NPS-AM-26-301

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

From Separation to Participation: Predicting Reserve Service Outcomes Using Machine Learning in the Australian Defence Force

March 2026

Maj Darby A. Nelson, Royal Australian Air Force

Thesis Advisors: Dr. Chad W. Seagren, Senior Lecturer
Dr. Ruriko Yoshida, Professor

Department of Acquisition, Finance and Manpower

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

The research presented in this report was supported by the Acquisition Research Program of the Department of Acquisition, Finance and Manpower Naval Postgraduate School.

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact the Acquisition Research Program (ARP) via email, arp@nps.edu or at 831-656-3793.

ABSTRACT

The Australian Government has directed the Australian Defence Force (ADF) to increase its Reserve workforce by 1,000 additional members by 2030; however, current Reserve recruitment and full-time-to-Reserve transition rates remain insufficient to meet this target. This thesis uses individual-level demographic, geographic, and economic data to train and evaluate machine learning models that predict Service Category (SERCAT) choice, subsequent Reserve participation and service intensity among separating full-time ADF members from FY2016–17 to FY2024–25. The models demonstrate strong performance in classifying SERCAT choice but substantially weaker performance for downstream participation and intensity outcomes. Institutional and geographic characteristics are most influential at the point of separation, whereas personal and life-stage characteristics are more strongly associated with post-separation Reserve engagement. These findings demonstrate how machine learning can inform targeted Reserve outreach strategies to support directed workforce growth, while also revealing structural and individual constraints that limit realized participation after separation.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

I would like to begin by thanking my co-advisors, Dr. Rudy Yoshida and Dr. Chad Seagren, whose support made it possible for me to pursue my interest in applying machine learning within the manpower domain.

None of this would have been possible without the support of my wonderful wife, Olivia. Sharing our time at the Naval Postgraduate School was such a gift, and I am deeply grateful for your unwavering encouragement throughout the many months of the thesis process, whether we were together or separated by distance. I am also grateful to Major Sam Lam, who let me pressure him into undertaking the machine learning research with me. Your partnership throughout the journey made the experience far more rewarding. To Flight Lieutenant Nicholas Harvey, thank you for your steadfast encouragement all the way from Australia. I am excited to be reunited.

I am especially thankful to Flight Lieutenant Jennifer Woodside, whose support for Olivia and me, beginning before our departure, continuing throughout my program, and extending to our return to Australia, has been invaluable. I owe sincere thanks to Group Captain Mark Powell for sparking my interest in the manpower field and encouraging me along the path that led me here. In the same spirit, I appreciate Squadron Leader Vivienne Clark and Squadron Leader Daniel Tyson for sharing their experiences and workforce insights, which helped shape my understanding of manpower in the ADF context. Thank you to Group Captain Peter Davis and Wing Commander Tony Smith from the ADF Total Workforce System Centre of Expertise for their generosity in sharing Reserve-specific knowledge and perspectives.

Thank you to my family. I am so glad we were able to spend so much time together during my time in Monterey.

Finally, I would like to thank the Royal Australian Air Force for not only granting me the opportunity to pursue this education, but also for the greater privilege of serving as part of your force. I look forward with excitement to what the future holds.

THIS PAGE INTENTIONALLY LEFT BLANK

NPS-AM-26-301

ACQUISITION RESEARCH PROGRAM
SPONSORED REPORT SERIES

From Separation to Participation: Predicting Reserve Service Outcomes Using Machine Learning in the Australian Defence Force

March 2026

Maj Darby A. Nelson, Royal Australian Air Force

Thesis Advisors: Dr. Chad W. Seagren, Senior Lecturer
Dr. Ruriko Yoshida, Professor

Department of Acquisition, Finance and Manpower

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

ABS Australian Bureau of Statistics

ADF Australian Defence Force

AUC area under the curve

DoD Department of Defense

FY financial year

GCCSA Greater Capital City Statistical Areas

HR human resources

MAE mean absolute error

ML machine learning

OSCA Occupation Standard Classification for Australia

RMSE root mean squared error

SA4 Statistical Areas Level 4

SERCAT Service Category

SHAP SHapley Additive exPlanations

THIS PAGE INTENTIONALLY LEFT BLANK

Disclosures

We use generative artificial intelligence tools, specifically ChatGPT and Claude, to support the development of the machine learning models in this thesis (Anthropic, 2026; OpenAI, 2026). These tools assisted us with hyperparameter tuning and generation of figures based on the tuned models and their associated results. All outputs generated by these tools were reviewed and validated by the authors, who retain full responsibility for the methodology, analysis, and conclusions presented in this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

Following a comprehensive review of the Australian Defence Force (ADF) Reserve Force, the Australian Government approved a recommendation of the *Strategic Review of the Australian Defence Force Reserves* (hereafter, the Review) to increase the Reserve Force by 1,000 additional members by 2030 (Department of Defence, 2024). This target represents growth from a Reserve Force of 41,717 as of 1 March 2024, of whom 32,311 provided service and 9,406 did not, as reported in the Review. These figures reflect a point-in-time snapshot and may vary over time.

Defence can achieve this growth through Reserve recruitment or Permanent-to-Reserve service transfers; however, the Review notes that the majority of job roles within the Royal Australian Navy and Royal Australian Air Force do not have direct-to-Reserve pathways through initial recruitment (Department of Defence, 2024). Throughout this thesis, the terms *Permanent* and *full-time* are used interchangeably to refer to members serving in full-time Service Categories (SERCATs) (SERCAT 6 and SERCAT 7). As a result, Permanent-to-Reserve service transfers represent the most viable mechanism for achieving the directed growth target.

1.1 Problem Statement

Realizing the Government-directed growth in the ADF Reserves requires a clearer understanding of how full-time members transition into and engage in Reserve service. At the point of separation from full-time service, members are required to select a Reserve SERCAT (SERCAT 2, SERCAT 3, or SERCAT 5), formalizing their intended level of Reserve service commitment. This thesis seeks to predict that initial SERCAT selection and subsequent participation outcomes. Although administrative personnel data capture extensive information about members at the point of separation, Defence currently lacks the analytical tools required to translate these data into actionable insights regarding Reserve service choice, participation, and intensity. As a result, Reserve growth initiatives largely rely on aggregated trends and post hoc analysis rather than prospective, data-informed assessment.

This challenge is compounded by the fact that Reserve engagement is not a single decision, but a sequence of related behaviors that evolve over time as circumstances, constraints, and preferences change, from service category selection at separation to post-separation participation. This complexity limits the usefulness of traditional descriptive approaches and motivates the exploration of methods capable of capturing non-linear relationships and non-uniform distributions of outcomes. Addressing this gap is necessary to support more informed workforce planning, targeted outreach, and effective allocation of resources intended to influence Reserve participation at the point of separation.

1.2 Research Questions

To address this problem, this thesis investigates the following research questions:

1. Can machine learning (ML) models be developed to predict the Reserve service behavior of separating full-time ADF members with accuracy exceeding chance-level baselines?
2. Using ML methods trained on financial year (FY)16–17 to FY22–23 ADF personnel data, with what degree of certainty can we predict the following outcomes for separating full-time ADF members:
    - the Reserve service category choice;
    - whether a member completes any Reserve service in the FY following separation, conditional on being predicted to join SERCAT 3 or SERCAT 5; and
    - the annual number of hours of Reserve service in the FY following separation, conditional on predicted SERCAT assignment and participation.
3. What are the key features driving the predictions generated by the ML models?

We find that ML models can predict several dimensions of Reserve service behavior with performance exceeding naive baselines, though prediction strength varies across outcomes. The model predicting Reserve service category choice substantially outperforms random classification, achieving approximately 74 percent accuracy compared to a 33 percent random baseline across the three SERCATs. Geographic mobility, institutional structure, and economic characteristics emerge as the most influential predictors of service category outcomes, reflecting how members anticipate post-separation location, career continuity, and institutional fit at the point of separation.

In contrast, the model predicting whether a separating member completes Reserve service in the following FY demonstrates more modest performance, attaining 65 percent accuracy relative to a 50 percent random baseline in a binary classification setting. Participation outcomes are driven primarily by institutional affiliation and life-stage characteristics, suggesting that realized engagement reflects how members' post-separation circumstances evolve during their first year outside full-time service.

Finally, the model predicting annual Reserve service hours improves upon a naive mean benchmark, reducing root mean squared error (RMSE) from 660 to 512, a 22.4 percent reduction in prediction error. The model achieves an R^2 of 15.7 percent, indicating that it explains a meaningful, though limited, share of the variance in post-separation service intensity. Age and economic features emerge as the strongest predictors of service intensity, consistent with members adjusting their level of engagement according to their career progression and life stage following separation.

These findings demonstrate that SERCAT choice is meaningfully predictable at the point of separation, while participation and intensity are characterized by greater predictive uncertainty. The chapters that follow provide the institutional background, methodological design, and analysis through which these conclusions are established.

1.3 Structure

This thesis comprises six chapters, including this introduction (Chapter 1). Chapter 2 provides organizational background information about the ADF workforce system, separation process, and Reserve growth challenge. Chapter 3 reviews the literature on Reserve service participation, ML models, and ML applications across civilian and military human resources (HR). Chapter 4 outlines the data and three-stage methodology we use in this research. Chapter 5 reviews the results of the models by stage. Lastly, Chapter 6 discusses the results, outlines practical implications of the modeling framework, discusses limitations and constraints, and suggests further research areas.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 2:
Background

This chapter outlines the institutional and administrative context underpinning the research questions we examine in this thesis. Understanding these ADF-specific practices informs the methodology and interpretation of results. Accordingly, this chapter describes the workforce structure, the seven Service Categories, the separation process, targeted Reserve growth, and the challenges associated with achieving that growth.

2.1 Australian Defence Force Workforce Structure

Prior to 2016, *Defence (Personnel) Amendment Regulations 2002 (No. 1) (Cth)* defined the ADF's workforce system (Defence (Personnel) Amendment Regulations 2002 (No. 1) (Cth), 2002) . Under this framework, the workforce consisted of the Permanent Force and five categories of the Reserve Force: High Readiness Active Reserve, High Readiness Specialist Reserve, Active Reserve, Specialist Reserve, and the Standby Reserve category.

In FY 2016–17, the ADF transitioned to the Total Workforce System, which restructured the previous Permanent and Reserve categories into seven SERCATs across three broad classifications (Department of Defence, 2017). SERCAT 1 comprises civilian employees, SERCATs 2–5 constitute the Reserve component, and SERCATs 6–7 comprise the Permanent Force (Department of Defence, 2025a). This transition provides nine complete FYs of data under the Total Workforce System.

2.2 Service Categories

This section outlines the seven SERCATs defined under the Total Workforce System, as described by Department of Defence (2025a).

- **SERCAT 1**: Australian Public Service employees who serve on operations contributing to ADF capability. Individuals in this category are not members of the ADF; rather, SERCAT 1 provides a mechanism for civilians to perform roles within an area of operation in support of Defence capability.

- **SERCAT 2**: ADF Reserve members who do not render service. This category represents a strategic reserve of former SERCATs 3–7 members who have completed their service obligation. Members in SERCAT 2 have no military obligations, except in circumstances of call out by the Governor-General under section 28 of the *Defence Act 1903*. Individuals in this category may apply to transfer to another SERCAT should they wish to resume service.
- **SERCAT 3**: This SERCAT comprises two workforce components:
  - ADF Reserve members who identify their ability to contribute to capability on a contingency basis. These members form a capability pool that Services and Groups may draw upon at short notice to meet short-term or emergent workforce demands.
  - ADF Reserve members who render service toward a specific task within a given FY, such as short-duration assignments or the provision of specialized skills and qualifications.

  Members in SERCAT 3 may continue to render service within this category, transfer to SERCAT 4 or 5, or volunteer for a period of continuous full-time service, depending on the nature and duration of the service requirement. Examples of activities SERCAT 3 members may undertake include short-term project work, job sharing, or the provision of specialized skills not readily available within the Permanent Force.
- **SERCAT 4**: ADF Reserve members who provide a contingent capability at short notice. Members in this category enter into a written undertaking with their Service to participate for a specified minimum period and must comply with conditions determined by that Service. Roles within SERCAT 4 typically reflect qualifications that a Service may need to call upon at short notice to assist or augment their Permanent Force.
- **SERCAT 5**: ADF Reserve members who provide an enduring part-time capability. Individuals undertake a formal agreement with their Service to participate in a specific work pattern across multiple financial years. This category offers stable, longer-term contributions from Reserve members while supporting the Permanent Force through improved resourcing and access to specialized skills not readily available within the Permanent Force.
- **SERCAT 6**: ADF Permanent members rendering a pattern of service less than full-time. Individuals undertake a formal agreement with their Service to participate in

an agreed pattern of service, such as specified days per fortnight, weeks per month, or months per year.

- **SERCAT 7**: ADF Permanent members rendering full-time service. This category represents the standard and maximum service obligation within the Permanent Force.

Our subsequent analysis focuses on separating full-time (SERCAT 6 and SERCAT 7) ADF members who transition into Reserve SERCATs 2, 3, and 5. SERCAT 1 is excluded as it consists of civilian employees rather than members of the ADF. SERCAT 4 is excluded due to its specialized contingent nature, which limits eligibility to a narrow set of roles and does not reflect broadly accessible Reserve participation opportunities. Notably, only one observation in the dataset transferred from the Permanent Force to SERCAT 4, further supporting its exclusion from our analysis.

Having outlined the structural differences between SERCATs, it is also necessary to understand the incentives associated with Reserve service, as these may shape members' decisions at separation.

2.3 Benefits of Reserve Service

Reserve service offers a combination of professional, financial, and lifestyle benefits that may influence members' decisions to affiliate and participate in Reserve service after separation. ADF recruitment material describes Reserve service as a rewarding additional dimension of an individual's professional and personal life (ADF Careers, n.d.). Recruitment messaging emphasizes opportunities to undertake challenging and meaningful work, travel, develop new skills, and exercise flexibility in choosing when to render service (ADF Careers, n.d.).

In addition to these intrinsic benefits, Reserve members receive tax-free remuneration and may be eligible for financial support such as a health support allowance and subsidies toward home loan repayments (Department of Defence, n.d.-d). The Department of Defence (n.d.-f) further specifies that Reserve members are remunerated on a daily basis and are paid only for days worked, in contrast to Permanent members who receive a continuous salary, employer superannuation contributions, and accrue paid leave. Although Reserve pay is tax-free, members do not receive employer superannuation contributions while serving in a Reserve capacity (Department of Defence, n.d.-f).

Together, these arrangements indicate that Reserve service provides a flexible form of supplementary employment rather than a primary career pathway. The relative attractiveness of these incentives may vary across members depending on age, rank, civilian employment prospects, geographic location, and life stage. As such, Reserve benefits form an important contextual backdrop for understanding the economic and personal trade-offs modeled in this thesis.

2.4 Separation Process

In this research, separation is defined as the transition of ADF members from the Permanent Force (i.e., SERCAT 6 or SERCAT 7) to the Reserve Force (i.e., SERCATs 2, 3, or 5). This definition excludes involuntary separations that do not result in a transfer to the Reserve Force, such as separations due to being medically unfit for service, retention not in the Service's interest, or training failure. Within this study, the defined separation transition is effected administratively through the AC853 *Application to Transfer Within or Separate from the ADF* form (Department of Defence, 2025b). The AC853 specifies three months as the minimum period required for a member to provide notice of their intention to separate. In this study, this notice period is used to align observed separation decisions with prevailing economic conditions at the time the SERCAT decision is made. Although the ADF may approve separations with less than three months' notice in limited circumstances, this study uses a uniform three-month notice proxy for all members to ensure analytical consistency.

In the *Details of request* section of the form, members are required to nominate the SERCAT to which they wish to transfer. The three available options are outlined below as described by Department of Defence (2025b):

- **SERCAT 2**: Selection of SERCAT 2 prompts advisory text informing the member that, if they intend to render service within two years of their transfer date, they should consider applying to transfer to SERCAT 3 or 5 instead.
- **SERCAT 3**: Selection of SERCAT 3 generates advisory text indicating that members who do not intend to provide service within two years of their transfer date should apply to transfer to SERCAT 2. In addition, several SERCAT 3-specific fields are generated, including:

  - A yes/no selection indicating whether the application is in response to an expression of interest or call for nominations for a particular job, tasking, or position;
  - A text field for preferred position location or region;
  - A selection between posting to a SERCAT 3 pool position at the higher headquarters where service is most likely to be rendered, or posting to the nearest base unit;
  - A yes/no selection indicating whether Reserve service has been arranged; and
  - A text field specifying the number of days available to work in each financial year.
- **SERCAT 5**: Selection of SERCAT 5 does not prompt advisory text but similarly generates category-specific fields, including:
  - A yes/no selection indicating whether the application is in response to an expression of interest or call for nominations for a particular job, tasking, or position;
  - A text field for preferred position location or region;
  - A yes/no selection indicating whether the member has contacted the desired unit and confirmed the availability of work;
  - A text field specifying the preferred unit and position details; and
  - A text field specifying the number of days available to work in each financial year.

The information collected through this section of the AC853 *Application to Transfer Within or Separate from the ADF* form illustrates the degree of advance planning required of members when notifying their Service of an intended transition to Reserve service.

2.5 Annual Permanent to Reserve Transfer Magnitude

The Review provides insight into the scale of transitions from full-time service to the Reserve Force (Department of Defence, 2024). Using data spanning FY2017–18 to FY2022–23, the Review reports the annual average volume of transitions from SERCAT 6 and SERCAT 7 to SERCATs 2, 3, and 5. The transition rates exclude members assessed as ineligible for Reserve service due to factors such as medical limitations, involuntary separation, or reaching retirement age (Department of Defence, 2024). We reproduce this figure in Table 2.1.

Table 2.1. This table reflects the average annual full-time to Reserve transition volume between FY2017–18 and FY2022–23, excluding ineligible members. Source: Department of Defence (2024).

Service	SERCAT 2	SERCAT 3	SERCAT 5	Total
Navy	153	502	6	661
Army	474	544	851	1,869
Air Force	68	516	165	749
Total	695	1,526	1,022	3,279

These statistics indicate that the majority of separating full-time members, approximately 78 percent, transition into SERCAT 3 or SERCAT 5. However, there is substantial variation in Reserve participation across the three Services. The Navy's Reserve Force is predominantly concentrated within SERCAT 3, while the Army has a comparatively large population of SERCAT 5 members. We consider these differences alongside the required growth of the Reserve Force.

2.6 Directed Reserve Growth

Although there are moderate transitions from the Permanent Force to SERCATs 3 and 5, existing transition rates are insufficient to meet future workforce objectives. Accordingly, the Australian Government accepted a recommendation from the Review to increase the SERCATs 3–5 components of the Reserve Force by at least 1,000 additional members by 2030. The Review's target represents net growth rather than a replacement of routine attrition, and imposes additional demand on both full-time-to-Reserve transition pathways and direct Reserve recruitment (Department of Defence, 2024). We anticipate that achieving this growth objective will be challenging under current full-time-to-Reserve transition patterns and persistent Reserve recruitment shortfalls.

2.7 Reserve Recruitment Challenge

The Strategic Review further observes that approximately 90 percent of Army SERCAT 3–5 members are recruited directly from the civilian population, in contrast to the Navy and Air Force, which rely heavily on transitions from full-time service to sustain their Reserve Forces (Department of Defence, 2024). This structural distinction heightens the importance for the Navy and Air Force to identify, engage, and retain separating full-time members within the Reserves.

Table 2.2 illustrates the challenging Reserve recruitment environment faced by the ADF in FY2023–24 (Department of Defence, 2024). Across all Services, Reserve recruitment fell well short of targets and achieved less than half of the required inflow. These shortfalls mirror those experienced in Permanent Force recruitment and are compounded by systemic factors such as extended delays between expressions of interest and hiring. Prolonged recruitment times increase the likelihood that prospective reservists withdraw from recruitment and pursue alternative civilian employment, further constraining Reserve force generation (Department of Defence, 2024).

Table 2.2. This table illustrates the challenging Reserve recruitment environment faced by the ADF in FY2023–24. Source: Department of Defence (2024).

Service	Target	Forecast Result	Shortfall	Achievement (%)
Navy	88	20	−68	22.7
Army	1,869	936	−933	50.1
Air Force	132	55	−77	41.7
Total Reserves	2,089	1,011	−1,078	48.4

This chapter describes the workforce structure, service categories, separation process, growth targets, and recent transition and recruitment outcomes. Through this discussion, we establish the institutional and administrative context within which Reserve participation decisions are made. While a considerable number of full-time members separate each year with eligibility for Reserve service, observed transition patterns and recruitment outcomes are not fulfilling the agreed requirement to expand the Reserve Force by 2030. This challenge is particularly acute for Services that rely heavily on full-time-to-Reserve transitions to sustain capability, compounded by recruitment shortfalls and competition from the civilian labor market.

These conditions highlight a fundamental weakness in existing engagement approaches, which offer limited capacity to distinguish between individuals who are likely to affiliate with the Reserves, those whose participation may be influenced through targeted intervention, and those who are unlikely to be affected. In relation to the agreed-upon Reserve growth targets, this limitation constrains Defence's ability to efficiently generate additional Reserve capability from the existing separating workforce.

This thesis is motivated by the need to address this gap in targeted outreach through a data-driven approach to achieve the agreed-upon Reserve workforce growth. By applying ML techniques to administrative, military, geographic, and economic data, this research seeks to identify individual-level patterns associated with Reserve affiliation and participation. These results support more targeted, timely, and efficient workforce interventions prior to separation, contributing to improved Reserve generation under current transition and recruitment constraints. The next chapter reviews the relevant literature on Reserve affiliation behavior and predictive modeling approaches that inform this research.

CHAPTER 3:
Literature Review

In this chapter, we review prior literature regarding Reserve service participation, ML models, and ML applications across civilian and military HR. We structure this chapter in six sections. First, we introduce the conceptual framework of this thesis. Second, we outline the three ML models we use. Third, we examine civilian HR applications of ML, followed by military applications in the fourth section. Subsequently, we identify influential factors from prior service Reserve research. Finally, we assess the current ADF Reserve research and highlight the gap this thesis addresses.

3.1 Conceptual Framework

We identify a gap in recent academic literature identifying characteristics of prior service individuals who choose to join the Reserves, compared to the extensive body of work analyzing civilians' decisions to join the military in full-time or Reserve roles. The most relevant prior service-specific Reserve literature analyzes U.S. military services and their members.

Among the studies relevant to this research, Jeremy Arkes and Rebecca Kilburn (2005) identified characteristics of prior service Reserve members to support the development of Reserve recruiting models for all U.S. Department of Defense (DoD) services. Their prior service Reserve recruiting model provides the conceptual foundation for this thesis.

Arkes and Kilburn (2005) adapted Charles Manski's (1977) Random Utility Model, which defines an individual's decision to join the Reserves as occurring when the utility associated with Reserves employment is greater than the utility associated with all other discrete alternatives. Arkes and Kilburn's (2005) framework can be expressed by the following inequality:

$$U_{ir} > U_{ij} \quad \text{for } j = 1, 2, \ldots, J, \tag{3.1}$$

where U_{ir} denotes the utility individual i derives from joining the Reserves option r, and U_{ij} denotes the utility individual i derives from any alternative option j.

Arkes and Kilburn (2005) extend the Random Utility Model concept to a labor-supply problem by incorporating Robert Shishko and Bernard Rostker's (1976) Moonlighting Model. The Moonlighting Model provides a framework for determining how much labor an individual supplies to a second employer, such as Reserve service in the ADF.

In Shishko and Rostker's (1976) moonlighting framework, an individual's utility is a function of consumption and leisure. Consumption is derived by the combination of wages from the primary employment and a moonlighting (Reserve) job (w_0, w_m), working hours from each role (L_0, L_m), and non-labor income such as pensions, inheritance, or spousal income (A_0). Leisure is the residual of total time after working ($N - L_0 - L_m$). The individual chooses the hours of labor to supply to their employers that maximize the individual's total utility (Shishko & Rostker, 1976). Therefore, the individual will allocate hours to the Reserves only if the utility of combined primary and Reserve work exceeds the utility from primary employment alone. For this condition to hold, the benefit in increased utility from Reserve income must outweigh the decrease of utility from reduced leisure. Following Arkes and Kilburn (2005), who incorporate the moonlighting framework of Shishko and Rostker (1976), the moonlighting decision is characterized by the following inequality:

$$U(w_0 L_0 + w_m L_m + A_0,\ N - L_0 - L_m) > U(w_0 L_0 + A_0,\ N - L_0)\,, \tag{3.2}$$

where:

- $w_0 L_0 + w_m L_m + A_0$ denotes income earned from both primary and moonlighting work,
- $N - L_0 - L_m$ denotes remaining leisure time after primary and moonlighting work,
- $w_0 L_0 + A_0$ denotes income earned from only primary work, and
- $N - L_0$ denotes remaining leisure time after only primary work.

In order to empirically use these conceptual frameworks, observable individual characteristics serve as proxies for unobservable preferences towards consumption and leisure (Arkes & Kilburn, 2005). An individual's utility of alternatives can be represented by a function containing the individual's observable characteristics, such as age, rank, gender, length of service, and education. Arkes and Kilburn (2005) propose a utility function that can be used to estimate an individual's propensity for Reserve participation based on measurable traits:

$$U_{ik} = f_k(X_i) + \varepsilon_{ik}, \tag{3.3}$$

where:

- U_{ik} denotes the utility individual i derives from option k,
- $f_k(X_i)$ denotes the portion of utility for option k explained by individual i's observable characteristics, and
- ε_{ik} denotes a random error term specific to option k.

Accordingly, Arkes and Kilburn's (2005) model shows that the decision to join the Reserves can be expressed as the probability that an individual's utility from joining the Reserves exceeds that of alternative options, such as only civilian employment or non-employment:

$$P_i(Reserves) = P(U_{i,Reserves} > U_{i,CivilianEmployment} \text{ and } U_{i,Reserves} > U_{i,NoEmployment}). \tag{3.4}$$

While the Arkes and Kilburn's (2005) model treats Reserve service as a binary choice, participate or not participate, we extend their framework by modeling the probability that an individual selects a particular SERCAT based on the utility associated with each option. For example, the probability that an individual joins SERCAT 2 is equal to the probability the individual has a greater preference for the SERCAT 2 income and leisure combination compared to the SERCATs 3 and 5 income and leisure combinations, such that:

$$P_i(SERCAT_2) = P(U_{i,SERCAT_2} > U_{i,SERCAT_3} \text{ and } U_{i,SERCAT_2} > U_{i,SERCAT_5}). \tag{3.5}$$

While these economic frameworks provide the theoretical foundation for Reserve participation decisions, ML methods offer an empirical mechanism for estimating these probabilities using observable characteristics of separating ADF members.

3.2 Foundational Machine Learning Methods

We use three types of ML methods to model Reserve SERCAT choice probabilistically. Following the standard supervised learning framework (Freiesleben & Molnar, 2024), each supervised model is trained on a set of examples where $\boldsymbol{X}_i = \langle X_{i1}, \ldots, X_{in} \rangle$ denotes the vector of observed features describing an observation (e.g., characteristics of a separating ADF member) and y_i represents the associated outcome variable, the member's SERCAT choice. These models employ numerous decision trees of varying depths, each individually learning a set of "if-then rules" from the features to predict the associated outcomes. Once trained, the models are evaluated on a separate set of observations $\boldsymbol{X}_k = \langle X_{k1}, \ldots, X_{kn} \rangle$. The models' performance is evaluated based on how accurately they predict each observation's outcome $\hat{y}_k$, the predicted SERCAT, compared to the known outcome y_k, the SERCAT selected (Freiesleben & Molnar, 2024).

Yoav Freund and Robert Schapire (1997) developed the first model we use, Adaptive Boosting. Adaptive Boosting creates a strong classifier by combining multiple weak learners (shallow decision trees akin to a "rule-of-thumb"), whose individual performance is only marginally better than random guessing. During sequential training, each learner improves by learning from past mistakes because the model assigns higher weights to observations that were previously misclassified. The final classifier is formed through a weighted majority vote, where the weight of each weak learner is proportional to its accuracy. Although each learner is weak in isolation, Adaptive Boosting's emphasis on learning from misclassifications and the combination of many learners decreases classification error exponentially to achieve strong predictive performance.

The second model, Random Forest, introduces randomness within feature and data selection to reduce overfitting (Breiman, 2001). Random Forest creates a strong classifier by combining multiple deep decision trees. Each tree is trained using a random subset of the observations in $\boldsymbol{X}_i$ with replacement, allowing an individual ADF member to appear in multiple trees. At each decision tree split, a random subset of features from $\boldsymbol{X}_i$ is considered. This random subsetting allows each tree to learn from diverse information. The final classification is determined by majority vote across all trees. The introduction of random subsetting reduces correlation among individual trees, which enhances generalization performance on unseen data.

Tianqi Chen and Carlos Guestrin (2016) developed the third model, eXtreme Gradient Boosting. Similar to Adaptive Boosting, eXtreme Gradient Boosting builds a collection of shallow decision trees that work together to improve prediction accuracy. However, rather than reweighting incorrectly classified observations like Adaptive Boosting, eXtreme Gradient Boosting sequentially fits each new tree to correct the mistakes made by the previous trees. This iterative process reduces prediction error throughout the model's development, which enhances the model's ability to generalize to new data. Having outlined the foundations of Adaptive Boosting, Random Forest, and eXtreme Gradient Boosting, we examine how these methods have been applied in civilian and military HR contexts to justify their use in this study in the following sections.

3.3 Civilian Human Resources Applications of Machine Learning

We select Adaptive Boosting, Random Forest, and eXtreme Gradient Boosting for this thesis based on their demonstrated predictive success in civilian HR applications. Saeed Nosratabadi et al. (2022) conducted a systematic literature review of 23 academic papers examining how ML models assist HR managers across the five phases of the employee life cycle: recruitment, onboarding, employability and benefits, retention, and offboarding. The offboarding phase, which is most closely aligned with this thesis, occurs when employees separate from an organization due to job change, retirement, personal reasons, or dismissal. The authors identified all three models as high-performing within offboarding-related HR applications.

We also note that ML models are used in the civilian literature to identify the features that most strongly influence employee outcomes. Shahin Manafi Varkiani et al. (2025) tested 10 models, including Random Forest, to explain the determinants of employee attrition using HR data from 5,767 employees at an Italian financial company collected between January 2012 and June 2019. The authors identified Random Forest as the most effective model. Out of 27 employee features, the study found that compensation, age, tenure, and gender are the most influential in predicting attrition within their dataset. The authors find that higher compensation reduces attrition risk, while younger employees with shorter tenures exhibit greater mobility and were therefore more likely to leave. Moreover, men were found to be less likely to attrite than women. Although these demographic findings are not directly transferable to the ADF context, these findings illustrate the predictive capability of ML models and their interpretive value for understanding workforce dynamics, providing direct methodological motivation for this study. While civilian organizations have widely adopted these approaches, their application within military HR management remains comparatively limited but is steadily expanding.

3.4 Military Applications of Machine Learning

In this section, we review four studies that apply ML to inform military recruitment and attrition. Kelsey Hassin (2019) used ML methods to predict the impact of conduct waivers on first-term attrition for 286,696 U.S. Army soldiers who enlisted between financial years 2009 and 2013. The study compared Logistic Regression, Random Forest, Classification and Regression Tree, Support Vector Machine, and Adaptive Boosting. While Hassin (2019) acknowledges that none of the models demonstrated strong performance based on the reported metrics, the study found Random Forest had the highest accuracy and Logistic Regression had the highest area under the curve compared to the other models.

Johnathan Thornton (2023) studied determinants that distinguish Marine Corps recruits who successfully enter basic training from those who reneged. The author used 70 demographic variables on 437,989 Marine Corps applicants from financial years 2015 to 2022 to develop his Random Forest model. This model achieves 94 percent accuracy in predicting individuals who attended or reneged from training. Four variables describing the applicant's weight, run time, and contracting timing details were highly influential to his strongly predictive model.

Andrew Born (2024) employed Automated Machine Learning, which automates the processes of feature selection, feature engineering, and hyperparameter tuning, to optimize recruiter-to-school allocations for the Marine Corps. His models were trained on a dataset combining 2016 to 2022 data from the Marine Corps recruiting system, a U.S. elementary and secondary school database, and American Community Survey results. The AutoML model predictions outperformed the existing recruiter-to-school allocation process and highlight the potential in using ML models to optimize scarce resource allocation.

In contrast, Aaron Falk (2023) compared a Random Forest model to the current method used to predict U.S. Marine attrition at both the aggregate and individual levels. The Random Forest models were trained on monthly snapshots of 470 predictors describing all enlisted Marines from October 2009 to October 2021. While his aggregate-level model outperformed the current Marine Corps prediction process, the study highlights the limited predictive power of his model at the individual level. His individual-level model achieved an average classification accuracy of 45 percent, and he was required to use a high-performance computer due to the size of the data.

These four studies highlight strengths and weaknesses of ML applications in military HR. The strengths include Thornton (2023) using ML determinants to inform influential factors on applicants attending basic training and Born (2024) highlighting the ability of ML to optimize allocation of scarce resources. The studies identify limitations in the models developed by Hassin (2019), including limited predictive power regarding waiver impact, while Falk (2023) reports weak individual-level attrition predictions and significant computational cost requiring high-performance computing. Although results across studies are mixed, these applications demonstrate the growing potential of ML to identify patterns within military personnel data and inform workforce planning decisions. To adapt these analytical methods to Reserve participation modeling, we first examine the underlying factors that influence an individual's decision to join the Reserves.

3.5 Factors Contributing to Reserve Service Decision

Building on our justification for using ML, we draw on economic literature to determine which factors should be incorporated as features in the study's models. Existing research on prior service Reserve participation provides an important empirical foundation for identifying variables relevant to predicting Reserve choice in the ADF. In particular, Arkes and Kilburn (2005) and Jennie Wenger et al. (2016) highlight four influential categories: demographic, economic, geographic, and prior full-time experience. These characteristics are included as proxies for the unobservable factors influencing individuals' preferences for income and leisure and, ultimately, their likelihood of participating in Reserve service.

Wenger et al. (2016) updated and focused Arkes and Kilburn's research by analyzing prior service member uptake into the U.S. Army's Reserve Components. Wenger et al. (2016) quantitatively studied 520,000 enlisted and 43,000 officers who separated from full-time service between financial years 2003 to 2010 using logistic regression. The quantitative study is complemented with qualitative results from nine focus groups of seven mixed rank participants across three U.S. Army bases.

The most influential demographic factors identified in previous Reserve research include ethnicity, age, gender, and education level (Arkes & Kilburn, 2005; Wenger et al., 2016). Wenger et al. (2016) observe that job role is strongly correlated with affiliation in the Reserves. Minority groups are more likely to join the Reserves than non-minority groups while older service members are less likely to affiliate than younger service members. This study finds no gendered difference in officer probability of Reserve affiliation, but enlisted men are approximately two percentage points more likely to participate in the Reserves compared to enlisted women. These categories are included in this thesis data; however, the influence on Reserve prediction may not directly align due to different workforce structures and demographics within the U.S. military compared to the ADF (Wenger et al., 2016).

Relevant economic factors influencing prior service member association with the Reserves include state factors such as company size, unemployment, and median high school wage. States with larger company workforces are associated with higher prior service accessions (Arkes & Kilburn, 2005). Arkes and Kilburn (2005) find that increases in state unemployment rate and median wages of high school graduates are associated with an increase in prior service accessions. In contrast, Wenger et al. (2016) finds that higher unemployment rates lead to accession into the Reserves at lower rates. This finding was supported qualitatively within the focus groups, which identified the perception that the Reserves would disadvantage individuals in job competitions with their civilian peers due to periodic absences for military duties. Focus group participants also noted that economic benefits such as tuition assistance, health care, and bonus pay were a driving influence towards accession in the Reserves. Reserve recruiters commented that applicants see the Reserves as an avenue for supplemental income (Wenger et al., 2016).

Wenger et al. (2016) emphasize the importance of geographic factors to an individual's decision to join the Reserves. This is because an individual can only participate in the Reserves if there is a position available in the area in which they choose to live. Their study finds that there was no evidence that individuals' Reserve availability influenced their geographic preferences when separating from full-time service. Most focus group respondents confirmed individuals considered Reserve affiliation only after relocation. Wenger et al. (2016) find that individuals who leave full-time service in states with more Reserve positions available are more likely to join the Reserves. The authors attribute this to knowledge about Reserve opportunities which drive individuals to affiliate with the Reserves after relocation. Their study recommends the most effective Reserve recruitment efforts should be targeted to individuals leaving full-time service (Wenger et al., 2016).

Further, Wenger et al. (2016) also discuss individuals' experience in full-time service across the categories of overall positive or negative opinion of their full-time service, job applicability in the civilian sector, and separation reasons. They find that accession in the Reserves is positively correlated with being deployed. Moreover, individuals with a longer deployment history are more likely to affiliate with the Reserves more quickly after separating from full-time service. Qualitatively, the recruiter focus groups indicate individuals with poor opinion on their full-time service are harder to hire into the Reserves. Recruiters comment that applicants may use the Reserves to gain employment or civilian job skills while others may be unhappy after leaving full-time service and use the Reserves as an avenue to return to full-time service (Wenger et al., 2016).

3.6 Australian Defence Force Reserves

Having identified key features influencing Reserve participation in U.S. military studies, we next consider the practical value of predicting Reserve outcomes within the context of Australian Defence workforce planning. The ADF's Review emphasized the importance of a well-resourced Reserve workforce, while identifying Australian labor market challenges as contributing to an understrength force (Department of Defence, 2024).

Despite extensive documentation of Reservists' contributions in Australian military history, researchers agree that there remains a lack of reservist-focused analysis, particularly amid declining Reserve force manning across Western militaries (Armstrong, 2020; West, 2018; West & Healy, 2025). Lording's (2015) qualitative research into Australian Army reservists found that Reserve participants perceive their service encompasses elements of both paid employees and unpaid volunteers. The study finds Reservists participate for volunteer-driven benefits such as identity and social network purposes while also being motivated by employee-related benefits, such as remuneration and training. Lording (2015) recommends Defence develop Reserve-focused HR policies to accommodate Reservists' intrinsic and extrinsic motivators.

While prior research begins to describe Reservists, there exists a gap in Australian literature that documents the characteristics of those who choose to join and actively participate in the Reserve force. The gap underscores the need for a predictive framework capable of identifying which separating members are most likely to affiliate and actively participate in the ADF Reserves.

3.7 Summary

In this chapter, we establish the conceptual and empirical foundations for modeling ADF Reserve participation. We show how proxy characteristics can estimate the economic trade-offs that influence multiple job holding, drawing on the Random Utility Model and Moonlighting Model. We then review three ML models, Adaptive Boosting, Random Forest, and eXtreme Gradient Boosting, and summarize evidence of their effectiveness in civilian and military HR contexts. We also review previous U.S. Reserve service research, which identifies four categories of influential characteristics that inform feature selection in our models. Lastly, we highlight a significant gap in Australian literature analyzing who participates in the ADF Reserves. To address this gap, the next chapter applies the selected ML models to assess the ADF's capacity to predict Reserve participation and identify key determinants of Reserve workforce engagement.

CHAPTER 4: Methodology

We use existing ADF and Australian Bureau of Statistics (ABS) data to predict a separating full-time ADF member's SERCAT choice and subsequent Reserve service behavior. This chapter outlines the methodological approach and justifications used to generate these predictions.

4.1 Methodology Overview

Figure 4.1 illustrates the methodology we use to test the effectiveness of ML models to predict separating ADF members' SERCAT choice and Reserve participation behavior. As depicted in Figure 4.1, this chapter is organized according to the following steps:

1. Acquire ADF datasets
2. Pre-process data
3. Conduct feature engineering
4. Split data into training and test sets
5. Train then test models
6. Analyze model results

Steps one through five are discussed in the following sections, with step six, Analyze model results, addressed in Chapter 5.

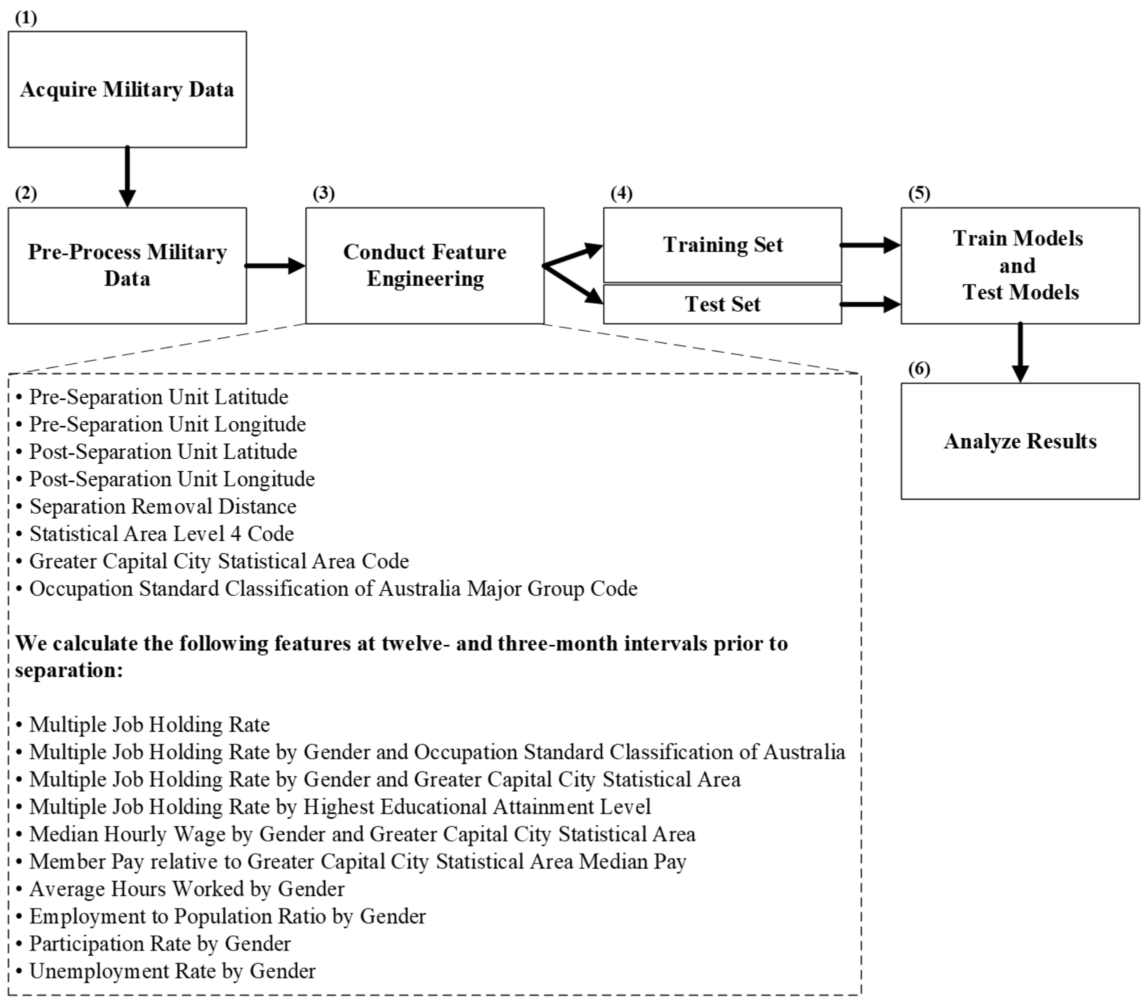

Figure 4.1. This research methodology comprises six steps. First, we acquire the military data from the ADF. Next, we merge, clean, and filter the data. Third, we create 28 features using complementary data sources. In the fourth step, the complete dataset is split into training and test sets. In the fifth step, the models are trained and then evaluated. Lastly, the results are analyzed.

4.2 Data Acquisition

The dataset we use for this thesis was provided by Defence People Group within the ADF from information contained within the ADF's Management, Analysis and Reporting Solution. Five comma-separated values files were provided, each containing distinct characteristics about the observations in the sample. The Decision dataset includes the demographic, military, and Reserve SERCAT choice information; the Operations dataset provides the deployment history; the Education dataset contains academic information; the Pay dataset contains hourly wage rates; and the Reserve Attendance data records the number of days and hours of completed Reserve service. We merge these datasets to construct the initial sample prior to pre-processing.

The dataset contains all SERCAT 6 or SERCAT 7 ADF personnel who separate from full-time service between 01 July 2016 and 30 June 2025 (FY 2016–17 to FY 2024–25)[1]. The raw dataset contains 48,422 rows, with each row corresponding to a distinct de-identified ADF member. Each observation is assigned a unique identification number, `Party_ID`, generated by the Management, Analysis and Reporting Solution. This identification number is not associated with their service number. Each observation is described by 84 variables across the five files, providing demographic, geographic, and military service information at the point of separation.

4.3 Data Pre-Processing

We begin the analysis by importing the five datasets into Python using the Pandas package (Pandas Development Team, 2026). The datasets are merged on the `Party_ID` variable. This creates a single dataset in which each row represents a separated ADF member and the columns encompass all the relevant observable characteristics provided by the Management, Analysis and Reporting Solution.

With a single raw dataset, we filter the sample to retain the Reserve SERCATs of interest. We drop 22,054 observations whose separation SERCAT choice is SERCAT 7, SERCAT 4,

[1] An Australian Government FY is a 12-month period commencing on 01 July, as defined in the *Income Tax Assessment Act 1997* (s. 995.1) (e.g., FY 2016/17 begins on 01 July 2016 and ends on 30 June 2017).

or SRC N/A.[2] This filtering restricts the sample to the Reserve SERCATs 2, 3, and 5. We additionally remove two observations with a `Gender_Code` of `X` to allow binary coding of the gender variable. The resulting processed dataset contains 26,366 observations of separated ADF members across the Royal Australian Navy, Australian Regular Army, and the Royal Australian Air Force.

4.4 Feature Engineering

We created 28 new features by manipulating the existing data and incorporating new data. This section outlines the process undertaken to manipulate existing data, create geographic features, and incorporate economic data.

4.4.1 Existing Data Manipulation

We create eight variables by manipulating categorical and temporal data provided by the Management, Analysis and Reporting Solution. We construct a binary `Religious` variable from the 59-category variable `Religion_Short_Description`, indicating whether an observation reports a religious preference. We add the `Job_Family_Change` variable to indicate whether an observation had a change in job category at separation.

Using the job role mappings published by Department of Defence (n.d.), we aggregate the 248 ADF job roles to eight higher-level strategic workforce segments, producing the `Segment_Long_Description` variable. We map categorical educational attainment levels ordinally in accordance with the Australian Qualifications Framework (Australian Qualifications Framework Council, 2013). We then pivot the data to identify the highest educational attainment level, `Max_Degree_Value`, and number of recorded educational attainments, `Count_Degrees`, for each observation.

Lastly, we transform date variables to reflect the number of months between each event and separation, resulting in variables capturing the time between marriage and separation (`Months_Between_Marriage_Separation`), the latest promotion and separation (`Months_Between_Promotion_Separation`), and enlistment and separation (`Months_`

[2] A `SERCAT_Code` of SERCAT 7 indicates that the member was not required to complete Reserve service. The majority of these observations reflect involuntary separations, including being medically unfit for service, retention not in service interest, or training failure.

Between_Enlistment_Separation). These features reduce dimensionality and enable the model to identify additional patterns in the data.

4.4.2 Geographic Feature Creation

We create five features to investigate geographic patterns and enable the incorporation of region-specific economic data. To generate these features, we rely on two variables in the dataset that describe the city and state of the military unit to which a member is assigned. `Prior_Actual_Location_Long_Description` represents the city and state of the unit a member was assigned to prior to separation, while `Actual_Location_Long_Description` is the city and state of the unit to which the member is assigned within Reserve service. We compile a list of unique city-state combinations using these two variables.

Next, we use the Geocodify application programming interface to match each city-state combination to the closest geographic location and obtain standardized location names, latitudes, and longitudes (Geocodify, n.d.). This process yields proxy geographic coordinates for all military unit locations. We merge these coordinates back to each observation using `Prior_Actual_Location_Long_Description` and `Actual_Location_Long_Description`, creating four new features: `Prior_Lat`, `Prior_Lon`, `Actual_Lat`, and `Actual_Lon`. Using these coordinates, we calculate `Separation_Removal_Distance` with the Haversine formula. This records the great-circle distance between each member's pre-separation and post-separation unit locations.

We then use these coordinates to spatially link each observation to region-specific ABS data defined by Australian Statistical Geography Standard digital boundaries. The ABS defines a seven-level hierarchical structure for these geographic classifications (Australian Bureau of Statistics, 2021b). The highest level begins with Australia, followed by States and Territories (S/T), progressively smaller Statistical Areas Levels 4 through 1, and concludes with Mesh Blocks as the smallest functional geographic unit. Table 4.1 outlines the number of geographic units at each level of the hierarchy by region. We use Statistical Areas Level 4 (SA4) and Greater Capital City Statistical Areas (GCCSA) Statistical Areas classifications based on the availability of relevant ABS statistical releases. These geographic levels also support the proxy assumption that a member resides within the same SA4 and GCCSA as their assigned military unit.

Table 4.1. This table presents the number of geographic units in the Australian Statistical Geography Standard Edition 3 by state and territory across the seven hierarchical geographic levels. The bold rows represent the geographic unit levels we use to incorporate ABS data. Source: Australian Bureau of Statistics (2021b).

Region	NSW	Vic.	Qld	SA	WA	Tas.	NT	ACT	OT	Aust.	Total
S/T	1	1	1	1	1	1	1	1	9	10	
GCCSA	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**3**	**3**	**34**	**35**
SA4	**30**	**19**	**21**	**9**	**12**	**6**	**4**	**3**	**3**	**107**	**108**
SA3	94	68	84	30	36	17	11	12	6	358	359
SA2	644	524	548	176	267	101	70	136	6	2,472	2,473
SA1	19,750	15,482	12,549	4,329	6,352	1,482	649	1,229	22	61,844	61,845
MB	112,738	88,739	71,862	28,431	43,322	13,034	3,361	6,662	136	368,285	368,286

Note: The columns reflect the abbreviations of Australia's states and territories. OT represents Other Territories. The Total column includes one residual geographic category not allocated to any other regions.

These SA4 and GCCSA codes form the basis for integrating region-based ABS economic and labor market data, as discussed in the following section. We use the shapefile provided by the ABS (2021a) to incorporate these spatial levels within our dataset. We load the ABS SA4 shapefile into Python using the GeoPandas package and spatially assign each observation to an SA4 code and GCCSA code based on the pre-separation unit's geographic coordinates (Australian Bureau of Statistics, 2021a; GeoPandas Developers, 2025).

4.4.3 Economic Feature Creation

We created 20 features based on data sourced from the ABS. Similar to the use of SA4, we map individuals' military roles to an Occupation Standard Classification for Australia (OSCA) major group to enable linkage with selected ABS datasets. The major group is an eight-category level representing the broadest classification within the OSCA framework.

Under the OSCA, ADF members are generally classified by rank group rather than job role, which limits the suitability of these classifications for direct alignment with civilian occupations (e.g., 149231 Commissioned Defence Force Officer, 451131 Defence Force Member - Other Ranks) (Australian Bureau of Statistics, 2024b, 2024c). To address this limitation, we map the 292 unique `Category_Short_Description` job roles in the dataset to their corresponding OSCA major groups. This mapping is conducted using the OSCA Search tool, which enables alignment of military roles with their closest civilian occupational equivalents (Australian Bureau of Statistics, 2024d).

After adding SA4 and OSCA, we integrate economic and labor data from the ABS. We generate all economic features at two pre-separation reference points: twelve months and three months prior to each individual's separation date. We select these periods to proxy labor market and economic conditions at decision stages. One year represents when individuals may begin actively considering separation. Three months typically represents the minimum separation notification period required by the ADF, during which members select their Reserve SERCAT. To identify statistics corresponding to these periods, we construct placeholder date variables set twelve and three months prior to the recorded separation date and align each observation to the temporally closest relevant ABS values using these dates.

We use seven tables from three ABS statistical releases to create economic features within the study's dataset:

- 6217.0 *Multiple job-holders* (Australian Bureau of Statistics, 2025a):
  - Table 1 Multiple job-holders rates and levels
  - Table 5.2 Employed persons rates, by occupation, by job-holding status, by sex (2006 - present)
  - Table 5.7 Employed persons rates, by Greater Capital City Statistical Area (1998 - present)
  - Table 5.8 Employed persons levels and rates, by educational qualifications (2015 - present)
- 6337.0 *Employee Earnings* (Australian Bureau of Statistics, 2024e):
  - Data 2 Median earnings for employees by demographic characteristics
- 6291.0.55.001 *Labour Force, Australia, Detailed* (Australian Bureau of Statistics, 2025b):
  - Table 09. Employed persons by hours actually worked in all jobs and sex
  - Table 16. Labour force status by Labour market region and Sex

We use four tables from the quarterly published ABS *Multiple job-holders* statistical release (Australian Bureau of Statistics, 2025a). We record the nation-level multiple job-holding rate from Table 1 of this statistical release. We use Table 5.2 to extract the multiple job-holding rate corresponding to each observation's gender and OSCA classification. Similarly, we use Table 5.7 to incorporate the multiple job-holding rate matching each observation's GCCSA region and gender. Finally, Table 5.8 provides the multiple job-holding rate by highest educational attainment level for each observation. These four variables capture variation in multiple job-holding rates across demographic characteristics aligned with each observation.

We use one annually published data table from the *Employee Earnings* statistical release (Australian Bureau of Statistics, 2024e), which provides median hourly wage rate disaggregated by gender and GCCSA. Using this, we create two variables. First, we assign the median hourly wage rate to each individual based on their gender and GCCSA. Second, we construct a variable capturing the differential between an individual's hourly ADF pay rate and the corresponding gender-region median. This variable represents an individual's relative economic position compared with the median worker in their geographic area.

Finally, the monthly *Labour Force, Australia, Detailed* provides two relevant tables (Australian Bureau of Statistics, 2025b). Table 09 reports the average hours worked across all jobs per employed person, disaggregated by gender. This measure captures the typical workload faced by an observation's gender cohort when considering separation and subsequent employment time demands. Meanwhile, Table 16 provides national gender-specific measures of the employment-to-population ratio, labor force participation rate, and unemployment rate. Together, these features reflect prevailing labor market conditions at the time an observation determines their SERCAT choice. After incorporating these economic features, we transition to splitting the dataset into a training set and a test set.

4.5 Data Split

We split the data to train the models on one portion of the dataset and to evaluate their performance on a separate, unseen test set that is held constant across all three models. A common convention in ML model development is to retain a random 20% of a dataset for testing while using the remaining 80% for training. However, there is no prescriptive guidance on an optimal split ratio, and alternative ratios such as 70:30, 60:40, or 50:50 are commonly used depending on the data and approach (Joseph, 2022).

To better assess the managerial application of predicting SERCAT choice and Reserve participation behavior for future separations, we instead opt to use a temporal data split. This process mirrors the operational context in which the models would be applied. We train the models on observations from the first seven financial years of data ($n = 21,633, 82\%$) and reserve the most recent two financial years as the test set ($n = 4,733, 18\%$).

4.6 Train and Test Models

We use both classification and regression models to answer the two primary research questions. First, we use classification models to predict SERCAT choice and binary Reserve participation. We implement these models using the `AdaBoostClassifier` and `RandomForestClassifier` classes from the `sklearn.ensemble` module and the `XGBClassifier` class from the `xgboost` package (Scikit-learn Developers, 2011; XGBoost Developers, 2011).

Second, we use regression models to predict the number of hours of Reserve service a member completes in the first full financial year following separation. We implement these models using the `AdaBoostRegressor` and `RandomForestRegressor` classes from the `sklearn.ensemble` module and the `XGBRegressor` class from the `xgboost` package (Scikit-learn Developers, 2011; XGBoost Developers, 2011).

To answer the research questions, we apply a three-stage modeling framework illustrated in Figure 4.2. First, the SERCAT Choice model predicts a separating member's Reserve category (i.e., SERCAT 2, 3, or 5). Second, conditional on being predicted as SERCAT 3 or 5, the Participation model predicts whether a member completes any Reserve service in the first financial year after separating. Third, conditional on participation, the Intensity model predicts the number of Reserve service hours completed.

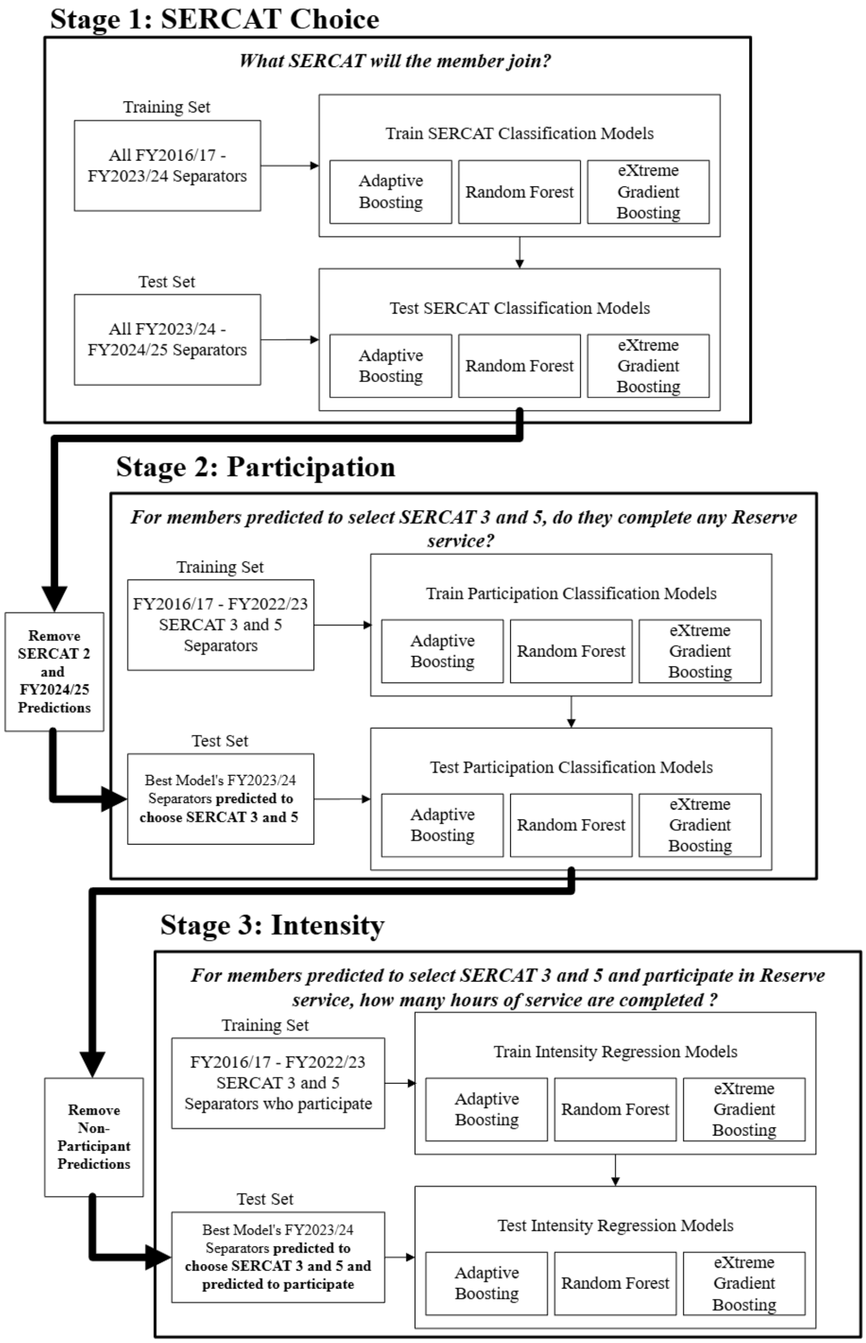

Figure 4.2. This figure illustrates a three-stage modeling framework we use to predict separating members' Reserve SERCAT choice, Reserve participation conditional on category choice, and the intensity of Reserve service conditional on participation. The best model is determined using model-specific evaluation metrics.

4.6.1 Stage 1: SERCAT Choice Classification Model

We use the SERCAT Choice classification model to predict Reserve service category choice. We train the classification models using 70 features of all separated ADF members between FY2016–17 and FY2022–23, with `SERCAT_Code` as the outcome variable. We then evaluate the trained models using all separated ADF members within FY2023–24 and FY2024–25 by comparing their predicted `SERCAT_Code` values with recorded `SERCAT_Code` values. Among Adaptive Boosting, Random Forest, and eXtreme Gradient Boosting, we carry forward predictions from the model with the highest macro-averaged one-versus-rest area under the curve (AUC) into the subsequent stage of the framework.

4.6.2 Stage 2: Participation Classification Model

We use the Participation classification model to predict a member's likelihood of completing any Reserve service in the first FY following separation, conditional on being predicted to join SERCAT 3 or 5. For model training, SERCAT 2 observations are excluded, restricting the sample to members entering SERCAT 3 or SERCAT 5 between FY2016–17 and FY2022–23. The full feature set is retained. The response variable is defined as 1 if `Unit_Result_Total` > 0 and 0 if `Unit_Result_Total` = 0.

For the test set, we filter the predictions from the SERCAT Choice model by excluding SERCAT 2 predictions and all FY2024–25 observations. The same binary response definition is applied. Observations from FY2024–25 are excluded at this stage due to the absence of complete Reserve participation data. Among Adaptive Boosting, Random Forest, and eXtreme Gradient Boosting, we carry forward predictions from the model with the highest binary AUC into the final stage of the framework.

We include this interim classification step because 60.7 percent of members complete no Reserve service in this first FY post-separation, resulting in a highly imbalanced response variable. Without this stage, the models exhibit limited predictive power for Reserve service intensity.

4.6.3 Stage 3: Intensity Regression Model

We use the Intensity regression model to predict the number of hours a member completes in Reserve service, conditional on being predicted to join SERCAT 3 or 5 and to participate in the Reserves. We train this model using SERCAT 3 or 5 members from FY2016–17 to FY2022–23 who completed at least one hour of Reserve service in the first full FY following separation, excluding observations with no recorded participation. The response variable is the continuous measure of Reserve service hours, `Unit_Result_Total`. The trained models are evaluated using the subset of FY2023–24 SERCAT 3 and 5 members who are predicted to participate by the Participation classification model. The results from this model, together with those from the preceding classification stages are analyzed in Chapter 5.

4.7 Data Limitations

Prior to the discussion of results, we acknowledge several limitations related to the data. First, data entry errors warrant consideration, as the manner in which they are handled may influence the study's findings. Within the raw Reserve attendance data, 140 person-FY observations record negative Reserve service hours. As negative participation is not meaningful in this context, these values are recorded as zero and retained in the dataset. In addition, approximately six percent of SERCAT 2 members record non-zero Reserve service hours, `Unit_Result_Total`. This is inconsistent with ADF guidance, which requires members to transfer into SERCAT 3 or 5 in order to complete Reserve service. These observations are retained; however, they are expected to have a minimal impact on the results, as SERCAT 2 members are excluded between the SERCAT Choice and Participation modeling stages.

Second, we are unable to obtain home location data for observations in the dataset due to the lack of historical address retention within ADF systems. As a result, we assume that members reside within the same SA4 and GCCSA as their assigned military units when incorporating ABS economic data. This proxy may fail in cases where members work remotely from their assigned unit, do not update their address following separation, or are assigned to Reserve units geographically distant from their place of residence. Given the available data, we are unable to identify or correct instances in which this assumption does not hold.

Third, ten observations list an unknown value for their pre-separation unit location. Without this information, these observations cannot be matched to relevant ABS geographic data. Given the limited scope of this issue, we assign each affected observation's post-separation unit location as its pre-separation unit location for the purpose of incorporating economic variables.

Fourth, the ABS OSCA categorizes ADF personnel by rank group rather than by occupational role. Prior to the release of OSCA in 2024, the ABS used the Australian and New Zealand Standard Classification of Occupations, under which the ADF had internally mapped its job roles (Australian Bureau of Statistics, 2024a). As no equivalent mapping currently exists between ADF job categories and OSCA, we rely on informed judgment to assign ADF roles to the eight OSCA major groups.

Finally, this dataset is limited by its lack of demand-side information. As a result, we are unable to account for how organizational changes, such as Reserve workforce growth, policy adjustments, or unit-level demand, may have influenced individual Reserve service decisions over the observation period. The analysis reflects only supply-side behavior, as observable in individual characteristics and choices, rather than the interaction between individual preferences and institutional demand.

This chapter outlines the methodology employed in this research. We describe the acquisition and processing of the data, the engineering of additional features, the data-splitting approach, and the three-stage modeling framework. Having established the methodological design and acknowledged the data limitations, the following chapter presents and discusses the results of all three models.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5: Results

This chapter presents and analyzes the results of the three-stage modeling framework. For each stage, we report performance metrics across all three candidate model types, identify the model selected for subsequent analysis, and examine the results of the selected model in greater depth. We also assess the feature importance at each stage to highlight the primary drivers of model predictions. We begin by outlining the performance metrics used to evaluate the classification models in Stage One and Stage Two.

We assess classification performance using accuracy, area under the curve (AUC), precision, recall, and F1 scores following definitions by Manning et al. (2008). Precision measures the proportion of predicted positive observations that are true positives, while recall measures the proportion of actual positive observations that are correctly classified. The F1 score summarizes the trade-off between precision and recall and is sensitive to both false positives and false negatives. Accuracy measures the proportion of correctly classified observations across all classes. AUC summarizes the classification model's ability to discriminate between classes across all possible thresholds by measuring the area under a receiver operating characteristic curve, which plots the true positive rate against the false positive rate (Manning et al., 2008).

We assess regression performance using root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) following definitions by Kuhn and Johnson (2013) and Willmott and Matsuura (2005). RMSE measures the magnitude of prediction error in the units of the outcome variable and penalizes larger errors more heavily by squaring deviations before averaging. It can be interpreted as the average distance between the observed values and the predicted values (Willmott & Matsuura, 2005). MAE measures the average absolute prediction error in the units of the outcome variable, treating all errors equally and reflecting typical model performance (Willmott & Matsuura, 2005). The coefficient of determination, R^2, summarizes the proportion of variance in the dependent variable explained by the model (Kuhn & Johnson, 2013).

Lastly, we use the SHapley Additive exPlanations (SHAP) framework, proposed by Lundberg and Lee (2017) and interpreted by Molnar (2025), to generate feature importance and summary plots. SHAP provides a mechanism to interpret complex models by attributing a model's prediction for a given observation to the contributions of its individual features.

SHAP feature importance plots rank features according to their mean absolute Shapley values across observations, where larger values indicate greater overall influence on model prediction. SHAP summary (beeswarm) plots extend this by displaying the distribution of Shapley values for each feature. In these plots, each point represents the Shapley value of a feature for a single observation, and features are ordered on the y-axis by their average absolute contribution to the model's prediction. We use the `shap` Python package to generate these plots (Lundberg & Lee, 2025).

To support interpretation of SHAP results for categorical features, variables are one-hot encoded using `pd.get_dummies` with `drop_first=True` (Pandas Development Team, 2026). This encoding omits one reference category per categorical variable to avoid perfect multicollinearity. Consequently, SHAP values for one-hot encoded features are interpreted relative to the omitted reference category.

5.1 Stage One: SERCAT Choice Classification Model Results

Stage One begins by training three classification models, Adaptive Boosting, Random Forest, and eXtreme Gradient Boosting, using all separating members from FY2016–17 to FY2022–23. We tune the model hyperparameters using 100 iterations of scikit-learn's `RandomizedSearchCV` with five-fold cross validation, optimized for one-versus-rest AUC (Scikit-learn Developers, 2011). We then evaluate the tuned models on the response variable, `SERCAT_Code`, with the test set comprising separating members from FY2023–24 to FY2024–25.

5.1.1 Model Comparison and Selection

Table 5.1 presents a comparative summary of model performance when predicting among the three SERCAT categories. Overall, the three models perform very similarly across all evaluation metrics. Classification accuracy differs by only 2.2 percentage points, suggesting that each model captures broadly similar patterns in the underlying data. All three models achieve accuracy in the low-70 percent range, substantially exceeding the 33 percent accuracy expected under random classification across three outcome categories.

eXtreme Gradient Boosting demonstrates a modest but consistent performance advantage by achieving the highest values across all evaluation metrics. Its highest F1 score indicates a balanced trade-off between false positives and false negatives compared to the other models. Given its consistent performance across all metrics, we select eXtreme Gradient Boosting as the preferred Stage One model. Predictions generated by this model are carried forward into Stage Two. The following sections provide additional metrics regarding the eXtreme Gradient Boosting model's performance and predictions.

Table 5.1. In Stage One, we report test set performance using hyperparameters optimized via cross-validated one-versus-rest AUC. Precision, recall, and F1 scores are macro-averaged across all classes. We carry forward eXtreme Gradient Boosting's predictions to inform Stage Two based on its superior performance across accuracy, AUC, and macro-averaged classification metrics.

Model	Accuracy	AUC	Precision	Recall	F1
Adaptive Boosting	0.723	0.860	0.721	0.723	0.714
Random Forest	0.727	0.858	0.725	0.727	0.716
eXtreme Gradient Boosting	**0.745**	**0.879**	**0.744**	**0.745**	**0.734**

5.1.2 Confusion Matrix and Classification Metrics

Figure 5.1 is the confusion matrix for the eXtreme Gradient Boosting model, summarizing the relationship between predicted and observed SERCAT outcomes. The model demonstrates strong classification performance for SERCAT 3 and SERCAT 5, with correct classification rates of approximately 85 and 77 percent respectively. This indicates the model is particularly effective at identifying members who transition into Reserve SERCATs associated with active participation.

In contrast, the classification performance for SERCAT 2 is notably weaker. Only 36 percent of true SERCAT 2 observations are correctly classified, with the majority misclassified as SERCAT 3. This pattern indicates the model struggles to distinguish between SERCAT 2 and SERCAT 3 observations at the point of separation. As shown later through feature importance analysis, this difficulty reflects substantial overlap in the observable characteristics associated with these two service categories. By comparison, misclassifications between SERCAT 2 and SERCAT 5 are rare: approximately ten percent of true SERCAT 2 observations are classified as SERCAT 5, while fewer than one percent of true SERCAT 5 observations are classified as SERCAT 2.

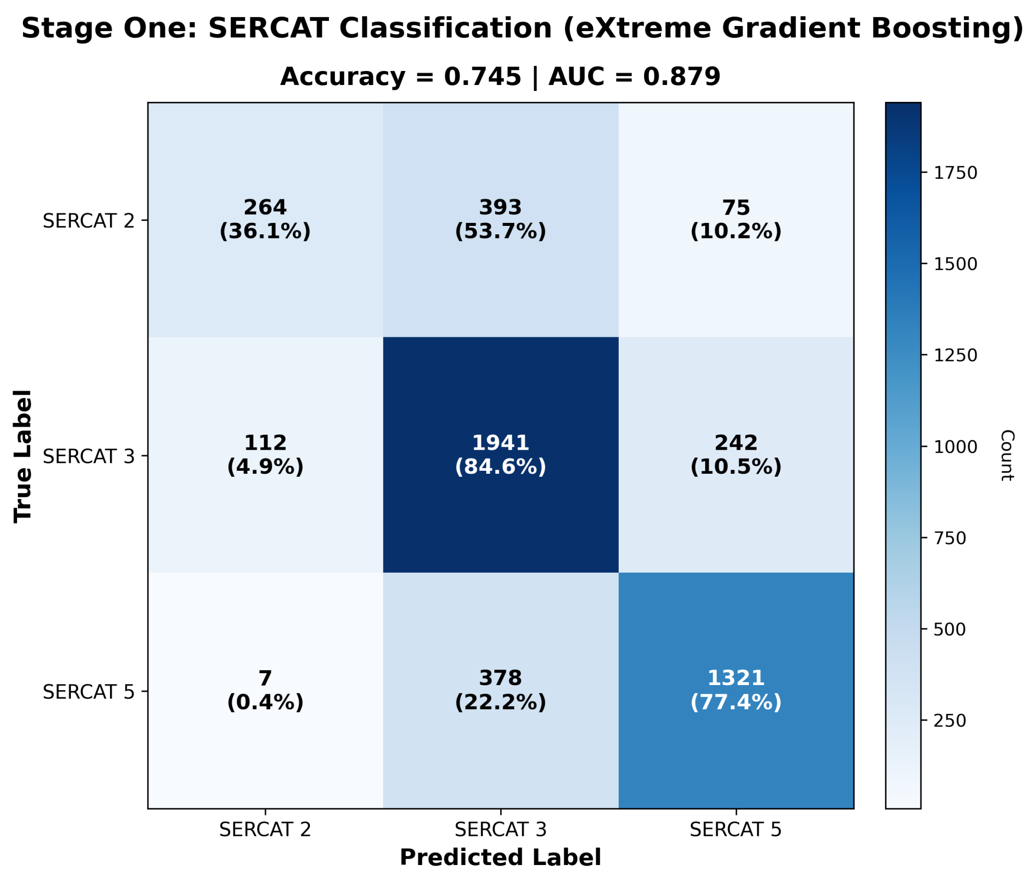

Figure 5.1. This confusion matrix illustrates the eXtreme Gradient Boosting model's strong classification performance for SERCAT 3 and SERCAT 5, alongside substantial misclassification between SERCAT 2 and SERCAT 3.

These same classification patterns are also reflected in Figure 5.2, which presents the per-class performance metrics of the eXtreme Gradient Boosting model. The model predicts SERCAT 2 conservatively, as indicated by a relatively high precision score alongside its low recall score. This pattern suggests that while predictions labeled as SERCAT 2 are often correct, the model fails to identify a substantial proportion of the true SERCAT 2 observations, instead misclassifying them as other categories. In cases where observations lie near the decision boundary and exhibit limited discriminative signal, the model tends to assign them to an adjacent service category, most commonly SERCAT 3, rather than predicting SERCAT 2. This tendency may be reinforced by the higher prevalence of SERCAT 3 in the training data.

In contrast, the model achieves high recall for SERCAT 3, successfully identifying the majority of the true SERCAT 3 members. Performance for SERCAT 5 is more balanced, with comparatively high precision and recall, indicating that the model both correctly identifies most SERCAT 5 observations and produces relatively few false-positive predictions for this category.

Together, the confusion matrix and per-class metrics indicate that classification errors are concentrated between the SERCAT 2 and SERCAT 3 categories, rather than being evenly distributed across categories. This pattern reflects the conceptual and administrative proximity of these two service categories.

The difficulty in distinguishing between SERCAT 2 and SERCAT 3 may also reflect aspects of the separation process itself. Members selecting either category often exhibit highly similar observable characteristics, as examined in more detail in the feature importance section. Institutional guidance explicitly encourages members who initially select SERCAT 2 to consider transferring to SERCAT 3 if they intend to render service within two years of their transfer date. Similarly, members selecting SERCAT 3 are advised to consider SERCAT 2 if they do not intend to provide service within that period. This administrative nudging may further blur distinctions between these two categories at separation, contributing to systemic misclassifications between SERCAT 2 and SERCAT 3 in the model outputs.

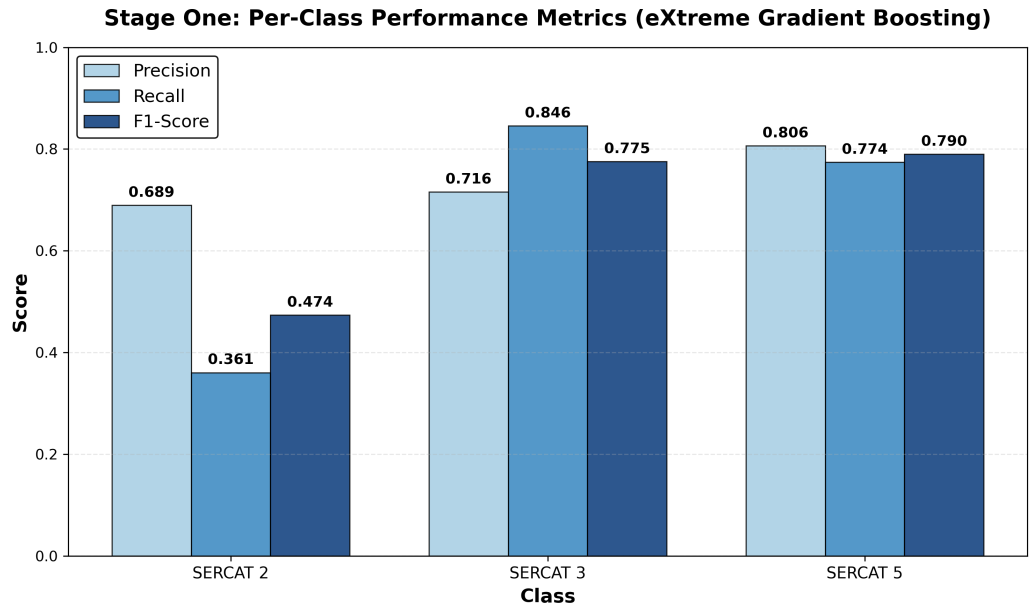

Figure 5.2. This figure reports precision, recall, and F1 score by SERCAT class for the eXtreme Gradient Boosting model. The model achieves weaker recall for SERCAT 2 and stronger, more balanced performance for SERCAT 5.

5.1.3 Feature Importance Analysis

Using the results of the eXtreme Gradient Boosting model, we next calculate the mean absolute Shapley values to assess feature importance, as illustrated in the Feature Importance Stacked by Service Category in Figure 5.3 and the SHAP beeswarm plot in Figure 5.4.

This feature importance analysis helps explain the classification patterns observed in the Stage One confusion matrix (Figure 5.1), particularly the concentration of errors between SERCAT 2 and SERCAT 3. Figure 5.3 illustrates that SERCAT 5 is characterized by larger SHAP value magnitudes across the most influential features, indicating stronger and more consistent effects in the model's predictions. In contrast, predictions for SERCAT 2 and SERCAT 3 are driven by the same core features, including separation move distance, time in last SERCAT, pre-separation Service affiliation, geographic location, and pay, but with weaker effects. For these two categories, these corresponding SHAP values exhibit similar magnitudes and substantial overlap rather than distinct directional patterns, limiting the model's ability to cleanly separate SERCAT 2 and SERCAT 3.

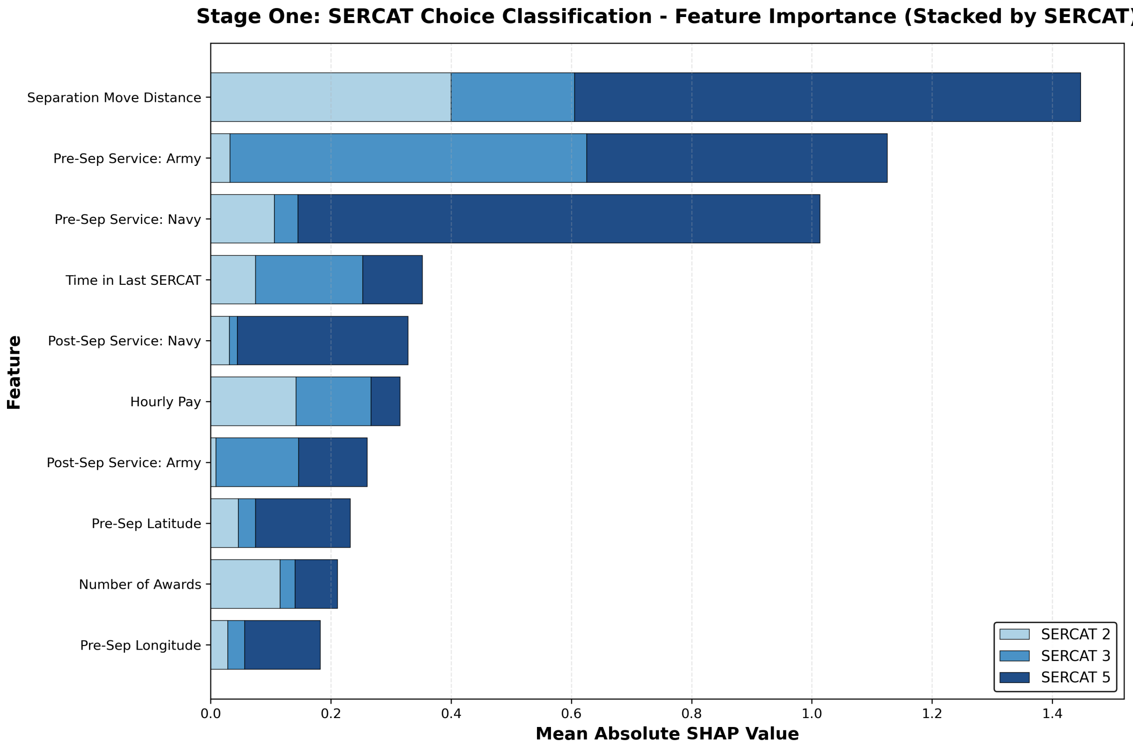

Figure 5.3. This figure plots the mean absolute SHAP values for the Stage One SERCAT Choice model, stacked by predicted SERCAT outcome. Each bar represents a feature's total contribution to model predictions, while the stacked segments show how that contribution is distributed across SERCAT 2, SERCAT 3, and SERCAT 5. This figure highlights that SERCAT 5 is associated with strong feature effects, whereas SERCAT 2 and SERCAT 3 rely on similar features with overlapping magnitudes, helping explain the concentration of misclassification between these two categories.

Figure 5.4 presents the same set of influential features while additionally illustrating how feature values relate to the magnitude of contributions to the model's predictions across SERCATs. Specific geographic characteristics, such as the distance of move after separation and pre-separation geographic location, emerge as dominant predictors in the sample. The long-tail distribution of SHAP values for these variables indicates diversity in their influence on members, with specific cohorts having a larger impact than the majority. This pattern suggests that anticipated geographic relocation and pre-separation unit location are strongly associated with predicted SERCAT outcomes.

Institutional characteristics, including pre- and post-separation Service, also exhibit high importance in the beeswarm. As these variables are binary, their SHAP values reflect discrete differences in predicted SERCAT outcomes by Service rather than gradual effects. Service affiliation is encoded using one-hot indicators with Air Force serving as the omitted reference category; accordingly, SHAP values for Army and Navy Service reflect differences relative to the Air Force. Service affiliation in the Navy produces relatively concentrated and directionally consistent SHAP values, indicating a stable contribution to model predictions. In contrast, the effects associated with Army Service are more dispersed, suggesting greater variability in how Army affiliation influences predicted SERCAT choices across individuals.

The difference in dispersion is consistent with structural differences in separation and Reserve participation by Service. The model results are consistent with the transition rates illustrated in Table 2.1. Historical separation data show that approximately 76 percent of Navy members transition to SERCAT 3, whereas Army members transition across all three SERCATs in a more balanced manner (Department of Defence, 2024). These patterns help contextualize the Service-specific effects observed in the feature importance analysis.

The economic variable of hourly pay similarly exhibits non-uniform effects in the global beeswarm (Figure 5.4). Higher hourly pay rates are associated with larger SHAP values for a subset of higher-paid members, while contributing minimally to the model's predictions for the remainder of the population. This pattern suggests that higher remuneration is informative for some individuals but not universally predictive across the sample.

Disaggregating these effects by SERCAT reveals asymmetric SHAP patterns across SERCATs. Hourly pay contributes positively to predictions of SERCAT 3 while pushing predictions away from SERCAT 2 (Figure C.1 and Figure C.2). These effects are concentrated among higher-paid members and are consistent with pay acting as a proxy for correlated rank- and role-related characteristics. This pattern is not necessarily intuitive, as higher pay might be expected to correlate with stronger civilian employment opportunities and weaker incentives for continued Reserve affiliation. The observed SHAP effects highlight an empirical association, suggesting that the relationship between remuneration and SERCAT choice is more complex than that of a simple financial trade-off.

As a whole, these results suggest that SERCAT choice reflects the anticipated constraints and opportunities when planning for separation, such as geographic choice and Service-specific pathways. Additional SHAP visualizations illustrating by-SERCAT SHAP plots are provided in Appendix C.

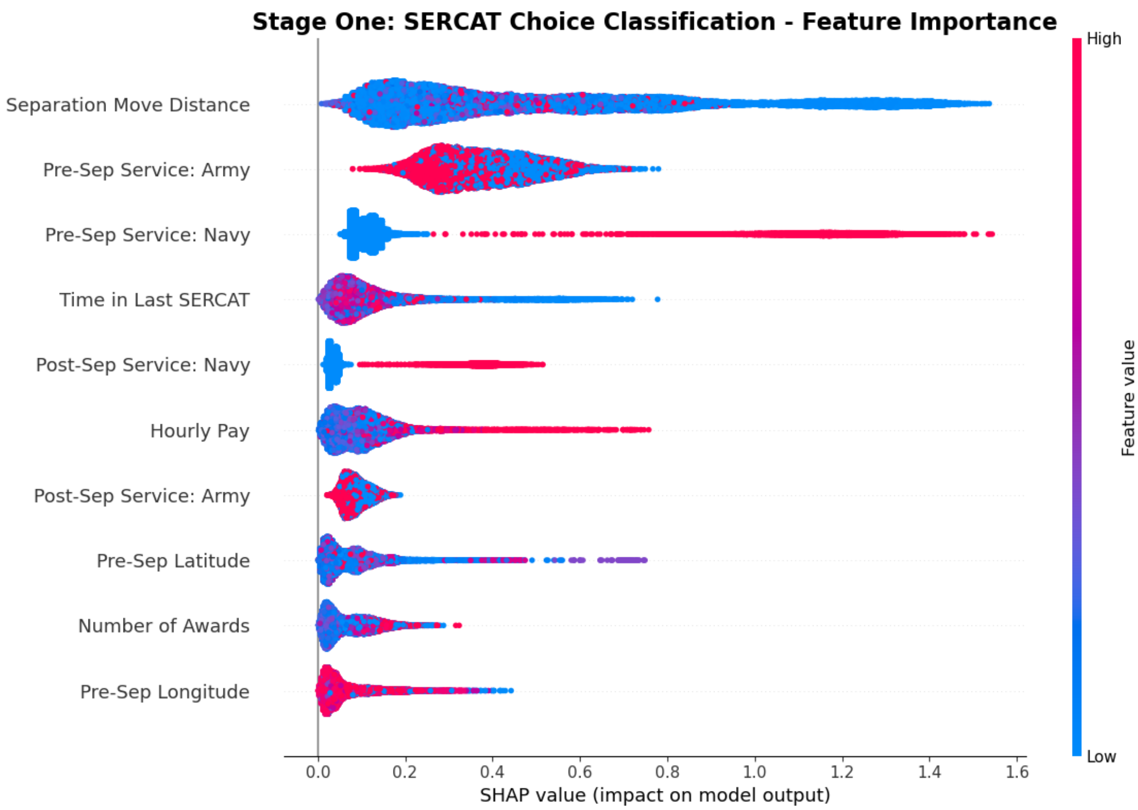

Figure 5.4. This visualization displays the distribution of SHAP values across the most influential features in the SERCAT Choice model. In the multiclass setting, SHAP values reflect the magnitude of each feature's contribution to the model's predicted probabilities across SERCATs rather than towards any single category in isolation. Geographic mobility measures, Service affiliation, and remuneration-related characteristics emerge as the most influential predictors of SERCAT outcomes at the point of separation.

5.2 Stage Two: Participation Classification Model Results

Stage Two uses the same three types of classification models to predict whether a member completes any Reserve service, conditional on being predicted as SERCAT 3 or SERCAT 5. This stage is necessary given the substantial proportion of non-participants even within these SERCATs.

Figure 5.5 illustrates the distribution of annual Reserve hours in the financial year following separation (`Unit_Result_Total`), which is characterized by a large mass at zero and pronounced right tail. Approximately 45 percent of the training set and 43 percent of the test set record no Reserve service in the FY following separation. In the absence of Stage Two, the Stage Three regression model is dominated by non-participants, resulting in performance worse than a model predicting mean Reserve hours for all observations. Stage Two therefore functions as a filtering mechanism, removing non-participants and enabling Stage Three to model the intensity of Reserve participation conditional on participation.

Because participation is observed one FY after separation, Stage Two captures the extent to which intentions formed at separation translate into realized behavior as members' post-separation constraints and opportunities evolve.

Stage Two begins by filtering the training data to include only SERCAT 3 and SERCAT 5 observations within FY2016–17 to FY2022–23. Predictions from the Stage One eXtreme Gradient Boosting model are similarly filtered to retain only FY2023–24 SERCAT 3 and SERCAT 5 predictions. This ensures that the test set contains an observed response variable for Reserve service participation in the FY following separation.

As in Stage One, we tune the models' hyperparameters using 100 iterations of scikit-learn's `RandomizedSearchCV` with five-fold cross validation, optimized for AUC (Scikit-learn Developers, 2011). We then evaluate the tuned models on a binary response variable derived from `Unit_Result_Total`, using the FY2023–24 SERCAT 3 and SERCAT 5 prediction set generated by the Stage One classification model.

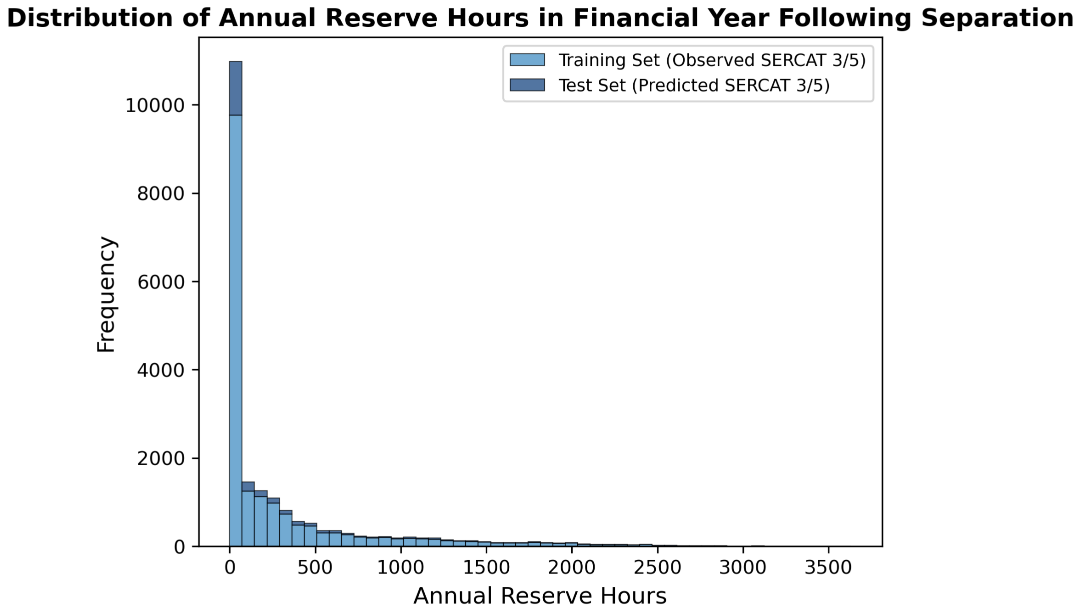

Figure 5.5. This figure shows the large mass at zero and right-skewed distribution of annual Reserve hours in the FY following separation, with approximately 45 percent of the training set and 43 percent of the test set recording zero hours. Given this concentration of non-participation, Stage Two explicitly models participation prior to estimating service intensity in Stage Three. Observations with high annual hours are retained, as they capture meaningful variation in Reserve service intensity, despite the potential for error in recorded data.

5.2.1 Model Comparison and Selection

Table 5.2 presents a comparative summary of model performance for predicting whether a member completes any Reserve service, defined as `Unit_Result_Total` > 0. As in Stage One, performance differences across models are small, with accuracies differing by less than two percentage points. Given the potential asymmetric consequences of misclassifying participants and non-participants, accuracy alone provides a limited summary of model performance, even with relatively balanced class proportions.

While achieving a mid-range accuracy relative to the other models, eXtreme Gradient Boosting attains the highest F1 score and ties Adaptive Boosting for the highest AUC. Given the importance of correctly identifying participating members, the F1 score and

AUC provide more informative measures of performance than accuracy alone. On this basis, we select eXtreme Gradient Boosting as the Stage Two Participation model and carry its predictions into Stage Three.

We conducted threshold sensitivity analysis across high-precision, high-recall, and F1-optimized classification thresholds. In the absence of clearly defined institutional costs associated with false positives and false negatives, the F1-optimized threshold provided the most balanced performance across evaluation metrics.

Table 5.2. For Stage Two, we report test set performance using validation-set F1-optimized thresholds and carry forward eXtreme Gradient Boosting's predictions based on its highest F1 score and tied highest AUC.

Model	Accuracy	AUC	Precision	Recall	F1
Adaptive Boosting	0.664	0.719	0.644	0.902	0.752
Random Forest	0.645	0.701	0.620	0.956	0.752
eXtreme Gradient Boosting	**0.651**	**0.719**	**0.625**	**0.956**	**0.756**

5.2.2 Confusion Matrix

The confusion matrix in Figure 5.6 illustrates the model's strong performance in identifying members who subsequently participate in the Reserve service, correctly classifying approximately 96 percent of true participants. This high recall reflects the model's effectiveness in capturing participation among eligible members.

In contrast, performance is notably weaker for identifying non-participants. The model correctly classifies only around 26 percent of members who do not participate, with the majority instead predicted to participate. This asymmetry indicates a tendency for the model to favor participation predictions, resulting in a high false-positive rate for non-participation. This behavior is consistent with model selection based on F1 score optimization, which balances precision and recall rather than prioritizing either in isolation, and therefore tolerates higher false-positive rates when this improves overall classification performance. Given that the Stage Two predictions are used to condition the subsequent intensity model, this trade-off favors coverage of true participants over precise exclusion of non-participants.

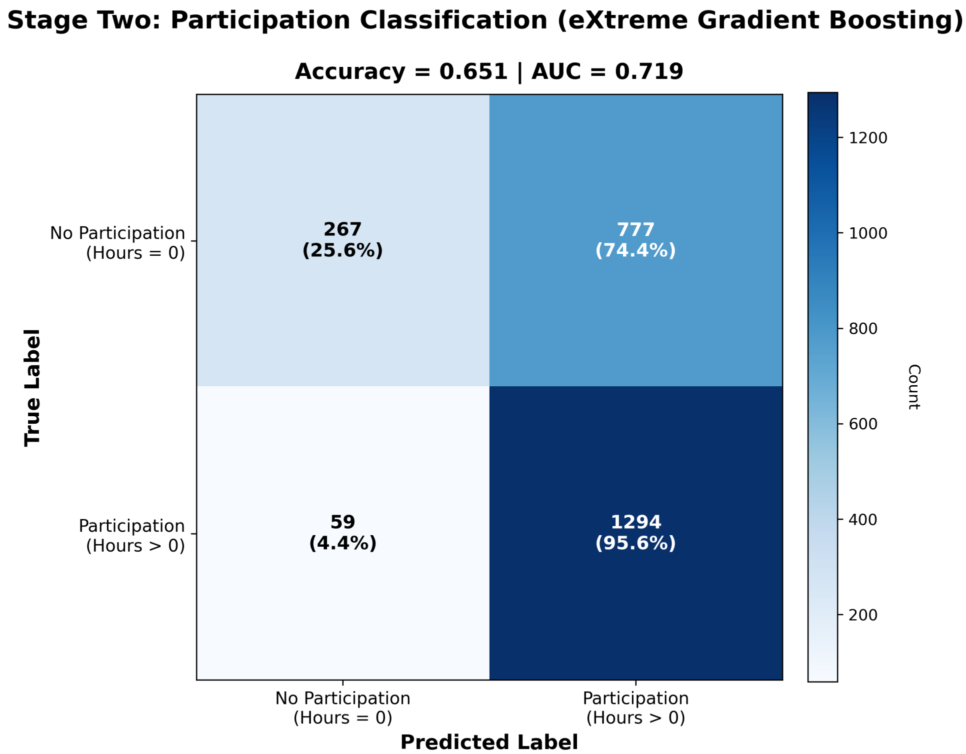

Figure 5.6. This Stage Two confusion matrix illustrates the model's strong performance in correctly identifying Reserve service participants, alongside weaker discrimination of non-participants.

5.2.3 Feature Importance Analysis

We next conduct SHAP analysis using the eXtreme Gradient Boosting Participation model, as illustrated in Figure 5.7. In this binary setting, positive SHAP values indicate that a feature increases the model's predicted likelihood of Reserve participation relative to the model's baseline prediction, while a negative SHAP value reduces it. The most influential feature is institutional in nature, with pre-separation Navy Service predominantly associated with negative SHAP values, indicating a lower predicted likelihood of participation.

In contrast, the remaining influential features are largely individual- and life-stage-related. Achievement signals (number of awards), household circumstances, and remuneration exhibit diverse SHAP effects, suggesting that these factors influence participation in a non-uniform manner across members.

While some influential features in Stage Two overlap with those identified in the SERCAT choice model, others appear to be specific to Reserve participation itself. Characteristics such as Service affiliation and pay influence both the SERCAT decision and subsequent participation. In contrast, several individual-level features exhibit greater importance in the Participation model, suggesting that different factors shape initial choice and observed participation behavior.

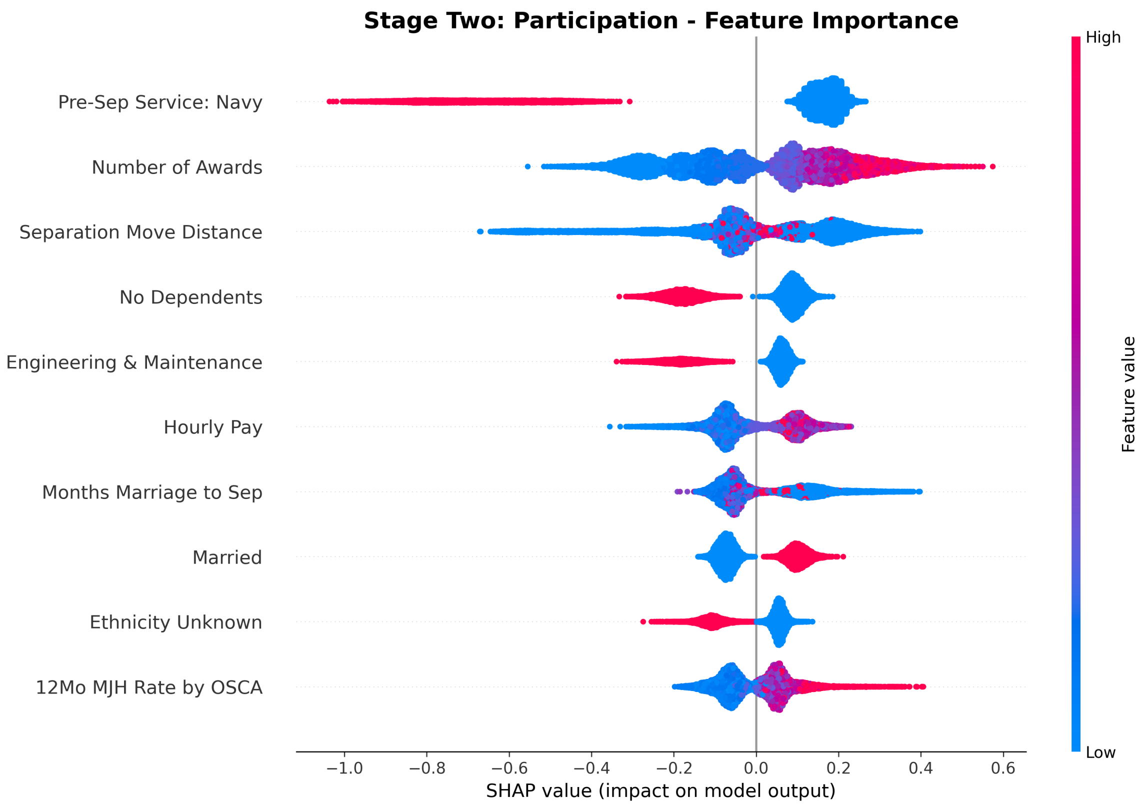

Figure 5.7. This visualization displays the distribution of SHAP values across the most influential features in the Participation model. In this binary classification setting, positive SHAP values increase the predicted likelihood of Reserve participation, while negative values decrease it. Institutional affiliation and life-stage characteristics emerge as the most influential predictors of participation during the first financial year following separation.

5.3 Stage Three: Intensity Regression Model Results

Stage Three applies regression variants of the same model classes to predict annual Reserve service hours, conditional on members being predicted as SERCAT 3 or SERCAT 5 in Stage One and predicted to participate in Stage Two. This stage extends the framework by modeling participation intensity among predicted participants. It begins by filtering the training data to include only SERCAT 3 and SERCAT 5 observations who participated in the Reserve service in their first year after separation within FY2016–17 to FY2022–23.

We tune the models' hyperparameters using 100 iterations of scikit-learn's `RandomizedSearchCV` with five-fold cross validation, optimized for RMSE. We then evaluate the tuned models on the continuous response variable, `Unit_Result_Total`, using Stage One's predicted FY2023–24 SERCAT 3 and SERCAT 5 observations that the Stage Two model classifies as participants.

As shown in Figure 5.5, Stage Three is intentionally trained only on true participants to support learning within the participating population and to avoid distortion from the large mass of non-participants. However, the Stage Three test set necessarily includes some true non-participants due to error propagation from Stage Two. At the point of managerially relevant prediction, these observations are indistinguishable from true participants, as Reserve service outcomes are only observed after the completion of the financial year following separation.

Approximately 38 percent of observations carried forward into Stage Three are false positives, members predicted to participate who ultimately record no Reserve service participation. Including these observations reflects a realistic model deployment scenario, in which downstream models must operate on upstream predictions without access to future outcomes. Excluding them would introduce post-outcome information and overstate Stage Three performance. This design choice reflects realistic operational model deployment conditions rather than a modeling limitation, as downstream models must rely on upstream predictions without access to realized outcomes. This error profile is consistent with the Stage Two model's F1 score optimization, which balances false positives and false negatives rather than minimizing either in isolation.

5.3.1 Model Comparison

Table 5.3 presents a comparative summary of regression model performance for predicting annual Reserve service hours. To contextualize performance, we benchmark results against a naive mean-prediction baseline, which yields a RMSE of 660. Relative to this baseline, all three ML models reduce RMSE to approximately 512, representing a 22.4 percent reduction in prediction error. This improvement indicates that the models capture systematic structure in Reserve service intensity beyond predictions based solely on the sample mean.

Despite this improvement over baseline, overall explanatory power remains modest. The R^2 values, all below 16 percent, indicate that most variation in annual Reserve service hours remains unexplained.

All three models perform similarly, with RMSE, MAE, and R^2 values differing only marginally across models. However, Random Forest marginally outperforms the other models, achieving the lowest RMSE and MAE and the highest R^2. Although these differences are marginal, we select Random Forest for subsequent analysis as it consistently performs best across all evaluation metrics.

Table 5.3. For Stage Three, we report test set performance using hyperparameters optimized via cross-validated RMSE. All models perform comparably with modest results, reflecting significant variance in the data unexplained by the models. The Random Forest results are carried forward for further analysis.

Model	RMSE	MAE	R^2
Adaptive Boosting	513.015	339.993	0.154
Random Forest	**511.966**	**337.999**	**0.157**
eXtreme Gradient Boosting	512.131	337.148	0.157

5.3.2 Regression Metrics

Figure 5.8 presents a predicted versus actual plot for the Stage Three Random Forest model. The majority of points fall below the 1:1 line, indicating systematic underprediction of Reserve service hours, particularly among members with higher observed participation.

Predicted values remain relatively compressed across the range of outcomes, while actual hours vary substantially, contributing to the model's low explanatory power. This compression reflects the model's tendency to minimize average error under RMSE optimization, resulting in regression-to-the-mean behavior that dampens predictions for high-intensity outcomes.

Prediction error increases as actual hours increase, indicating greater uncertainty when modeling high-intensity participation. This pattern is consistent with the highly skewed distribution shown in Figure 5.5, where only approximately 17 percent of observations record more than 500 annual Reserve hours. The limited representation of high-intensity participants constrains the model's ability to learn distinct predictive patterns for this subgroup, leading to systematic underestimation for high-intensity observations.

More broadly, the difficulty in accurately predicting Reserve service intensity likely reflects both individual- and institutional-level uncertainty that emerges after separation. Although the outcome window is limited to the first full FY following separation, the model relies exclusively on characteristics observed at the point of separation. During this period, members' capacity to render service may evolve over the first FY following separation as personal circumstances, employment, and preferences change.

At the institutional level, observed participation is likely shaped by factors not captured in the data, including unit-level demand, funding availability, or tasking opportunities. Together, these dynamics underscore the inherent limitations of forecasting post-separation service intensity from pre-separation characteristics, particularly for high-intensity outcomes and over the first full FY following separation.

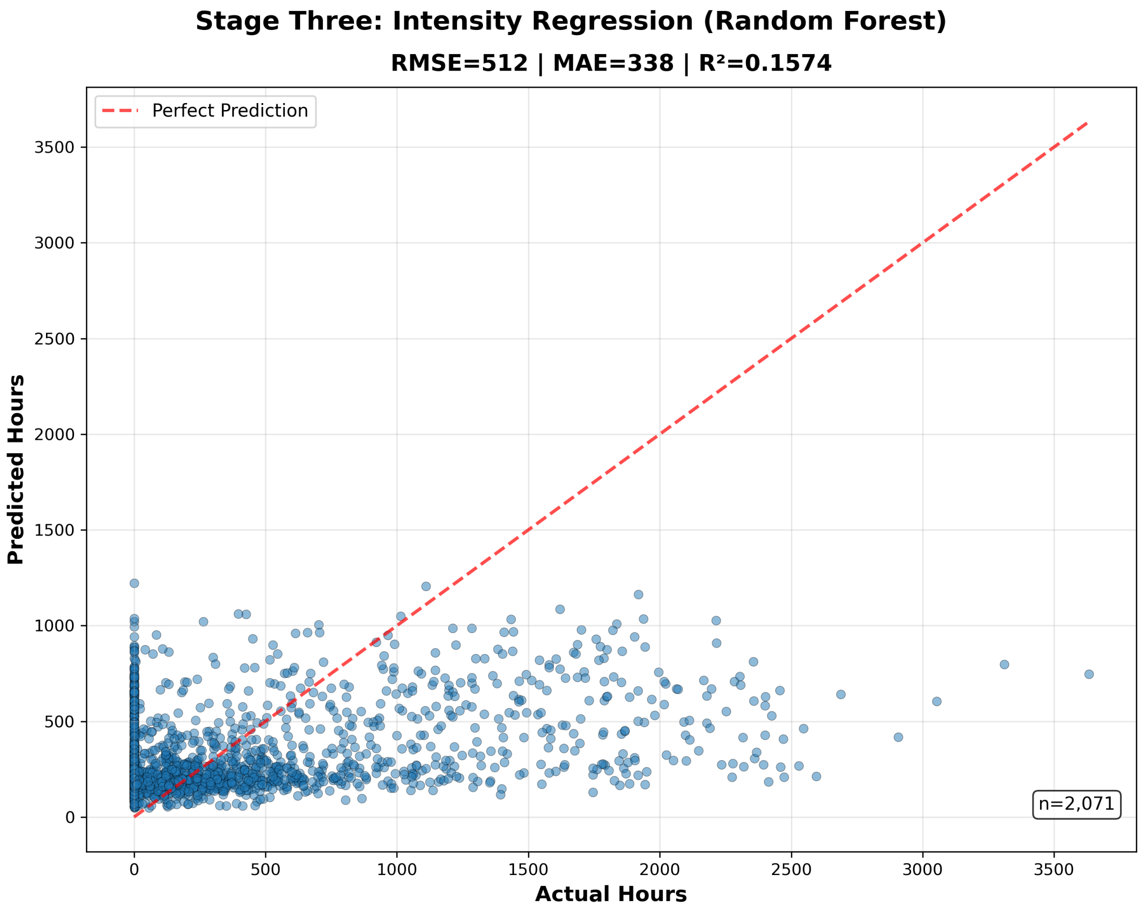

Figure 5.8. This figure compares predicted and observed annual Reserve service hours for the Random Forest intensity model. The model exhibits systematic underprediction, particularly among high-participation members. The limited representation of observations above 500 hours constrains the model's ability to accurately predict high-intensity service.

5.3.3 Feature Importance Analysis

While the model explains a limited share of the overall variance in Reserve service, the SHAP beeswarm plot in Figure 5.9 provides insight into which features the model relies on when generating predictions. The figure shows that life-stage, economic, and other individual-level characteristics dominate the most influential features in the Random Forest regression model, indicating that individual circumstances play a larger role than institutional factors in shaping predicted service intensity among participants.

Age emerges as the most influential feature, with the distribution of SHAP values indicating that higher age is generally associated with higher predicted Reserve service intensity. This pattern suggests that life-stage factors are strongly associated with predicted variation in service rendered once participation occurs. This may be due to older members exhibiting greater stability in civilian employment, having fewer competing career establishment or familial demands, or more predictable availability, all of which can support higher levels of Reserve engagement relative to younger members.

When considered alongside hourly pay and the pay differential features, this result is consistent with systematic differences in how Reserve service fits into members' broader employment and life circumstances across age cohorts. However, the model does not disentangle the relative contributions of age, remuneration, and institutional incentives, and these effects should therefore be interpreted as correlated rather than causal.

In this context, it is useful to note that the ADF has transitioned through three major retirement schemes, with eligibility determined by date of entry into the service: the Defence Force Retirement and Death Benefits scheme, the Military Superannuation and Benefits Scheme, and the current superannuation contributions (Department of Defence, n.d.-a, n.d.-b, n.d.-c). These schemes differ in how post-separation earnings and service affect retirement outcomes, and may shape the perceived value of Reserve service across cohorts. However, this methodology cannot isolate retirement incentives from broader life-stage effects and would require causal approaches beyond the scope of this study.

More broadly, most features exhibit SHAP values that are tightly concentrated around zero, indicating that while these variables influence predictions, their individual effects on predicted Reserve service intensity are small. This pattern highlights the highly individualized nature of Reserve service intensity and underscores the difficulty of predicting post-separation behavior using observed pre-separation characteristics alone.

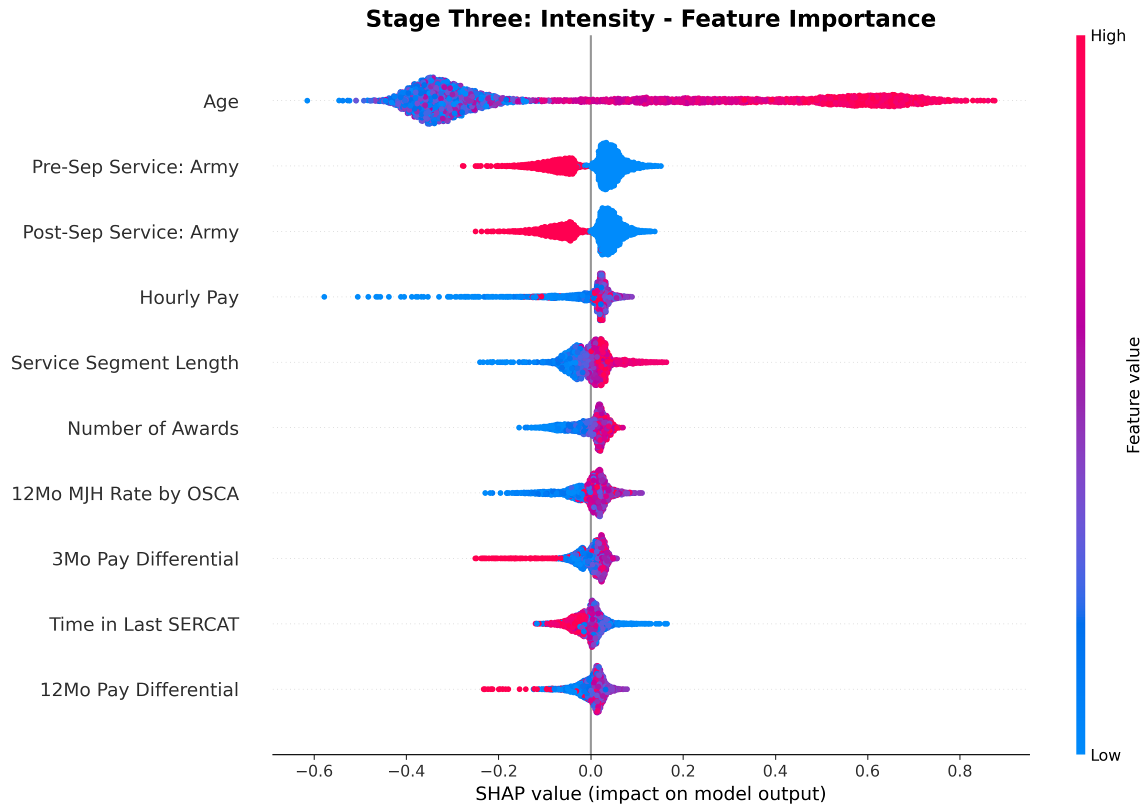

Figure 5.9. This visualization displays the distribution of SHAP values across the most influential features in the intensity regression model. SHAP values represent each feature's contribution to predicted annual Reserve service hours. Age and remuneration-related characteristics emerge as the strongest predictors of service intensity, consistent with variation in engagement across career and life stages.

In conclusion, this chapter presented the results of the three-stage modeling framework. The Stage One SERCAT choice model performed substantially better than random classification, though it exhibited difficulty distinguishing between SERCAT 2 and SERCAT 3 members. The Stage Two Participation model demonstrated more modest performance and carried a significant number of false positives into Stage Three. Consistent with this, the Stage Three Intensity model exhibited limited explanatory power, particularly for members with high levels of Reserve participation.

Across stages, SHAP feature importance highlights a progression in the determinants of Reserve behavior. At separation, SERCAT choice is driven primarily by geographic mobility, institutional structure, and economic position, factors that shape anticipated post-separation opportunities. One year later, participation is more strongly associated with institutional affiliation and life-stage circumstances, consistent with plans adapting during transition. Conditional on participation, service intensity is most strongly predicted by age and remuneration-related characteristics, reflecting variation in engagement across career and life stages. Overall, these model and feature importance results support the broader synthesis and interpretation that follow in the concluding chapter.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 6:
Conclusion

This thesis applies ML methods to administrative personnel data to model SERCAT choice, Reserve service participation, and intensity within the ADF. We conclude our research by analyzing the overall results, outlining practical implications, discussing limitations and constraints, and identifying avenues for future research.

6.1 Analysis of Results

When considered together, the feature importance results and model performance metrics provide insight into Reserve behavior across the separation-to-participation pipeline. The Stage One models demonstrate strong performance when classifying members against the three SERCATs, substantially outperforming random classification. Because these predictions rely on characteristics observed at the point of separation, the influential features capture conditions that are salient at the time the decision is made. Institutional structure, geographic mobility, and economic position are strongly associated with predicted SERCAT choice, reflecting how individuals anticipate post-separation opportunities and how the Reserve service fits within those plans.

These predicted SERCAT outcomes are then carried forward in the Participation models, which assess observed behavior one FY after separation. In contrast to Stage One, these models perform only marginally better than random classification. Threshold selection in Stage Two prioritizes a balance between precision and recall, reflecting the relatively low institutional cost of both false positives and false negatives in this exploratory setting. Using the same pre-separation characteristics to predict post-separation behavior, the feature importance analysis again highlights institutional affiliation and life-stage characteristics. This pattern suggests that while initial intentions may be formed when submitting for separation, realized participation reflects how post-separation plans evolve as members encounter changing constraints, opportunities, and preferences.

Lastly, the Stage Three regression models yield modest explanatory power when predicting annual Reserve hours among eligible participants, with particularly weak performance for high-intensity participation outcomes. These results indicate that characteristics observed at separation are insufficient to explain the wide variation in Reserve service intensity one year later, especially given the magnitude of life-stage and professional change that accompanies transition from full-time service. Although the models explain only a limited share of the total variance, the feature importance results consistently emphasize individual-level characteristics, aligned with the interpretation that members may adjust their level of Reserve engagement as they come to understand how the institution fits into their post-separation career and life-stage circumstances.

These findings suggest several implications for policy and practice. First, the prominence of geographic features in the SERCAT choice model, particularly separation move distance and pre-separation geographic location, highlights the importance of anticipated post-separation location in shaping SERCAT choice. Policies that expand geographically flexible Reserve opportunities, including roles supported through remote work arrangements or available across a wider range of locations, may reduce the extent to which geographic constraints influence SERCAT selection and subsequent engagement.

Second, given the observed overlap between SERCAT 2 and SERCAT 3 characteristics, policy frameworks that allow smoother or quicker transitions between these categories may better support participation as separated members' motivations and circumstances evolve over time.

Third, realized participation appears to diverge from intentions at separation. As a result, policies that engage members during their first FY after separation, such as structured check-ins, targeted offers, or individual-position matching, may be more effective than relying on engagement driven solely by member initiative after separation.

Lastly, Reserve benefits are largely structured in a uniform manner and may not align with members' preferences at different life stages. Targeted qualitative or mixed-methods research could help identify how Reserve policies and benefits might be adapted to better support engagement across diverse career and life stages.

6.2 Practical Implications

Although predictive performance varies across stages, the staged framework provides a structured lens for understanding Reserve behavior over time. In particular, the staged framework demonstrates how administrative data available at separation can be used to identify broad patterns in likely Reserve engagement, even when individual outcomes remain difficult to precisely predict. From an institutional perspective, this approach could support the ADF by functioning as a screening and resource-planning tool during the separation process.

If SERCAT choice and the related follow-on fields were temporarily removed from the initial separation application, the ADF could instead process individual records through the model to generate an indicative profile of likely Reserve behavior. These profiles could guide where additional outreach, information, or incentives may be most effectively applied to target members with lower participation probabilities. While there is inherent uncertainty in using historic data to anticipate future outcomes, this use case could be empirically tested by selectively sampling members predicted to choose SERCAT 2 into enhanced outreach or resourcing groups and evaluating subsequent effects. The model would also require routine retraining on recently separated cohorts to capture evolving institutional conditions and individual-level trends. Importantly, such use would aim to inform the allocation of institutional resources rather than to predict or constrain individual outcomes.

Beyond operational considerations, the results also provide a useful foundation for qualitative inquiry into Reserve behavior. As demonstrated throughout this thesis, the models extract patterns only from observable characteristics and cannot capture unmeasured motivations, constraints, or institutional frictions experienced after separation. The features identified as influential therefore serve as a starting point for qualitative research design, helping to guide interviews toward the mechanisms that likely shape Reserve choice, participation, and intensity but remain unobserved in administrative data.

6.3 Limitations and Model Constraints

We explicitly highlight five limitations and constraints associated with the data and modeling approach. First, the dataset and associated features reflect members' characteristics at the point of separation, which may not align with their post-separation realities. This temporal mismatch likely constrains the predictive ceiling of the models, particularly for outcomes observed one FY after separation.

Second, the models primarily attribute Reserve participation to individual choice. In practice, participation is also shaped by supply-side constraints such as funding availability, tasking demand, geographic mismatch between members and units, and administrative friction. These institutional factors are not directly observed in the data and may substantially limit or prevent participation, independent of individual intent.

Third, Stage Two and Stage Three rely on the Reserve hours variable, which has inherent limitations. A value of zero conflates members who lack interest with those who lack opportunity, experience administrative delays, or are impacted by recording errors. At the upper end of the distribution, members seeking to complete more than 1,200 Reserve hours in a year must be in a role providing a critical capability and require senior-level approvals, raising interpretive questions about the approximately seven percent of observations above this threshold (Department of Defence, n.d.-e).

Fourth, the Stage Two and Stage Three models are sensitive to threshold selection. Although we optimize thresholds using the F1 score to balance precision and recall, alternative thresholding objectives such as minimizing false positives would yield different performance outcomes. As such, the reported metrics and downstream results should be interpreted as conditional on this deliberate modeling choice.

Lastly, the use of ML limits causal interpretation. While SHAP values provide insight into how features contribute to model predictions, they do not identify causal relationships. This limitation is particularly salient in Stage Three, where overall explanatory power is modest and observed feature effects should be interpreted as model reliance rather than causal drivers of Reserve service intensity. Collectively, these limitations indicate that the models are best interpreted as descriptive and exploratory tools rather than definitive predictors of individual behavior.

6.4 Future Work

Several avenues for future work could improve both the prediction performance and interpretability of the modeling framework presented in this thesis. First, the inclusion of additional individual-level performance data may enhance predictions of SERCAT choice. Although we intended to incorporate ordinal performance reporting metrics, these data were not available in a form that could be reliably matched to the de-identified dataset used in this study. Nonetheless, performance reporting is plausibly correlated with SERCAT outcomes, a relationship indirectly suggested by the consistent importance of the number of awards variable across feature importance analyses. This pattern indicates that formal performance evaluations may capture information relevant to Reserve selection that is only imperfectly proxied by observed awards.

Second, future research would benefit from improved data on members' post-separation careers and personal circumstances. Historically, limited information has been available regarding post-separation employment, occupational field, and geographic location following separation from full-time service. Incorporating such data would support the construction of features that better proxy preferences for labor supply, opportunity cost, and lifestyle, which are likely to influence both Reserve participation and intensity.

Finally, the geographic features engineered in this thesis could be further refined by distinguishing between Reserve roles performed remotely and those requiring in-person attendance. At present, approval for remote Reserve work is managed outside the personnel system and is not systematically captured in administrative data. Incorporating this information would improve the measurement of geographic constraints and better reflect the practical feasibility of participation for separating members. In combination, these extensions would not only support improved modeling of Reserve behavior, but also enhance the ADF's ability to conduct personnel planning and contingency analysis.

This thesis demonstrates how machine learning methods can be applied to administrative personnel data to illuminate Reserve choice, participation, and intensity decisions within the ADF. By structuring analysis across the separation-to-participation pipeline, the framework distinguishes between intentions formed at separation, realized engagement in the first post-separation year, and variation in service intensity across career and life stages. While predictive performance varies across stages, the framework clarifies where administrative data provide strong signals and where evolving post-separation circumstances introduce inherent limits to prediction. In doing so, this work contributes empirical insight into post-separation Reserve dynamics and establishes a foundation for future research, policy experimentation, and institutional planning aimed at supporting Government-directed Reserve workforce growth.

APPENDIX A: Data Dictionary

Table A.1. This table provides variable names and descriptions for all variables included in the dataset, including both source variables and engineered features.

Variable	Description
	Decision Dataset
`Actual_Location_Long_Description`	Post-separation assigned base location (214-category nominal variable). *Not included in modeling after derivation of geographic features.*
`Actual_State_Code`	Post-separation unit's geographic state location (10-category nominal variable).
`Actual_Unit_Long_Description`	Post-separation assigned unit (420-category nominal variable). *Not included in modeling to avoid post-separation assignment information that is not available at the point of separation.*
`Age_(years)`	Age of observation at separation (discrete integer).
`Award/Achievement_Number`	Number of recorded awards or achievements (discrete integer).
`CALD_Flag`	Binary indicator for culturally and linguistically diverse (CALD) status: 0 = not CALD, 1 = CALD.
`Category_Short_Description`	Job category at separation (271-category nominal variable). *Not included in modeling after derivation of* `Segment_Long_Description`.
`Disabled_Indicator`	Binary indicator for disability status: 0 = not disabled, 1 = disabled. *Not included in modeling due to small sample size.*

Continued on next page. . .

Variable	Description
	Decision Dataset
`Enlistment_Date`	Date joined the ADF (datetime). *Not included in modeling after derivation of `Months_Between_Enlistment_Separation`.*
`Ethnic_Group_Short_Description`	Self-declared ethnicity (7-category nominal variable).
`Family_Long_Description`	Job family at separation (80-category nominal variable). *Not included in modeling after derivation of `Segment_Long_Description`.*
`Gender_Code`	Binary indicator for gender: 0 = male, 1 = female.
`Indigenous_Status_Flag`	Binary indicator for Indigenous status. *Dropped due to coverage from `Ethnic_Group_Short_Description`.*
`Job_Family_Change`	Binary indicator for a change between `Prior_Family_Long_Description` and `Family_Long_Description` at separation: 0 = no change, 1 = change.
`Language_Number`	Number of registered languages spoken (discrete integer).
`Length_of_Service_(years)`	Length of service between enlistment and separation in whole years (discrete integer).
`Length_of_Service_Segment_(days)`	Total number of days of service (discrete integer).
`Marital_Status_Date`	Date of marital status change (datetime). *Not included in modeling after derivation of `Months_Between_Marriage_Separation`.*
`Marital_Status_Short_Description`	Partnership status (5-category nominal variable).
`Member_with_Dependents_Short_Description`	Family status at separation (3-category nominal variable).
`Months_Between_Enlistment_Separation`	Months between enlistment date and separation date (discrete integer).
`Months_Between_Marriage_Separation`	Months between marital status date and separation date (discrete integer).

Continued on next page. . .

Variable	Description
	Decision Dataset
Months_Between_Promotion_Separation	Months between last promotion and separation (discrete integer).
Party_ID	Unique identifier for each observation (discrete integer).
Prior_Actual_Location_Long_Description	Pre-separation base location (207-category nominal variable). *Not included in modeling after derivation of geographic features.*
Prior_Actual_State_Code	Pre-separation unit's geographic state location (10-category nominal variable).
Prior_Actual_Unit_Long_Description	Pre-separation unit (674-category nominal variable). *Not included in modeling due to number of categories.*
Prior_Category_Short_Description	Pre-separation job category (271-category nominal variable). *Not included in modeling after derivation of `Segment_Long_Description`.*
Prior_Family_Long_Description	Pre-separation job family (80-category nominal variable). *Not included in modeling after derivation of `Job_Family_Change`.*
Prior_Loc_Missing	Binary indicator for observations with missing `Prior_Actual_Location_Long_Description`: 0 = not missing, 1 = missing.
Prior_Rank_Code	Substantive rank preceding most recent promotion (22-category nominal variable).
Prior_Rank_Officer_Indicator	Binary indicator for officer status of substantive rank preceding most recent promotion: 0 = non-officer, 1 = officer. *Not included in modeling due to similarity with `Prior_Worn_Rank_Officer_Indicator`.*
Prior_Service_Short_Description	Pre-separation Service affiliation (3-category nominal variable).
Prior_Worn_Rank_Code	Worn rank preceding most recent promotion (22-category nominal variable).
Prior_Worn_Rank_Officer_Indicator	Binary indicator for officer status of worn rank preceding most recent promotion: 0 = non-officer, 1 = officer.

Continued on next page. . .

Variable	Description
	Decision Dataset
`Promotion_Date`	Date of last promotion (datetime). *Not included in modeling after derivation of* `Months_Between_Promotion_Separation`.
`Rank_Code`	Substantive rank attained at most recent promotion (22-category nominal variable).
`Rank_Officer_Indicator`	Binary indicator for officer status of substantive rank attained at most recent promotion: 0 = non-officer, 1 = officer.
`Religious`	Binary indicator for declared religious affiliation: 0 = not religious, 1 = religious.
`Religion_Short_Description`	Single declaration of religious affiliation (59-category nominal variable). *Not included in modeling after derivation of* `Religious`.
`Segment_Long_Description`	Broad strategic workforce segments grouping `Family_Long_Description` (8-category nominal variable).
`Separation_Action_Reason_Grouping_Short_Description`	Voluntary or involuntary separation label (2-category nominal variable).
`Separation_Action_Reason_Long_Description`	Detailed reason for separation (3-category nominal variable).
`Separation_Date`	Date separated from full-time service (datetime). *Not included in modeling after derivation of* `Months_Between_Marriage_Separation` *and* `Months_Between_Promotion_Separation`.
`Service_Short_Description`	Service affiliation post-separation (3-category nominal variable).
`SERCAT_Code`	Service category post-separation (3-category nominal variable).
`Time_in_Last_Rank_(days)`	Days spent in previous rank (discrete integer).

Continued on next page. . .

Variable	Description
Decision Dataset	
`Time_in_Last_SERCAT_(days)`	Number of days in previous service category (discrete integer).
`Time_in_Rank_(years)`	Years spent in most recent rank at separation (discrete integer).
`Time_to_Rank_(days)`	Number of days from attainment of preceding rank to attainment of most recent rank (discrete integer).
`Worn_Rank_Code`	Worn rank at separation (22-category nominal variable).
`Worn_Rank_Officer_Indicator`	Binary indicator for officer status of worn rank for most recent promotion: 0 = non-officer, 1 = officer.
Education Dataset	
`Count_Degrees`	Number of degrees recorded (discrete integer).
`Highest_Education_Level_Long_Description`	Highest educational attainment on record (8-category nominal variable). *Not included in modeling after derivation of `Max_Degree_Value`.*
`Max_Degree_Value`	Highest educational achievement level (7-category ordinal variable).
Operations Dataset	
`NWL_DEPL_COUNT`	Number of non-warlike deployments (discrete integer).
`NWL_TOTAL_TIME`	Total days on non-warlike deployments (discrete integer).
`NWL_WSA_DAYS`	Total days in non-warlike service areas (discrete integer).
`Peace_DEPL_COUNT`	Number of peace deployments (discrete integer).
`Peace_TOTAL_TIME`	Total days on peace deployments (discrete integer).
`Peace_WSA_DAYS`	Total days in peace service areas (discrete integer).
`WARLIKE_DEPL_COUNT`	Number of warlike deployments (discrete integer).
`WARLIKE_TOTAL_TIME`	Total days on warlike deployments (discrete integer).
`WARLIKE_WSA_DAYS`	Total days in warlike service areas (discrete integer).

Continued on next page. . .

Variable	Description
	Pay Dataset
Hourly_Pay	Hourly wage from final full-time payslip at separation (continuous).
	Reserve Attendance Dataset
Reserve_Attendance_Date_Value	Number of Reserve days attended in the first FY post-separation (discrete integer).
Unit_Result_Total	Number of Reserve hours completed in the first FY post-separation (discrete integer).
	Engineered Features
Actual_Lat	Latitude of the member's post-separation unit location (continuous). *Not included in modeling after derivation of Separation_Removal_Distance.*
Actual_Lon	Longitude of the member's post-separation unit location (continuous). *Not included in modeling after derivation of Separation_Removal_Distance.*
GCC_NAME21	Greater Capital City Statistical Area associated with the member's pre-separation unit location (categorical).
OSCA	Occupational Skill Classification Area major group representing the member's broad civilian occupational skill domain, mapped from their military role (categorical).
Prior_Lat	Latitude of the member's pre-separation unit location (continuous).
Prior_Lon	Longitude of the member's pre-separation unit location (continuous).
SA4_CODE21	Statistical Area Level 4 regional classification associated with the member's pre-separation unit location (categorical).
Separation_Removal_Distance	Haversine distance between pre- and post-separation unit locations, measuring geographic relocation at separation (continuous).

Continued on next page. . .

Variable	Description
	Engineered Features
Three_Month_PreSep_Average_Hours_Worked	Average weekly hours worked in the pre-separation unit local labor market at three months prior to separation (continuous).
Three_Month_PreSep_Education_MJH_Rate	Multiple job holding rate among individuals with the same education level as the member in the labor market three months prior to separation (continuous).
Three_Month_PreSep_GCC_Median_Pay	Median full-time pay within the pre-separation unit's GCCSA at three months prior to separation (continuous).
Three_Month_PreSep_GCC_Median_Pay_Differential	Difference between the member's observed pay at separation and the median pay for individuals of the same gender within the pre-separation unit's GCCSA, measured three months prior to separation (continuous).
Three_Month_PreSep_GCC_MJH_Rate	Multiple job holding rate within the pre-separation unit's GCCSA at three months prior to separation (continuous).
Three_Month_PreSep_Multiple_Job_Holding_Rate	National multiple job holding rate at three months prior to separation (continuous).
Three_Month_PreSep_OSCA_MJH_Rate	Multiple job holding rate among individuals within the same gender and OSCA as the member in the labor market three months prior to separation (continuous).
Three_Months_PreSep_Employment_to_Population_Ratio	Employment-to-population ratio among individuals of the same gender as the member in the pre-separation unit's SA4 labor market at three months prior to separation (continuous).
Three_Months_PreSep_Participation_Rate	Labor force participation rate among individuals of the same gender as the member in the pre-separation unit's SA4 labor market three months prior to separation (continuous).
Three_Months_PreSep_Unemployment_Rate	Unemployment rate among individuals of the same gender as the member in the pre-separation unit's SA4 labor market at three months prior to separation (continuous).

Continued on next page. . .

Variable	Description
Engineered Features	
Twelve_Month_PreSep_Average_Hours_Worked	Average weekly hours worked in the pre-separation unit local labor market at twelve months prior to separation (continuous).
Twelve_Month_PreSep_Education_MJH_Rate	Multiple job holding rate among individuals with the same education level as the member in the labor market twelve months prior to separation (continuous).
Twelve_Month_PreSep_GCC_Median_Pay	Median full-time pay within the pre-separation unit's GCCSA at twelve months prior to separation (continuous).
Twelve_Month_PreSep_GCC_Median_Pay_Differential	Difference between the member's observed pay at separation and the median pay for individuals of the same gender within the pre-separation unit's GCCSA, measured twelve months prior to separation (continuous).
Twelve_Month_PreSep_GCC_MJH_Rate	Multiple job holding rate within the pre-separation unit's GCCSA at twelve months prior to separation (continuous).
Twelve_Month_PreSep_Multiple_Job_Holding_Rate	National multiple job holding rate at twelve months prior to separation (continuous).
Twelve_Month_PreSep_OSCA_MJH_Rate	Multiple job holding rate among individuals within the same gender and OSCA as the member in the labor market twelve months prior to separation (continuous).
Twelve_Months_PreSep_Employment_to_Population_Ratio	Employment-to-population ratio among individuals of the same gender as the member in the pre-separation unit's SA4 labor market at twelve months prior to separation (continuous).
Twelve_Months_PreSep_Participation_Rate	Labor force participation rate among individuals of the same gender as the member in the pre-separation unit's SA4 labor market twelve months prior to separation (continuous).
Twelve_Months_PreSep_Unemployment_Rate	Unemployment rate among individuals of the same gender as the member in the pre-separation unit's SA4 labor market at twelve months prior to separation (continuous).

APPENDIX B: Mapping of Figure Labels to Data Dictionary Variables

Table B.1. This table maps figure labels used in the visualizations to the variable names within the Data Dictionary at Table A.1.

Figure label	**Variable**
3Mo MJH Rate	`Three_Month_PreSep_Multiple_Job_Holding_Rate`
3Mo Pay Differential	`Three_Month_PreSep_GCC_Median_Pay_Differential`
12Mo MJH Rate by OSCA	`Twelve_Month_PreSep_OSCA_MJH_Rate`
12Mo Pay Differential	`Twelve_Month_PreSep_GCC_Median_Pay_Differential`
Age	`Age_(years)`
Annual Reserve Hours	`Unit_Result_Total`
Engineering & Maintenance	`Segment_Long_Description_Engineering, Maintenance and Construction`
Ethnicity Unknown	`Ethnic_Group_Short_Description_NA`
Hourly Pay	`Hourly_Pay`
Married	`Marital_Status_Short_Description_Married`
Months Enlistment to Sep	`Months_Between_Enlistment_Separation`

Continued on next page. . .

Figure label	**Variable**
Months Marriage to Sep	`Months_Between_Marriage_Separation`
No Dependents	`Member_with_Dependents_Short_Description_MBR`
Number of Awards	`Award/Achievement_Number`
OSCA Labourers	`OSCA_8.0`
Post-Sep Service: Army	`Service_Short_Description_Army`
Post-Sep Service: Navy	`Service_Short_Description_Navy`
Pre-Sep Latitude	`Prior_Lat`
Pre-Sep Longitude	`Prior_Lon`
Pre-Sep Service: Army	`Prior_Service_Short_Description_Army`
Pre-Sep Service: Navy	`Prior_Service_Short_Description_Navy`
Separation Move Distance	`Separation_Removal_Distance`
Service Segment Length	`Length_of_Service_Segment_(days)`
Time in Last Rank	`Time_in_Last_Rank_(days)`
Time in Last SERCAT	`Time_in_Last_SERCAT_(days)`
Time to Current Rank	`Time_to_Rank_(days)`

APPENDIX C: Stage One Expanded Result Figures

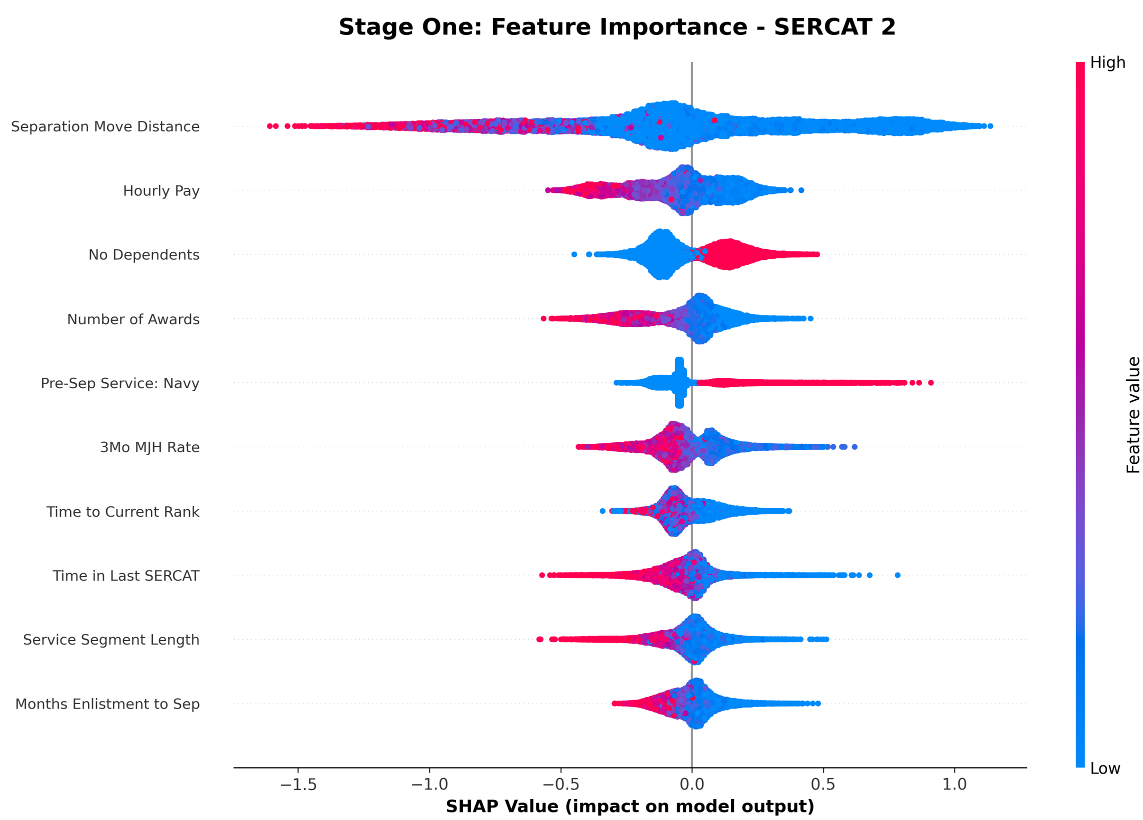

Figure C.1. This class-specific SHAP beeswarm plot displays the most influential features for the Stage One SERCAT choice model when predicting membership in SERCAT 2. Positive values increase the model's predicted likelihood of SERCAT 2, while negative values decrease it. Features are ordered by their mean absolute SHAP values. This figure complements the Stage One feature importance results presented in Chapter 5.

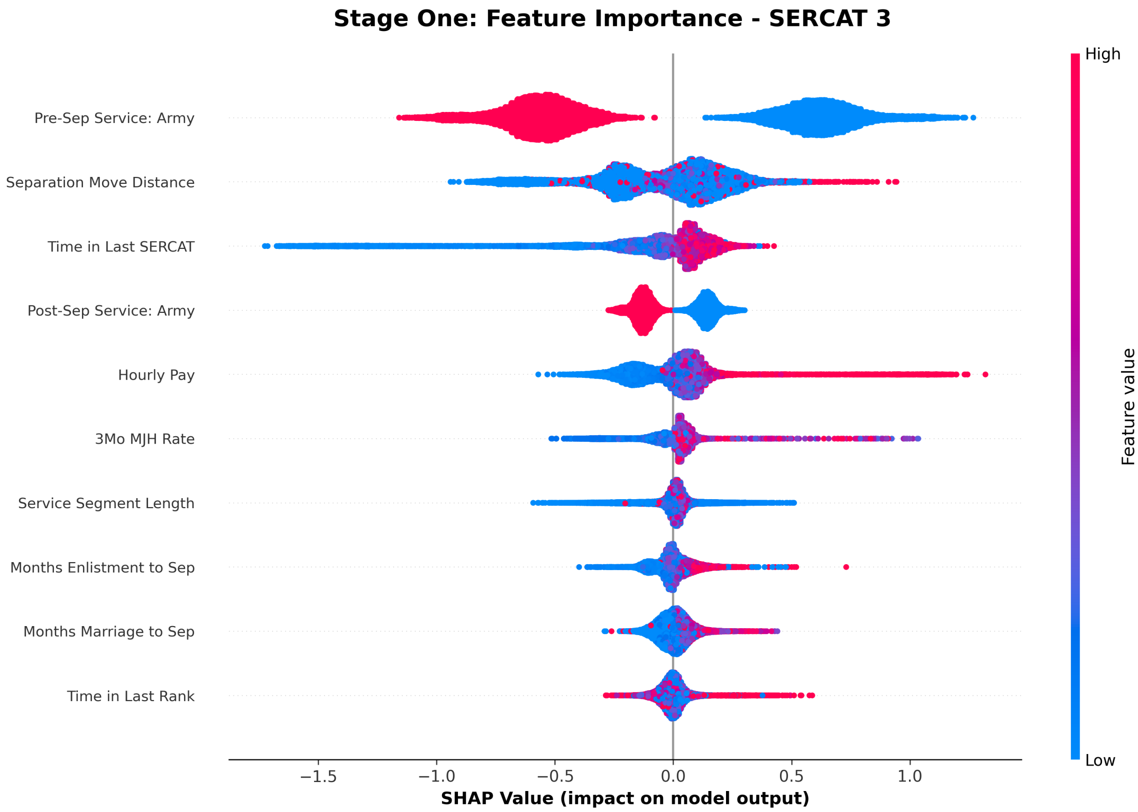

Figure C.2. This class-specific SHAP beeswarm plot displays the most influential features for the Stage One SERCAT choice model when predicting membership in SERCAT 3. Positive values increase the model's predicted likelihood of SERCAT 3, while negative values decrease it. Features are ordered by their mean absolute SHAP values. This figure complements the Stage One feature importance results presented in Chapter 5.

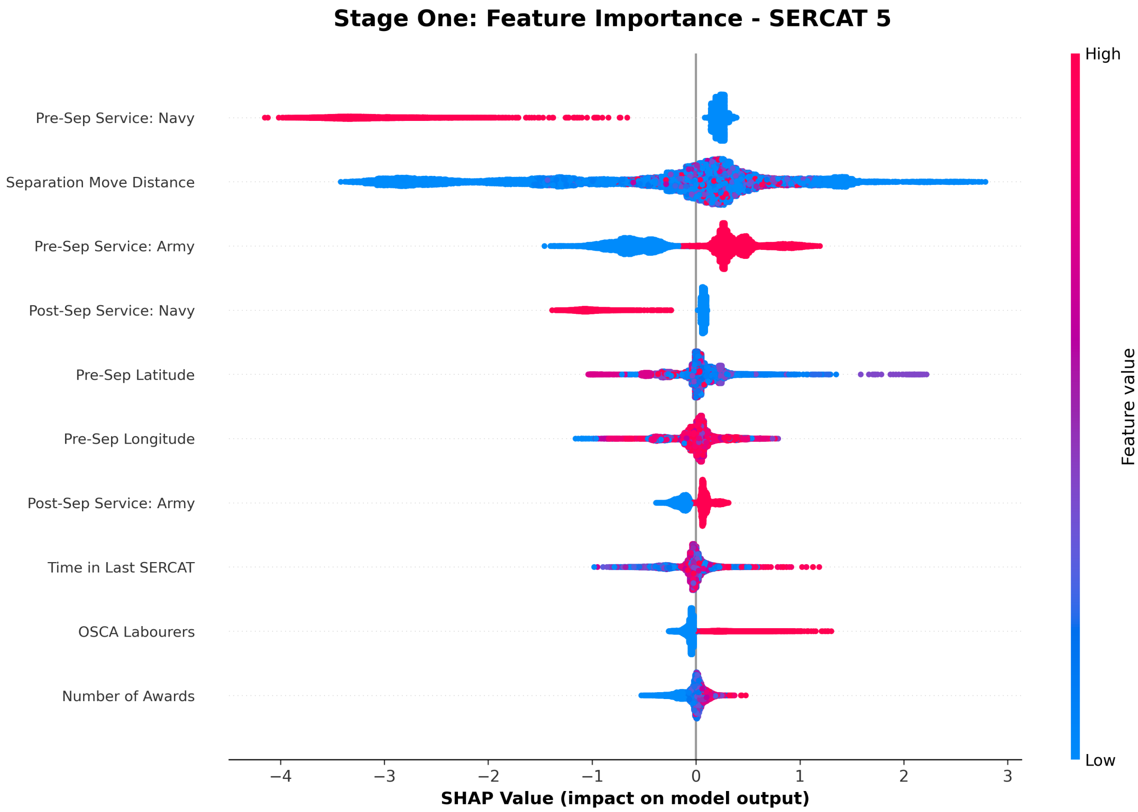

Figure C.3. This class-specific SHAP beeswarm plot displays the most influential features for the Stage One SERCAT choice model when predicting membership in SERCAT 5. Positive values increase the model's predicted likelihood of SERCAT 5, while negative values decrease it. Features are ordered by their mean absolute SHAP values. This figure complements the Stage One feature importance results presented in Chapter 5.

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

ADF Careers. (n.d.). *ADF Reserves*. Retrieved February 10, 2026, from https://www.adfcareers.gov.au/careers/reserves

Anthropic. (2026). *Claude* [Large language model]. https://claude.ai

Arkes, J., & Kilburn, M. R. (2005). *Modeling Reserve recruiting: Estimates of enlistments* (Report No. MG-202-OSD). RAND. https://www.rand.org/pubs/monographs/MG202.html

Armstrong, M. (2020). Pursuing the total force: Strategic guidance for the Australian Defence Force Reserves in Defence White Papers since 1976. *The Institute for Regional Security*, *16*(2), 71–87. https://regionalsecurity.org.au/security_challenge/strategic-guidance-for-the-australian-defence-force-reserves-in-defence-white-papers-since-1976/

Australian Bureau of Statistics. (2021a, July 20). Australian Statistical Geography Standard (ASGS) Edition 3 digital boundary files (Reference period: July 2021 – June 2026). https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/access-and-downloads/digital-boundary-files

Australian Bureau of Statistics. (2021b, July 20). *Main structure and greater capital city statistical areas* (Reference period: July 2021 – June 2026). https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/main-structure-and-greater-capital-city-statistical-areas

Australian Bureau of Statistics. (2024a, June 12). *History of occupation classifications used in Australia* (Reference period: 2024, Version 1.0). https://www.abs.gov.au/statistics/classifications/osca-occupation-standard-classification-australia/2024-version-1-0/history-occupation-classifications-used-australia

Australian Bureau of Statistics. (2024b, June 12). *Occupation 149231 Commissioned Defence Force Officer* (Reference period: 2024, Version 1.0). https://www.abs.gov.au/statistics/classifications/osca-occupation-standard-classification-australia/2024-version-1-0/browse-classification/1/14/149/1492/149231

Australian Bureau of Statistics. (2024c, June 12). *Occupation 451131 Defence Force Member – Other Ranks* (Reference period: 2024, Version 1.0). https://www.abs.gov.au/

statistics/classifications/osca-occupation-standard-classification-australia/2024-version-1-0/browse-classification/4/45/451/4511/451131

Australian Bureau of Statistics. (2024d, June 12). *OSCA – Occupation Standard Classification for Australia* (Reference period: 2024, Version 1.0). https://www.abs.gov.au/statistics/classifications/osca-occupation-standard-classification-australia/latest-release

Australian Bureau of Statistics. (2024e, September 12). *Employee earnings* (Report No. 6337.0). https://www.abs.gov.au/statistics/labour/earnings-and-working-conditions/employee-earnings/aug-2024

Australian Bureau of Statistics. (2025a, September 5). *Multiple job-holders* (Report No. 6217.0). https://www.abs.gov.au/statistics/labour/jobs/multiple-job-holders/jun-2025

Australian Bureau of Statistics. (2025b, September 25). *Labour Force, Australia, Detailed* (Report No. 6291.0.55.001). https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/aug-2025

Australian Qualifications Framework Council. (2013). *Australian qualifications framework (2nd ed.)* https://www.aqf.edu.au/publication/aqf-second-edition

Born, A. A. (2024). *An automated machine learning approach for more efficient Marine Corps recruiter prospecting* [Master's thesis, Naval Postgraduate School]. https://calhoun.nps.edu/server/api/core/bitstreams/31459a35-7d4e-4c7a-835c-c51d976a4582/content

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

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

Defence (Personnel) Amendment Regulations 2002 (No. 1) (Cth) (2002). https://www.austlii.edu.au/au/legis/cth/num_reg_es/dar200212002n279408.html

Department of Defence. (n.d.-a). *Choice superannuation funds*. ADF Pay and Conditions. Retrieved February 10, 2026, from https://pay-conditions.defence.gov.au/benefits/superannuation/choice-superannuation-funds

Department of Defence. (n.d.-b). *Defence Force Retirement and Death Benefits (DFRDB) scheme*. ADF Pay and Conditions. Retrieved February 10, 2026, from https://pay-conditions.defence.gov.au/benefits/superannuation/defence-force-retirement-and-death-benefits-dfrdb-scheme

Department of Defence. (n.d.-c). *Military Superannuation and Benefits Scheme (MSBS)*. ADF Pay and Conditions. Retrieved February 10, 2026, from https://pay-conditions.defence.gov.au/benefits/superannuation/military-superannuation-and-benefits-scheme-msbs

Department of Defence. (n.d.-d). *Pay and financial benefits in the Reserves*. ADF Pay and Conditions. Retrieved February 10, 2026, from https://www.defence.gov.au/adf-members-families/benefits-conditions-service/pay-financial-benefits/reserves

Department of Defence. (n.d.-e). *Reserve service day allocation*. Defence People Group. https://dpeintranet-dpg.defence.gov.au/careers-conditions/pay-benefits/reserve-service-day-allocation

Department of Defence. (n.d.-f). *Reserves*. ADF Pay and Conditions. Retrieved February 10, 2026, from https://pay-conditions.defence.gov.au/benefits/pay/how-your-pay-worked-out/reserves

Department of Defence. (2017, October 12). *Defence annual report 2016—17*. https://www.defence.gov.au/sites/default/files/2021-10/AR-2016-17.pdf

Department of Defence. (2024, December 18). *Strategic review of the Australian Defence Force Reserves*. https://www.defence.gov.au/sites/default/files/2024-12/Strategic-Review-of-the-Australian-Defence-Force-Reserves.pdf

Department of Defence. (2025a). *ADF Total Workforce System*. ADF Pay and Conditions. https://pay-conditions.defence.gov.au/adf-total-workforce-system#service-category-sercat

Department of Defence. (2025b, June 6). *ADF application to transfer within or separate from the ADF* [Internal form]. https://formsportal.dpe.protected.mil.au/content/forms-portal/content/smart-form.html?AC853

Department of Defence. (n.d.). Current category and occupation mapping to Defence strategic workforce segments [Internal document]. dpeintranet-dpg.defence.gov.au/strategy-policy/workforce-strategy/workforce-analytics/defence-strategic-workforce-segments

Falk, A. (2023). *Who leaves? Individual-based predictive modeling of non-end of active service attrition for enlisted Marines* [Master's thesis, Naval Postgraduate School].

https://calhoun.nps.edu/server/api/core/bitstreams/78fe7652-dd72-4092-8209-565357c73878/content

Freiesleben, T., & Molnar, C. (2024). *Supervised machine learning for science: How to stop worrying and love your black box*. https://ml-science-book.com/supervised-ml.html

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. https://doi.org/10.1006/jcss.1997.1504

Geocodify. (n.d.). *Geocodify API* (Version 2.0.0.1) [Application programming interface]. https://geocodify.com

GeoPandas Developers. (2025). *Geopandas* (Version 1.1.1) [Computer software]. https://geopandas.org/en/v1.1.1/index.html

Hassin, K. (2019). *Analysis of the effect of conduct waivers on U.S. Army enlisted soldier first-term attrition* [Master's thesis, Naval Postgraduate School]. https://calhoun.nps.edu/server/api/core/bitstreams/811da299-6414-47f2-915a-cc3104a2b81c/content

Income Tax Assessment Act 1997 (1997). https://www5.austlii.edu.au/au/legis/cth/consol_act/itaa1997240/s995.1.html#definition

Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *15*(4), 531–538. https://doi.org/10.1002/sam.11583

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer New York. https://doi.org/10.1007/978-1-4614-6849-3

Lording, J. (2015). Paid volunteers: Experiencing Reserve service and resignation. *Australian Army Research Centre*, *12*(1). https://researchcentre.army.gov.au/library/australian-army-journal-aaj/volume-12-number-1-winter/paid-volunteers-experiencing-reserve-service-and-resignation

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Lundberg, S. M., & Lee, S.-I. (2025). *SHAP: SHapley Additive exPlanations* (Version 0.47.2) [Computer software]. https://shap.readthedocs.io/en/latest/

Manafi Varkiani, S., Pattarin, F., Fabbri, T., & Fantoni, G. (2025). Predicting employee attrition and explaining its determinants. *Expert Systems with Applications*, *272*, Article 126575. https://doi.org/10.1016/j.eswa.2025.126575

Manning, C., Raghavan, P., & Schuetze, H. (2008). *Introduction to information retrieval*. Cambridge University Press. https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, *8*(3), 229–254. https://doi.org/10.1007/BF00133443

Molnar, C. (2025). *Interpretable machine learning* (3rd ed.). Leanpub. https://christophm.github.io/interpretable-ml-book/

Nosratabadi, S., Khayer Zahed, R., Ponkratov, V. V., & Kostyrin, E. V. (2022). Artificial intelligence models and employee lifecycle management: A systematic literature review. *Organizacija*, *55*(3), 181–198. https://doi.org/10.2478/orga-2022-0012

OpenAI. (2026). *ChatGPT* [Large language model]. https://chatgpt.com

Pandas Development Team. (2026). *Pandas* (Version 2.3.3) [Computer software]. https://pandas.pydata.org/pandas-docs/version/2.3/index.html#

Scikit-learn Developers. (2011). *Scikit-learn* (Version 1.7.2) [Computer software]. https://scikit-learn.org

Shishko, R., & Rostker, B. (1976). The economics of multiple job holding. *The American Economic Review*, *66*(3), 298–308. https://www.jstor.org/stable/1828164

Thornton, J. L. (2023). *Analysis of Marine Corps recruits in the delayed entry program* [Master's thesis, Naval Postgraduate School]. https://hdl.handle.net/10945/72402

Wenger, J. W., Orvis, B. R., Stebbins, D., Apaydin, E., & Syme, J. (2016). *Strengthening prior service–civil life gains and continuum of service accessions into the Army's Reserve components* (Report No. RR-1376-A). RAND. https://www.rand.org/pubs/research_reports/RR1376.html

West, B. (2018). The future of Reserves: In search of a social research agenda for implementing the Total Workforce Model. *Australian Army Research Centre*, *14*(1). https://researchcentre.army.gov.au/library/australian-army-journal-aaj/volume-14-number-1-autumn/future-reserves-search-social-research-agenda-implementing-total-workforce-model

West, B., & Healy, J. (2025). *Drawing on Reserves* (Occasional Paper No. 28). Australian Army Research Centre. https://doi.org/10.61451/267514

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*, 79–82. https://doi.org/10.3354/cr030079

XGBoost Developers. (2011). *XGBoost* (Version 3.0.5) [Computer software]. https://xgboost.readthedocs.io/en/stable/

Acquisition Research Program
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net