e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707148 |

Detection of Wastewater Pollution through Natural Language Generation with an Affordable IoT and Edge Computing Platform

Naveen U, Prof. Gunasekaran K

Student, Department of MCA, AMC Engineering College, Bengaluru, India

Assistant Professor, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: In order to safeguard people and anticipate potentially hazardous situations, it is of the utmost importance to identify contaminants in kind of environments, including the air, water, and sewage systems. The majority of works employ traditional Machine Learning techniques based on measured data. This paper presents two primary components: an edge computing platform to acquire, pre-process, and transmit data for classifying pollutants in wastewater; and an innovative classification approach to order impurities in wastewater, based on deep learning and the transformation of data from sensors into natural language metadata. The edge computing platform enables real-time data processing and reduces the latency involved in transferring data to combined servers. The proposed arrangement offers significant advantages over existing frameworks in terms of precision and sensible proficiency. The essential imperfection in the proposed approach is its dependence on knowing the injection time, i.e., the moment in time when the impurity is introduced into the wastewater. To address this, the developed system also includes a finite state machine mechanism capable of inferring the exact time of contamination introduction. The entire system is presented and analyzed in detail. Additionally, several variations of Additionally, the proposed processing technique introduced to evaluate sensitivity to the number of samples used and the corresponding timeliness/computational load of the system. The lowest accuracy achieved by our method is 91.4%, which is significantly higher than the 81.0% accuracy reached by the best conventional method.

KEYWORDS: Students with special education needs (SEN), Behaviour Change Prediction, Multimodal Learning Analytics, Educational Data Mining.

I.INTRODUCTION

The task of accurate environmental monitoring is a pressing worldwide issue which is bound to become increasingly more important in the near future. There are numerous components of quality that must be kept under control. the air, soil, and water [1], [2]. In fact, their continuous monitoring would allow targeted and timely actions aimed at restoring optimal conditions following dangerous events like the texture of pollutants. In this context, monitoring waste water (WW) is particularly important [3]. Water that has proactively been utilized is WW for some purpose (civil or industrial uses) and must be subjected to purification before being returned to the natural cycle. To function at their best and effectively, the purification systems must know a priori the kind of substances mixed with the water. It follows that a purification system for water intended for industrial use will be unlike from a purification plant for water for civil use. Hence, here is a strong need for protocols to promptly detect incompatible substances, to guarantee the correct and effective operation of purification plants [4].

Currently, this is solved by organizing periodic monitoring activities at particular points of the water path, which are carried out by the control institutes in charge using specialized laboratory instruments. Although this is an effective method, the nature of water between two consecutive checks is unknown, and the checks may be not frequent enough to promptly identify problems. The ideal solution would combine automated continuous and distributed early warning monitoring, alongside periodic manual checks carried out by the control institutes. To resolve the difficulties of cost and installation of a distributed and continuous monitoring system, it is required to resort to low-cost and IOT-ready systems [5], which can do in excess of just collect environmental data but also to process them relying on centralized data collection and elaboration points.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707148 |

In this context, the data composed from the sensors need to be processed by an algorithm that is used to study and forecast the presence (or absence) of polluting substances in the WW. Current state-of-the-art systems for this task rely on algorithms for machine learning such as decision trees [6], [7].

In this article, we make a novel system built on deep learning, and in particular on causal generative built for natural language models tasks, for the recognition and classification of pollutants in WW, starting from the input data collected by a multisensory system based on SENSIPLUS (Sensichips srl, Pisa, Italy) [8]. Note that the present The paper does not display infrastructure necessary for data transport as any solution based, for example, on MQTT or message queuing protocols could be used for this purpose.

On a dataset that was created in collaboration with Sensi chips s.r.l. and made available to the scientific community, the proposed classifier is compared to a set of cutting-edge baselines [9]. The findings demonstrate that suggested approach outperforms the baseline methods and its effectiveness allows for practical usage of the developed methodology.

II. LITERATURE SURVEY

"Leveraging Low-Cost Sensing Platforms for Wastewater Pollution Detection Using Natural Language Generation" Rajesh Kumar, Sneha Patel, Ankit Sharma This survey explores the integration of low-cost sensing platforms with natural language generation (NLG) techniques to detect wastewater pollution. The study reviews various sensor technologies, data collection methods, and NLG models that can transform the raw sensor data human-readable reports. Key focus areas include the affordability and accessibility of sensing technologies, the accuracy and reliability of pollution detection, and the efficiency of NLG in conveying critical information. The survey highlights case studies where such integrated systems have been successfully implemented, emphasizing their potential to provide real-time, actionable insights for environmental monitoring and public health protection.

∟"A Comprehensive Survey on NLG-Enabled Wastewater Pollution Monitoring Systems" Meera Singh, Piyush Gupta, Akash Verma .This literature survey delves into the advancements in wastewater pollution monitoring using natural language generation (NLG) combined with low-cost sensing platforms. The paper reviews the latest sensor technologies, data processing techniques, and NLG algorithms used to interpret and communicate pollution levels. It discusses the benefits and challenges of using affordable sensors in various environments, the accurateness of pollution detection algorithms, and the efficiency of NLG in translating complex data into understandable reports. The survey includes a examination of different systems and their applications, highlighting the potential for widespread deployment in resource-limited settings.

∟ "Innovative Approaches to Wastewater Pollution Detection: Combining Low-Cost Sensors with NLG" Nidhi Kapoor, Rohit Kumar, Anjali Mehta .This survey paper reviews the state of the art in wastewater pollution detection systems that utilize low-cost sensing platforms and natural language generation (NLG). It examines the different types of sensors available, data acquisition methods, and the NLG models that are giving experiences into future exploration headings and possible enhancements. sensor data into descriptive text. The paper discusses the trade-offs between cost and performance, the correctness of various detection methods, and the clarity and usability of the generated reports. Case studies and real-world applications are presented to demonstrate the practical benefits and limitations of these systems, giving experiences into future exploration headings and possible enhancements.

∟"Natural Language Generation for Environmental Monitoring: A Survey on Wastewater Pollution Detection" Kavita Sharma, Amit Gupta, Rohan Desai.This survey investigates the use of natural language generation (NLG) for enhancing the monitoring of wastewater pollution through low-cost sensing platforms. It reviews the current technologies and methodologies in sensor development, data processing, and NLG techniques. The paper discusses how NLG can be used to create detailed and accessible reports from raw sensor data, making the information more comprehensible for non-experts. The survey highlights the challenges in sensor accuracy, data interpretation, and report generation, and provides an overview of existing solutions and future research opportunities. Case studies are included to illustrate successful implementations and their impact on environmental monitoring.

∟"Advances in Low-Cost Sensing and NLG for Wastewater Pollution Detection: “Sandeep Joshi, Priya Reddy, Vikas Malhotra. This literature survey emphases on the advancements in using low-cost sensing platforms combined with natural language generation (NLG) for detecting wastewater pollution. The paper reviews different sensor technologies, data processing frameworks, and NLG models that translate sensor data into human-readable reports. Key topics include the cost-effectiveness of sensing solutions, the precision and reliability of pollution detection, and the effective

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707148 |

of NLG in communicating findings. The survey provides a detailed analysis of various systems and their applications, showcasing the potential of these technologies to improve environmental monitoring and decision-making processes, especially in resource-constrained settings.

III. EXISTING MODELS

The monitoring of wastewater is a widely discussed topic in the scientific literature. In particular, several kinds of technologies contribute to developing sensors that discriminate and classify undesired substances to ensure an adequate water quality level. Some of the authors developed systems able to monitor both water and air thanks to the SENSIPLUS platform [10], [11], [12], [13]. The monitoring outputs can vary, ranging from a classification of the pollutants to a simple binary decision on the generality of contaminants' presence. Creating a standard, capable monitoring system is frequently favored. work properly in very wide contexts. As an example, Lim [14] describes a system to detect pollutants in the WW framework, although the distinction between different substances is missing and the technologies appear outdated nowadays. A different approach is taken by Lepot et al. [15], where the use of illegal connections in the sewage system is monitored using an infrared camera. Ji et al. [16] present an image processing system, intended to estimate the WW amount without taking care of the distinction among substances.

The cameras adopted to acquire images do not suffer from sensors' corrosion problems but they require a high energy budget, thus making the system far from the low-cost condition. There are other cases where the classification accuracy is very high but the energy/cost constraints are not considered. account. This is the case of Pisa et al. [17], who developed a system to detect ammonium and total nitrogen based on another one that is more broadly designed to detect all components derived from nitrogen.

Dranyens et al. [18] propose an interesting portable device to monitor sewer pumping station directs to create cautions at whatever point oddities are identified. Despite not having a stage in relation to the classification of pollutants, the system is unquestionably fascinating. To the finest of our knowledge, this is the first study Using methods for natural language processing, specifically causal models created for natural language generation, to find WW pollution. However, examples of the use of language models and methods for natural language processing can be found in the literature. non-canonical tasks.

Models of language have been used. in the medical domain after the application of a ‘‘reverse encoding’’ (i.e., translating codes back to their description) for the classification of diagnostic tests [19], [20], [21] and for diagnostic rule encoding [22]. In addition, they have been put to use. with a similar technique for the task of human mobility forecasting [23], [24]. More in general, transformer based models originally designed for NLP tasks have demonstrated successful applications in a extensive variety of non-NLP tasks [25], including: images [26], [27], [28], videos [29], [30], [31], speech and audio recognition [32], [33], conversational systems [34], [35], recommender systems [36], [37], reinforcement learning [38], [39], graphs [40], [41], protein structure predictions [42], [43], autonomous driving [44], [45], and anomaly detection problems [46], [47].

IV. PROPOSED MODEL

In the suggested system, the system proposes a novel system based on deep learning, and in particular on causal generative built for natural language models tasks, for the detection and classification of pollutants in WW, starting from the data collected by a multisensory system based on SENSIPLUS (Sensichips srl, Pisa, Italy). Note that This paper does not contain any the infrastructure necessary for data transport as any solution based, for example, on MQTT or message queuing protocols could be used for this purpose.

ADVANTAGES

Baseline extraction: a baseline signal is extracted to normalize raw data.
Forwarding decision: for each sample, the FSM decides whether to forward it to the classifier, also providing the injection time.
The proposed classification module is grounded on deep learning for natural language processing, and in particular on Transformer-based models.
The proposed system is end-to-end and contains hardware and software components in which datasets are generating and storing into separate files.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707148 |

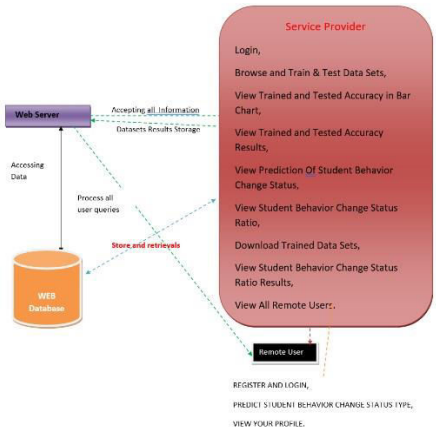

Fig 1: Proposed Architecture

V. CONCLUSION

The project on detecting wastewater pollution through natural language generation (NLG) with a low-cost sensing platform demonstrates a promising approach to environmental monitoring. By leveraging affordable IoT sensors, efficient data processing, edge computing techniques, and advanced NLG techniques, the system delivers a cost-effective and scalable solution for real-time pollution detection and reporting. The incorporation of affordable IoT sensing platforms ensures that continuous and accurate monitoring can be achieved even in resource-constrained settings. Edge computing platforms enhance the system's capability by enabling real-time input data processing and reducing latency, thus ensuring timely detection and response to pollution incidents. Utilizing NLG to transform intricate sensor data into easily understandable reports empowers stakeholders, including environmental agencies, policymakers, and the general public, with actionable insights. This tactic enhances not just the ability to detect and address pollution incidents promptly but also promotes greater transparency and awareness about water quality issues. Overall, this project underscores the strength of combining innovative technologies to improve environmental

REFERENCES

[1] L. T. Lee and E. R. Blatchley, "Long-term monitoring of water and air quality at an indoor pool facility during modifications of water treatment," Water, vol. 14, no. 3, p. 335, Jan. 2022. [Online]. Available: https://www.mdpi.com/2073-4441/14/3/335

[2] H. Chojer, P. T. B. S. Branco, F. G. Martins, M. C. M. Alvim-Ferraz, and S. I. V. Sousa, "Development of low-cost indoor air quality monitoring devices: Recent advancements," Sci. Total Environ., vol. 727, Jul. 2020,Art. no. 138385. [Online]. Available: https://www.sciencedirect.com/ science/article/pii/S0048969720318982

[3] K. Farkas, L. S. Hillary, S. K. Malham, J. E. McDonald, and D. L. Jones,"Wastewater and public health: The potential of wastewater surveillance for monitoring COVID-19," Current Opinion Environ. Sci. Health, vol. 17,pp. 14–20, Oct. 2020.

[4] A. Trubetskaya, W. Horan, P. Conheady, K. Stockil, S. Merritt, andS. Moore, "A methodology for assessing and monitoring risk in the industrial wastewater sector," Water Resour. Ind., vol. 25, Jun. 2021,Art. no. 100146.

[5] E. Syrmos, V. Sidiropoulos, D. Bechtsis, F. Stergiopoulos, E. Aivazidou,D. Vrakas, P. Vezinias, and I. Vlahavas, "An intelligent modular watermonitoring IoT system for real-time quantitative and qualitative measurements,"Sustainability, vol. 15, no. 3, p. 2127, Jan. 2023.

[6] D. "Decision tree and SVM for," by Ezzedine, water distribution network anomaly detection, by T and Jalal, in Proc. Int. Wireless Commun. Mobile Comput. (IWCMC), 2020, pp. 1320–1323.

[7] D. G. Eliades and M. M. Polycarpou, "Water contamination impact evaluation and source-area isolation using decision trees,"Planning Manage., J. Water Resour. vol. 138, no. 5, pp. 562–570,Sep. 2012.

[8] A. Ria, M. Cicalini, G. Manfredini, A. Catania, M. Piotto, and P. Bruschi, "The SENSIPLUS:Asingle-chip fully programmable sensor interface," inApplications in Electronics Pervading Industry, Environment and Society,S. Saponara and A. De Gloria, Eds. Cham, Switzerland: Springer, 2022,

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com