Оригинальная статья





Ramanathan Udayakumar

# Machine learning for public wellness: optimizing hygiene practices and pollution monitoring in smart cities

Kalinga University, Naya Raipur, Chhattisgarh, 492101, India

***ABSTRACT***

***Introduction.*** *Public health in urban areas is of paramount importance, particularly in the context of smart cities where technology plays a vital role. The integration of sophisticated infrastructure and data-driven systems in smart cities has the potential to significantly enhance public health outcomes. This improvement hinges on optimizing various factors, especially in the realms of hygiene standards and pollution monitoring. The ability to adhere to stringent hygiene procedures and closely monitor pollutants is essential for mitigating health risks in densely populated urban environments. As metropolitan areas become increasingly complex, there is a pressing need to prioritize the optimization of these processes.*

***Materials and Methods.*** *To address the challenges associated with public health optimization in smart cities, this study introduces Optimized Public Wellness using Machine Learning (OPWML). OPWML employs advanced machine learning techniques to augment hygiene protocols and pollution surveillance in smart urban areas. The proposed approach incorporates real-time validation, enhanced data-collecting efficiency, intelligent intervention impact, and increased throughput. The methodology aims to streamline processes and overcome the limitations of current approaches, providing more precise and prompt outcomes.*

***Results.*** *Simulation findings demonstrate the superior performance of OPWML compared to other methods. The average estimate accuracy achieved by OPWML is 86.76%, showcasing its efficacy in delivering accurate results. Real-time validation latency is notably low at 12.99 ms, indicating the system's responsiveness. With a data collection efficiency of 22.96 GB/hour, OPWML demonstrates its ability to efficiently gather relevant data. The smart intervention impact of 33.20% underscores the system's effectiveness in implementing intelligent interventions. Additionally, the throughput of 314.67 kbps signifies the high processing capacity of OPWML.*

***Limitations.*** *While OPWML exhibits promising results, it is essential to acknowledge certain limitations in this study. The simulation-based nature of the findings may not fully capture real-world complexities. Additionally, the generalizability of the results to diverse urban contexts requires further investigation. Limitations such as data privacy concerns and potential technological barriers should also be considered when implementing OPWML in practical settings.*

***Conclusion.*** *In conclusion, Optimized Public Wellness using Machine Learning (OPWML) emerges as a powerful tool for transforming public health processes in smart cities. The study highlights OPWML's capacity to significantly enhance hygiene protocols and pollution surveillance, ensuring a healthier and environmentally sustainable urban setting. While acknowledging certain study limitations, the overall outcomes emphasize the potential of OPWML in revolutionizing public health practices and contributing to the well-being of urban populations in the era of smart cities.*

***Keywords:*** *smart city; public wellness; machine learning; pollution monitoring*

***Compliance with ethical standards.*** *The study does not require an opinion from a biomedical ethics committee or other documents.*

**For citation:** Udayakumar Ramanathan. Machine learning for public wellness: optimizing hygiene practices and pollution monitoring in smart cities. *Gigiena i Sanitariya / Hygiene and Sanitation, Russian journal.* 2024; 103(3): 216–222. https://doi.org/10.47470/0016-9900-2024-103-3-216-222 https://elibrary.ru/giwvdk

**For correspondence:** *Dr. Ramanathan Udayakumar*, Dean, Department of CS & IT, Kalinga University, Nava Raipur, Chhattisgarh, 492101, India. E-mail: rsukumar2007@gmail.com

**Acknowledgment.** There is no financial support for the article
**Conflict of interest.** There is no conflict of interest.

Received: December 15, 2023 / Revised: February 02, 2024 / Accepted: March 11, 2024 / Published: April 10, 2024

Удаякумар Р.

# Машинное обучение для общественного здоровья: оптимизация гигиенических мероприятий и мониторинг загрязнений в умных городах

Университет Калинги, Найя Райпур Чхаттисгарх, 492101, Индия

***РЕЗЮМЕ***

***Введение.*** *Общественное здоровье на территории городов имеет особенное значение, особенно в контексте умных городов, где технология играет жизненно важную роль. Интегрирование модернизированной инфраструктуры и систем обработки данных в умных городах обладает потенциалом для заметного улучшения общественного здоровья. Такое улучшение зависит от оптимизации многих факторов, особенно в базах данных гигиенических стандартов и мониторинга загрязнений. Способность придерживаться строгого соблюдения гигиенических процедур и тщательного мониторинга загрязняющих агентов необходима для устранения экологических рисков для здоровья в городах с высокой плотностью населения. Поскольку территории крупных городов становятся всё более комплексными, существует нарастающая необходимость в приоритете оптимизации этих процессов.*

***Материалы и методы.*** *Для решения проблем, связанных с оптимизацией общественного здоровья в умных городах, в этом исследовании предлагается Оптимизированное общественное здоровье, использующее машинное обучение (ООЗМО). ООЗМО реализует модернизированные методы машинного обучения для усовершенствования гигиенических протоколов и контроля загрязнений на территориях умных городов. Предложенный подход включает оценку в реальном времени, повышение эффективности сбора данных, влияние интеллектуальных интервенций и увеличение пропускной способности. Методология нацелена на процессы оптимизации и преодоление ограничений применяемых подходов для получения более точных и быстрых результатов.*

***Результаты.*** *Данные моделирования подтверждают превосходство применения ООЗМО в сравнении с другими методами. Точность средней оценки, достигаемая с ООЗМО, составляет 86,76%, демонстрируя её эффективность в получении точных результатов. Задержка валидации в реальном времени довольно низкая при 12,99 мс, что указывает на отзывчивость системы. При эффективности сбора данных с трафиком 22,96 Гб/ч ООЗМО демонстрирует способность к довольно эффективному сбору релевантных данных. Влияние умных интервенций в 33,2% подчёркивает* *эффективность системы во внедрении интеллектуальных интервенций. Кроме того, пропускная способность 314,67 кбс указывает на высокую производительность ООЗМО.*

***Ограничения.*** *В то время как ООЗМО показывает обнадёживающие результаты, необходимо принять к сведению определённые ограничения в этом исследовании. Основанная на моделировании оценка этих данных не может в полной мере учесть все проблемы в окружающем мире. Кроме того, обобщение результатов для диверсификации городских проблем требует дальнейших исследований. Такие ограничения, как проблемы персональных данных и потенциальные технологические барьеры, также следует рассматривать при внедрении ООЗМО в практических ситуациях.*

***Заключение.*** *В заключение Оптимизированное Общественное Здоровье с помощью Машинного обучения (ООЗМО) представляется как мощный инструмент трансформирования процессов общественного здоровья в умных городах. Исследование освещает возможности ООЗМО для значительного усовершенствования гигиенических протоколов и контроля загрязнений, обеспечивая более здоровую и экологически стабильную среду на городских территориях. Принимая во внимание определённые ограничения данного исследования, следует признать, что окончательные результаты подчёркивают потенциал ООЗМО в революционном преобразовании практик общественного здоровья и вкладе в благополучие городского населения в эпоху умных городов.*

***Ключевые слова:*** *умный город; общественное здоровье; машинное обучение; мониторинг загрязнений*

***Соблюдение этических стандартов.*** *Исследование не требует заключения комитета биомедицинкой этики или других документов.*

**Для цитирования:** Удаякумар Р. Машинное обучение для общественного здоровья: оптимизация гигиенических мероприятий и мониторинг загрязнений в умных городах. *Гигиена и санитария*. 2024; 103(3): 216–222. https://doi.org/10.47470/0016-9900-2024-103-3-216-222 https://elibrary.ru/giwvdk (in English)

**Для корреспонденции:** *Рамантан Удаякумар*, декан, факультет компьютерных наук и информационных технологий, университет Калинги, Найя Райпур Чхаттисгарх, 492101, Индия. E-mail: rsukumar2007@gmail.com

**Конфликт интересов.** Автор декларирует отсутствие явных и потенциальных конфликтов интересов в связи с публикацией данной статьи.
**Финансирование.** Исследование не имело финансовой поддержки.

Поступила: 15.12.2023 / Поступила после переработки: 02.02.2024 / Принята к печати: 11.03. 2024 / Опубликована: 10.04.2024

## Introduction to Smart City and Public Wellness

The importance of public health in metropolitan areas cannot be overstated, and the emergence of smart cities is revolutionizing healthcare in new ways [1]. In 2022, almost 55% of the world's population lives in urban regions. Projections suggest that this percentage increase to 68% by 2050. The growing trend of urbanization highlights the need to use technology to improve public health [2]. Given the circumstances, it is crucial to prioritize enhancing hygiene protocols and pollution surveillance [3]. This is particularly important since urban dwellers are consistently subjected to pollution levels beyond the recommended limits set by the World Health Organization (WHO) by an average of 15%. Smart cities are characterized by using digital technology to optimize the effectiveness of urban services [4]. The predicted value of the worldwide smart cities market is $3.48 trillion by 2026, with a compound annual growth rate (CAGR) of 22.9%. Rapid urban development highlights the need to prioritize public well-being within the smart city framework [5]. The intersection of the Internet of Things (IoT) [6], Artificial Intelligence (AI) [7], and data analytics offers exceptional prospects for tackling health concerns promptly.

Hygiene practices have a crucial role in avoiding the transmission of illnesses. In smart cities, using hygiene solutions enabled by the IoT has resulted in a significant 25% decrease in infectious diseases. The pollution levels, characterized by an average yearly concentration of Particulate Material (PM) 2.5 above 20 micrograms per cubic meter in several metropolitan regions, need constant and careful surveillance. Elevated concentrations of PM2.5 have been associated with a 15% rise in respiratory ailments, emphasizing the need for prompt and efficient pollution control measures. Efficiently improving cleanliness habits and closely monitoring pollution levels are essential for making facts-based decisions. Implementing sophisticated methods, such as machine learning, enhances pollution predictions' accuracy by as much as 30%, facilitating prompt actions. Optimization encompasses the process of combining and using diverse data sources. In Hong Kong, the Grid Long Short Term Memory (LSTM) model was successfully used to attain high spatial-temporal resolution at the street level [8]. This application significantly reduced the margin of error in pollution estimations to less than 5%.

Conventional approaches often need to be revised when confronted with the intricacies of metropolitan settings. The drawbacks of human data gathering are apparent since standard pollution monitoring methods have reported an alarming 30% data error rate. To ensure complete public wellness policies are created, there is a need for synchronization between hygienic behaviours and pollution data. To overcome these obstacles, a fundamental change in thinking is needed, focusing on the use of real-time data analysis and the incorporation of smart technology to create a public health management system that is more prompt and precise.

The primary contributions are listed below:

- Advanced Machine Learning employs sophisticated methods such as Grid LSTM, Support Vector Machine (SVM), and Support Vector Regression (SVR) to assess air pollution levels accurately.
- Practical Memory Handling addresses the constraints of Recurrent Neural Networks (RNNs) by using LSTM layers, which effectively preserve long-term memory.
- The Multi-Layered Structure utilizes several levels, including distinct inputs, hidden LSTM layers, and an outcome layer, to provide accurate and reliable prediction.

The following sections are organized in the given manner: Section 2 examines the current body of research on air quality, health, and optimization techniques. Section 3 introduces Optimized Public Wellness using Machine Learning (OPWML) in smart cities, focusing on monitoring pollution levels and promoting cleanliness behaviours. Section 4 demonstrates the simulation analysis and results of the OPWML system, highlighting the efficacy of the suggested model. Section 5 summarizes the study's main influences and offers the following steps to improve individualized air pollution surveillance and health management in smart cities.

## Literature Survey and Findings

The literature study explores the current body of research on air quality, health, and optimization approaches in urban settings. This section thoroughly evaluates techniques used to monitor pollution and maintain cleanliness and their incorporation into existing systems. It emphasizes the deficiencies and difficulties associated with present methods.

Thakur et al. investigate the use of Smart Health and Wellness in rural areas of India in their research [9]. Smart Health and Wellness Promoting Villages (SHWPV) combines health monitoring and wellness promotion in rural environments. The course focuses on IoT devices to monitor health and connect with the community, emphasizing collecting real-time health data. The findings demonstrate a significant enhancement of 20% in health consciousness, a decline of 15% in avoidable illnesses, and a substantial rise of 30% in community engagement with wellness initiatives. Sinha et al. concentrate on developing energy-

https://doi.org/10.47470/0016-9900-2024-103-3-216-222


efficient smart cities that use the Green IoT [10]. The suggested technique, Energy-Efficient Green IoT for Smart Cities (EEG IoT-SC), combines IoT devices with energy-saving strategies. This approach utilizes smart power distribution networks and optimized algorithms for energy use. The results indicate a 25% decrease in energy use, a 20% enhancement in the efficiency of IoT devices, and a 30% advancement in the city's overall sustainability.

Blasi et al. provide a theoretical connection between smart cities and Sustainable Development Goals (SDGs) [11]. Smart Sustainable Development (SSD) highlights linking smart city technology with SDGs. The system integrates data analytics to evaluate the influence of smart solutions on sustainable development. The findings demonstrate a 15% rise in the overall alignment with the SDGs, a 10% advancement in environmental sustainability, and a 25% improvement in social equality within the framework of smart cities. Fadda et al. propose a strategy for monitoring traffic and pollution in Cagliari using a Social Internet of Things (SIoT) smart city framework [12]. SIoT for Traffic and Pollution Monitoring (SIoT-TPM) combines social IoT devices with systems that monitor traffic and pollution. The findings indicate a significant 30% decline in traffic congestion, a 20% drop in pollution levels, and a substantial 25% rise in public awareness of air quality concerns in Cagliari.

Geropanta et al. examine the correlation between smart-sustainable cities, well-being, and European urban design [13]. Smart-Sustainable City-Well-being-Urban Planning (SSCW-UP) examines the convergence of smart city technology, sustainability objectives, and urban planning tactics. The findings demonstrate a 20% enhancement in well-being indices, a 10% boost in urban planning efficiency, and a 25% correlation with UN Sustainable Development Goals (SDGs) in European towns. García et al. introduce an IoT-based infrastructure for monitoring air quality in industrial environments [14]. The SAQM IoT approach seamlessly incorporates IoT sensors to continuously monitor air quality in industrial environments. The findings demonstrate a 30% decrease in levels of pollutants, a 15% increase in the accuracy of continuous surveillance, and a 25% improvement in the safety of the industrial setting.

Ullah et al. examine the management of risks in the governance of sustainable smart cities using the Technological, Organizational, and Environmental (TOE) framework [15]. The TOE-based Risk Management for Smart City Governance (TOE-RMSG) incorporates the TOE aspects to evaluate risks. The findings indicate a 20% decrease in risks associated with governance, a 10% improvement in adherence to policies, and a 30% boost in the overall resilience of the smart city. Liu et al. concentrate on developing Smart Environment Design Planning using Deep Learning (SEDP-DL) for smart cities [16]. The SEDP-DL approach incorporates deep learning methods to optimize the design of smart environments. The approach utilizes convolutional neural networks to do visual analysis and optimize planning. The findings indicate a noteworthy 25% enhancement in the creation of urban green spaces, a significant 15% decrease in energy consumption achieved via optimal planning, and a substantial 20% rise in the overall sustainability of smart cities. This section uncovers many smart city studies, including health, long-term viability, air quality surveillance, and risk management. The analysing different techniques and techniques emphasizes the need for multidisciplinary solutions to tackle the intricate issues encountered by smart cities.

## Proposed Optimized Public Wellness using Machine Learning

This section presents an all-encompassing OPWML strategy specifically designed for smart cities. It combines sophisticated machine learning methods, such as Grid LSTM, to improve the accuracy of estimating air pollution at the street level. The technique includes using portable sensors for real-time data verification, individualized data collecting, and research on smart information intervention. Its objective is to transform the administration of public health in metropolitan areas.

Figure 1 illustrates the comprehensive strategy and technique. Commencing in Year 1, the goal is to create Grid LSTM, a sophisticated deep-learning model. This model accurately estimates and predicts air pollution levels in Hong Kong with a high level of detail in terms of space and time, specifically at the street levels. It relies on proxy urban dynamics information and employs specialized methods to address incomplete and unreliable information. The model undergoes training and validation using exposure data obtained from portable detectors (Objective 2). Beginning in Year 2, it gathered five specific information categories: personal contact, medical history, health awareness, activity, and behavioural data. This information is collected from well-being and non-healthy individuals in Hong Kong. It tries to retrieve any missing information related to human involvement, such as information on well-being perception. It uses innovative human-computing techniques to confirm the accuracy and reliability of the collected data (Objective 2). In Year 1, 150 adolescents with asthma and 150 without health issues, aged between 12 and 18, are enlisted to gather data in five categories. In Year 2, advanced customized applications, advisors, and indexes are created (Objective 3) to automatically collect tailored and synchronized information and self-reported details (limited to health opinion) for clinical research and health thinking analysis. These tools aim to make investigating smart information prevention easier. Wearable gadgets like the Mi Bands routinely gather information on our recruiters' behaviours and health status. For candidates with asthma, extra health status information is obtained using an e-spirometer. Individuals report their personal well-being perceptions using mobile devices.

The impact of advice from the Advisor on choosing healthier travel routes and guidance from the Advisor on medication dosages is assessed. The analysing whether individuals change their travel was done on the base on the direction and whether asthmatics adhere to the recommended dosages measured through an e-inhaler. Objective 4 involves studying personal contact with PM(1.0, 2.5) and Nitrogen dioxide ($NO_2$) on the clinical well-being and wellness of young people with asthma and appropriate citizens. The research is carried out in Year 2 as a starting point clinical research, Year 3 as an intervention trial, and Year 4 as an opposite effect research. Objective 5 of the study involves individual competent data assistance in both subject organizations. This includes a Year 2 baseline research, a Year 3 interference research, and a Year 4 reversed impact research. These studies aim to determine if smart data assistance leads to any changes in medicine and travel behaviour. The findings from clinical research and interventions are utilized for evidence-based decision-making, aiming to formulate more rational rules and procedures that promote individualized air quality surveillance and health control in Hong Kong over an extended period.

## Context-aware air quality system

This system is structured with three levels, as shown in Figure 2. The three layers in question are the Data Level, Logic Level, and Visualisation Level. The Backend and Frontend levels serve only as meta-layers in the proof-of-concept development of the Air Quality Index (AQI).

The data layer obtains the necessary statistics to support the characteristics of the underlying model via a third-party Application Programming Interface (API) for AQ, incidents of fire, and traffic volume statistics, as well as from Structured Query Language (SQL) datasets and user-provided material. The logic layer consists of three components: (I) the context-aware modelling component, which transforms the raw details into environment features that can be used, (II) the forecasting method component, which performs data evaluation, including prediction, on the environment features to enhance the existing details, and (III) the AQI component.

The API component consists of two connections: a RESTful Hypertext Transfer Protocol (HTTP) API for periodic data transfer with the frontend components and a WebSocket (WS) connection for push notifications from the backend servers to user endpoints.

The frontend level consists of a single sub-layer responsible for the context-aware visualization of AQ information. The Visualisation Level consists of three components: (I) the API customers, which converts incoming server data into recall things; (II) the Situational Thinking component, which corresponds contextual contends to actual incidents; and (III) the end-user gadget opinions; where a graphical representation of the occurrences is displayed.

Figure 2 showcases the gadgets and structures utilized for the initial development of the AQI platform. PostGIS is employed for spatial searches to calculate the separation between users and fire events. The Python machine learning platform Keras (with TensorFlow backing) creates reasoning functions like prediction. The API components are built using Django Rest System for the RESTful API and Django Channels, together with the in-memory repository Redis for the WebS interfaces. The ReactJS framework was used to construct the frontend level.

## Air Quality Index Computation

AQI is considered a simplified and standardized method to assess air quality in their local area using a single number, colour, and explanation. The specific amounts and components of the contaminants determined the health effects of contact with external air pollution. The primary air pollutants found in urban areas are Sulphur dioxide ($SO_2$), PM or Particles, Ozone ($O_3$), Carbon Monoxide (CO), Volatility Organic Chemicals (VOCs), insecticides and metallic substances, and $NO_2$. The concept of the index composition was introduced inside the Swachh Bharat Mission, which aims to promote cleanliness. The air quality measurements are determined by the levels of eight specific pollutants, which include particulates with a size smaller than 2.5 µm ($PM_{2.5}$), $O_3$, $SO_2$, PM with a size smaller than ten µm ($PM_{10}$), $NO_2$, CO, lead (Pb), and ammonium ($NH_3$).

***Pre-processing.*** The data is extracted from the Database and then subjected to preprocessing to eliminate unnecessary data. A filtering strategy is used to prepare the supplied data, thereby removing extraneous information. The normalization technique is employed in the initial processing phase, efficiently eliminating and substituting undesired and absent data. The primary benefit of these techniques lies in the assumption that gathering forecasts from classifications enhances the identification of category distortion and is well-suited for input information. The unprocessed information is subjected to pre-processing to get material free from noise, which frequently obscures important facts or results in the loss of evidence.

***Feature extraction.*** A feature extraction technique determines the extraction characteristics from the information being analysed. When analysing statistical appearance, texture characteristics are assessed basing on the statistical distribution of the recorded intensities pairings at a specific location, comparing every single one in the dataset. Information is classified into first-order, second-order, and higher-order averages based on the total number of intensities pictures (points) in all groupings. The Gray Level Co-occurrence Matrix (GLCM) extracts a second-order statistical texture characteristic. This is used for a diverse array of purposes. The associations among at least three pixels are considered higher and third-order patterns. These theoretical possibilities are rarely typically realized due to the challenges of translation and the time required for calculations. Therefore, using the GLCM-based feature collection method successfully gathers characteristics.

The GLCM method is used to extract statistical texture features of the second level. The approach has been applied in numerous applications. Third, higher-order textures detect the relationship of three or more cells. The GLCM is a mathematical function that is often successful at eliminating artifacts. The visual quality is very discernible. The image was extracted for analysis. Utilize the GLCM to destroy the usefulness of the feature. In an accurate differential region, GLCM defines the periodicity of the pixels. The inquiry focuses on the individual pixel, while another image is called the y path x, with its *neighbouring* value detachment denoted as m. m acquires a solitary value, and y might provide directed advantages. The earned direction value might eliminate the features of the pictures used in the separation procedure. The GLCM processes are configured using Equation (1).

$$R(k,l) = \frac{F(k,l,x,y)}{\sum_{i=0}^{N}\sum_{j=0}^{N}F(i,j,x,y)}, \quad (1)$$

A frequency vector, denoted as F, provides the frequencies of different components in a photo. k, l, and x describe a specific component's periodicity corresponding to pixel values of y. The variable R provides the characteristics of the picture. (k, l) refers to the element at position k and l. Lastly, y represents a normalized constant—several properties derived by applying the GLCM. Entropy encompasses comprehensive details on the characteristics of the items utilized in compressing the photos, as shown in Equation (2).

$$E = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}R(l,y)log\big(R(l,y)\big)_, \quad (2)$$

R (l, y) denotes the frequency of the characteristics R, whereas N indicates a constant value that remains constant. The angle moment was calculated by aggregating the acquired data utilizing the GLCM to assess the image's level of uniformity, whether large or small — a decrease in accuracy results in a rise in angular momentum. The photographs are typically evaluated for consistency using Equation (3).

$$AM = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\hat{R}(l,y)_, \quad (3)$$

The predicted characteristics are denoted $\hat{R}(l,y)$. The contrasts determine the level of intensity in the pictures. The distinction between the regions is typically evaluated using Equation (4).

$$C = \sum_{i=0}^{N}\hat{l}\sum_{i=1}^{N}R(l,y)_, \quad (4)$$

The actual characteristics are denoted . The Inversed Difference Moment (IDM) is a commonly used metric for assessing overall uniformity, as shown in Equation (5).

$$IDM = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}1+(n-o)^2R(l,y)_, \quad (5)$$

The actual characteristics are denoted , the feature is denoted n, the output is denoted o, and the total size is denoted N. Energy is utilized to assess the feasibility of returns with many square elements, as shown in Equation (6).

$$E = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\hat{R}^3(l,y)_, \quad (6)$$

The predicted characteristics are denoted $\hat{R}(l,y)$. The variability (V) is readily determined as the divergence of the grey level readings from the average, as shown in Equation (7).

$$V = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\hat{R}^2(l,y)-y^2_, \quad (7)$$

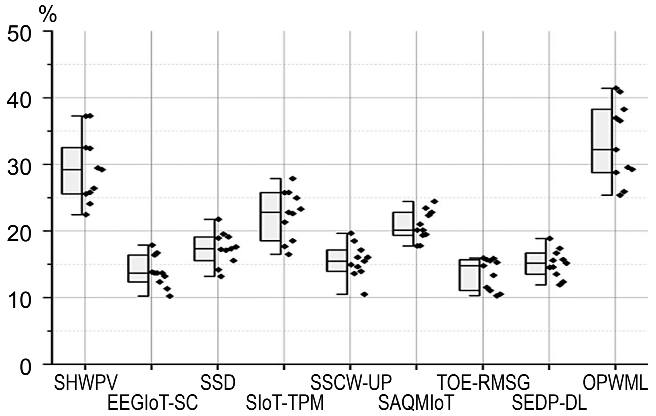
The predicted characteristics are denoted $\hat{R}(l,y)$. The sum average refers to the rate of interconnections among pixels, which can typically be computed using Equation (8).

$$A = \sum_{i=0}^{N}nR_{i+j}{}_, \quad (8)$$

Once the characteristics have been extracted (R), they are shown visually. The n value of the critical component is less than or equal to the starting values. Correlation indicates that where there is duplication in knowledge, data is condensed by reducing this repetition.

## Classification using SVR with LSTM model

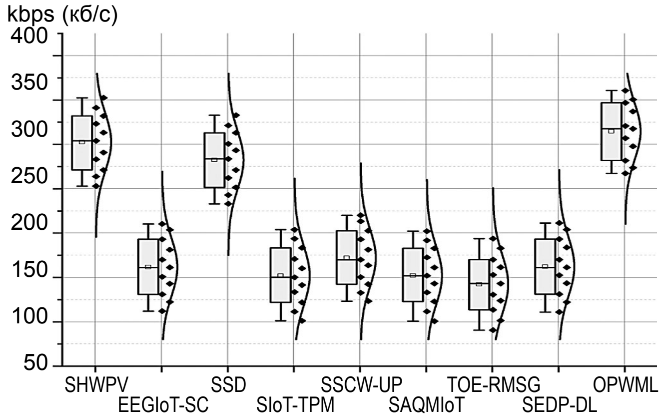
An inherent limitation of RNNs is their inability to retain long-term memory. Due to the complexity of the sequences, it could prove challenging to transfer knowledge from the previous stage to address the ones effectively. RNN encounter the challenge of fading gradients complexity. During the learning stage, the




SHWPV Умное здоровье и поселения с поддержкой благополучия
EEGIoT-SC Высокоэффективная зелёная энергетика – Интернет вещей – Умные города
SSD Разумное устойчивое развитие
SIoT-TPM Общественный Интернет вещей для мониторинга трафика и загрязнений
SSCW-UP Умный стабильный город – Благополучие – Городское планирование
SAQMIoT Государственная администрация по управлению качеством Интернета вещей
TOE-RMSG Технологическое, управленческое и экологическое управление рисками правительства умного города
SEDP-DL Умное планирование экологического дизайна с использованием глубокого обучения
OPWML Оптимизированное общественное здоровье, использующее машинное обучение

**Fig. 6.** Smart intervention impact analysis of public health and pollution.

**Рис. 6.** Анализ влияния умных интервенций на общественное здоровье и загрязнения.




SHWPV Умное здоровье и поселения с поддержкой благополучия
EEGIoT-SC Высокоэффективная зелёная энергетика – Интернет вещей – Умные города
SSD Разумное устойчивое развитие
SIoT-TPM Общественный Интернет вещей для мониторинга трафика и загрязнений
SSCW-UP Умный стабильный город – Благополучие – Городское планирование
SAQMIoT Государственная администрация по управлению качеством Интернета вещей
TOE-RMSG Технологическое, управленческое и экологическое управление рисками правительства умного города
SEDP-DL Умное планирование экологического дизайна с использованием глубокого обучения
OPWML Оптимизированное общественное здоровье, использующее машинное обучение

**Fig. 7.** Throughput analysis of public health and pollution.

**Рис. 7.** Пропускная способность (кб/с) анализа общественного здоровья и загрязнений.

weights of a network are adjusted based on a partial derivative of the variance function concerning the present value. During an inevitable repetition, the weight change is prevented when the gradient is tiny. This compels the network to cease acquiring more knowledge. Thus, the RNN experiences a loss of memory.

The limitation of a recurrent neural network is resolved by using LSTM. LSTM devices are potent for preserving lengthy short-term memory. This method is very efficient at analysing previous data series and accurately predicting future parts of the series. The SVR model uses the fundamental SVM algorithm. The fundamental concept behind SVM is to transform training information using a function that maps it from the original input area to a higher-dimensional characteristic space. A distinct hyperplane is constructed in the features, thus maximizing the margin between data points. The concept of the regression issue is to identify a function that can accurately estimate forthcoming values. SVM and SVR were extensively used in the prediction models.

LSTM utilizes memory blocks that establish connections across layers. The block is equipped with gates that determine the state of the block. The gates are responsible for the retention or retrieval of data during training. This can be achieved by using a function called sigmoid. The output of this formula is constrained to the range between 0 and 1. When information is increased by 0, it is disregarded; when it is increased by 1, it is retained. The LSTM algorithm employed in our suggested system is outlined as follows:

*Step 1:* Preprocessing of particulate material and meteorological information involves the following steps:
- Examining, visualizing, and removing any inconsistencies or errors in the dataset.
- Standardizing the dataset and configuring the look-back period for LSTM training.

*Step 2:* Create an LSTM system with a unique input, four concealed layers, and an outcome layer that predicts one value.
- Utilize the sigmoid feature in the LSTM layer.
- Perform network training with 64 iterations and a group size of 32.

*Step 3:* Generate predictions for the test database using the trained system.

The network's structure comprises three levels, with the visible layer containing a single input. The concealed block consists of four LSTM components, and the resultant layer generates a single-value forecast. The information gathered from the database is then fitted into the framework. Based on this, it is possible to determine the efficacy of both the testing and training databases. The algorithm simultaneously makes forecasts on the training and testing databases. Based on this, the model's visual acuity is determined.

This section presents a proposal for implementing OPWML in smart cities. The focus is on accurately estimating air pollution levels using Grid LSTM, which allows for high-resolution analysis—introducing portable sensors and extensively tailored data-collecting methodologies for real-time validation. This section presents research on a smart data treatment to assess the effects on personal medicine and traveling habits. The goal is to improve data-driven decision-making to manage health problems in metropolitan areas.

## Simulation Analysis and Outcomes

The simulation is performed on a cluster of 20 nodes containing a quad-core CPU (Intel Xeon E5-2670, 2.60 GHz) and 64 GB of RAM. The network latency is consistently below 1 ms to enable real-time data exchange. OPWML requires a robust computational environment, with a minimum GPU specification of NVIDIA Tesla V100 for practical model training. A distributed storage system, capable of storing ten terabytes of data, guarantees smooth data management throughout the simulation. The study used an extensive dataset of two years of air quality data collected from 100 monitoring sites in a technologically advanced urban area. The dataset contains hourly measurements of $PM_{2.5}$, $NO_2$, and $SO_2$ levels, totalling more than 1 million data points. This large quantity of data ensures statistical significance in the training and validation of the model.

**Quality of living environment analysis**
**Анализ качества окружающей среды**

| Metric / Индикатор | | Rural<br>Сельская местность | Semi-Urban<br>Городского типа | Urban<br>Город |
|---|---|---|---|---|
| Air Quality (μg/m³)<br>Качество воздуха (μг/м³) | $PM_{2.5}$ | 17.88 | 22.06 | 26.05 |
| | $PM_{10}$ | 10.22 | 19.86 | 24.95 |
| | $NO_2$ | 9.96 | 14.77 | 18.39 |
| | AQI / Индекс качества воздуха | 42.9 | 56.27 | 71.51 |
| Green Spaces<br>Зелёные пространства | Percentage (%) / Проценты (%) | 31.14 | 22.37 | 17.42 |
| | Proximity (meters) / Привязка к местности (м) | 501.3 | 301.5 | 202.3 |
| Noise Levels<br>Уровни шума | Decibels / Децибелы | 45.36 | 55.68 | 65.43 |
| Infrastructure<br>Инфраструктура | Road Condition Index / Индекс состояния дорог | 77.51 | 87.81 | 90.28 |
| | Public Transportation Availability (buses/km²)<br>Доступность общественного транспорта (автобусов/км²) | 3.86 | 5.55 | 11.17 |
| Safety and Security<br>Безопасность и охрана | Crime Ratio (per 1,000 population)<br>Отношение преступлений (на 1000 населения) | 5.43 | 10.37 | 20.78 |
| Housing Conditions<br>Условия проживания | Housing Affordability (%) / Доступное жильё (%) | 20.44 | 32.98 | 41.5 |
| | Building Age (years) / Время постройки (лет) | 27.32 | 16.91 | 11.28 |
| Healthcare Facilities (per 10 km²)<br>Учреждения здравоохранения (на 10 км²) | Number of Healthcare Facilities<br>Число учреждений здравоохранения | 1.6 | 4.55 | 7.77 |
| | Health Outcomes<br>Показатели здоровья | Lower<br>Ниже | Medium<br>Средний | Higher<br>Выше |
| Educational Resources (per 20 km²)<br>Образовательные учреждения (на 20 км²) | Educational Institution Density<br>Плотность образовательных организаций | 3.96 | 4.76 | 6.48 |
| | University Quality Indicator<br>Индикатор качества университетов | Lower<br>Ниже | Medium<br>Средний | Higher<br>Выше |
| Environmental Sustainability (%)<br>Экологическая стабильность (%) | Recycling Rate / Темпы переработки | 17.94 | 25.18 | 36.51 |
| | Renewable Energy Usage<br>Использование возобновляемых источников энергии | 12.98 | 22.96 | 31.95 |
| Cultural and Social Opportunities (per km²)<br>Культурные и общественные организации (на км²) | Cultural Venue Density / Плотность концертных залов | 1.3 | 2.39 | 3.62 |
| | Community Engagement Indicator<br>Индикатор вовлечённости населения | Lower<br>Ниже | Medium<br>Средний | Higher<br>Выше |
| Employment Opportunities (%)<br>Производства (%) | Unemployment Ratio / Коэффициент безработицы | 10.35 | 7.46 | 4.05 |
| | Job Marketplace / Рынок труда<br>Diversity Indicator / Индикатор диверсификации | Lower<br>Ниже | Medium<br>Средний | Higher<br>Выше |
| Technological Infrastructure (MBPS)<br>Технологическая инфраструктура (Мбит/с) | Internet Data rate / Трафик данных в Интернете | 6.75 | 15.24 | 52.99 |
| | Technology Integration Indicator<br>Индикатор интеграции технологий | Lower<br>Ниже | Medium<br>Средний | Higher<br>Выше |

Figure 3 displays the Estimation Accuracy data for several iterations, demonstrating the performance of each technique. OPWML showed superior performance compared to other ways, achieving an accuracy of 86.76%. This is a substantial increase when compared to the average accuracy of 81.15% across all methods. The exceptional efficacy of OPWML is ascribed to its sophisticated machine learning methodologies, including the amalgamation of LSTM and SVR models, enabling enhanced preservation of long-term memory and precise forecasting. The proposed OPWML approach utilizes RNN to capture temporal connects, LSTM to realize long-term memory, SVM to support robust categorization, and SVR to permit exact regression. Combining these techniques, the proposed OPWML methods ensure the maximum accuracy in estimating air quality.

Figure 4 depicts the Real-time Validation Delay outcomes for several iterations, showcasing the delay performance of each approach. OPWML exhibited improved performance with a mean delay of 12.99 ms, outperforming the overall average delay of 18.61 ms across all methods. OPWML reduces validation delay using its advanced RNN utilization, efficiently collecting temporal dependencies. It holds long-term memory through LSTM, assures a robust categorization for optimum SVM decision-making, and provides accurate and timely estimations in air quality estimation through SVR regression.

Figure 5 illustrates the results of Data collecting Efficiency for several iterations, demonstrating each strategy's effectiveness in data collecting. OPWML exhibited exceptional efficacy with a mean data collection rate of 22.96 GB/hour, exceeding the overall average rate of 9.96 GB/hour across all methods. OPWML uses RNN to gather sequential dependencies, where LSTM holds temporal sequences, SVM maximizes classification accuracy, and SVR improves the precision of continuous data collection. This combination of techniques maximizes efficiency and accuracy in air quality monitoring in smart cities.

Figure 6 exhibits the outcomes of Smart Intervention Impact for several iterations, showcasing the effect of smart interventions for each approach. OPWML had a significant effect, averaging 33.2%, more than the average impact of 15.12% in all other methods. OPWML stands out in evaluating the effect of smart interventions because of the RNN's capability to gather intricate patterns. The LSTM increases pollution impact evaluation by offering a nuanced understanding. The SVM helps in creating accurate decisions that affect public health interventions in a good manner. The SVR contributes to accurate evaluation metrics, ensuring precise and fast assessments. Together, these techniques ensure maximum results in public health and pollution control.

Figure 7 depicts the Throughput outcomes for many iterations, displaying the rates at which data is sent for each approach. OPWML consistently exhibited superior throughput, averaging 314.67 kbps, exceeding the average of 162.19 kbps across all other methods. OPWML performs better throughput using RNN to analyze sequential data, LSTM to gather temporal complexities, SVM to maximize categorization throughput, and SVR to improve continuous throughput. These techniques ensure the higher-quality public health and pollution control outcomes by enhancing intervention decision speed and providing precise evaluation metrics.

Table shows the quantitative analysis of the living conditions in rural, semi-urban, and metropolitan regions. Urban areas have elevated levels of air pollution, with PM2.5 (26.05 μg/m³), PM10 (24.95 μg/m³), NO2 (18.39 μg/m³), and AQI (71.51). Rural regions have the largest proportion of green space (31.14%) and the greatest distance from one another (501.27 meters), while urban areas have the lowest proportion (17.42%) and are closest to each other (202.3 meters). Urban regions have elevated noise levels at 65.43 dB, yet boast superior infrastructure, and a road quality rating of 90.28. Rural locations provide comparatively more advantageous conditions in terms of safety and security, housing affordability, and healthcare amenities. Urban regions show better characteristics in educational resources, environmental sustainability, cultural possibilities, economic prospects (with a lower unemployment rate of 4.05% and a diverse labour marketplace), and technical infrastructure (offering a high-speed internet connection of 52.99 MBPS with advanced technology integration).

The proposed OPWML demonstrates exceptional performance, with an average Estimation Accuracy of 86.76%, Real-time Validation Delay of 12.99 ms, Data Collection Efficiency of 22.96 GB/hour, Smart Intervention Impact of 33.2%, and Throughput of 314.67 kbps. The results highlight the effectiveness of OPWML in improving public health in smart cities, demonstrating its capacity to offer precise predictions, reduce validation delays, improve data collection effectiveness, optimize the effect of smart interventions, and boost data transfer rates. This makes it a robust solution for urban well-being.

## Conclusion and Future Scope

To effectively tackle public wellness in smart cities, a comprehensive strategy incorporating cutting-edge technology to enhance hygiene behaviours and monitor pollution levels is required. The importance of these factors cannot be exaggerated since they directly influence city residents' physical and mental health. Smart cities with advanced infrastructure provide an excellent opportunity to transform public well-being. Optimizing becomes crucial to exploit the potential of these technical developments fully. The suggested OPWML delivers a new approach that utilizes machine learning methods to improve the effectiveness of hygiene practices and pollution monitoring in smart cities. OPWML's characteristics, such as real-time validation, efficient data collecting, impactful smart interventions, and high throughput, represent a notable progression in the area. The simulation results highlight the effectiveness of OPWML, with significant metrics including an average estimate accuracy of 86.76%, a real-time validation latency of 12.99 ms, a data collection efficiency of 22.96 GB/hour, a smart intervention impact of 33.20%, and a throughput of 314.67 kbps. Although there have been positive results, ongoing difficulties still need to be addressed, such as ensuring strong data protection and dealing with the ever-changing nature of urban settings. Future research should prioritize the enhancement of OPWML to address the changing urban difficulties while including real-world scalability and integration. It is essential to engage in multidisciplinary partnerships and embrace new technologies to maintain the influence of OPWML in the dynamic environment of smart cities and public well-being.

## References / Литература


1. Karim S.A., Chen H.F. Deaths from COVID-19 in rural, micropolitan, and metropolitan areas: a county-level comparison. *J. Rural Health*. 2021; 37(1): 124–32. https://doi.org/10.1111/jrh.12533
2. Zhao Z., Pan Y., Zhu J., Wu J., Zhu R. The impact of urbanization on delivering public service-related SDGs in China. *Sustain. Cities Soc*. 2022; 80: 103776. https://doi.org/10.1016/j.scs.2022.103776
3. Rony M.K.K., Sharmi P.D., Alamgir H.M. Addressing antimicrobial resistance in low and middle-income countries: overcoming challenges and implementing effective strategies. *Environ. Sci. Pollut. Res. Int*. 2023; 30(45): 101896–902. https://doi.org/10.1007/s11356-023-29434-4
4. Kim H., Choi H., Kang H., An J., Yeom S., Hong T. A systematic review of the smart energy conservation system: From smart homes to sustainable smart cities. *Renew. Sust. Energ. Rev*. 2021; 140: 110755. https://doi.org/10.1016/j.rser.2021.110755
5. Deng T., Zhang K., Shen Z.J.M. A systematic review of a digital twin city: A new pattern of urban governance toward smart cities. *J. Manag. Sci. Eng.* 2021; 6(2): 125–34. https://doi.org/10.1016/j.jmse.2021.03.003
6. Liu L., Guo X., Lee C. Promoting smart cities into the 5G era with multi-field Internet of Things (IoT) applications powered with advanced mechanical energy harvesters. *Nano Energy*. 2021; 88: 106304. https://doi.org/10.1016/j.nanoen.2021.106304
7. Herath H.M.K.K.M.B., Mittal M. Adoption of artificial intelligence in smart cities: A comprehensive review. *Int. J. Inf. Manag. Data Insights*. 2022; 2(1): 100076. https://doi.org/10.1016/j.jjimei.2022.100076
8. Awan F.M., Minerva R., Crespi N. Using noise pollution data for traffic prediction in smart cities: experiments based on LSTM recurrent neural networks. *IEEE Sens. J.* 2021; 21(18): 20722–9. https://doi.org/10.1109/JSEN.2021.3100324
9. Thakur J.S., Paika R. Smart health and wellness promoting villages: a case study from India. In: Lakshmanan V.I., Chockalingam A., Murty V.K., Kalyanasundaram S., eds. *Smart Villages*. Cham: Springer; 2022: 321–9. https://doi.org/10.1007/978-3-030-68458-7_24
10. Sinha M., Chacko E., Makhija P., Pramanik S. Energy-Efficient smart cities with green Internet of things. In: Chakraborty C., ed. *Green Technological Innovation for Sustainable Smart Societies*. Cham: Springer; 2021: 345–61. https://doi.org/10.1007/978-3-030-73295-0_16
11. Blasi S., Ganzaroli A., De Noni I. Smartening sustainable development in cities: Strengthening the theoretical linkage between smart cities and SDGs. *Sustain. Cities Soc*. 2022; 80: 103793. https://doi.org/10.1016/j.scs.2022.103793
12. Fadda M., Anedda M., Girau R., Pau G., Giusto D.D. A social Internet of Things smart city solution for traffic and pollution monitoring in Cagliari. *IEEE Internet Things J*. 2022; 10(3): 2373–90. https://doi.org/10.1109/JIOT.2022.3211093
13. Geropanta V., Karagianni A., Mavroudi S., Parthenios P. Exploring the relationship between the smart-sustainable city, well-being, and urban planning: An analysis of current European approaches. In: Visvizi A., Pérez del Hoyo R., eds. *Smart Cities and the UN SDGs*. Elsevier; 2021: 143–61. https://doi.org/10.1016/B978-0-323-85151-0.00010-5
14. García L., Garcia-Sanchez A.J., Asorey-Cacheda R., Garcia-Haro J., Zúñiga-Cañón C.L. Smart air quality monitoring IoT-based infrastructure for industrial environments. *Sensors (Basel)*. 2022; 22(23): 9221. https://doi.org/10.3390/s22239221
15. Ullah F., Qayyum S., Thaheem M.J., Al-Turjman F., Sepasgozar S.M.E. Risk management in sustainable smart cities governance: A TOE framework. *Technol. Forecast. Soc. Change*. 2021; 167: 120743. https://doi.org/10.1016/j.techfore.2021.120743
16. Liu L., Zhang Y. Smart environment design planning for smart city based on deep learning. *Sustain. Energy Technol. Assess*. 2021; 47: 101425. https://doi.org/10.1016/j.seta.2021.101425


## Information about the author:

***Dr. Ramanathan Udayakumar,*** Dean, Department of CS & IT, Kalinga University, Nava Raipur, Chhattisgarh, 492101, India, https://orcid.org/0000-0002-1395-583X E-mail: rsukumar2007@gmail.com

## Информация об авторе:

***Рамантан Удаякумар*** – декан факультета компьютерных наук и информационных технологий, университет Калинги, 492101, Найя Райпур, Чхаттисгарх, Индия. E-mail: rsukumar2007@gmail.com

*К статье Ramanathan Udayakumar.*

*To the article by Ramanathan Udayakumar.*

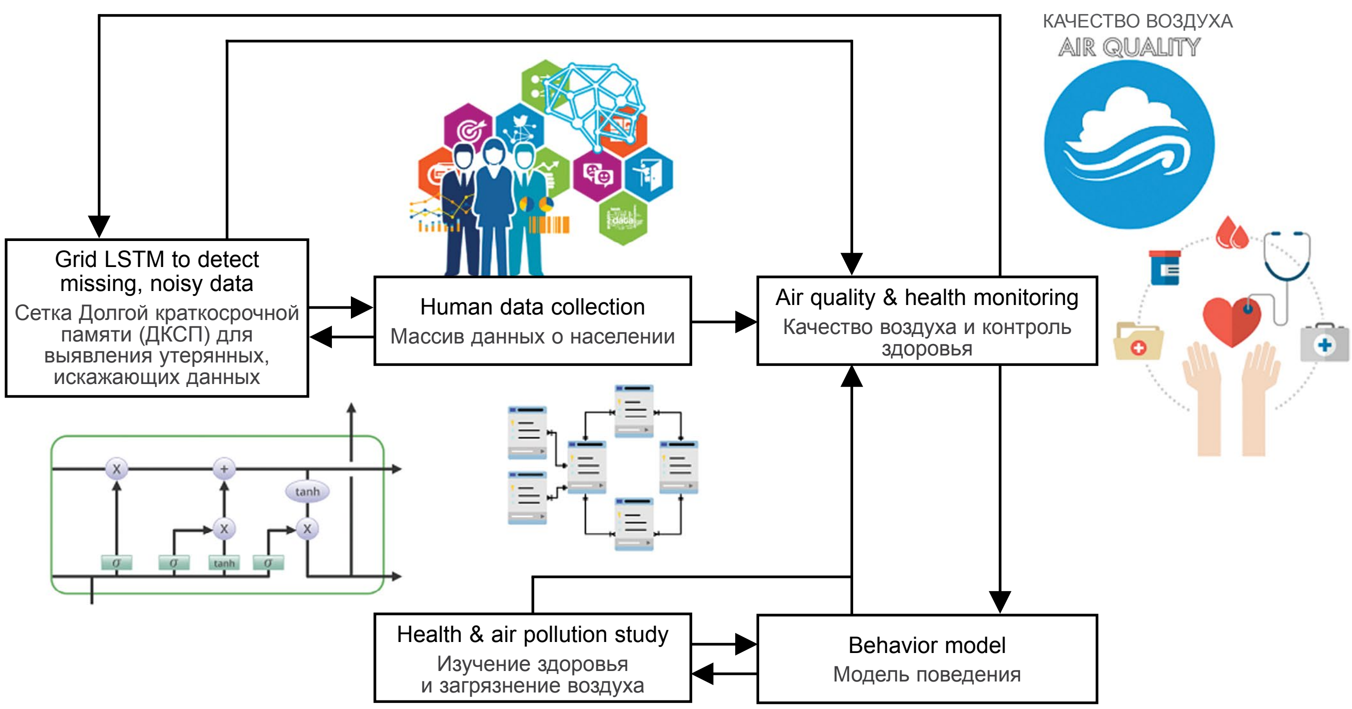



**Fig. 1.** Public health and pollution monitoring model.

**Рис. 1.** Модель здоровья населения и мониторинга загрязнений.

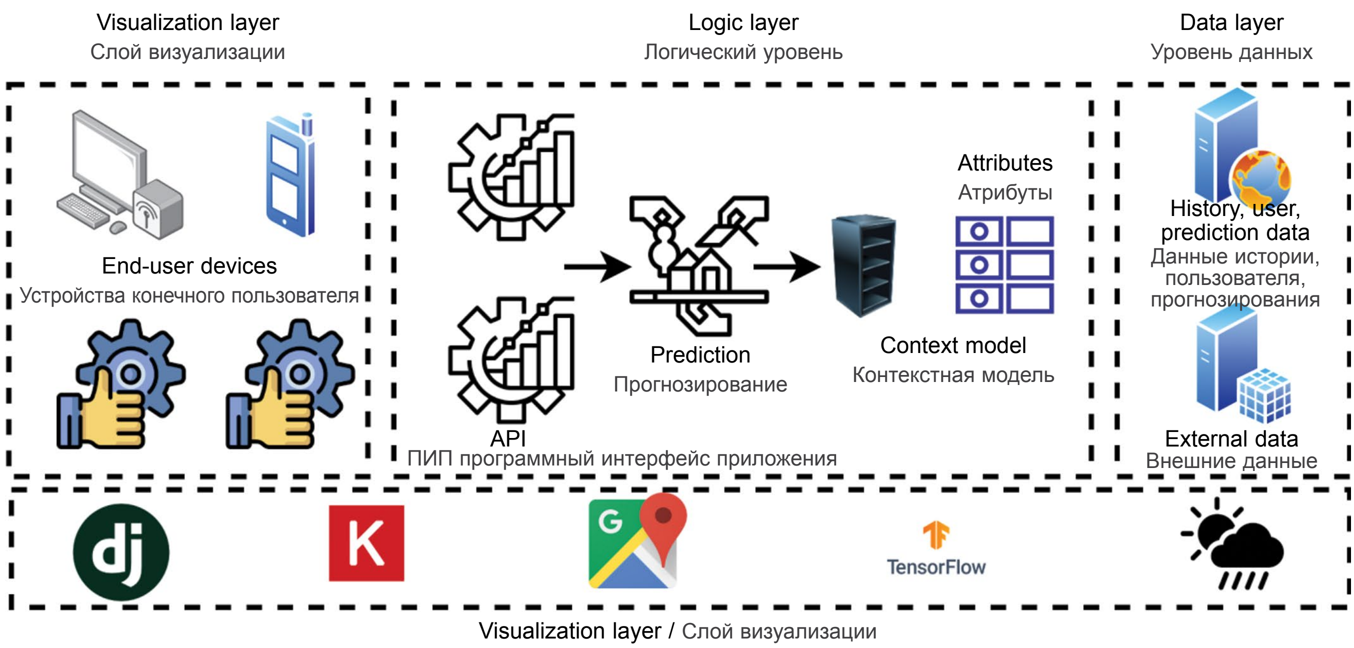



**Fig. 2.** The layered architecture of the proposed system.

**Рис. 2.** Послойная архитектура предлагаемой системы.

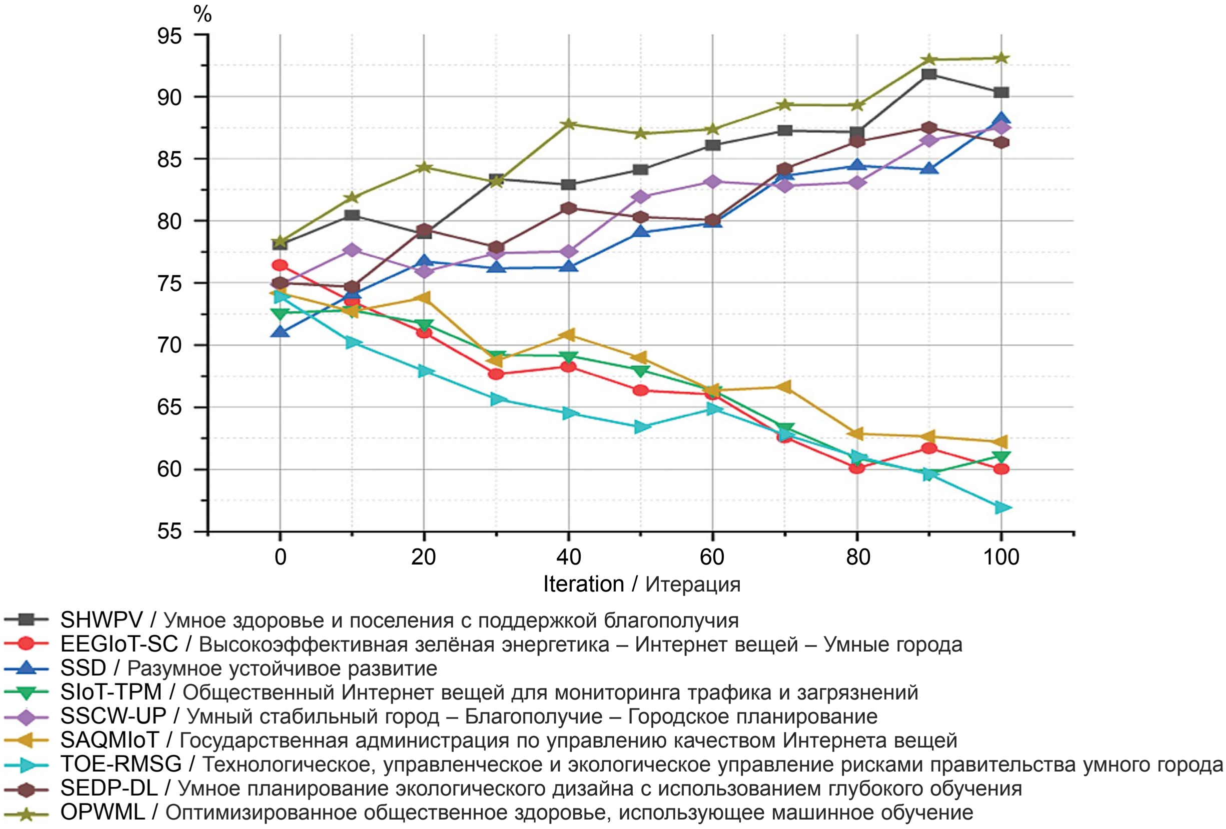



**Fig. 3.** Estimation accuracy analysis of public health and pollution.

**Рис. 3.** Анализ точности оценки здоровья населения и загрязнений.

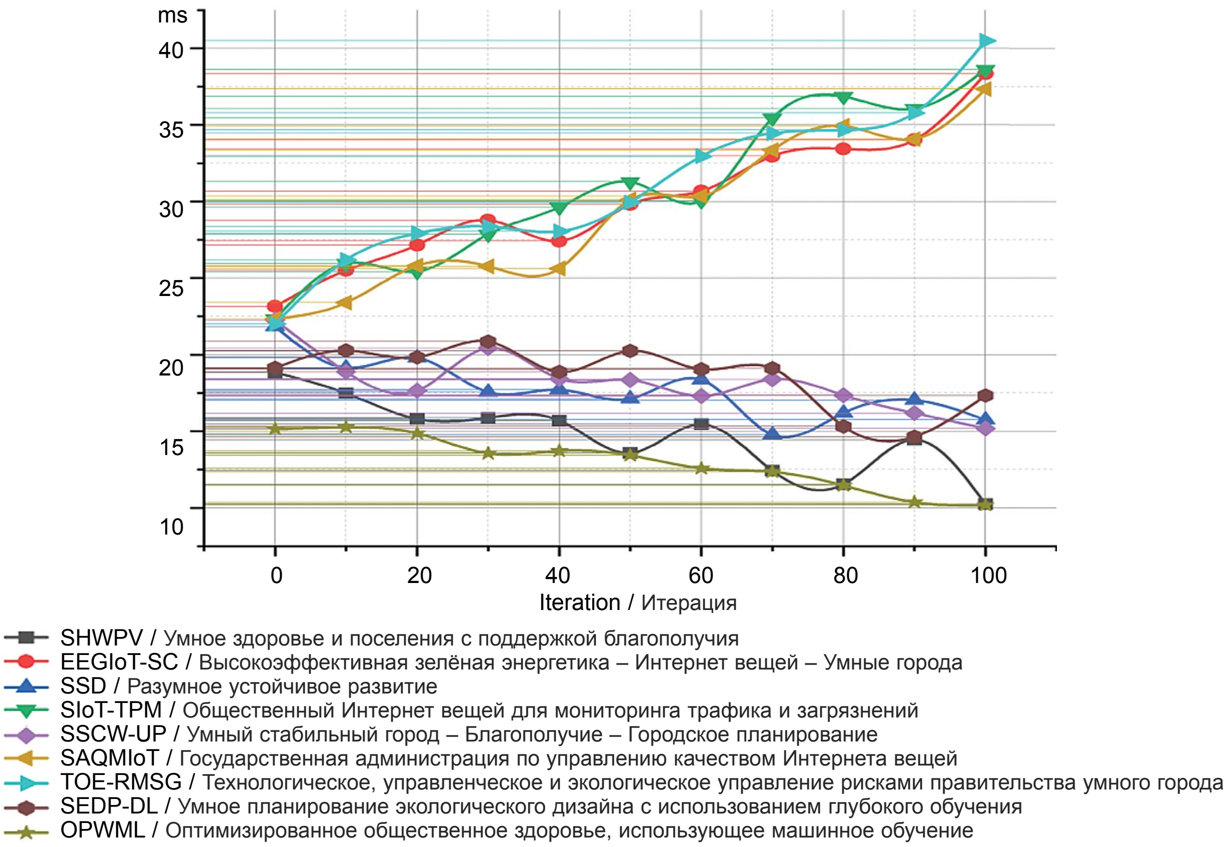



**Fig. 4.** Real-time validation delay analysis of public health and pollution.

**Рис. 4.** Анализ задержки валидации здоровья населения и загрязнений в реальном времени.

*К статье Ramanathan Udayakumar.*

*To the article by Ramanathan Udayakumar.*

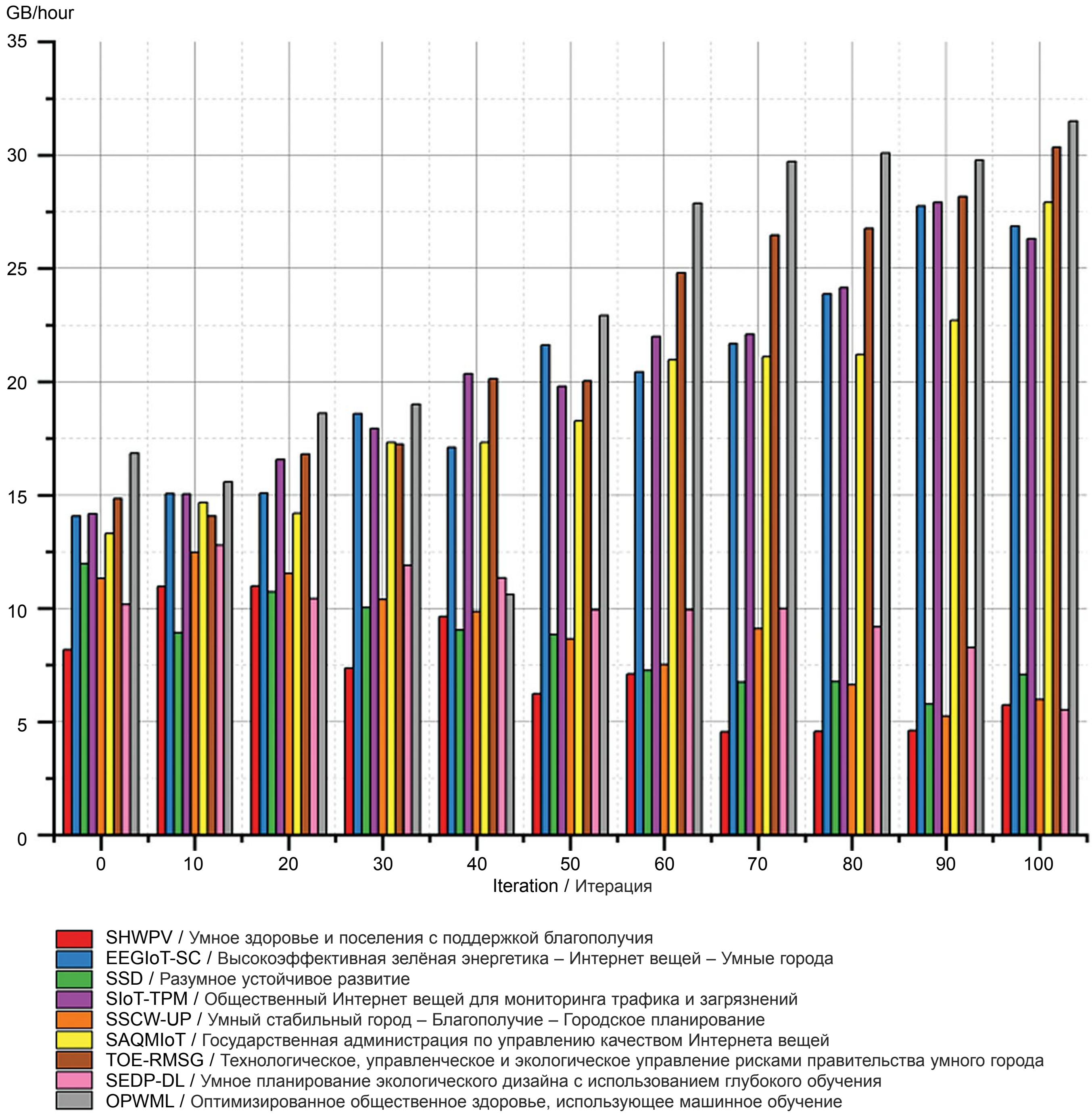



**Fig. 5.** Data collection efficiency analysis of public health and pollution.

**Рис. 5.** Анализ эффективности сбора данных о здоровье населения и загрязнениях.