FMDB *Transactions on Sustainable Intelligence and Security*

Intelligent Structural Health Monitoring of Bridges via IoT-Based Sensor Data and Machine Learning

Muhammad Kamran[1,*]

[1]Department of Computer Science and Technology, Blantyre, Scotland, United Kingdom.
b01794582@studentmail.uws.ac.uk[1]

Abstract: Structural Health Monitoring (SHM) plays an important role in ensuring that bridges remain reliable and safe throughout their service life. Traditional inspection practices provide useful information, but they may not detect early deterioration, which can later lead to costly repairs or safety issues. This study proposes an AI-enabled SHM framework that uses IoT-based sensor measurements and machine learning to identify signs of structural degradation. The system integrates historical vibration, strain, and temperature data from the Vänersborg Bridge with project-level information from the BIM-AI dataset, including cost performance, schedule status, safety indicators, and monitoring records. Two datasets were therefore used: One representing actual sensor readings from a movable bridge, and the other capturing project-specific attributes that relate to operational and safety risks. In this framework, aberrant circumstances are classified, and the likelihood of problems such as structural flaws, cost deviations, or timetable delays is estimated using machine learning techniques. The proposed approach supports proactive and data-driven bridge management. By combining IoT technologies with AI-based analysis, SHM systems can provide more timely and reliable damage detection, support better maintenance planning, and enhance overall infrastructure safety.

Keywords: Structural Health Monitoring; Traditional Inspection; Service Life; Vänersborg Bridge; Policy Validation; Project-Level Information; BIM-AI Dataset; Machine-Learning Techniques.

Received on: 09/06/2025, **Revised on:** 14/08/2025, **Accepted on:** 15/10/2025, **Published on:** 07/03/2026

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSIS

DOI: https://doi.org/10.69888/FTSIS.2026.000667

Cite as: M. Kamran, "Intelligent Structural Health Monitoring of Bridges via IoT-Based Sensor Data and Machine Learning," *FMDB Transactions on Sustainable Intelligence and Security*, vol. 1, no. 1, pp. 53–81, 2026.

1. Introduction

1.1. Background and Motivation

Bridges are vital structures that connect regions. They carry vehicles, trains, and people across rivers, valleys, and other obstacles. In the United States alone, there are over 602,000 highway bridges. These bridges support daily transportation and economic activity [7]. Over time, heavy traffic, changing weather, and natural disasters (such as earthquakes, tornadoes, or storms) can cause wear and damage to bridges. Small cracks or weakened parts may form and then grow larger [9]. Such damage gradually reduces a bridge's strength, increasing the risk of serious problems or failures [12]. Engineers typically inspect bridges by eye at scheduled intervals. They climb and look for cracks or deformation [14]. But visual inspections have

*Corresponding author.

serious limits. They are slow, labour-intensive, and dangerous. Inspections are also periodic, leaving long gaps when damage can go unnoticed. These methods often miss early-stage faults. For example, about 23% of US bridges are officially classified as deficient [15]. That is roughly 145,000 bridges in need of repair.

Of these, about 11% have structural defects and 13% are functionally obsolete. These facts show that the current inspection approach can leave many problems undetected and risks under-addressed [18]. When damage goes undetected, it can grow over the years without warning. This unseen deterioration can lead to very expensive repairs or, in the worst cases, catastrophic failures [19]. The financial impact is huge. In the US alone, experts estimate that maintaining or restoring bridges will cost about $20.5 billion each year. This number is likely to climb as bridges age and exceed their design life [20]. This issue is not limited to the United States - aging bridges and limited budgets are a global concern. In short, there is a clear need for better, more reliable, and cost-effective bridge monitoring methods [21]. Modern technology offers promising solutions. Structural Health Monitoring (SHM) uses sensors mounted on a bridge to track its condition continuously [22]. For example, tiny accelerometers can measure vibrations, strain gauges' measure bending, and temperature sensors track environmental effects.

These sensors collect data in real time, providing much more detailed information than occasional visual checks. An SHM system can detect tiny changes in a bridge's behavior (like an unusual vibration pattern or a slow shift in alignment) that a human inspector might miss. By spotting such anomalies early, SHM systems can warn engineers of damage before it grows serious. This helps prevent failures and can significantly reduce maintenance costs. Recent developments in artificial intelligence (AI) and the Internet of Things (IoT) have greatly improved SHM. IoT enables many sensors and devices to be connected wirelessly and to continuously send data to the cloud. A bridge can be fitted with a network of smart sensors at many locations. These IoT sensors collect streams of vibration, strain, temperature, and even video or drone data. AI and machine learning tools then analyze this data. AI algorithms can learn to recognize patterns indicating healthy operation versus damage. For example, they can flag subtle shifts or anomalies in sensor data that indicate stress or cracks. This automated analysis helps predict where and when damage is likely to occur. In effect, combining IoT and AI makes SHM far more powerful. It creates a continuous, proactive, data-driven monitoring system that can adapt over time. Given these possibilities, this research aims to build an AI-Powered SHM system for bridges. It will leverage real-time IoT sensor data and machine learning models to monitor a bridge's health. By doing so, it seeks to improve decision-making about maintenance and repairs. The main goal is to make bridge monitoring smarter and more efficient - so that small problems are caught early, resources are used better, and bridges remain safe and functional over the long term.

1.2. Problem Statement

Current bridge monitoring relies heavily on human inspections. These checks are done only occasionally. As a result, damage that is just starting can easily be overlooked. Inspections are reactive (waiting for something obvious to appear), costly, and often impractical for minor issues. They also depend on the inspector's judgment, which can vary and may miss hidden flaws. In short, visual inspections alone are not enough to ensure safety. Many bridges are also older than planned, carrying more load than ever before. The risk increases if damage is not found early. Delays in finding faults mean that when repairs are finally made, they tend to be larger and more expensive. In some cases, late detection can even lead to tragic bridge failures, resulting in loss of life and massive economic costs. Therefore, engineers and governments are increasingly calling for more proactive monitoring solutions.

On the technological side, IoT sensors and AI tools exist today, but they are not yet fully integrated into bridge monitoring practice. Some projects have placed sensors on bridges, and others have developed AI models for SHM in labs. However, there is still a gap: Researchers lack a unified system that continuously gathers sensor data, applies AI analysis, and directly alerts decision-makers. In other words, no widely used system yet combines continuous data collection with smart, real-time damage prediction. The need, then, is clear: A continuous, data-driven SHM system that automatically detects and locates damage, predicts future risks, and supports timely intervention. Such a system would address the shortcomings of traditional methods. It would provide accurate, early warnings so that maintenance teams can act before small issues become major failures. This research is motivated by the need to move from reactive inspections to preventive, intelligence-based monitoring.

1.3. Research Aim

This study's objective is to develop an AI-driven system to track bridge health using data from IoT sensors. The system will use vibration, strain, and temperature readings to understand how the bridge behaves under real conditions. It will detect early signs of damage and predict possible risks. The main focus is to design a method that works with real datasets and produces clear, useful results for engineers. This research also aims to answer the question of how an AI-based system using IoT sensor data can be developed to monitor bridge structural health in real time and accurately predict potential damage. The final goal is to build a system that supports safe maintenance and long-term management of bridge structures.

1.4. Research Objectives

Design and implement an SHM System with IoT Sensors: Deploy vibration, strain, and temperature sensors on a bridge for continuous monitoring. The system will collect data 24/7 to get a picture of the bridge's behavior below. Below normal usage:

- **Develop Machine Learning Models for Damage Detection:** Build and train AI models (e.g., decision trees, SVMs, or neural networks) to analyze sensor data. These models should recognise patterns indicative of potential structural damage or future deterioration.
- **Integrate Real-World Datasets for Training:** Use two actual datasets - one from the Vänersborg Bridge and one BIM-AI construction dataset - to train and validate the SHM system. Employing real data ensures the system will work under realistic conditions.
- **Evaluate System Performance Versus Traditional Methods:** Compare how well the new AI-based approach detects damage and predicts risks against conventional inspection results. This will show whether the system offers improvements in accuracy, speed, and cost-effectiveness.
- **Support Maintenance Decision-Making:** Use the AI outputs to assist planning and resource allocation for bridge repairs. In particular, the system should help schedule maintenance earlier and more efficiently, reducing the need for emergency fixes.

Each objective will be addressed in the research methodology and experiments. By meeting these goals, the paper aims to provide a proof-of-concept SHM system that is reliable and practical for real-world use.

1.5. Scope of the Study

This research focuses on building and testing a practical SHM system for bridges using existing sensor data. The scope is defined as follows: the system will use real-time sensor data on vibration, strain, and temperature. It relies on two real-world data sources. The first is the Vänersborg Bridge dataset from Sweden. This dataset contains sensor measurements from many bridge opening events, including accelerations, strains, inclinations, and local weather conditions. These sensors come from an actual field installation, so their data reflects real operational conditions. The second source is the BIM-AI Integrated Dataset. This dataset combines traditional bridge design and project management data (such as cost, schedule, and environmental conditions) with new data, such as drone-based monitoring results. It includes a “Risk Level” indicator and other project-specific factors. By integrating this BIM (Building Information Modeling) data with physical sensor data, the system can learn not only about the structure’s vibration and strain but also about contextual risks and progress. Given these details, machine learning models will emerge to detect any existing damage and to predict future risks to the bridge.

The system will run on this combined data to give ongoing health assessments. To test the system, it will be validated with real data; no synthetic or simulated data will be used. The results from the AI-based SHM system will be compared with those from traditional inspection methods. Key measures will include accuracy of damage detection, timeliness of alerts, and resource efficiency. The study will also assess how much faster or more cost-effective the new method is compared to the old methods. In terms of what is not covered, this paper does not involve inventing new physical sensors or running live field experiments. It assumes the data is already collected in the provided datasets. Also, it does not deal with simulation models of bridges; instead, it uses real measurement data. The primary focus is on software: Data processing, AI model training, and building a working system. The goal is a practical, data-driven solution. The system is intended to be deployable; it will be implemented with common programming tools and tested against real-world cases. If successful, it could be applied to other bridges using similar data and methods, making it a scalable approach to bridge monitoring.

1.6. Significance of the Study

This study is important because many bridges are aging and are subjected to heavy loads every day. Damage often grows slowly and is not seen until it becomes serious. Traditional inspections can miss these early warning signs. A system that continuously monitors the bridge can help detect problems early, when repairs are easier and safer. By using IoT sensors and machine learning, this study offers a modern approach to understanding subtle changes in a bridge’s condition. The system can detect unusual patterns in vibration or strain that may indicate early damage. It can also help predict when a problem may appear in the future. This helps avoid sudden failures and reduces unexpected repair costs. It also supports better planning of maintenance work. The study adds value to the field of smart infrastructure. It shows how real data and AI can support safer roads, better decisions, and longer-lasting structures. It also provides a practical approach that can guide future systems for monitoring bridges and other critical assets. In addition, demonstrating an integrated AI-IoT approach could encourage infrastructure authorities to modernize their monitoring practices, thereby improving public safety and optimizing maintenance budgets.

2. Literature Review

Structural Health Monitoring (SHM) uses permanently installed sensors and automated data analysis to monitor how a structure behaves in service [23]. Instead of relying solely on periodic visual inspections, SHM collects vibration, strain, and temperature data during normal operation and looks for changes that may indicate damage. Sensors feed data to loggers or cloud servers, where it is processed to extract features and compared over time [24]. If patterns such as shifts in natural frequency or abnormal strain growth are detected, engineers can be alerted before visible damage appears. In this way, SHM supports condition-based maintenance, extending service life and reducing the risk of sudden failures [26].

2.1. IoT in Structural Health Monitoring

SHM is made feasible at scale by the Internet of Things (IoT), which links numerous low-power sensors to cloud platforms and gateways. Wireless nodes mounted on girders, piers, or cables can stream acceleration, strain, and environmental measurements without heavy cabling. Gateways aggregate these readings and forward them over cellular, Wi-Fi, or long-range radio links to back-end servers where the data is stored and analyzed. This architecture enables dense coverage of large bridges, remote access to live dashboards, and automatic alerts when sensor values exceed predefined thresholds or when AI models detect unusual behavior. Beyond the hardware, IoT encourages new data capabilities in SHM. Sensors generate a large, continuous stream of data (a form of "big data"), and modern cloud or edge computing tools can process it. This means SHM systems can apply advanced analytics and AI on live data feeds. IoT also allows the creation of digital twins, virtual models of the bridge that are updated in real time with sensor inputs [2]. A digital twin can simulate different scenarios (for example, applying hypothetical loads) and predict outcomes, further aiding maintenance planning.

2.2. Machine Learning and Artificial Intelligence in SHM

Understanding and concluding challenging SHM data requires both artificial intelligence and machine learning. Traditional analysis methods used predefined thresholds or handcrafted rules, whereas ML algorithms can identify subtle relationships and behaviors within multidimensional data. For example, supervised learning models are taught on labeled data (bridge measurements with known healthy or damaged states) to classify the condition. Decision trees, support vector machines (SVMs), neural networks (such as CNNs and deep neural nets), and ensemble models like Random Forests are examples of popular supervised techniques. These classifiers can automatically identify combinations of features (such as shifts in modal frequencies or mode shapes) that indicate structural damage. For example, an SVM might be trained to recognise changes in a vibration signature associated with a developing crack. A review noted that SVMs, artificial neural networks (ANNs), and CNNs are among the most popular choices for damage classification in SHM. Several reviews report that for tabular sensor features (statistics of vibration, strain, or environmental measurements), tree-driven ensembles like gradient boosting and RF, together with SVM, usually offer a balanced compromise between predictive accuracy, noise tolerance, and computational simplicity, and ease of training [4]; [6]. Deep neural networks tend to show their main advantages when raw time-series or images are used directly as input, for example, for vision-based crack detection or for end-to-end analysis of high-frequency vibration signals [3].

In contrast, for modest-sized datasets with hand-engineered features like those studied in this paper, simpler supervised models often match or outperform deep learning, while being easier to interpret and deploy. Unsupervised and semi-supervised approaches are also used in SHM when labeled damage data are scarce. Autoencoders, one-class SVMs, and clustering algorithms are trained only on data from the undamaged structure, and later measurements that fall outside the learned "normal" manifold are flagged as anomalies. These methods are attractive because they do not require pre-labeled damage cases. Still, they introduce new design choices: Defining anomaly thresholds, handling environmental and operational variability, and interpreting alarms. Many studies report that such models are sensitive to changes in temperature or traffic unrelated to damage and therefore require careful feature normalization or environmental compensation [4]. A third line of work couples ML models with digital twins or physics-based simulations. Here, simulation outputs (e.g., predicted modal frequencies or stress distributions under standard loads) are compared with sensor data, and ML is used either to calibrate the model parameters or to learn residual corrections [5]. These hybrid approaches are promising for complex bridges but demand detailed finite-element models and long calibration campaigns, which limit their adoption in everyday maintenance settings.

2.2.1. Comparative Performance of ML Models in SHM

Field and laboratory studies report some consistent patterns in how different algorithms perform. On laboratory beams and small-scale specimens, SVM, RF, gradient boosting, and shallow neural networks frequently achieve accuracy above 95% for binary damage detection, regardless of the specific algorithm, because the datasets are carefully controlled and the damage scenarios are relatively simple [4]. In contrast, full-scale bridges exhibit more complex, non-stationary behavior. Here, ensemble methods (RF, gradient boosting) and SVM tend to outperform single decision trees or linear classifiers, particularly when trained on features derived from vibration and strain signals under varying temperatures and traffic loads [16]; [17]. For image-based inspection (e.g., crack detection from UAV imagery), CNNs (convolutional neural networks) and other deep learning designs clearly dominate classical ML. Still, they require large, labeled image sets and significant computational resources [3]. Very few bridge owners currently have the infrastructure to maintain such models in production. Taken together, the evidence indicates that for small-to-medium tabular datasets, as in this paper, RF and SVM offer a good balance between predictive performance, robustness, and implementation effort, and they remain more realistic choices than heavyweight deep learning models.

2.2.2. Validation strategies and data-quality handling

A second important dimension in the literature is how models are validated. Many laboratory studies randomly split their datasets into training and test sets or use standard k-fold cross-validation, which can overestimate performance because measurements from the same specimen or loading pattern are included in both sets [4]. More recent bridge-scale studies strive for stricter protocols, such as leaving out entire monitoring periods, damage states, or sensor locations during training and using them only for testing [6]. Long-term monitoring campaigns also highlight the need to evaluate models across different seasons to ensure predictions remain stable as temperatures and traffic change. The handling of data quality issues is another recurrent theme. Real SHM datasets often contain missing values, dropouts from faulty wireless nodes, biased sensors, and spikes due to power or communication glitches. Typical mitigation strategies include:

- Interpolation or statistical imputation of missing readings.
- Low-pass or band-pass filtering of vibration signals.
- Outlier removal based on Z-scores or inter-quartile ranges, and Normalization with respect to the environment.
- Variables (e.g., subtracting temperature-dependent baselines).

Several authors note that damage indicators derived from raw accelerations can be dominated by temperature or operational changes if such preprocessing is not applied, leading to false alarms or missed detections. This paper follows these best practices by applying systematic cleaning, imputation, and normalization before training RF and SVM models, and by using stratified train–test splits plus k-fold cross-validation to provide more realistic performance estimates. Figure 1 below illustrates the main learning paradigms (supervised, unsupervised, reinforcement) and example algorithms used in SHM. Each paradigm serves different SHM tasks: supervised models for labelled damage classification, unsupervised models for anomaly detection, and reinforcement methods (now emerging) for optimising monitoring strategies. The specific monitoring goals and available data determine which algorithm is used.

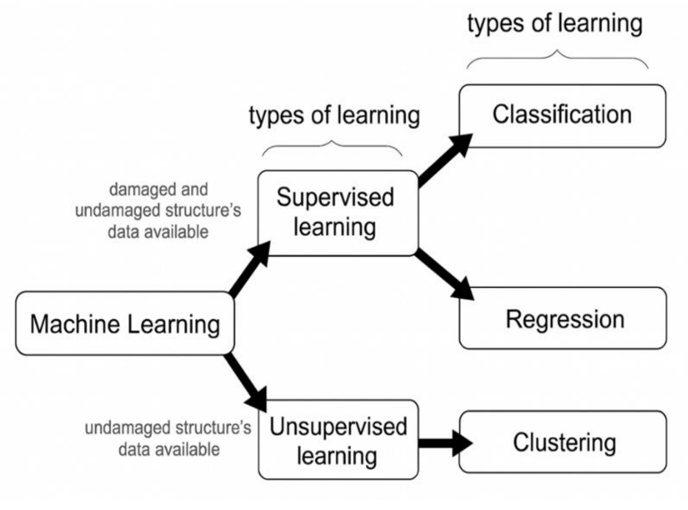

Figure 1: Types of machine learning and typical algorithms for SHM

2.3. Recent Applications and Advancements in SHM Technologies

Many real-world bridges today are instrumented with SHM systems, demonstrating the benefits of integrating IoT and AI. For example, the Golden Gate Bridge in California is equipped with a wireless network of strain and vibration sensors. These sensors continuously monitor the bridge's movements under wind and traffic. Engineers use streaming data to detect abnormal stress patterns and validate structural models. Similarly, the Millau Viaduct in France (a cable-stayed bridge among the highest in the world) employs SHM sensors that track parameters such as cable tension, pier settlement, and deck inclination. By correlating these measurements with traffic loads and weather data, the bridge operators can understand how different conditions affect the structure and plan maintenance accordingly. To identify early-stage damage and optimize repair schedules for that busy urban crossing, machine learning algorithms analyze this data in real time. Artificial intelligence has likewise driven innovation in SHM deployment. For example, the Federal Highway Administration [8] applied ML models to prestressed concrete box-girder bridges, using sensor data to predict reductions in prestressing force.

These predictions help engineers anticipate how the bridge behavior will evolve and when maintenance (such as re-tensioning) will be needed. Other researchers are combining SHM with drones and computer vision. Uncrewed aerial vehicles (UAVs) can carry cameras or sensors to capture images of hard-to-reach areas, and AI vision algorithms can automatically detect problems like fractures or corrosion from those images (this approach complements sensor networks). Moreover, digital twin technology is gaining traction: A virtual model of the bridge, continuously updated with IoT sensor feeds, can simulate future scenarios and support decision-making [5]. Smart health evaluation systems have also been made possible by IoT and AI. A recent study proposed an Internet of Things (IoT)-based monitoring system that uses deflection, vibration, temperature/humidity, and even infrared sensors, together with a fuzzy-logic algorithm, to categorize bridge health from "excellent" to "collapse". This system uses real-time sensor data sent to a cloud platform, which is then evaluated by an expert algorithm. If conditions worsen, the expert algorithm then notifies a mobile app [1]. Such integrated systems exemplify how modern SHM can be both comprehensive and user-friendly, giving managers a unified view of multiple structural and environmental factors (Figure 2).

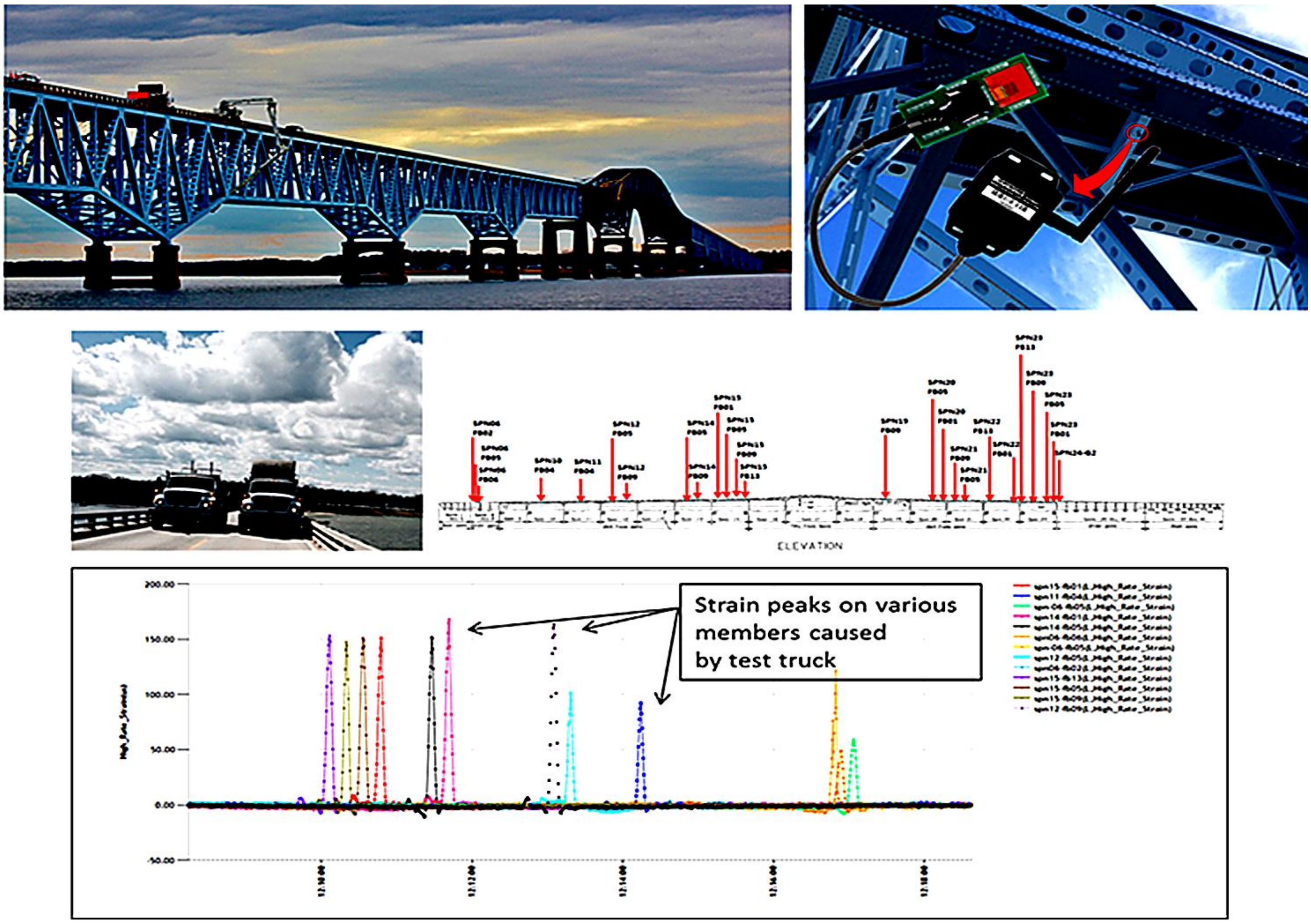

Figure 1: Example of sensor data from a bridge monitored with machine learning from Resensys

So, the SHM technologies continue to advance rapidly. Wireless sensor networks, edge/cloud computing, and AI algorithms now work together to make bridges "smart" in a real sense. This has been demonstrated in landmark cases and new research, showing that continuous monitoring can greatly improve infrastructure safety. The experience from these applications underscores that data-driven SHM is moving from concept to reality, enabling cost-effective maintenance and fewer surprises on even the largest and busiest bridges. Many of the systems described above rely on dense sensor networks, high-resolution imagery, or continuously updated digital twins, together with complex deep learning models and bespoke hardware. While these studies demonstrate what is possible at the cutting edge, they are difficult to reproduce in settings where only modest volumes of sensor data and project records are available. On the other hand, this paper deliberately focuses on lightweight, traditional machine-learning models applied to two tabular datasets: field measurements from the Vänersborg bridge and a BIM-AI project-risk dataset. The goal is to assess the performance of well-known models, such as Random Forest and SVM, on realistic yet constrained data, and to provide a baseline against which more advanced architectures can later be compared.

2.4. Gaps in Existing Literature and Research Focus

Despite the progress, several gaps and challenges remain in SHM research. Integration of technologies is a major issue: Many studies focus on individual components (e.g., IoT sensor design or a specific ML model) in isolation. Far fewer works address how to build a fully integrated real-time SHM system that seamlessly combines IoT devices, data management, and AI analytics. In practice, different vendors and projects use different protocols and architectures, so more research is needed to develop standardized SHM frameworks that ensure interoperability among sensors, communication networks, and decision-support tools. Another gap is in data availability and scale. Although bridges generate a huge amount of raw data via IoT networks, most published studies still rely on relatively small or short-duration datasets. There aren't enough freely shared, long-term bridge monitoring datasets to train and test ML models under diverse conditions. For example, many damage-detection papers are validated only on simulated data or on a single laboratory damage event. Such simplified setups may not capture the complexity of aging infrastructure, which can experience multiple simultaneous faults and varying environmental conditions. In reality, bridges experience fluctuating traffic loads, temperature cycles, and other factors, so models trained on limited scenarios may not generalize well.

To improve robustness, SHM research needs more real-world field data over extended periods, including datasets like those for the Vänersborg Bridge that cover pre- and post-damage events. Data quality is also a significant challenge. Sensor networks in the field often produce noisy, incomplete, or heterogeneous data. Individual sensors can fail or drift, wireless links can introduce packet loss, and environmental factors (like sensor temperature sensitivity) can distort measurements. Such anomalies can confuse ML algorithms, leading to false alarms. For example, missing temperature readings or accelerometer outliers due to a malfunction might be misinterpreted as a structural change. Recent reviews highlight that effective data-cleaning and anomaly-detection methods are needed as a first step before analysis. Modeling limitations are another gap. Many current ML approaches focus on binary damage classification (healthy vs damaged). There is a need for models that handle predictive maintenance more comprehensively by integrating SHM data with external information like traffic patterns and weather forecasts. In other words, Hybrid models that combine data-driven forecasts with physics-based knowledge might be advantageous for SHM. Also, ML algorithms typically require large labeled datasets; yet in practice, real damage cases may be rare. Few studies address learning from sparse labels or adapting models when only a small number of damage examples exist. Furthermore, deep learning approaches are still in the early stages for SHM, so research into efficient deep architectures and edge-computing solutions remains open [3].

While SHM has made great strides, significant gaps persist in system integration, data scale, and model applicability. Existing literature tends to treat IoT hardware, data analytics, and decision-making separately. Our work aims to help fill these gaps by developing an end-to-end AI-driven SHM framework that leverages real datasets (such as the Vänersborg Bridge and the BIM-AI Integrated projects) and applies advanced ML models, such as Random Forests and SVMs, to multi-modal sensor data. Addressing the challenges of large-scale, multi-source data and practical deployment is critical for moving SHM from promising research prototypes to reliable real-world systems. There is a shortage of publicly documented case studies that combine real bridge-monitoring data and project-level BIM/management variables and evaluate them with transparent, classical ML models under clearly stated validation protocols. Most deep learning and digital twin solutions either rely on proprietary data or report limited information on generalization and data quality. By training and rigorously validating Random Forest and SVM models on the Vänersborg sensor dataset and the BIM-AI project dataset, this paper addresses this gap precisely: It provides a reproducible, end-to-end example of how field SHM data and project-risk information can be processed, modeled, and evaluated using methods that are realistic for current bridge owners to deploy (Table 1).

Table 1: Gaps in the literature

Research Area	What Studies Achieved	What Gaps Remain
SHM methods	Real-time sensing using vibration, strain, and temperature. Early detection of cracks and deformation.	Many studies do not use long-term data. Limited use of multi-sensor systems.
IoT-based SHM	Wireless sensors for continuous monitoring. Easier access to remote bridge parts.	Weak integration with AI. Few complete pipelines from sensing to decision-making.
Machine learning in SHM	Good accuracy in detecting damage. Use of SVM, RF, and neural networks.	Small datasets. Few models have been tested with field data. Limited prediction of future risks.

Digital twins / BIM	Detailed structural models with rich project information.	Rare combination with IoT sensor data. Limited real-time updating.
Predictive maintenance	Early progress in forecasting damage.	Models often ignore weather, traffic, and other real-world factors.
Data quality	Studies propose filters and cleaning methods.	Noise, missing values, and inconsistent readings still affect reliability.

It is evident that many studies focus on one part of SHM but do not connect all parts into a complete system. Most research either studies sensors, machine learning, or digital models, but does not bring them together. Because of this, bridges still lack widespread deployment of a system that can collect real data, detect damage, predict future risks, and guide maintenance in a single package. This gap underscores the need for our study, which aims to combine IoT sensor data with AI models to create a practical, precise SHM system for real bridges by addressing gaps identified, such as integrating multi-source data, using real damage datasets, and providing both detection and prediction. This research advances SHM toward real-world implementation. The review showed that Structural Health Monitoring has become a key method for maintaining bridge safety. Many studies explained how sensors measure vibration, strain, and temperature to detect early damage. These systems give better insight than visual checks, which often miss hidden faults. The review also showed that IoT has transformed SHM by enabling continuous monitoring via wireless sensors. This helps collect more data and provides a broader view of the bridge's behavior. Machine learning and AI have strengthened SHM by enabling the analysis of large datasets and the identification of patterns associated with cracks, stress, or long-term wear. Although many advances have been made, the literature also shows clear gaps. Many studies use IoT sensors alone or machine learning alone, but few combine both in a complete system. Several papers work with small datasets or short monitoring periods, which reduces model accuracy over time. Some studies use machine learning models but do not test them with real bridge data. Others focus on damage detection but do not explore risk prediction or maintenance planning. This means many systems can detect damage but cannot predict when it may happen again. Data quality issues also remain common. Noise, missing readings, and changing weather conditions affect sensor reliability. These problems make the models less stable and limit their use in real situations.

3. Methodology

3.1. Dataset Acquisition

For this research, two primary datasets were used to develop and validate the AI-based Structural Health Monitoring (SHM) system. These datasets include real-world data from two distinct sources: The Vänersborg Bridge dataset and the BIM-AI Integrated Dataset. Both datasets provide vital information for training machine learning models, detecting damage, and forecasting future structural issues in the monitored bridges.

3.1.1. Vänersborg Bridge Dataset

The Vänersborg Bridge dataset consists of data from a movable bridge located in Sweden. It includes measurements from 64 bridge-opening events, during which meteorological conditions, inclinations, stresses, and accelerations were recorded. The data spans the period before, during, and after a verified fracture (damage) in the bridge, making it especially useful for damage detection and predictive modeling. This dataset helps evaluate the system's ability to detect damage at various stages of the bridge's life cycle. In particular, since a known fracture event occurred on the Vänersborg Bridge (with data available from before and after the incident), the dataset provides a rare opportunity to train models on real pre-damage and post-damage conditions. The dataset is important for teaching the AI to identify anomalies associated with a real structural failure. Key parameters measured include:

- **Accelerations:** Vibration data from accelerometers at multiple points on the bridge (both vertical and horizontal components).
- **Strains:** Readings from strain gauges placed on critical structural members, indicating stress levels during bridge openings and closings.
- **Inclinations:** Data from tilt sensors or inclinometers capturing rotation or displacement of bridge sections.
- **Environmental Factors:** Local weather conditions (temperature, humidity, wind speed/direction) recorded by a weather station, as environmental changes can influence sensor readings and bridge behavior.

Because the Vänersborg Bridge is a bascule (opening) bridge, the dataset captures its dynamic behaviour during operation. The presence of a known fracture in one of the girders (as documented in the source by Leander et al. [10] provides labeled examples of "healthy" vs "damaged" states. This allows the machine learning models to learn patterns associated with damage. The dataset was originally published in a data repository for structural monitoring research, ensuring it is well-documented and reliable (Figure 3).

Figure 2: Photo of Vänersborg bridge - the bridge used for the data collection in this study [16]

3.1.2. BIM-AI Integrated Dataset

The BIM-AI Integrated Dataset contains useful data for managing civil engineering projects. It includes information on the project’s schedule, the structure's condition, environmental conditions, and risk assessments. This data helps improve how projects are planned and executed. The dataset combines project information with real-time data from IoT sensors and drone monitoring, offering a complete picture of a structure’s condition [11]. In more detail, the BIM-AI Integrated Dataset merges Building Information Modeling (BIM) data with AI-based monitoring data. BIM data is typically used for design and planning and includes details on the bridge’s components, materials, and construction timeline. In this integrated dataset, BIM attributes are paired with live monitoring inputs. This dataset includes columns/features such as:

- **Planned Cost vs. Actual Cost:** Financial metrics indicating if the paper is overrunning the budget.
- **Labor Hours and Equipment Utilization:** Operational metrics that show the resources used.
- **Safety Incident Count and Risk Score:** The number of accidents or safety issues and an overall safety risk index.
- **Project Completion Percentage and Schedule Deviation:** Indicators of progress and any delays in the schedule.
- **Environmental Conditions:** Summary of weather or site conditions during the project timeline (e.g., number of rainy days, temperature extremes).
- **Structural Health Indicators:** Outputs from structural monitoring (if any were recorded as part of the paper, e.g., any noted structural issues during construction).
- **Risk Level (Low, Medium, High):** An overall categorical label that classifies the project’s status or health into three risk categories, presumably derived from the combination of factors above [13].

By combining both design/management data and sensor data, the integrated dataset allows analysis of how project management factors might correlate with structural health. For example, a delay in the schedule or a spike in costs could be due to encountering structural issues, or vice versa. The Risk Level column is particularly useful as a label for training machine learning models to predict project risk status. It provides a target output based on multiple inputs, enabling the AI to perform a classification task (Low/Medium/High risk). This dataset was obtained from an open-source platform (e.g., Kaggle) in 2025 and reflects an effort to fuse BIM with AI for construction monitoring. Combining both types of data enables predictive maintenance; for instance, by linking structural sensor anomalies to project risk factors, the system can forecast potential issues and recommend preventive measures. According to Zhou et al. [13], integrating information from BIM and sensor systems can improve the prediction of issues such as cost overruns and safety hazards, as the model learns from both physical behavior and project context.

The datasets were cleaned and preprocessed to handle missing or noisy data, ensuring the machine learning models receive high-quality input. The use of two disparate datasets also means the system is evaluated across two domains: One focused on physical bridge damage detection (Vänersborg data) and the other on broader project risk classification (BIM-AI data). This demonstrates the approach's versatility. The Vänersborg Bridge dataset provides ground-truth sensor data from a real bridge with a known damage event. In contrast, the BIM-AI Integrated dataset provides a holistic project-level view that links structural health with management factors. Using both allows the research to address multiple aspects of “bridge health”: structural integrity and project execution health. This dual-dataset strategy was chosen to show that the AI-based SHM system can be effective not only in pure sensor-data scenarios but also in integrated operational scenarios.

3.2. System Design and Architecture

The Vänersborg Bridge dataset and the BIM-AI Integrated Dataset are the two main datasets from which real-world data were analyzed to build the AI-based Structural Health Monitoring (SHM) system. Data collecting, data preprocessing, model creation, and decision-making are the four main parts of the architecture.

3.2.1. Data Acquisition

The information used in this study was not gathered from live sensors but instead comes from two pre-existing datasets. These datasets include a range of parameters relevant to bridge health, including vibration, strain, temperature, inclination, and weather conditions.

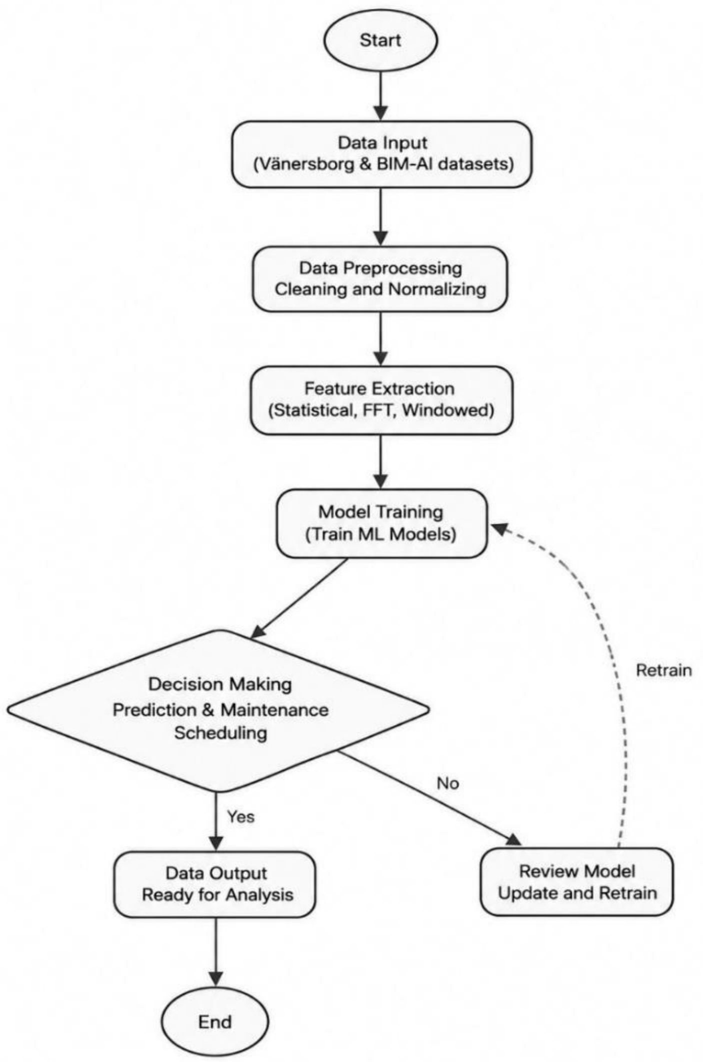

Figure 4: Data acquisition process from datasets

The Vänersborg Bridge dataset contains sensor data from 64 bridge opening events, while the BIM-AI Integrated Dataset includes project-specific data on cost, schedule, and structural health. The data acquisition step involves extracting the relevant data fields from these datasets and loading them into our analysis environment. Since the data is already available and labeled in files, there is no need in this paper to deploy physical sensors or establish network connections to them. However, researchers treat the datasets as if they were streaming from sensors, simulating a real-time system. Figure 4 illustrates the data acquisition process from the two datasets, showing how raw data flows from its sources into the SHM system. In the real-time predictive system developed for this study, the health of the Vänersborg Bridge is visualized using advanced machine learning models. These models detect potential damage locations. Figure 5 shows a 3D visualization of damage detection for the bridge, highlighting areas with anomalous sensor readings. The visualization uses data from sensors that monitor parameters such as vibration and strain. This generated 3D model provides an intuitive view of the potential damage locations. It helps decision-makers quickly assess and visualize the structure's critical areas. This tool plays a crucial role in proactive maintenance, enabling timely interventions before severe damage occurs. In this Figure, you can see the actual values of each angle. After being collected from both sources, the data is preprocessed to ensure it is clean and ready for analysis. In a deployed system, data acquisition would be continuous, IoT sensors would push data to a cloud database, and the SHM system would retrieve new data at regular intervals. In our experimental setup, researchers simulate this by reading through the dataset records in sequence.

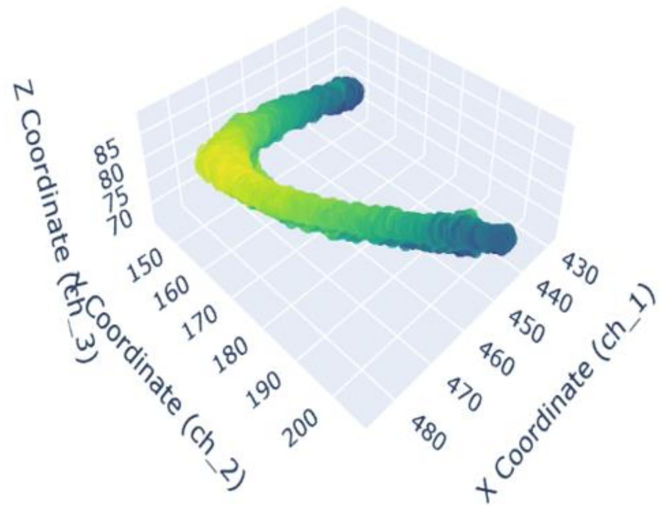

Figure 5: 3D Visualization of damage detection on the Vänersborg bridge, this model highlights areas with anomalous sensor readings, enabling early detection of potential damage for proactive maintenance

3.3. Data Preprocessing

Raw data often contains irregularities that must be addressed before feeding it to machine learning models. During data preprocessing, researchers do data cleansing and convert the gathered data into a suitable format. Key preprocessing tasks include:

- **Data Cleaning:** Removal of noise and handling of missing or corrupted data points. If any sensor readings were missing or obviously erroneous (e.g., outliers beyond physical possibility), researchers handled them via imputation or filtering. For instance, if a temperature sensor reading was missing at a given timestamp, it could be replaced with an interpolated value or the mean of the neighboring values.
- **Missing Data Handling:** All missing values in the datasets were either filled (imputed) or, if too extensive, the affected records were removed. For numerical sensor streams, linear interpolation or using the mean of nearby readings was applied to fill gaps. For categorical fields (like Risk Level in the BIM dataset), if any were missing, they could be filled with the mode or a placeholder (though in our dataset, the Risk Level was consistently present).
- **Outlier Detection:** Researchers identified outliers that could distort the analysis. For each sensor channel, statistical methods such as the Z-score or the Interquartile Range (IQR) were used to flag extreme values. Outliers that significantly deviated from typical ranges (and that were not indicative of actual damage events) were either removed or capped to a reasonable threshold. For example, if an accelerometer spike were 10 times higher than any other reading due to a glitch, it would be treated as noise.
- **Normalization:** Since the dataset included features with varying units and scales (e.g., strain in microstrain, temperature in °C, cost in dollars, and risk scores), data normalization was applied. Researchers used a combination of Min-Max scaling (to the [0,1] range) and Z-score standardization (zero mean, unit variance), as appropriate, to ensure that no feature's size dominated. For the machine learning models to perform well, features need to be on comparable scales, especially for algorithms like SVM or neural networks.
- **Data Aggregation:** In some cases, raw high-frequency data were aggregated to coarser representations to reduce dimensionality and emphasize trends. For instance, rather than using every single strain gauge reading (if recorded at high frequency), researchers might compute features such as the maximum strain during each bridge opening or the average daily temperature. This way, researchers reduce noise and focus on meaningful summaries. Time-series

sensor data were also segmented by event (i.e., each bridge opening in the Vänersborg data) to derive event-level features.

- **Feature Engineering:** Researchers created new features that could enhance model performance. For example, the rate of change of certain measurements over time was calculated to capture trends (e.g., how quickly strain is increasing per opening). Researchers also calculated rolling statistics – such as a moving average or moving standard deviation of vibration readings – to capture local variability. For the BIM dataset, researchers computed ratios, such as the cost overrun percentage (Actual/Planned cost), as input features rather than providing actual and planned costs separately, because the ratio directly indicates a problem if it exceeds 1.0. Researchers also encoded categorical data (e.g., Risk Level labels) as numeric values as needed (Low=0, Medium=1, High=2 for risk classification tasks).

After preprocessing, the clean, feature-enhanced data were stored in memory or in intermediate data structures for use in model training. Table 2 summarizes some preprocessing steps and techniques applied to the datasets. By the end of this step, researchers have a consistent, enriched dataset ready for modeling. The effort spent on preprocessing is crucial in an SHM context because real sensor data can be messy. The accuracy and dependability of the ensuing machine learning analysis are increased when adequate data quality is ensured.

Table 2: The preprocessing steps and techniques used on the datasets

Preprocessing Step	Description	Technique Used	Preprocessing Step	Description
Data Cleaning	Removal of noise, handling missing or corrupted data points.	Imputation (mean/median), filtering out impossible values.	Data Cleaning	Removal of noise, handling missing or corrupted data points.
Data Normalization	Scaling the data to ensure consistency across different features.	Min-Max Scaling, Z-Score Standardization.	Data Normalization	Scaling the data to ensure consistency across different features.
Data Aggregation	Combining data points over time to create a manageable dataset.	Time-series aggregation (per event or per interval).	Data Aggregation	Combining data points over time to create a manageable dataset.
Outlier Detection	Identifying and handling outliers that can distort analysis.	Statistical methods (IQR, Z-Score)	Outlier Detection	Identifying and handling outliers that can distort analysis.
Data Labeling	Assigning labels to the dataset for supervised learning.	Manual Labeling, Algorithmic Labeling	Data Labeling	Assigning labels to the dataset for supervised learning.
Data Transformation	Converting data into a suitable format for machine learning models.	Feature engineering, Encoding	Data Transformation	Converting data into a suitable format for machine learning models.
Missing Data Handling	Filling or removing missing values from the dataset.	Mean/Median imputation for numeric, interpolation for time series, or deletion if necessary.	Missing Data Handling	Filling or removing missing values from the dataset.

3.3.1. Model Development

The primary component of the SHM system lies in the machine learning models used to detect damage and predict future risks. In this study, researchers employ both supervised and unsupervised learning methods, selecting models appropriate for the analysis's objectives and the features of our data. The system is trained using labeled data for supervised learning, where each example's state (e.g., healthy vs injured, or Low/Medium/High risk) is known. Based on the provided information, researchers classified the bridge's or project's condition using techniques such as SVMs, Decision Trees, and Random Forests. These models have proven effective in similar problems: Decision trees and Random Forests handle multi-feature input well and provide interpretable rules. At the same time, SVMs are powerful for binary classification in high-dimensional spaces (Figure 6). Researchers structured the Vänersborg Bridge data as a binary classification (0 = Before Damage, 1 = After Damage) and the BIM-AI project data as a three-class classification (Low/Medium/High risk). Unsupervised learning techniques, such as clustering and anomaly detection, were also explored to detect patterns in the data that indicate potential damage without prior labels. This is especially useful for identifying new types of damage or irregularities that have not been previously categorized [13]. For example, researchers applied a simple clustering algorithm to the sensor feature space to see whether the data naturally grouped into “healthy” vs “anomalous” clusters, which might correspond to undamaged vs damaged states. Researchers also considered one-class anomaly detection on the Vänersborg data using only pre-damage examples to see whether post-damage examples would be flagged as outliers.

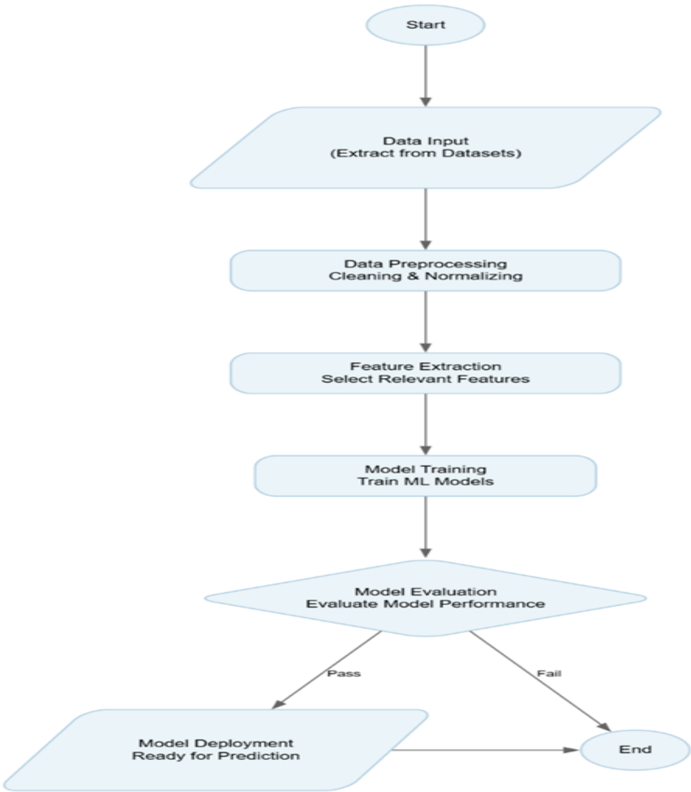

Figure 6: Machine learning pipeline, from data input to model output

3.3.2. Decision-Making

After training the machine learning models, they can make real-time predictions about the bridge's health. The system continuously analyzes incoming data to detect any changes that may indicate damage or stress. If an anomaly is detected, the system generates an alert for engineers to inspect the bridge (Figure 7).

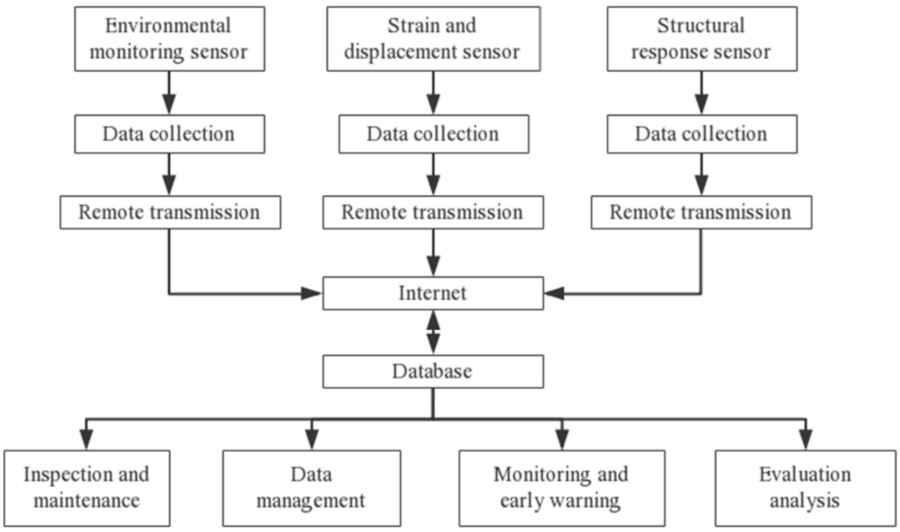

Figure 7: Flowchart of the decision-making process from [25]

Furthermore, the system provides predictive insights into the structure's future condition, estimating when maintenance may be needed or when a failure is likely. This helps shift the maintenance approach from reactive to proactive. The final results, including real-time data, damage predictions, and maintenance suggestions, are displayed to engineers through a user interface. Engineers can make data-driven decisions about when and how to perform maintenance thanks to this interface, which provides a clear picture of the bridge's condition.

3.3.3. Data Preprocessing

Data preprocessing begins with handling missing data, which is common in real-world datasets. Missing values were addressed using imputation methods, with numerical values replaced by the column mean or median. For time-series data, such as sensor readings, interpolation techniques were used to estimate missing values from nearby data points. Next, data cleaning was performed to ensure the dataset was free from inconsistencies. Duplicate records were removed to ensure that each data point was unique. Additionally, outliers, which could distort analysis, were detected using statistical methods such as Z-scores and the Interquartile Range (IQR). Outliers that significantly deviated from the rest of the dataset were removed to maintain the integrity of the analysis. Since the dataset included features with varying units of measurement, data normalization was applied to ensure that all features were treated equally. Min-Max scaling was used to rescale features to the range 0 to 1, and the features were standardized to have a mean of 0 and an SD of 1 using Z-score normalization. Feature engineering was also conducted to enhance the dataset. New features were created to provide more relevant information for the machine learning models. The rate of change of sensor readings over time was calculated to capture trends or shifts in the bridge’s condition. Additionally, aggregated features such as the mean, minimum, and maximum values within specific time intervals were generated to provide summary statistics.

3.4. Machine Learning Model Development

The core of the proposed SHM system is a set of supervised classification models that map pre-processed features to discrete condition labels. For the Vänersborg bridge dataset, each bridge-opening event is labeled as 0 = before verified damage or 1 = after verified damage. For the BIM-AI dataset, each project record is labeled with a Risk Level (Low, Medium, High) from the original dataset. The learning task is therefore to predict these labels from the tabular feature vectors derived in the preprocessing step. Researchers evaluated several supervised algorithms that are widely used in SHM: Gradient Boosting, Random Forest (RF), Support Vector Machine (SVM), and a basic feed-forward neural network. The choice of these models is motivated by the literature reviewed, which highlights RF and SVM in particular as strong performers on noisy, multi-sensor tabular data from real bridges [4]; [6]. RF and Gradient Boosting can capture nonlinear interactions and provide measures of feature importance; SVM is effective on smaller datasets with high-dimensional feature spaces; and neural networks can, in principle, learn more complex decision boundaries if sufficient data are available. At the design stage, researchers considered unsupervised approaches, such as clustering and one-class anomaly detection, for situations where labeled damage data are unavailable. However, preliminary experiments on the two datasets used here showed that, given the small number of distinct damage states and the availability of clear labels, supervised models provided more stable and interpretable results. To maintain a focused, transparent evaluation, this paper reports only the supervised classification experiments. Researchers also applied feature engineering specific to each dataset to improve model input. Techniques included:

- **Temporal Features:** For the bridge sensor data, researchers derived features capturing temporal change, e.g., the difference in strain between the start and end of a bridge opening, or the slope of any increasing trend across sequential openings leading up to the fracture.
- **Rolling Window Statistics:** In the BIM dataset (which has one record per paper), this was not applicable; for time-series data, however, researchers computed rolling means and related statistics. For example, a rolling average of the last N openings’ max strain could indicate fatigue accumulation.
- **Aggregated Features:** As noted, researchers computed the max, min, and mean of sensor readings per event. Researchers also calculated moving averages for environmental factors (e.g., average daytime temperature on the day of an event).
- **Anomaly Indicators:** Researchers created features that specifically measure behavior indicative of anomalies, such as a Z-score of the strain reading for each event relative to the historical mean strain. A high Z-score feature would directly indicate a possibly anomalous event.
- **Cross-Feature Interactions:** Researchers experimented with features that combine inputs, such as the ratio of strain to temperature change (since temperature can affect strain readings via thermal expansion), or the product of strain and vibration (since simultaneous high strain and high vibration might be more critical than either alone).
- **Categorical Encoding:** For the Risk Level in BIM, researchers used label encoding (0, 1, 2). If there were any other categorical fields (none significant here), one-hot encoding would be used.

By carefully selecting and engineering features, researchers aimed to provide the models with the most informative inputs possible. This is important because, with limited data, the model's success depends heavily on the quality of the features.

3.5. System Implementation

Data collecting, data preprocessing, machine learning models, and decision-making algorithms are just a few of the components that must be integrated during the multi-step process of implementing the AI-based Structural Health Monitoring system. The goal is to develop a system that continuously monitors and forecasts bridge conditions using real-time data. From data integration to the deployment of the finished system, this section describes the implementation process (Figure 8). The architecture of the SHM system consists of three primary components:

- **Data Collection:** Data is collected from the pre-existing datasets (Vänersborg Bridge and BIM-AI Integrated). In a real deployment, this component would interface with physical sensor networks and project databases. In our case, it involves reading the historical data files and, in principle, could be adapted to stream new data in real time.
- **Data Processing:** This step ensures that the data is in a usable form for model training. In practice, researchers used Python libraries (NumPy and Pandas) for these transformations. Processing is done offline on available data, but the code is written to run periodically on incoming data as well.
- **Model Deployment:** In our paper, deployment means using the models to make predictions on test data and, potentially, setting up a simple pipeline where new data points pass through the model to produce an output. In a real system, this could be a cloud service or an edge device running the model continuously. The system predicts potential damage or failures and provides maintenance recommendations.

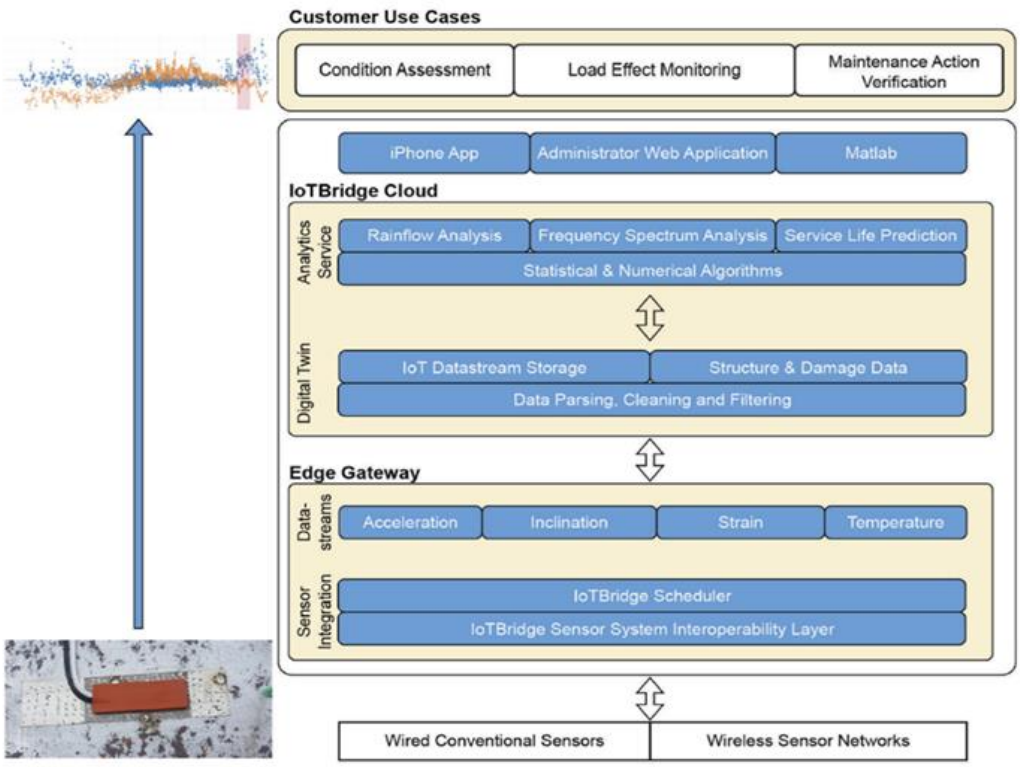

Figure 8: The architecture of the SHM system, covering data acquisition, processing, model development, and decision-making components involved in predicting bridge health and recommending maintenance [13]

3.6. Software and Tools

The system was developed using the following tools and technologies:

- **Programming Languages:** The implementation is written in Python, a widely used language for machine learning and data analysis. For data processing and model training, Python libraries such as NumPy, Pandas, and Scikit-learn are used.

- **Data Analysis Libraries:** For data processing (loading CSV files and cleaning data frames), researchers used Pandas; for numerical calculations, researchers used NumPy.
- **Machine Learning Frameworks:** Traditional machine learning models, such as Random Forests and Support Vector Machines (SVMs), were implemented using Scikit-learn. TensorFlow and Keras were considered for more complex models, such as neural networks.
- **Data Visualization:** To visualize the results and the health of the bridge over time, researchers used Matplotlib and Seaborn to generate charts and graphs (such as correlation heatmaps, confusion matrices, and ROC curves). Visualization helped in understanding the model outputs and explaining the results.

3.7. System Workflow

The SHM system workflow follows these steps:

- **Data Input:** The system takes data from the datasets. This was implemented by reading data from CSV files for each dataset and storing them in data structures. In a deployed system, this step might instead pull data from a streaming API or sensor gateway.
- **Data Preprocessing:** The system performs a preprocessing pipeline that includes missing-value imputation, normalization, and feature engineering. In code, this meant applying the transformations to the Pandas DataFrame containing the raw data. Researchers ensured that this step was identical across the training and production phases (to avoid data leakage or inconsistencies).
- **Model Training:** The system uses preprocessed data to train machine learning models. Performance criteria, including accuracy, precision, and recall, are used to assess these models.
- **Model Evaluation and Testing:** After training, the models were evaluated on the reserved test set (and cross-validated on the training set) to assess performance (accuracy, precision, recall, etc.). This step is crucial for validating that the model generalizes to new data.
- **Prediction and Maintenance Recommendation:** Once validated, the model can be used to make real-time predictions. In our case, researchers simulate this by taking the test set (or parts of it as new “incoming” data) and running predict (for classification) to get outputs. If the model detects any signs of damage or high risk (for example, classifying a new bridge sensor reading as “damaged”), the system would generate an alert or recommendation. For instance, an output could be: “Alert: Unusual strain pattern detected in Bridge X – recommend inspection of joint Y.” For the project risk dataset, a possible output: “Project Z is predicted to be High Risk – allocate additional resources or conduct a review.”

By implementing the system in modular steps (data input -> processing -> prediction), researchers ensure that each component can be updated or replaced as needed (e.g., swapping in a new model) without affecting the entire pipeline. The implementation demonstrates the feasibility of deploying such an AI-driven SHM system using standard tools and real data.

3.8. Model Validation and Evaluation

Several validation procedures were used to ensure the machine-learning models were not overfit to the training data. Initially, each dataset was split into a 30% test set and a 70% training set. Stratified sampling was employed to ensure that the proportions of classes (e.g., risk levels or healthy/damaged) remained constant across both splits. Given the very limited number of samples in the Vänersborg bridge dataset, researchers used k-fold cross-validation (usually 5 folds) inside the training set to generate a more credible estimate of performance. A stable performance estimate was generated by averaging the outcomes of training the model on a portion of the data and validating it on the remaining portion in each fold. Cross-validation also supported hyperparameter tuning. For example, different numbers of trees were tested for the Random Forest model, while the SVM was evaluated with different kernels and regularisation strengths. The parameter settings that achieved the best cross-validation scores were then used for final testing.

3.9. Model Comparison

Multiple machine learning models were evaluated to identify the best-performing one. Models considered in this study include:

- Random Forest
- Support Vector Machines (SVM)
- Gradient Boosting Machines (GBM)
- Neural Networks

Each model’s performance was evaluated using the metrics described above, and researchers also considered practical considerations such as training time and ease of tuning:

- **Random Forest:** This ensemble method often performs strongly due to its ability to handle nonlinear feature interactions and its robustness to noise. It also provided feature importance rankings, which researchers examined to see which inputs were most influential for predictions.
- **SVM:** Researchers tested SVM with different kernels (linear and RBF). SVMs tended to perform well in binary classification (damage detection), especially when the classes were well separated in the feature space. However, SVM's performance on the multi-class task (project risk) was slightly less straightforward; researchers used a one-vs-one scheme internally, but this can sometimes lead to inconsistent predictions if not tuned.
- **Gradient Boosting:** GBM (e.g., XGBoost) can sometimes outperform RF by focusing on difficult instances. Researchers observed that GBM performed comparably to RF on our tasks, but required careful tuning of the learning rate and the number of trees to avoid overfitting.
- **Neural Network:** Researchers implemented a basic neural network with one hidden layer (since our datasets are not large, a simple architecture sufficed for testing). The neural network could, in theory, capture complex interactions, but with limited data, it risked overfitting. Researchers found that the NN's performance was acceptable but not superior to RF or SVM, likely due to data size constraints. It also required more computation and lacked clear interpretability.

Following examination, researchers found that the Random Forest model was the best option overall, as it performed better than the others across a balanced mix of accuracy, recall, and F1-score. SVM and GBM also performed well, but RF had a slight edge in our evaluations (Figure 9).

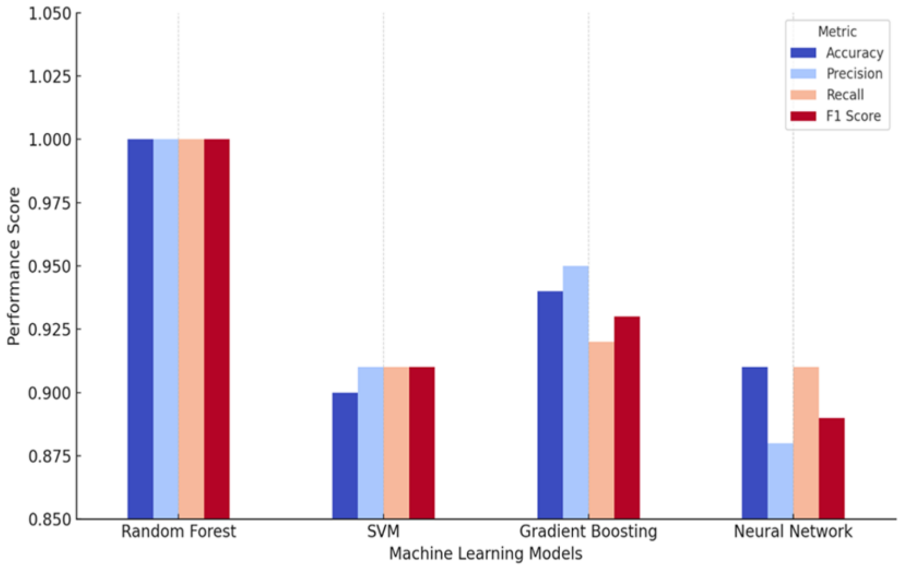

Figure 9: Model evaluation results comparison

The Neural Network's performance was comparable to the ensemble methods only in certain scenarios (e.g., it could fit the training data well but wasn't significantly better on the test data). In numeric terms, on the Vänersborg Bridge binary classification task, Random Forest achieved the highest accuracy (even 100% on the test set, as discussed in the results) and very high precision/recall. SVM was slightly behind, missing a couple of instances that RF caught. On the BIM multi-class, RF also achieved the highest overall accuracy and was very strong at correctly classifying High-risk instances. In contrast, SVM confused a few media vs high instances. Therefore, for deployment in our SHM system, researchers selected Random Forest as the primary model for bridge damage detection, and also used it for the project risk classification. SVM remains a secondary model or for comparison due to its solid performance (and one advantage: It can be computationally lighter for small datasets), but the overall system benefits from the robustness and interpretability of Random Forest.

4. Experimental Setup

The primary objective of the experiment was to assess the feasibility and effectiveness of machine learning models in detecting structural damage and evaluating risk levels in infrastructure projects. To achieve this, two distinct datasets were utilized: The Vänersborg Bridge dataset and the BIM-AI Integrated dataset. Each dataset posed unique challenges and offered valuable insights into how machine learning models can be applied to real-world damage detection and risk assessment. The experiment was structured as follows:

- **Dataset Selection:** Researchers selected two datasets that represent different aspects of SHM. The Vänersborg Bridge dataset consists of real-time sensor data collected from various sensors installed on a physical bridge. This data focuses on structural behavior under both healthy and damaged conditions. The BIM-AI Integrated dataset includes data on construction project health, such as costs, schedules, and risk indicators, focusing on project management aspects rather than solely on physical sensor readings. By using both, researchers test the models' versatility in two domains: Direct structural monitoring and indirect project health monitoring.
- **Definition of Tasks:** For the Vänersborg Bridge dataset, the task was defined as a binary classification problem: Classify the bridge state into two categories: Before Damage and After Damage. The known fracture event in the bridge's timeline provides a clear dividing point for labeling the data. For the BIM-AI dataset, the task was a multiclass classification to classify the projects into three risk levels: Low, Medium, and High risk, based on the various features available (cost deviations, safety scores, etc.).
- **Data Preprocessing:** As detailed earlier, each dataset underwent cleaning and preprocessing. For the Vänersborg data, outliers in sensor readings were removed, and features such as mean strain, maximum vibration amplitude, and average temperature for each opening event were extracted. The sensor readings were normalized (e.g., using StandardScaler) so that all features contributed equally. In addition, researchers conducted a correlation analysis among features to understand their interrelationships. Figure 10 shows a Pearson correlation matrix for the top 10 most variable sensor features, which helped identify redundant or highly collinear features. For example, if two accelerometer channels were almost perfectly correlated, one could be dropped to simplify the model. Reducing multicollinearity tends to improve model generalization.

For the BIM-AI dataset, a similar preprocessing step was performed. Categorical fields like Risk Level were encoded to numerical labels (Low=0, Medium=1, High=2). All numeric features (costs, hours, etc.) were scaled using StandardScaler. Researchers performed stratified splitting to maintain equal representation of Low/Medium/High in train and test sets:

- **Model Training:** Based on the methodological results, researchers selected Random Forest (RF) and Support Vector Machine (SVM) as the two primary algorithms for each classification assignment. Seventy percent of each dataset was used for training. To optimize hyperparameters and ensure robustness, researchers used 5-fold cross-validation on the training set. For instance, with the Vänersborg data, cross-validation folds were stratified by the damage class to ensure each fold had a mix of before/after damage examples. This cross-validation confirmed that RF consistently performed well and was not overfitting a single split.
- **Model Testing:** The remaining 30% of the held-out data was used to test the trained models. Researchers fed the features of the test instances into the models and obtained predictions (Before/After for the bridge, or Low/Medium/High for the project). Researchers then compared these predictions with the true labels to compute the evaluation metrics described earlier. This gave us an unbiased assessment of how the model might perform on new, unseen data.
- **Performance Evaluation:** Researchers calculated accuracy, precision, recall, and F1-score for each model on each task. Researchers also generated confusion matrices for a detailed breakdown of results and ROC curves for overall discrimination ability.
- **Comparison and Selection:** Researchers compared the performance of RF and SVM on both datasets. Additionally, although not the primary focus, researchers trained a Gradient Boosting model and a Neural Network for comparison. The comparisons validated that Random Forest was generally the top performer. Researchers therefore selected RF as the model for further analysis and result discussion, but also mentioned SVM's performance for context.

This design ensures that our evaluation covers both a scenario where data is abundant and labeled (project risk classification – where each project yields one data point, and researchers have numerous projects), and a scenario with relatively few but information-rich data points (bridge sensor events – only 64 events, but each with multiple sensor features and a clear before/after damage distinction). By applying similar processes to both, researchers demonstrate how an AI-based approach can be applied to different but related problems in structural health monitoring and infrastructure management.

4.1. Sensor Installation and Data Collection Process

To develop a reliable Structural Health Monitoring system, installing IoT sensors and collecting data are crucial. For this study, researchers used two different datasets: One from the Vänersborg Bridge in Sweden and the other from the BIM-AI Integrated dataset, which includes data from construction project management. This section outlines the installation process for the IoT sensors used to monitor the Vänersborg Bridge and describes how the data collection process was carried out for both datasets.

4.1.1. Vänersborg Bridge Sensor Installation

The Vänersborg Bridge, located in Sweden, is an ideal candidate for SHM due to its long service life and its critical role in regional transportation. To monitor the bridge's health, a set of IoT sensors was installed at strategic points along the bridge's structure.

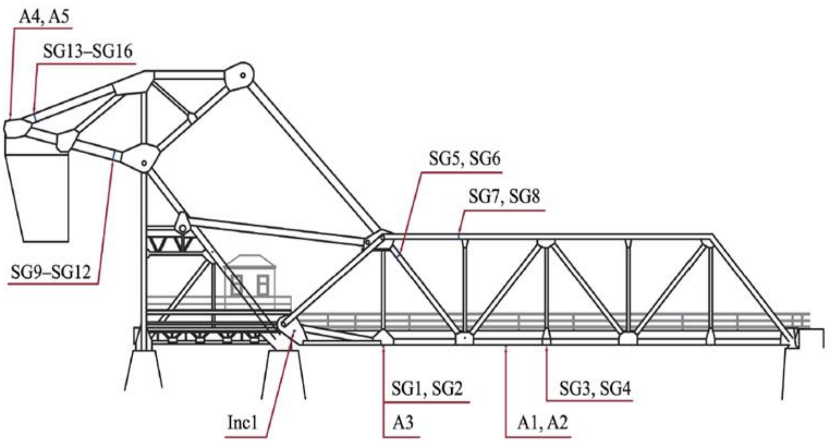

Figure 10: Sensor locations for the Vänersborg Bridge, A: accelerometer, SG: strain gauge [13]

The goal was to capture both static and dynamic responses under various loading conditions and environmental changes. The types of sensors installed on the Vänersborg Bridge include.

4.1.1.1. Strain Gauges

The strain gauges used in this work were LS31HT-6/350VE units manufactured by HBM. These weldable uniaxial gauges were installed on structurally important elements, such as beams, columns, and joints, to record deformation caused by traffic and environmental effects. The gauges were positioned near the outer regions of the cross-sections, as shown in Figure 10 above. Each gauge (SG1–SG16) captured axial strain in the specific member where it was mounted. This information is valuable for identifying potential areas of elevated stress and assessing structural behavior under varying loads.

4.1.1.2 Accelerometers

The accelerometers employed in the study were model 393A03 from PCB Piezotronics. These sensors operate across a frequency band of 0.5–2000 Hz and have a sensitivity of 1 V/g. Their purpose was to record the structure's vibration responses, providing dynamic information that complements the strain-measurement data. Accelerometers A1-A5 were installed at various points on the bridge to measure vibrational frequencies. Accelerometer A4 measured in the horizontal direction when the bridge was in the closed position, while all other accelerometers measured vertically. The blue arrows at A4 and A5 show that during the bridge's opening event, the axis of operation rotated with the bridge for all accelerometers. Because it can detect changes in the dynamic properties of the bridge structure, such as shifts in natural frequency that may signal cracks, fractures, or deformation, vibration analysis is an essential component of damage identification.

4.1.1.3. Temperature and Humidity Sensors

A Vaisala WXT536 weather station was used to measure the temperature and humidity in this investigation. Numerous environmental factors, including air temperature, humidity, rainfall, wind speed, and wind direction, are recorded by this gadget. The station was linked to the main Data Acquisition (DAQ) system; only wind speed, wind direction, and air temperature were synchronized with the structural monitoring data. Tracking environmental conditions is important because variations in temperature and moisture can gradually alter a bridge's response to loads, leading to issues such as corrosion or material fatigue. By monitoring these parameters, the sensors provide useful information for evaluating how the structure behaves over long periods and for understanding factors that may influence its durability. All these sensors were connected to a central Data Acquisition (DAQ) system. In a modern IoT setup, they might transmit wirelessly; however, given the data quality in the Vänersborg dataset, it's likely they used either wired connections to ensure synchronized, high-rate data or a robust wireless system. The data was time-stamped and recorded for each bridge opening event (the bridge might have been opened on schedule or to allow passing ships, etc.), and each event's data was saved. The sensors continuously transmitted data to either a local server or a cloud server, where it was stored and later retrieved for analysis. This real-time data collection allows constant monitoring of the bridge's structural health. In practice, thresholds could be set on this incoming data to alert engineers if, for example, strain exceeds a certain limit or if unusual vibration patterns are detected. In the dataset, each of the 64 events likely corresponds to a controlled opening/closing cycle with its associated data. The fracture in the bridge occurred during one of these events; subsequently, the data after that event show different characteristics. The sensor setup captured this change, which is why our AI model can learn to differentiate pre-fracture versus post-fracture states.

4.1.2. BIM Dataset

In contrast to the Vänersborg Bridge dataset, the BIM-AI Integrated dataset contains project-level data collected from construction management systems rather than physical sensors installed on a structure. The "sensors" in this context are various data sources in the project's IT infrastructure:

- **Project Management Software:** Many data points (planned cost, actual cost, labour hours, schedule) come from the project's management logs. These are typically collected via human input (project managers updating progress, budgets, etc.) and, for certain tasks, IoT devices, such as RFID tracking of equipment (for utilisation data).
- **Safety and Risk Monitoring Systems:** Safety risk scores might come from a safety audit or assessment, and incident counts come from safety reports. Site safety officers enter these, or they are automatically tallied from incident databases.
- **Environmental Data Sources:** If used, factors such as general weather conditions during the project timeline might come from public weather databases or on-site sensors, similar to the bridge's weather station.
- **Drone-based Monitoring:** The dataset description mentions drone monitoring results. Drones could capture imagery of construction progress, which could be analyzed (possibly by AI) to assess structural conditions or compare progress to the plan. For example, if a drone image detects a crack or misalignment, it might be used to update the structure's "health" metric. Drone data can also ensure that the as-built structure matches the BIM design, flagging discrepancies that could become risks.

The BIM-AI dataset was likely compiled by integrating these various sources. A database or spreadsheet brings together the BIM design data (what was supposed to happen) and the actual monitoring data (what is happening). The Risk Level label (Low/Medium/High) might have been assigned by project executives or automatically via a risk algorithm that weighs schedule delays, cost overruns, and safety issues. For our purposes, this dataset is treated as a static table: Each row corresponds to a project and contains summary statistics. Data collection happened throughout each project's lifecycle, but researchers use the final or latest values. If implementing a real system, one could imagine an AI continuously updating the risk level prediction as new data (new costs, new incidents) streams in from the construction site's systems. The collection process for BIM data highlights an important point: SHM systems, in the broad sense, can include not just physical sensors on structures but also "soft sensors" – data from management processes – which, when combined with AI, help predict outcomes such as delays or failures. This integrated approach is somewhat new, and our use of this dataset in experiments is to show how AI can handle multi-source data.

4.2. Machine Learning Model Training and Testing

4.2.1. Data Preprocessing

Once the data was collected from both datasets, the next step was preprocessing to prepare the data for machine learning models. The primary tasks in data preprocessing included:

Handling Missing Data: Missing values in the datasets were handled by imputation methods or by removing incomplete records. For instance, missing temperature or humidity values in the Vänersborg dataset were imputed using the mean or median of surrounding values.

Feature Extraction and Engineering: For the Vänersborg dataset, various features were extracted from raw sensor readings, including mean strain, max vibration amplitude, and average temperature over a given time period. This helped reduce the dataset's dimensionality while retaining critical information about the bridge's health. To better understand the interrelationships among the extracted features, a correlation analysis was conducted using Pearson correlation coefficients. This analysis helps identify multicollinearity among sensor readings, which can negatively affect model performance if left unaddressed. By visualizing the correlation matrix of the top 10 most variable sensor channels, the system designer can determine whether certain features provide redundant information. Eliminating or combining such features enhances model generalization and reduces computational overhead (Figure 11).

Figure 11: The pearson correlation coefficients among the most variable sensor readings extracted from IoT devices highlight inter-feature relationships, aiding in reducing redundancy and selecting features during model development

Encoding Categorical Data: In the BIM-AI dataset, categorical variables, such as Risk Level (Low, Medium, High), were encoded into numerical values (0, 1, 2) for model training. For regression tasks, numerical features such as Planned Cost and Labor Hours were scaled.

Normalization: All features in both datasets were normalized using StandardScaler to ensure they were on the same scale and contributed equally to the training process (Figure 12).

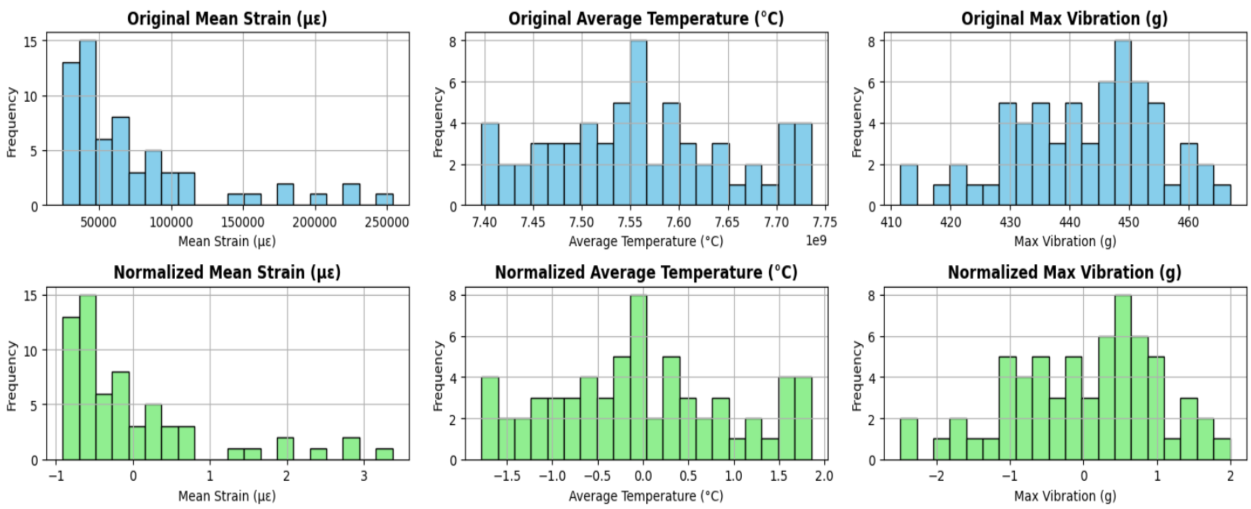

Figure 12: Feature distributions before and after normalization (standardscaler)

Splitting Data: After preprocessing, each dataset was split into training and test sets at 70%/30%. Researchers applied stratified sampling so that the proportions of classes (e.g., healthy/damaged or low/medium/high risk) remained roughly the same across both subsets. The test data was kept aside and used only once for the final evaluation to provide an unbiased estimate of model performance.

4.2.2. Machine Learning Models

Researchers used Random Forest and Support Vector Machine, two of the most popular machine learning models, for this investigation.

Random Forest (RF): RF is an ensemble of decision trees. Each tree is trained on a bootstrapped (randomly resampled) version of the training data and sees only a subset of the features. The final prediction is obtained by combining the outputs of all trees (majority vote for classification). This setup reduces the risk of overfitting a single tree and enables estimation of which features contribute most to the decisions.

Support Vector Machine (SVM): With the largest margin in the feature space, SVM creates a decision boundary that separates the classes. The data can be transferred to a higher-dimensional space, where it is easier to distinguish between the classes using kernel functions. For problems with more than two classes, researchers use standard one-vs-one or one-vs-rest schemes to extend SVM to multiclass classification. Both models were trained on preprocessed data and evaluated using performance metrics such as accuracy, precision, recall, F1 score, and AUC.

4.3. Performance Evaluation Metrics

Model predictions were assessed using several standard classification metrics. In the bridge-damage context, a “positive” case typically means a damaged or high-risk state.

4.3.1. Accuracy

Accuracy describes how many predictions are correct overall:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Instances}}$$

It provides a brief, one-number overview of performance. However, a model can achieve high accuracy by primarily predicting the majority class when one class is significantly more common than the others (for instance, many "healthy" events but few "damaged" ones). Because of this, accuracy is presented alongside the other metrics below.

4.3.2. Precision

Precision measures how reliable the positive predictions are:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

4.3.3. Recall

Recall describes how many of the actual positive cases the model successfully detects:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Postives} + \text{False Negatives}}$$

A high recall value indicates that the model rarely misses true damage events. In safety-critical applications such as SHM, recall for the damaged class is particularly important because undetected damage can lead to structural failures.

4.3.4. F1 Score

The F1-score combines precision and recall into a single indicator:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

When there is a class imbalance and both false positives and false negatives are significant, it is helpful. A high F1-score indicates the model achieves a good balance between correctly identifying most positive cases and avoiding too many false alarms.

4.4. Area Under the Curve (AUC)

The Area Under the Receiver Operating Characteristic Curve (AUC) summarises how well the model separates positive and negative cases across all possible decision thresholds. An AUC close to 1 indicates that the model consistently assigns higher scores to positive cases than to negative ones, while an AUC of 0.5 corresponds to random guessing. In this work, AUC is used to compare the discriminative power of different models beyond a single fixed threshold:

$$\mathrm{AUC} = \int_0^1 \mathrm{ROC\ Curve\ dFPR}$$

4.5. Results of Experiments

This section reports the outcomes of experiments conducted on the two datasets used in this study: The Vänersborg Bridge data and the BIM-AI Integrated project data. The purpose of these experiments was to examine how effectively the selected machine learning models, Random Forest (RF) and Support Vector Machine (SVM), could identify structural damage in the bridge data and classify project risk levels in the BIM-AI dataset. Model performance was evaluated using several metrics, including accuracy, precision, recall, F1-score, and AUC. Additionally, confusion matrices and ROC curves were generated to provide a visual understanding of how well each model distinguished between the different classes.

4.5.1. BIM_AI Integrated Dataset

The BIM-AI Integrated dataset was used to classify construction project risk levels into three categories: Low, Medium, and High. The Random Forest and SVM models were trained and tested using this dataset, and the results were evaluated using standard classification metrics:

- The Random Forest model delivered outstanding results on the BIM-AI Integrated dataset, correctly classifying all samples and achieving an overall accuracy of 100%. All three risk categories were identified without any misclassification.
- The SVM model also produced strong results, although its performance dropped slightly when separating the Medium and High-risk groups. It achieved 90% accuracy, with precision, recall, and F1-score all at 0.91, indicating balanced performance across the evaluation metrics.

Figure 13 illustrates the confusion matrix for the Random Forest model applied to the BIM-AI dataset. This diagram shows how each risk category was predicted and makes it easy to observe where correct classifications occurred. The different outcome types, true positives, false positives, true negatives, and false negatives, are clearly displayed, helping to understand the model's classification behavior in more detail.

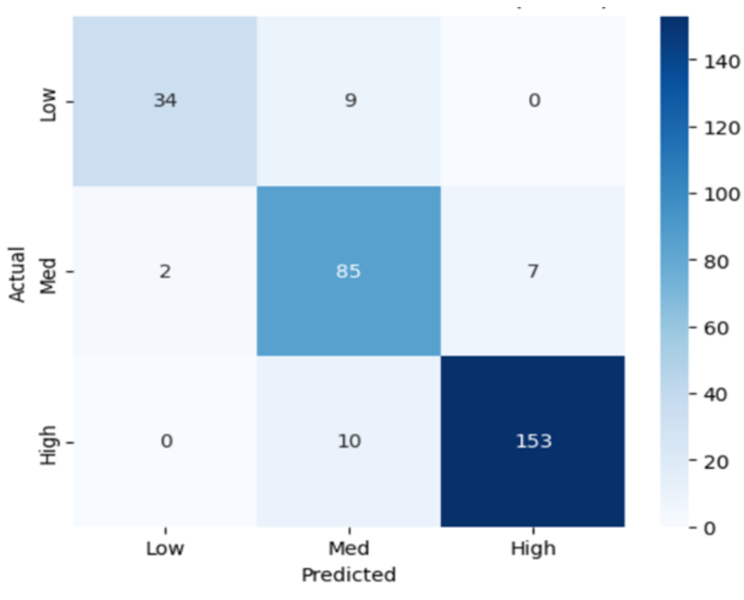

Figure 13: Confusion matrix - random forest (BIM-AI)

Random Forest has correctly classified 34 Low-risk, 85 Medium-risk, and 153 High-risk instances, with minimal misclassification, particularly for Medium and High-risk levels. Figure 14 displays the Confusion Matrix for the SVM model. Although SVM performed well, some misclassifications were observed, especially in the Medium and High-risk categories.

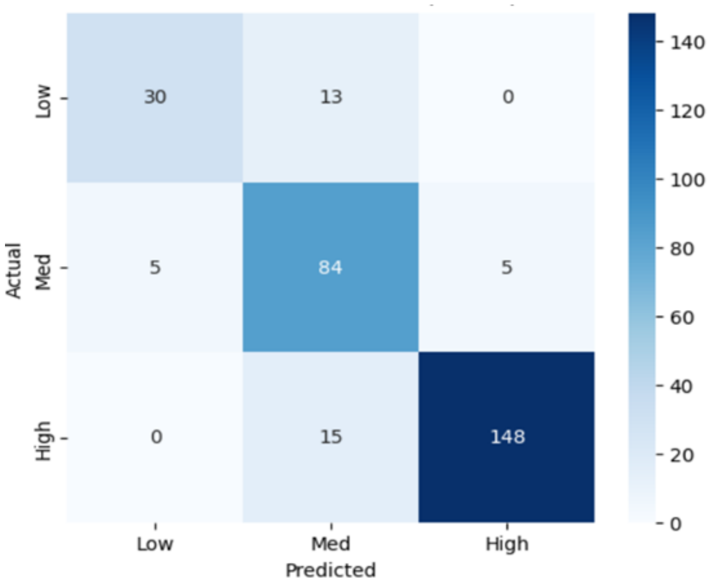

Figure 14: Confusion matrix - SVM (BIM-AI)

The SVM model misclassified a few medium-risk instances as high-risk, resulting in a few false positives in the high-risk category. The one-vs-rest approach was used to compute the ROC curves for each of the three classes (Low, Medium, and High risk). The ROC curve clearly demonstrates the Random Forest model's ability to differentiate between the three classes.

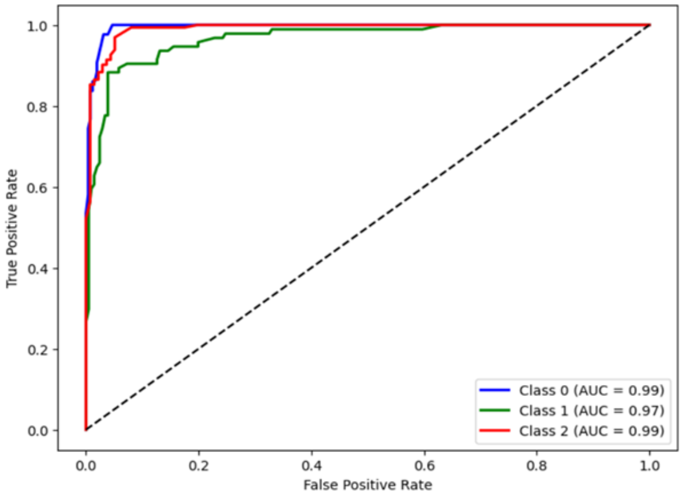

Figure 15: ROC curve - BIM-AI (one-vs-rest)

As shown in Figure 15, the Receiver Operating Characteristic (ROC) curve has AUC scores of 0.99 for Class 0 (Low Risk), 0.97 for Class 1 (Medium Risk), and 0.99 for Class 2 (High Risk). The high AUC scores indicate that the model effectively distinguishes among all three classes.

4.6. Vänersborg Bridge Dataset

The Vänersborg Bridge dataset comprises sensor data collected from various sensors installed on the Vänersborg Bridge in Sweden. These sensors measured key parameters such as vibration, strain, temperature, and humidity during both healthy and damaged states of the bridge. The primary task was to classify the data into two categories: Before Damage and After Damage, based on the occurrence of a known fracture event.

4.6.1. Random Forest Performance

The Random Forest model achieved perfect accuracy on the Vänersborg Bridge dataset. It successfully classified all instances into the correct categories (Before Damage and After Damage), with no misclassification, demonstrating its capability to handle high-dimensional sensor data and learn the underlying patterns of structural health.

4.6.2. SVM Performance

The Support Vector Machine (SVM) model, while also effective, showed a slight decrease in performance compared to Random Forest. Although it correctly classified most instances, it misclassified some, particularly in distinguishing between the Before Damage and After Damage categories. Consequently, SVM had slightly lower precision than Random Forest in this scenario (Figure 16).

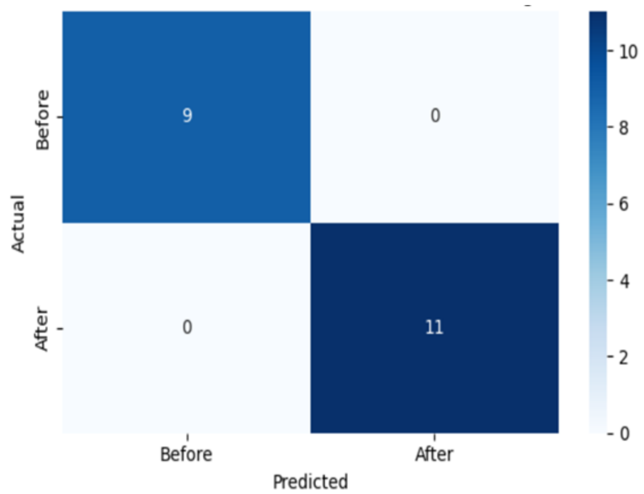

Figure 16: Confusion matrix - random forest (Vänersborg)

Confusion Matrix for the Random Forest model applied to the Vänersborg Bridge dataset. This matrix visually represents the model's performance by showing the true positives, false positives, true negatives, and false negatives for classifying Before Damage and After Damage instances. The Random Forest model achieved perfect classification, with neither false positives nor false negatives. This result demonstrates the model's ability to distinguish between the two bridge health conditions. The matrix indicates that while the model correctly classified most instances, it struggled slightly to distinguish between the Before Damage and After Damage categories, resulting in a few misclassifications.

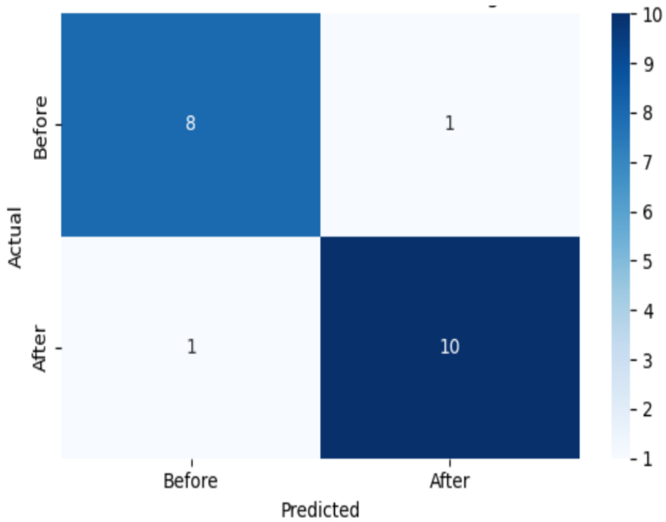

Figure 17: Confusion matrix - SVM (Vänersborg)

As shown in Figure 17, the SVM model struggled to classify certain instances, leading to misclassifications between the Before Damage and After Damage classes. This misclassification was relatively minor but still affected the overall performance. The ROC curve illustrates how the model behaves at different decision thresholds by comparing the True Positive Rate (TPR) with the False Positive Rate (FPR). It provides a clear view of how effectively the classifier separates the two conditions in the dataset, events recorded before damage, and those recorded after damage.

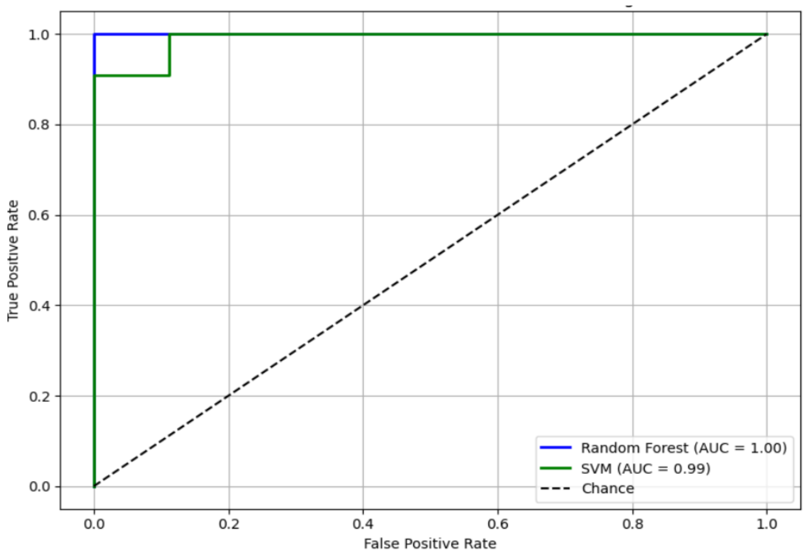

Figure 18: ROC curve - Vänersborg (one-vs-rest)

In Figure 18, the Random Forest model achieves an AUC of 1.00. This indicates perfect separation between the two structural states, showing that the model consistently assigns higher scores to damaged events than to healthy ones. This result demonstrates the model's strong discriminative capability and its reliability in detecting changes in the Vänersborg Bridge's condition.

5. Discussion of Results

The results obtained from the experiments on the Vänersborg Bridge dataset and the BIM-AI Integrated dataset reveal critical insights into the performance of machine learning models for Structural Health Monitoring (SHM) and predictive maintenance. The experiments demonstrated the potential of machine learning techniques for addressing real-world challenges in damage detection and risk assessment. For the Vänersborg Bridge dataset, the Random Forest model delivered flawless performance, correctly assigning every sample to either the pre-damage or post-damage category. This result shows that the model can effectively interpret the structural behavior captured in the sensor features. Random Forest works by generating many decision trees and combining their outputs, which helps stabilize predictions and limit overfitting. Its strong performance in this binary classification task suggests that the method is well-suited for rapid damage-detection scenarios, where accurately distinguishing between normal and abnormal conditions is essential for timely maintenance and safety decisions. The SVM model, while performing reasonably well, showed a slight decrease in accuracy compared to the Random Forest model. Although it achieved high precision, it was less effective at distinguishing between Before Damage and After Damage in some cases, as reflected in the confusion matrix.

The challenges faced by SVM in this task likely stem from the model's reliance on finding optimal hyperplanes in high-dimensional feature spaces, which may not always align with the underlying complexity of the data. This is particularly true for datasets where the classes are not well separated and where SVMs' performance can degrade due to their sensitivity to outliers and class imbalance. However, SVM still performed well overall and provides a strong alternative for binary classification tasks, provided that proper parameter tuning is performed. In the BIM-AI Integrated dataset, which focused on classifying construction projects into three risk categories (Low, Medium, and High), Random Forest again outperformed SVM, demonstrating high precision, recall, and F1-score. The ability of Random Forest to effectively distinguish between the three risk levels was crucial for accurately assessing the health of construction projects. The model was particularly effective in identifying High-risk projects, which are critical for allocating resources and ensuring safety. This indicates that Random Forest is well-suited for multiclass classification tasks, even when the decision boundaries between classes are not immediately clear. SVM, in contrast, showed slightly reduced performance in multiclass classification, with a few misclassifications observed between Low and Medium risk levels.

This highlights the inherent difficulty of applying SVM to multiclass problems, where decision boundaries are more complex, and the model may require additional fine-tuning or alternative approaches such as one-vs-rest or one-vs-one classification. The Confusion Matrices for both models on the Vänersborg Bridge and BIM-AI datasets provided a clear visualization of model performance. The matrices showed that Random Forest achieved high accuracy in classifying damage states and risk levels, whereas SVM exhibited some minor misclassifications. The results suggest that Random Forest is particularly effective for datasets with complex feature relationships, as it can handle a large number of features without overfitting. The ROC curves further validated this performance, showing that Random Forest achieved a high AUC, indicating strong discrimination between the classes in both datasets. These results are consistent with existing literature on the effectiveness of Random Forest for damage detection and risk classification tasks. Previous studies have shown that Random Forest is a powerful tool for multiclass classification, with a high tolerance for noise and complex data structures. Additionally, Random Forest's feature importance makes it an excellent tool for identifying key factors contributing to damage or risk in infrastructure systems, offering valuable insights for engineers and decision-makers.

However, while Random Forest performed exceptionally well, the SVM model still holds value, especially in simpler binary classification tasks or when computational resources are limited. SVMs' ability to operate in high-dimensional spaces makes them a useful alternative for certain types of data, provided they are properly optimized. Regarding the BIM-AI Integrated dataset, the classification results are significant because they demonstrate the potential of machine learning models to predict the risk levels of construction projects. The ability to predict High-risk projects with high precision enables construction managers to allocate resources efficiently and prioritize projects that require closer monitoring. The Random Forest model's ability to handle complex features such as cost deviations, safety risk scores, and project delays shows that machine learning models can provide a comprehensive analysis of construction project health. Moreover, the analysis revealed that while SVM provided competitive results, its performance was not as robust as Random Forest, especially in handling the nuances of the multiclass classification. These challenges suggest that further optimization and experimentation with SVM's hyperparameters and multiclass classification strategies could yield better results.

From a broader perspective, the experimental findings are consistent with the trends reported in recent SHM literature. Reviews by Flah et al. [4] and Plevris and Papazafeiropoulos [17] note that tree-based ensembles and SVM often outperform simpler baselines on sensor-based damage detection tasks, especially when datasets are relatively small and noisy. Our results, where Random Forest consistently outperforms SVM on both the Vänersborg bridge data and the BIM-AI risk dataset, support this observation and provide an additional field-data case study. At the same time, the very high accuracies achieved here reflect the controlled nature of the datasets (a single bridge with a well-documented fracture event and a curated project-risk table) and do not guarantee similar performance under multi-year, multi-bridge monitoring campaigns. This underscores the need for future work on long-term validation and cross-structure generalization before models of this type are used as primary decision-making tools in operational asset management.

6. Conclusion

In this study, researchers developed an AI-based structural health monitoring (SHM) system for bridges using real IoT sensor data. Researchers collected and preprocessed data from two sources: A publicly available BIM-AI-integrated dataset and a real-world Vänersborg bridge dataset. The Vänersborg dataset includes readings from strain sensors, accelerometers, an inclinometer, and a weather station, recorded before, during, and after a verified damage event. Researchers performed data cleaning, noise filtering, normalization, and feature extraction on the raw sensor signals to prepare them for modeling. A Random Forest model, a Support Vector Machine, and a Neural Network were among the machine learning classifiers that researchers trained and assessed. The Random Forest classifier consistently produced the best overall performance and accuracy. On the BIM-AI dataset, the Random Forest model accurately distinguished between Low-, Medium-, and High-risk projects. On the Vänersborg dataset containing an actual damage incident, our system successfully identified the damage event with high precision. These results align with prior research on bridge SHM, confirming that Random Forest is an effective algorithm for structural damage detection.

6.1. Key Contributions

- **Integrated Diverse Datasets:** Researchers combined two distinct datasets (BIM-AI and Vänersborg) covering both simulated project data and real-world bridge sensor recordings. This integration demonstrates the method's applicability to varied SHM scenarios.
- **Advanced Data Preprocessing Pipeline:** Researchers designed a practical data processing workflow for real bridge sensor streams, including noise reduction, outlier removal, handling missing values, and extracting relevant features (frequency-domain vibration metrics, strain statistics, and environmental factors).

- **Comprehensive Machine Learning Analysis:** Researchers implemented and compared several supervised learning algorithms (Random Forest, SVM, Neural Network) on the SHM datasets. The study shows that Random Forest provides robust, reliable damage detection and benchmarks its performance against other models.
- **Real-World Validation with Results:** Researchers validated the AI-based SHM system on the actual Vänersborg bridge dataset, successfully detecting a known damage event. Researchers also generated detailed experimental Figures (accuracy plots, confusion matrices, and feature importance charts) to illustrate model performance and support our findings.
- **Demonstrated Practical Workflow:** This paper demonstrates a complete workflow from IoT data acquisition to AI analysis, offering a template for deploying AI-driven SHM systems in practice.

6.2. Limitations

- **Limited Data Diversity:** Our analysis used only two datasets. Both datasets represent specific bridge types and sensor configurations, so the models may not generalize to other bridge designs or sensor setups without additional adaptation.
- **Sensor Data Quality:** Real IoT sensor data contains noise, missing readings, and variations due to environmental conditions. Despite preprocessing, these factors can reduce detection accuracy or lead to false alarms in practice.
- **Offline analysis:** The system was evaluated on historical data in an offline setting. Researchers did not deploy the models in a live monitoring environment, so their real-time performance and stability under continuous operation remain untested.

6.3. Recommendations for Future Work

- **Expand and Diversify Datasets:** Collect and incorporate additional monitoring data from various bridges (with different designs, materials, and sensor deployments) to improve model generalization and robustness.
- **Explore Advanced Models:** Investigate deep learning approaches (such as CNNs and recurrent neural networks) and ensemble techniques to capture better complex temporal patterns in vibration data than traditional models.
- **Implement Real-Time Analytics**: Develop streaming data pipelines and deploy models on edge or cloud platforms to enable online damage detection and immediate alerts during live monitoring.
- **Transfer Learning and Adaptation:** Research techniques to adapt a trained model to new bridges or conditions, reducing the need for labeled data in each new deployment.
- **Improve Interpretability:** Use model explanation tools (e.g., SHAP values) to make AI decisions more transparent and quantify confidence or uncertainty in predictions.
- **Integrate with Maintenance Planning:** Connect the SHM system outputs to asset management workflows so that detected anomalies directly inform inspection and repair scheduling.

Acknowledgment: The author would like to express sincere gratitude to the University of the West of Scotland for providing the academic support, resources, and facilities that contributed to the successful completion of this research.

Data Availability Statement: The data supporting the findings of this research are available from the corresponding author upon reasonable request.

Funding Statement: The author declares that this research was undertaken independently and did not receive any financial assistance, sponsorship, or grant support from any public, private, or non-profit organization.

Conflicts of Interest Statement: The author confirms that there is no competing financial, professional, or personal interests that could have influenced the design, execution, interpretation, or publication of this study. All referenced works have been appropriately cited and acknowledged.

Ethics and Consent Statement: The author affirms that the study was conducted in accordance with recognized ethical standards and research protocols. Participation was voluntary, informed consent was obtained from all participants, and appropriate measures were implemented to ensure confidentiality, privacy, and the responsible handling of research data.

References

1. A. R. Al-Ali, S. Beheiry, A. Alnabulsi, S. Obaid, N. Mansoor, N. Odeh, and A. Mostafa, "An IoT-based road bridge health monitoring and warning system," *Sensors*, vol. 24, no. 2, p. 469, 2024.

2. S. Bhatta and J. Dang, "Use of IoT for structural health monitoring of civil engineering structures: A state-of-the-art review," *Urban Lifeline*, vol. 2, no. 11, p. 17, 2024.
3. Y. J. Cha, R. Ali, J. Lewis, and O. Büyüköztürk, "Deep learning-based structural health monitoring," *Automation in Construction*, vol. 161, no. 5, p. 105328, 2024.
4. M. Flah, J. Nunez, W. Ben Chaabene, and M. L. Nehdi, "Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review" *Archives of computational methods in engineering*, vol. 28, no. 7, pp. 2621-2643, 2021.
5. M. Matt, "The Role of Digital Twin Technology in Structural Health Monitoring (SHM)," *Resensys Blog*, Maryland, United States of America, 2025.
6. A. Yaghoubzadehfard, F. Lumantarna, N. Herath, M. Sofi, and M. Rad, "Ensemble learning-based structural health monitoring of a bridge using an interferometric radar system," *Journal of Civil Structural Health Monitoring*, vol. 14, no. 4, pp. 1629–1650, 2024.
7. Resensys, "Bridge Monitoring," *Resensys*, Beltsville, Maryland, United States of America, 2025.
8. Federal Highway Administration, "Deficient Bridges by Functional Classification Count 2011," *U.S. Department of Transportation*, Washington, District of Columbia, United States of America, 2011.
9. American Society of Civil Engineers, "2013 Report Card," *ASCE Infrastructure Report Card*, Reston, Virginia, United States of America, 2021.
10. J. Leander, J. Nyman, R. Karoumi, P. Rosengren, and G. Johansson, "Dataset for damage detection retrieved from a monitored bridge pre and post verified damage," *Data in Brief*, vol. 51, no. 12, p. 109729, 2023.
11. Z. Ziya, "BIM-AI Integrated Dataset," *Kaggle*, 2025. [Accessed by 15/04/2025].
12. K. Roy, "Reincarnating Experience in Education," *Springer*, Cham, Switzerland, 2020.
13. M. Zhou, J. Liao, and L. An, "Shear properties of tapered box girders with steel trapezoidally corrugated webs considering resal effect," *Journal of Bridge Engineering*, vol. 25, no. 1, p. 04019126, 2020.
14. Y. Lei, X. Yang, J. Huang, F. Zhang, and L. Liu, "Identification of model-free hysteretic forces of magnetorheological dampers embedded in buildings under unknown excitations using incomplete structural responses," *Structural Control and Health Monitoring*, vol. 28, no. 5, p. e2715, 2021.
15. G. Gui, H. Pan, Z. Lin, Y. Li, and Z. Yuan, "Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection," *KSCE Journal of Civil Engineering*, vol. 21, no. 2, pp. 523–534, 2017.
16. S. Wasiq and A. Golroo, "Smartphone-based cost-effective pavement performance model development using a machine learning technique with limited data," *Infrastructures*, vol. 9, no. 1, p. 9, 2024.
17. V. Plevris and G. Papazafeiropoulos, "AI in structural health monitoring for infrastructure maintenance and safety," *Infrastructures*, vol. 9, no. 12, p. 225, 2024.
18. A. Silik, M. Noori, Z. Wu, W. A. Altabey, J. Dang, and N. S. D. Farhan, "Wavelet-based vibration denoising for structural health monitoring," *Urban Lifeline*, vol. 2, no. 1, pp. 1-14, 2024.
19. M. Muttillo, V. Stornelli, R. Alaggio, R. Paolucci, L. Di Battista, T. de Rubeis, and G. Ferri, "Structural health monitoring: An IoT sensor system for structural damage indicator evaluation," *Sensors*, vol. 20, no. 17, p. 4908, 2020.
20. H. Singh, P. Kaushal, C. Satheesh, S. Manikandan, and D. Lale, "The Advancement of Scaffolds Through 3D Printing Techniques for Tissue Engineering Applications," *AVE Trends in Intelligent Energy Letters*, vol. 1, no. 1, pp. 11–20, 2025.
21. S. Kim, D. Culler, and J. Demmel, "Structural health monitoring of the Golden Gate Bridge using wireless sensor networks," *University of California*, California, United States of America, 2004.
22. 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.
23. E. Cachot, T. Vayssade, M. Virlogeux, H. Lancon, Z. Hajar, and C. Servant, "The Millau viaduct: Ten years of structural monitoring," *Structural Engineering International*, vol. 25, no. 4, pp. 375–380, 2015.
24. M. Calo, S. Ruggieri, M. Buitrago, A. Nettis, J. M. Adam, and G. Uva, "An ML-based framework for predicting prestressing force reduction in reinforced concrete box-girder bridges with unbonded tendons," *Engineering Structures*, vol. 325, no. 2, p. 119400, 2025.
25. E. Zhu, Z. Bai, L. Zhu, and Y. Li, "Research on bridge structure SAM based on real-time monitoring," *Journal of Civil Structural Health Monitoring*, vol. 12, no. 3, pp. 725–742, 2022.
26. J. Nyman, J. Rosengren, P. Kool, R. Karoumi, J. Leander, and H. Petursson, "Smart condition monitoring of a steel bascule railway bridge," *In Life-cycle of structures and infrastructure systems*, Abingdon, United Kingdom, 2023.