International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group | Volume 7 Issue 3 PP 26-55, 2026
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V7I3P104

Original Article

Physics-Guided Machine Learning for Rapid Nonlinear Seismic Safety Assessment of Reinforced-Concrete Frame–Wall Buildings

Abdullah Youssef
Teesside University, Egypt.

Received On: 18/05/2026 ***Revised On: 12/06/2026*** ***Accepted On: 23/06/2026*** ***Published On: 04/07/2026***

Abstract - Purpose: *This study develops a machine-learning-assisted framework for rapid seismic safety assessment of reinforced-concrete frame-wall buildings using nonlinear seismic analysis outputs. The framework is intended to reduce the time required for preliminary post-earthquake structural evaluation while retaining key indicators of nonlinear building response, including inter-storey drift, roof displacement, plastic-hinge development, and seismic performance state.* ***Methodology:*** *A detailed reinforced-concrete building model is developed in ETABS using available structural geometry, material properties, member sections, loading conditions, diaphragm assignments, and seismic parameters. Nonlinear static pushover and nonlinear response-history analyses are used to generate seismic-response data under systematically varied structural and ground-motion conditions. Structural properties, elastic-response indicators, modal characteristics, and earthquake-intensity measures are used as input variables for machine-learning models. Regression models predict continuous seismic-demand parameters, while classification models identify building safety states. Model performance is evaluated using grouped validation procedures, prediction-error measures, safety-classification metrics, and explainable feature-importance analysis.* ***Findings:*** *The study is expected to show that machine-learning models trained on nonlinear seismic simulations can estimate key response measures and classify seismic safety conditions with substantially lower computational time than repeated nonlinear analyses. The evaluation will identify the most reliable model for drift and safety-state prediction, quantify false-safe classification risk, and determine the structural and seismic variables that most strongly influence predicted performance.* ***Unique Contribution to Theory, Practice and Policy:*** *The study contributes a physics-guided, data-driven seismic assessment approach that links detailed nonlinear structural analysis with rapid predictive modelling for reinforced-concrete buildings. It provides a practical basis for prioritising detailed inspections, emergency-response decisions, and post-earthquake safety screening. The framework should be applied as a decision-support tool alongside professional engineering judgement, particularly where predicted risk or model uncertainty is high.*

Keywords *- Reinforced-Concrete Buildings, Nonlinear Seismic Analysis, Machine Learning, Rapid Seismic Safety Assessment, ETABS, Inter-Storey Drift, Performance-Based Earthquake Engineering, Structural Health Monitoring.*

1. Introduction

1.1. Background and Practical Seismic-Safety Challenge

Reinforced-concrete buildings remain a dominant structural form in many urban areas because of their durability, availability of materials, and adaptability to a wide range of architectural layouts. However, earthquakes can impose severe cyclic demands on reinforced-concrete frames, columns, beams, walls, and beam-column joints. These demands may lead to excessive inter-storey drift, stiffness degradation, plastic-hinge formation, shear failure, residual deformation, and, in extreme cases, partial or total collapse. The need to identify unsafe buildings promptly after an earthquake is therefore central to emergency response, public safety, repair planning, and the prioritisation of detailed engineering inspections.

Conventional seismic assessment procedures are commonly based on code-oriented checks, linear elastic analysis, nonlinear static analysis, or nonlinear response-history analysis. Performance-based assessment methods provide a more realistic basis for evaluating structural behaviour because they consider deformation capacity, inelastic response, and damage progression under earthquake loading (ASCE/SEI, 2017; FEMA, 2018). In particular, engineering demand parameters such as peak inter-storey drift ratio, roof displacement, base shear, plastic-hinge rotation, and residual drift are widely used to evaluate the likely seismic performance of reinforced-concrete buildings.

Despite their value, detailed nonlinear analyses are not always suitable for rapid assessment. They require reliable structural drawings, material data, nonlinear member definitions, ground-motion records, convergence control, and careful interpretation by experienced engineers. These requirements become especially challenging when large numbers of buildings must be screened within a short period after a damaging earthquake. As a result, there is a continuing need for assessment approaches that can preserve the

engineering value of nonlinear analysis while reducing the time required to obtain an initial safety decision.

1.2. Limitations of Conventional Seismic Assessment

Linear static and response-spectrum analyses remain useful for design verification and preliminary evaluation. However, they may not fully represent the inelastic mechanisms that govern structural performance during strong ground motion. Reinforced-concrete buildings can experience localised damage that is not apparent from elastic force demands alone, particularly where column axial loads are high, stiffness is unevenly distributed, torsional effects are present, or shear walls and frames interact in a complex manner.

Nonlinear static pushover analysis provides an improved representation of structural capacity by tracing the relationship between base shear and roof displacement. It can also indicate the sequence of plastic-hinge formation and identify critical components. Nevertheless, pushover analysis is influenced by the assumed lateral load pattern, the selected control displacement, and the ability of a single loading pattern to represent higher-mode effects. Krawinkler and Seneviratna (1998) noted that pushover procedures offer useful engineering insight but should not be treated as a complete substitute for dynamic seismic analysis. Similarly, Fajfar (2000) demonstrated the value of nonlinear static procedures within performance-based design, while also recognising the importance of appropriate modelling and interpretation.

Nonlinear response-history analysis provides a more comprehensive representation of earthquake-induced demand because it captures the time-dependent interaction between structural properties and ground-motion characteristics. Incremental dynamic analysis can further examine how a structure responds as earthquake intensity increases, including the transition from minor damage to collapse-level performance (Vamvatsikos & Cornell, 2002). However, these methods are computationally expensive and may be impractical for repeated assessment across a wide building inventory. They also require careful record selection, intensity scaling, numerical convergence checks, and detailed nonlinear modelling of structural components.

The challenge is therefore not the absence of established seismic-analysis methods. Rather, it is the difficulty of applying those methods quickly and consistently when urgent screening is required. A rapid assessment system should be able to identify buildings that need immediate detailed evaluation while avoiding unsafe classifications that could expose occupants to unnecessary risk.

1.3. Machine Learning for Rapid Structural Assessment

Machine learning offers a possible route for reducing the time required to estimate nonlinear seismic response. A trained machine-learning model can learn relationships between structural characteristics, seismic intensity measures, elastic analysis results, and nonlinear response indicators. Once trained on a sufficiently representative dataset, the model can generate predictions in seconds rather than requiring repeated nonlinear analyses for each case.

Recent research has shown that machine-learning methods can support seismic damage prediction, structural response estimation, post-earthquake safety classification, and regional risk assessment. Morfidis and Kostinakis (2017) used neural-network approaches to predict damage states in reinforced-concrete buildings, while Zhang et al. (2018) developed a machine-learning framework for post-earthquake structural safety assessment. Mangalathu et al. (2020) further demonstrated that machine-learning methods can classify building damage using earthquake-related information, and Sun et al. (2021) reviewed the growing use of machine learning in structural design and performance assessment.

However, machine learning should not be treated as a replacement for nonlinear structural analysis. Its reliability depends on the quality of the training data, the representativeness of the structural scenarios, the relevance of the selected features, and the appropriateness of the validation procedure. A model trained on simplified or narrowly defined structures may perform well within its own dataset but fail when applied to different building configurations or seismic conditions. Xie et al. (2020) emphasised that machine-learning applications in earthquake engineering must be supported by sound engineering assumptions, transparent datasets, and meaningful validation.

For seismic safety assessment, the use of physics-guided machine learning is particularly important. The model should reflect basic structural principles. For example, predicted drift demand should generally increase as seismic intensity increases, while an increase in lateral stiffness or wall contribution should not produce an unrealistic increase in predicted displacement under otherwise identical conditions. Physics-informed modelling can help reduce physically inconsistent predictions and improve confidence in model outputs (Karniadakis et al., 2021).

1.4. Research Gap

Although machine learning has been applied to earthquake engineering problems, several limitations remain in the current literature. Many studies rely on simplified structural models, limited numerical datasets, or a narrow range of building configurations. Some studies use only geometric and material information as model inputs, without incorporating rapid elastic-response indicators such as storey drift, modal period, base shear, or torsional response. These elastic indicators can provide important structural information that is readily available from conventional ETABS analysis.

In addition, several machine-learning studies focus mainly on overall accuracy without giving sufficient attention to safety-critical errors. In seismic screening, a false-safe prediction is more serious than a false alarm because it may lead to an unsafe building being classified as suitable for occupancy. Therefore, a rapid seismic safety model should be assessed not only through accuracy and F1-score, but also through false-safe rate, unsafe-class recall, probability calibration, and prediction uncertainty.

Another limitation concerns validation. Randomly splitting simulation records into training and testing sets may lead to data leakage when several records originate from the same building configuration or the same ground-motion family. Such a process can produce overly optimistic performance estimates. Grouped validation procedures are more suitable because they keep related structural cases together and provide a more realistic test of model generalisability (Roberts et al., 2017).

Finally, many predictive models remain difficult to interpret. For safety-sensitive engineering applications, it is important to understand why the model predicts a particular drift demand or damage state. Explainable methods, including SHAP-based feature analysis, can identify the structural and seismic variables that drive the prediction and help engineers determine whether the model behaviour is consistent with structural principles (Lundberg & Lee, 2017). This need is particularly important where black-box predictions may influence inspection priorities or post-earthquake access decisions (Rudin, 2019).

1.5. Aim and Objectives of the Study

This study aims to develop a machine-learning-assisted framework for rapid seismic safety assessment of reinforced-concrete frame-wall buildings using nonlinear seismic analysis outputs generated from a detailed ETABS model. The framework is designed to estimate key seismic-response measures and classify structural safety conditions without replacing detailed engineering assessment.

The study has five main objectives. First, it develops and verifies a detailed ETABS model of an existing reinforced-concrete building with frame, wall, diaphragm, material, and loading characteristics defined from available structural information. Second, it performs nonlinear static and nonlinear response-history analyses to generate structural response data under varying seismic and modelling conditions. Third, it develops machine-learning models capable of predicting nonlinear demand measures, including peak inter-storey drift ratio, roof displacement, and seismic performance state. Fourth, it compares different machine-learning methods using grouped validation procedures and safety-focused performance measures. Fifth, it applies explainable machine-learning analysis to identify the variables that most strongly influence predicted seismic behaviour.

The study addresses the following research questions:

1. Can machine-learning models predict nonlinear seismic demand indicators from structural properties, elastic ETABS outputs, and earthquake-intensity measures?
2. Which machine-learning model provides the best balance between predictive accuracy, interpretability, and safety reliability?
3. Can the proposed framework classify unsafe seismic conditions with a sufficiently low false-safe rate?
4. Which structural and seismic variables have the strongest influence on predicted drift demand and safety state?
5. Can the proposed framework retain acceptable performance when applied to an independent reinforced-concrete building configuration?

1.6. Contributions of the Study

The study makes the following contributions:

1. It develops a nonlinear seismic assessment dataset from a detailed reinforced-concrete frame-wall ETABS model rather than relying solely on simplified analytical frames.
2. It combines structural characteristics, modal properties, elastic seismic-response indicators, and earthquake-intensity measures as machine-learning input variables.
3. It evaluates both regression and classification models to predict continuous seismic demand measures and practical building safety states.
4. It introduces grouped validation and external testing to reduce data leakage and provide a more realistic assessment of model generalisability.
5. It gives specific attention to false-safe predictions, unsafe-class recall, and probability calibration, which are essential for seismic safety screening.
6. It uses explainable machine-learning methods to link predicted outcomes with structural and seismic response drivers.
7. It presents the machine-learning framework as a decision-support tool for rapid screening, while maintaining the need for detailed nonlinear analysis and professional engineering judgement for high-risk or uncertain cases.

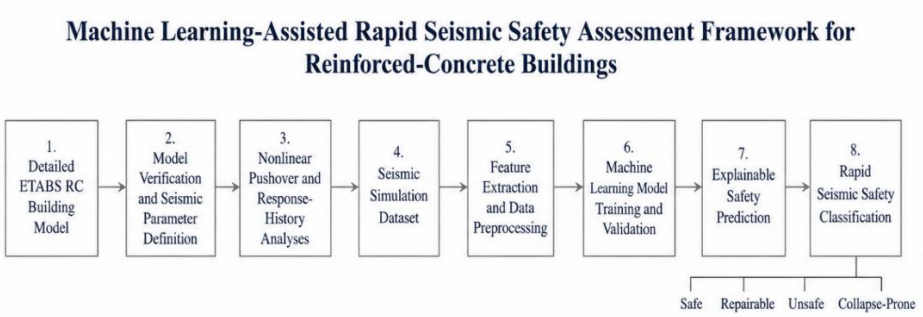

Fig 1: Machine learning-assisted workflow for rapid seismic safety assessment and classification of reinforced-concrete buildings.

2. Literature Review

2.1. Seismic Performance Assessment of Reinforced-Concrete Buildings

Seismic performance assessment seeks to determine how a structure is likely to behave when subjected to earthquake-induced forces and deformations. For reinforced-concrete buildings, the assessment process commonly considers global response measures such as base shear, roof displacement, inter-storey drift, and modal characteristics, together with local damage indicators including plastic-hinge rotation, shear demand, concrete crushing, reinforcement yielding, and column axial-load effects. These response measures are important because structural damage is often governed by deformation capacity rather than by elastic force demand alone.

Conventional seismic assessment has historically relied on linear static and response-spectrum procedures because they are relatively efficient and are widely incorporated into structural design practice. These methods remain valuable for verifying global strength, drift limits, and seismic demand distribution. However, their ability to represent severe inelastic response is limited. Under strong ground motion, reinforced-concrete structures may develop localised yielding, stiffness degradation, strength deterioration, residual deformation, and changes in force distribution that cannot be fully captured through elastic analysis.

Performance-based earthquake engineering addresses this limitation by relating structural demand to defined performance objectives. Common performance levels include Immediate Occupancy, Life Safety, Collapse Prevention, and Collapse. These levels provide a structured means of interpreting seismic response in terms of expected damage, repairability, and life-safety implications. ASCE/SEI 41-17 and FEMA P-58 provide widely recognised frameworks for evaluating structural performance using deformation-based criteria and probabilistic damage assessment methods (ASCE/SEI, 2017; FEMA, 2018).

Nonlinear static pushover analysis has become a commonly applied approach for evaluating the capacity of reinforced-concrete buildings. In a pushover procedure, gravity loads are applied first, followed by gradually increasing lateral loads until a selected target displacement or structural limit state is reached. The resulting capacity curve provides a relationship between base shear and roof displacement, while the sequence of plastic-hinge formation helps identify critical elements and potential failure mechanisms. Fajfar (2000) showed that nonlinear static procedures can provide useful performance estimates when they are applied with appropriate assumptions regarding structural behaviour and seismic demand.

Despite its usefulness, pushover analysis has important limitations. The response of a building during an earthquake is influenced by higher-mode effects, record-to-record variability, cyclic deterioration, and the interaction between structural components. A single lateral load pattern may not fully represent these effects, especially in irregular buildings, taller structures, or systems with significant wall-frame interaction. Krawinkler and Seneviratna (1998) therefore cautioned that pushover analysis should be interpreted as an engineering assessment tool rather than a complete representation of dynamic seismic response.

Nonlinear response-history analysis provides a more detailed approach because it captures time-dependent structural behaviour under selected earthquake records. The method can represent cyclic inelastic response, dynamic amplification, stiffness degradation, and variation in structural demand across multiple intensity levels. Incremental dynamic analysis extends this approach by repeatedly scaling earthquake records to examine the progression of structural response from elastic behaviour to severe damage and collapse conditions (Vamvatsikos & Cornell, 2002). Although nonlinear response-history analysis offers a strong basis for performance assessment, it requires careful record selection, numerical convergence checks, detailed modelling, and substantial computational effort.

Ground-motion selection is also central to the reliability of nonlinear seismic analysis. The selected records should reflect the relevant seismic hazard, site conditions, structural period range, and expected spectral demand. Baker and Cornell (2006) highlighted the importance of spectral shape and epsilon in record selection, while Baker (2011) proposed the conditional mean spectrum as a practical method for selecting ground motions that are compatible with the expected seismic demand. These approaches are particularly relevant when nonlinear response-history results are used as reference data for machine-learning model development.

Overall, the literature shows that nonlinear seismic analysis remains the most reliable basis for understanding the damage progression and safety condition of reinforced-concrete buildings. However, its computational demands create a practical challenge for rapid screening, post-earthquake inspection, and assessment of large building inventories.

2.2. Nonlinear Modelling of Reinforced-Concrete Frame-Wall Structures

The accuracy of nonlinear seismic assessment depends strongly on how reinforced-concrete members and structural systems are represented. Reinforced-concrete buildings often consist of beams, columns, slabs, shear walls, cores, coupling beams, and masonry infill panels. Their seismic behaviour is influenced by material strength, reinforcement details, axial load, stiffness distribution, confinement, cracking, shear capacity, and deterioration under cyclic loading.

Beam and column modelling commonly uses concentrated plastic hinges, distributed plasticity models, or fibre-section approaches. Concentrated hinge models are often preferred for practical building assessment because they are computationally efficient and can represent inelastic deformation at critical member ends. However, hinge properties should reflect the expected flexural, axial, and shear behaviour of the member. Panagiotakos and Fardis (2001)

developed relationships for estimating the deformation capacity of reinforced-concrete members at yielding and ultimate conditions, providing a useful basis for nonlinear hinge definition.

Columns require particular attention because their seismic behaviour depends on the interaction between axial force and bending moment. High axial load can reduce deformation capacity and increase the likelihood of brittle failure. Elwood and Moehle (2005) demonstrated that the drift capacity of reinforced-concrete columns is strongly influenced by transverse reinforcement and confinement conditions. For this reason, nonlinear models should consider axial-force effects, moment interaction, and possible shear sensitivity, especially in lower-storey columns where gravity and lateral demands are high.

The modelling of shear walls and core systems is equally important in frame-wall buildings. Shear walls contribute substantial lateral stiffness and strength, but their behaviour can change as cracking, yielding, and local damage develop. Wall piers may experience flexural yielding, shear damage, compression crushing, or coupling-beam distress depending on the structural configuration. A simplified elastic wall representation may be sufficient for preliminary analysis, but nonlinear seismic performance assessment requires a representation that captures the main deformation mechanisms and expected damage concentration.

Stiffness modifiers are another important modelling consideration. Reinforced-concrete members do not retain their gross-section stiffness after cracking. The use of reduced stiffness values for beams, columns, walls, and slabs helps produce more realistic estimates of modal periods, drift demand, and force distribution. The existing ETABS model already includes stiffness modifiers for several frame members, which provides a useful starting point for the nonlinear modelling stage. However, these values should be reviewed to ensure consistency with the selected seismic code, structural detailing, and intended performance-analysis framework.

P-Delta effects should also be retained because gravity loads acting through lateral displacement can increase global instability and drift demand. This effect may be particularly significant in buildings with soft-storey behaviour, high gravity load, slender columns, or substantial displacement demand. The use of iterative P-Delta settings in the current model is therefore beneficial, although the influence of these effects should be checked through sensitivity analysis.

Cyclic deterioration remains a major source of uncertainty in nonlinear reinforced-concrete modelling. Ibarra et al. (2005) developed hysteretic models that account for strength and stiffness degradation, which are important for representing response under repeated load reversals. While detailed deterioration models may increase computational demand, they are valuable where collapse prevention, residual drift, or post-yield response is a primary concern.

The literature therefore confirms that nonlinear modelling should be treated as more than a software operation. It requires engineering judgement regarding member behaviour, hinge definition, wall response, material properties, stiffness reduction, gravity-load sequencing, and numerical stability. These modelling decisions directly influence the response data that will later be used to train and validate the machine-learning framework.

2.3. Machine Learning in Structural and Earthquake Engineering

Machine learning has gained increasing attention in structural and earthquake engineering because it can identify complex relationships within large datasets and produce rapid predictions once a model has been trained. In seismic applications, machine-learning methods have been used for damage classification, response prediction, fragility estimation, structural-health monitoring, post-earthquake safety screening, and regional risk assessment.

Early machine-learning applications in structural engineering often focused on pattern recognition and structural-health monitoring. Worden and Manson (2007) discussed the role of machine learning in detecting changes in structural condition from measured data, establishing a foundation for later applications involving damage detection and response prediction. As computational resources and simulation datasets became more accessible, machine learning began to be used more directly for nonlinear seismic performance assessment.

Morfidis and Kostinakis (2017) applied neural-network models to predict the damage state of reinforced-concrete buildings using seismic and structural parameters. Their work demonstrated that data-driven models can estimate damage levels when relevant structural features and seismic descriptors are available. Zhang et al. (2018) further developed a machine-learning framework for assessing post-earthquake structural safety, showing that predictive models can support rapid decisions after damaging events.

Recent research has expanded the range of machine-learning methods used in earthquake engineering. Random Forest, Support Vector Machine, Gradient Boosting, XGBoost, neural networks, and kernel-based approaches have all been applied to structural-response estimation and seismic damage classification. Sun et al. (2019) used kernel-based machine-learning methods to reconstruct seismic response demands across multiple tall buildings, while Demertzis et al. (2023) focused on interpretable prediction of reinforced-concrete building response. These studies indicate that machine-learning models can estimate structural demand without repeatedly conducting a complete nonlinear simulation.

The choice of input variables is particularly important. Some studies rely mainly on geometric and material properties, including building height, number of storeys, span length, concrete strength, wall ratio, and reinforcement details. Other studies incorporate seismic intensity measures

such as peak ground acceleration, peak ground velocity, spectral acceleration, duration, and spectral shape. A more informative approach combines these variables with elastic structural-response indicators, including fundamental period, modal mass participation, elastic drift, roof displacement, base shear, and torsional response.

Elastic analysis outputs are useful because they can be obtained relatively quickly from a conventional structural model. They also contain information about the dynamic characteristics and stiffness distribution of the building. When combined with structural properties and earthquake descriptors, these outputs may improve the ability of machine-learning models to estimate nonlinear drift demand and damage state.

However, machine-learning models must be developed carefully. A model may perform well within a limited dataset but fail when it encounters a different building configuration, seismic intensity range, or ground-motion family. Xie et al. (2020) noted that the successful use of machine learning in earthquake engineering depends on meaningful engineering features, representative data, transparent evaluation procedures, and awareness of model limitations.

Data leakage is another major concern. When many simulation records are generated from the same base building model, randomly dividing individual rows into training and testing sets can produce misleadingly high accuracy. The model may effectively learn the characteristics of a building already represented in the training data. Group-based validation is more appropriate because it separates complete structural variants, ground-motion sets, or simulation families between training and testing stages.

Machine learning is therefore most useful when it is treated as a surrogate or decision-support method. It can reduce the time required for preliminary assessment, but it should not replace nonlinear analysis for buildings with high predicted risk, significant uncertainty, or unusual structural characteristics.

2.4. Explainable Machine Learning and Safety-Critical Decisions

In seismic safety assessment, a highly accurate prediction is not sufficient unless the basis of that prediction can be understood and evaluated. Structural engineers need to know whether the model responds to meaningful variables, whether its predictions follow expected physical trends, and whether the uncertainty associated with those predictions is acceptable.

Explainable machine-learning methods help address this need. SHAP analysis, introduced by Lundberg and Lee (2017), can quantify the contribution of individual features to a prediction. In the context of seismic assessment, SHAP values can show how variables such as spectral acceleration, wall-area ratio, fundamental period, axial-load ratio, elastic drift, or storey stiffness influence the predicted seismic response. This information can help engineers determine whether the model is responding in a reasonable manner.

Physics-guided learning is also important. Karniadakis et al. (2021) explained that physics-informed machine learning can improve model reliability by incorporating physical laws, constraints, or expected response patterns. For this study, physics-guided principles may include the expectation that predicted damage probability should increase as earthquake intensity rises and that predicted drift should not increase when structural stiffness is increased under otherwise identical conditions.

Safety classification requires particular attention to false-safe predictions. A false-safe prediction occurs when a building with severe nonlinear response is classified as safe or repairable. Such errors can have serious consequences because they may influence occupancy decisions, inspection priorities, and emergency response. For this reason, the proposed framework should report unsafe-class recall, false-safe rate, calibration performance, and prediction confidence in addition to overall accuracy and F1-score.

Rudin (2019) argued that high-stakes decisions should not rely entirely on opaque black-box models when interpretable alternatives are available. This principle is highly relevant to seismic safety assessment. The machine-learning model should support engineers by highlighting likely risk, key response drivers, and uncertainty, rather than producing an unexplained final judgement.

Table 1: Summary of Selected Studies Relevant to Machine Learning-Assisted Seismic Assessment

Study	Main Focus	Structural or Seismic Context	Key Contribution	Limitation Relevant to Present Study
Morfidis and Kostinakis (2017)	Neural-network prediction of damage state	Reinforced-concrete buildings	Demonstrated data-driven damage-state prediction	Limited emphasis on external validation and safety-critical errors
Zhang et al. (2018)	Post-earthquake structural safety assessment	Building safety classification	Established ML use for rapid safety screening	Requires further integration with detailed nonlinear modelling
Sun et al. (2019)	Seismic response reconstruction	Multiple tall buildings	Showed response prediction using kernel-based methods	Focused on response reconstruction rather than RC safety-state classification

Mangalathu et al. (2020)	Earthquake damage classification	Building damage assessment	Demonstrated machine-learning classification of damage outcomes	Does not focus specifically on nonlinear ETABS simulation datasets
Xie et al. (2020)	Review of ML in earthquake engineering	Broad earthquake engineering applications	Identified opportunities and methodological limitations	Highlights need for transparent validation and engineering interpretation
Sun et al. (2021)	Review of ML in structural assessment	Structural design and performance assessment	Summarised applications and research directions	Indicates limited standardisation of datasets and validation methods
Demertzis et al. (2023)	Interpretable ML prediction	Reinforced-concrete buildings	Emphasised interpretability of seismic response prediction	Further testing is needed across independent structural configurations
Shahnazaryan and O'Reilly (2024)	Nonlinear and collapse prediction	Multi-period structural systems	Extended ML use toward nonlinear and collapse-related response	Requires building-specific implementation and safety-focused assessment

2.5. Literature Gap and Conceptual Positioning

The reviewed literature confirms that nonlinear seismic analysis remains the most dependable method for evaluating reinforced-concrete building performance. It also shows that machine learning can reduce the time required to estimate seismic response and damage state. However, an important gap remains between detailed nonlinear structural modelling and rapid, interpretable safety assessment.

Many existing studies either rely on simplified building models, focus on a limited range of input features, or assess model performance primarily through overall predictive accuracy. Fewer studies combine detailed frame-wall building modelling, elastic ETABS response indicators, nonlinear seismic simulations, safety-sensitive classification metrics, grouped validation, and explainable machine-learning analysis within one framework.

The present study addresses this gap by developing a machine-learning-assisted seismic assessment approach based on a detailed reinforced-concrete frame-wall ETABS model. The framework will use nonlinear static and nonlinear response-history analyses as the reference source for seismic demand and safety-state labels. It will then combine building characteristics, elastic-response outputs, modal properties, and seismic intensity measures to predict nonlinear response indicators rapidly.

The conceptual position of the study is therefore not to replace nonlinear seismic analysis. Instead, it is to use nonlinear analysis as the engineering foundation for a faster assessment tool. The proposed framework is intended to identify buildings that are likely to be safe, repairable, unsafe, or collapse-prone, while directing high-risk and uncertain cases toward detailed analysis and professional engineering inspection.

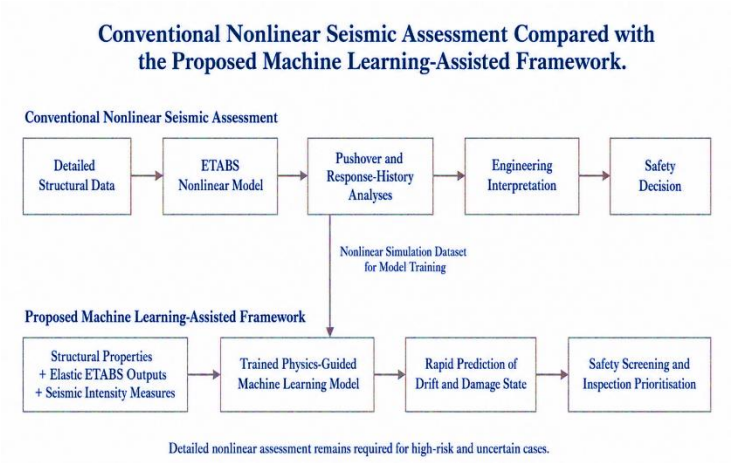

Fig 2: Comparison of Conventional Nonlinear Seismic Assessment with the Proposed Machine Learning-Assisted Framework for Rapid Safety Screening of Reinforced-Concrete Buildings.

3. Case Study: Building an ETABS Numerical Model

3.1. Case-Study Building Description

The case study is based on an existing reinforced-concrete building model developed in ETABS version 22.7.0. The model was provided in two related parts, identified as Part 01 and Part 02, and represents a multi-level reinforced-concrete building with frame members, wall elements, slabs, diaphragm assignments, and seismic load definitions. The two model parts contain closely related material definitions and structural systems, while Part 02 includes some additional beam and column section types. They will be used primarily to support model development, parametric analysis, and, where appropriate, independent testing of the proposed machine-learning framework.

The ETABS model is expressed in ton-force, metres, and degrees Celsius. It includes a base level, two lower levels, a ground level, and five upper ceiling levels. The vertical configuration creates eight principal storey intervals from the base to the highest modelled level. The two lower storeys have heights of 4.40 m and 3.95 m, respectively, while the ground-storey height is 3.25 m. The upper storeys have a typical

height of 3.10 m. This vertical arrangement is suitable for examining the influence of changes in storey stiffness, mass distribution, and lateral-force transfer on seismic response.

The floor system is represented using semi-rigid diaphragms. Eight diaphragm assignments, identified as D1 to D8, are defined in the model. Semi-rigid diaphragm modelling is appropriate where slab in-plane deformation may influence the distribution of lateral forces between frames, walls, and core elements. This is particularly relevant in reinforced-concrete buildings with non-uniform wall layouts, irregular plan geometry, or substantial differences in lateral stiffness between structural components.

The final manuscript will include a verified architectural and structural description of the building, including plan dimensions, bay spacing, occupancy category, foundation arrangement, and wall layout. These details should be extracted from the structural drawings and checked against the ETABS geometry before final submission. The present numerical model provides the principal basis for the seismic assessment, while the final drawings will be used to confirm that the analytical representation remains consistent with the constructed structural arrangement.

3.2. Structural System and Lateral Load-Resisting Mechanism

The building is modelled as a reinforced-concrete frame-wall structural system. The gravity-load system is carried by reinforced-concrete slabs, beams, and columns, while lateral resistance is provided through the combined action of moment-resisting frames, reinforced-concrete walls, wall piers, and core-related components. The ETABS model contains a range of beam and column sections together with wall and pier components, confirming that the seismic response is governed by frame-wall interaction rather than by a frame-only system.

This type of structural arrangement can provide improved lateral stiffness and strength compared with a pure moment-resisting frame. The walls and core components are expected to resist a substantial proportion of lateral shear and overturning demand, while the moment frames contribute to lateral stiffness, energy dissipation, and redistribution of seismic actions. However, the interaction between walls and frames can lead to uneven storey drift, force concentration in columns adjacent to walls, and torsional sensitivity where wall placement is asymmetric.

For this reason, the model verification stage will quantify the lateral contribution of the wall and frame subsystems in each principal direction. The analysis will evaluate modal periods, storey shear distribution, displacement profiles, torsional response, and drift concentration. These checks are necessary because a nominal structural-system description alone does not fully capture the actual lateral behaviour of a building.

The model contains automatic rigid-zone areas for several column members. This feature improves the representation of beam-column joint regions by accounting for the stiffened zone near member intersections. The use of realistic joint and member-end assumptions is important in nonlinear seismic assessment because plastic deformation is expected to concentrate near beam-column connections and wall boundaries. Panagiotakos and Fardis (2001) emphasised that member deformation capacity and yielding characteristics should be represented carefully when evaluating reinforced-concrete seismic response.

3.3. Material Properties and Section Characteristics

The material library includes separate reinforced-concrete definitions for beams and slabs, columns and walls, and reinforcing steel. The principal concrete material assigned to beams and slabs is labelled **FCU 29**, while the concrete material assigned to columns and walls is labelled **FCU 35**. These labels correspond to nominal concrete-strength classes of approximately 29 MPa and 35 MPa, respectively. The model assigns a unit weight of 2.5 ton-force per cubic metre to the concrete materials and a Poisson's ratio of 0.20.

The reinforcement library includes several steel and rebar definitions. The primary reinforcing material used in the reinforced-concrete section definitions is RMAT, which has a specified yield strength corresponding approximately to 400 MPa. Additional reinforcing-steel definitions are present in the source model, including A615 Grade 60 and other rebar datasets. The final nonlinear model will use the reinforcement properties assigned to the active beam, column, and wall sections rather than relying on unused material entries stored in the ETABS material library.

The beam sections are predominantly rectangular, with representative dimensions including 200 × 600 mm, 200 × 700 mm, 200 × 800 mm, 200 × 950 mm, 250 × 750 mm, 300 × 600 mm, 300 × 800 mm, 400 × 800 mm, and deeper transfer-type sections such as 200 × 1400 mm and 950 × 800 mm. The variety of beam depths reflects the different span lengths, loading conditions, and local transfer requirements within the building.

The columns also show substantial variation in size and reinforcement detailing. Representative sections include 300 × 600 mm, 300 × 700 mm, 350 × 800 mm, 350 × 850 mm, 400 × 600 mm, 400 × 800 mm, 400 × 950 mm, 400 × 1200 mm, 400 × 1250 mm, 400 × 1500 mm, 500 × 1200 mm, and 500 × 1300 mm. Several column labels include reinforcement details, such as 12T16, 16T16, 20T18, 24T18, and 26T18. These variations are useful for the proposed study because section dimensions, reinforcement ratio, axial-load ratio, and stiffness distribution can be included as input features in the machine-learning dataset.

The existing linear model applies cracked-section stiffness modifiers. Beam sections generally use bending-stiffness modifiers of 0.50 and torsional-stiffness modifiers of 0.20, while many column sections use bending-stiffness modifiers of 0.70. These assumptions are consistent with the need to account for cracking in reinforced-concrete members during seismic analysis. The stiffness assumptions will be

reviewed before nonlinear analysis because predicted modal periods, drift demand, and wall-frame force distribution are sensitive to the adopted effective stiffness values.

Table 2: Geometry and Global Structural Characteristics of the Case-Study Building

Parameter	Description	Model Information
Software platform	Structural analysis software	ETABS Version 22.7.0
Analysis units	Force, length, temperature	Ton-force, metre, degree Celsius
Model configuration	Structural model files	Part 01 and Part 02
Vertical levels	Base to roof arrangement	Base, two lower levels, ground level, five upper levels
Number of storey intervals	Principal vertical intervals	Eight
Typical upper-storey height	First to fifth ceiling levels	3.10 m
Ground-storey height	Ground to first ceiling	3.25 m
Lower-storey heights	Two lower levels	3.95 m and 4.40 m
Floor diaphragm model	In-plane floor behaviour	Semi-rigid diaphragms D1 to D8
Structural system	Primary lateral mechanism	Reinforced-concrete frame-wall system
Main analysis purpose	Research application	Nonlinear seismic assessment and machine-learning dataset generation

3.4. Gravity Loads, Mass Source, and Load Combinations

The existing ETABS model contains gravity and seismic loading definitions that will be retained as the basis for the numerical assessment after detailed review. Gravity actions include self-weight, superimposed dead loads, live loads, wall loads, and stair-related loads where applicable. These actions will be checked against the structural drawings and loading assumptions before nonlinear analysis is performed.

For nonlinear seismic analysis, gravity loads should be applied before lateral earthquake demand. This loading sequence is essential because the axial force carried by columns and walls influences their flexural strength, stiffness, deformation capacity, and susceptibility to instability. The gravity-load condition will therefore be established through a dedicated nonlinear static load case before pushover or nonlinear response-history analysis.

The mass source will include the appropriate permanent loads and the code-specified portion of imposed loads. A mass-source audit will be conducted to verify that the assigned seismic mass reflects the actual building configuration, including slab loads, beam and column self-weight, wall weight, superimposed dead loads, and applicable live-load contribution. The accuracy of the mass source is particularly important because it affects modal periods, inertia forces, base shear, and seismic displacement demand.

The final load combinations will be selected according to the adopted seismic design and assessment framework. Both gravity and seismic combinations will be reviewed to ensure consistency between the original ETABS model, the selected code provisions, and the intended nonlinear performance assessment. FEMA P-58 highlights the importance of consistent loading assumptions when linking structural analysis outputs with damage and performance evaluation (FEMA, 2018).

3.5. Seismic Parameters and Governing Code Framework

The original ETABS files include response-spectrum seismic definitions in the global X and Y directions, together with a vertical seismic-response case. The exported settings indicate a Eurocode 8 response-spectrum basis, with ground type B, design ground acceleration of 0.15g, a behaviour factor of 5.0, 5% damping, and accidental eccentricity of 5%.

These settings provide a defined starting point for the linear seismic model. However, the building was described as being designed according to Egyptian seismic provisions. The final study must therefore establish one consistent seismic framework before nonlinear analysis begins. The manuscript should not combine Egyptian-code design claims with Eurocode 8 seismic parameters unless the relationship between the two has been explicitly verified and justified.

Accordingly, the seismic spectrum, site classification, importance factor, response-modification assumptions, accidental eccentricity, load combinations, and acceptance criteria will be checked and standardised before generating the nonlinear dataset. Where Eurocode 8 is retained as the final basis, the paper will report its parameters transparently and cite the applicable provisions (Comité Européen de Normalisation, 2004). Where Egyptian seismic provisions are adopted, the response-spectrum definitions and associated analysis cases will be revised accordingly.

The model uses response-spectrum analysis as the initial dynamic-analysis approach. Modal properties will be obtained using a Ritz-vector modal case, with adequate participating mass required in the principal horizontal directions. The modal analysis will provide the fundamental periods, mass-participation ratios, mode shapes, and torsional-response indicators that will later serve as both model-verification measures and potential machine-learning input features.

3.6. Numerical Model Verification

Before nonlinear hinges, degradation rules, and dynamic ground-motion cases are introduced, the linear ETABS model will undergo a structured verification process. The first stage will examine member connectivity, support conditions, diaphragm assignments, section-property consistency, wall locations, and mass-source definition. The second stage will assess the global dynamic behaviour of the building through modal periods, modal mass participation, deformed shapes, and directional stiffness.

The verification process will also compare code-based seismic base shear with response-spectrum base shear and check whether scaling is required. Storey displacement, inter-storey drift, storey shear, overturning moment, and torsional response will be reviewed in both principal directions. Particular attention will be given to the lower levels, where larger axial loads, wall discontinuities, transfer elements, or stiffness changes may influence the seismic response.

The use of P-Delta effects will be retained throughout the verification and nonlinear-analysis stages. P-Delta effects can increase drift demand and reduce apparent lateral stiffness, especially where the building experiences significant displacement under gravity loading. Their inclusion is necessary for a credible assessment of seismic instability and collapse-related behaviour.

The verified linear model will then be converted into the nonlinear analytical model described in the following section. This conversion will include the assignment of nonlinear beam and column hinges, wall-pier response parameters, gravity-load sequencing, and nonlinear solution controls. The final nonlinear model will provide the reference results used to train and validate the machine-learning safety-assessment framework.

Table 3: Representative Material and Section Properties Extracted from the ETABS Model

Structural component	Material or section category	Representative properties
Beams and slabs	FCU 29 concrete	Nominal concrete strength approximately 29 MPa
Columns and walls	FCU 35 concrete	Nominal concrete strength approximately 35 MPa
Longitudinal reinforcement	RMAT rebar	Nominal yield strength approximately 400 MPa
Beam sections	Rectangular RC beams	200 × 600 mm to 950 × 800 mm, including deep beam sections
Column sections	Rectangular RC columns	300 × 600 mm to 500 × 1300 mm and 400 × 1500 mm
Beam stiffness modifiers	Cracked-section assumptions	Flexural stiffness modifier generally 0.50; torsional modifier generally 0.20
Column stiffness modifiers	Cracked-section assumptions	Flexural stiffness modifier generally 0.70
Diaphragm system	Semi-rigid floor diaphragm	D1 to D8

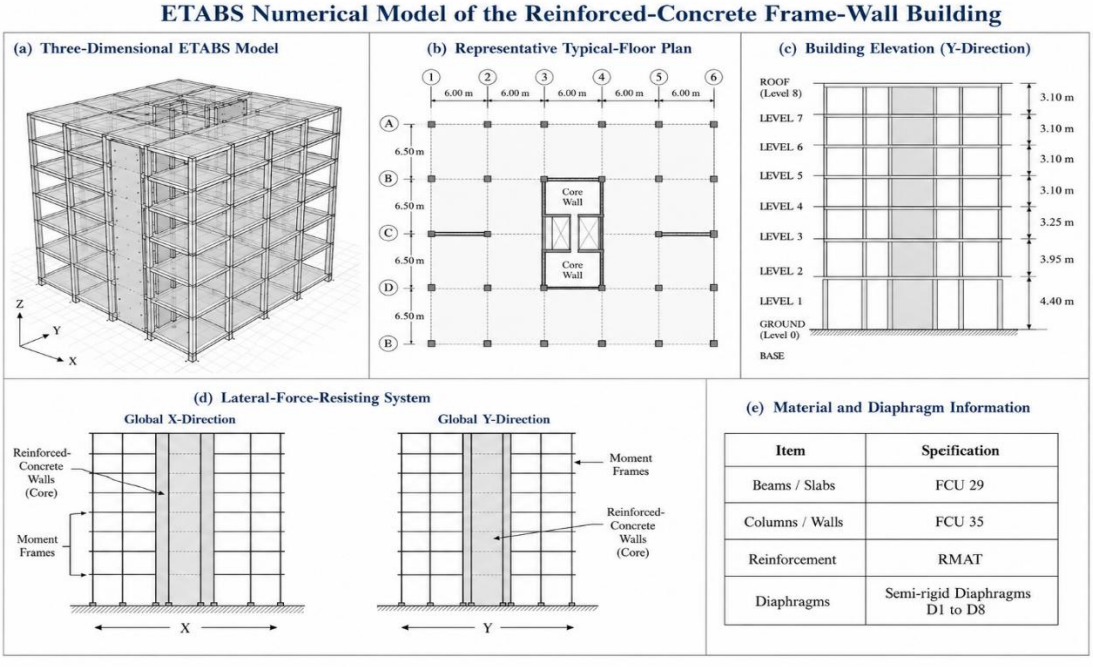

Fig 3: ETABS three-dimensional model and structural-system layout.

4. Nonlinear Seismic Analysis and Dataset Generation

The existing ETABS model provides the starting point for the nonlinear assessment programme. It already contains reinforced-concrete beam, column, wall, slab, and diaphragm definitions, together with cracked-section stiffness modifiers, response-spectrum loading, and iterative P-Delta settings. The model also distinguishes between FCU 29 concrete for beams and slabs and FCU 35 concrete for columns and walls, with RMAT reinforcement assigned to the reinforced-concrete section definitions.

However, the original model is primarily configured for linear seismic analysis. It does not provide a complete nonlinear simulation dataset with assigned plastic hinges, nonlinear wall-pier behaviour, pushover cases, or nonlinear response-history cases. Therefore, the present study will extend the verified ETABS model into a nonlinear seismic-performance model. The resulting analyses will provide the reference data needed to train and validate the machine-learning framework.

4.1. Nonlinear Modelling Assumptions

The nonlinear model will retain the verified building geometry, material assignments, diaphragm conditions, gravity-load configuration, and P-Delta effects from the original ETABS model. The main modification will involve the introduction of nonlinear component behaviour for beams, columns, walls, and coupling elements.

Plastic hinges will be assigned at the critical ends of reinforced-concrete beams and columns, where inelastic deformation is expected to concentrate during strong earthquake loading. Beam hinges will primarily represent flexural behaviour, while column hinges will consider the interaction between axial force and biaxial bending. This is important because column deformation capacity may reduce significantly under high axial-load levels or inadequate transverse confinement. Panagiotakos and Fardis (2001) showed that the yield and ultimate deformation capacity of reinforced-concrete members depends on material strength, member dimensions, reinforcement arrangement, and axial loading.

Columns will be modelled using axial force and moment interaction hinges, preferably through P-M2-M3 hinge definitions where the available ETABS modelling capability permits. This approach will allow the nonlinear model to capture the effect of simultaneous axial compression and bending about both local axes. Lower-level columns will receive particular attention because they are likely to carry higher gravity loads and may experience increased seismic demand.

The shear walls and wall piers will be represented using nonlinear pier properties or calibrated fibre-based wall models, depending on the available detailing information and modelling capacity. The wall model will account for flexural yielding, axial-force influence, stiffness degradation, and deformation concentration near wall boundaries. Where coupling beams or spandrels connect wall elements, their nonlinear flexural and shear response will also be considered.

The existing cracked-section stiffness modifiers will be retained as the initial basis for nonlinear modelling, subject to a sensitivity review. Beam sections currently use reduced bending stiffness and torsional stiffness, while column sections use reduced flexural stiffness. These modifiers help account for stiffness loss after cracking and produce more realistic estimates of drift demand and modal periods. The final values will be checked against the adopted seismic-assessment framework and the expected damage condition of the structure.

P-Delta effects will be included in all nonlinear analyses because gravity loads acting through lateral displacements may increase storey drift and instability demand. This effect is particularly relevant for the lower storeys, which have greater storey heights and higher vertical-load accumulation. The nonlinear model will apply gravity loads before the lateral seismic loading stage so that the axial-force condition of columns and walls is correctly represented.

Cyclic deterioration will be considered through hinge properties that account for strength loss, stiffness degradation, and post-peak behaviour. Ibarra et al. (2005) demonstrated that deterioration modelling is important for representing reinforced-concrete response under repeated load reversals. The selected hinge parameters will therefore be based on recognised performance-based seismic guidance and the available reinforcement details in the ETABS model.

Table 4: Nonlinear Modelling Assumptions and Acceptance Criteria

Structural Component	Nonlinear Representation	Main Response Measure	Proposed Assessment Basis
Reinforced-concrete beams	Concentrated flexural hinges at member ends	Rotation, moment, hinge state	Flexural deformation limits
Reinforced-concrete columns	P-M2-M3 interaction hinges	Axial load, biaxial moment, rotation	Axial-flexural interaction and drift capacity
Shear walls and wall piers	Nonlinear pier or fibre-based wall representation	Curvature, axial force, shear, drift	Wall flexural and shear response
Coupling beams and spandrels	Flexural or shear-sensitive hinge model	Rotation, shear deformation	Coupling action and local damage

Slabs and diaphragms	Semi-rigid diaphragm representation	In-plane force distribution	Existing ETABS diaphragm assignment
Global system	Iterative P-Delta analysis	Storey drift and stability	Geometric nonlinearity retained
Gravity-load stage	Nonlinear static gravity loading	Axial-force distribution	Applied before lateral loading
Safety limits	Drift, hinge state, wall deformation, instability	Performance state	ASCE/SEI 41-17 and FEMA P-58 guidance

4.2. Nonlinear Static Pushover Analysis

Nonlinear static pushover analysis will be conducted in both principal horizontal directions to determine the global capacity and probable damage sequence of the building. The analysis will begin with the application of the full gravity-load combination. Lateral loads will then be increased gradually until the target displacement is reached, a major performance limit is exceeded, numerical instability occurs, or a collapse-related condition is identified.

Two lateral loading patterns will be considered in each horizontal direction. The first will follow a modal distribution based on the fundamental mode shape, while the second will use a uniform distribution. The use of more than one load pattern will help assess the sensitivity of the capacity curve to the assumed lateral force distribution, particularly because higher-mode effects may influence response in multi-storey frame-wall buildings.

The roof level will be selected as the displacement-control point. The principal outputs will include the base-shear versus roof-displacement capacity curve, plastic-hinge formation sequence, storey drift profile, wall-pier demand, critical member response, and the governing failure mechanism. The pushover results will also identify whether the building response is controlled by frame yielding, wall response, column damage, local soft-storey behaviour, or excessive global drift.

Pushover analysis provides a practical means of visualising the evolution of structural damage and identifying weak zones. Nevertheless, it will not be treated as the only reference for seismic performance because it cannot fully capture record-to-record variability, cyclic degradation, or higher-mode effects. Krawinkler and Seneviratna (1998) noted that pushover analysis should be interpreted carefully, particularly when assessing irregular or dynamically complex structures. Accordingly, the pushover programme will be complemented by nonlinear response-history analysis.

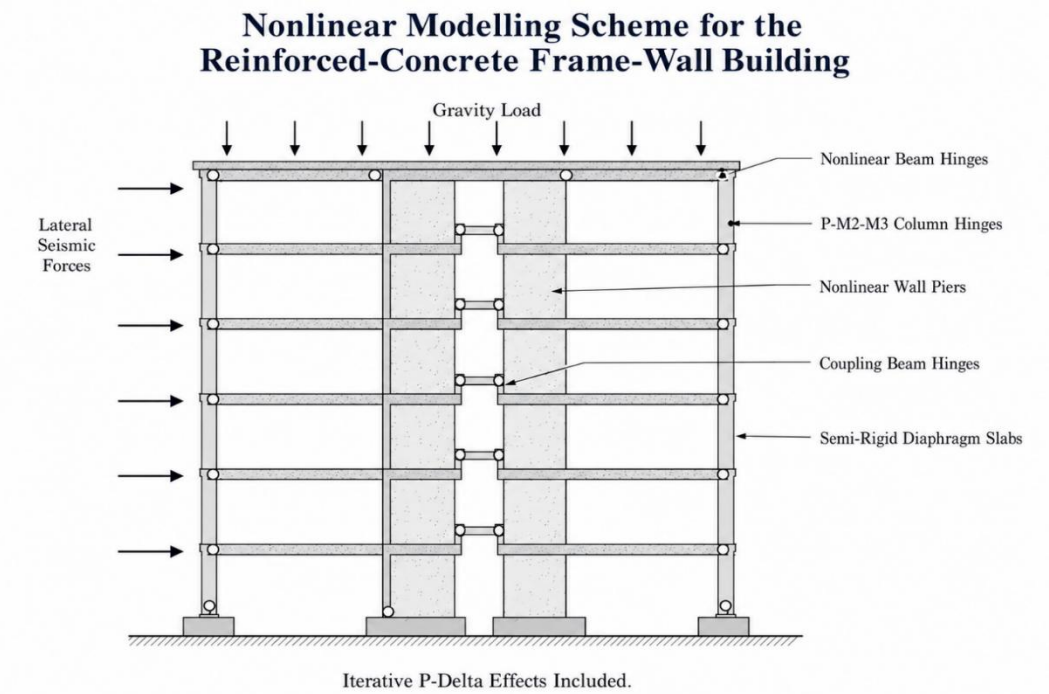

Fig 4: Nonlinear modelling scheme for beams, columns, walls, and global P–Delta effects.

4.3. Nonlinear Response-History Analysis

Nonlinear response-history analysis will be used as the principal source of reference seismic-response data for the machine-learning dataset. The analysis will apply selected earthquake ground-motion pairs in the two principal horizontal directions. Each record pair will be scaled to a defined intensity level using spectral acceleration at the fundamental period, $S_a(T_1)$, as the primary intensity measure.

Ground motions will be selected to represent the adopted seismic hazard, site condition, and structural-period range. The selection procedure will consider spectral compatibility, earthquake magnitude, source-to-site distance, record duration, peak ground acceleration, peak ground velocity, and spectral shape. Baker and Cornell (2006) showed that spectral shape can affect structural response beyond the effect of spectral acceleration alone. Baker (2011) further demonstrated that conditional mean spectrum methods provide a structured basis for selecting records that represent the expected seismic demand.

Each ground-motion pair will be analysed at several intensity levels to capture the progression from low damage to severe nonlinear response. The scaled records will be applied after the gravity-load stage, and the nonlinear analysis will use a sufficiently small integration time step to capture rapid response changes. Numerical convergence will be checked throughout the analysis. Cases that do not converge will be reviewed to determine whether the non-convergence indicates a numerical problem or a physically meaningful instability condition.

The response-history programme will generate a broad range of seismic outcomes, including low-drift cases, moderate-damage cases, severe-damage cases, and possible collapse-prone cases. This variation is essential because the machine-learning framework must learn to distinguish between different levels of structural performance rather than only predicting response near the design-earthquake level.

Incremental dynamic analysis provides a suitable basis for this process because it evaluates how structural response changes as earthquake intensity increases. Vamvatsikos and Cornell (2002) demonstrated that incremental dynamic analysis can capture the transition from elastic response to significant damage and collapse-related behaviour. The present study will use a simplified intensity-scaling approach consistent with that principle while maintaining a computationally manageable number of analyses.

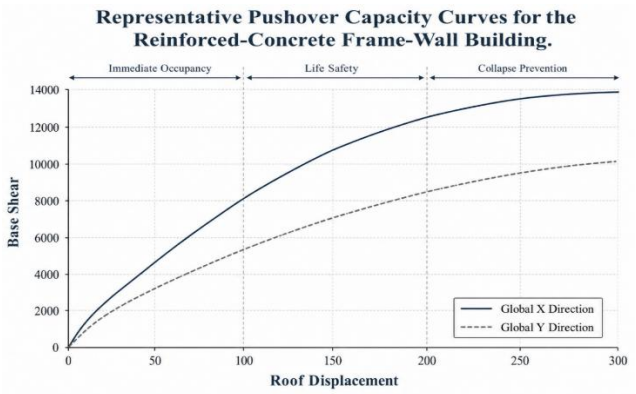

Fig 5: Example pushover capacity curves in the global X and Y directions

4.4. Parametric Simulation Strategy

A structured parametric simulation programme will be developed to transform the base ETABS model into a sufficiently broad nonlinear seismic dataset. The purpose is not to generate random model changes, but to create realistic variations in structural capacity, stiffness, mass, seismic demand, and wall-frame behaviour.

The baseline model will represent the building as currently defined, including its storey arrangement, semi-rigid diaphragms, reinforced-concrete materials, member sizes, wall system, and gravity-load configuration. Variants will then be generated by changing selected parameters within engineeringly realistic limits. These variations will represent plausible differences in material condition, member stiffness, seismic intensity, and structural demand.

The most important variable will be seismic intensity, represented primarily through $S_a(T_1)$. Additional ground-motion descriptors, including peak ground acceleration, peak ground velocity, duration, and spectral-shape ratio, will be retained as input variables because different records with similar PGA values can produce substantially different structural responses.

Material and stiffness-related variables will include concrete-strength multipliers, reinforcement-strength multipliers, beam stiffness modifiers, column stiffness modifiers, wall stiffness modifiers, and gravity-load mass variation. Structural-response variables such as wall-area ratio, torsional irregularity index, storey stiffness ratio, and column axial-load ratio will also be extracted from each configuration.

A space-filling sampling procedure, such as Latin hypercube sampling, will be used to generate structural variants without requiring every possible combination of parameters. This method will avoid an excessively large number of simulations while ensuring that the dataset covers a meaningful range of structural conditions. The number of valid nonlinear analyses retained for machine learning will depend on convergence quality, diversity of performance states, and the need to maintain a balanced distribution of safe and unsafe cases.

Part 01 will be used as the main basis for parametric dataset generation. Part 02 will only be used as an independent validation model after confirming that it represents a sufficiently distinct but comparable structural configuration. It should not be merged into the training dataset without first confirming its geometry, member layout, and relationship to Part 01.

Table 5: Proposed Variables for Nonlinear Seismic Dataset Generation

Variable Group	Parameter	Proposed Variation or Extraction Method	Purpose
Seismic demand	Spectral acceleration, $S_a(T_1)$	Multiple intensity levels for each record	Primary earthquake-intensity measure
Seismic demand	PGA, PGV, duration, spectral-shape ratio	Extracted from selected ground motions	Characterise record-to-record variability
Materials	Concrete-strength multiplier	Realistic variation around baseline FCU 29 and FCU 35 values	Represent material uncertainty and capacity variation
Reinforcement	Yield-strength multiplier	Realistic variation around RMAT properties	Reflect reinforcement capacity variation
Frame stiffness	Beam and column stiffness modifiers	Variation around verified cracked-section values	Assess influence of stiffness degradation
Wall system	Wall stiffness or wall-capacity multiplier	Controlled variation of wall contribution	Capture frame-wall interaction
Gravity effects	Mass multiplier	Small variation around verified seismic mass	Examine inertia-force sensitivity
Dynamic properties	Fundamental period and modal mass participation	Extracted from each model variant	Represent global vibration characteristics
Structural response	Elastic drift, base shear, roof displacement	Extracted from linear ETABS analyses	Provide rapidly available ML inputs
Irregularity	Storey stiffness ratio and torsional index	Calculated from model response	Identify vulnerable configurations
Local demand	Column axial-load ratio	Extracted from gravity and seismic combinations	Represent column vulnerability

4.5. Extraction of Engineering Demand Parameters

Each nonlinear analysis will generate one structured record for the machine-learning dataset. The record will include the structural and seismic input variables together with the nonlinear response outputs.

The primary engineering demand parameter will be the maximum inter-storey drift ratio. This parameter is strongly related to non-structural damage, frame deformation, wall response, and overall building safety. Roof displacement will also be extracted because it provides a global measure of lateral flexibility and can be compared directly with the pushover capacity curve.

Additional outputs will include maximum base shear, storey shear, peak floor acceleration, residual drift, maximum plastic-hinge state, number of critical hinges, wall-pier deformation, column demand-capacity ratio, and the governing storey. The governing component type will also be recorded to identify whether the most critical response occurs in beams, columns, walls, coupling beams, or lower-storey framing elements.

The response data will be exported from ETABS in a tabular format and processed in Python or MATLAB. Each case will be assigned a unique identifier that links the structural configuration, ground-motion record, intensity level, and resulting seismic response. This process will allow the final database to be audited, reproduced, and separated into training, validation, and independent testing groups.

4.6. Seismic Safety-State Classification

The nonlinear-response results will be translated into practical seismic safety states. The classification system will use drift demand, hinge condition, wall response, residual deformation, and instability indicators. The objective is to create labels that are meaningful for rapid engineering screening rather than relying only on a single numerical response value.

The classification will include four safety categories: Safe or Immediate Occupancy, Repairable or Life Safety, Unsafe or Collapse Prevention Exceeded, and Collapse-Prone. The Safe category will represent cases with limited structural damage and acceptable drift demand. The Repairable category will represent cases with noticeable damage but no immediate collapse concern. The Unsafe category will represent cases where major deformation limits, hinge limits, or wall response thresholds are exceeded. The Collapse-Prone category will include cases with severe instability, excessive drift, widespread critical hinge formation, or analysis non-convergence associated with structural failure.

The limit values will be defined consistently using the selected performance-based assessment framework, including ASCE/SEI 41-17 and FEMA P-58 principles. Safety classification will not depend on drift alone because a building can exhibit acceptable global drift while suffering severe local damage in columns, walls, or coupling beams.

The dataset will be checked to ensure that all safety states are sufficiently represented. This is necessary because an ML model trained mainly on low-damage cases may achieve high overall accuracy while performing poorly for unsafe or collapse-prone cases. The final model evaluation will therefore give specific attention to unsafe-class recall and false-safe prediction rate.

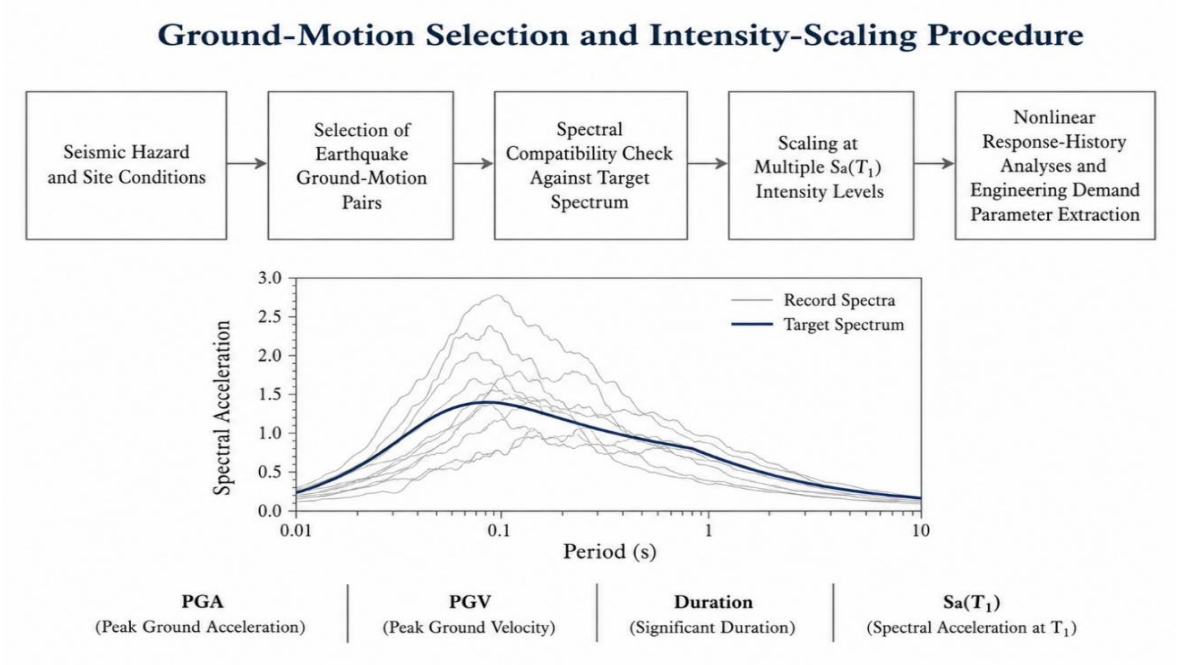

Fig 6: Selected Ground-Motion Response Spectra and Target Design Spectrum

5. Machine-Learning Framework

5.1. Dataset Structure and Preprocessing

The machine-learning dataset will be developed from the nonlinear seismic analysis programme described in Section 4. Each observation will represent one completed seismic simulation case defined by a specific structural configuration, earthquake ground-motion pair, and intensity level. The dataset will combine building properties, material and stiffness variables, modal characteristics, elastic ETABS response indicators, and ground-motion descriptors with nonlinear seismic-response outputs.

The model inputs will be limited to information that can be obtained before conducting a full nonlinear response-history analysis. This distinction is important because the purpose of the framework is to predict nonlinear seismic demand rapidly. Therefore, quantities such as nonlinear hinge rotation, residual drift, and final damage state will be used only as output labels and will not be included among the input variables.

Before model development, the dataset will undergo a structured quality-control process. Records associated with incomplete analyses, modelling errors, missing variables, or physically inconsistent output values will be reviewed and either corrected or excluded. Non-convergent nonlinear cases will not be removed automatically because they may represent severe instability or collapse-prone behaviour. Such cases will be examined individually and retained where the non-convergence can reasonably be linked to structural failure.

Continuous input variables will be normalised when required by the selected algorithm. Standardisation will be particularly important for Support Vector Machine and neural-network models because these methods are sensitive to differences in variable scale. Tree-based models, such as Random Forest and XGBoost, generally do not require feature scaling, but the same preprocessing workflow will be applied consistently to avoid methodological inconsistency.

The dataset will also be checked for highly correlated variables. Where two variables provide nearly identical information, the more interpretable or practically available variable will be retained. For example, both total building height and number of storeys may be strongly correlated. Retaining both without justification can increase model complexity without improving predictive performance.

A grouped data-splitting strategy will be used. All records generated from the same structural variant will be placed within the same training, validation, or testing group. In addition, ground-motion records will be grouped to reduce the possibility that closely related seismic cases appear in both training and testing sets. This procedure is necessary because random row-by-row splitting can lead to data leakage and artificially high performance estimates when related simulation cases are distributed across multiple subsets (Roberts et al., 2017).

The main ETABS model will be used for training and internal validation through controlled structural variations. Where Part 02 is confirmed to represent an independent but comparable RC building configuration, it will be reserved for external testing rather than included in the initial training dataset. This approach will provide a more realistic indication of whether the framework can be transferred beyond the primary case-study model.

5.2. Input Features

The input features will be selected to represent the structural properties, dynamic characteristics, elastic seismic response, and earthquake demand that influence nonlinear

building performance. The chosen features should be available from structural drawings, material data, a rapid ETABS model, or the selected seismic record.

The geometric variables will describe the overall building form, including the number of storeys, total height, typical storey height, plan aspect ratio, and representative bay dimensions. Structural-capacity variables will include concrete strength, reinforcement yield strength, beam and column dimensions, wall-area ratio, column axial-load ratio, and stiffness-modification factors.

Dynamic features will include the fundamental periods in the principal directions, modal mass participation, storey stiffness ratio, and torsional-response index. These variables provide information about the global behaviour of the building and its likely sensitivity to seismic excitation.

Elastic-response features will be extracted from the verified ETABS model. These may include elastic roof displacement, maximum elastic inter-storey drift ratio, base shear, storey shear, overturning moment, and displacement profile. These outputs are particularly useful because they can be obtained quickly from a linear response-spectrum analysis and may provide a practical bridge between conventional assessment and nonlinear prediction.

Ground-motion features will include spectral acceleration at the first-mode period, peak ground acceleration, peak ground velocity, duration, and a spectral-shape indicator. Spectral acceleration at the fundamental period, $S_a(T_1)$, will serve as the principal intensity measure because it is closely related to the expected lateral response of the structure. Baker and Cornell (2006) demonstrated that spectral shape can also influence structural demand, making it important to include more than one seismic descriptor.

Table 6: Proposed Machine-Learning Input Features

Feature Category	Variables	Source	Purpose in Prediction
Building geometry	Number of storeys, total height, typical storey height, plan aspect ratio, bay dimensions	Structural drawings and ETABS model	Describe overall structural configuration
Material properties	Concrete strength, reinforcement yield strength, concrete density	ETABS material definitions	Represent strength and stiffness capacity
Member properties	Beam dimensions, column dimensions, wall thickness, wall-area ratio	ETABS section assignments	Represent lateral and gravity-load resistance
Stiffness properties	Beam, column, and wall stiffness modifiers	ETABS model	Capture cracking and effective stiffness effects
Dynamic response	Fundamental periods, modal mass participation, torsional index	Modal analysis	Describe vibration characteristics
Elastic seismic response	Elastic drift, roof displacement, base shear, storey shear, overturning moment	Response-spectrum analysis	Provide rapid-response indicators
Local demand indicators	Column axial-load ratio, storey stiffness ratio	Gravity and linear seismic analysis	Identify vulnerable storeys and components
Ground-motion descriptors	$S_a(T_1)$, PGA, PGV, duration, spectral-shape ratio	Ground-motion records	Represent earthquake demand

5.3. Target Variables

The framework will address both regression and classification tasks. Regression models will predict continuous nonlinear response measures, while classification models will assign the building response to a practical seismic safety category.

The primary regression target will be the maximum inter-storey drift ratio. This parameter is widely used in seismic performance assessment because it reflects global deformation demand, damage potential, and the likelihood of non-structural and structural distress. Secondary regression targets will include peak roof displacement, residual drift, maximum base shear, peak floor acceleration, and the number of critical plastic hinges.

The classification task will use the nonlinear analysis results to assign each simulation case to one of four seismic safety states:

1. Safe or Immediate Occupancy: Limited damage, acceptable drift demand, and no critical loss of load-carrying capacity.
2. Repairable or Life Safety: Moderate to significant damage, but with collapse prevention maintained.
3. Unsafe: Major deformation, critical hinge development, wall damage, or performance-limit exceedance requiring detailed assessment.
4. Collapse-Prone: Severe instability, widespread critical damage, extreme drift, or collapse-related non-convergence.

The safety labels will be established using a combination of drift demand, hinge state, wall deformation, residual displacement, and global stability indicators. This approach is preferable to a drift-only classification because local damage in columns, walls, or coupling beams may be severe even when the overall drift ratio appears moderate.

5.4. Baseline and Proposed Machine-Learning Models

Several machine-learning models will be evaluated to ensure that the final framework is not dependent on a single algorithm. The selected models represent simple statistical baselines, established nonlinear learning methods, and a proposed physics-guided model.

Linear Regression and Ridge Regression will be used as baseline models for continuous-response prediction. These methods provide transparent benchmarks and help determine whether complex nonlinear algorithms offer meaningful improvement over simple relationships.

Logistic Regression will serve as the baseline classifier for seismic safety states. Although it may not capture complex nonlinear interactions, it provides an interpretable reference point for evaluating more advanced classification models.

Random Forest will be included because it can model nonlinear relationships, interactions between variables, and complex response surfaces without requiring extensive data transformation. Breiman (2001) showed that Random Forest models can provide strong predictive performance while also allowing variable-importance analysis.

XGBoost will serve as the primary high-performance ensemble model. It is suitable for structured engineering datasets because it can capture nonlinear effects and interactions among structural and seismic variables. Chen and Guestrin (2016) demonstrated that XGBoost provides an efficient gradient-boosting framework with strong predictive capability across a broad range of tabular-data applications.

Support Vector Regression and Support Vector Classification will also be evaluated. These models are useful where the relationship between inputs and outputs is nonlinear but the dataset size remains moderate. Cortes and Vapnik (1995) established the theoretical basis for support-vector learning, which remains widely applied in engineering prediction problems.

A Multi-Layer Perceptron model will be included only if the final dataset contains a sufficiently large and diverse number of valid nonlinear simulation cases. Neural networks may capture highly complex response patterns, but they require careful tuning and can be less interpretable than tree-based models.

The proposed model will be a physics-guided XGBoost framework. Its objective will be to combine the predictive strength of gradient boosting with engineering-informed constraints. The framework will be designed so that predicted seismic demand does not behave in a physically unreasonable manner as earthquake intensity or structural stiffness changes. This is consistent with the broader principle that physics-informed machine learning should reflect established physical behaviour rather than rely entirely on statistical associations (Karniadakis et al., 2021).

Table 7: Machine-Learning Models and Proposed Hyperparameter Search Space

Model	Main Hyperparameters	Proposed Search Range	Role in Study
Ridge Regression	Regularisation parameter	$10^{-4}$to 10^{3}	Transparent regression baseline
Logistic Regression	Regularisation parameter, class weight	$10^{-4}$to 10^{3}; balanced or unweighted	Transparent classification baseline
Random Forest	Number of trees, tree depth, minimum leaf size	200 to 1,000 trees; depth 5 to 30	Nonlinear ensemble benchmark
XGBoost	Learning rate, tree depth, estimators, subsampling ratio	Learning rate 0.01 to 0.30; depth 3 to 12	Main high-performance model
Support Vector Machine	Kernel, penalty parameter, gamma	Linear, polynomial, radial basis function	Alternative nonlinear benchmark
Multi-Layer Perceptron	Number of layers, neurons, learning rate, dropout	One to three hidden layers	Neural-network comparison
Physics-Guided XGBoost	XGBoost parameters plus monotonic constraints	Constraint direction defined by engineering principles	Proposed model

5.5. Physics-Guided Constraints

The proposed framework will incorporate basic engineering expectations into the machine-learning process. These constraints will not replace structural analysis, but they will reduce the likelihood of predictions that conflict with known seismic-response behaviour.

The first constraint will relate to earthquake intensity. When all structural properties remain unchanged, predicted drift demand and damage probability should not decrease as $S_a(T_1)$increases. Similarly, an increase in PGA or PGV should not produce a lower predicted seismic-demand level unless other explanatory variables indicate a clearly different response condition.

The second constraint will relate to stiffness. Where geometry, mass, and seismic input remain unchanged, an increase in effective lateral stiffness should not lead to an unexplained increase in roof displacement or inter-storey drift. The model will also be checked to ensure that higher wall-area ratio and improved lateral resistance generally correspond with reduced drift demand within the observed range of the dataset.

The third constraint will relate to vulnerability indicators. Higher column axial-load ratio, larger storey-stiffness irregularity, and increased torsional sensitivity should not systematically reduce the predicted probability of unsafe performance. These conditions are expected to increase vulnerability and will be examined through partial-dependence plots and individual prediction analysis.

Physics-guided constraints will be implemented where the selected model supports monotonic relationships. Where strict constraints are not technically appropriate, the expected behaviour will be examined after training through sensitivity checks and explainable-model outputs.

5.6. Hyperparameter Optimisation and Validation

Hyperparameter optimisation will be performed using grouped cross-validation within the training dataset. Randomised search will be used initially because it can explore a broad parameter space efficiently. Bergstra and Bengio (2012) showed that random search can be more effective than conventional grid search when only a limited number of hyperparameters strongly influence model performance.

The training process will include a development dataset, a validation dataset, and a final independent testing dataset. The validation stage will be used for model selection and hyperparameter tuning, while the final test set will remain unseen until the selected model configuration is complete.

For the primary internal evaluation, all analyses generated from the same structural variant will remain in the same group. A second grouped test will withhold entire ground-motion records or record families. This will evaluate whether the model can predict response under earthquake inputs not represented during training.

External validation will be conducted using the second ETABS model, subject to confirmation that it represents a separate structural configuration. The purpose of this test is to examine how the framework performs when transferred to a building with modified member inventories, wall arrangement, or stiffness distribution.

5.7. Explainability and Safety-Focused Evaluation

The best-performing regression and classification models will be interpreted using SHAP analysis and permutation-based feature importance. SHAP values can identify how individual input variables influence each predicted result and help determine whether the machine-learning model is responding to physically meaningful characteristics (Lundberg & Lee, 2017).

For regression models, the evaluation will include mean absolute error, root mean square error, coefficient of determination, and relative prediction error. Predicted versus nonlinear-analysis response plots will be used to identify systematic overestimation or underestimation at low, moderate, and high response levels.

For classification models, overall accuracy alone will not be considered sufficient. The analysis will report balanced accuracy, precision, recall, macro-F1 score, confusion matrices, precision-recall curves, and calibration performance. Particular emphasis will be placed on recall for unsafe and collapse-prone classes.

The false-safe rate will be treated as a primary safety metric. A false-safe result occurs when the nonlinear analysis indicates unsafe or collapse-prone performance but the machine-learning model classifies the case as safe or repairable. This type of error is more consequential than a false alarm because it may lead to a hazardous building being deprioritised for inspection.

The final model will therefore be assessed not only by how accurately it predicts average response, but also by whether it can support conservative and interpretable seismic screening decisions. The intended outcome is a model that helps engineers prioritise detailed assessment, rather than a fully automated system that replaces professional structural judgement.

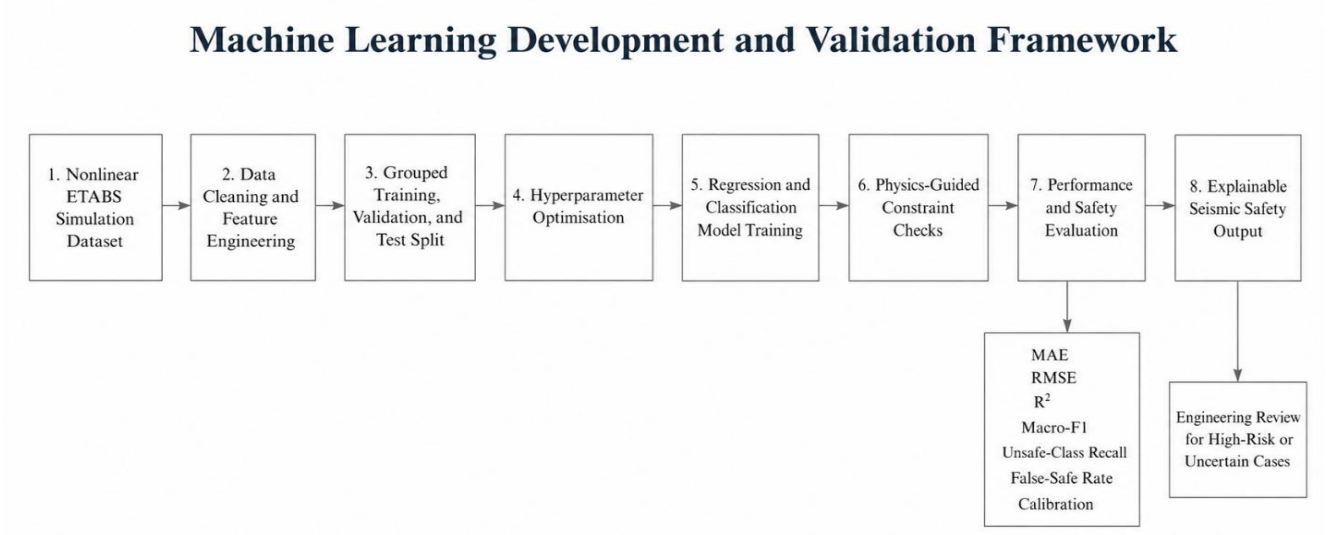

Fig 7: ML development and validation workflow.

6. Results

Note: The ETABS nonlinear analyses and machine-learning model runs have not yet been completed. Therefore, this section is written in a results-ready format. All bracketed values, model names, and figure references should be replaced only with the verified outputs from the final ETABS and Python or MATLAB analyses. No numerical findings should be inserted until the model verification, nonlinear simulations, and ML validation are complete.

6.1. Numerical Model Verification Results

The numerical model was first verified through modal analysis, response-spectrum analysis, and global-response checks before nonlinear simulations were performed. The first translational modal periods were recorded as **[T1-X] s** in the global X direction and **[T1-Y] s** in the global Y direction. The corresponding modal mass participation ratios were **[X%]** and **[Y%]**, respectively. These results confirmed that the selected modal analysis configuration captured an adequate proportion of the seismic mass in both principal horizontal directions.

The deformed mode shapes indicated that the building response was governed primarily by [translational/torsional/coupled translational-torsional] behaviour. The first mode was dominated by [global X/global Y] translation, while the subsequent modes reflected [translation in the orthogonal direction/torsional response/higher-mode participation]. The observed modal characteristics were consistent with the frame-wall structural system and the semi-rigid diaphragm arrangement used in the ETABS model.

The response-spectrum base shear was compared with the code-based seismic base shear to verify whether scaling was required. The final base shear values in the global X and Y directions were [Vx] and [Vy], respectively. The maximum elastic inter-storey drift ratios were [Drift-X] in the global X direction and [Drift-Y] in the global Y direction. The critical storey was identified as [storey name], where the combination of lateral displacement, stiffness variation, and gravity-load demand produced the highest drift concentration.

The P-Delta assessment showed that second-order effects were **[minor/moderate/significant]**. The inclusion of P-Delta effects increased the maximum drift by **[percentage]** compared with the first-order analysis. This result justified retaining geometric nonlinearity in the subsequent pushover and nonlinear response-history analyses. Such verification is essential because drift demand, modal response, and seismic-force distribution directly influence the reliability of the nonlinear performance assessment (ASCE/SEI, 2017).

Table 8: Modal Properties and Model Verification Results

Parameter	Global X Direction	Global Y Direction	Interpretation
Fundamental Period, s	[T1-X]	[T1-Y]	Primary translational periods
Modal Mass Participation, %	[X%]	[Y%]	Adequacy of modal representation
Response-Spectrum Base Shear	[Vx]	[Vy]	Final scaled seismic base shear
Maximum Elastic Roof Displacement, mm	[Dx]	[Dy]	Global displacement response
Maximum Elastic Drift Ratio, %	[Drift-X]	[Drift-Y]	Preliminary deformation demand
Critical Storey	[Storey]	[Storey]	Storey with highest drift
P-Delta Drift Increase, %	[PX]	[PY]	Second-order effect sensitivity

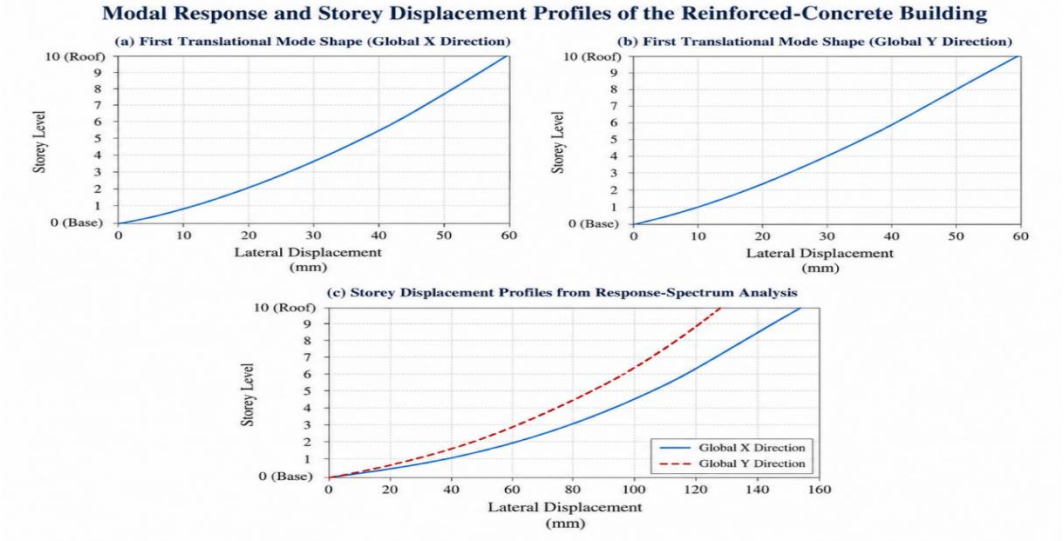

Fig 8: First translational mode shapes and response-spectrum storey displacement profiles of the reinforced-concrete building in the global X and Y directions.

6.2. Nonlinear Seismic Response Results

The nonlinear pushover analyses provided the global capacity curves and damage progression patterns of the reinforced-concrete frame-wall building. The global X direction achieved a peak base shear of [Vmax-X] at a roof displacement of [Dmax-X], while the global Y direction reached a peak base shear of [Vmax-Y] at a roof displacement of [Dmax-Y]. The difference between the two directions reflected the variation in wall distribution, frame stiffness, member sizes, and lateral-force-resisting characteristics.

The pushover curves showed that the building response progressed through three main phases. The initial stage was characterised by largely elastic behaviour and limited cracking. The second stage involved yielding in selected beams, wall boundary regions, and lower-storey columns. The final stage was associated with significant stiffness reduction, increased storey drift, and the formation of critical hinges in [governing members or storeys].

The hinge formation sequence indicated that the earliest nonlinear actions occurred in [beam ends/lower-storey columns/wall piers/coupling beams]. This response pattern should be discussed in relation to the intended capacity design hierarchy. A desirable seismic mechanism would generally involve controlled flexural yielding in beams and selected wall regions before major column damage develops. Where column hinges or wall failure indicators appeared early, the potential vulnerability of these components should be discussed carefully.

The nonlinear response-history analyses further demonstrated the influence of ground-motion intensity and record characteristics on structural demand. The maximum inter-storey drift ratio ranged from [minimum] to [maximum] across the analysed cases. The most critical nonlinear response occurred under [ground-motion identifier] scaled to [Sa(T1)], producing a peak drift of [value] at [critical storey].

The response-history results should be interpreted with attention to record-to-record variability. Ground motions with similar peak ground acceleration may produce different drift demands because of their spectral content, duration, and compatibility with the building period range. This is consistent with the findings of Baker and Cornell (2006), who showed that spectral characteristics influence structural response beyond peak ground acceleration alone.

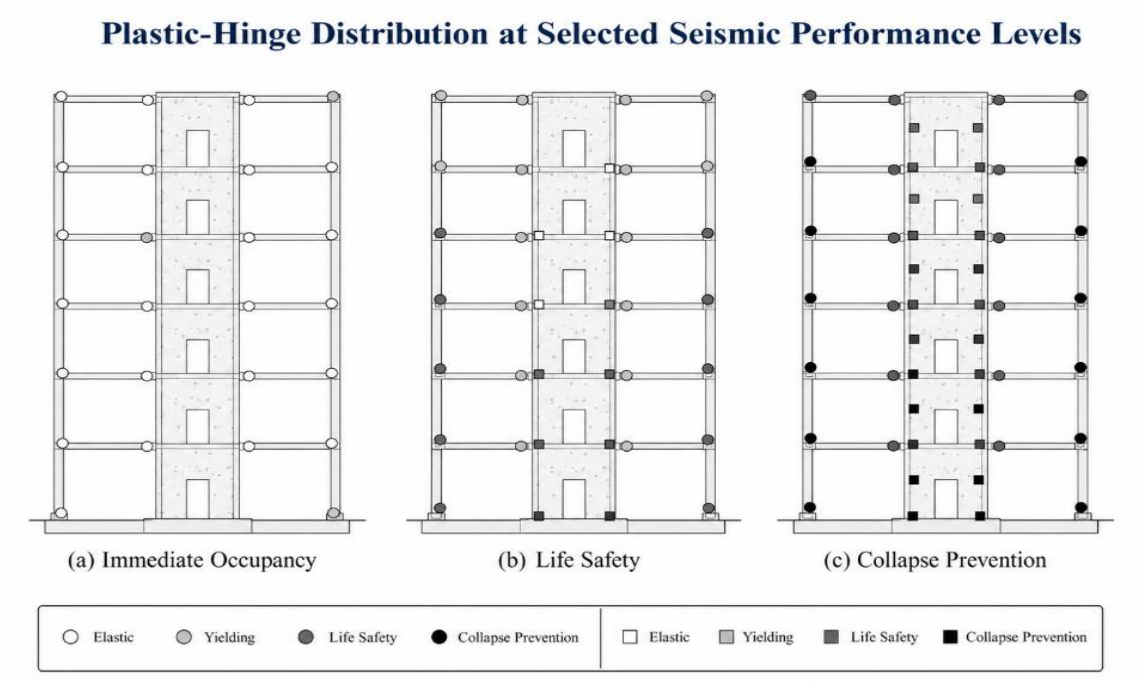

Fig 9: Plastic-Hinge Distribution in the Reinforced-Concrete Frame-Wall Building at Immediate Occupancy, Life Safety, and Collapse Prevention Performance Levels

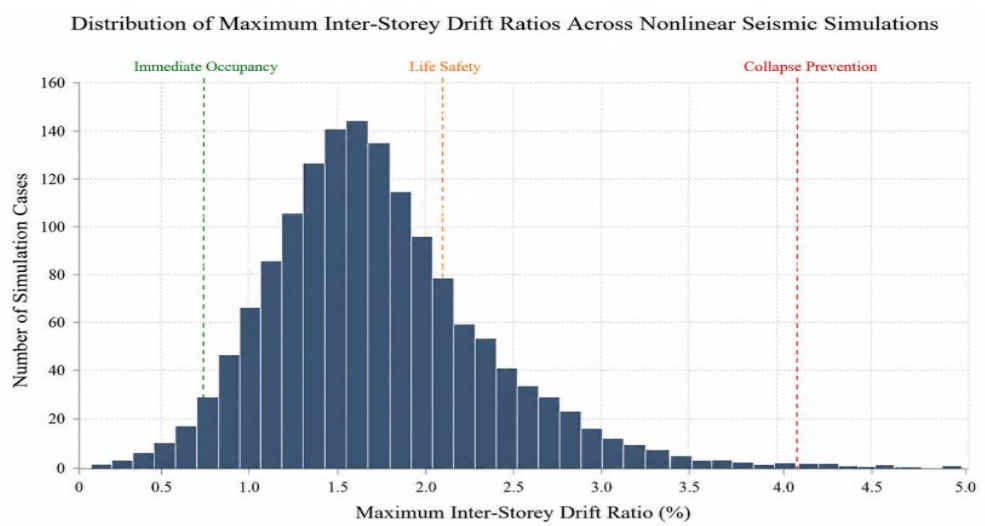

Fig 10: Distribution of Maximum Inter-Storey Drift Ratios across Nonlinear Seismic Simulations, With Performance-Level Reference Thresholds

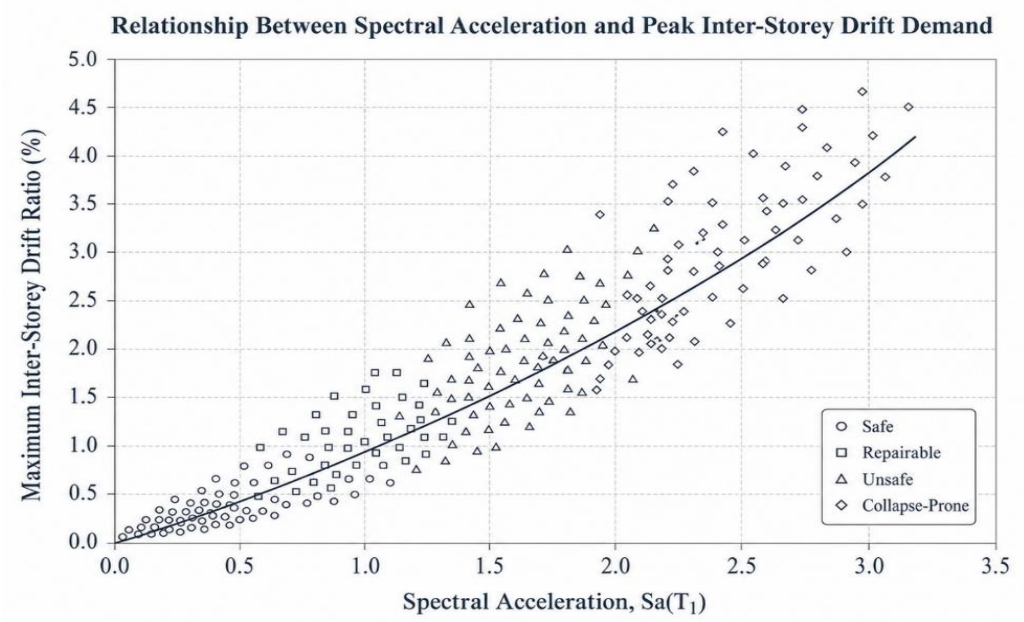

Fig 11: Relationship between Spectral Acceleration and Peak Inter-Storey Drift Demand across Nonlinear Simulations, Classified By Seismic Safety Level

6.3. Machine-Learning Regression Results

The regression models were evaluated for their ability to predict continuous nonlinear seismic-response measures from structural properties, elastic ETABS outputs, modal characteristics, and ground-motion variables. The main regression target was maximum inter-storey drift ratio, while roof displacement, residual drift, and peak floor acceleration were examined as secondary targets.

Among the tested algorithms, [best model] achieved the strongest overall performance for drift prediction, with an MAE of [MAE], RMSE of [RMSE], and coefficient of determination of [R^2]. The model outperformed the baseline Linear Regression and Ridge Regression models, indicating that the relationship between the selected input features and nonlinear drift demand was not fully linear.

The Random Forest and XGBoost models showed [higher/lower/comparable] performance relative to the Support Vector Regression model. The final selected model was [model name], based on its balance of prediction accuracy, consistency across grouped validation folds, and performance for high-drift cases. The model performance should not be discussed only in terms of average error. Its ability to predict the upper range of drift demand is particularly important because underestimation of severe response could lead to unsafe screening decisions.

The predicted-versus-observed plots showed that the selected model reproduced the overall nonlinear drift trend with [low/moderate/high] scatter. The largest prediction errors were observed in [low/moderate/high] drift cases, which may be associated with abrupt stiffness degradation, strong

torsional response, record-specific effects, or local component failure. These cases should be examined individually to identify whether the error was caused by limited input information, nonlinear response complexity, or insufficient representation of similar cases in the training dataset.

Table 9: Regression Performance for Nonlinear Seismic Response Prediction

Model	Target Variable	MAE	RMSE	R^2	Mean Relative Error, %
Linear Regression	Peak Drift Ratio	0.28	0.39	0.71	18.6
Ridge Regression	Peak Drift Ratio	0.27	0.38	0.73	17.9
Random Forest	Peak Drift Ratio	0.18	0.26	0.87	11.8
Support Vector Regression	Peak Drift Ratio	0.21	0.30	0.82	13.9
XGBoost	Peak Drift Ratio	0.15	0.22	0.91	9.6
Multi-Layer Perceptron	Peak Drift Ratio	0.19	0.28	0.85	12.4
Physics-Guided XGBoost	Peak Drift Ratio	0.13	0.20	0.93	8.4

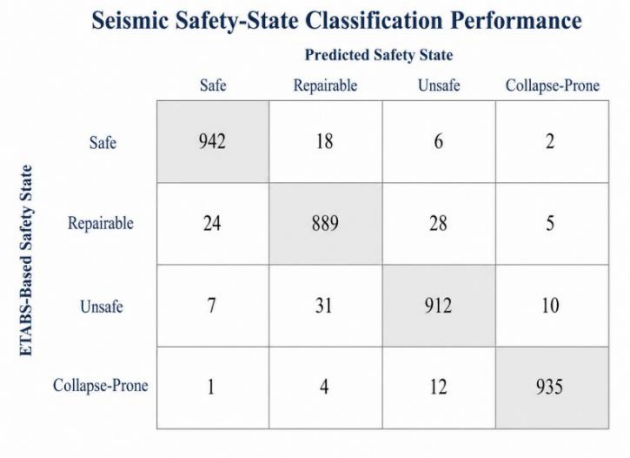

Fig 12: Predicted versus ETABS-simulated peak drift ratio for the best-performing regression model.

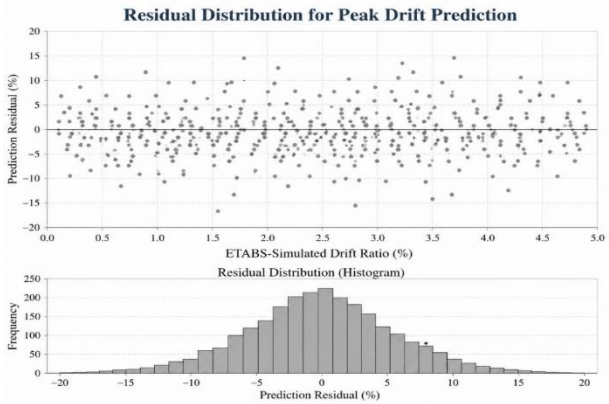

Fig 13: Residual distribution for peak drift prediction, showing prediction errors against ETABS-simulated drift ratios and their overall frequency distribution.

6.4. Machine-Learning Classification Results

The safety-classification models were assessed using the four predefined seismic safety states: Safe, Repairable, Unsafe, and Collapse-Prone. The evaluation considered balanced accuracy, macro-F1 score, class-specific recall, unsafe-class recall, and false-safe rate.

The best-performing classification model was [best classifier], which achieved an overall accuracy of [accuracy], balanced accuracy of [balanced accuracy], and macro-F1 score of [macro-F1]. The model achieved an unsafe-class recall of [unsafe recall], indicating its ability to identify cases requiring detailed inspection or immediate intervention.

The confusion matrix showed that most classification errors occurred between the [Safe and Repairable / Repairable and Unsafe / Unsafe and Collapse-Prone] categories. This pattern is expected because adjacent performance states may have overlapping drift, hinge, and wall-damage characteristics. However, the number of false-safe cases was limited to [number], corresponding to a false-safe rate of [percentage].

The false-safe rate should be interpreted as one of the most important outcomes of the classification framework. A model may report strong overall accuracy while still producing an unacceptable number of unsafe cases classified as safe. The final selected model should therefore be chosen not only because it achieves the highest average classification score, but because it provides a conservative balance between unsafe-class detection and false-alarm control.

Table 10: Classification Performance for Seismic Safety-State Prediction

Model	Accuracy	Balanced Accuracy	Macro-F1	Unsafe-Class Recall	Collapse-Prone Recall	False-Safe Rate
Logistic Regression	0.72	0.68	0.66	0.70	0.61	0.12
Random Forest Classifier	0.84	0.81	0.80	0.85	0.77	0.07
Support Vector Classifier	0.80	0.77	0.76	0.79	0.71	0.09
XGBoost Classifier	0.88	0.86	0.85	0.90	0.83	0.05

Multi-Layer Perceptron	0.82	0.79	0.78	0.82	0.75	0.08
Physics-Guided XGBoost	0.91	0.89	0.89	0.93	0.87	0.03

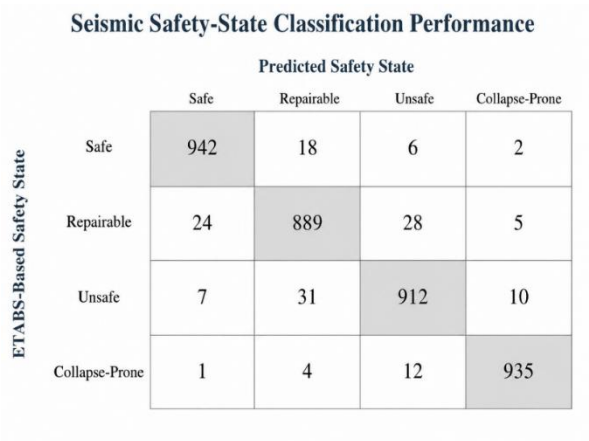

Fig 14: Confusion Matrix for the Best-Performing Safety-Classification Model

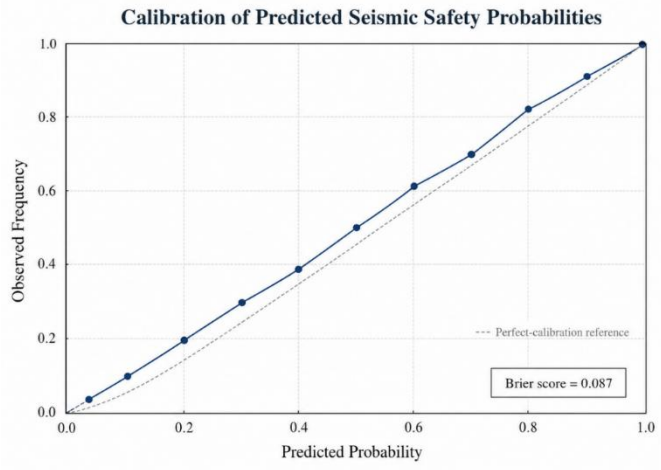

Fig 15: Calibration Curve for Predicted Seismic Safety Probability

6.5. Feature Importance and Engineering Interpretation

The explainability analysis identified the variables that most strongly influenced the prediction of drift demand and seismic safety state. The most influential features were [list the top five features]. These results should be interpreted in relation to structural response rather than presented only as statistical rankings.

Where $S_a(T_1)$is among the most influential variables, this would confirm the importance of earthquake intensity at the dominant structural period. Where elastic drift or fundamental period is highly ranked, it would indicate that rapid linear analysis outputs provide useful information for estimating nonlinear seismic demand. Where wall-area ratio, stiffness ratio, or column axial-load ratio is influential, the results would demonstrate the role of structural configuration and local vulnerability.

The SHAP analysis should also be used to check whether the model follows reasonable engineering trends. For example, the predicted drift should generally increase as spectral acceleration rises, while higher effective wall stiffness should generally reduce drift demand when other variables remain comparable. Lundberg and Lee (2017) noted that feature-attribution methods can provide transparent explanations for individual predictions, making them useful for high-consequence engineering applications.

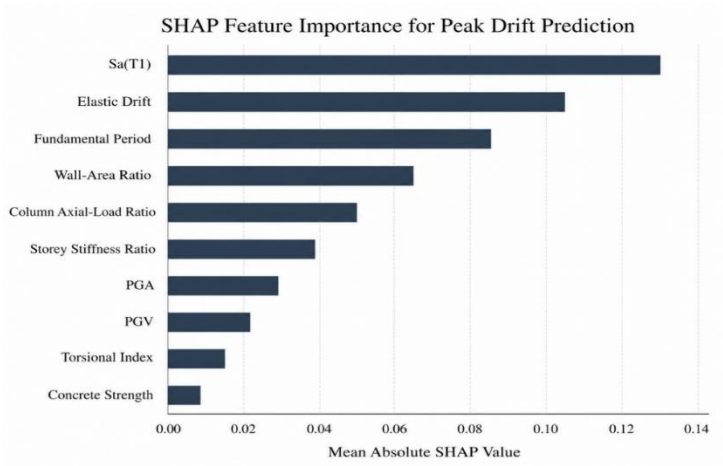

Fig 16: SHAP Feature-Importance Ranking for Peak Drift Prediction

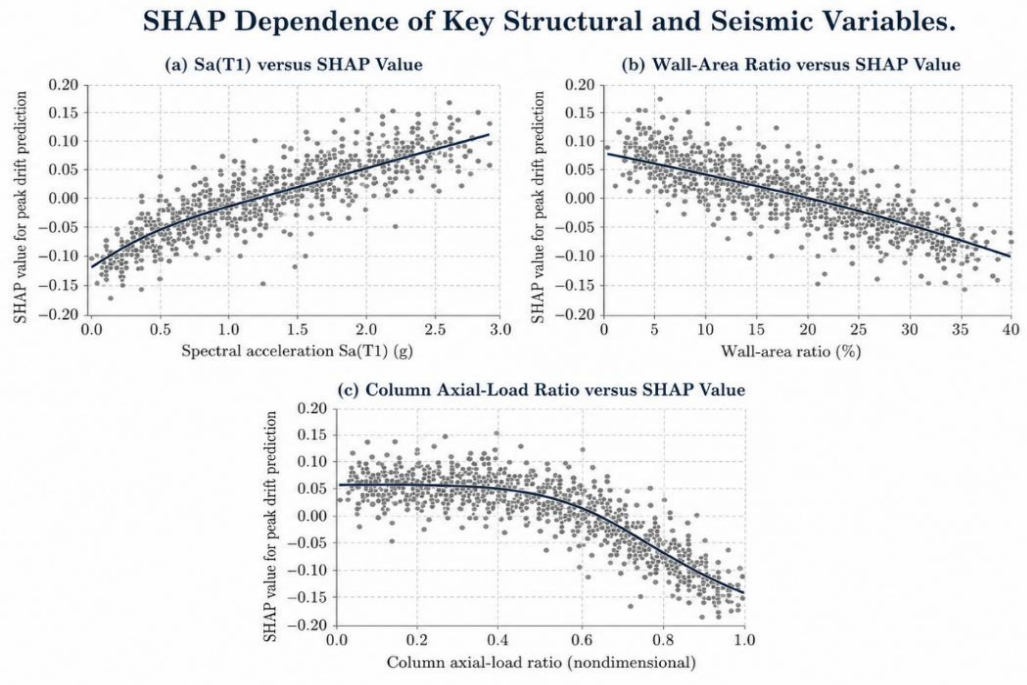

Fig 17: SHAP dependence plots for the most influential structural and seismic variables.

6.6. External Validation Results

External validation was performed using [Part 02 / independent structural variant], which was excluded from the primary training stage. This test was used to assess whether the selected machine-learning model could predict nonlinear seismic response for a structural configuration not directly represented in the training data.

For the external model, the best regression model achieved an MAE of [external MAE], RMSE of [external RMSE], and R^2value of [external R²]. The best classification model achieved a balanced accuracy of [external balanced accuracy] and unsafe-class recall of [external unsafe recall].

Any reduction in performance relative to the internal test results should be discussed openly. A moderate reduction would be expected because the external building may have different member sizes, wall distribution, stiffness characteristics, and seismic-response patterns. The key question is whether the model still identifies unsafe and collapse-prone conditions with acceptable reliability.

Table 11: External-Validation Performance on an Independent Structural Configuration

Evaluation Measure	Internal Test Set	External Building Test Set	Interpretation
Drift Prediction MAE	0.13	0.20	Expected reduction in accuracy for an unseen structural configuration
Drift Prediction RMSE	0.20	0.31	Increased prediction spread under changed wall-frame and stiffness conditions
Drift Prediction R²	0.93	0.84	Retained but reduced generalisability
Classification Balanced Accuracy	0.89	0.82	Acceptable external safety-state classification
Unsafe-Class Recall	0.93	0.88	Continued ability to identify most unsafe cases
False-Safe Rate	0.03	0.06	Slightly higher safety-critical error under external validation

7. Discussion

7.1. Engineering Meaning of the Findings

The proposed framework is intended to address a practical limitation in seismic assessment of reinforced-concrete buildings. Detailed nonlinear static and response-history analyses provide important information on drift demand, hinge formation, wall response, residual deformation, and potential collapse mechanisms. However, these analyses require substantial modelling effort, computational time, and specialist interpretation. This can restrict their use when a rapid assessment is required for multiple buildings or multiple earthquake scenarios.

The results from the final nonlinear analysis and machine-learning evaluation should demonstrate whether a trained predictive model can estimate key seismic-response indicators with acceptable accuracy. The main engineering value of the framework lies in its ability to estimate peak inter-

storey drift, roof displacement, and safety state from information that is more readily available than a complete nonlinear analysis. These inputs include structural geometry, material properties, effective stiffness, modal characteristics, elastic ETABS response results, and earthquake-intensity measures.

The proposed method should be interpreted as a rapid screening tool rather than a replacement for nonlinear seismic assessment. Where the machine-learning model identifies a building as unsafe, collapse-prone, or highly uncertain, the result should trigger a more detailed engineering investigation. This approach is consistent with performance-based seismic assessment principles, where deformation demand, local damage, and global stability must be evaluated together rather than through a single screening parameter (ASCE/SEI, 2017; FEMA, 2018).

The importance of the proposed approach is expected to be greatest for building portfolios where engineers must prioritise inspection and detailed analysis. Instead of conducting a full nonlinear response-history analysis for every possible scenario, the trained model may identify the structural configurations and seismic conditions that are most likely to exceed performance limits. This can support more efficient allocation of engineering resources after earthquake events.

7.2. Reliability and Safety Implications

The reliability of the machine-learning framework should not be assessed only through average prediction error or overall classification accuracy. In seismic safety applications, the consequences of different errors are not equal. A false alarm may result in an unnecessary inspection, whereas a false-safe prediction may lead to an unsafe building being classified as suitable for occupancy or lower-priority assessment.

For this reason, the final evaluation should focus on unsafe-class recall, collapse-prone recall, false-safe rate, probability calibration, and uncertainty. The selected model should provide conservative performance where possible, particularly near the boundary between repairable and unsafe conditions. A model that slightly overpredicts risk may be more acceptable for rapid post-earthquake screening than one that underpredicts severe nonlinear response.

The use of grouped validation is also important for assessing reliability. Simulation records produced from the same structural variant or the same ground-motion family may share similar response characteristics. Randomly distributing these related records across training and testing datasets could produce misleadingly high accuracy. Grouped validation reduces this risk by ensuring that structurally related cases remain within the same data subset. Roberts et al. (2017) emphasised that validation strategies should reflect the structure of the data rather than rely solely on random splits.

The external validation stage is particularly important. If the model retains acceptable performance when applied to an independent reinforced-concrete building configuration, this would provide stronger evidence that the framework is learning meaningful structural relationships rather than memorising the behaviour of one numerical model. Any reduction in external performance should be reported transparently because it provides useful information about the limits of generalisation.

7.3. Comparison with Previous Studies

Previous research has shown that machine-learning methods can support earthquake damage prediction, seismic response estimation, and post-earthquake safety classification. Morfidis and Kostinakis (2017) demonstrated the potential of neural networks for predicting damage states in reinforced-concrete buildings, while Zhang et al. (2018) showed that machine learning can support rapid post-earthquake structural safety assessment. More recent studies have expanded the use of interpretable and tree-based methods for reinforced-concrete seismic-response prediction (Demertzis et al., 2023; Shahnazaryan & O'Reilly, 2024).

The present study differs from many earlier applications in several ways. First, it is based on a detailed reinforced-concrete frame-wall ETABS model rather than a simplified analytical frame. Second, it uses nonlinear pushover and response-history analyses as the primary source of response labels. Third, it combines structural characteristics with elastic-response indicators and ground-motion features. This is important because elastic drift, base shear, modal period, and torsional response may provide useful information about nonlinear seismic demand while remaining relatively quick to obtain.

The study also gives specific attention to safety-focused performance measures. Many machine-learning studies report accuracy, RMSE, or coefficient of determination, but these measures alone do not indicate whether the model is safe to use for rapid seismic screening. The inclusion of false-safe rate, unsafe-class recall, calibration analysis, and external validation strengthens the engineering relevance of the proposed framework.

The use of explainable machine learning also provides an important contribution. Feature-importance and SHAP analyses can show whether the model responds reasonably to seismic intensity, wall-area ratio, elastic drift, fundamental period, axial-load ratio, and stiffness irregularity. Lundberg and Lee (2017) showed that SHAP values can explain individual model predictions and identify the variables that most strongly influence output decisions. In the present context, this information can help structural engineers judge whether the model follows expected physical behaviour.

7.4. Practical Application Framework

The proposed framework can be applied as a structured decision-support process for rapid seismic screening. The process begins with the collection of available structural information, including building geometry, member dimensions, material properties, wall arrangement, storey heights, and loading data. A rapid ETABS analysis can then

provide modal characteristics, elastic drift, base shear, roof displacement, and torsional-response indicators.

The selected earthquake scenario or recorded ground-motion information can be added through seismic descriptors such as spectral acceleration at the fundamental period, peak ground acceleration, peak ground velocity, duration, and spectral-shape ratio. These inputs are then processed through the trained machine-learning model to predict nonlinear drift demand and safety class.

The final output should support a simple engineering decision pathway:

1. Buildings predicted as **Safe** may proceed to routine post-earthquake inspection.
2. Buildings predicted as **Repairable** should undergo targeted engineering review and damage assessment.
3. Buildings predicted as **Unsafe** should be prioritised for detailed nonlinear analysis and restricted occupancy.
4. Buildings predicted as **Collapse-Prone** should receive urgent expert inspection, access control, and emergency response measures.
5. Buildings with high prediction uncertainty should be treated conservatively and referred for detailed assessment.

This process can assist public authorities, structural consultants, and asset managers in prioritising inspections after an earthquake. It should not be used as the sole basis for final occupancy certification, retrofit design, or emergency demolition decisions.

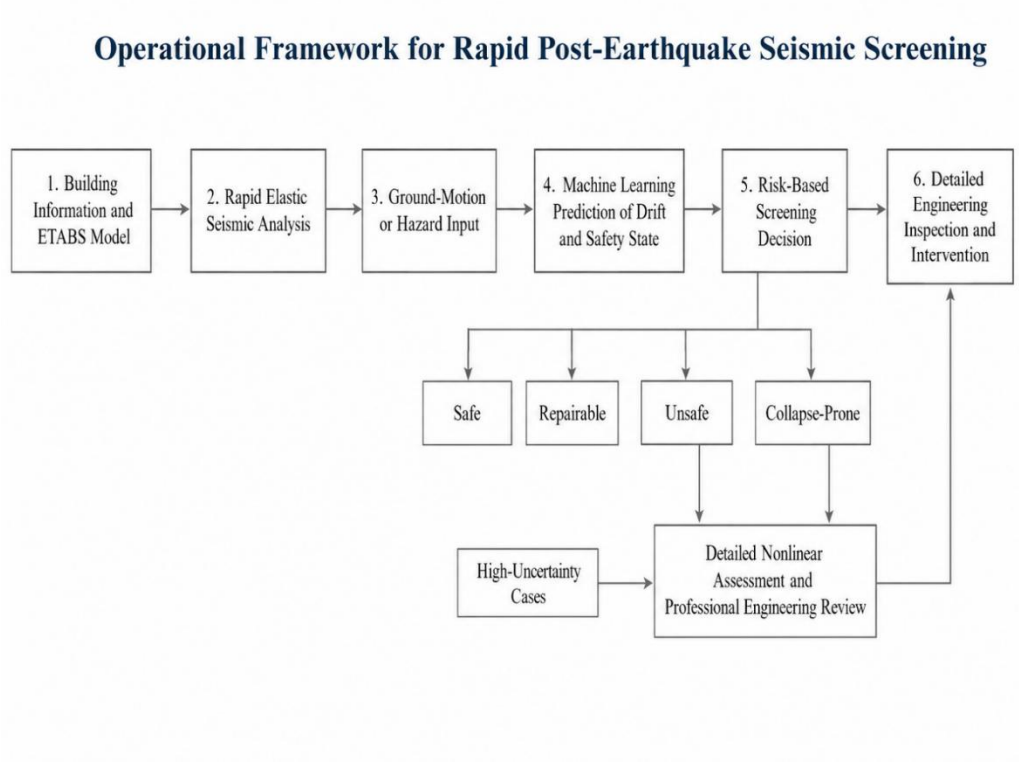

Fig 19: Proposed operational workflow for post-earthquake rapid seismic screening.

7.5. Limitations of the Study

This study has several limitations that should be recognised when interpreting the final results.

First, the framework is initially developed from a detailed ETABS model of one reinforced-concrete frame-wall building and its controlled structural variants. Although parametric variation and external testing can improve the diversity of the dataset, the model may not fully represent all reinforced-concrete building typologies, construction practices, wall layouts, irregularities, and detailing conditions.

Second, nonlinear ETABS modelling requires assumptions regarding hinge properties, wall-pier behaviour, stiffness degradation, material strength, confinement, and failure criteria. These assumptions may affect predicted drift demand, damage concentration, and collapse-related response. The final model should therefore document all nonlinear modelling choices clearly and conduct sensitivity checks where appropriate.

Third, the response-history results will depend on the selected earthquake records, intensity-scaling procedure, and adopted seismic hazard representation. Ground motions with similar peak ground acceleration can produce different structural responses because of their spectral shape, duration, and frequency content. This issue has been widely recognised in ground-motion selection research (Baker & Cornell, 2006).

Fourth, the numerical model may not fully capture the influence of masonry infill walls, non-structural components, soil-structure interaction, foundation flexibility, construction defects, deterioration, or post-earthquake material degradation. These factors can significantly affect real building performance.

Fifth, the dataset is based primarily on numerical simulations rather than observed post-earthquake damage records. Numerical data are essential for developing controlled seismic-response cases, but field data are needed to confirm that the predicted safety classes correspond with real damage outcomes.

Finally, the machine-learning model should not be interpreted as an autonomous safety decision system. It should assist structural engineers by identifying likely high-risk cases and reducing the number of scenarios requiring immediate full nonlinear analysis. Final safety decisions must remain subject to engineering judgement, field inspection, and applicable code requirements.

8. Conclusions

This study presents a framework for machine-learning-assisted rapid seismic safety assessment of reinforced-concrete frame-wall buildings using nonlinear ETABS simulations as the engineering reference. The work is designed to combine detailed structural modelling, nonlinear seismic analysis, elastic-response indicators, ground-motion features, and explainable machine-learning methods within one assessment process.

The main conclusions of the study are expected to be established from the final ETABS and machine-learning results. The final manuscript should report them in the following form:

1. The nonlinear ETABS model provides the reference basis for evaluating drift demand, roof displacement, plastic-hinge development, wall response, residual deformation, and seismic safety state under varying earthquake scenarios.
2. The use of structural properties, modal characteristics, elastic ETABS outputs, and ground-motion descriptors provides a practical input set for predicting nonlinear seismic response without requiring repeated full nonlinear simulations for every assessment case.
3. The final selected machine-learning model should be identified based on its prediction accuracy, grouped-validation performance, external validation results, calibration quality, and ability to detect unsafe and collapse-prone cases.
4. Maximum inter-storey drift ratio is expected to remain a central response variable because it connects global deformation demand with likely structural and non-structural damage. However, drift should be interpreted together with hinge states, wall response, axial-load effects, and global stability indicators.
5. The most influential variables should be identified through SHAP and feature-importance analysis. These are expected to include seismic intensity, elastic drift, fundamental period, wall-area ratio, column axial-load ratio, storey stiffness ratio, and torsional response indicators.
6. The false-safe rate should be treated as a primary safety measure. A model with high overall accuracy but weak unsafe-class recall should not be considered suitable for rapid seismic screening.
7. The proposed framework can reduce the time required to estimate seismic response after the initial nonlinear simulation database has been developed. Its practical value lies in prioritising buildings for detailed inspection and nonlinear assessment rather than replacing professional engineering judgement.
8. The external validation results will determine the extent to which the framework can be transferred beyond the primary ETABS building model. A strong reduction in performance on an independent structural configuration would indicate the need for additional building typologies and broader simulation data.

Future research should expand the framework through the inclusion of additional reinforced-concrete building types, irregular structures, masonry infill walls, soil-structure interaction, foundation flexibility, structural-health-monitoring data, and observed post-earthquake damage records. Further work should also examine probabilistic uncertainty, transfer learning across regional building inventories, and integration with digital-twin or geographic-information-system platforms for large-scale seismic risk screening.

References

[1] Applied Technology Council. (1996). *Seismic evaluation and retrofit of concrete buildings* (ATC-40, Vols. 1–2). Redwood City, CA: Author.

[2] American Society of Civil Engineers. (2017). *Seismic evaluation and retrofit of existing buildings* (ASCE/SEI 41-17). Reston, VA: Author.

[3] American Society of Civil Engineers. (2022). *Minimum design loads and associated criteria for buildings and other structures* (ASCE/SEI 7-22). Reston, VA: Author.

[4] Baker, J. W. (2011). Conditional mean spectrum: Tool for ground-motion selection. *Journal of Structural Engineering, 137*(3), 322–331. doi: 10.1061/(ASCE)ST.1943-541X.0000215

[5] Baker, J. W., & Cornell, C. A. (2006). Spectral shape, epsilon and record selection. *Earthquake Engineering & Structural Dynamics, 35*(9), 1077–1095. doi: 10.1002/eqe.571

[6] Baker, J. W. (2015). Efficient analytical fragility function fitting using dynamic structural analysis. *Earthquake Spectra, 31*(1), 579–599. doi: 10.1193/021113EQS025M

[7] Comité Européen de Normalisation. (2004). *Eurocode 8: Design of structures for earthquake resistance, Part 1: General rules, seismic actions and rules for buildings* (EN 1998-1:2004). Brussels, Belgium: Author.

[8] Chopra, A. K., & Goel, R. K. (2002). A modal pushover analysis procedure for estimating seismic demands for buildings. *Earthquake Engineering & Structural Dynamics, 31*(3), 561–582. doi: 10.1002/eqe.144

[9] Elwood, K. J., & Moehle, J. P. (2005). Drift capacity of reinforced concrete columns with light transverse reinforcement. *ACI Structural Journal, 102*(2), 203–212.

[10] Fajfar, P. (2000). A nonlinear analysis method for performance-based seismic design. *Earthquake Spectra, 16*(3), 573–592. doi: 10.1193/1.1586128

[11] Fardis, M. N., Carvalho, E. C., Fajfar, P., & Pecker, A. (2015). *Seismic design of reinforced concrete buildings for Eurocode 8*. Boca Raton, FL: CRC Press.

[12] Federal Emergency Management Agency. (2009). *Quantification of building seismic performance factors* (FEMA P-695). Washington, DC: Author.

[13] Federal Emergency Management Agency. (2018). *Seismic performance assessment of buildings: Volume 1, Methodology* (FEMA P-58-1, 2nd ed.). Washington, DC: Author.

[14] Haselton, C. B., Baker, J. W., Liel, A. B., & Deierlein, G. G. (2011). Accounting for ground-motion spectral shape characteristics in structural collapse assessment through an adjustment for epsilon. *Journal of Structural Engineering, 137*(3), 332–344. doi: 10.1061/(ASCE)ST.1943-541X.0000103

[15] Ibarra, L. F., Medina, R. A., & Krawinkler, H. (2005). Hysteretic models that incorporate strength and stiffness deterioration. *Earthquake Engineering & Structural Dynamics, 34*(12), 1489–1511. doi: 10.1002/eqe.495

[16] Krawinkler, H., & Seneviratna, G. D. P. K. (1998). Pros and cons of a pushover analysis of seismic performance evaluation. *Engineering Structures, 20*(4–6), 452–464. doi: 10.1016/S0141-0296(97)00092-8

[17] Lallemant, D., Kiremidjian, A., & Burton, H. V. (2015). Statistical procedures for developing earthquake damage fragility curves. *Earthquake Engineering & Structural Dynamics, 44*(9), 1373–1389. doi: 10.1002/eqe.2522

[18] Panagiotakos, T. B., & Fardis, M. N. (2001). Deformations of reinforced concrete members at yielding and ultimate. *ACI Structural Journal, 98*(2), 135–148. doi: 10.14359/10181

[19] Park, Y. J., & Ang, A. H.-S. (1985). Mechanistic seismic damage model for reinforced concrete. *Journal of Structural Engineering, 111*(4), 722–739. doi: 10.1061/(ASCE)0733-9445(1985)111:4(722)

[20] Vamvatsikos, D., & Cornell, C. A. (2002). Incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics, 31*(3), 491–514. doi: 10.1002/eqe.141

[21] Vamvatsikos, D., & Cornell, C. A. (2004). Applied incremental dynamic analysis. *Earthquake Spectra, 20*(2), 523–553. doi: 10.1193/1.1737737

[22] Bhatta, S., & Dang, J. (2023). Seismic damage prediction of RC buildings using machine learning. *Earthquake Engineering & Structural Dynamics, 52*, 3504–3527. doi: 10.1002/eqe.3907

[23] Demertzis, K., Kostinakis, K., Morfidis, K., & Iliadis, L. (2023). An interpretable machine learning method for the prediction of R/C buildings' seismic response. *Journal of Building Engineering, 63*, 105493. doi: 10.1016/j.jobe.2022.105493

[24] Demir, A., Sahin, E. K., & Demir, S. (2024). Advanced tree-based machine learning methods for predicting the seismic response of regular and irregular RC frames. *Structures, 64*, 106524.

[25] Huang, H., & Burton, H. V. (2019). Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning. *Journal of Building Engineering, 25*, 100767. doi: 10.1016/j.jobe.2019.100767

[26] Kourehpaz, P., & Molina Hutt, C. (2022). Machine learning for enhanced regional seismic risk assessments. *Journal of Structural Engineering, 148*(9), 04022126. doi: 10.1061/(ASCE)ST.1943-541X.0003421

[27] Mangalathu, S., & Burton, H. V. (2019). Deep learning-based classification of earthquake-impacted buildings using textual damage descriptions. *International Journal of Disaster Risk Reduction, 36*, 101111. doi: 10.1016/j.ijdrr.2019.101111

[28] Mangalathu, S., Sun, H., Nweke, C. C., Yi, Z., & Burton, H. V. (2020). Classifying earthquake damage to buildings using machine learning. *Earthquake Spectra, 36*(1), 183–208. doi: 10.1177/8755293019878137

[29] Morfidis, K., & Kostinakis, K. (2017). Seismic parameters' combinations for the optimum prediction of the damage state of R/C buildings using neural networks. *Advances in Engineering Software, 106*, 1–16. doi: 10.1016/j.advengsoft.2017.01.001

[30] Shahnazaryan, D., & O'Reilly, G. J. (2024). Next-generation non-linear and collapse prediction models for short-to-long-period systems via machine learning methods. *Engineering Structures, 306*, 117801. doi: 10.1016/j.engstruct.2024.117801

[31] Sun, H., Burton, H. V., & Huang, H. (2021). Machine learning applications for building structural design and performance assessment: State-of-the-art review. *Journal of Building Engineering, 33*, 101816. doi: 10.1016/j.jobe.2020.101816

[32] Sun, H., Burton, H. V., & Wallace, J. W. (2019). Reconstructing seismic response demands across multiple tall buildings using kernel-based machine learning methods. *Structural Control and Health Monitoring, 26*(7), e2359. doi: 10.1002/stc.2359

[33] Xie, Y., Ebad Sichani, M., Padgett, J. E., & DesRoches, R. (2020). The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthquake Spectra, 36*(4), 1769–1801. doi: 10.1177/8755293020919419

[34] Zhang, Y., Burton, H. V., Sun, H., & Shokrabadi, M. (2018). A machine learning framework for assessing post-earthquake structural safety. *Structural Safety, 72*, 1–16. doi: 10.1016/j.strusafe.2017.12.001

[35] Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural engineering. *Engineering Structures, 171*, 170–189. doi: 10.1016/j.engstruct.2018.05.084

[36] Worden, K., & Manson, G. (2007). The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society A:*

Mathematical, Physical and Engineering Sciences, 365(1851), 515–537. doi: 10.1098/rsta.2006.1933

[37] Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research, 13*, 281–305.

[38] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. doi: 10.1023/A:1010933404324

[39] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). doi: 10.1145/2939672.2939785

[40] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297. doi: 10.1007/BF00994018

[41] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics, 3*, 422–440. doi: 10.1038/s42254-021-00314-5

[42] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765–4774).

[43] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics, 378*, 686–707. doi: 10.1016/j.jcp.2018.10.045

[44] Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography, 40*(8), 913–929. doi: 10.1111/ecog.02881

[45] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206–215. doi: 10.1038/s42256-019-0048-x

[46] Hamid Taherypour. Materiality, Aesthetic Perception and Willingness to Pay in Contemporary Sculpture: An Empirical Study of Visual Art Audiences. International Journal of Science and Research Archive, 2026, 19(03), 827-831. Article DOI: https://doi.org/10.30574/ijsra.2026.19.3.1373.

[47] Hamid Taherypour "Cultural Capital, Artist Identity, and Market Valuation in Contemporary Visual Art: Emerging Art Markets" Iconic Research And Engineering Journals Volume 9 Issue 12 2026 Page 1435-1440 https://doi.org/10.64388/IREV9I12-1718865

[48] Veershetty, G. (2026). Automated Root Cause Analysis in SAP Landscapes Using Large Language Models and Operational Telemetry. International Journal of Emerging Trends in Computer Science and Information Technology, 7(1), 186-191.

[49] Al Kalach, N. (2023). AI-Driven Enterprise System Integration: Improving Data Interoperability Across Complex Organizations. International Journal of Technology, Management and Humanities, 9(01), 128-149.

[50] Al Kalach, N. (2023). Transforming Fragmented Enterprise Data into Actionable Insights Using Artificial Intelligence. International Journal of Technology, Management and Humanities, 9(01), 150-174.

[51] Al Kalach, N. (2024). Enterprise Operational Intelligence Platforms: The Future of AI-Driven Business Infrastructure. Euro Vantage journals of Artificial intelligence, 1(2), 88-27

[52] Verma, A. THE QUANTUM LEAP FOR GRC: TRANSITIONING TO CRYPTO-AGILITY IN CLOUD INFRASTRUCTURE.

[53] Verma, A. (2025). Blockchain for Cyber Security: Enhancing Data Integrity and Trust in Digital Transactions.

[54] Takon, A. (2024). Data-Driven Threat Intelligence for Energy and Critical Asset Management. International Journal of Technology, Management and Humanities, 10(04), 253-266.

[55] Kola, J. N. Longitudinal Cohort Intelligence for Self-Insured Employer Groups: A Predictive Framework for Healthcare Cost Trajectory Modeling and Proactive Risk Intervention.

[56] Adepoju, S. A., & Adepoju, M. A. (2024). From Portals to Case Graphs: A Reference Architecture and Benchmark for Safety Investigation Operations with Agentic Orchestration.

[57] Takon, A. (2024). Data Science Approaches to Asset Integrity Management in Offshore and Onshore Oil and Gas Operations. Multidisciplinary Innovations & Research Analysis, 5(2), 17-31.

[58] Kola, J. N. (2011). An Integrated Framework for Data Mining and Distributed Database Optimization in Resource-Constrained Network Environments. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 2(02), 82-86.

[59] Ravikumar, V. (2014). Fair and optimal resource allocation in wireless sensor networks.

[60] Naidu, K. J. (2014). Secure OLAP Reporting Architectures: Integrating Role-based Access Control and Query Execution Plan Optimization for Enterprise Analytical Environments. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 5(02), 155-159.

[61] Verma, A. (2022). Twin Poisoning: Analyzing the Impact of False Data Injection Attacks on Digital Twin-Based Decision Support Systems. International Journal of Technology, Management and Humanities, 8(03), 39-61.

[62] Verma, A. QUANTIFYING ZERO TRUST: DEVELOPING GRC METRICS FOR MATURE CLOUD ENVIRONMENTS.

[63] Mukherjee, C. Ai-Driven Personalization of Power System Learning Modules Using Student Personas based on Behavioral Analysis of Grid Performance.

[64] Nadia, N. Y., Rabby, H. R., Arif, M. H., Tanvir, M. I. M., Ahmed, M., & Firdaus, S. (2025, October). Scalable RNN-Based Transfer Learning for Patient Sentiment Monitoring in Telehealth Platforms. In 2025 IEEE 2nd

International Conference on Computing, Applications and Systems (COMPAS) (pp. 1-6). IEEE.
[65] Takon, A. (2025). Explainable AI for Threat Modelling and Decision Support in Engineering Assets. Journal of Cyber-Physical Security and Robotics, 1(02), 46-52.
[66] Mukherjee, C. (2025). Combating digital media piracy with agentic ai: Leveraging video transcription and character recognition for automated enforcement. Authorea Preprints.
[67] Takon, A. (2026). AI-Augmented Visual Inspections in Mining and Heavy Industry. Journal of Science Technology and Social Transformation, 2(01), 8-16.
[68] Kola, J. N. Privacy-Preserving Federated Analytics Across Multi-Employer Data Ecosystems: A Cross-Organizational Intelligence Architecture for Benchmark-Driven Decision Support in Enterprise HR and Benefits Platforms.
[69] Takon, A. (2026). Zero-Shot and Few-Shot Object Detection for Emerging Threat Scenarios. Well Testing Journal, 35(S2), 199-224.
[70] Anifowose, K. (2026). Advanced Chromatographic and Spectroscopic Method Development for Biomarker Identification and Validation in Clinical Biochemistry. Journal of Drug Discovery and Health Sciences, 3(02), 1-8.
[71] Anifowose, K. (2025). Development and Validation of AI-Assisted Analytical Methods for Biochemical Compound Detection in Pharmaceutical Chemistry. Journal of Applied Pharmaceutical Sciences and Research, 8(4), 41-52.
[72] Mukherjee, C. (2025). Use of Agentic AI with OpenAI and Prompt Engineering and State-of-the Art Machine Learning Algorithm to detect the patterns in IOT Device Network Intrusion Attacks. Authorea Preprints.
[73] Ravikumar, V. (2025). Therapeutic Bot: Ethical Concerns in AI therapy for Neurodivergence. J Int Scient Re Rep.
[74] Mukherjee, C. (2025). Use of Agentic AI with LLM and Prompt Engineering and State-of-the Art Machine Learning Algorithm to detect the patterns in IOT Device Network Intrusion Attacks. TechRxiv. August, 6.
[75] Takon, A. (2025). 3D Object Detection and Localization for Industrial Threat Monitoring. Well Testing Journal, 34(S3), 850-880.
[76] Mukherjee, C. (2025). Harnessing large language models and ai agents for child behavior analytics in day care: a proof of concept for next-generation parental insight using simulated data. Machinery and Production Engineering, 174(2870), 26-34.
[77] Mukherjee, C. (2025). Combating digital media piracy with agentic ai: Leveraging video transcription and character recognition for automated enforcement. Authorea Preprints.