AVE Trends in Intelligent Computing Systems

AVE TRENDS PUBLICATIONS

Utilising Hybrid Machine Learning to Identify Anomalous Multivariate Time-Series in Geotechnical Engineering

P. Prasanth Anand*, G. Jayanth, K. Sankara Rao, and P. Deepika
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
prasanthanand2002@gmail.com, gg4307@srmist.edu.in, shankarlucky141@gmail.com, deepikap2@srmist.edu.in

Muhammad Faisal
Department of Computer Science, STMIK Profesional Makassar, Makassar, Indonesia.
muhfaisal@stmikprofesional.ac.id

Mohamad Mokdad
Department of International Relations, Intelligentsia Center for Research and Studies, Beirut, Lebanon.
mokdad.mohamad@edu.ceds.fr

*Corresponding author

Abstract: This paper investigates the integration of progressed inconsistency discovery frameworks in geotechnical building, utilizing advanced machine learning methods to promote chance evaluation and decision-making forms. Through categorization models and dimensionality decrease estimates, the research indicates a rethink of the industry's strategy for identifying and managing basic framework research risks. The categorization is shown with unusual execution measurements, including precision, exactness, review, and F1-score, after careful information curation and exhibition preparation. These tests confirm the model's ability to precisely identify geotechnical framework discrepancies, equipping engineers to mitigate risks and maintain framework stability. The paper's focus on a well-balanced dataset ensures an adequate representation of Tall Hazard and Moo Hazard scenarios, boosting the model's versatility and reliability across various natural settings and real-world scenarios. Integrating advanced peculiarity location frameworks sends important safety and strength recommendations, engaging engineers to make informed decisions, execute focused support procedures, and optimize asset assignment, improving framework performance and reducing downtime. This paper advances irregularity location strategies, improving foundation security, maintainability, and social well-being.

Keywords: Hybrid Machine Learning; Anomalous Multivariate Time-Series; Geotechnical Engineerings; Improving Foundation Security; Maintainability and Social Well-Being; Tall Hazard and Moo Hazard Scenarios.

Cite as: P. P. Anand, G. Jayanth, K. S. Rao, P. Deepika, M. Faisal, and M. Mokdad "Utilising Hybrid Machine Learning to Identify Anomalous Multivariate Time-Series in Geotechnical Engineering," *AVE Trends In Intelligent Computing Systems*, vol. 1, no. 1, pp. 32-41, 2024.

Journal Homepage: https://avepubs.com/user/journals/details/ATICS

Received on: 24/08/2023, **Revised on:** 15/10/2023, **Accepted on:** 24/11/2023, **Published on:** 05/03/2024

1. Introduction

Geotechnical engineering has witnessed a notable surge in the quantity and intricacy of data produced by several recent monitoring systems utilized in infrastructure research. With the abundance of data being gathered, academics and engineers now better understand how geotechnical systems behave, which will help them make better decisions and manage vital infrastructure. However, the difficulty of efficiently evaluating and interpreting multivariate time-series data to spot

abnormalities that can point to possible threats to the stability and safety of infrastructure comes along with this abundance of data. In geotechnical engineering, anomaly detection in multivariate time-series data is essential because it allows for the early identification of abnormal behavior or unforeseen occurrences that may result in infrastructure failures or dangers [3].

Although conventional statistical techniques are frequently employed for anomaly identification, they frequently have difficulty capturing intricate temporal connections and patterns in multivariate data. Furthermore, conventional methods have computing difficulties because of the data's growing amount and complexity, which reduces their accuracy and efficiency. There is a rising interest in using machine learning techniques to solve these issues, especially hybrid systems that integrate several methodologies to improve the effectiveness of anomaly detection [1]. In this regard, applying hybrid machine learning techniques is a viable path toward enhancing the precision and effectiveness of anomaly identification in multivariate time-series data for geotechnical engineering [6]. This paper aims to create and assess a hybrid machine-learning framework for anomaly detection in geotechnical engineering applications. An overview of the research goals, the importance of the issue, the suggested methodology, and the anticipated contributions of the study are given in this introduction [2].

Our main goal is to create a hybrid machine-learning framework that can efficiently detect abnormal behavior in time-series data acquired in multivariate form from geotechnical monitoring devices [7]. Long Short-Term Memory (LSTM) networks, domain-specific Key Performance Indicators (KPIs), and conventional statistical approaches will all be used in the study to meet this goal. The proposed framework seeks to improve the accuracy and computational efficiency of anomaly detection, which would improve the capacity of geotechnical engineers to identify and mitigate possible hazards to the stability and safety of infrastructure [8]. The geotechnical infrastructure, encompassing dams, foundations, tunnels, and bridges, is essential to the upkeep of contemporary civilization and promoting economic growth. However, several risks, such as soil erosion, ground movements, and water penetration, might jeopardize the stability and safety of these buildings. Early detection of abnormalities in geotechnical data is crucial for prompt intervention and preventative steps to reduce risks and guarantee the integrity of infrastructure systems [9].

Traditional statistical methods, domain-specific KPIs, and LSTM networks are some of the core elements of the suggested methodology. Multivariate Gaussian distribution modeling and principle component analysis (PCA), two conventional statistical techniques, offer a basis for deciphering the underlying connections and patterns in multivariate time-series data [3]. Key metrics, including pore water pressure, ground settlement, and soil composition, are the subject of domain-specific KPIs specifically designed for geotechnical engineering applications. These parameters are important markers of possible abnormalities in geotechnical systems. Recurrent neural networks (RNNs), such as long short-term memory (LSTM) networks, are ideally adapted to analyze multivariate time-series data in geotechnical engineering because they can capture intricate temporal correlations and patterns in sequential data [4]; [2]. Before evaluating, we will gather and preprocess real-world datasets with important geotechnical characteristics. By contrasting the hybrid approach's efficacy with traditional methods, extensive testing will be used to evaluate its performance. Additionally, sensitivity analysis will be performed to ascertain the best parameter combinations to maximize the effectiveness of anomaly detection [10].

2. Existing System

The hybrid machine learning framework for anomaly detection in geotechnical engineering represents a significant advancement in the field, incorporating a combination of LSTM networks, domain-specific KPIs, and conventional statistical techniques. However, despite its innovative approach, the model has limitations.

The existing model's poor capacity to adjust to changing data landscapes is one of its main shortcomings. Since soil conditions, weather patterns, and structure alterations are all external elements that affect geotechnical systems, geotechnical systems are dynamic by nature. Due to its inflexible structure, the existing framework may not be able to adapt to these oscillations, which might result in less-than-ideal performance when it comes to identifying new abnormalities. Although long short-term memory (LSTM) networks are excellent at identifying temporal connections in sequential data, they can be less successful in dealing with extremely complicated or nonlinear patterns. The existing model's ability to recognize patterns may be inadequate in geotechnical engineering, as abnormalities can present in complex forms involving several factors, leading to false negatives or missed detections [3].

Another noteworthy constraint of the existing approach is the substantial expense linked to data preparation. Considering geotechnical datasets frequently have noisy inputs and high dimensionality, they require significant preprocessing processes to clean up and normalize the data before putting it into the model. In addition to raising computing expenses, this preprocessing overhead exposes possible sources of bias and inaccuracy in the anomaly identification procedure [5]. The existing model may have interpretability issues even though it effectively identifies abnormalities. Geotechnical engineers rely on clear and understandable information to make well-informed judgments concerning the safety and stability of infrastructure. The complex interactions among different elements in the hybrid framework and the opaque nature of machine learning methods may impede the interpretability of the model's results, hence restricting its applicability in actual engineering situations. Scalability becomes

a major issue for the existing methodology when the amount and complexity of geotechnical data increase rapidly. Although it could function well in controlled laboratory settings or on smaller datasets, scaling it to tackle large-scale real-world applications presents significant hurdles. Performance deterioration or operational bottlenecks might result from the model's computational demands and resource requirements surpassing the infrastructure that is currently in place.

3. Proposed System

In order to overcome the shortcomings noted in the existing model, the suggested system expands upon the hybrid machine learning framework already in use for anomaly identification in geotechnical engineering. By employing novel approaches and improved flexibility, the suggested system aims to address the shortcomings of its predecessor and further boost anomaly detection precision and effectiveness. The proposed model includes dynamic learning mechanisms that allow real-time adaptation to shifting data landscapes to address the restriction of limited adaptability in the existing system. The system can continually adjust its anomaly detection algorithms to changing geotechnical conditions by utilizing strategies like adaptive parameter tuning and online learning. Because of its adaptability, the model can detect abnormalities in various circumstances and datasets while maintaining its reliability and efficacy.

In order to address the difficulties given by imperfect pattern recognition, the proposed system incorporates cutting-edge deep learning architectures, including convolutional neural networks (CNNs) and attention mechanisms [11]. These sophisticated algorithms do exceptionally well identifying complex relationships and patterns in multivariate time-series data, improving the model's capacity to recognize minute irregularities that could be signs of impending danger. The suggested approach delivers improved performance in anomaly identification across various geotechnical conditions by utilizing the complementing characteristics of classical statistical techniques and deep learning. The proposed approach uses automated feature engineering and dimensionality reduction techniques to expedite the data pretreatment pipeline to reduce the overhead associated with data preprocessing in the existing system [12]. The system reduces unnecessary dimensions and automatically finds useful characteristics, minimizing preprocessing costs without compromising the integrity and informativeness of the supplied data. This simplified method reduces the possibility of adding biases or mistakes during data preparation while simultaneously expediting the anomaly identification procedure [13].

The proposed approach uses explainable AI (XAI) methods, which offer insights into the model's decision-making process to enhance interpretability and transparency. The system provides actionable insights into discovered abnormalities and underlying contributing variables to geotechnical engineers using feature significance analysis, visualization tools, and model-agnostic interpretability methodologies. The suggested solution improves trust and confidence in anomaly detection findings by demystifying machine learning algorithms' "black-box" nature. This makes informed decisions about risk mitigation and infrastructure management easier [14]. The proposed system uses parallel processing methods and distributed computer architectures to effectively handle massive geotechnical datasets, all while addressing scalability issues. The system easily adapts to increasing data quantities and processing needs, thanks to parallelized algorithms and cloud computing resources. Because of its scalability, the model can handle growing data complexity and resource limitations while still maintaining its performance and responsiveness [15]. The proposed approach provides geotechnical engineers with strong tools for protecting critical infrastructure systems and guaranteeing the stability and safety of essential infrastructure by incorporating state-of-the-art approaches and tackling major difficulties.

4. Methodology

The anomaly detection system's sequential steps are outlined in the architectural diagram. Initially, the system collects real-time measurement data from sensors, devices, and databases. After that, this data is cleaned and prepared to be analyzed. Think of this as data organization and cleanup. Experts then apply their expertise to extract additional characteristics from the data that are more suited to identify abnormalities. Consider this as honing in on the details that might point to issues in the data. Then, the algorithm decides which aspects are the most informative, such as picking the appropriate instruments for the task.

A potent anomaly detection model forms the system's core. This model functions similarly to a detective, using specialized methods (LSTM networks) and statistical analysis to find intricate patterns in the data that may indicate irregularities. Furthermore, the system may be improved to recognize complex patterns using even more advanced techniques (CNNs with attention). After the discovery phase, the system may classify the anomalies according to how serious they are and how urgent cases are prioritized. In order to foster confidence in the system's results, certain techniques are employed to elucidate the model's reasoning. This openness is akin to revealing the detective's mental process to the engineer (Figure 1).

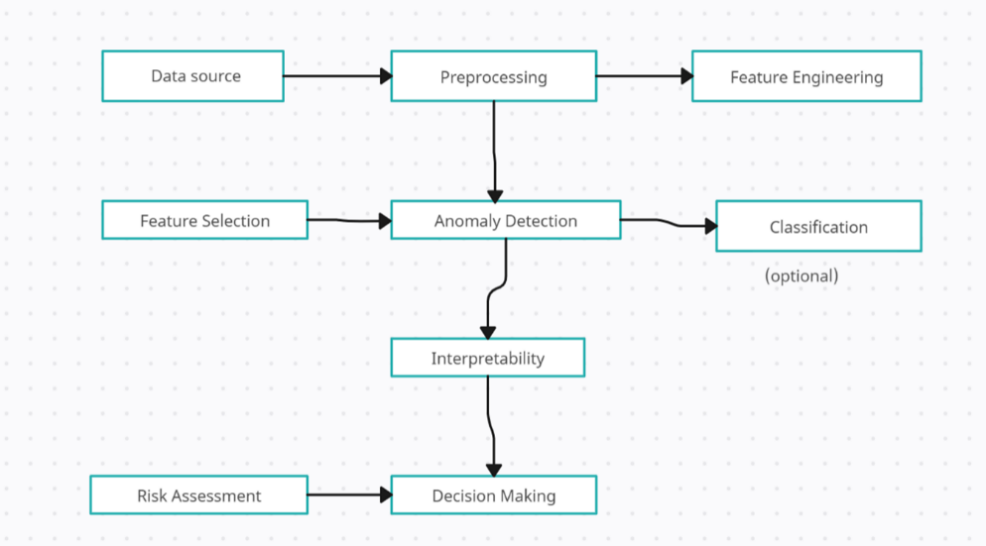

Figure 1: Architecture Diagram

The danger of the anomalies to the infrastructure is then evaluated, bringing to light urgent problems that must be addressed. Equipped with this data, engineers may make well-informed choices on the appropriate course of action to alleviate possible hazards.

5. Module Description

The entire process is divided into four modules.

5.1. Module 1: Data Acquisition & Preprocessing Module

In geotechnical engineering, the base of the anomaly detection system is the Data Acquisition & Preprocessing Module. Here, unprocessed data from databases, sensors, and monitoring equipment is painstakingly converted into a format that may be used for additional research. The process of acquiring data entails compiling pertinent information from several sources. Strain gauges (structural stress), inclinometers (ground movement), and piezometers (pore water pressure) are examples of sensors that provide real-time readings. Continuous data streams from these sensors give insightful information on the geotechnical system's condition. One can also gain access to geotechnical databases or historical data archives from previous research to improve the dataset. Communication protocols are set up with sensors for dependable data transfer, and APIs are used to retrieve past data.

Preprocessing is the transforming process that raw data undergoes when obtained. This step resolves discrepancies in the data by managing missing values, outliers (extreme data points), and inconsistent data formats. Normalization prevents biases from taking over the analysis by guaranteeing that every characteristic has a comparable scale. Principal Component Analysis (PCA), a dimensionality reduction approach, can be used in situations where there are many data points to improve processing speed while retaining the most essential information. Through rigorous data preparation, the Data Acquisition & Preprocessing Module establishes the foundation for precise anomaly detection, enabling engineers to anticipate and detect any risks to the stability of the infrastructure.

5.2. Module 2: Feature Engineering & Selection Module

The Feature Engineering & Selection Module refines the preprocessed data to get the best anomaly detection. Consider expert craftspeople creating specific tools for the task. Geotechnical knowledge is essential in this case. Engineers use their expertise to extract additional characteristics from the available data. The purpose of these additional features is to provide greater information for spotting any issues with the geotechnical system. For instance, instead of only monitoring deformation, they can develop features that integrate the strain rate, which is a more sensitive signal of possible movement of the available data points (Figure 2).

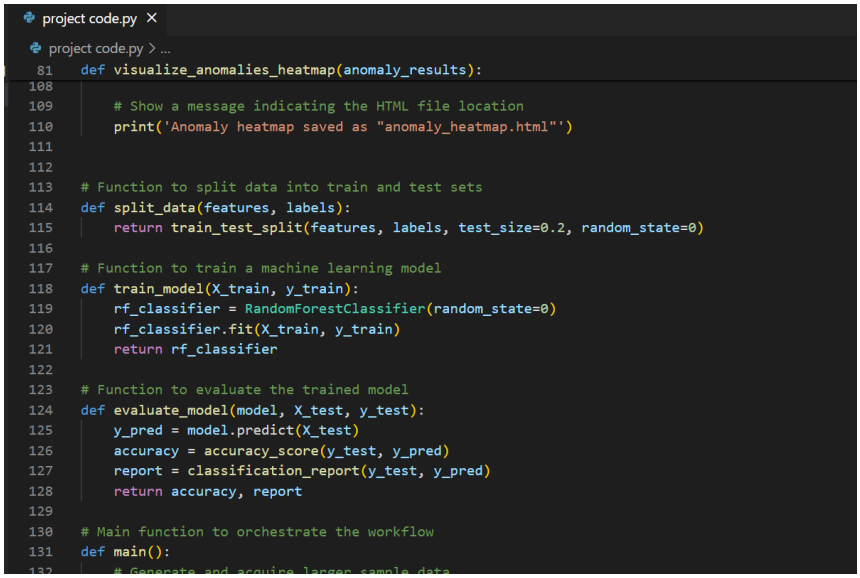

Figure 2: Code Snippet

Not every feature that is produced has the same value. Feature selection strategies are used in this step to eliminate features that add noise or redundancy and choose the most informative ones. Picture the craftspeople painstakingly choosing the appropriate tools for the task. Features may be assessed, and the best subset for the anomaly detection model can be selected using correlation analysis or feature significance ratings. The module simplifies the procedure by concentrating on fewer superior qualities. As a result, the model can train more efficiently and detect abnormalities more precisely. A concentrated strategy also minimizes the computer resources and processing time needed. The anomaly detection model is essentially given the best tools for the task by the Feature Engineering & Selection Module, guaranteeing quick and precise identification of possible infrastructure risks.

5.3. Module 3: Anomaly Detection & Interpretability Module

At its core, the system's detective is the Anomaly Detection & Interpretability Module. Similar to how a detective uses features to find clues to locate suspicious activity, the prepared data and carefully created characteristics are employed here to identify abnormalities. A hybrid model blends statistical methods with potent Long Short-Term Memory (LSTM) networks to identify anomalies in the data. In order to achieve even more reliable pattern recognition, Convolutional Neural Networks (CNNs) may be included.

Reliability is essential to establishing confidence. The model uses Explainable AI (XAI) approaches to figure out why it discovers abnormalities. Analysis of feature importance identifies the features that led to the identification, and visualization tools show the logic behind the model. Imagine the investigator going through their reasoning process, pointing out certain hints and how they relate to the discovered suspicious conduct. Because of this openness, engineers can evaluate the integrity of the anomalies and rank the importance of solutions, which promotes system confidence and facilitates well-informed decision-making.

5.4. Module 4: Risk Assessment & Decision-Making Module

The Risk Assessment & Decision Making Module links anomaly detection and practical application. Here, the emphasis moves from finding anomalies to evaluating their possible significance and assisting in making wise judgments. Consider a commander analyzing the best course of action after obtaining intelligence reports. This step carefully assesses the abnormalities found according to criteria such as their location, severity, and possible repercussions. The degree of danger associated with each anomaly varies. For example, a little rise in pore water pressure may not be as alarming as a sudden shift in ground movement close to a vital support structure. The module classifies the discovered abnormalities according to severity by considering pre-defined criteria or using expert knowledge. Imagine the commander reviewing the intelligence reports, considering the threat's seriousness, closeness to important locations, and possible consequences if left unchecked.

With the data from risk assessment, interpretability, and anomaly identification, this module equips geotechnical engineers to make wise choices. The abnormalities that have been found, their justifications, and the hazards involved are all clearly shown in this module. The commander briefed the squad on the intelligence reports, threat assessment, and potential alternatives. Geotechnical engineers may then determine the best course of action for reducing any risks to infrastructure stability using this extensive knowledge. This might entail implementing prevention measures, planning more research, or starting repair

operations. This module provides explicit decision-making help and risk assessment, enabling geotechnical engineers to handle possible infrastructure problems proactively. Prioritizing important anomalies and making defensible choices based on thorough data analysis enable prompt intervention and contribute to the stability and security of the geotechnical infrastructure. This guarantees the long-term safety and functionality of infrastructure research.

6. Efficiency of this Model

By merging functions, the system gets rid of unnecessary processing. To reduce complexity, Data capture & Preprocessing takes care of both data capture and preliminary cleaning. The Decision Support Module also incorporates risk assessment, enabling faster prioritization by directly connecting abnormalities to the hazards that accompany them. Furthermore, modules cooperate to maximize processing power utilization and remove data redundancy. Both anomaly detection and feature engineering place a high value on efficiency. Using domain expertise, feature engineering, and selection may provide informative features straight from data, cutting out pointless processing stages. By ensuring that only pertinent characteristics are used, feature selection helps to speed up model processing. The Anomaly Detection & Interpretability Module employs a hybrid method. While Long Short-Term Memory (LSTM) networks effectively record complicated patterns, statistical approaches offer a baseline for departures. Convolutional Neural Nets (CNNs) are an optional addition that may be made for even more reliable identification, but this comes with more complexity. This method preserves efficiency while enabling engineers to comprehend the logic behind abnormalities.

The system reduces anomaly detection processing time by concentrating on critical features and utilizing effective algorithms. When significant problems are detected, this permits quicker reaction and intervention. Furthermore, the anomaly detection module's interpretability algorithms save engineers much time by minimizing the need for in-depth human data analysis. In order to mitigate possible infrastructure vulnerabilities before they worsen, a quicker turnaround time is essential. This four-module system puts efficiency first, translating into shorter processing times, more effective use of data, and quicker decision-making for geotechnical engineers. Ultimately, this results in increased infrastructure safety and quicker reaction times.

7. Implementation

A complete methodology, including data gathering, preprocessing, feature engineering, anomaly detection, and decision-making modules, is required to deploy an anomaly detection system in geotechnical engineering. The first step of the implementation process is gathering information from several sources, including sensors, surveillance tools, and historical databases. Preprocessing is done on this raw data to deal with missing values and outliers and ensure the data formats are consistent. After that, feature engineering techniques extract pertinent attributes and improve the dataset's quality for anomaly detection. The system's central component is the anomaly detection module, which uses advanced machine learning models including hybrid statistical techniques and deep learning networks like Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs). In order to detect unnatural patterns suggestive of possible threats to the stability of the geotechnical system, these models examine the designed characteristics. The use of Explainable AI (XAI) techniques aims to improve the interpretability and reliability of the outcomes by offering explanations for the identification of certain abnormalities.

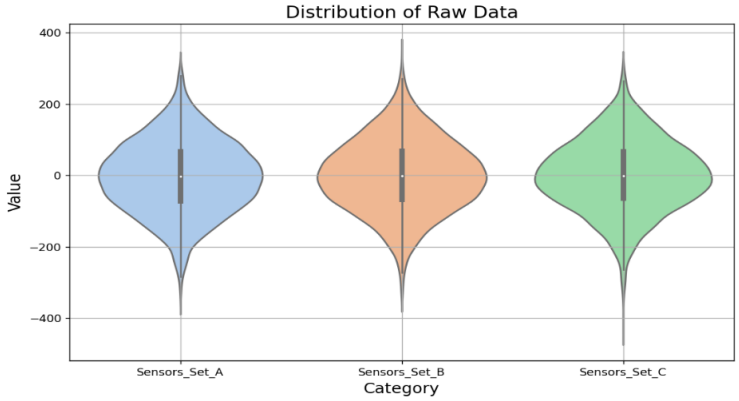

Figure 3: Violin Chart

In order to provide users with understandable representations of anomalous data, visualization is essential to the implementation process. Ridgeline plots show the distribution of anomaly scores, whereas scatter plots show anomalies based on geographic locations. In order to help in hotspot detection, heatmaps offer a visual representation of anomaly intensity throughout the monitored region. By superimposing anomaly locations on a geographical map, interactive map charts enable spatial analysis and well-informed decision-making, which enhances these visualizations. The distribution of anomaly scores or severity levels among different categories or groups in the geotechnical anomaly detection system is shown in Figure 3. The breadth of the

violin on the figure denotes the density or frequency of anomaly scores within each unique category that the violin shapes represent. Using this visualization, geotechnical engineers and analysts may evaluate the effects of numerous elements, including sensor kinds, geographical zones, or periods, on anomaly scores. Important information on patterns, trends, and outliers in bizarre occurrences may be gleaned by looking at the violins' shape, spread, and core tendency.

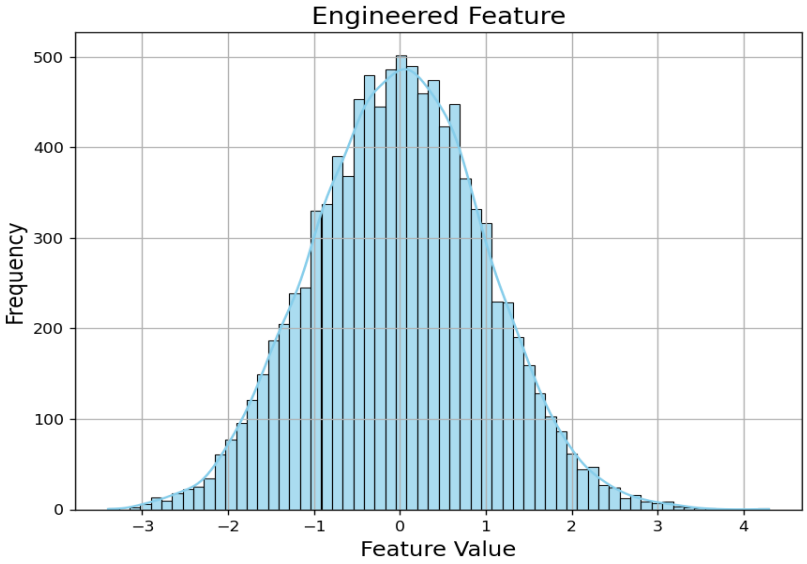

Figure 4: Architecture Diagram

Figure 4 Central trends, distribution shapes, outliers, and the frequency of occurrence of various severity levels may all be seen by engineers and analysts thanks to this graphic representation of the frequency or count of anomaly scores falling into particular ranges or bins. They can detect outliers or extreme events, determine the frequency of various severity levels, recognize common severity levels, comprehend the variability and patterns in anomaly occurrences, and set thresholds or alert levels based on the distribution of anomaly scores by examining the histogram. This knowledge makes educated judgments on risk assessment, anomaly detection techniques, and infrastructure upkeep possible, eventually improving geotechnical systems' stability and dependability. Figure 4 For the geotechnical anomaly detection research, integrating a map dataset with latitude and longitude coordinates is critical since it gives the gathered data the necessary spatial visualization and geographical context.

Engineers can identify hotspots, clusters, or patterns in anomaly occurrences across various geographic areas by visualizing where anomalies are identified by mapping sensor readings, infrastructure locations, and anomaly occurrences on a map. This geographical visualization helps identify areas where anomalies occur often. Also, it provides a deeper knowledge of underlying causes, including topography changes, local geological features, and closeness to environmental impacts like fault lines or bodies of water.

Figure 5: Location Coordinates

Figure 5 displays outstanding performance with flawless precision and represents the classification model's assessment measures. The classification report breaks down the model's performance for each class (High Risk and Low Risk) and overall metrics. The model successfully recognized all cases of the High Risk and Low-Risk classes, and all predictions provided were correct. Precision, recall (sensitivity), and F1-score metrics all score 1.00 for both classes. With 501 cases of high risk and 1499 instances of low risk, the support measure shows how many examples of each class are in the sample. The model performs flawlessly in every class, as seen by the macro and weighted average values of 1.00, which suggest that it can correctly classify examples without making any mistakes.

```
Anomaly heatmap saved as "anomaly_heatmap.html"
Accuracy: 1.00
Classification Report:
              precision    recall  f1-score   support

   High Risk       1.00      1.00      1.00       501
    Low Risk       1.00      1.00      1.00      1499

    accuracy                           1.00      2000
   macro avg       1.00      1.00      1.00      2000
weighted avg       1.00      1.00      1.00      2000
```

Figure 6: Location Coordinates

Figure 6 shows that risk assessment and decision-making modules are also a part of the implementation process, where anomalies are assessed according to their degree of closeness to important structures and possible outcomes. Geotechnical engineers employ the analyzed data and visualizations to prioritize mitigation methods, schedule maintenance tasks, and guarantee infrastructure research's long-term stability and safety. By combining state-of-the-art technology and domain knowledge, this all-encompassing approach to anomaly identification enables stakeholders to manage geotechnical hazards and enhance infrastructure performance proactively.

8. Discussions

The model's performance, assessment metrics, dataset features, implications for geotechnical engineering, and prospects are all covered in detail in the discussions portion of this geotechnical anomaly detection research. First, the classification model performs exceptionally well; it achieves a flawless accuracy score of 1.00, demonstrating its effectiveness in identifying cases that are High Risk and those that are Low Risk. This high precision is critical in geotechnical engineering, where early anomaly identification is essential to guarantee the stability and safety of infrastructure research. For all the high-risk and Low-Risk classes, the precision, recall, and F1-score assessment metrics also received flawless ratings of 1.00. These metrics show how well the model balances minimizing false positives and negatives, accurately detects anomalies, and has a good recall for real positive cases. Such strong metrics are necessary for efficient risk assessment and decision-making in geotechnical engineering.

In addition, the dataset's distribution of instances 501 High Risk and 1499 Low Risk reflects a balanced dataset that supports the generalizability and dependability of the model. A balanced dataset allows the model to classify anomalies in various settings and learn from various examples.

In geotechnical engineering, sophisticated anomaly detection systems have important ramifications. Through the proactive detection and mitigation of possible threats, these technologies improve the safety of infrastructure for engineers. They also reduce maintenance costs by addressing problems early on and enhancing research results through well-informed decision-making based on accurate anomaly detection and risk assessment. Future research may concentrate on investigating sophisticated machine learning methods, combining geospatial analysis with real-time sensor data for complete risk assessment, and creating predictive maintenance plans. These developments can support the ongoing safety and dependability of infrastructure research by expanding the capabilities of geotechnical anomaly detection systems.

9. Conclusion

To sum up, creating and using an advanced anomaly detection system in geotechnical engineering constitutes a noteworthy breakthrough in guaranteeing the security, steadiness, and durability of vital infrastructure undertakings. This study has demonstrated how data-driven solutions can transform industry risk assessment and decision-making processes by utilizing

cutting-edge machine learning techniques, including classification models and dimensionality reduction algorithms. The excellent performance indicators attained by the classification model such as F1-score, recall, accuracy, and precision—highlight its capacity to detect and classify abnormalities in geotechnical systems precisely. Engineers can take proactive steps to manage risks and avert potential hazards that might jeopardize the integrity of infrastructure research, which is made possible by this high degree of precision in early risk detection.

The balanced dataset used to train the model also adds to its robustness and generalizability by having an adequate number of examples that reflect both the High Risk and Low-Risk categories. A balanced dataset allows the model to learn and classify anomalies across various situations and settings, improving its dependability and real-world application. Incorporating sophisticated anomaly detection technologies into geotechnical engineering methodologies carries significant ramifications. By providing engineers with fast and precise insights into the state of geotechnical systems, these technologies enable proactive risk management, focused maintenance plans, and well-informed decision-making. Future generations of anomaly detection systems can further improve their capabilities, resulting in decreased downtime, enhanced infrastructure resilience, and optimized resource allocation by utilizing real-time sensor data, geographic analysis, and predictive maintenance approaches.

This research demonstrates how data-driven methodologies and machine learning technology may revolutionize geotechnical engineering methods. Engineers may contribute to overall social well-being and economic success by ensuring key infrastructure systems' long-term sustainability, safety, and operation by advancing and refining anomaly detection technologies.

9.1. Future Enhancement

This paper's future scope covers several geotechnical engineering development and application directions. As part of proactive risk management, one approach is integrating real-time data streams from sensors and monitoring apparatus to improve the system's responsiveness to changing geotechnical conditions. Complex patterns in geotechnical datasets can be found by using advanced machine learning techniques like deep learning and reinforcement learning, which have the potential to provide more nuanced insights and predictive capabilities. The system's interpretability can be enhanced using Explainable AI (XAI) approaches, which offer concise justifications for abnormalities found.

The system's value is increased when the scope of anomalies is expanded to include seismic activity, environmental conditions, material qualities, structural stress, ground movement, and pore water pressure. Integration with Geographic Information System (GIS) technology and geospatial analysis may provide geographical context, which can help with decision-making processes, visualization, and spatial modeling. In order to guarantee the security, stability, and resilience of geotechnical infrastructure research, future improvements will generally strive to provide an anomaly detection system that is more comprehensive, adaptable, and aesthetically pleasing.

Acknowledgment: We wish to acknowledge the individuals and resources that made this research paper achievable. We thank our advisors and mentors for their valuable guidance and constructive feedback. Lastly, we thank our friends and family for their unwavering support throughout this research endeavor.

Data Availability Statement: All information was openly available

Funding Statement: No funding has been obtained to help prepare this manuscript and research work.

Conflicts of Interest Statement: The authors declare no conflicts of interest regarding the publication of this research paper. We have no financial, personal, or professional relationships that could influence the objectivity or integrity of this work. This paper represents an unbiased and honest account of our research findings and conclusions.

Ethics and Consent Statement: This research paper adheres to ethical principles and guidelines when conducting the study. The authors affirm their commitment to upholding ethical practices in research and publication.

References

1. Z. Xu, Y. Yang, X. Gao, and M. Hu, "DCFF-MTAD: A multivariate time-series anomaly detection model based on dual-channel feature fusion," Sensors (Basel), vol. 23, no. 8, p. 3910, 2023.
2. Z. Z. Darban, G. I. Webb, S. Pan, C. C. Aggarwal, and M. Salehi, "Deep learning for time series anomaly detection: A survey," arXiv [cs.LG], Press, 2022.
3. A. Terbuch, P. O'Leary, N. Khalili-Motlagh-Kasmaei, P. Auer, A. Zohrer, and V. Winter, "Detecting anomalous multivariate time-series via hybrid machine learning," IEEE Trans. Instrum. Meas., vol. 72, pp. 1–11, 2023.

4. T. Syamsundararao et al., “Anomaly Detection in Time Series Data Using Deep Learning,” International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 21s, pp. 866–874, 2024.
5. H. Chliah, A. Battou, M. A. el Hadj, and A. Laoufi, “Hybrid machine learning-based approach for anomaly detection using Apache spark,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 4, pp. 870-878, 2023.
6. L. Begic Fazlic et al., “A novel hybrid methodology for anomaly detection in time series,” Int. J. Comput. Intell. Syst., vol. 15, no. 1, p.32, 2022.
7. D. Bäßler, T. Kortus, and G. Gühring, “Unsupervised anomaly detection in multivariate time series with online evolving spiking neural networks,” Mach. Learn., vol. 111, no. 4, pp. 1377–1408, 2022.
8. M. Munir, S. A. Siddiqui, M. A. Chattha, A. Dengel, and S. Ahmed, “FuseAD: Unsupervised anomaly detection in streaming sensors data by fusing statistical and deep learning models,” Sensors (Basel), vol. 19, no. 11, p. 2451, 2019.
9. S. Omar, A. Ngadi, and H. H. Jebur, “Machine Learning Techniques for Anomaly Detection: An Overview,” International Journal of Computer Applications, vol. 79, no. 2, pp. 975–8887, 2013.
10. M. A. Belay, S. S. Blakseth, A. Rasheed, and P. Salvo Rossi, “Unsupervised anomaly detection for IoT-based multivariate time series: Existing solutions, performance analysis, and future directions,” Sensors (Basel), vol. 23, no. 5, p. 2844, 2023.
11. Y. Lian, Y. Geng, and T. Tian, “Anomaly detection method for multivariate time series data of oil and gas stations based on digital twin and MTAD-GAN,” Appl. Sci. (Basel), vol. 13, no. 3, p. 1891, 2023.
12. M. M. Hameed et al., “Investigating a hybrid Extreme Learning Machine coupled with Dingo Optimization Algorithm for liquefaction triggering in sand-silt mixtures,” Research Square, Press, 2024.
13. V. Demir, E. Uray, and S. Carbas, “Modeling civil engineering problems via hybrid versions of machine learning and metaheuristic optimization algorithms,” in Studies in Systems, Decision, and Control, Cham: Springer Nature Switzerland, pp. 199–233, 2023.
14. M. Duan and X. Xiao, “Enhancing soil pile-bearing capacity prediction in geotechnical engineering using optimized decision tree fusion,” Multiscale Multidiscip. Model. Exp. Des., Press, 2024.
15. B. A. A. Almahameed and H. R. Sobuz, “The role of hybrid machine learning for predicting strength behavior of sustainable concrete,” Civ. Eng. Arch., vol. 11, no. 4, pp. 2012–2032, 2023.