CAPSTONE PROJECT 1

Planning Document

Concrete Strength Analysis and Prediction for Quality Control Using Machine Learning

by

Ong Ze Ming

14044101

Bachelor of Software Engineering (Hons)

Supervisor: Dr. Richard Wong Teck Ken

Semester: September 2024

Date: 7 August, 2025

School of Computing and Artificial Intelligence

Faculty of Engineering and Technology

Sunway University

1.0 Introduction

1.1 Background

Concrete is the most fundamental material in global civil engineering and construction, providing the structural foundation for modern infrastructure (Tak et al., 2025). A crucial property for ensuring the safety and long-term stability of these structures is its compressive strength, which measures the material's ability to withstand loads without failure. The accurate and timely prediction of this strength is therefore essential for optimizing structural performance, ensuring project safety, and meeting stringent engineering standards (Tak et al., 2025).

The construction industry has long been driven by the need for precise and prompt strength data, as critical decisions such as formwork removal and the application of subsequent loads depend on the concrete reaching a specific strength threshold.1 However, traditional assessment techniques, which typically rely on a 28-day curing period for laboratory-cast specimens, often lack the immediate feedback required in dynamic construction environments (Nithurshan & Elakneswaran, 2023). This inherent delay can lead to compromised safety, significant project delays, and increased costs. Over the past decade, the limitations of these conventional methods, coupled with the complex, non-linear nature of concrete, have motivated a significant shift in material science towards advanced, data-driven approaches leveraging artificial intelligence (AI) and machine learning (ML).

1.2 Problem Statement

Traditional methods for assessing concrete strength are both time-consuming and costly, introducing significant delays into construction workflows and failing to provide the real-time feedback necessary for modern project management (DeWalt, n.d.). A more fundamental limitation is that laboratory-controlled test specimens often do not accurately represent the actual strength of concrete within a structure. The final in-situ strength is profoundly affected by variable field conditions, such as ambient temperature, humidity, and the quality of on-site compaction, which are not reflected in idealized lab samples (DeWalt, n.d.).

Furthermore, existing empirical and mathematical models struggle to account for the complex, non-linear interactions between the multiple variables that determine concrete strength, including cement content, water, aggregates, and chemical admixtures (Tak et al., 2025). These cumulative limitations highlight a critical need for a more efficient, accurate, and reliable approach to concrete strength prediction that can overcome the challenges of the conventional paradigm.

1.3 Project Aim

This project aims to analyze and predict concrete strength for quality control purposes by leveraging machine learning techniques. The goal is to develop predictive models that can overcome the limitations of traditional testing methods, providing more accurate and timely strength assessments.

1.4 Project Objective

To achieve the project's aim, the following objectives have been set:

- To conduct a comprehensive review of existing literature on concrete properties, traditional strength testing methods, and the application of machine learning in this field.
- To identify the key algorithms used in concrete strength prediction and evaluate their performance.
- To explore and delineate the persistent challenges and research gaps in the current body of knowledge.
- To provide a solid foundation for a final year project focused on developing a machine learning model for concrete strength analysis and prediction.

1.5 Project Scope

This literature review will analyze existing research on concrete strength prediction. The review will begin by examining the fundamental properties of concrete and the traditional methods used for its testing. We will then delve into the principles of machine learning and its advanced applications in predicting concrete strength. The scope includes identifying key algorithms, evaluating their performance, and outlining the remaining challenges in the field to establish a comprehensive basis for a final year project on this topic.

2.0 Literature Review

2.1 Introduction

Concrete, a cornerstone of modern infrastructure, plays an indispensable role in civil engineering and construction globally. Its inherent strength, particularly its compressive strength, is a fundamental mechanical property and a critical parameter in the design and development of structures, directly influencing their safety and long-term stability (Thapa, 2024). The ability to accurately and promptly predict concrete strength is paramount for optimizing the performance of structural components, ensuring overall structural integrity, and adhering to stringent engineering specifications. The need for precise and timely predictions underscores a critical demand within the construction industry. Traditional methods for assessing concrete strength, although foundational, often fall short in providing the immediate feedback required for dynamic construction environments, thus creating a compelling motivation for exploring advanced predictive techniques.

2.1.1 Significance of Concrete Strength in Civil Engineering and Construction Quality Control

The compressive strength of concrete is essential to stable structural design, with both linear and non-linear characteristics of concrete typically evaluated through this key mechanical property (Thapa, 2024). The emphasis on precise and prompt strength prediction highlights a crucial need in contemporary construction. The consequences of inaccurate or delayed strength data can range from compromised structural safety to significant project delays and cost overruns. For instance, decisions regarding formwork removal or the application of subsequent loads are directly contingent on the concrete achieving a specified strength. If strength prediction is crucial for optimizing performance and ensuring structural integrity, then any method that improves this prediction offers substantial value. This inherent requirement implicitly critiques traditional methods that, despite their accuracy, may not offer the necessary timeliness, thereby establishing the practical relevance and urgency of research into advanced predictive methodologies.

2.1.2 Evolution of Concrete Strength Assessment: From Traditional Methods to Data-Driven Approaches

Historically, the determination of concrete strength has predominantly relied on destructive testing methods conducted in controlled laboratory environments (Altuncı, 2024). These empirical and traditional mathematical models have consistently faced challenges in accurately forecasting actual concrete strength. This limitation stems from the complex, non-linear nature of concrete itself, coupled with the inability of simplified empirical equations to precisely account for the intricate interplay of various independent variables. Such inadequacies have often led to decreased accuracy in strength estimations.

The last decade has witnessed a profound transformation in this domain, driven by the rapid advancements in artificial intelligence (AI) and machine learning (ML). This technological evolution offers promising new avenues to overcome the limitations, enabling the determination of accurate concrete compressive strength (CS) through the utilization of real-time databases (Thapa, 2024). This transition from empirical and traditional methods to data-driven ML approaches represents a significant paradigm shift in material science and civil engineering. The fundamental limitation of older models lies in their inability to adequately capture the inherent complexity and non-linearity of concrete behavior. Machine learning's strength, conversely, lies precisely in its capacity to discern and model such complex, non-linear relationships, positioning it as a necessary advancement rather than merely an alternative solution to long-standing challenges in concrete strength prediction.

2.1.3 Overview of Machine Learning's Role in Civil Engineering

Machine learning is rapidly revolutionizing diverse facets of civil engineering, extending its influence far beyond concrete strength prediction. Its applications now encompass predicting material properties like concrete compressive strength, automating inspection processes, and optimizing various construction workflows (Dzięcioł, 2024). These models possess a unique ability to capture complex, non-linear relationships within data, often surpassing the predictive accuracy of conventional analytical models (Nyokum & Tamut, 2025).

The widespread adoption of ML is evident in its diverse applications, which include predicting infrastructure service life, forecasting traffic volumes, estimating soil properties from limited tests, classifying pavement distress, and enhancing overall construction safety and efficiency (Nyokum & Tamut, 2025). This broad integration of ML across various civil engineering domains signals a pervasive trend towards data-driven decision-making and intelligent automation throughout the industry. The successful application of machine learning in areas such as structural health monitoring, traffic management, and soil analysis demonstrates its general capability to effectively handle complex engineering data. This broader context positions the proposed project not as an isolated application but as an integral part of a larger, transformative movement within civil engineering, applying a powerful and broadly applicable methodology to a specific, high-impact problem within a field undergoing significant technological change.

2.1.4 Scope and Structure of the Literature Review

This literature review aims to systematically analyze existing research pertaining to concrete strength prediction. The discussion will progress from an examination of fundamental material properties and traditional testing methodologies to an exploration of the underlying principles and advanced applications of machine learning in this context. The review will identify key algorithms, evaluate their performance, and delineate the persistent challenges that remain in the field. By highlighting these research gaps and future directions, this document seeks to provide a comprehensive foundation for a final year project focused on concrete strength analysis and prediction for quality control using machine learning.

2.2 Fundamentals of Concrete and Strength Development

A comprehensive understanding of concrete's basic composition and the intricate chemical process of hydration is essential for developing effective predictive models. This section delves into these fundamentals, followed by a detailed discussion of the critical factors that govern its compressive strength.

2.2.1 Concrete Composition and the Hydration Process

At its most fundamental level, concrete is a composite material comprising a paste and aggregates, which include both fine aggregate sand and coarse aggregate rock. The paste, formed by the chemical reaction between cement and water, envelops the surfaces of these aggregates, binding them together into a hardened, rock-like mass (National Ready Mixed Concrete Association, 2021).

The cornerstone of concrete's development is the hydration process, a nonreversible chemical reaction where cement reacts with water to form a strong, stiff, and impermeable hydrated cement paste. This process is responsible for concrete's setting and hardening, as it leads to the formation of microscopic crystals that interlock and bind all the components together, ultimately contributing to its strength gain (Reyes, 2025).

Supplementary Cementitious Materials (SCMs), such as fly ash, slag cement (Ground Granulated Blast Furnace Slag), and silica fume, are increasingly incorporated into concrete mixes. These materials enhance concrete properties through various chemical reactions: hydraulic reactions and/or pozzolanic reactions (with calcium hydroxide, a byproduct of cement hydration) (UltraTech, 2021).

- **Fly Ash** is the most widely utilized SCM, often substituting 15-40% of the total cementitious material. While it may delay setting time and decrease early strengths, it generally leads to increased long-term strengths due to its slower but more prolonged reaction rates (National Concrete Pavement Technology Center, n.d.).
- **Ground Granulated Blast Furnace Slag (GGBFS)**, a byproduct of iron production, is known to improve concrete's mechanical characteristics and enhance its resilience to weak acids and salts. Research indicates that optimal

concrete strength can be achieved when GGBFS substitutes approximately 10% of the cement content (UltraTech, 2021).

- **Silica Fume**, a byproduct from the production of silicon metal and ferrosilicon alloys, is a highly reactive pozzolan. Its incorporation significantly enhances concrete's durability and strength, notably by reducing its permeability (UltraTech, 2021).

The hydration process is a complex, time-dependent chemical reaction influenced by multiple factors, with the presence and type of SCMs adding significant layers of complexity. The non-linear nature of these interactions, for instance, the delayed early strength development observed with SCMs, contrasted with their contribution to increased long-term strength which makes traditional empirical modeling particularly challenging. This inherent complexity underscores why data-driven machine learning approaches are more suitable for accurately capturing these intricate and dynamic relationships. The process of crystal formation and intermeshing is not a simple additive function of ingredients but rather it is a dynamic system where the rate and extent of reactions are influenced by temperature, moisture, and the chemical composition of all cementitious materials. This fundamental understanding of concrete's behavior highlights the need for advanced modeling techniques that can learn from and predict outcomes in such a complex system.

2.2.2 Critical Factors Influencing Concrete Compressive Strength

Concrete strength is a multifaceted property, shaped by a confluence of interacting factors. These include the quality of raw materials, the water-cement ratio, the properties of aggregates, the age of the concrete, the degree of compaction achieved, and the prevailing temperature and humidity conditions during curing (Alem, 2022).

2.2.3 Water-Cement Ratio and its Impact

The water-cement (W/C) ratio stands as a paramount determinant of concrete strength where a lower ratio generally correlates with stronger concrete. This relationship holds because a reduced W/C ratio leads to a denser cement paste with fewer voids, thereby directly contributing to higher strength. The W/C ratio fundamentally dictates the porosity of the hardened cement paste at any given stage of hydration (Li, 2021).

However, this relationship is not without its complexities. If the W/C ratio falls too low, the workability of the fresh concrete is significantly diminished, making it difficult to properly compact. Insufficient compaction can lead to the entrapment of air, which, despite a theoretically stronger paste, ultimately reduces the actual strength of the concrete (Li, 2021). This presents a classic engineering optimization challenge: while minimizing water is desirable for strength, adequate water is necessary to ensure workability and proper compaction. This trade-off, where an optimal range exists rather than a simple linear correlation, demonstrates why machine learning approaches are particularly adept at identifying such non-linear optima. Traditional linear models would struggle to capture this nuanced relationship where both extremes (too high or too low W/C) can be detrimental to final strength.

2.2.4 Role of Aggregates

Aggregates constitute a substantial portion of concrete, typically occupying 60-75% of its total volume (Mohammed & Al-Mashhadi, 2020). While the intrinsic strength of the aggregate itself is generally not the primary limiting factor for concrete strength (as aggregates are often stronger than the surrounding cement paste), other characteristics significantly influence the final product. These include the aggregate's size, shape, surface texture, grading, and mineralogical composition (Mohammed & Al-Mashhadi, 2020).

- **Size:** Larger aggregate particles may reduce the mixing water required for a given consistency. However, they can also lead to weaker transition zones within the concrete, which are prone to forming more micro-cracks around the larger pieces, potentially lowering overall strength. For high-strength concrete applications, coarse aggregate size is commonly restricted to a maximum of 19 mm (ASTM, 1991).
- **Shape and Texture:** Cubical or spherical aggregate shapes are considered ideal for maximizing concrete strength. Rough textured or crushed aggregates can contribute to higher early strength due to an improved physical bond with the hydrated cement paste. Conversely, flat and elongated particles should be minimized due to their detrimental effects on workability and strength (Qureshi et al., 2015).

- **Grading:** The proper grading of aggregates, referring to the distribution of particle sizes, is crucial for producing workable and economical concrete mixtures. Aggregates that lack a balanced distribution of sizes can lead to harsh or uneconomical mixes.

The influence of aggregates on concrete strength is often subtle and indirect. Their impact is not typically a direct contribution of strength but rather an effect on the efficiency of the hydration process, the formation of micro-cracks, or the overall density of the cement paste. This complex and indirect influence is challenging to quantify accurately using simple empirical rules, but it can be effectively learned by machine learning models from large datasets that capture these intricate relationships. This highlights that concrete strength is a system-level property, where the interactions between components are as important as the individual properties of the components themselves.

2.2.5 Influence of Admixtures and Supplementary Cementitious Materials (SCMs)

Admixtures, whether in liquid or powder form, are intentionally added to concrete mixes to modify or enhance both their fresh and hardened properties. These improvements can span workability, strength development, surface finish, and long-term durability (UltraTech, 2021).

- **Chemical Admixtures:** These primarily function to improve the workability of fresh concrete. For instance, plasticizers and superplasticizers reduce the water demand necessary to achieve a desired slump, thereby enabling lower water-cement ratios and consequently higher strength concrete. Beyond workability, chemical admixtures can also influence the rate of cement hydration (either accelerating or decelerating it) and improve particle dispersion within the mix (UltraTech, 2021).
- **Mineral Admixtures (SCMs):** As discussed previously, SCMs like fly ash, silica fume, and ground granulated blast-furnace slag (GGBS) are incorporated to enhance strength and durability while simultaneously reducing the environmental footprint of concrete production by partially replacing cement

(UltraTech, 2021). Their contribution to strength primarily occurs through pozzolanic or hydraulic reactions, which form additional cementitious compounds.

Admixtures introduce additional layers of complexity into concrete mix design. Their effects are often highly dependent on dosage and can interact with other mix components and curing conditions in non-linear ways. For example, while superplasticizers generally decrease water demand for higher strength, some studies on high-strength concrete (HSC) have found them to be the least sensitive input parameter influencing compressive strength, despite their overall positive impact (Qi et al., 2022). This context-dependent behavior underscores the need for machine learning's ability to model multi-variable interactions and discover nuanced, context-specific relationships that simple empirical rules might overlook. Such capabilities enable the optimization of admixture use for specific performance targets.

2.2.6 Effects of Curing Conditions, Temperature, Humidity, and Age

Curing is a critical post-placement process that is absolutely essential for proper cement hydration and the subsequent development of concrete strength. It involves maintaining adequate moisture levels (e.g., through moist curing, pond curing, or the application of curing compounds) and controlling temperature conditions for a specified period after concrete placement (Auwal Ibrahim et al., 2024). Premature drying, if allowed, can halt the hydration process, significantly reducing the concrete's potential strength.

The age of concrete is directly correlated with its strength development. Concrete continuously gains strength over time as hydration progresses, although the rate of gain diminishes over longer periods. While 28 days is a widely accepted benchmark for achieving design strength, concrete can continue to gain strength for many years (De De Jesus et al., 2025). Early age strength, such as at 7 days, typically reaches 60-75% of the 28-day strength.

Temperature plays a pivotal role in the curing process. Hotter temperatures can accelerate early hydration and strength gain but may lead to lower long-term strength if not properly managed. This is due to rapid moisture loss and the formation of less well-structured bonds prematurely (Zeyad et al., 2022). Conversely, colder temperatures significantly slow down the hydration process, delaying strength development (Wang et al., 2025).

Humidity is equally important, as high humidity levels (around 85%) are crucial for ensuring a continuous supply of moisture necessary for hydration (The Concrete Countertop Institute, 2021). Low humidity, particularly when combined with high wind speeds, drastically increases evaporation rates, which can adversely affect concrete properties and lead to issues like plastic shrinkage cracking (Meyer et al., 2022).

The dynamic interplay between age, temperature, and humidity introduces a complex temporal dimension to concrete strength development. Accurately predicting strength therefore requires accounting for the entire curing history, not merely the initial mix proportions. This complex interaction highlights why curing time or age is consistently identified as a critical input feature in machine learning models across numerous studies (Thapa, 2024). This reinforces the need for ML models that can incorporate time-series data or age as a critical input, moving beyond static mix design parameters to capture the full-strength development profile.

2.2.7 Importance of Compaction

Compaction is a critical step in concrete placement, serving the primary purpose of eliminating entrapped air voids from the fresh concrete mixture (ACI Committee, 2000). The presence of these voids, even in small percentages, can significantly compromise the concrete's strength. For instance, concrete with 5-10% voided space due to poor compaction can experience a substantial reduction in strength, ranging from 30% to 40% (Alem, 2022).

Vibration is a highly effective method for achieving proper compaction, and studies have shown that even short periods of vibration (e.g., 15 seconds) can remove a significant amount of entrapped air (e.g., 3%) (Tuncan et al., 2007).

Compaction, a physical process, introduces a critical source of variability that can profoundly impact the actual in-situ strength of concrete, even when the mix design is theoretically perfect. This highlights a notable gap between the strength predicted by laboratory mix designs and the real-world performance of concrete in a structure. Machine learning models, if provided with relevant field data (e.g., data from NDT methods that are sensitive to compaction levels), could potentially bridge this gap. This suggests an opportunity for ML models to integrate field-level quality control parameters, such as the quality of compaction, into their strength predictions, thereby moving beyond predictions based solely on laboratory mix designs.

2.3 Traditional Methods for Concrete Strength Testing and Quality Control: Capabilities and Limitations

This section provides a review of established methods for assessing concrete strength and ensuring quality, highlighting their procedures, applications, and inherent limitations that underscore the growing motivation for adopting machine learning approaches.

2.3.1 Destructive Testing Methods

Destructive testing methods provide direct and definitive measurements of concrete strength, serving as the "ground truth" for structural design and quality verification.

- **Compressive Strength Testing:** The most prevalent method involves subjecting cylindrical concrete specimens (commonly 6x12-inch or 4x8-inch) to failure in a compression-testing machine (Hearns, 2022). The compressive strength is calculated by dividing the maximum load at failure by the cross-sectional area of the specimen, typically reported in pounds per square inch (psi) or megapascals (MPa) (ASTM et al., 2003).
- **Tensile Strength Testing:** Concrete inherently possesses low tensile strength. Various methods are employed to assess this property, including direct tension tests, flexure tests, and splitting tensile tests.
- **Splitting Tensile Strength (Indirect or Brazilian Test):** This method involves placing a cylindrical concrete specimen horizontally between two parallel platens and applying a compressive force perpendicular to its longitudinal axis until it fractures. This induces tensile stresses that cause the specimen to split. This test is particularly important in the design of structural lightweight concrete members (Daneshvar et al., 2022).
- **Flexural Strength Testing (Modulus of Rupture - MR):** This test measures the ability of an unreinforced concrete beam or slab to resist failure under bending loads. It is performed by loading 6x6-inch (150x150 mm) concrete

beams with a span length at least three times their depth (National Ready Mixed Concrete Association, n.d.).

The invasiveness, time-consuming nature (e.g., waiting 28 days for full strength development), and resource intensity of destructive tests inherently limit their utility for real-time quality control and continuous monitoring in dynamic construction environments. While these tests provide direct and reliable strength measurements, which are critical for design, the delay in obtaining results creates a fundamental need for predictive alternatives that can provide timely strength data without destroying the actual structure or requiring prolonged curing periods.

2.3.2 Non-Destructive Testing (NDT) Methods

Non-destructive testing (NDT) methods offer a valuable alternative to destructive tests by assessing the quality and integrity of concrete structures without causing physical damage. These methods are widely employed for fault investigation, verifying construction quality, and informing decisions regarding repair or demolition (Jedidi et al., 2014).

- **Rebound Hammer Testing (Schmidt Hammer Test):** This method estimates the surface hardness of concrete, which can be correlated to its compressive strength. It involves pressing a spring-loaded hammer against the concrete surface and measuring the distance it rebounds. However, this test primarily assesses surface properties and may not accurately reflect deeper structural issues. Its results can also be influenced by factors such as moisture content, surface roughness, and carbonation. Consequently, it is often utilized as an initial screening method rather than a definitive assessment tool and is generally more reliable for older concrete (Jedidi et al., 2014).
- **Ultrasonic Pulse Velocity (UPV) Testing:** UPV is a more advanced NDT method that evaluates the internal quality of concrete by measuring the speed at which ultrasonic waves travel through the material. Slower pulse velocities may indicate the presence of cracks, voids, or areas of poor compaction within the concrete (Jedidi et al., 2014). This method assesses the entire cross-section between transducers, making it valuable for detecting internal flaws not visible on the surface. It is commonly applied to assess bridge decks, tunnel

linings, and structural slabs, particularly after events like flooding or fire exposure (Format NDT Ltd, 2025). The dynamic modulus of elasticity, derived from UPV, is observed to increase with curing time and decrease with a higher water-cement ratio (Jedidi et al., 2014).

- **Cover Meter and Rebar Location:** These devices utilize electromagnetic fields to detect embedded steel reinforcement (rebar) within concrete and measure the depth of the concrete cover above it. Knowing the precise location, spacing, and depth of rebar is crucial for confirming compliance with design specifications, planning core drilling or retrofit installations, and assessing the risk of corrosion (Format NDT Ltd, 2025).
- **Half-Cell Potential Testing:** This technique assesses the likelihood of corrosion activity in embedded steel reinforcement by measuring the electrical potential at various points across the concrete surface relative to a reference electrode connected to the rebar. Lower potential readings typically indicate active corrosion (Format NDT Ltd, 2025).
- **Ground Penetrating Radar (GPR):** GPR employs electromagnetic waves to generate subsurface profiles of concrete. It can detect voids, locate rebar, assess slab thickness, and identify buried utilities, providing real-time results with minimal surface preparation (Format NDT Ltd, 2025).

While NDT methods offer non-invasive and generally faster assessments compared to destructive tests, they often provide indirect measures of strength (e.g., surface hardness, pulse velocity) that require empirical correlation to actual strength. This introduces a degree of uncertainty and can limit their definitive use for strength acceptance. This inherent limitation creates a clear opportunity for machine learning to integrate multiple NDT inputs, potentially providing more robust and accurate strength predictions from non-invasive data. By learning the complex, non-linear relationships between various NDT parameters and actual concrete strength, ML could enhance the reliability and utility of these methods for comprehensive quality control.

2.3.3 Inherent Challenges and Limitations of Traditional Approaches for Real-time Quality Control

The cumulative limitations of traditional concrete strength assessment methods pose significant challenges for modern construction, where efficiency, cost-effectiveness, and real-time decision-making are paramount.

- **Time and Cost Constraints:** Traditional laboratory experiments, such as the standard 28-day compressive strength tests, are inherently both costly and time-consuming (Qi et al., 2022). These methods introduce considerable delays in construction workflows, particularly when early strength data is critical for making timely decisions related to formwork removal, load application, or scheduling subsequent construction phases (Marchewka et al., 2025).
- **Lack of Real-time Feedback:** A major drawback of traditional methods is their inability to provide continuous, in-situ monitoring and real-time feedback on concrete strength progression. They offer only periodic and delayed results, which is insufficient for the dynamic demands of modern construction practices (Liu, 2024).
- **Inability to Reflect Field Conditions:** Laboratory test specimens are produced under highly controlled conditions, which often fail to accurately reflect the actual strength development of concrete structures in the field. Real-world structures are exposed to a wide range of varying environmental factors, curing conditions, and temperature distributions caused by the heat of hydration, all of which directly impact strength development and can lead to discrepancies between theoretical predictions and actual performance (Ryu et al., 2024).
- **Complexity of Variable Interactions:** Traditional empirical approaches and simplified mathematical models struggle to capture the complex, non-linear interactions among the numerous variables influencing concrete compressive strength. These variables include cement content, water content, supplementary cementitious materials (SCMs), and curing age (Thapa, 2024). Manually calculating and accounting for the effects of these intricate interactions is exceedingly time-consuming and complex (Qi et al., 2022).
- **Dataset Size Limitations for Empirical Models:** Empirical equations often rely on a limited number of independent variables, which inherently restricts

their accuracy and generalizability when applied to diverse concrete mixes and conditions (Thapa, 2024).

This comprehensive set of limitations collectively serves as the primary justification for focusing on machine learning in concrete strength prediction. These challenges demonstrate that the current state-of-the-art in traditional concrete strength assessment is inadequate for meeting the demands of contemporary construction, thereby establishing a clear problem that machine learning can effectively address. Machine learning is not merely an incremental improvement but a necessary innovation to enhance efficiency, reduce costs, and improve safety in the construction industry.

2.4 Machine Learning Principles for Regression Analysis

This section establishes the theoretical foundation for the machine learning component of the project. It elucidates the fundamental concepts of regression, outlines essential data preprocessing steps, introduces key algorithms, and details the metrics used for performance evaluation.

2.4.1 Introduction to Supervised Learning and Regression

Machine learning encompasses various paradigms for training algorithms, among which supervised learning is prominent. In this paradigm, a model learns from a labeled dataset, meaning that during the training phase, both the input features (independent variables) and their corresponding correct output values (dependent variables) are explicitly provided to the algorithm. The model then identifies patterns and relationships within this labeled data to make predictions on new, unseen data.

Regression is a specific type of supervised learning technique. Its primary purpose is to model the relationship between input features and a *continuous* output variable. Unlike classification algorithms, which predict discrete categories (e.g., "strong" or "weak" concrete), regression models aim to predict numerical values, such as concrete strength, house prices, or stock market trends. The objective of a regression model is to uncover the underlying relationships between variables, thereby enabling accurate forecasting and estimation for new data points.

The selection of regression, as opposed to classification, is fundamental for concrete strength analysis because concrete compressive strength is a continuous numerical property, typically measured in units like megapascals (MPa). The problem is not merely about categorizing concrete as "strong" or "weak," but rather about predicting its precise strength value, which is crucial for detailed engineering design and rigorous quality control. This choice clarifies the mathematical nature of the problem and justifies the exclusive selection of regression algorithms, emphasizing the need for models that can output continuous numerical predictions to meet the specific demands of civil engineering applications.

2.4.2 Essential Data Preprocessing Techniques for Machine Learning Datasets

Raw data, regardless of its source, is seldom in a format immediately suitable for direct analysis by machine learning algorithms. It often contains inconsistencies, missing values, erroneous inputs, or formats that are not conducive to effective modeling. Consequently, data preprocessing is an indispensable phase to transform raw data into a clean, consistent, and usable format, which is critical for ensuring data quality and enhancing the accuracy of model results (Sah & Hong, 2024).

2.4.2.1 Data Cleaning

Data cleaning is a foundational component of the data preprocessing pipeline, focused on identifying and rectifying errors or inconsistencies to ensure high-quality data for analysis and model training (Content Studio, 2024).

- **Handling Missing Values:** Missing values (often represented as NaN or NULL) are a common occurrence in real-world datasets and can adversely affect model performance. Identification typically involves descriptive statistics or visualizations. Strategies for handling them include removal of rows or columns with missing data (if the percentage is small and non-critical) or, more commonly, imputation. Imputation involves filling missing values using statistical measures like the mean, median, or mode for numerical features, or the most frequent category for categorical features. More advanced imputation methods include regression imputation or k-nearest neighbors' imputation (Sah & Hong, 2024).
- **Handling Outliers:** Outliers are data points that deviate significantly from other observations in the dataset. They can disproportionately influence statistical analyses and machine learning models, leading to skewed results. Detection often involves visual methods such as box plots, histograms, or scatter plots, or statistical methods like the Interquartile Range (IQR) rule. Handling strategies include removal (if due to data entry errors), data transformation (e.g., log or square root transformations to normalize distributions), or employing machine learning models that are inherently more robust to outliers, such as Support Vector Machines (SVM) or Random Forests (Content Studio, 2024).

- **Handling Duplicates:** Duplicate records, if not addressed, can inflate certain patterns or biases in the data, leading to inaccurate model training. Identification typically involves checking for identical rows. If these duplicates are redundant and provide no additional information, they are removed to ensure data quality (Content Studio, 2024).

Data cleaning is not merely a technical step but a crucial foundation for model reliability. Uncleaned data can introduce biases, reduce accuracy, and lead to misleading conclusions, thereby undermining the entire predictive effort. This process adheres to the "garbage in, garbage out" principle in machine learning, emphasizing that the quality of the input data directly dictates the quality and trustworthiness of the model's output, particularly in safety-critical applications like civil engineering.

2.4.2.2 Feature Engineering and Data Transformation

Feature engineering is the art and science of creating new, more informative features from existing raw data using domain knowledge. This process can significantly enhance the performance of machine learning models by uncovering non-linearities, complex interactions, and hidden patterns that might not be apparent in the original features (Tak et al., 2025). For example, in concrete strength prediction, while individual quantities of water and cement are input features, calculating the water-cement ratio as a new feature is a powerful application of domain expertise. The W/C ratio is a well-established critical factor influencing concrete strength. By explicitly providing this ratio to the model, it can "see" and leverage this fundamental relationship more effectively than if it had to infer it solely from raw component quantities, thereby improving predictive power and interpretability.

Data transformation is a broader process of converting data into a suitable format for analysis and downstream processes. This can involve converting categorical text attributes into numerical representations (e.g., one-hot encoding), or reducing continuous data into a smaller number of categorical intervals or bins (discretization) to simplify the data and reduce noise (Saiwa, 2023). Feature engineering, therefore, represents a direct interface where civil engineering knowledge of concrete properties directly enhances machine learning model development, making the models more effective and their predictions more aligned with physical principles.

2.4.2.3 Data Scaling and Normalization

Data scaling (also known as feature scaling or normalization) involves rescaling feature values within a dataset to a similar range without distorting their original differences. Common techniques include Min-max scaling, which transforms values to a range between 0 and 1, and Z-score normalization (standardization), which transforms values to have a mean of 0 and a standard deviation of 1 (Thapa, 2024). This step is crucial for many machine learning algorithms, particularly those that rely on distance calculations or gradient descent optimization (e.g., neural networks, Support Vector Machines, linear regression, logistic regression). Without proper scaling, features with larger numerical ranges (e.g., coarse aggregate weight in kg/m^3) could disproportionately influence the model's learning process and parameter updates compared to features with smaller ranges (e.g., superplasticizer dosage in kg/m^3). This can lead to suboptimal model performance, slower convergence during training, and biased learning (Thapa, 2024). Data scaling ensures the fairness of feature contribution during model training, allowing all features to contribute equally to the model's learning process. This highlights a common pitfall in machine learning implementation and underscores the importance of meticulous data preparation for building robust and accurate predictive models.

2.4.2.4 Data Splitting

Data splitting is a fundamental practice in machine learning that involves dividing the dataset into independent subsets to ensure correct model evaluation and prevent overfitting (Alooba, n.d.). This process is critical for assessing a model's generalizability, its ability to perform well on new, unseen data, which is the ultimate goal in real-world prediction.

- **Training Set:** This subset of data is used to train the machine learning model. The model learns patterns, relationships, and structures from this dataset. Typically, around 80% of the available data is allocated for training (Mistry, 2025).
- **Validation Set:** This set is used during the model development phase for tuning hyperparameters and assessing intermediate model performance. It helps in

selecting the best model configuration and preventing overfitting to the training data. Generally, 10-20% of the data is reserved for validation (Mistry, 2025).

- **Test Set:** The test set is reserved exclusively for evaluating the final model's performance on completely unseen data. It provides an unbiased estimate of the model's generalization capability. Typically, 10-20% of the data is allocated for testing (Mistry, 2025).

A good strategy for data splitting is cross-validation, such as K-Fold Cross-Validation (e.g., 5-fold). This method divides the dataset into K subsets (folds) and trains the model K times, using a different fold as the validation set in each iteration. This approach provides a more reliable and less biased estimate of model performance compared to a single train-test split (Chugani, 2024).

Proper data splitting is paramount for assessing a model's generalizability. Without a clear separation of training, validation, and test sets, models can appear highly accurate on the data they were trained on but fail dramatically in practical applications due to **overfitting** (where the model memorizes the training data, including noise, rather than learning underlying patterns). This rigorous methodology is essential for building truly reliable predictive models, ensuring that the reported performance accurately reflects real-world applicability and trustworthiness.

2.4.3 Key Regression Algorithms Applied in Concrete Strength Prediction

Regression algorithms form the bedrock of machine learning for predicting continuous values, effectively uncovering relationships between input features and a target variable (GeeksforGeeks, 2025). In the context of concrete strength prediction, several algorithms have demonstrated varying degrees of success and utility.

2.4.3.1 Linear Regression and its Regularized Variants

Linear Regression is a fundamental algorithm that assumes a linear relationship between the independent input variables (x) and the continuous output variable (y). Its objective is to fit a straight line (or hyperplane in higher dimensions) that minimizes the error (residuals) between predicted and actual values (Aakash, 2024). It is valued for its simplicity and interpretability, often serving as a baseline model for comparison. However, it is sensitive to outliers and, crucially, its assumption of linearity limits its effectiveness when the underlying relationship is non-linear. Given the inherently complex and non-linear nature of concrete strength development (due to intricate hydration processes, aggregate interactions, and admixture effects), simple linear models often prove insufficient.

To address issues like overfitting and multicollinearity (where input variables are highly correlated), linear regression has **regularized variants**:

- **Ridge Regression (L2 Regularization):** This method adds a penalty term to the cost function that discourages large coefficients. This helps in reducing overfitting and performs well when features are correlated (Aakash, 2024).
- **Lasso Regression (L1 Regularization):** Lasso adds an L1 penalty to the cost function, which encourages sparsity in the model by driving some coefficients precisely to zero. This effectively performs feature selection by eliminating irrelevant features and reduces model complexity (Aakash, 2024).
- **Elastic Net:** This combines both L1 and L2 penalties, balancing the strengths of Ridge and Lasso regression. It is particularly useful when features are highly correlated (Aakash, 2024).

While regularization techniques can mitigate overfitting in linear models, they do not fundamentally resolve the assumption of linearity. This limitation in capturing the inherently non-linear relationships in concrete strength highlights the necessity for more advanced, non-linear machine learning algorithms in this domain.

2.4.3.2 Support Vector Regression (SVR)

Support Vector Regression (SVR) is a robust machine learning method designed for forecasting continuous results, drawing its principles from Support Vector Machines (SVM) which are typically used for classification (GeeksforGeeks, 2024). The core objective of SVR is to find an optimal hyperplane in a high-dimensional feature space that accurately represents the data. Unlike classification SVMs that aim to separate classes, SVR seeks to find a hyperplane that best fits the data points in a continuous space, maximizing the margin (distance) between the hyperplane and the closest data points, while simultaneously minimizing the prediction error (GeeksforGeeks, 2024).

SVR employs kernel functions (such as Radial Basis Function (RBF), polynomial, and sigmoid) to effectively handle complex, non-linear relationships within the data without explicitly transforming the data into higher dimensions, thereby avoiding increased computational cost (Pal et al., 2024). This makes SVR particularly well-suited for noisy and complex concrete datasets, and it can be less sensitive to outliers compared to some other regression models. However, a notable aspect of SVR, like many complex ML models, is its "black-box" nature. The difficulty in directly interpreting the hyperplane and the precise contribution of individual features can be a limitation for engineers who require transparent insights into the decision-making process for trust and actionable design changes. This represents a common trade-off between model accuracy and interpretability in machine learning.

2.4.3.3 Decision Tree Regression

Decision Tree Regression is a non-parametric machine learning technique that predicts numerical values using a hierarchical, tree-like structure. The algorithm operates by recursively dividing the dataset into smaller subsets based on features that best reduce prediction error, typically measured by Mean Squared Error (MSE) (Baladram, 2024). Each internal node in the tree represents a decision point based on a specific feature, branching out to child nodes until it reaches a "leaf node," which contains the final predicted numerical value (often the average of the target values within that subset of data).

Decision trees are highly valued for their interpretability. They are considered "white box" models because the decision-making process can be easily followed and closer to human language (Loyola-González, 2019). They can handle both numerical and categorical data, require relatively little data preparation, and can perform well with large datasets. To prevent overfitting, a common issue where the tree becomes too complex and memorizes the training data, pruning techniques are employed. These include pre-pruning (halting tree growth early based on criteria like maximum depth or minimum samples per split) and post-pruning (allowing the tree to grow fully and then trimming back branches, such as through Cost-Complexity Pruning) (Jain, 2024).

While individual decision trees offer high interpretability, which is a significant advantage in engineering where understanding *why* a prediction is made is crucial for trust and actionable insights, they can be prone to overfitting and may exhibit lower generalization performance when applied to new, unseen data. This limitation serves as a primary motivation for the development and use of ensemble methods like Random Forest, which build upon the strengths of decision trees while mitigating their weaknesses.

2.4.3.4 Artificial Neural Networks (ANN) for Regression

Artificial Neural Networks (ANNs), also known as simulated neural networks (SNNs), are computational models inspired by the structure and function of biological neural networks. They are designed to predict a numerical output variable from a set of input features (IBM, 2021).

An ANN is composed of interconnected "artificial neurons" or nodes, typically organized into multiple layers: an input layer (receiving external data), one or more hidden layers (processing information), and an output layer (producing the final result) (IBM, 2021). Each neuron processes weighted inputs through a non-linear activation function to produce an output, which then serves as input for subsequent neurons. The strength of connections between neurons is determined by "weights," which are adjusted during the learning process.

ANNs are highly capable of learning and modeling complex, non-linear relationships within data. This capability comes from the use of non-linear activation functions and the hierarchical, interconnected structure of their layers, which allows them to capture patterns that simpler linear models cannot capture.

ANNs are trained through an iterative process of empirical risk minimization. During training, they learn from labeled data by adjusting their parameters (weights and biases) to minimize a defined loss function (e.g., Sum of Squared Errors - SSE for regression). This optimization is commonly achieved using gradient-based methods like backpropagation, which calculates the gradient of the loss function with respect to the weights and uses it to update the weights in the direction that minimizes error (GeeksforGeeks, 2025). The performance of an ANN is significantly influenced by its hyperparameters, which are parameters set before the training process begins.

While ANNs excel at capturing highly complex, non-linear relationships, making them powerful tools for concrete strength prediction, their "black-box" nature remains a significant limitation for practical adoption in safety-critical civil engineering applications. The difficulty in understanding the internal decision processes of ANNs means engineers may struggle to trust or explain *why* a particular prediction was made.

2.4.3.5 Ensemble Learning Methods: Random Forest and Gradient Boosting

Ensemble learning is a powerful machine learning paradigm that combines the predictions of multiple individual "base learners" (often decision trees) to produce a more accurate, robust, and generalizable prediction than any single model could achieve alone (Murel & Kavlakoglu, 2025). This approach effectively addresses the bias-variance trade-off, leading to a reduced overall error rate. Ensemble methods are broadly categorized into parallel methods (like bagging, where base learners are trained independently) and sequential methods (like boosting, where new learners correct errors of previous ones).

- **Random Forest (RF):** Random Forest as shown in Figure 2.1, proposed by Leo Breiman and Adele Cutler (Qi et al., 2022). It constructs multiple decision trees independently. Each tree is trained on a different random subset of the original data (sampled with replacement, known as bootstrapping) and, crucially, at each split within a tree, it considers only a random subset of the available features. For regression tasks, the final prediction is obtained by averaging the predictions from all individual trees. This dual source of randomness helps to decorrelate the individual trees, reducing the overall variance of the ensemble. RF models are highly robust to overfitting, can efficiently handle missing data, are relatively easy to implement and tune, and are highly scalable, making them suitable for large datasets.

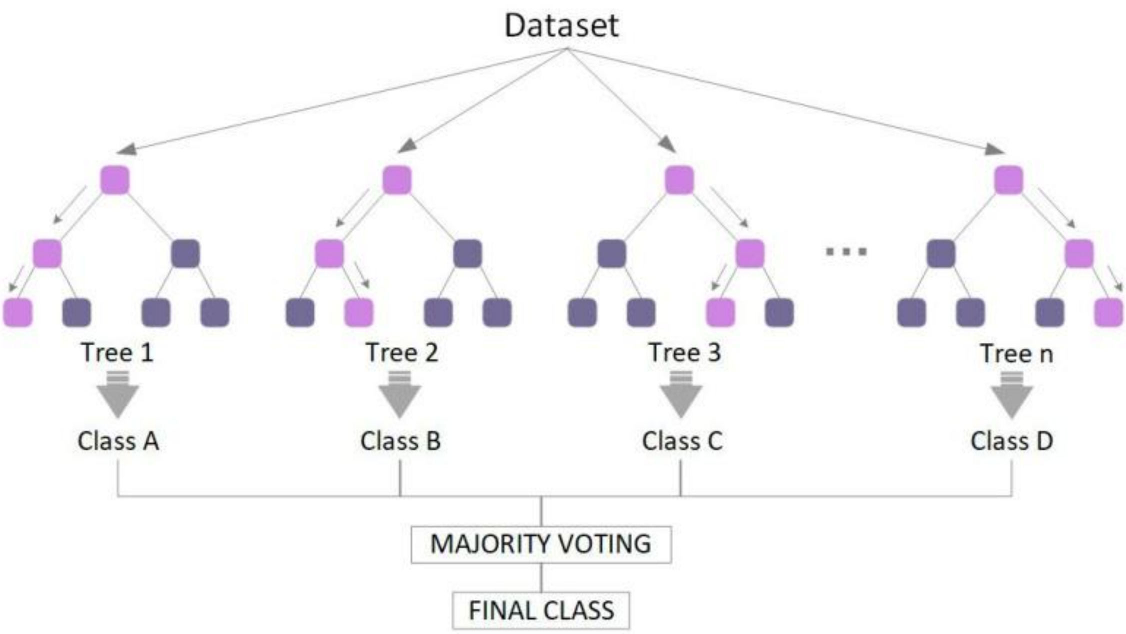

Figure 2.0.1

- **Gradient Boosting (GB):** Gradient Boosting (Figure 2.2) is a sequential ensemble method. Instead of building independent trees, it constructs decision trees one after another, with each new tree specifically designed to correct the *residual errors* (the differences between current predictions and actual values) of the preceding trees. The process iteratively optimizes a defined loss function through a gradient descent approach.

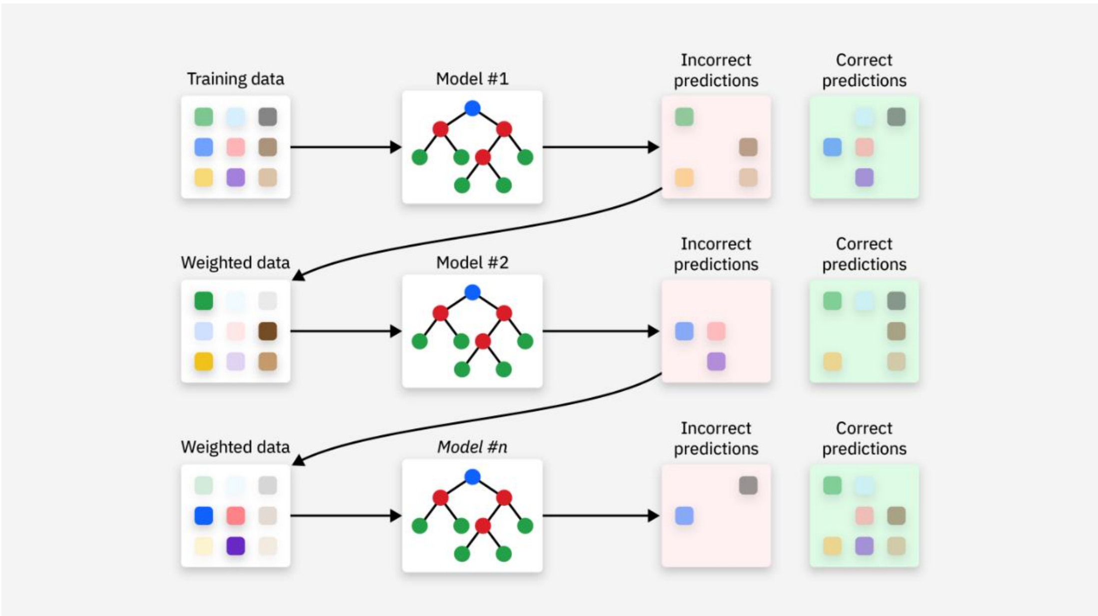

Figure 2.2

- **XGBoost (eXtreme Gradient Boosting):** This is a highly optimized and widely used implementation of gradient boosting. It is renowned for its computational efficiency, accuracy, and scalability, incorporating features like regularization (L1/L2) to prevent overfitting and supporting parallel processing. XGBoost is particularly effective for large datasets and complex, non-linear relationships, making it a strong performer in concrete strength prediction.
- **CatBoost:** Another gradient boosting variant, notable for its robust handling of categorical features and its use of ordered boosting to mitigate target leakage, which can be beneficial for specific types of concrete datasets. Gradient Boosting models generally offer high predictive power and often achieve higher accuracy than Random Forests, especially when carefully tuned. They are flexible, supporting various loss functions, and can handle imbalanced datasets effectively.

Compared to Random Forest, Gradient Boosting can be more prone to overfitting if not properly tuned, may have longer training times due to its sequential nature, and is more sensitive to the precise tuning of its hyperparameters (GeeksforGeeks, 2024).

Ensemble methods represent the state-of-the-art in achieving high predictive accuracy for complex, non-linear problems such as concrete strength prediction. The distinction between bagging (RF) and boosting (GB, XGBoost) highlights different strategies for error reduction—RF primarily reduces variance, while GB focuses on reducing bias by iteratively correcting errors. The consistent superior performance of these models in numerous concrete prediction studies makes them prime candidates for advanced predictive modeling in this domain.

2.4.4 Performance Evaluation Metrics for Regression Models

Evaluating the performance of regression models is crucial for understanding how well they predict continuous outcomes. This involves quantifying accuracy by measuring the difference between predicted and actual values. No single metric is universally "best" where the optimal choice depends on the specific context and objectives of the analysis (e.g., robustness to outliers vs. penalizing large errors). A comprehensive evaluation typically requires considering multiple metrics to gain a holistic understanding of model performance and limitations.

- **R-squared (R^2 / Coefficient of Determination):** This metric quantifies the proportion of the variance in the dependent variable that can be predicted from the independent variables by the model. R-squared values range from 0 to 1, with higher values (closer to 1) indicating a better fit and more variance explained by the model. It is a scale-independent metric, making it useful for comparing models across different datasets or scales (Fernando, 2024).
- **Mean Absolute Error (MAE):** MAE measures the average magnitude of the errors between predicted and actual values, without considering their direction. MAE is less sensitive to outliers compared to MSE or RMSE and offers intuitive interpretability as it represents the average error in the same units as the target variable (Sabbha, 2024).
- **Mean Squared Error (MSE):** MSE calculates the average of the squares of the errors, representing the average squared difference between estimated and actual values. Due to the squaring of errors, MSE penalizes larger errors more heavily (Sabbha, 2024).
- **Root Mean Squared Error (RMSE):** RMSE is the square root of MSE. This transformation brings the error scale back to the same units as the target variable, making it more interpretable than MSE. Like MSE, RMSE is also sensitive to large errors (Sabbha, 2024).
- **Mean Absolute Percentage Error (MAPE):** MAPE expresses the error as a percentage of the actual value. This provides a relative and scale-independent measure of error, which is particularly valuable for comparing the performance of models across datasets with different scales or units (Roberts, 2023).

The selection of evaluation metrics is critical for a thorough assessment. Each metric offers a distinct perspective on model performance. For instance, MAE is preferred when robustness to outliers is important, while RMSE is chosen when larger errors need to be penalized more significantly. Therefore, a comprehensive evaluation necessitates considering multiple metrics to gain a holistic understanding of the model's predictive capabilities and limitations. This approach ensures that performance claims are well-supported and contextually appropriate for engineering applications.

2.5 Machine Learning Applications in Concrete Strength Prediction and Quality Control

This section details the application of machine learning to predict concrete strength, compares the performance of various ML models, discusses the datasets and input features commonly used, and explores the benefits and opportunities for real-time quality control and sustainable mix design.

2.5.1 Comparative Analysis of Machine Learning Models for Concrete Compressive Strength Prediction

Numerous studies have extensively evaluated the applicability and predictive performance of various machine learning regression models for forecasting concrete strength, often conducting direct comparisons between different algorithms. The models most commonly evaluated for this task include:

- Random Forest (RF)
- K-Nearest Neighbors (KNN)
- Support Vector Machine (SVM) / Support Vector Regression (SVR)
- Decision Tree (DT)
- Artificial Neural Networks (ANN)
- Gradient Boosting Regressor (GBR)
- Extreme Gradient Boosting (XGBoost)
- CatBoost
- Bagging
- Stacking
- Multi-Linear Regression (MLR)

General findings from this extensive research consistently indicate that ensemble methods (such as Random Forest, Gradient Boosting, XGBoost, and Stacking) frequently demonstrate superior predictive capabilities compared to individual, standalone models.

- **Random Forest (RF):** This model consistently achieves high R-squared values, and reported in the range of 0.91 (Thapa, 2024), 0.9726 (Shaqadan, 2016) and 0.96 (Farooq et al., 2020).
- **XGBoost:** This gradient boosting variant is frequently cited for its superior accuracy. Reported R-squared values for XGBoost models include 0.95 (Paudel et al., 2023), 0.954 (Qi et al., 2022) and 0.91 (Elhishi et al., 2023). It consistently demonstrates low error metrics across various datasets.
- **Artificial Neural Networks (ANN):** ANNs show strong agreement with experimental results, with R-squared values reaching as high as 0.9796 (Akande et al., 2014) .They have also been shown to outperform traditional regression analysis in concrete strength prediction.
- **Support Vector Machine (SVM) / SVR:** These models achieve good R-squared values, such as 0.84 (Thapa, 2024), and notably 0.992 (Akande et al., 2014). They are generally considered robust for this application.
- **Decision Tree (DT):** While generally performing lower than ensemble methods e.g., R-squared 0.78 (Thapa, 2024), Decision Trees are valued for their inherent interpretability.

The consistent outperformance of ensemble models, particularly XGBoost and Random Forest, across a multitude of studies points to a robust and emerging consensus on the most effective machine learning approaches for concrete strength prediction. This pattern of superior predictive accuracy and reliability suggests that any new project in this domain should prioritize these advanced techniques, as they are more likely to yield highly accurate and dependable results that meet the demands of practical engineering applications.

Table 1: Comparative Performance of Selected Machine Learning Models for Concrete Compressive Strength Prediction

Model Name	**R-squared (R^2)**	**MAE (MPa)**	**MSE (MPa^2)**	**RMSE (MPa)**	**MAPE (%)**	**Dataset Size (Samples)**	**Source Snippets**
Random Forest (RF)	0.91	2.66	14.7	3.83	9.14	776	(Thapa, 2024)

K-Nearest Neighbors (KNN)	0.88	2.91	20.2	4.5	9.74	776	(Thapa, 2024)
Support Vector Machine (SVM)	0.84	3.59	27.3	5.22	11.38	776	(Thapa, 2024)
Decision Tree (DT)	0.78	3.75	34.53	5.87	13.38	776	(Thapa, 2024)
Artificial Neural Network (ANN)	0.712	4.443	-	6.092	15.112	1030	(Sah & Hong, 2024)
Multiple Linear Regression (MLR)	-0.57	8.1824	-	10.45	32.51	1030	(Sah & Hong, 2024)
Gradient Boosting Regressor (GBR)	0.973	1.901	-	2.664	7.2	1030	(Mustapha et al., 2024)
XGBoost	0.93	2.31	11.29	3.36	7.44	776	(Thapa, 2024)

Note: Some metrics may not be available in all cited snippets, indicated by '-'.

This table provides a concise, quantitative overview of the performance of different ML models, allowing for direct comparison and highlighting the superior performance of ensemble methods. It visually reinforces the findings discussed in the text, making the academic rigor of the literature review immediately apparent. The aggregation of scattered performance metrics from various sources into a single, digestible format serves as a strong evidence base for claims about model superiority and helps justify the selection of specific models for the project.

2.5.2 Datasets and Input Features Utilized in ML-based Concrete Studies

Machine learning models developed for concrete strength prediction predominantly rely on experimental datasets. These datasets are typically compiled from past research papers or accessed through public repositories, such as the UCI Machine Learning Repository. The size of these datasets can vary significantly, ranging from hundreds to several thousands of entries.

A noteworthy observation across numerous studies is the remarkable consistency in the selection of independent input variables used to predict concrete compressive strength. These commonly utilized variables reflect a widely accepted understanding within civil engineering of the primary drivers of concrete strength are, cement, water, sand, coarse aggregate, fly ash, superplasticizer, curing time, blast furnace slag. The target variable in all these studies is consistently concrete compressive strength.

The remarkable consistency of input features across numerous studies indicates a widely accepted and robust understanding of the primary factors influencing concrete strength. This convergence of research validates the selection of these parameters as essential inputs for any new machine learning-based prediction project. However, the varying dataset sizes and sources across different studies also highlight practical challenges in data availability and the generalizability of models across diverse real-world conditions. While laboratory-collected data provides a controlled environment, real-world construction projects introduce variability in material quality, mixing, placement, and curing conditions that may not be fully captured in existing datasets. This implies that while these features are fundamental, the quality and representativeness of the dataset remain crucial for building truly robust and generalizable models.

2.5.3 Feature Importance Analysis and Model Interpretability

For machine learning models to gain widespread adoption in safety-critical fields like civil engineering, it is not enough for them to be accurate as understanding the influence of different input variables on predicted outcomes is equally crucial. This understanding makes ML models more transparent and explainable, fostering trust among engineers and facilitating actionable insights for design and quality control (Wiratsin & Ragkhitwetsagul, 2025).

The "black-box" nature of many powerful machine learning models, such as complex neural networks or ensemble methods, presents a significant barrier to their practical implementation in civil engineering. Engineers require not just a prediction, but an explanation of *why* a particular strength value was predicted, enabling them to make informed decisions about mix design adjustments or construction practices. This is where Explainable AI (XAI) techniques become invaluable.

- **SHAP (SHapley Additive exPlanations):** SHAP is a unified, game theory-based approach to explaining individual predictions by quantifying how much each feature contributes to the model's output. It provides a fair allocation of the contribution of each input feature to the overall outcome, offering both local (for individual predictions) and global (overall model behavior) interpretations (Wiratsin & Ragkhitwetsagul, 2025). SHAP values help engineers understand the impact of each feature in the model's decision-making process.
- **Partial Dependence Plots (PDPs):** PDPs illustrate the marginal effect of one or two features on the predicted outcome of a machine learning model. They show the relationship between a feature and the predicted outcome while holding other features constant, which is useful for understanding how a feature influences the model's prediction in isolation (Mistry, 2025).

Importance analysis, often conducted using SHAP or other sensitivity analyses, consistently identifies several key influential factors for concrete compressive strength:

- **Curing Age/Time:** This is almost universally identified as the most critical variable influencing concrete compressive strength. Strength typically increases sharply from day 0 to approximately 50 days, reflecting the critical curing phase (Tak et al., 2025).
- **Cement Content:** Consistently ranked as a highly important factor. Higher cement amounts generally have a positive effect on strength by enhancing cohesion and bonding.
- **Water Content / Water-Cement Ratio:** Water content is frequently identified as a sensitive parameter, sometimes even the most sensitive in high-strength concrete (HSC). The water-cement ratio is a crucial factor, with lower ratios

leading to increased strength and higher ratios reducing it due to increased porosity (Tak et al., 2025).

- **Superplasticizer Dosage:** This admixture is consistently found to be influential, with its contribution noted in various analyses.
- **Fly Ash (FA) and Blast Furnace Slag (BFS):** The influence of these supplementary cementitious materials varies, but they often contribute significantly to strength development.
- **Aggregates (Coarse and Fine):** While generally contributing less than cement, water, or age, aggregates remain influential factors.

The consistent identification of these key parameters by feature importance analysis, particularly through techniques like SHAP, provides valuable validation for the machine learning models. It demonstrates that the model's behavior aligns with established civil engineering domain knowledge, thereby enhancing trust among practitioners. This interpretability is vital for enabling informed decision-making in concrete mix design optimization and promoting sustainable construction practices.

3.0 Methodology

This section outlines the approach that will be taken to develop and evaluate machine learning models for predicting concrete compressive strength. The overall strategy is a correlational case study. It is a case study because it provides an in-depth investigation of a single entity which is the local partner company. It is correlational as its core aim is to determine the nature and strength of the relationship between a set of predictor variables (7-day strength, weather) and a target variable (28-day strength).

3.1 Data Acquisition and Sourcing

The foundation of any data-driven project is the quality and relevance of its data. This study employs a dual-source data acquisition strategy to construct a comprehensive dataset that accurately reflects the operational reality of the partner company.

3.1.1 Primary Data: Company Production Records

The primary dataset will consist of historical production and quality control records provided directly by the local concrete manufacturer. This data represents the "ground truth" of the company's output. Obtaining this data will be the first critical step, potentially requiring a formal data-sharing agreement or Non-Disclosure Agreement (NDA) to protect the company's proprietary information.

The key variables to be extracted from the company's records are:

- **7-Day Compressive Strength (MPa):** This is the principal independent variable. It represents an early indicator of the concrete's ultimate performance and is the most critical input for the predictive model.
- **28-Day Compressive Strength (MPa):** This is the dependent target variable. The entire project is oriented around predicting this value accurately. It serves as the benchmark for quality control and structural design compliance.
- **Casting Date (DD/MM/YYYY):** This seemingly simple piece of information is the key information of the entire study. It acts as the primary key that enables

the fusion of internal production data with external environmental data, thereby creating a richer, more context-aware dataset.

3.1.2 Secondary Data: Historical Environmental Records

Given that the company's raw materials (e.g., sand, aggregates) are stored in open yards, their physical properties are susceptible to environmental influence. For instance, high ambient temperatures can affect the rate of cement hydration, while heavy rainfall can alter the moisture content of the aggregates, which in turn influences the effective water-cement ratio of the mix. Therefore, incorporating weather data is not merely an academic exercise but a necessary step to account for a significant source of real-world variability.

Historical weather data corresponding to the casting dates of the concrete batches will be sourced from a reputable and official authority to ensure accuracy and reliability. The primary source will be the Malaysian Meteorological Department (MetMalaysia).

The following daily weather variables will be acquired for the specific geographic location of the company's production plant:

- **Daily Average Temperature (°C):** Influences the rate of hydration and workability.
- **Daily Minimum and Maximum Temperature (°C):** Provides insight into the temperature fluctuation throughout the day.
- **Daily Average Humidity (%):** Affects the rate of evaporation from the concrete surface and the moisture content of aggregates.
- **Total Daily Rainfall (mm):** A direct indicator of moisture being introduced to exposed material stockpiles.

3.2 Data Integration, Preprocessing, and Transformation

Raw data is rarely in a format suitable for direct use in machine learning models. This stage, which can be thought of as an Extract, Transform, Load process, is dedicated to cleaning, merging, and engineering the data to maximize its predictive power.

3.2.1 Data Merging and Integration

The two different datasets (company strength data and MetMalaysia weather data) will be integrated into a single, unified data frame. This will be accomplished by performing a "join" operation using the Date as the common key. The outcome will be a single table where each row represents a unique concrete batch, containing its 7-day and 28-day strength values alongside the full weather profile of the day it was cast. Potential challenges, such as inconsistencies in date formats (e.g., YYYY-MM-DD vs. DD/MM/YY), will be programmatically resolved.

3.2.2 Data Cleaning and Outlier Handling

The integrated dataset will be rigorously inspected for anomalies.

- **Missing Values:** Records with missing strength or weather values will be analysed. Depending on the extent of the missing data, strategies may include deletion (if only a very small percentage of records are affected) or imputation (e.g., filling a missing humidity value with the monthly average), with careful consideration not to introduce significant bias.
- **Outlier Detection:** Outliers, or data points that deviate significantly from the norm, can disproportionately affect model training. They will be identified using both statistical methods and visualization techniques. Any identified outliers will be investigated to determine if they are data entry errors (to be corrected or removed) or legitimate but extreme values (to be retained).

3.2.3 Feature Engineering

This is a creative and critical step where domain knowledge is used to construct new, more informative features from the existing data.

- **Lagged Weather Features:** The weather on the exact day of casting might not tell the whole story. The condition of the raw materials is a result of the weather over the preceding days. Therefore, new features will be created, such as the 3-day and 7-day rolling averages for temperature and humidity. For example, the 7_day_avg_temp will represent the average temperature of the week leading up to the casting date, providing a better proxy for the thermal state of the aggregate stockpiles.

- **Temporal Features:** To capture potential seasonal effects or operational cycles (e.g., differences between weekday and weekend production), temporal features will be extracted from the casting date, such as the month, quarter, and day of the week.
- **Categorical Features:** To simplify complex numerical data, a Had_Rain feature (1 for yes, 0 for no) will be created and can be derived from the daily rainfall column, which may prove to be a powerful, simple predictor.

3.3 Predictive Modelling Strategy

The core of the project is the development of predictive models. The choice of algorithms is informed by the literature review and the nature of the problem, which involves capturing potentially complex and non-linear relationships.

3.3.1 Model Selection and Justification

Three distinct regression algorithms will be implemented and compared:

1. **Multiple Linear Regression (MLR):** This model will serve as a crucial baseline. It assumes a simple linear relationship between the predictors and the target. By establishing its performance first, we create a benchmark against which the more complex models can be judged. If the advanced models do not significantly outperform MLR, it may indicate that the underlying relationships are simpler than initially hypothesized.
2. **Random Forest (RF):** As an ensemble of decision trees, RF is highly effective and robust. Its mechanism of training many trees on different subsets of data (bagging) and considering only a random subset of features at each split makes it highly resistant to overfitting and noise, which is invaluable when dealing with real-world industrial data.
3. **Extreme Gradient Boosting (XGBoost):** This is a state-of-the-art implementation of gradient boosting. Unlike RF, which builds trees

independently, XGBoost builds them sequentially, with each new tree specifically designed to correct the errors made by the previous ones. This iterative learning process makes it exceptionally powerful at uncovering subtle and complex patterns within the data, often leading to superior predictive accuracy.

3.3.2 Model Training and Hyperparameter Optimization

The models will be trained exclusively on the pre-processed training set.

- **K-Fold Cross-Validation:** To ensure the model's performance is stable and not the result of a lucky train-test split, 10-fold cross-validation will be employed. The training data is divided into 10 subsets where the model is trained on 9 of them and tested on the 10th, repeating this process 10 times. The average performance across all 10 folds provides a much more reliable estimate of the model's generalization ability.
- **Hyperparameter Tuning:** Machine learning models have internal settings called hyperparameters that are not learned from the data. For RF and XGBoost, these include parameters like the number of trees and the maximum depth of each tree. To find the optimal combination of these settings, a Grid Search Cross-Validation strategy will be implemented. This process systematically builds a model for every possible combination of hyperparameters and evaluates them using cross-validation, ultimately selecting the combination that yields the best performance.

3.4 Model Evaluation Framework

To objectively assess and compare the performance of the trained models, a suite of standard regression evaluation metrics will be calculated on the held-out test set. Using multiple metrics provides a holistic and nuanced understanding of each model's strengths and weaknesses.

- **Coefficient of Determination (R^2):** R-squared measures the proportion of the variance in the 28-day strength that is predictable from the input features. A value of 0.90, for example, would imply that 90% of the variability in the final strength can be explained by the model's inputs. It is a scale-free score, which makes it good for general interpretation.

- **Mean Absolute Error (MAE):** This metric represents the average absolute difference between the predicted values and the actual values. It is easily interpretable as it is in the same unit as the target variable (MPa). For the partner company, this translates to the average magnitude of the prediction error, for example, on average, the model's prediction is off by 2.5 MPa.
- **Root Mean Squared Error (RMSE):** Like MAE, RMSE is in the same unit (MPa), but by squaring the errors before they are averaged, it gives a disproportionately higher weight to large errors. In an engineering context where large prediction errors can be more critical than small ones, RMSE is a particularly important metric. A model with a low RMSE is one that avoids making significant mispredictions.

3.5 Model Interpretability and Insights Generation

For this project to deliver true value, it must produce more than just predictions. It must generate insights. In a safety-critical field like civil engineering, a "black box" model, no matter how accurate, is of limited use if its decision-making process cannot be understood, trusted, and verified. Therefore, a significant focus will be placed on model interpretability using state-of-the-art Explainable AI (XAI) techniques.

- **SHAP (SHapley Additive exPlanations):** SHAP is a game theory-based approach that explains the output of any machine learning model by computing the contribution of each feature to a particular prediction. It provides a powerful and consistent way to understand model behavior at both a global and local level.
  - **Global Interpretation:** A SHAP summary plot (beeswarm plot) will be generated to rank the features by their overall importance. This will definitively show which variables (7-day strength, average temperature, humidity, etc.) have the most impact on the model's predictions across the entire dataset.

    - **Local Interpretation:** SHAP dependence plots will be used to visualize the marginal effect of a single feature on the predicted outcome, illustrating complex relationships (e.g., how the effect of temperature on strength might not be linear). Furthermore, SHAP force plots will be generated for individual predictions, showing exactly how each feature contributed to pushing a specific prediction higher or lower. This capability is invaluable for performing diagnostic analyses and explaining the model's reasoning to project stakeholders at the partner company.

3.6 Related Tools

This section outlines the software and libraries that will be instrumental in implementing the data analysis, modeling, and interpretation phases of this project.

- **Python:** This will serve as the primary programming language for the project. Python's extensive ecosystem of libraries for data science and machine learning makes it the ideal choice.
- **Pandas:** A powerful Python library essential for data manipulation and analysis. It will be used for loading, merging (company production records with weather data), cleaning, and structuring the datasets into data frames for easy handling.
- **NumPy:** The fundamental library for numerical computing in Python. It will be used for efficient array and matrix operations, which are crucial for handling feature vectors and preparing data for machine learning models.
- **Scikit-learn:** This is a core machine learning library for Python. It will be utilized for multiple key tasks, including:
    - Implementing the Multiple Linear Regression (MLR) and Random Forest (RF) models.
    - Splitting the dataset into training and testing sets.
    - Performing hyperparameter optimization using GridSearchCV.
    - Evaluating model performance with metrics like R^2, MAE, and RMSE
- **XGBoost:** A dedicated library for the eXtreme Gradient Boosting algorithm. It will be used to implement the XGBoost model, which is renowned for its high performance and scalability.

- **SHAP (SHapley Additive exPlanations):** This library is essential for model interpretability. It will be used to generate SHAP summary and dependence plots to explain the output of the predictive models and quantify the impact of each feature.
- **Matplotlib & Seaborn:** These are Python libraries for data visualization. They will be used to create plots for exploratory data analysis, outlier detection, and visualizing the results from the SHAP analysis.

4.0 Work Plan

4.1 Capstone Project 1 Work Plan

4.1.1 Phase 1: Project Initiation & Planning (Week 1-5)

Activity	Description	Deliverables	Predecessor and Successor	Risk Factors
Topic Finalization	Solidify the project's focus on predicting concrete strength using machine learning for quality control purposes.	Finalized project title and a brief project description.	Predecessor: N/A Successor: Define Problem Statement & Objectives.	The project scope could be too ambitious for the timeframe.
Define Problem Statement, Aim & Objectives	Clearly articulate the limitations of traditional methods and set the specific aim and objectives for the ML-based approach.	Problem statement, project objectives, and scope of work.	Predecessor: Topic Finalization Successor: Initial Work Plan.	Objectives might not be specific or measurable, leading to a lack of focus.
Produce initial work plan and Gantt chart	Draft a high-level work plan and Gantt chart illustrating the timeline for all major phases	Initial version of the work plan and Gantt chart.	Predecessor: Define Problem Statement & Objectives Successor:	Overly optimistic scheduling can lead to delays later in the project.

	and activities in both Capstone 1 and 2.		Literature Review.	

4.1.2 Phase 2: Literature Review & Requirement Analysis (Week 6 – 10)

Activity	Description	Deliverables	Predecessor and Successor	Risk Factors
Review Concrete Fundamentals & Testing Methods	Conduct a detailed review of concrete composition, strength factors (W/C ratio, aggregates, curing), and traditional destructive/NDT methods.	A comprehensive literature review on concrete technology and testing.	Predecessor: Initial Work Plan Successor: Review ML Literature.	Difficulty in finding consolidated and up-to-date sources on specific materials
Review ML for Concrete Strength Prediction	Analyze literature on ML regression algorithms (Linear, SVR, Trees, ANN, Ensembles) and their specific application in concrete strength prediction,	Literature review on ML applications, comparative model performance tables , and key input features.	Predecessor: Review Concrete Fundamentals Successor: Methodology Planning.	Performance metrics in literature may be inconsistent, making direct comparisons difficult.

	including feature importance analysis			

4.1.3 Phase 3: Architecture & Workflow Planning (Week 11-13)

Activity	**Description**	**Deliverables**	**Predecessor and Successor**	**Risk Factors**
Plan Data Acquisition & Integration Strategy	Define the strategy to acquire primary data from the partner company and secondary weather data from MetMalaysia. Plan the data merging process using 'Casting Date' as the key.	A detailed data acquisition and integration plan.	Predecessor: Literature Review Successor: Plan Modeling Strategy.	The partner company might have restrictions on data sharing (NDA required). Weather data might not be available for the exact location.
Plan Data Preprocessing & Feature Engineering	Outline the steps for data cleaning, outlier handling, and creating new features like lagged weather variables (e.g.,	A detailed plan for data preprocessing and feature engineering.	Predecessor: Plan Data Acquisition Successor: Plan Modeling Strategy.	The impact of engineered features is unknown until implementation.

	7-day rolling averages) and temporal features.			
Plan Modeling & Evaluation Strategy	Select the specific models for implementation (MLR, RF, XGBoost) and the evaluation metrics (R^2, MAE, RMSE). Plan the hyperparameter tuning and model interpretation (SHAP) approach.	A complete modeling, evaluation, and interpretation plan.	Predecessor: Plan Data Preprocessing Successor: Finalize Documentation.	The chosen models might require significant computational resources for tuning.

4.1.4 Phase 4: Finalizing Documentation (Week 14)

Activity	Description	Deliverables	Predecessor and Successor	Risk Factors
Finalize Capstone 1 Report	Compile all sections (Introduction, Literature Review, Methodology, Work Plan) into a coherent document. Refine and format for final submission.	A complete and finalized Capstone Project 1 document.	Predecessor: All previous phases Successor: Begin Capstone 2.	Time constraints may rush the final review and proofreading process.

4.2 Capstone Project 2 Work Plan

4.2.1 Phase 1: Data Implementation (Week 1-3)

Activity	Description	Deliverables	Predecessor and Successor	Risk Factors
Data Acquisition and Merging	Obtain the company production records and MetMalaysia historical weather data. Integrate them into a single data	A unified, raw dataset containing both strength and weather data.	Predecessor: Capstone 1 Completion Successor: Data Cleaning.	Delays in receiving data from the partner company; data quality issues.

	frame using the 'Casting Date'.			
Data Cleaning and Feature Engineering	Execute the preprocessing plan: handle missing values and outliers, and programmatically create lagged weather and temporal features.	A clean, processed, and feature-engineered dataset.	Predecessor: Data Acquisition Successor: Model Training.	The feature engineering process may be more complex than anticipated.
Data Splitting	Split the finalized dataset into training and testing sets to prepare for model development and ensure unbiased evaluation.	Separate training and testing data files.	Predecessor: Data Cleaning Successor: Model Training.	An improper split could lead to data leakage and overly optimistic results.

4.2.2 Phase 2: Model Development & Training (Week 4-7)

Activity	Description	Deliverables	Predecessor and Successor	Risk Factors
Train Baseline Model	Implement and train the Multiple Linear Regression model to serve as a	A trained model.	Predecessor: Data Splitting Successor: Train Ensemble Models.	Baseline performance may be very low, confirming the need for complex models.

	performance benchmark.			
Train & Tune Ensemble Models (RF, XGBoost)	Implement the Random Forest and XGBoost models. Use 10-fold cross-validation and Grid Search to find the optimal hyperparameters for each.	Fully trained and optimized RF and XGBoost models.	Predecessor: Data Splitting Successor: Model Evaluation.	Hyperparameter tuning can be computationally expensive and time-consuming.

4.2.3 Phase 3: Model Evaluation & Interpretation (Week 8-10)

Activity	**Description**	**Deliverables**	**Predecessor and Successor**	**Risk Factors**
Model Performance Evaluation	Test the final trained models on the held-out test set. Calculate R^2, MAE, and RMSE for each model to compare their predictive accuracy.	A comprehensive report with performance metrics for all models.	Predecessor: Model Training Successor: Model Interpretation.	The chosen metrics might not fully capture the business needs (e.g., cost of large errors).
Model Selection & Interpretation	Select the best-performing model. Apply SHAP to	The selected final model and a full interpretation report with	Predecessor: Model Performance Evaluation Successor:	SHAP results may be complex to explain to a

	generate summary and dependence plots to understand feature impacts and model behavior.	SHAP visualizations.	Finalize Report.	non-technical audience.

4.2.4 Phase 4: Final Documentation & Submission (Week 10-12)

Activity	Description	Deliverables	Predecessor and Successor	Risk Factors
Compile Final Report	Write the results, discussion, and conclusion sections. Integrate all parts of the project into the final report.	A complete, submission-ready Final Year Project report.	Predecessor: Model Interpretation Successor: Project Submission.	Insufficient time for thorough writing and revision of the final document.
Prepare for Final Presentation	Create presentation slides summarizing the project's aim, methodology, key findings, and conclusions.	A complete presentation deck and project submission package.	Predecessor: Compile Final Report Successor: N/A.	Technical difficulties during the final presentation.

	Rehearse for the final presentation.			

4.3 Gantt Chart

Activities/Tasks	Week													
	Capstone 1													
	Week													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Project Initiation and Planning														
Finalize Project Topic and Scope														
Define Problem Statement, Aim, and Objectives														
Develop Initial Work Plan and Gantt Chart														
Litreature Review and Requirement Analysis														
Conduct Literature Review on Concrete Properties & Testing Methods														
Conduct Literature Review on ML Applications for Strength Prediction														
Literature review on integration methods														
Methodology & Workflow Planning														
Plan Data Preprocessing and Feature Engineering Workflow														
Plan Modeling Strategy (MLR, RF, XGBoost)														
Plan Model Evaluation and Interpretation Strategy														
Document Finalisation														
Compile and Finalize Capstone 1 Report														

Figure 4.0.1 Gantt Chart for Capstone 1

Activities/Tasks	Week													
	Capstone 2													
	Week													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Data Implementation														
Acquire dataset														
Preprocess and clean dataset														
Perform Feature Engineering														
Model Development & Training														
Train Baseline Model: Multiple Linear Regression (MLR)														
Train and Tune Random Forest (RF) Model														
Train and Tune XGBoost Model														
Model Evaluation & Interpretation														
Evaluate All Models on Test Set (using R^2, MAE, RM														
Compare Model Performances and Select Best Moc														
Document Finalisation														
Final model optimisation														
Finalise report														

Figure 4.0.2 Gantt Chart for Capstone 2

5.0 References

Aakash. (2024, November 28). Supervised Learning: Regression Techniques - Aakash R - Medium. *Medium*. https://medium.com/@aakash013/master-regression-techniques-in-supervised-learning-352ffd5f69cc

ACI Committee. (2000). *Guide to Concrete Construction — Part V-Section 13 — Compaction*. https://www.ccaa.com.au/common/Uploaded%20files/CCAA/Publications/Technical%20Publications/PART_V_-_13_-_COMPACTION_GTCC_2020.pdf

Akande, K. O., Owolabi, T. O., Twaha, S., & Olatunji, S. O. (2014). Performance comparison of SVM and ANN in predicting compressive strength of concrete. *IOSR Journal of Computer Engineering*, *16*(5), 88–94. https://doi.org/10.9790/0661-16518894

Alem. (2022, October 4). *6 important factors that affect the strength of concrete*. Al Manaratain. https://almanaratain.com/6-important-factors-that-affect-the-strength-of-concrete/

Alooba. (n.d.). *Data splitting: Everything you need to know when assessing data splitting skills*. https://www.alooba.com/skills/concepts/deep-learning/data-splitting/

Altuncı, Y. T. (2024). A comprehensive study on the estimation of concrete compressive strength using machine learning models. *Buildings*, *14*(12), 3851. https://doi.org/10.3390/buildings14123851

ASTM. (1991). *Strength*. https://www.engr.psu.edu/ce/courses/ce584/concrete/library/materials/aggregate/Strength.htm

ASTM, NRMCA, Portland Cement Association, & American Concrete Institute. (2003). *Testing compressive strength of concrete*.

Auwal Ibrahim, M., I., D, S., U, A., S, U., I., U, S. S., Z, I., I., & A, I., I. (2024). A Review on the Curing of Concrete using Different Methods. *International Journal of Mechanical and Civil Engineering*, *7*(2), 15–25. https://doi.org/10.52589/ijmce-4envmzox

Baladram, S. (2024, November 30). Regression Tree | TDS Archive. *Medium*. https://medium.com/data-science/decision-tree-regressor-explained-a-visual-guide-with-code-examples-fbd2836c3bef

Chugani, V. (2024). *A Comprehensive Guide to K-Fold Cross Validation*. Datacamp. https://www.datacamp.com/tutorial/k-fold-cross-validation

Content Studio. (2024, August 8). *What is data preprocessing for machine learning?* Pure Storage. https://www.purestorage.com/uk/knowledge/what-is-data-preprocessing.html

Daneshvar, D., Behnood, A., & Robisson, A. (2022). Interfacial bond in concrete-to-concrete composites: A review. *Construction and Building Materials*, *359*, 129195. https://doi.org/10.1016/j.conbuildmat.2022.129195

De De Jesus, W. S., De Almeida, T. M. S., Da Silva, S. F. M., Souza, M. T., Leal, E. S., Souza, R. S., Sacramento, L. A., Allaman, I. B., & De Castro Pessôa, J. R. (2025). Leveraging delayed strength gains in supplementary cementitious material concretes: rethinking mix design for enhanced cost efficiency and sustainability. *Journal of Composites Science*, *9*(3), 110. https://doi.org/10.3390/jcs9030110

Dzięcioł, J. (2024). Machine Learning in Civil Engineering - on the example of prediction of the coefficient of permeability. *Acta Scientiarum Polonorum Architectura*, *22*, 184–191. https://doi.org/10.22630/aspa.2023.22.18

Elhishi, S., Elashry, A. M., & El-Metwally, S. (2023). Unboxing machine learning models for concrete strength prediction using XAI. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-47169-7

Farooq, F., Amin, M. N., Khan, K., Sadiq, M. R., Javed, M. F. F., Aslam, F., & Alyousef, R. (2020). A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC). *Applied Sciences*, *10*(20), 7330. https://doi.org/10.3390/app10207330

Fernando, J. (2024, November 13). *R-Squared: Definition, Calculation, and interpretation*. Investopedia. https://www.investopedia.com/terms/r/r-squared.asp

Format NDT Ltd. (2025, May 7). *What is Non-Destructive Testing of Concrete? | Format NDT*. Format NDT. https://www.formatndt.co.uk/what-is-non-destructive-testing-of-concrete/

GeeksforGeeks. (2024a, April 9). *Gradient Boosting vs Random Forest*. GeeksforGeeks. https://www.geeksforgeeks.org/machine-learning/gradient-boosting-vs-random-forest/

GeeksforGeeks. (2024b, August). *Support Vector Machine (SVM) algorithm*. https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm/

GeeksforGeeks. (2025a, January 13). *Regression in machine learning*. GeeksforGeeks. https://www.geeksforgeeks.org/machine-learning/regression-in-machine-learning/

GeeksforGeeks. (2025b, July). *Backpropagation in neural network*. https://www.geeksforgeeks.org/machine-learning/backpropagation-in-neural-network/

Hearns, A. (2022, June 27). 7 Methods for testing concrete strength. *For Construction Pros*. https://www.forconstructionpros.com/concrete/article/21072546/giatec-scientific-inc-7-methods-for-testing-concrete-strength#:~:text=The%20most%20common%20method%20for,party%20lab%20at%20various%20stages.

IBM. (2021, June). *Neural Networks*. https://www.ibm.com/think/topics/neural-networks

Jain, A. (2024, November 16). Pre-Pruning and Post-Pruning in Decision Trees: A Comprehensive guide. *Medium*. https://medium.com/@abhishekjainindore24/pre-pruning-and-post-pruning-in-decision-trees-a-comprehensive-guide-391fd3682883

Jedidi, M., Kaouther, M., National Engineering School of Tunis, & High institute for technological studies of rades, tunis,tunisia. (2014). Destructive and non-destructive testing of concrete structures. *Jordan Journal of Civil Engineering*, *8*(4), 432–434. https://www.researchgate.net/publication/279547128

Li, W. (2021). Analysis of the influence of water-cement ratio on concrete strength. *E3S Web of Conferences*, *283*, 01016. https://doi.org/10.1051/e3sconf/202128301016

Liu, J. (2024). A review of research on prediction methods for compressive strength of concrete. *Frontiers in Science and Engineering*, *4*(2), 31–35. https://doi.org/10.54691/n1f9hj06

Loyola-González, O. (2019). *Black-Box vs. White-Box: Understanding Their Advantages and Weaknesses From a Practical Point of View*. IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/8882211

Marchewka, A., Ziolkowski, P., & Galán, S. G. (2025). Real-time prediction of early concrete compressive strength using AI and hydration monitoring. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-97060-w

McClements, D., & Keane, P. (2025, August 5). *ASTM C39/C39M: Concrete Cylinder Testing*. Xometry. https://www.xometry.com/resources/certifications/astm-compressive-strength/

Meyer, M., Van Zyl, J., & Combrinck, R. (2022). The influence of temperature on the cracking of plastic concrete. *MATEC Web of Conferences*, *364*, 02018. https://doi.org/10.1051/matecconf/202236402018

Mistry, R. (2025, March 14). Data splitting (Train-Test-Validation) in machine learning. *Medium*. https://medium.com/@rohanmistry231/data-splitting-train-test-validation-in-machine-learning-2d5d1927fa69

Mohammed, G. A., & Al-Mashhadi, S. a. A. (2020). Effect of maximum aggregate size on the strength of normal and high strength concrete. *Civil Engineering Journal*, *6*(6), 1155–1165. https://doi.org/10.28991/cej-2020-03091537

Murel, J., PhD, & Kavlakoglu, E. (2025, July 22). *Ensemble learning*. IBM. https://www.ibm.com/think/topics/ensemble-learning

Mustapha, I. B., Abdulkareem, M., Jassam, T. M., AlAteah, A. H., Al-Sodani, K. a. A., Al-Tholaia, M. M. H., Nabus, H., Alih, S. C., Abdulkareem, Z., & Ganiyu, A. (2024). Comparative Analysis of Gradient-Boosting ensembles for estimation of compressive strength of quaternary blend concrete. *International Journal of Concrete Structures and Materials*, *18*(1). https://doi.org/10.1186/s40069-023-00653-w

National Concrete Pavement Technology Center. (n.d.). *Cementitious materials*. https://www.cptechcenter.org/cementitious-materials/

National Ready Mixed Concrete Association. (n.d.). *Flexural Strength Concrete — What, why, & how?* Nevada Ready Mix. https://www.nevadareadymix.com/concrete-tips/flexural-strength-concrete/

National Ready Mixed Concrete Association. (2021, October 26). *About concrete*. NRMCA. https://www.nrmca.org/about-nrmca/about-concrete/

Nyokum, T., & Tamut, Y. (2025). Artificial intelligence in civil engineering: emerging applications and opportunities. *Frontiers in Built Environment*, *11*. https://doi.org/10.3389/fbuil.2025.1622873

Pal, N. S., Trang, N. L. H., Hieu, N. V. T., Nguyen, N. D. D., Vu, N. D. Q., & Prakash, N. I. (2024). Investigation of Support Vector Machines with Different Kernel Functions for Prediction of Compressive Strength of Concrete. *Journal of Science and Transport Technology*, 55–68. https://doi.org/10.58845/jstt.utt.2024.en.4.2.55-68

Paudel, S., Pudasaini, A., Shrestha, R. K., & Kharel, E. (2023). Compressive strength of concrete material using machine learning techniques. *Cleaner Engineering and Technology*, *15*, 100661. https://doi.org/10.1016/j.clet.2023.100661

Qi, C., Huang, B., Wu, M., Wang, K., Yang, S., & Li, G. (2022a). Concrete strength prediction using different machine learning processes: effect of slag, fly ash and superplasticizer. *Materials*, *15*(15), 5369. https://doi.org/10.3390/ma15155369

Qi, C., Huang, B., Wu, M., Wang, K., Yang, S., & Li, G. (2022b). Concrete strength prediction using different machine learning processes: effect of slag, fly ash and superplasticizer. *Materials*, *15*(15), 5369. https://doi.org/10.3390/ma15155369

Qureshi, M. A., Aslam, M., Shah, S. N. R., & Otho, S. H. (2015). Influence of aggregate characteristics on the compressive strength of normal weight concrete. *Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan*, *20*, III–2015. https://www.researchgate.net/publication/283501237

Reyes, J. (2025, January 20). *Concrete Curing Time Chart: Achieving maximum compressive strength*. HeatAuthority. https://heatauthority.com/blogs/news/concrete-curing-time-chart-achieving-maximum-compressive-strength

Roberts, A. (2023, February 10). *Mean Absolute percentage error (MAPE): what you need to know*. Arize AI. https://arize.com/blog-course/mean-absolute-percentage-error-mape-what-you-need-to-know/

Ryu, G., An, G., Yoon, Y., Kim, J., & Choi, S. (2024). Evaluation of concrete compressive strength prediction using the maturity method incorporating various curing temperatures and binder compositions. *Materials*, *17*(23), 5794. https://doi.org/10.3390/ma17235794

Sabbha. (2024, August 16). *Understanding MAE, MSE, and RMSE: Key metrics in Machine Learning*. DEV Community. https://dev.to/mondal_sabbha/understanding-mae-mse-and-rmse-key-metrics-in-machine-learning-4la2

Sah, A. K., & Hong, Y. (2024). Performance comparison of machine learning models for concrete compressive strength prediction. *Materials*, *17*(9), 2075. https://doi.org/10.3390/ma17092075

Saiwa. (2023). *Data Preprocessing in Machine Learning*. https://saiwa.ai/blog/data-preprocessing-in-machine-learning/

Shaqadan, A. (2016). Prediction of concrete mix strength using random forest model. In Zarqa University, *International Journal of Applied Engineering Research* (Vols. 11–11, pp. 11024–11029). Research India Publications. https://www.researchgate.net/publication/311797168

Tak, M. S. N., Feng, Y., & Mahgoub, M. (2025). Advanced machine learning techniques for predicting concrete compressive strength. *Infrastructures*, *10*(2), 26. https://doi.org/10.3390/infrastructures10020026

Tegegne, D. A., Kitaw, D., & Berhan, E. (2022). Advances in statistical quality control chart techniques and their limitations to cement industry. *Cogent Engineering*, *9*(1). https://doi.org/10.1080/23311916.2022.2088463

Thapa, J. (2024a). Machine learning approaches for predicting concrete compressive strength. *Journal of Advanced Research in Civil and Environmental Engineering*, *11*(1), 9–20. https://doi.org/10.24321/2393.8307.202404

Thapa, J. (2024b). Concrete Compressive Strength Prediction by Ensemble Machine Learning Approach. *Journal of Engineering and Sciences*, *3*(1), 66–73. https://doi.org/10.3126/jes2.v3i1.66249

The Concrete Countertop Institute. (2021, October 18). *What is Temperature & Relative Humidity for Concrete Countertop*. https://concretecountertopinstitute.com/free-training/temperature-and-relative-humidity-what-they-mean-for-you-and-your-concrete-countertops/

Tuncan, M., Arioz, O., Ramyar, K., Karasu, B., Tuncan, A., Kilinc, K., Anadolu University, Ege University, & Eskisehir Technical University. (2007). *Effect of compaction on assessed concrete strength*.

UltraTech. (2021). *Guide on Admixtures in Concrete: 10 Types and their Uses*. https://www.ultratechcement.com/for-homebuilders/home-building-explained-single/descriptive-articles/a-guide-on-admixtures-in-concrete

Wang, Y., Wu, Z., Zheng, X., Li, K., & Liu, C. (2025). A review of concrete exposed to low-temperature environments at early ages: mechanical properties and durability. *Journal of Building Engineering*, 113122. https://doi.org/10.1016/j.jobe.2025.113122

Wiratsin, I., & Ragkhitwetsagul, C. (2025). Effectiveness of Explainable Artificial intelligence (XAI) techniques for improving human Trust in machine learning Models: A systematic literature review. *IEEE Access*, 1. https://doi.org/10.1109/access.2025.3575022

Zeyad, A. M., Tayeh, B. A., Adesina, A., De Azevedo, A. R., Amin, M., Hadzima-Nyarko, M., & Agwa, I. S. (2022). Review on effect of steam curing on behavior of concrete. *Cleaner Materials*, *3*, 100042. https://doi.org/10.1016/j.clema.2022.100042