in Depth* (1st ed). Manning Publications Co. LLC.
- [4] Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- [5] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems*, 30. [https://papers.nips.cc/paper\\_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html](https://papers.nips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html)
- [6] Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. *Future Generation Computer Systems*, 97, 849–872. <https://doi.org/10.1016/j.future.2019.02.028>
- [7] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *2017 IEEE International Conference on Computer Vision (ICCV)*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- [8] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- [9] Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (No. arXiv:1705.07874). arXiv. <https://doi.org/10.48550/arXiv.1705.07874>
- [10] Zhang, G., Chang, F., Jin, J., Yang, F., & Huang, H. (2024). Multi-objective deep reinforcement learning approach for adaptive traffic signal control system with concurrent optimization of safety, efficiency, and decarbonization at intersections. *Accident Analysis & Prevention*, 199, 107451. <https://doi.org/10.1016/j.aap.2023.107451>
- [11] Meepokgit, T., & Wisayataksin, S. (2024). Traffic Signal Control with State-Optimizing Deep Reinforcement Learning and Fuzzy Logic. *Applied Sciences*, 14(17), Article 17. <https://doi.org/10.3390/app14177908>
- [12] Yan, Y., Wen, H., Deng, Y., Chow, A. H. F., Wu, Q., & Kuo, Y.-H. (2024). A mixed-integer programming-based Q-learning approach for electric bus scheduling with multiple termini and service routes. *Transportation Research Part C: Emerging Technologies*, 162, 104570. <https://doi.org/10.1016/j.trc.2024.104570>
- [13] Li, D., Zhu, F., Wu, J., Wong, Y. D., & Chen, T. (2024). Managing mixed traffic at signalized intersections: An adaptive signal control and CAV coordination system based on deep reinforcement learning. *Expert Systems with Applications*, 238, 121959. <https://doi.org/10.1016/j.eswa.2023.121959>
- [14] Schreiber, L., Ramos, G., & Bazzan, A. (2021, July 24). Towards Explainable Deep Reinforcement Learning for Traffic Signal Control. *LatinX in AI at International Conference on Machine Learning 2021*. LatinX in AI at International Conference on Machine Learning 2021. <https://doi.org/10.52591/lxai2021072414>
- [15] Louati, A., Louati, H., Kariri, E., Neifar, W., Hassan, M. K., Khairi, M. H. H., Farahat, M. A., & El-Hoseny, H. M. (2024). Sustainable Smart Cities through Multi-Agent Reinforcement Learning-Based Cooperative Autonomous Vehicles. *Sustainability*, 16(5), Article 5. <https://doi.org/10.3390/su16051779>
- [16] Jiang, X., Zhang, J., & Wang, B. (2022). Energy-Efficient Driving for Adaptive Traffic Signal Control Environment via Explainable Reinforcement Learning. *Applied Sciences*, 12(11), Article 11. <https://doi.org/10.3390/app12115380>
- [17] Ding, W., Alrashdi, I., Hawash, H., & Abdel-Basset, M. (2024). DeepSecDrive: An explainable deep learning framework for real-time detection of cyberattack in in-vehicle networks. *Information Sciences*, 658, 120057. <https://doi.org/10.1016/j.ins.2023.120057>
- [18] Shabbir, A., Cheema, A. N., Ullah, I., Almanjahie, I. M., & Alshahrani, F. (2024). Smart City Traffic Management: Acoustic-Based Vehicle Detection Using Stacking-Based Ensemble Deep Learning Approach. *IEEE Access*, 12, 35947–35956. IEEE Access. <https://doi.org/10.1109/ACCESS.2024.3370867>
- [19] Kostić, Z., Angus, A., Yang, Z., Duan, Z., Seskar, I., Zussman, G., & Raychaudhuri, D. (2022). *Smart City Intersections: Intelligence Nodes for Future Metropolises* (No. arXiv:2205.01686). arXiv. <https://doi.org/10.48550/arXiv.2205.01686>
- [20] Algami, A., & Thayananthan, V. (2022). Autonomous Vehicles: The Cybersecurity Vulnerabilities and Countermeasures for Big Data Communication. *Symmetry*, 14(12), Article 12. <https://doi.org/10.3390/sym14122494>
- [21] Ghazal, T. M., Hasan, M. K., Alshurideh, M. T., Alzoubi, H. M., Ahmad, M., Akbar, S. S., Al Kurdi, B., & Akour, I. A. (2021). IoT for Smart Cities: Machine Learning Approaches in Smart

- Healthcare—A Review. *Future Internet*, 13(8), Article 8. <https://doi.org/10.3390/fi13080218>
- [22] Rahman, A., Debnath, T., Kundu, D., Khan, M. S. I., Aishi, A. A., Sazzad, S., Sayduzzaman, M., Band, S. S., Rahman, A., Debnath, T., Kundu, D., Khan, M. S. I., Aishi, A. A., Sazzad, S., Sayduzzaman, M., & Band, S. S. (2024). Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges and opportunities. *AIMS Public Health*, 11(1), Article publikealth-11-01-004. <https://doi.org/10.3934/publichealth.2024004>
- [23] Li, W., Chai, Y., Khan, F., Jan, S., Verma, P., Menon, V., . K., & Li, X. (2021). A Comprehensive Survey on Machine Learning-Based Big Data Analytics for IoT-Enabled Smart Healthcare System. *Mobile Networks and Applications*, 26. <https://doi.org/10.1007/s11036-020-01700-6>
- [24] Moujahid, H., Cherradi, B., Al-Sarem, M., Bahatti, L., Eljialy, A. B. A. M. Y., Alsaeedi, A., & Saeed, F. (2021). Combining CNN and Grad-Cam for COVID-19 Disease Prediction and Visual Explanation. *Intelligent Automation & Soft Computing*, 32(2), 723–745. <https://doi.org/10.32604/iasc.2022.022179>
- [25] Lama, L., Wilhelmsson, O., Norlander, E., Gustafsson, L., Lager, A., Tynelius, P., Wärvik, L., & Östenson, C.-G. (2021). Machine learning for prediction of diabetes risk in middle-aged Swedish people. *Heliyon*, 7(7). <https://doi.org/10.1016/j.heliyon.2021.e07419>
- [26] Asthana, S., Megahed, A., & Strong, R. (2017). A Recommendation System for Proactive Health Monitoring Using IoT and Wearable Technologies. *2017 IEEE International Conference on AI & Mobile Services (AIMS)*, 14–21. <https://doi.org/10.1109/AIMS.2017.11>
- [27] Khan, M. F., Ghazal, T. M., Said, R. A., Fatima, A., Abbas, S., Khan, M. a., Issa, G. F., Ahmad, M., & Khan, M. A. (2021). An IoT-Enabled Smart Healthcare Model to Monitor Elderly People Using Machine Learning Technique. *Computational Intelligence and Neuroscience*, 2021(1), 2487759. <https://doi.org/10.1155/2021/2487759>
- [28] Khan, M. S. I., Rahman, A., Debnath, T., Karim, M. R., Nasir, M. K., Band, S. S., Mosavi, A., & Dehzangi, I. (2022). Accurate brain tumor detection using deep convolutional neural network. *Computational and Structural Biotechnology Journal*, 20, 4733–4745. <https://doi.org/10.1016/j.csbj.2022.08.039>
- [29] Srinivasu, P. N., Sandhya, N., Jhaveri, R. H., & Raut, R. (2022). From Blackbox to Explainable AI in Healthcare: Existing Tools and Case Studies. *Mobile Information Systems*, 2022(1), 8167821. <https://doi.org/10.1155/2022/8167821>
- [30] Khodabandehloo, E., Riboni, D., & Alimohammadi, A. (2021). HealthXAI: Collaborative and explainable AI for supporting early diagnosis of cognitive decline. *Future Generation Computer Systems*, 116, 168–189. <https://doi.org/10.1016/j.future.2020.10.030>
- [31] *Cyber Security Report 2025*. (n.d.). Check Point Software. Retrieved February 24, 2025, from <https://www.checkpoint.com/security-report/>
- [32] El Majdoubi, D., El Bakkali, H., Sadki, S., Maqour, Z., & Leghmid, A. (2022). The Systematic Literature Review of Privacy-Preserving Solutions in Smart Healthcare Environment. *Security and Communication Networks*, 2022(1), 5642026. <https://doi.org/10.1155/2022/5642026>
- [33] Alzu'bi, A., Alomar, A., Alkhaza'leh, S., Abuarqoub, A., & Hammoudeh, M. (2024). A Review of Privacy and Security of Edge Computing in Smart Healthcare Systems: Issues, Challenges, and Research Directions. *Tsinghua Science and Technology*, 29(4), 1152–1180. <https://doi.org/10.26599/TST.2023.9010080>
- [34] Karunaratne, S. M., Saxena, N., & Khan, M. K. (2021). Security and Privacy in IoT Smart Healthcare. *IEEE Internet Computing*, 25(4), 37–48. <https://doi.org/10.1109/MIC.2021.3051675>
- [35] Lee, S., Xie, L., & Choi, D.-H. (2021). Privacy-Preserving Energy Management of a Shared Energy Storage System for Smart Buildings: A Federated Deep Reinforcement Learning Approach. *Sensors*, 21(14), Article 14. <https://doi.org/10.3390/s21144898>
- [36] Aljohani, A. (2024). Deep learning-based optimization of energy utilization in IoT-enabled smart cities: A pathway to sustainable development. *Energy Reports*, 12, 2946–2957. <https://doi.org/10.1016/j.egy.2024.08.075>
- [37] Zekić-Sušac, M., Mitrović, S., & Has, A. (2021). Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities. *International Journal of Information Management*, 58, 102074. <https://doi.org/10.1016/j.ijinfomgt.2020.102074>
- [38] Suanpang, P., & Jamjuntr, P. (2024). Machine Learning Models for Solar Power Generation Forecasting in Microgrid Application Implications for Smart Cities. *Sustainability*, 16(14), Article 14. <https://doi.org/10.3390/su16146087>
- [39] Shafiq, M., Bhavani, N. P. G., Venkata Naga Ramesh, J., Veerasha, R. K., Talasila, V., & Sulaiman Alfurhood, B. (2024). Thermal modeling and Machine learning for optimizing heat transfer in smart city infrastructure balancing energy efficiency and Climate Impact. *Thermal Science and Engineering Progress*, 54, 102868. <https://doi.org/10.1016/j.tsep.2024.102868>
- [40] Chang, Q., Ma, X., Chen, M., Gao, X., & Dehghani, M. (2021). A deep learning based secured energy management framework within a smart island. *Sustainable Cities and Society*, 70, 102938. <https://doi.org/10.1016/j.scs.2021.102938>
- [41] Ragab, M., Ashary, E. B., Alghamdi, B. M., Aboalela, R., Alsaadi, N., Maghrabi, L. A., & Allehaibi, K. H. (2025). Advanced artificial intelligence with federated learning framework for privacy-preserving cyberthreat detection in IoT-assisted sustainable smart cities. *Scientific Reports*, 15(1), 4470. <https://doi.org/10.1038/s41598-025-88843-2>
- [42] Sulaiman, A., Nagu, B., Kaur, G., Karuppaiah, P., Alshahrani, H., Reshan, M. S. A., AlYami, S., & Shaikh, A. (2023). Artificial Intelligence-Based Secured Power Grid Protocol for Smart City. *Sensors*, 23(19), Article 19. <https://doi.org/10.3390/s23198016>
- [43] Xin, Q., Alazab, M., Díaz, V. G., Montenegro-Marin, C. E., & Crespo, R. G. (2022). A deep learning architecture for power management in smart cities. *Energy Reports*, 8, 1568–1577. <https://doi.org/10.1016/j.egy.2021.12.053>
- [44] Rožanec, J. M., Novalija, I., Zajec, P., Kenda, K., Tavakoli, H., Suh, S., Veliou, E., Papamartzivanos, D., Giannetsos, T., Menesidou, S. A., Alonso, R., Cauti, N., Meloni, A., Recupero, D. R., Kyriazis, D., Sofianidis, G., Theodoropoulos, S., Fortuna, B., Mladenčić, D., & Soldatos, J. (2022). *Human-Centric Artificial Intelligence Architecture for Industry 5.0 Applications* (No. arXiv:2203.10794). <https://doi.org/10.48550/arXiv.2203.10794>
- [45] Ghahramani, M., Qiao, Y., Zhou, M., OHagan, A., & Sweeney, J. (2020). *AI-based Modeling and Data-driven Evaluation for Smart Manufacturing Processes* (No. arXiv:2008.12987). <https://doi.org/10.48550/arXiv.2008.12987>
- [46] Lan, R., Awolusi, I., & Cai, J. (2024). Computer Vision for Safety Management in the Steel Industry. *AI*, 5(3), Article 3. <https://doi.org/10.3390/ai5030058>
- [47] Sandini, G., Sciutti, A., & Morasso, P. (2024). Collaborative Robots with Cognitive Capabilities for Industry 4.0 and Beyond. *AI*, 5(4), Article 4. <https://doi.org/10.3390/ai5040092>
- [48] Luo, Z., Li, Q., Qi, R., & Zheng, J. (2025). Designing Channel Attention Fully Convolutional Networks with Neural Architecture Search for Customer Socio-Demographic Information Identification Using Smart Meter Data. *AI*, 6(1), Article 1. <https://doi.org/10.3390/ai6010009>
- [49] Tapia-Mandiola, S., & Araya, R. (2024). From Play to Understanding: Large Language Models in Logic and Spatial Reasoning Coloring Activities for Children. *AI*, 5(4), Article 4. <https://doi.org/10.3390/ai5040093>
- [50] Alzahrani, A., Alshehri, A., Alamri, M., & Alqithami, S. (2025). AI-Driven Innovations in Tourism: Developing a Hybrid Framework

- for the Saudi Tourism Sector. *AI*, 6(1), Article 1. <https://doi.org/10.3390/ai6010007>
- [51] Panagiotou, C., Faliagka, E., Antonopoulos, C. P., & Voros, N. (2025). Multidisciplinary ML Techniques on Gesture Recognition for People with Disabilities in a Smart Home Environment. *AI*, 6(1), Article 1. <https://doi.org/10.3390/ai6010017>
- [52] Shankarappa, R., Prasad, N., Guddeti, R. M. R., & Mohan, B. R. (2024). Bio-Inspired Hyperparameter Tuning of Federated Learning for Student Activity Recognition in Online Exam Environment. *AI*, 5(3), Article 3. <https://doi.org/10.3390/ai5030051>
- [53] Duran, K., Cakir, L. V., Ozdem, M., Gursu, K., & Canberk, B. (2024). *Generative AI-enabled Digital Twins for 6G-enhanced Smart Cities* (No. arXiv:2411.14222). arXiv. <https://doi.org/10.48550/arXiv.2411.14222>
- [54] Lee, A. V. Y., Teo, C. L., & Tan, S. C. (2024). Prompt Engineering for Knowledge Creation: Using Chain-of-Thought to Support Students' Improvable Ideas. *AI*, 5(3), Article 3. <https://doi.org/10.3390/ai5030069>
- [55] Yuan, H., Liu, Y., Huang, L., Liu, G., Chen, T., Su, G., & Dai, J. (2025). Real-time detection of urban gas pipeline leakage based on machine learning of IoT time-series data. *Measurement*, 242, 115937. <https://doi.org/10.1016/j.measurement.2024.115937>
- [56] Jaradat, S., Elhenawy, M., Ashqar, H. I., Paz, A., & Nayak, R. (2025). Leveraging Deep Learning and Multimodal Large Language Models for Near-Miss Detection Using Crowdsourced Videos. *IEEE Open Journal of the Computer Society*, 6, 223–235. IEEE Open Journal of the Computer Society. <https://doi.org/10.1109/OJCS.2025.3525560>
- [57] Dubey, P., Dubey, P., Iwendi, C., Biamba, C. N., & Rao, D. D. (2025). Enhanced IoT-Based Face Mask Detection Framework Using Optimized Deep Learning Models: A Hybrid Approach With Adaptive Algorithms. *IEEE Access*, 13, 17325–17339. IEEE Access. <https://doi.org/10.1109/ACCESS.2025.3532764>
- [58] Zhang, Z., Wu, J., Song, W., Zhuang, Y., Xu, Y., Ye, X., Shi, G., & Zhang, H. (2025). ARDs-YOLO: Intelligent detection of asphalt road damages and evaluation of pavement condition in complex scenarios. *Measurement*, 242, 115946. <https://doi.org/10.1016/j.measurement.2024.115946>
- [59] Khokhar, F. A., Shah, J. H., Saleem, R., & Masood, A. (2025). Harnessing deep learning for faster water quality assessment: Identifying bacterial contaminants in real time. *The Visual Computer*, 41(2), 1037–1048. <https://doi.org/10.1007/s00371-024-03382-7>
- [60] A Aldabagh, H., & Talal, R. (2025). Hybrid Intelligent Technique between Supervised and Unsupervised Machine Learning to Predict Water Quality. *International Journal of Computing and Digital Systems*, 17(1), 1–14. <https://doi.org/10.12785/ijcnds/1571031447>
- [61] Liu, H., Wang, N., Fang, H., Yu, X., & Du, W. (2025). Identification of the number of leaks in water supply pipes based on wavelet scattering network and Bi-LSTM model with Bayesian optimization. *Measurement*, 243, 116348. <https://doi.org/10.1016/j.measurement.2024.116348>
- [62] Reddy, P. D. K., Margala, M., Shankar, S. S., & Chakrabarti, P. (2024). Early fire danger monitoring system in smart cities using optimization-based deep learning techniques with artificial intelligence. *Journal of Reliable Intelligent Environments*, 10(2), 197–210. <https://doi.org/10.1007/s40860-024-00218-y>
- [63] Singh, V. B. P., Hemamalini, V., Muttipati, A. S., Raju, S. G., Shatil, A. H. M., & Sharma, A. (n.d.). Application of Machine Learning Predicting Injuries in Traffic Accidents through the Application of Random Forest. *Recent Patents on Engineering*, 19(2), 108–121. <https://doi.org/10.2174/0118722121248202231003064459>
- [64] Ansari, A., & Quaff, A. R. (2024). Advanced Machine Learning Techniques for Precise hourly Air Quality Index (AQI) Prediction in Azamgarh, India. *International Journal of Environmental Research*, 19(1), 15. <https://doi.org/10.1007/s41742-024-00684-5>
- [65] Yin, W.-X., Lv, J.-Q., Liu, S., Chen, J.-J., Wei, J., Ding, C., Yuan, Y., Bao, H.-X., Wang, H.-C., & Wang, A.-J. (2025). Microbial-Guided prediction of methane and sulfide production in Sewers: Integrating mechanistic models with Machine learning. *Bioresource Technology*, 415, 131640. <https://doi.org/10.1016/j.biortech.2024.131640>
- [66] Wang, R., Lam, S.-K., Wu, M., Hu, Z., Wang, C., & Wang, J. (2025). Destination intention estimation-based convolutional encoder-decoder for pedestrian trajectory multimodality forecast. *Measurement*, 239, 115470. <https://doi.org/10.1016/j.measurement.2024.115470>
- [67] Mohammed, B. H., Sallehudin, H., Satar, N. S. M., Murhg, H. D., Mohamed, S. A., Alaba, F. A., Rocha, A., & Bianchi, I. (2025). Anomaly Detection of Distributed Denial of Service (DDoS) in IoT Network Using Machine Learning. In R. Pereira, I. Bianchi, & A. Rocha (Eds.), *Digital Technologies and Transformation in Business, Industry and Organizations: Volume 3* (pp. 41–64). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-78412-5\\_3](https://doi.org/10.1007/978-3-031-78412-5_3)
- [68] An Intrusion Detection System Using a Machine Learning Approach in IOT-based Smart Cities. (2024). *ResearchGate*. <https://doi.org/10.58346/JISIS.2023.11.002>
- [69] Chohan, M. N., Haider, U., Ayub, M. Y., Shoukat, H., Bhatia, T. K., & Hassan, M. F. U. (2023). Detection of Cyber Attacks using Machine Learning based Intrusion Detection System for IoT Based Smart Cities. *EAI Endorsed Transactions on Smart Cities*, 7(2), Article 2. <https://doi.org/10.4108/eetsc.3222>
- [70] Alshahrani, M. M. (2023). A Secure and Intelligent Software-Defined Networking Framework for Future Smart Cities to Prevent DDoS Attack. *Applied Sciences*, 13(17), Article 17. <https://doi.org/10.3390/app13179822>
- [71] Said, D. (2023). Quantum Computing and Machine Learning for Cybersecurity: Distributed Denial of Service (DDoS) Attack Detection on Smart Micro-Grid. *Energies*, 16(8), Article 8. <https://doi.org/10.3390/en16083572>
- [72] *Securing smart cities through machine learning: A honeypot-driven approach to attack detection in Internet of Things ecosystems—Ahmed—2024—IET Smart Cities—Wiley Online Library*. (n.d.). Retrieved February 24, 2025, from <https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/smc2.12084>
- [73] Priyadarshini, I. (2024). Anomaly Detection of IoT Cyberattacks in Smart Cities Using Federated Learning and Split Learning. *Big Data and Cognitive Computing*, 8(3), Article 3. <https://doi.org/10.3390/bdcc8030021>
- [74] Ajao, L. A., & Apeh, S. T. (2023). Blockchain Integration with Machine Learning for Securing Fog Computing Vulnerability in Smart City Sustainability. *2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC)*, 1–6. <https://doi.org/10.1109/ICAISC56366.2023.10085192>
- [75] *SMART: A Secure Remote Sensing Solution for Smart Cities' Urban Areas | IEEE Journals & Magazine | IEEE Xplore*. (n.d.). Retrieved February 24, 2025, from <https://ieeexplore.ieee.org/document/10433153>

#### Contact information:

**Anja DELIĆ**, MSc, Teaching Assistant  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
delic.anja@uns.ac.rs  
<https://orcid.org/0009-0003-7506-3268>

**Jelena KOVAČ**, BSc, Teaching Assistant  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
kovac.e214.2024@uns.ac.rs  
<https://orcid.org/0009-0005-0287-4662>

**Nevena GLIGOROV**, BSc, Teaching Assistant  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
gligorov.e27.2024@uns.ac.rs  
<https://orcid.org/0009-0002-4366-060X>

**BranislavS. RISTIĆ**, MSc, Teaching Assistant  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
branislav.ristic@uns.ac.rs  
<https://orcid.org/0009-0004-1438-142X>

**Marko GORDIĆ**, Student  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
marko@gordic.rs  
<https://orcid.org/0009-0003-6306-4511>

**Radovan Turović**, MSc, Teaching Assistant  
Corresponding author  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
radovan.turovic@uns.ac.rs  
<https://orcid.org/0000-0003-2853-4875>

**Dinu DRAGAN**, PhD, Associate Professor  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
dinud@uns.ac.rs  
<https://orcid.org/0000-0001-8623-1923>

**Dušan GAJIĆ**, PhD, Associate Professor  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
dusan.gajic@uns.ac.rs  
<https://orcid.org/0000-0003-0495-8788>

**Veljko PETROVIĆ**, PhD, Docent  
University of Novi Sad, Faculty of Technical Sciences  
Trg Dositeja Obradovića 6, 21102 Novi Sad  
pveljko@uns.ac.rs  
<https://orcid.org/0000-0001-8662-1366>