Predictive Model for Consulting Project Success

Eeshan Waqar

September 2025

School of Mathematics,
Cardiff University

A dissertation submitted in partial fulfilment of the
requirements for **MSc Data Science and Analytics**
by taught programme, supervised by **Surya Thottam Valappil**.

Executive Summary

Consulting projects are vital for organizations that want to grow, change or improve how they work. Many of these projects adopt to new technologies and redesign their business processes. But many struggle to deliver what they promise. These failures originate from many factors some of them being higher than expected costs and disappointed clients. For consulting firms these failures can lead to financial losses and a threat to business reputation.

In this dissertation, modern data techniques are explored that can help predict which projects are likely to succeed and which are at the risk of failing. Some of the traditional approaches, such as project dashboards or expert judgements, are useful but only highlight the problems which have already occurred. We lay our emphasis on machine learning, a form of artificial intelligence, which utilizes past project data and predict early warning signs that can help managers act on time.

The study uses data from consulting projects that include financial projects, how teams are organised, and what kind of approach was taken to deliver the results. Furthermore, success was measured in two ways: how well the projects performed across different areas, and a simple pass-or-fail flag. Machine learning models were than tested to see how accurately they could predict the outcome.

The findings showed that financial and human level indicators of a project are the most important drivers of project success. Projects that have strong financial performance, healthy revenues and margins, were more likely to succeed. How teams are structured also plays an important role: where larger well-balanced teams delivered better results, while stable teams with experienced members were more effective. In some cases, variety in roles helped but too much diversity sometimes made coordination harder.

These predictive models performed well. Where in some cases they were able to forecast project outcomes with a high degree of accuracy. The advance models gave early-warning signals that could be built into project monitoring systems, while simpler models still offered valuable insights into why projects were succeeding or struggling.

The overall conclusion is that consulting firms can gain real value from adopting predictive analytics. By keeping track of financial health, team design, and continuity, and by using data-driven forecasts alongside human judgement, firms can manage risks more proactively.

This does not replace managers, but it equips them with stronger evidence to make decisions, intervene earlier, and improve the chances of project success.

In the long run, using tools like these could help consulting firms deliver more consistent results, strengthen client trust, and stay competitive in an industry where success rates are often uncertain.

Table Of Content

List Of Figures

List Of Tables

Chapter 1; Introduction

1.1 Context and Motivation

Organizational growth and change are anchored on projects. Projects, be it the adoption of new technologies, redesign of business processes, or the provision of customized client solutions, offer the mechanism through which strategic change is realized. This is more clearly seen in management consulting where not just immediate client value can be delivered, but also competitive advantage of firms in an economy that is becoming more knowledge based.

But even with such centrality, projects often fall short of expectations. Factual data brings out the scale of this challenge. In the Standish Group CHAOS report (2013), more than 60 percent of projects either face challenges or fail completely, and schedule overruns (74 percent), unmet requirements (69 percent), and budget overruns (59 percent) are the most frequent problems (Fauser et al., 2015). These failures signify a great loss of both money and reputation, as well as strategy, to consulting firms and their clients. In the worst-case scenario, failure of the project can be a threat to the very existence of companies (Bloch, 2011; Fauser et al., 2015).

These risks are especially susceptible to the consulting industry. In comparison to standardized industrial or construction projects, consulting engagements are highly variable, always dependent on client-specific needs, dynamic stakeholder expectations and an ever-evolving external environment. This variability adds to the uncertainty of results and makes the conventional project management much harder (Ajanaku, 2025).

The power to forecast success or failure in the early stages of the project process has great strategic value in this respect. Lagging indicators and expert judgement are however major practices in the monitoring and evaluation practices, which are traditional. Although these approaches provide a useful insight, they are usually subjective, disjointed, and reactive. As Deloitte has pointed out, predictive project analytics (PPA), in comparison with traditional risk evaluation, fail to isolate the exact deficiencies of governance and control that lead to the failure of projects (Fauser et al., 2015). This illustrates the fact that organizations have all the information available to them but lack insight into predicting the success of the project.

A possible solution is recent developments in data science, artificial intelligence (AI), and machine learning (ML). By utilizing predictive analytics, risks, optimizing processes, and making predictions, this application has succeeded in the sphere of finance, retail, and healthcare. Its use in project management is developing but not fully developed. For example, a study by Ajanaku (2025) showed that ML models, such as XGBoost and Random Forest, can attain high predictive accuracy for the estimation of cost on consulting projects. Furthermore, the research of dynamic prediction framework has proven that AI can adapt to the changes of project environment and improve the prediction accuracy during the project life cycle (Gao et al., 2024). These developments reflect the possibility of using predictive techniques not only for costs, but for measures of wider aspects of project success.

At the same time, organizational research points to the fact that project outcomes are affected by multidimensional factors other than cost and schedule. The collaboration of social, technical and organizational complexities is also emphasized in the international literature, and both quantitative metrics and qualitative drivers must be integrated into holistic predictive models (Fauser et al., 2015).

Combined, this literature indicates the presence of an opportunity and gap. All consulting projects retain a certain level of failure; however, predictive analytics and machine learning provide a set of tools that can be used to mitigate this situation. However, existing empirical studies rarely compare various techniques formulated and applied in a systematic way to identify the future success of consulting projects defined not merely in economic terms, but as a multi-dimensional concept embracing outcomes to do with delivery, customer, and strategic value.

1.2 Problem Statement

Despite the considerable size of project data repositories managed by consulting firms, the project data are not sufficiently used to make predictions. Conventional risk management depends on the expert opinions which are mostly subjective, inconsistent and reactive (Fauser et al., 2015). Academic research offers valuable research frameworks based on which the success factors may be determined but fails to present a predictive model that can be implemented at that scale (Alias et al. 2014). Consequently, managers do not have tools that can help them predict results in advance, track projects actively and respond efficiently in case of risks.

To fill this gap, this dissertation explains the way machine learning methods could be used in forecasting consulting project success. It attempts to use historical data with advanced analytics to go beyond the descriptive assessment to provide predictive insights that can guide managerial decisions.

1.3 Research Aims and Objectives

The primary aim of this research is to explore the application of machine learning to consulting project data to identify success drivers and build predictive models of project outcomes. Specifically, the study pursues four objectives:

1. Identify key drivers of project success by analysing historical consulting project data, including financial, organizational, and contextual variables.
2. Develop a machine learning model capable of predicting project success with a high degree of accuracy.
3. Evaluate the predictive performance of the model against established benchmarks, using metrics appropriate to both classification and regression tasks.
4. Generate practical recommendations for consulting firms on how predictive insights can be integrated into project monitoring, governance, and decision-making processes.

1.4 Research Questions

This dissertation is guided by the following questions:

1. What are the key factors that contribute to consulting project success?
2. How accurately can machine learning models predict project success using historical data?
3. What recommendations can be drawn from predictive insights to support proactive project monitoring and risk management?

1.5 Structure of the Dissertation

The dissertation is organized into seven chapters.

- Chapter 2 – Literature Review examines research on project success factors, the role of management support, and recent applications of predictive analytics and ML in project management.
- Chapter 3 – Methodology outlines the research design, including data sources, preprocessing, feature engineering, and the machine learning techniques employed.
- Chapter 4 – Data Exploration and Preparation present the exploratory analysis and the steps taken to clean, transform, and prepare data for modelling.
- Chapter 5 – Model Development and Evaluation describe the development of machine learning models, their training and validation, and a comparative analysis of predictive accuracy.

- Chapter 6 - Discussion, Conclusion and Recommendations discusses the implications of the findings for consulting practice while highlighting key success drivers and managerial implications, concludes with remarks and recommendations for future research and practitioners.

Chapter 2; Literature Review

2.1 Defining Project Success

The meaning of project success has proven to be one of the intellectual issues in the history of project management documents. Initial conceptualizations were characterized by the dominance of the so-called iron triangle/triple constraint model where success is strictly limited to projects being on-time, on-cost, and on-scope (Atkinson, 1999). This perspective, although apparently correct and practically useful, has been too reductionistic since it does not consider larger organizational, strategic and stakeholder implications of the project outcomes (Jugdev and Muller, 2005).

To answer this, scholars have suggested a multi-dimensional definition of success. Shenhar, Dvir, Levy and Maltz (2001) suggested a four-level framework that segregates into efficiency (meeting the usual constraints), customer-centred impact (satisfaction, meeting the needs), business achievements (profitability, The competitive advantage), and creating the future (capability development). This wider perspective understands that a project may be deemed as successful even when it overruns in one aspect, if it realizes strategic value in the long run. Similarly, Baccarini (1999) provided differentiation on both the project management success (time, cost and quality) and project product success (In terms of effectiveness of meeting objectives and delivery of beneficial ramifications).

Such extended perspectives are especially applicable to the consulting context. The nature of consulting projects is knowledge based, specific to the client, and relational, therefore, client satisfaction and strategic fit can be taken as key performance indicators. Researchers point out that consulting firms often measure the project success not only by its timely and within-budget outcome delivery, but also with references to the ways in which the engagement increases the strength of relationships with clients and produces opportunities (Alias et al., 2014).

Recent empirical studies still emphasize that success is multidimensional. Another example of this may be that Gao et al. (2024) also suggest that by measuring success as the single variable of cost and schedule performance, the idea of systemic drivers of project outcomes; that is, contract design, incentives, and governance mechanisms could be neglected.

Fauser et al. (2015) also emphasise the fact that predictive frameworks should consider those gaps in governance that control not only efficiency but also the sustainability of the entire project. Combined, these views imply that the success of a project should be viewed as a multi-dimensional construct that includes:

- Effectiveness: project and deliverable accomplishment.
- Efficiency: compliance with time, cost, and scope limits.
- Client and stakeholder satisfaction: above and beyond the expectation.
- Strategic congruence: support of long-term organizational goals.
- Monetary and relationship deliverables: profitability, margins, and client retention (especially in consulting).

The dissertation has taken the multi-dimensional approach, however with a keen focus on the consulting-specific dimensions of financial margins, effectiveness in delivery, and client engagement. This kind of strategy is consistent both with the scholarly evidence on successful projects and with the practical demands of consulting firms, where projects are both means to provide value quickly and to build strategic relationships.

2.2 Success Factors in Consulting Projects

Whereas setting the boundaries of project success creates parameters where projects are measured, it is also necessary to understand the driving factors of success. The literature on Critical Success Factors (CSFs) offers an adequate prism using which project performance can be examined and forecasted. Introduced by Rockart (1979) and further applied to project management by Pinto and Slevin (1987), the CSF approach centres the identification of a small number of variables that have an exceptionally large effect (having a disproportionate influence) on project performance. Over the years CSFs have been narrowed down into categories which can be described as both hard factors, which include planning and control processes, and soft factors, which include leadership, communication and stakeholder involvement (Ika, 2009).

The CSFs are of particular importance in consulting projects. In contrast to engineering or construction projects, which have a tangible output and a standardized process, consulting projects are typified by intangible output, excessive dependence on human expertise, and a dynamic relationship with the client.

2.2.1 Team Dynamics and Capabilities

Because consulting projects are resource intensive and knowledge intensive, project composition, skills and cohesion are an important determinant of success. Successful teams have a combination of technical skills and interpersonal skills like communication, flexibility and problem-solving. The literature emphasizes that role clarity, knowledge sharing, and teamwork are essential in consulting teams (Alias et al., 2014). On the other hand, team instability or inexperience-skill mismatch has been associated with delays, cost overruns, and client dissatisfaction.

2.2.2 Client Engagement and Relationship Quality

A client-consultant relationship is always discussed as a determining factor of success when consulting. The success of consulting is often measured not only on the immediate outputs of a project but also on whether the client-consultant relationship is strengthened by the project overall and subsequently impacts on repeat business and the reputation of the consultant (Fauser et al., 2015).

2.2.3 Gap Identification and Requirement Definition

One common aspect of project failure is the so-called scope creep, when the scope of the deliverables grows out of the original agreements, with no corresponding changes in time or resources. In consulting projects, where the requirements are not as tangible as in the technical industry, scope and contractual definition are especially important (Ika, 2009).

2.2.4 Delivery Metrics and Performance Tracking

Successful consulting projects have been known to use well-developed performance tracking systems to keep track of progress against milestones and deliverables. This implies the incorporation of both financial (e.g., margins, utilization rates) and operational (e.g., timely delivery, deliverable quality) metrics in operation monitoring systems (Alias et al., 2014).

2.2.5 Financial Results and Margin

One of the rare and less studied aspects of project success in consulting is profitability. Ajanaku (2025) illustrates that the following variables can be considered strong predictors of the cost and margin of a consulting project: person-months, team size and project duration. Although the key metric of a successful consulting firm is serving the needs of clients, the financial viability and payback of the effort is also considered. Therefore, achieving a balance between client satisfaction and profitability is a core CSF in the consulting engagements.

2.2.6 Top Management Support

In addition to the factors that are project-specific, organizational support has been revealed as a critical success determinant. Reports present very sound empirical evidence that top management support (TMS) in multiple dimensions, such as visible leadership, resources provided, and timely decision-making, have a significant impact on the outcome of projects. In consulting assignments, hierarchies are frequently traversed, and in such cases, the backing of top-level management within the consulting firm and the client organization is essential.

2.2.7 Integrating Social, Technical, and Environmental Contexts

Several researchers claim that CSFs cannot be thought of in isolation but as part of a multidimensional system in which social, technical, and environmental aspects interact. Other works, such as Fauser et al. (2015), report the significance of governance and control techniques, and Gao et al. (2024) discuss contextual factors, including contract incentives and market situations.

2.2.8 Critical Synthesis

The published sources on CSFs in consulting projects appear to indicate that there are several interrelated factors that determine success or failure in consulting projects and they include team capacity, client interaction, scope, financial margin, and top management support. These drivers are more established at a conceptual level, but empirical research in consulting is more fragmented. Other studies are sector-specific (e.g., IT, construction) or examine only one of various dimensions (cost or schedule) and therefore do not have high relevance to the consulting industry. Moreover, even though models like the multidimensional model expand the panorama, there is insufficient integration between financial, operations and human aspects of success.

In this dissertation, the CSF perspective offers the theoretical approach and methodology. By rendering CSFs as things that can be measured, one can test their ability to predict using the tools of machine learning and consequently close the gap between descriptive theory and practical prediction in consulting project management.

2.3 Existing Prediction Frameworks in Project Management

Several frameworks and tools to monitor and evaluate the performance of a project have been developed over the past three decades. Although these frameworks are useful, though, most of them are descriptive or diagnostic in nature, making it difficult to predict success or failure at the start of the project.

2.3.1 Scoring and Strategic Alignment

One of the most popular methods is Balanced Scorecard (BSC) of Kaplan and Norton (1996). The BSC was initially invented as a strategic management instrument that has been transformed to fit project management as a tool of integrating project objectives with those of organizations. The BSC integrates the financial, customer, internal process, and learning perspectives to expand the traditional iron triangle of time, cost, and scope to long-term strategic effect. With consulting projects, this expanded perspective can be useful, since engagements will not only be expected to deliver on an immediate product or service but also meet client satisfaction or relationship-building objectives (Alias et al., 2014). Nevertheless, the use of retrospective indicators does make the BSC more likely to explain past performance than it is to forecast future performance.

2.3.2 PMO Dashboards and Key Performance Indicators

Tracking of performance by using dashboards that show Key Performance Indicators (KPIs) has also become institutionalized by the Project Management Office (PMO). Some of the typical metrics that are consolidated into these dashboards include cost variance, schedule variance, resource utilization and milestone completion percentages (Ika, 2009). Whereas dashboards give insight and accountability, they are still lagging indicators. Risks are not always seen until they materialize by which time they are rendered visible on a dashboard.

2.3.3 Risk Assessment Frameworks

Conventional risk management frameworks are another form of predictive ambition. Probabilistic modelling, Monte Carlo simulations, and expert risk registers are among the techniques aimed at predicting the probability of adverse events (Hillson, 2002). Such techniques are limited by the fact that they are subjective and rely on expert judgement and predefined assumptions. The empirical evidence illustrates that risk registers are unable to give an indication of governance and control weaknesses that eventually sabotage projects (Fauser et al., 2015).

2.3.4 Predictive Project Analytics

One more developed method is Predictive Project Analytics (PPA) that is introduced by Deloitte and is characterized by Fauser et al. (2015). PPA uses historical project data together with statistical prediction to create early warnings of possible failure. In contrast with dashboards or risk registers, PPA focuses on the identification of control gaps in governance associated with an unfavourable outcome. According to case evidence, PPA has increased ability to identify risks in organizations earlier and enhances performance of project delivery. However, it is not widely used and the extent of it is limited by the quality and exhaustiveness of historical datasets. In addition, PPA is more focused on control mechanisms and systematically less accommodative to broader success drivers like client engagement, margins, or strategic alignment.

2.3.5 Critical Synthesis

Current frameworks are useful in learning about project monitoring but cannot be used to predict success in complex consulting settings. Balanced Scorecards and dashboards allow strategic alignment and visibility and are retrospective. Risk management frameworks provide probabilistic judgments but are constrained by subjectivity and range / focus of risks contemplated. Predictive Project Analytics is a significant evolutionary move but fails to include the multi-dimensional analysis of success factors, especially the financial and human variables that are significant in consulting.

Such a gap suggests that increasingly complex, more data-driven solutions are required. With machine learning, one can get beyond the descriptive monitoring to the predictive modelling that considers financial, operational, and human factors at once. This way, foresight in predictive models can give consulting firms insightful information on what to do proactively to allow more sustainable project results.

2.4 Machine Learning and Artificial Intelligence in Project Management

With a new rise of the machine learning (ML), artificial intelligence (AI), there are new prospects in project management, mainly related to augmenting the forecasting and decision-making. Whereas conventional frameworks like dashboards and risk registers offer descriptive or diagnostic information, the ML applications can reveal the hidden results in the historical project evidence and develop predictive knowledge of what could be expected in the future.

2.4.1 Applications Across Sectors

Predictive analytics in project management has been most prominently applied in construction, engineering, and IT project management problems where the volume of the data is large and thus support machine learning-based projects. As an example, Gao et al. (2024) established an up-to-date prediction models framework that automatically adjusts based on the new data collected during the project lifecycle and makes more accurate predictions.

Predictive analytics has been used in IT and software projects to help predict defects, the amount of time it takes to complete, and various risks of the project. The use of ML in agile settings is suggested to show capability of predicting sprint overrun or node congestion prior to any escalation (Menzies et al., 2012). These applications showcase the versatility of ML in different fields, but such applications mainly revolve around fields in which concrete deliverables and organized databases are prevalent.

2.4.2 Applications within Consulting

Research on the topic of consulting projects is less extensive but is increasing. As demonstrated in Ajanaku (2025), ML models like XGBoost and Random Forest in cost prediction of consulting projects can be more accurate than a conventional estimation method. In these models, the key predictive variables have been the team size, the number of persons, and the duration of the projects, and this shows how effort-related variables are important in consulting engagements. Nonetheless, these models are strikingly limited in their scope and though they capture cost prediction as a critical element of project success, other areas of project achievement, including client satisfaction, strategic fit and margins remain unexplored.

2.4.3 Strengths of ML Approaches

The advantages of ML in project management are due to the fact of its capabilities to:

- Work with large, multi-dimensional data and highlight variables of interest in a complex interplay.
- Adapt in a dynamic way as more data is available and so make better predictions over time.
- Automate the process of discovering patterns and lessen the recourse to subjective expert judgement.

2.4.4 Challenges and Limitations

Alongside these advantages, other factors limit the applicability of the ML in project management. The first issue is data availability/quality. Projects are still being treated as individual in nature, resulting in a disjointed dataset that lacks in volume and uniformity necessary to use as input to build a robust model (Fauser et al., 2015).

A second challenge is on model explainability. Most of the ML algorithms, especially those based on use of ensembles and neural networks are black boxes that give correct prediction but without explanations of their working pathways.

Lastly, the studies that have been conducted so far have a limited understanding of success in the form of cost and schedule. Although these are essential, other important dimensions are being neglected, including client satisfaction, quality of governance, and value-creation of strategy (Gao et al., 2024).

2.4.5 Critical Perspective

The literature shows the potential and the shortcomings of ML in project management. Although predictive models have demonstrated relevance in terms of cost and schedule projection, their utilization of the multi-dimensional success measures applicable to consulting project has not yet been well developed.

The dissertation addresses these drawbacks by adopting Critical Success Factors (CSFs), as a theoretical basis to select features. The operationalization of CSFs into measurable variables and the use of ML methods allow the study to produce both accurate and interpretable predictive models and overcome the twofold challenge of performance and transparency in consulting project success prediction.

2.5 Critical Appraisal and Research Gap

The studies above reveal an extensive literature on project success, important success factors, as well as predictive frameworks, but there is still room to fill in the gaps. Together, these shortcomings identify a need that this dissertation aims to fulfil: the lack of an effective model that is machine learning based and fitted to the specific oddities of consulting projects.

2.5.1 Limitations of Existing Definitions and Frameworks

Conventional measures of the success of a project, like the iron triangle of time, cost, and scope (Atkinson, 1999), give an incomplete picture of the results. Even though multi-dimensional models have extended the scope and covered client satisfaction as well as strategic alignment (Shenhar et al., 2001; Jugdev and Muller, 2005), they have not fully addressed operational issues of consulting environment. The consulting projects are especially dependent on intangible deliverables, trust-based partnerships, and financial feasibility, but these aspects are not included in the popular success models (Alias et al., 2014). Equally, monitoring tools like Balanced Scorecards and PMO dashboards enhance visibility and alignment although they are fundamentally backward looking with limited predictive capability (Kaplan and Norton, 1996).

2.5.2 Gaps in Understanding Success Factors

The Critical Success Factors (CSF) approach has solid theoretical background, but empirical studies on consulting projects are disintegrated. Some of their traditional CSFs, like planning, communication, and leadership (Pinto and Slevin, 1987; Ika, 2009) have been well established, however, consultancy specific factors, like client engagement, scope clarity, and financial margins are less tested in predictive models.

2.5.3 Shortcomings of Predictive Approaches

Emergent predictive models like Predictive Project Analytics (Fauser et al., 2015) are an improvement because they use statistical means to anticipate governance-related risks. Nevertheless, these strategies are limited in their abilities and focus on control mechanisms instead of lagging behind wider success drivers' profitability, the satisfaction of the clients and overall strategic accomplishments. The applicability of ML to project management, however, has not yet reached its potential, with most applications limited to cost or schedule forecasting (Gao et al., 2024; Ajanaku, 2025).

Another shortcoming is the question of explainability. Most ML models especially ensemble techniques and neural networks are black box models, and moreover they provide no transparent answers on the drivers. Predictions without interpretable input stand the danger of being discredited by being labelled as technically advanced but unusable in applications.

2.5.4 Research Gap

All the literature identifies three gaps an interrelation between them. To begin with, consulting project success is not measured in a holistic manner that would consider all three components, namely, financial, operational, and human, of a consulting project success in a coherent predictive model. Second, although the concept of CSFs is well-developed, it has not been operationalized in a systematic manner with regard to the context of machine learning.

Third, the need to have explainable ML models that balance the predictive accuracy with interpretability is lacking, an aspect that is a prerequisite in consulting businesses where both predictions and managerial legitimacy are necessary.

2.5.5 Contribution of this Dissertation

This dissertation attempts to fill these gaps, proposing a machine learning based framework to predict consulting project success, using the Critical Success Factors (CSF) approach. The research will determine the predictive nature of CSFs when grouped together with financial and operational measures. By doing so, the study not only advances academic knowledge regarding the success of projects, but also offers consulting firms practical, interpretive measurements of proactive monitoring and decision-making on projects.

Chapter 3; Contextual Background

3.1 The Consulting Industry and Project Success

Based on the literature review that identified the multidimensional nature of project success and the specific challenges of consulting engagement, the current study places itself in the consulting industry. Consulting projects are unlike technical or engineering projects because they often have intangible outputs such as organisational change, process redesign, or capacity building. These results are closely associated with the level of expected cooperation between consultants and clients, their expectation matching, and the capacity to carry on changes even after the project is completed.

Success in consulting projects cannot be distilled to the classical iron triangle of cost, time, and scope as discovered in Chapter 2. Efficiency is still imperative but effectiveness, client satisfaction, strategic alignment as well as financial viability are also decisive. In the case of consulting firms, a key indicator is margins and client retention, both of which imply short-term and long-term value. These complications also highlight the importance of analysing consulting projects through a framework that can capture both quantitative measures and qualitative dynamics.

3.2 Characteristics of Projects in Context (Clarasys Dataset)

In establishing this study, this research relies on a dataset of consulting projects offered by Clarasys. This data constitutes a wide cross-section of interactions, including financial, operational, and human aspects of delivery. It contains variables including:

- Monetary indicators: baseline income, margin rates, project costs, rates of utilisation.

- Operation descriptors: type of activity assignment, model of delivery engagement, demand predictions.
- Human, relational aspects: candidate status, staffing arrangements, and indicators of client engagement.

Combined, these aspects demonstrate the Critical Success Factors established in the literature, namely, financial viability, team dynamics, and client involvement but framed here in an empirical manner. The dataset thus offers an interface between theory and practice, which enables these dimensions to be systematically analysed.

The key feature of the dataset is that it represents project outcomes in two ways. The Success Score is a continuous measure providing gradations of performance both in the financial and delivery dimensions. Success Flag is binary (defines project as successful or not successful) based upon predetermined thresholds. This dichotomy reflects the strategic dilemma observed in the literature: managers need both subtle measures of performance intensity and a simplified signal to make their decisions.

Similar to most real-life project repositories, the dataset has flaws as well. Some variables, like forecasts of demand or candidate status, are missing, others face definitional ambiguity. As an example, the quality of engagement can have different meanings depending on a client and the situation.

3.3 Challenges in Defining and Measuring Success

These two expressions of success in the dataset mirror the wider conceptual challenges discussed in Chapter 2. The continuous Success Score embraces variability and makes the analysis finer grained. It helps to distinguish those projects which significantly surpassed expectations and the ones that did so only slightly. Nevertheless, such richness is achieved at the expense of interpretation. Its formulation is based on the weighting of various KPIs, some of which might favour financial performance at the expense of relational performance.

By contrast, the binary Success Flag can be instantly interpreted, and it fits executive decision-making: projects are either successful or fail. However, such clarity is achieved at the cost of nuance where complex realities get reduced to one dichotomy. Projects which might have provided some strategic value but failed to meet profitability goals, such as those, will also be classified as complete failures.

This duality reflects the fundamental difficulty of consulting project research: success is multi-dimensional and contentious. The outcomes cannot be summarized by a single indicator. This is why it is useful that the dataset gives both measures. It allows us to model success in two complementary ways, in line with the realities of consulting practice where a leader may require both fine-grained predictions and binary early warning indicators.

3.4 Rationale for Predictive Modelling

In this context, the predictive modelling can be used to counter the shortcomings of conventional monitoring systems mentioned above. Dashboards and scorecards are backward-looking because they reflect problems when they already take place.

This dissertation's modelling strategy leverages the two success measures of the dataset. By using Success Score as a continuous variable, regression models can be used to predict the level of success, which can be defined as how well a project is likely to perform. Using the Success Flag as a categorical target and applying classification methods like logistic regression, we obtain binary predictions representing early warning signals for the risk of failure. This two-track solution not only is technically feasible but is tactically harmonious with consulting practice that calls for both sophisticated performance forecasts as well as simple decision indicators.

From a strategic point of view, predictive modelling offers three benefits. First, it helps with proactive intervention, so that managers can spot at-risk projects sooner in their life cycle. Second, it takes many different aspects of success -- financial, operational, and human -- and combines them into one predictive model, filling gaps left by current descriptive tools. Third, it has granular forecasts that can be useful for analysts and binary risk signals useful for executives.

Hence, predictive modelling is not simply conceptualised as a technical exercise, but as a strategic capability. It builds on the insights of the literature on Critical Success Factors by operationalising them as quantitative predictors and addresses the research gap identified in Chapter 2 by using machine learning on the multi-dimensional notion of consulting project success.

Chapter 4; Methodology

4.1 Research Design

This chapter, based on the conceptual background laid out in Chapters 2 and 3, will outline the methodological framework of modelling consulting project success, using machine learning techniques. As illustrated in the literature, consulting project success is not outlined by mere efficiency measures of time, and cost, but effectiveness, client satisfaction as well as financial feasibility. This study therefore uses techniques that can both represent non-linear relations among variables and also produce both subtle forecasts as well as decision signals.

To solve this, the research strategy takes a two-fold modelling approach. The first strand considers project outcomes as a continuous variable, which is operationalised using the composite Success Score. Regression models are used to predict levels of success, therefore making it possible to distinguish projects that perform poorly, those that perform averagely, and those that do exceptionally well. The second strand redefines success as a binary event, which is represented as the Success Flag.

A stronger evaluation framework is used in this study. The regression strand uses a mixture of 80/20 hold-out test split, 5-Fold Cross-Validation (CV), and out-of-fold (OOF) diagnostics. Likewise, the classification strand relies on stratified hold-out split plus StratifiedKFold CV to maintain distributions of classes across folds. The methodological design of this nature guarantees that findings do not rely on the composition of one set of tests but rather on the stability and generalisability to different partitions of the data.

4.2 Data Preparation

4.2.1 Initial Cleaning

There was first a standard cleaning of all datasets to create consistency before they were split into two divergent classes of continuous and binary strands. The normalisation of column names to lower case, removal of whitespace and rearrangement of underscores to prevent parsing anomalies. The role, capability type and engagement stage text fields were harmonised to similar case formatting. Any missing values in important categorical columns were filled in with interpretable defaults (e.g., "unassigned" when user IDs were missing, "Unknown Role" when roles were undefined). In the assignmentdemandstatus field, there was one record that had a missing value. Since the distribution was heavily skewed with most of the cases being Fulfilled (354 out of 593), mode imputation was used. This was necessary to make sure that the missing entry was attributed to the most representative category, to maintain interpretability but not needlessly drop a row. Duplicate rows were eliminated to eliminate artificial inflation of signals.

4.2.2 Branching Rationale

Based on this shared foundation, the preparation process differed based on the modelling purpose. In the case of the continuous strand, in which the composite Success Score necessitated the incorporation of financial, tenure, and human results, a methodical Join

Discovery and Integration process was applied to develop only a single, combined base of analysis. In the binary strand, the emphasis was rather on generating engagement-level aggregates directly off source tables. This methodology offered small, interpretable features aligned to information provided to project managers in-flight, and minimized concept overlap with the composite score.

4.3 Continuous Strand

A systematic integration process was used to compile the set of features used in the Success Score. Join candidates were found by matching normalised column headers between datasets and calculating overlap statistics on a value level. To quantify similarity using Jaccard indices and coverage measures were used. To illustrate, there was a high overlap between capability type in the activity assignment and capability datasets and between record ID in the engagement resources and company-wide tables.

Using these diagnostics, joins were compiled with a coverage-sensitive policy. Precision was ensured by joining near-complete overlaps (≥99% coverage) with inner joins and keeping partial overlaps with left joins. To reduce the duplication of rows, unique keys were implemented before the merging. This generated consolidated datasets which merged assignment-level, engagement-level and organisational-level data into a unified base to be analysed.

Data Pair(key)	Overlap	Coverage A to B	Coverage B to A	Jaccard	Join Policy
activity_assignment. capability_type ↔ capabilities_with_user_id.capability_type	87	1.0	0.5878	0.5878	Inner (>=99% coverage)
engagements_resources.baseline_usage ↔ engagement_data.baseline_usage	356	0.5274	0.4456	0.3184	Left (preserve partial overlap)
activity_assignment.user_id ↔ capabilities_with_user_id.user_id	56	0.9492	0.2581	0.2545	Left (preserve partial overlap)

engagements_resources.record_id ↔ wholeCompany.record_id	11227	0.2415	1.0	0.2415	Left (preserve partial overlap)
engagements_resources.activity_assignment ↔ wholeCompany.activity_assignment	1395	0.5345	0.1106	0.1009	Left (preserve partial overlap)

Table 1: Insights on joins for analytical table

4.3.1 Missingness Profiling and Pruning

After integration the dataset was checked against a missing data audit. The absolute number and the percent of missing values were calculated, per variable. The findings found a heterogeneous presentation: the baseline financial indicators as well as utilisation rates were complete, whereas engagement-related (e.g., delivery status, risk level) and demand-related (e.g., proposals, required capabilities) indicators had a higher degree of missingness (over 90 percent).

A 50% missingness threshold was used to provide a balance between richness and usability of the data. Variables that had over half of their observations missing were omitted. This practical rule kept strong predictors and dropped off columns with few cases that were unlikely to carry useful signal. The filtered dataset served as a basis of feature engineering and target construction.

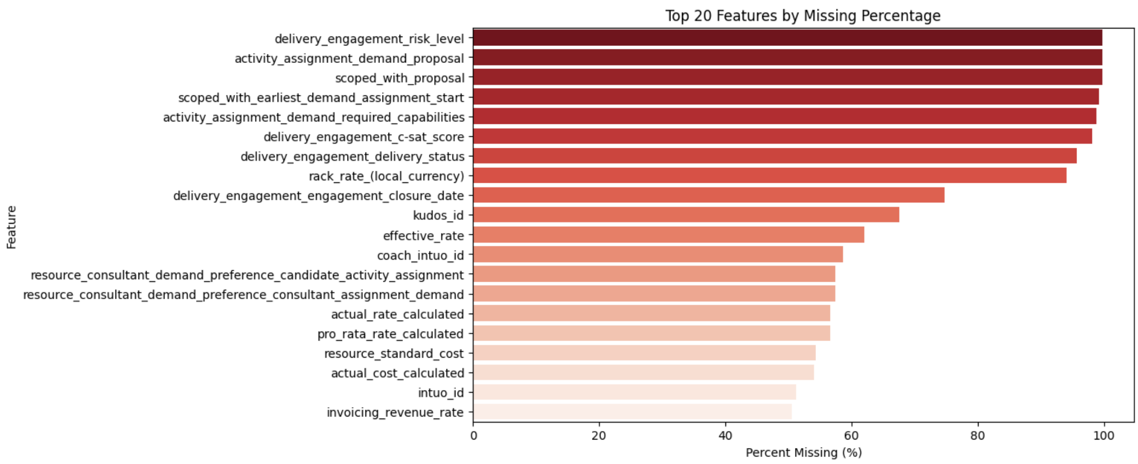

Figure 1: A Summary of top 20 features missing by percentage for context

4.3.2 Target Variable: Success Score

The Success Score, which is the first target variable, was modelled as a composite variable to reflect the multi-dimensionality of consulting project success. It was built in a systematic way considering a combination of focused imputation, diagnostic evaluation of financial indicators, and weighted integration across the dimensions.

4.3.3 Imputation of tenure-related variables

To prevent data loss caused by missing values, specific imputation strategies have been used. The mean was used to impute the variable years to promotion because it captures the average career path of consultants, whereas the median was used to impute the years since job start to minimize the impact of extreme values. This made tenure signals complete and represented with no distortive distributional properties.

4.3.4 Financial metric screening and anchor-based diagnostics.

An extensive scan of candidate financial fields was also performed to test the completeness, sparsity, and redundancy of the data. An anchor variable (assignment margin percentage) was selected to act as a reference and all the other fields related to margin were calculated against this variable. Diagnostic statistics (non-null percentage, zero proportion, mean and standard deviation) were computed in each case of candidate. Correlations with the anchor were then employed to prioritise both substantively covered and highly aligned variables. As an example, assignment_margin_percentage and margin_pct were perfectly correlated, which proves redundancy and variables with an extremely wide sparse coverage (such as assignmentpercentage with >99% 0s) were omitted.

	non_null_pct	zero_pct	mean	std	corr_with_anchor
delivery_engagement_contract_margin_amount	1.0	0.036	36225.87	236824.74	0.005344
margin_amount	1.0	0.067	4635.83	16264.70	0.006754
delivery_engagement_contract_margin_pct	1.0	0.156	12711.93	197885.54	0.001247
delivery_engagement_services_contract_margin_amount	1.0	0.164	31333.43	235680.80	0.002628
delivery_engagement_services_contract_margin_pct	1.0	0.272	-12.29	1368.84	-0.000170
delivery_engagement_baseline_contract_margin_amount	1.0	0.390	36141.05	92704.32	0.014059
assignment_margin_percentage	1.0	0.404	-3395.17	174907.78	1.000000
margin_pct	1.0	0.404	-3395.17	174907.78	1.000000
delivery_engagement_baseline_contract_margin_pct	1.0	0.483	-3293.18	169304.66	1.000000
baseline_margin_amount	1.0	0.957	11.58	4758.71	0.000066
baseline_margin_pct	1.0	0.984	0.72	5.78	0.002445
assignmentpercentage	1.0	0.996	0.29	5.24	0.001088

Figure 2: Profiling of margin variables

4.3.5 Integration of multidimensional components.

Based on this diagnostic step, four dimensions were included in the composite Success Score:

- Financial: realised margins (assignmentmarginpercentage, marginpct, marginamount), planned margins (deliveryengagementbaselinecontractmarginamount) and actual margins by engagement (deliveryengagementcontractmarginamount).
- Efficiency: utilisation rates and realisation rates at the engagement level.
- Human outcomes: binary indicators of whether a consultant has left the firm or received promotion in the course of the engagement, coded 0/1.
- Tenure and progression: imputed years to promotion and years of job start.

4.3.6 Pre-processing and scaling.

Binary variables were numerically coded, and all continuous variables were standardised to a range of 0-1 through Min-Max normalisation such that variables with differing magnitudes can be compared to each other. This was a significant move towards equalizing the role of financial values (which are often many orders of magnitude greater) and human and tenure variables.

4.3.7 Weighted composite construction.

The Success Score was then calculated as a weighted average of the four items 40 percent financial, 30 percent efficiency, 20 percent human and 10 percent tenure. This weighting framework embodied the primacy of financial sustainability in consulting projects, but also included delivery efficiency, consultant performance and career advancement as part of long-term performance.

4.3.8 Regression Modelling Approach

To avoid leakage of targets, all variables that are directly incorporated in the formation of the Success Score (financial, efficiency, human, and tenure factors) were removed in the set of predictors, as well as identifiers, including IDs, codes, and reference keys. The rest (quantitative and nominal) of the dataset was coded through one-hot encoding. This resulted in a high-dimensional and sparse design matrix containing 19,744 predictors on 1,215 observations following combined deletion of missing values in predictors and target. Since one-hot encoding added special characters in column names (e.g., [,], <), a sanitisation step was introduced to make it compatible with XGBoost but provided a mapping to the original names to interpret features.

Random Forest Regressor was chosen as a baseline model. The model was configured with 300 estimators and a maximum depth of 12 to offer the balance between predictive power and computational efficiency. Random forests were especially favourable to this problem domain, since they are resistant to multicollinearity, can capture complex non-linear interactions, and provide easy-to-interpret metrics of feature importance.

XGBoost was also applied to benchmark against a gradient boosting approach. The model was set to 2,000 boosting rounds, learning rate of 0.05, maximum depth of 6, subsampling rate of 0.9, column sampling rate of 0.9, and L2 regularisation. Such a set-up enabled XGBoost to capture finer-grained interactions whilst regularising to prevent overfitting.

4.3.9 Evaluation Strategy

To provide robustness and generalisability, the regression strand used a three-part evaluation framework:

1. Hold-out split (80/20): A baseline level of generalisability was measured by random Forest and XGBoost using R2 and RMSE on a fixed hold-out set.
2. 5-Fold Cross-Validation: The models were retrained on five shuffled folds and R2 and RMSE were compared as mean +- standard deviation to assess model stability. XGBoost was trained on the entire schedule of boosting and feature importances were averaged across folds to generate consistent rankings.
3. Out-of-Fold (OOF) diagnostics: With crossvalpredict, unbiased predictions were obtained on all records then compared with observed scores, resulting in a series of scatterplots and residual analyses that demonstrate systematic under- or over-estimation.

Collectively, these approaches struck a balance between single-split benchmarking and cross-validated stability and dataset-wide diagnostics. Random Forest provided interpretability and robustness and XGBoost provided better generalisation by modelling more subtle interactions.

4.4 Binary Strand

4.4.1 Target Variable: Success Flag

The study operationalises a binary outcome variable-Success Flag- to supplement the continuous Success Score that conforms to the executive decision requirements of early warning signals that are clear. This flag is based on the forecast status of the engagement, a management-facing metric common in the governance of consulting delivery. In particular, the statuses Firm (100%), Probable (90%), and Probable (60%) are associated with success (1): the rest of statuses are associated with non-success (0). Since an engagement may include many rows within source tables, the modal (most frequent) status per engagement is employed to establish the flag, which is resistant to transient updates and noisy records.

4.4.2 Feature Engineering

In the classification ABT (Analytical Base Table), the study engineers project-level aggregates that highlight team structure, human capital, resource utilisation, work-load scope, financial indicators, and operational cadence:

4.4.2.1 Team composition & diversity

Based on the activity assignment dataset, 3 engagement-level descriptors are created:

(i) Team size (unique user IDs),

(ii) Role diversity (role differences), and

(iii) Diversity in terms of activity/capability (different types of capability). These are the ways of breadth and mix of skills that correlate with success--proven CSFs in consulting situations.

4.4.3 Human capital depth

Out of the capabilities dataset, the number of unique capabilities per consultant is calculated and averaged at the engagement level (avgcapabilitiesperteammember). This estimates both breadth of collective skill and flexibility of deployment.

4.4.4 Resource utilisation intensity

With the resource's dataset, the number of different resources utilized on a (engagement, user) basis is determined and averaged across the team (avgresourcesusedperuser). This is an indication of the complexity of allocations and switching costs.

4.4.5 Workload scope & financial signal

Two resources-level indicators are included in the project: assignmentcount (number of distinct assignments to one project) as workload proxy, and avgassignmentmargin (mean assignment margin percentage) as an early financial performance indicator.

4.4.6 Temporal & cadence proxies

Based on the logs of engagement, the first proposaldeliverystart (projectstartdate) and the number of logs (totalengagementlogs) is obtained. These serve as crude proxies to timeline start and operational touchpoints. (In baseline models, dates are excluded to prevent time parsing but are included in case of future ablations).

Prior to modelling, common hygiene methods are used: column name normalisation, removal of redundant records, and pruning of non-modelling records (e.g., raw engagement IDs, unparsed date columns). The target is imposed through a single, integer-coded successflag to avoid schema drift.

This set has been chosen deliberately orthogonal to the continuous-target components to prevent leakage between strands. It puts much more emphasis on pre-delivery and in-flight indicators that can be realistically known by a PMO when risk calls are made to enhance external validity.

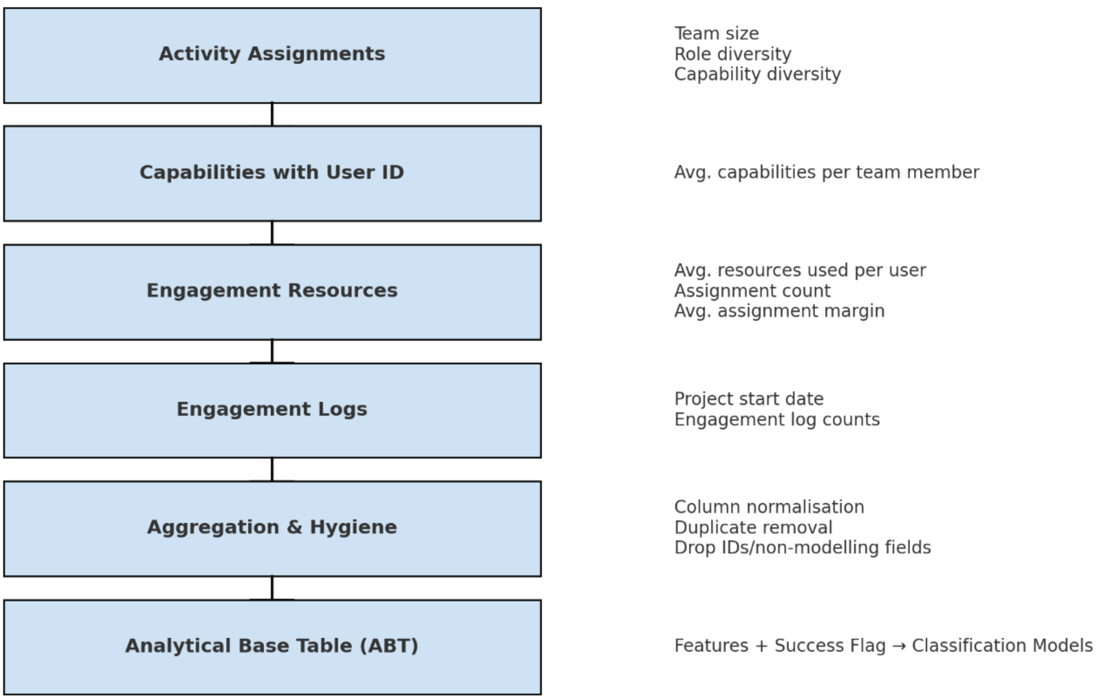

Figure 3: Workflow for constructing the Analytical Base Table (ABT)

4.4.7 Classification Modelling Approach

Two complementary pipelines were applied. The former used Logistic Regression with median imputation, standardisation and logistic regression implemented in one pipeline. Numeric variables with missing values were filled in with median imputation in a robust manner to skewed distributions and outliers. This arrangement gave a linear baseline that could be interpreted against which more flexible non-linear methods could be assessed.

The second model used a Gradient Boosting Classifier with a single pre-processing stage of median imputation. Gradient Boosting does not require scaling of features as well because it uses threshold splits and not distance measures, unlike linear models. Missing data were imputed using median, which felt very resistant to skew and outliers (which often dominated project-level data, e.g., team size and financial margins). The combination allowed the model to be able to represent non-linear interactions, while still being methodologically simple, interpretable, and resistant to irregular data distributions.

4.4.8 Evaluation Strategy

Evaluation of the binary strand was conducted using stratified hold-out split (65% train, 35% test) with StratifiedKFold CV on the training set in order to preserve the class balance. Performance was measured in terms of ROC-AUC, Average precision (PR-AUC), Accuracy, and F1. The training set also yielded out-of-fold predictions which allowed the calculation of the classification reports and overall AUCs. An external validation of generalization was obtained using a held-out test set.

Chapter 5; Results

5.1 EDA Highlights

The exploratory data analysis (EDA) offers a first impression of the Clarasys project data, both in terms of structural trends and possible predictors of success.

5.1.1 Correlation Analysis of Success Score

The most successful correlates of project success were financial indicators. The central role of realised revenues and profitability was confirmed by the dominant role of revenue rate ($r = 0.64$), contract revenue ($r = 0.63$), and invoiced revenue ($r = 0.62$). There were also high intercorrelations among these variables ($r > 0.95$) indicating redundancy in the feature space of financial features. The tenure-related factors including years since job start and years to promotion (r = 0.52 each) were moderately related positively, indicating that the stability of consultants supports results. In general, the analysis supports the multidimensional nature of the Success Score, where financial viability is the core factor to be accompanied by tenure indicators.

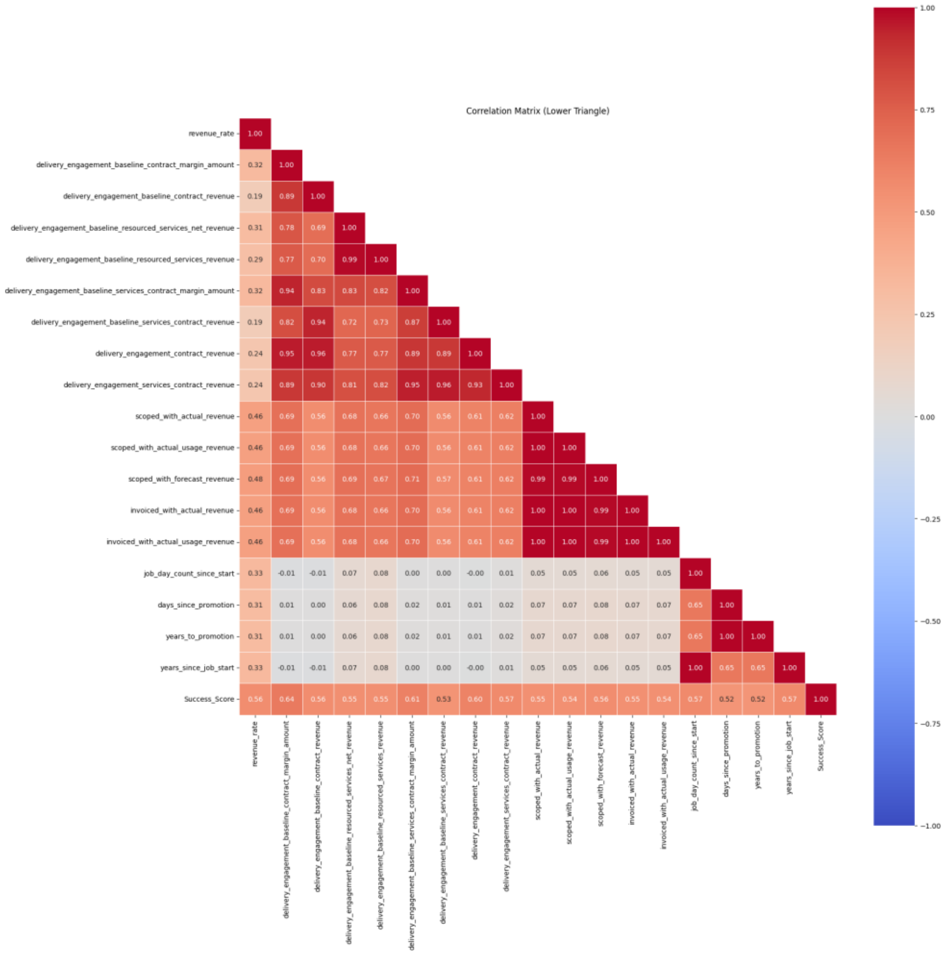

Figure 4: Correlation matrix of financial, utilisation, and tenure variables with Success Score

5.1.2 Distribution of Success Score

Success Score distribution is tightly concentrated in the range between 0.40 and 0.44 with the maximum of the distribution being close to 0.42 with a slight right skew that extends to approximately 0.52. This implies that most projects demonstrate relatively moderate-to-high performance with a smaller group performing better results. The stable distribution both facilitates its application as a continuous measure, and the derivation of a binary Success Flag to classify.

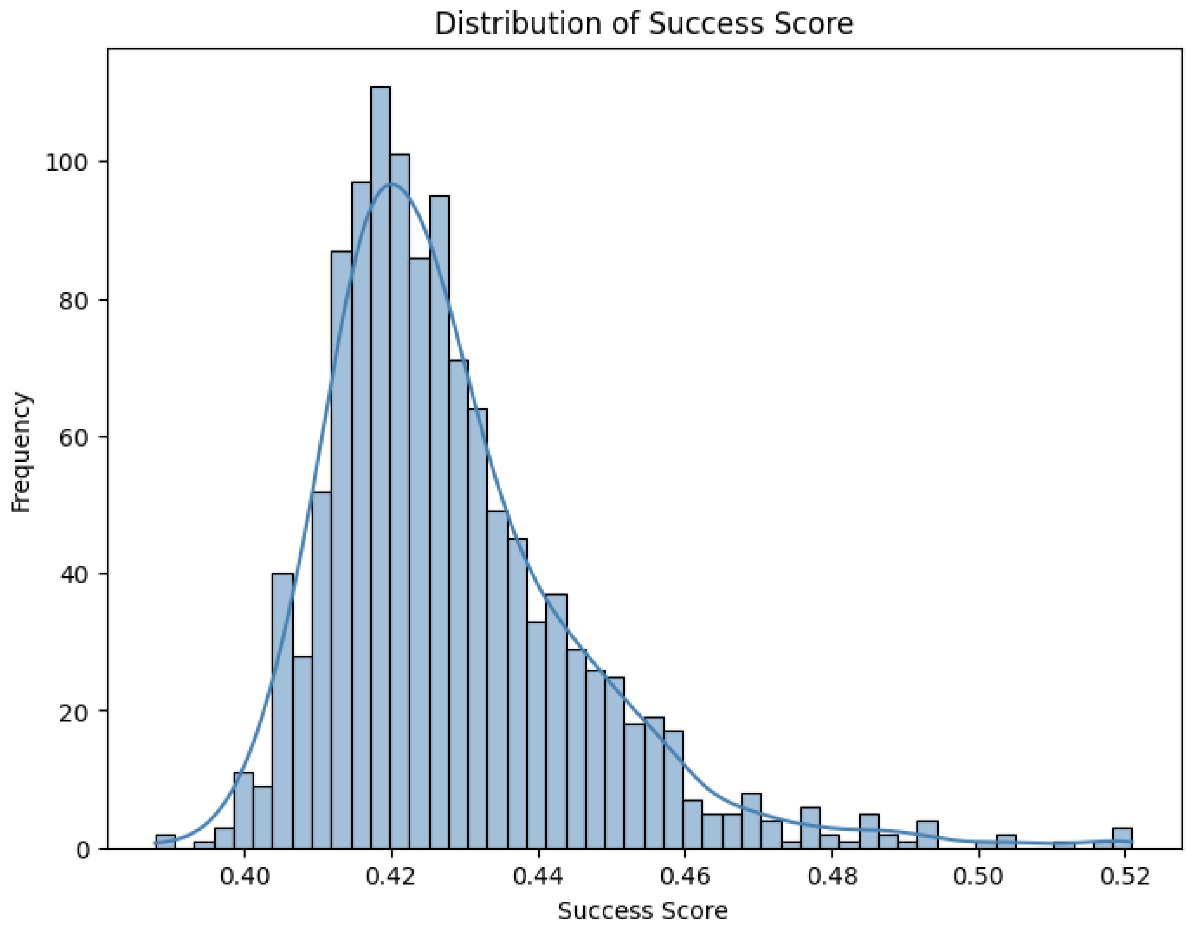

Figure 5: Distribution of Success Score across consulting projects

5.1.3 Team and Diversity Features

Figure 5.3 indicates that successful projects tend to be associated with more team size, justifying the consideration that more resources can help to increase delivery capacity. Such findings are consistent with CSF literature on capability breadth and role complementarity but show that diversity is not a sufficient factor but rather a facilitating factor in concert with other project management practices.

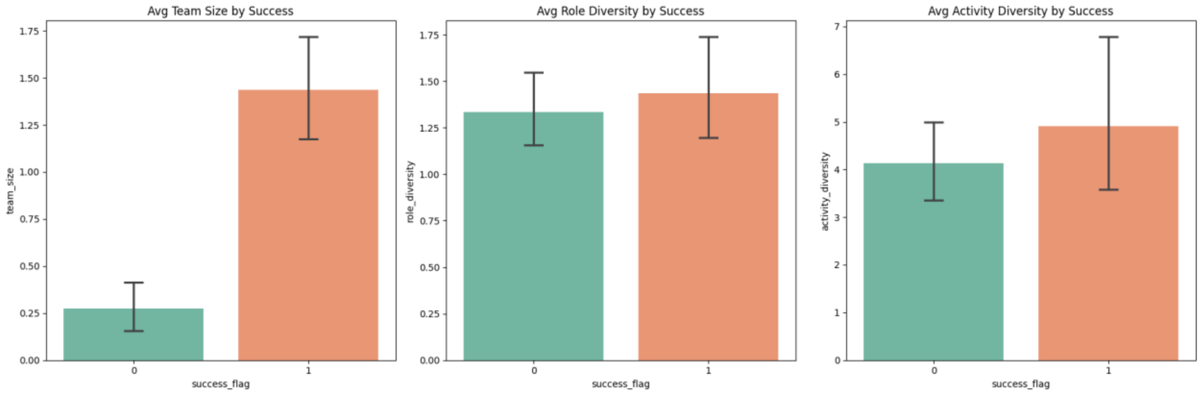

Figure 6: Team size, role diversity, and activity diversity by Success Flag.

5.1.4 Capability Breadth per Consultant

The average capabilities per team member are depicted in Figure 5.4. The distribution is long tailed: the average number of distinct capabilities per consultant is less than 20, although some have portfolios of very wide capabilities (>50). Such high capability consultants can have disproportionately significant roles on some projects. Although the correlation with success is not as direct, this characteristic supplement team diversity by representing the versatility of individual consultants.

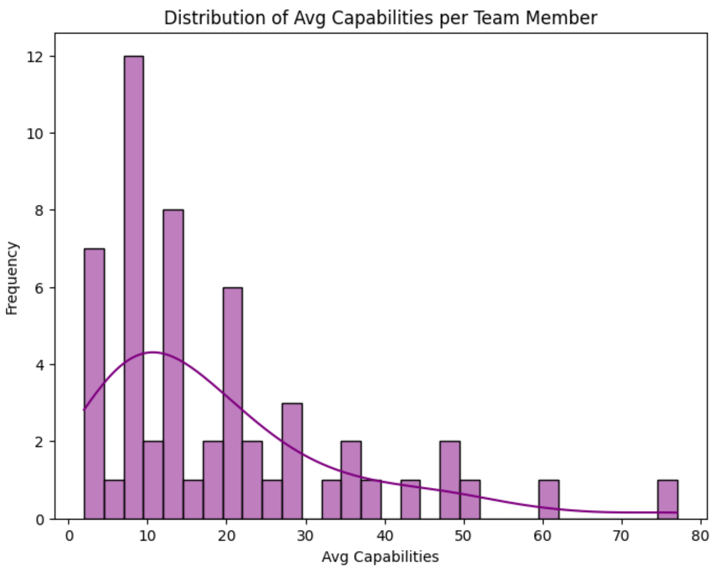

Figure 7: Distribution of average consultant capability breadth per project

5.2 Feature Engineering Outcomes

The feature engineering phase converted raw assignment, capability, and resource data into project level aggregates. The financial and structural signals, which appeared are depicted in Figure 5.5.

5.2.1 Assignment Margins by Success Flag

As indicated in figure 5.5, the median assignment margin in successful projects is a little bit higher than in unsuccessful projects, implying that being profitable leads to successful projects. The difference however is small and successful projects have a vast range of difference, ranging between very low and extremely high margins.

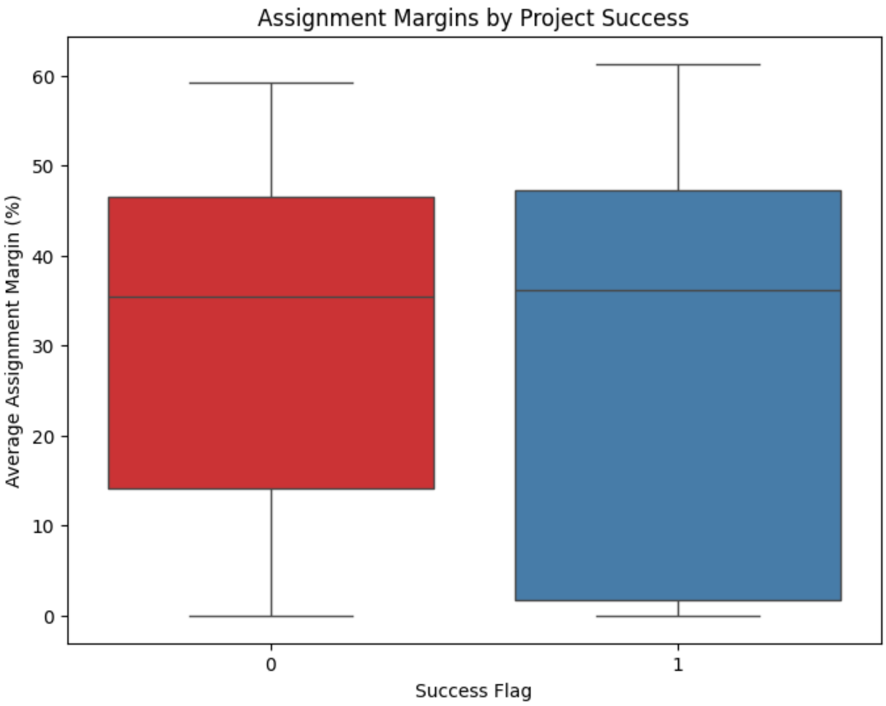

Figure 8: Average assignment margins by project success outcome.

5.2.2 Correlation with Success Flag

The binary correlation analysis (Figure 5.6) established that the strongest structural signal (r = 0.62) is the team size in which bigger projects have more chances to succeed. The assignment count (r = 0.19) and activity diversity (r = 0.09) yielded weaker positive correlations, and the role diversity effect was negligible (r = 0.06). Other characteristics, including capabilities per team member (r = -0.03) and assignment margin (r = -0.09) had weak or negative correlations, which indicated less prediction potential when used alone. On the whole, team size was pre-eminent with other features probably playing a role only in a multivariate model.

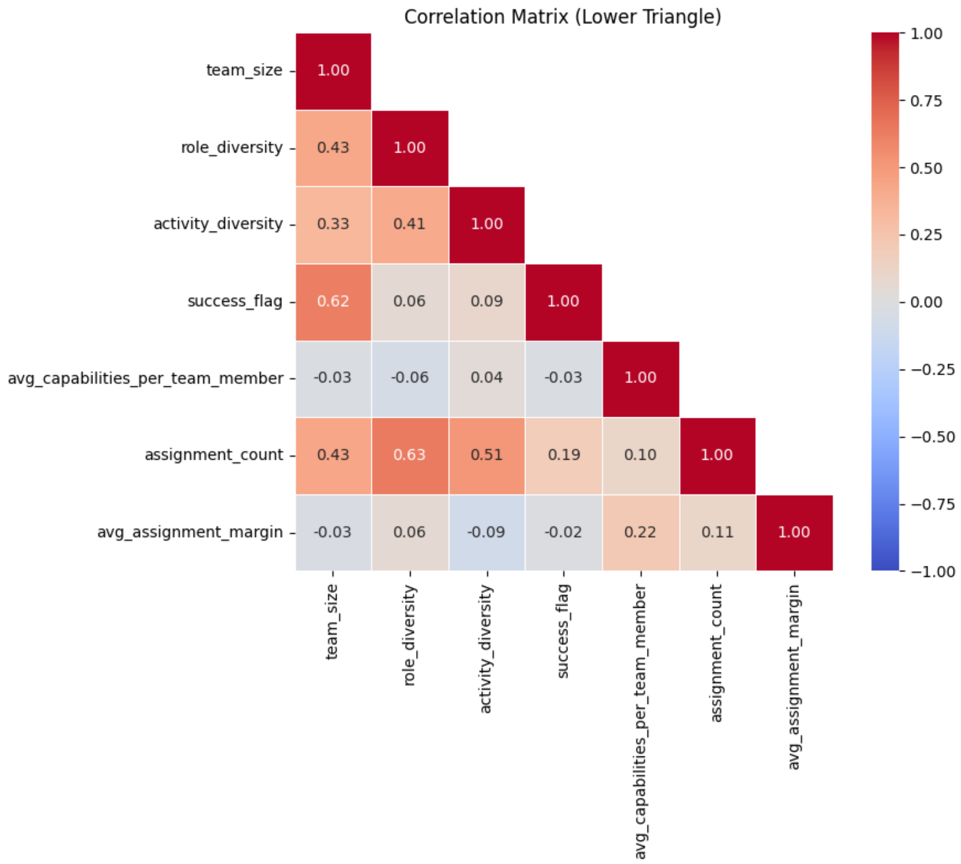

Figure 9: Correlation matrix of engineered feature with Success Flag

5.3 Model Performance

5.3.1 Regression (Success Score)

Random Forest had explained an average of 87 percent of the variation in the success of the projects in the hold-out set ($R^2 = 0.865$, RMSE = 0.0065) and was consistent in 5-fold CV ($R^2 = 0.873 \pm 0.066$, RMSE = 0.0062 ± 0.0017) and OOF evaluation ($R^2 = 0.871$, RMSE = 0.0064). Predictions were very close to observed scores and there was only slight dispersion at the upper tail (Figure 5.3).

XGBoost underperformed on the single hold-out ($R^2 = 0.832$, RMSE = 0.0073), but outperformed the Random Forest on the cross-validation ($R^2 = 0.904 \pm 0.057$, RMSE = 0.0053 ± 0.0018) and OOF ($R^2 = 0.904$, RMSE = 0.0055). This is an indication of greater generalisation of boosted ensembles even though they are sensitive to single-split composition.

In general, Simple partitioning was more stable with Random Forest, whereas XGBoost was able to model more subtle interactions and performed better at generalisation across folds.

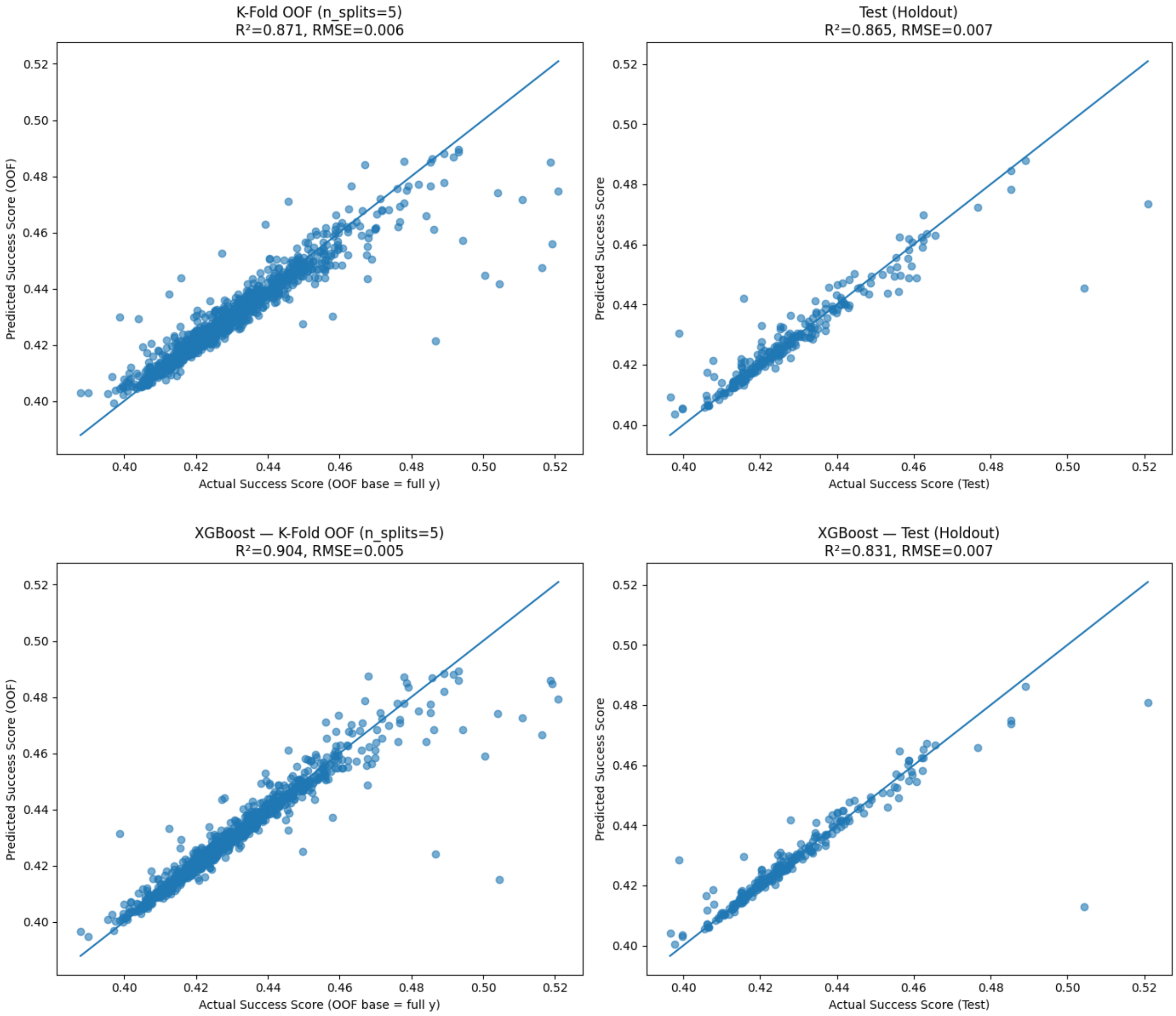

Figure 10: Predicted vs. Actual Success Score for XGBoost (Out-of-fold and Holdout Text Results).

5.3.1.1 Feature Importance Analysis

The outcomes of the feature importance showed the drivers of success on the project in the regression strand.

In the case of Random Forest, job tenure (jobdaycountsincestart) had the biggest influence (it could explain approximately 31%), then came the revenues of the contracts (17%), margin of the proposal (8%), and the revenues forecasted (7%). This means that the strongest predictors were project tenure, size of revenue, and contractual margins.

In the case of XGBoost, the importance was distributed evenly. Once again, contract revenue was of the highest rank (≈19%), but other finer-grained assignment-level and temporal variables (e.g., activity demand codes, scoped forecasted revenue) were also given priority by the model.

Comparatively, both models also constantly focused on contract revenues and margins, and this proves to be very critical. Random Forest preferred global structural indicators such as tenure, but XGBoost found operational information such as assignment-specific indicators and promotion changes.

This difference corresponds to the philosophy of the models: Random Forest emphasizes the predominant structural drivers, and XGBoost emphasizes the less obvious but interaction-based features.

5.3.2 Classification (Success Flag)

Cross-validation Logistic Regression reached mean ROC-AUC=0.850 (±0.085) and PR-AUC=0.832 (±0.116), with an accuracy and F1 score of about 0.82. Accuracy on the held-out test set was 0.765 with high recall on successful projects (0.94) and lower accuracy on unsuccessful projects (0.61). ROC-AUC and PR-AUC were 0.799 and 0.798, respectively.

Gradient Boosting was significantly higher and achieved 0.905 (±0.133) and PR-AUC of 0.911 (±0.122) in cross-validation. It was as accurate on the test set as Logistic Regression (0.765) but had more evenly distributed class-wise results, ROC-AUC 0.830, PR-AUC 0.844.

ROC and PR curve visualisations ensured the same behaviour of the model, with Gradient Boosting tending to yield better discrimination.

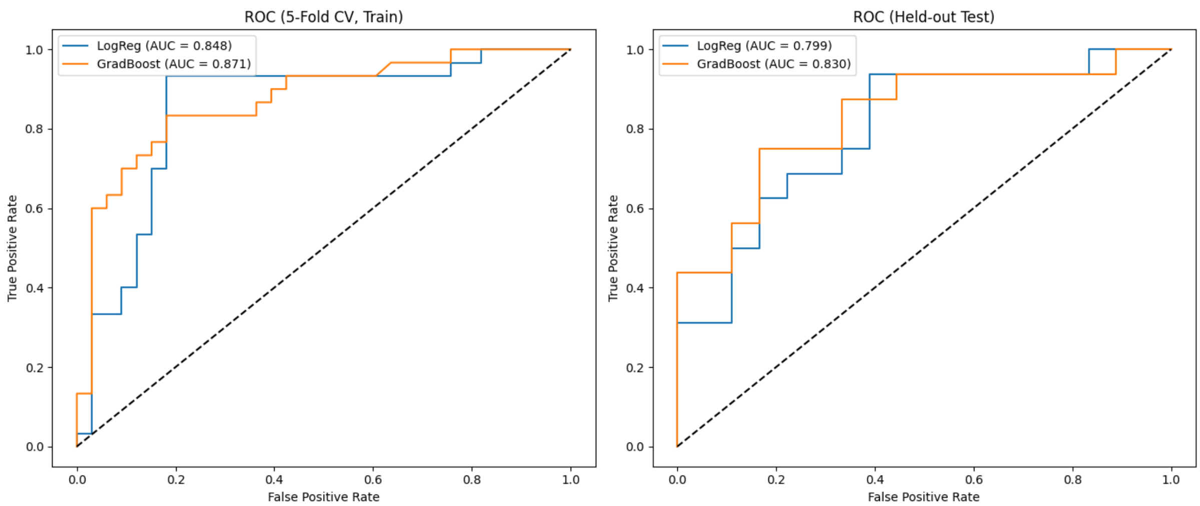

Figure 11: Precision-Recall curves for Logistic Regression and Gradient Boosting (5-Fold CV and Held-out Test)

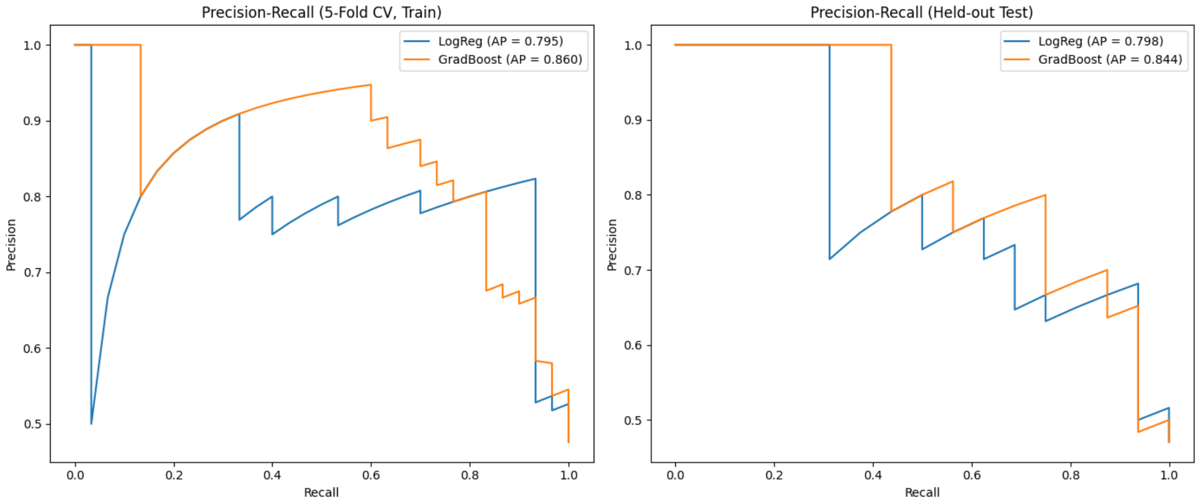

Figure 12: ROC curves for Logistic Regression and Gradient Boosting (5-fold CV and Held-out Test)

5.3.2.1 Feature Importance Analysis

The analysis of feature importance has identified the drivers of project success in the classification strand.

Team size was the most significant factor predictor (β ≈ 1.97) in the case of logistic regression, with bigger teams having a strong positive effect on success probability. Others that contributed positively were assignment margins, capabilities per team member, and the number of assignments but role diversity (β ≈ -0.48), and activity diversity (β ≈ -0.13) were negatively correlated which indicated that there was trouble in coordinating highly heterogeneous teams.

In gradient boosting, team size was again most influential (importance ≈ 0.56), then diversity of activities, number of assignments, and assignment margins. In contrast to logistic regression, boosting assumed diversity as a context-dependent factor in non-linear thresholds and interactions.

In all two models, the size of the team was continually identified as the most effective success factor, and financial margins have an effect. Divisions focused on diversity, as linear models focused on negative influences and increased the capturing of more subtle positions.

Combined, this indicates that a bigger, financially efficient, and capability-balanced team is systematically more successful, and the role of diversity is conditional upon circumstance.

Chapter 6; Discussion and Conclusion

6.1 Qualitative Evaluation

The findings validate that consulting project success is a multidimensional concept that can be successfully modeled with the help of machine learning. These two financial criteria and human or structural variables proved to be important as well as in line with the Critical Success Factor (CSF) literature.

Random Forest in the regression strand accounted for ~87 percent of the variance of Success Score on hold-out tests ($R^2 = 0.865$, RMSE = 0.0065), with consistent cross-validation and OOF. XGBoost was weaker in the single split ($R^2 = 0.832$, RMSE = 0.0073) but, in cross-validation ($R^2 = 0.904$, RMSE = 0.0053) and OOF evaluation, it showed better generalisation. These differences were also mirrored in feature analysis: Random Forest focused on structural drivers, including tenure and revenues, whereas XGBoost focused on smaller-scale operational signals, including assignment-level activity and promotion transitions.

The Logistic Regression achieved mean ROC-AUC and PR-AUC of 0.850 and 0.832 on cross-validation, respectively, with good recall on successful projects but lower on failures in the classification strand. Gradient Boosting performed consistently with the highest ROC-AUC of 0.905 (CV) and 0.830 on the test set and equal precision and recall. Convergence to team size was the strongest predictor in feature analysis, whereas role diversity was identified by Logistic Regression as a negative predictor, and activity diversity and assignment count were identified by Gradient Boosting as situation-specific drivers.

In general, the dual-strand design was effective: regression was able to give detailed information about the intensity of performance, and classification was able to give binary information about risk, which is directly correlated with managerial decisions. The stratified application of hold-out tests, cross-validation and OOF diagnostics validated strong and generalisable findings.

6.1.1 Model strengths

- Random Forest regression presented an explicable and stable baseline that can explain the success variance of approximately 87%.
- XGBoost outperformed other predictive models in generalisation, both in cross-validation and OOF, absorbing finer signals and the best predictive accuracy.
- Logistic Regression was interpretable, and the directional insights were clear as to drivers of success.

- Gradient Boosting performed better in non-linear classification with ROC-AUC values exceeding 0.90 in CV, which is applicable to warning system at an early stage.

6.1.2 Limitations

These findings are limited by several factors. First, the dataset, especially in the classification strand, was relatively small (97 projects), which limited generalisability. Variance was reduced by cross-validation; however, bigger more heterogeneous data sets are required to draw more powerful conclusions. Second, the Success Score weighting scheme (40/30/20/10) was both hypothetically based and subjective; another scheme would have different distributions. Third, Success Flag was based on thresholding a continuous outcome, which makes assumptions that cannot be generalised. Lastly, although XGBoost and Gradient Boosting were found to be superior predictive models, their lack of transparency makes them difficult to adopt relative to the transparency of linear models.

6.2 Possibility for Future Work

This dissertation shows that predictive modelling is an achievable tool in consulting project management and that it has the potential to be further developed:

• **Data enrichment.** It would be better to expand over time, companies, and countries. The inclusion of time aspects (e.g., changing margins over project lifecycle) might facilitate the use of survival models that predict when but not whether a project is at risk.

• **Alternative targets**. The Success Score may be improved in future work with data-driven weighting (e.g., PCA, factor analysis) or may use another dimension (e.g., client satisfaction or Net Promoter Scores). The binary Success Flag may be further extended to multi-class (e.g., high/medium/low success).

• **Advanced modelling.** While Random Forest and Gradient Boosting are currently the most popular algorithms for such tasks, experimenting with alternative and more specialized methods like LightGBM, CatBoost, or neural models, can also lead to some improvement.

• **Practical integration.** Predictive models could be integrated into risk governance frameworks or dashboards for the PMO to connect research and practice. This needs to be calibrated, interpretable and ethically safeguarded so that predictions can guide and not override the managerial decision.

• **Mixed-method approaches.** Integrating machine learning and qualitative results (e.g., interviews, surveys) may be able to reflect governance, leadership, and client relationship variables not reflected in the quantitative data, further dimensionalizing the multidimensional approach to consulting success.

6.3 Recommendations & Findings

6.3.1 For consulting firms:

1. **Prioritise financial monitoring.** Success is measured in terms of revenue rate, contract revenues, and margins, which should be monitored as part of the risk dashboards.
2. **Track team structure.** The size of the team, the variety of roles, and the range of capabilities are factors that affect the outcome of the project in a meaningful way and must be effectively regulated.
3. **Adopt predictive analytics**. Gradient Boosting and XGBoost had the best predictive performance, and hence both can be applied to early-warning systems, whereas Logistic Regression can be used to gain insight into decisions that are to be made in decision forums.
4. **Balance regression and classification views.** Continuous forecasts are useful when making plans to comply with margin, whereas binary outputs can be used to generate actionable risk signals to executives.
5. **Embed interpretability.** Adoption will be the most effective in cases when outputs are interpretable (coefficients, feature importances, SHAP values) and when they are combined with managerial stories.

6.3.2 For academia:

This dissertation empirically proves that Critical Success Factors may be operationalised into predictive models. The framework should be tested in other industries, alternative constructs of success experimented, and a combination of statistical and qualitative evidence should be used to study this problem more deeply in the future.

6.3.3 For methodology:

Hold-out splits, cross-validation and OOF diagnostics offer a reproducible evaluation framework. This multi-level design is both robust and clear, and ought to be thought of as best practice in predictive analytics research.

6.3.4 Conclusion

This dissertation has demonstrated that financial and human-related features in a machine learning model are enough to predict the success of consulting projects with a strong degree of accuracy. The dual-strand design was able to capture both continuous success intensity and binary outcomes and provide outputs that are specific to a particular decision requirement. Random Forest regression alone explained about 87 percent of the Success Score variance, and XGBoost attained still greater generalisation in the cross-validation. The classification models attained above 0.80 ROC-AUCs on held-out tests and above 0.90 cross-validation, and these results supported the application of the classification models as early-warning tools.

The results support profitability, contract revenues and human capital structures as the key Critical Success Factors, which complies with the empirical findings with theoretical assumptions. Meanwhile, dataset size, target construction and model interpretability are warning signs of generalisation. By offering pathways toward data enrichment, methodological advancement, and pragmatic integration, this dissertation offers a basis for the integration of predictive analytics into consulting project management.

In conclusion, while machine learning can by no means substitute for managerial judgment, it can certainly amplify it-from providing foresight, structure, and empirical grounding to the complex task of setting consulting projects up for success.

References

Rockart, J. F. (1979). *Chief executives define their own data needs.* Harvard Business Review, 57(2).

Available at: https://hbr.org/1979/03/chief-executives-define-their-own-data-needs

Pinto, J. K., & Slevin, D. P. (1987). *Critical factors in successful project implementation.* IEEE Transactions on Engineering Management, EM-34(1), 22–27.

Available at: https://doi.org/10.1109/TEM.1987.6498856

Kaplan, R. S., & Norton, D. P. (1996). *Using the balanced scorecard as a strategic management system.* Harvard Business Review, January–February, 75–85.

Available at: https://www.strimgroup.com/wp-content/uploads/pdf/KaplanNorton_Linking-the-BSC-to-Strategy.pdf

Atkinson, R. (1999). *Project management: cost, time and quality, two best guesses and a phenomenon, it's time to accept other success criteria.* International Journal of Project Management, 17(6), 337–342.

Available at: https://www.sciepub.com/reference/47710

Baccarini, D. (1999). *The logical framework method for defining project success.* Project Management Journal, 30(4), 25–32.

Available at: https://journals.sagepub.com/doi/10.1177/875697289903000405

Shenhar, A. J., Dvir, D., Levy, O., & Maltz, A. C. (2001). *Project success: a multidimensional strategic concept.* Long Range Planning, 34(6), 699–725.

Available at: https://doi.org/10.1016/S0024-6301(01)00097-8

Ika, L. A. (2009). *Project success as a topic in project management journals.* Project Management Journal, 40(4), 6–19.

Available at: https://journals.sagepub.com/doi/10.1002/pmj.20137

Standish Group International. (2013). *Chaos Manifesto 2013.* "Think Big, Act Small."

Available at: https://larlet.fr/static/david/stream/ChaosManifesto2013.pdf

Alias, Z., Zawawi, E. M. A., Yusof, K., & Aris, N. M. (2014). *Determining critical success factors of project management practice: a conceptual framework.* Procedia – Social and Behavioral Sciences, 153, 61–69.

Available at: https://doi.org/10.1016/j.sbspro.2014.10.041

Fauser, J., Schmidthuysen, M., & Scheffold, B. (2015). *The prediction of success in project management – Predictive project analytics.* Deloitte Consulting GmbH.

Available at: https://www.bdu.de/media/177649/predictive-project-analytics_artikel_gpm.pdf

Gao, X., Khorasani, S., Krishnan, V., Nittala, L., & Wang, S. (2024). *A machine learning framework for project performance prediction, interpretation, and inference.* SSRN.
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5051652

Ajanaku, A. G. (2025). *Practical applications of machine learning and artificial intelligence in predicting the cost of consulting projects.* SSRN.
Available at: https://ssrn.com/abstract=5096993