

ESCOLA DE GUERRA NAVAL

CC (T) ELAINE DE CARVALHO PONTES MARTINS

**A INTEGRAÇÃO ENTRE MINERAÇÃO DE DADOS E GESTÃO DO
CONHECIMENTO COMO SUPORTE À TOMADA DE DECISÃO NA
GESTÃO DE PESSOAL DA MARINHA DO BRASIL**

Rio de Janeiro

2025

CC (T) ELAINE DE CARVALHO PONTES MARTINS

**A INTEGRAÇÃO ENTRE MINERAÇÃO DE DADOS E GESTÃO DO
CONHECIMENTO COMO SUPORTE À TOMADA DE DECISÃO NA
GESTÃO DE PESSOAL DA MARINHA DO BRASIL**

Monografia apresentada à Escola de
Guerra Naval, como requisito parcial
para a conclusão do Curso Superior.

Orientador: CMG(RM1-T) Chiara Leão
A. de França Delgado de Freitas

Rio de Janeiro
Escola de Guerra Naval
2025

DECLARAÇÃO DA NÃO EXISTÊNCIA DE APROPRIAÇÃO INTELECTUAL IRREGULAR

Declaro que este trabalho acadêmico: a) corresponde ao resultado de investigação por mim desenvolvida, enquanto discente da Escola de Guerra Naval (EGN); b) é um trabalho original, ou seja, que não foi por mim anteriormente utilizado para fins acadêmicos ou quaisquer outros; c) é inédito, isto é, não foi ainda objeto de publicação; e d) é de minha integral e exclusiva autoria.

Declaro também que tenho ciência de que a utilização de ideias ou palavras de autoria de outrem, sem a devida identificação da fonte, e o uso de recursos de inteligência artificial no processo de escrita constituem grave falta ética, moral, legal e disciplinar. Ademais, assumo o compromisso de que este trabalho possa, a qualquer tempo, ser analisado para verificação de sua originalidade e ineditismo, por meio de ferramentas de detecção de similaridades ou por profissionais qualificados.

Os direitos morais e patrimoniais deste trabalho acadêmico, nos termos da Lei 9.610/1998, pertencem ao seu Autor, sendo vedado o uso comercial sem prévia autorização. É permitida a transcrição parcial de textos do trabalho, ou mencioná-los, para comentários e citações, desde que seja feita a referência bibliográfica completa.

Os conceitos e ideias expressas neste trabalho acadêmico são de responsabilidade do Autor e não retratam qualquer orientação institucional da EGN ou da Marinha do Brasil.

Assinatura digital gov.br

DEDICATÓRIA

Dedico este trabalho ao meu amado esposo, Jean, e ao meu filho, Davi, que são minha inspiração diária e minha maior motivação, por estarem ao meu lado com amor, paciência e apoio incondicional.

Em memória dos meus queridos pais, Edilça e Hamilton, que, embora ausentes fisicamente, permanecem vivos em meu coração e em cada conquista, guiando-me com os valores, o amor e os ensinamentos que deixaram como herança eterna.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por Sua presença constante em minha vida, por me conceder força, sabedoria e serenidade para seguir firme, mesmo diante dos desafios, iluminando meu caminho em cada decisão.

Aos meus queridos pais, Edilça e Hamilton (in memoriam), por todo o amor, dedicação e pelos valores que me transmitiram e que permanecem como guia em minha vida.

Ao meu amado esposo, Jean, e ao meu filho, Davi, por compreenderem minhas ausências, por seu amor incondicional e por serem minha maior motivação e inspiração para seguir adiante.

Aos amigos da Diretoria do Pessoal da Marinha, pelo apoio, incentivo e companheirismo durante este percurso.

À equipe e aos instrutores do Curso Superior 2025 da Escola de Guerra Naval, pelo apoio e pela dedicação demonstrados em todas as etapas do processo.

Por fim, à minha orientadora, CMG (RM1-T) Chiara Leão A. de França Delgado de Freitas, expressei meu reconhecimento e gratidão pela paciência, dedicação e pelas orientações sempre precisas. Sua confiança e incentivo foram essenciais para a conclusão desta pesquisa.

“O conhecimento e a informação são os recursos estratégicos para o desenvolvimento de qualquer país. Os portadores desses recursos são as pessoas.”

Peter Drucker

RESUMO

O cenário contemporâneo é marcado pelo expressivo aumento na produção e disponibilidade de dados, impondo às organizações o desafio de convertê-los em informações relevantes e conhecimento aplicável ao processo decisório. No âmbito da Marinha do Brasil (MB), a gestão da carreira dos militares, conduzida pela Diretoria do Pessoal da Marinha (DPM), apresenta elevada complexidade por exigir a análise integrada de múltiplas variáveis relacionadas à trajetória militar. Entre os desafios existentes, destaca-se a evasão, que representa um fator de atenção para o fluxo de carreira, o dimensionamento de efetivos e a preservação do conhecimento organizacional. Nesse contexto institucional, a gestão do conhecimento (GC) compreende processos estruturados para criar, organizar, compartilhar e aplicar o conhecimento organizacional, garantindo que informações estratégicas estejam disponíveis no momento oportuno para subsidiar decisões. De forma complementar, a mineração de dados (MD) aplica métodos estatísticos e computacionais para identificar padrões e tendências em grandes volumes de dados, oferecendo subsídios analíticos que ampliam a precisão e a fundamentação das escolhas. A integração entre GC e MD, portanto, configura-se como uma alternativa promissora para transformar dados históricos já disponíveis em conhecimento aplicável, contribuindo para a qualidade e a precisão das decisões na gestão de pessoal. Este estudo teve como objetivo analisar o potencial de aplicação de técnicas de MD no planejamento de carreira da MB, com ênfase em modelos preditivos de classificação para identificação de perfis com maior propensão à evasão. A pesquisa, de caráter bibliográfico e descritivo, abordou tarefas e técnicas de MD, o contexto institucional e a comparação de métodos de classificação. Como resultado, as árvores de decisão mostraram-se adequadas ao problema, por conciliarem acurácia preditiva e clareza interpretativa, possibilitando identificar variáveis críticas e subsidiar políticas de retenção mais eficazes. Espera-se que a adoção desse modelo fortaleça o processo decisório, otimize a gestão de efetivos e contribua para a eficiência institucional da MB.

Palavras-chave: Mineração de dados. Gestão do conhecimento. Planejamento de carreira. Evasão militar. Gestão de efetivos.

ABSTRACT

The Integration Between Data Mining and Knowledge Management as Support for Decision-Making in the Personnel Management of the Brazilian Navy

The contemporary scenario is characterized by a significant increase in data production and availability, which poses to organizations the challenge of converting them into relevant information and applicable knowledge for decision-making processes. Within the Brazilian Navy (MB), career management of military personnel, conducted by the Directorate of Naval Personnel (DPM), involves high complexity due to the integrated analysis of multiple variables related to military trajectories. Among the aspects that require attention, attrition stands out, as it influences career flow, personnel sizing, and the preservation of organizational knowledge. In this institutional context, Knowledge Management (KM) encompasses structured processes to create, organize, share, and apply organizational knowledge, ensuring that strategic information is available at the right time to support decisions. Complementarily, Data Mining (DM) applies statistical and computational methods to identify patterns and trends in large datasets, providing analytical insights that can enhance the accuracy and consistency of choices. The integration of KM and DM, therefore, emerges as a promising alternative to transform available historical data into applicable knowledge, contributing to the quality and reliability of personnel management decisions. This study aimed to analyze the potential application of DM techniques in career planning within the MB, with emphasis on predictive classification models to identify profiles with higher propensity for attrition. The research, of bibliographic and descriptive nature, addressed DM tasks and techniques, the institutional context, and the comparison of classification methods. As a result, decision trees proved to be suitable for the problem, as they combine predictive accuracy with interpretability, enabling the identification of critical variables and supporting more effective retention policies. It is expected that the adoption of this model will strengthen decision-making, optimize personnel management, and contribute to the institutional efficiency of the MB.

Keywords: Data mining. Knowledge management. Career planning. Military attrition. Personnel management.

LISTA DE ILUSTRAÇÕES

FIGURA 1	– Resumo das principais tarefas analíticas na Mineração de Dados.....	18
FIGURA 2	– Comparação de técnicas de classificação utilizadas na modelagem preditiva.....	29

LISTA DE TABELA

TABELA 1 – Exclusão de Oficiais da Marinha do Brasil por tipo e ano (2019–2023).....	22
---	----

LISTA DE ABREVIATURAS E SIGLAS

DDD	–	<i>Data-Driven Decision Making</i>
DPM	–	Diretoria do Pessoal da Marinha
GC	–	Gestão do Conhecimento
IPA	–	Índice de Possibilidade de Acesso
k-NN	–	K-vizinhos mais próximos
MB	–	Marinha do Brasil
MD	–	Mineração de Dados
PCOM	–	Plano de Carreira de Oficiais da Marinha
PCPM	–	Plano de Carreira de Praças da Marinha
SPP	–	Sistema de Planejamento de Pessoal
SVM	–	Máquinas de Vetores de Suporte

SUMÁRIO

1	INTRODUÇÃO.....	12
2	FUNDAMENTAÇÃO TEÓRICA.....	13
2.1	GESTÃO DO CONHECIMENTO: CONCEITOS E ABORDAGENS.....	14
2.2	MINERAÇÃO DE DADOS: FUNDAMENTOS, TÉCNICAS E APLICAÇÕES.....	16
2.3	A INTEGRAÇÃO ENTRE MINERAÇÃO DE DADOS E GESTÃO DO CONHECIMENTO.....	18
3	CONTEXTO INSTITUCIONAL E APLICAÇÃO NA MB	20
3.1	PLANEJAMENTO DA CARREIRA DOS MILITARES	20
3.2	EVASÃO DE MILITARES	22
3.3	POTENCIAL DA MINERAÇÃO DE DADOS NA GESTÃO DE CARREIRA.....	23
4	PROPOSTA DE APLICAÇÃO E RESULTADOS ESPERADOS.....	25
4.1	MODELO PREDITIVO PARA IDENTIFICAÇÃO DE PERFIS DE EVASÃO	25
4.2	ESCOLHA DA TÉCNICA DE CLASSIFICAÇÃO PARA MODELOS PREDITIVOS	29
4.3	TOMADA DE DECISÃO APOIADA POR MINERAÇÃO DE DADOS.....	32
4.4	ASPECTOS DE IMPLEMENTAÇÃO: DESAFIOS E OPORTUNIDADES...	34
5	CONSIDERAÇÕES FINAIS.....	35
	REFERÊNCIAS	38

1 INTRODUÇÃO

O cenário contemporâneo caracteriza-se pelo expressivo aumento na produção e na disponibilidade de dados, fenômeno que impõe às organizações o desafio de transformá-los em informações relevantes e conhecimento aplicável ao processo decisório (Davenport; Prusak, 1998). Nesse contexto, a literatura destaca a gestão do conhecimento e a mineração de dados como abordagens complementares capazes de potencializar a criação, a organização e o uso de informações estratégicas. A gestão do conhecimento, ao estruturar processos de captura, sistematização e disseminação do conhecimento, contribui para preservar e ampliar o patrimônio intelectual das instituições (Nonaka; Takeuchi, 2008). A mineração de dados, por sua vez, recorre a métodos estatísticos e computacionais para identificar padrões e tendências em grandes volumes de dados, ampliando o suporte analítico para decisões fundamentadas (Han; Pei; Tong, 2023).

Pesquisas de Siuko, Myllärniemi e Hellsten (2023) demonstram que, no setor público, a aplicação integrada dessas abordagens contribui para aumentar a eficiência, promover a transparência e garantir a continuidade das ações institucionais. Os autores evidenciam que o uso combinado de gestão do conhecimento e mineração de dados permite análises mais amplas de dados administrativos e indicadores de desempenho, favorecendo a formulação de políticas e estratégias baseadas em evidências, especialmente em contextos que exigem conciliar recursos limitados com demandas estratégicas.

No âmbito da Marinha do Brasil, a gestão da carreira dos militares, conduzida pela Diretoria do Pessoal da Marinha, envolve processos complexos que requerem a análise integrada de múltiplas variáveis relacionadas à trajetória profissional, como formação, desempenho, tempo de serviço, movimentações, cumprimento de requisitos de carreira, entre outros aspectos pertinentes à carreira. Entre os desafios apontados na literatura, a evasão de militares se sobressai por impactar diretamente o equilíbrio do fluxo de carreira, o dimensionamento dos efetivos e a preservação do conhecimento organizacional (Raimundo; Correia Neto, 2017).

A revisão bibliográfica evidencia que modelos preditivos vêm sendo explorados para apoiar a gestão de pessoas, permitindo identificar perfis e antecipar cenários que demandem ações específicas (Chen; Chiang; Storey, 2012). Tais

técnicas, quando associadas à gestão do conhecimento e à mineração de dados, apresentam potencial para fortalecer o planejamento de carreira, apoiar a tomada de decisão e direcionar medidas voltadas à retenção de profissionais qualificados.

Este estudo foi desenvolvido em etapas complementares, planejadas para garantir uma análise consistente e alinhada aos objetivos propostos. Inicialmente, procedeu-se a uma pesquisa bibliográfica voltada à identificação de conceitos, métodos e aplicações de gestão do conhecimento e mineração de dados, com ênfase em modelos preditivos aplicados à gestão de pessoas. Em seguida, realizou-se uma pesquisa descritiva para caracterizar o processo de planejamento de carreira na MB. Complementarmente, conduziu-se uma análise comparativa entre diferentes técnicas de classificação, buscando identificar aquelas mais adequadas ao problema investigado, considerando critérios de desempenho e interpretabilidade. Por fim, propôs-se a aplicação de um modelo preditivo ao contexto da Diretoria do Pessoal da Marinha, considerando as informações levantadas e as análises realizadas, como forma de indicar seu potencial enquanto ferramenta de apoio à decisão na gestão de efetivos.

Além desta introdução, o estudo organiza-se em cinco capítulos. O segundo capítulo apresenta o referencial teórico que fundamenta os conceitos de gestão do conhecimento e mineração de dados, suas abordagens e aplicações. O terceiro capítulo descreve o contexto institucional da Marinha do Brasil, detalhando o processo de planejamento de carreira e os desafios relacionados à evasão. O quarto capítulo expõe a proposta do modelo preditivo, justificando sua pertinência, apontando resultados esperados, bem como oportunidades de aplicação. Por fim, o quinto capítulo traz as considerações finais, apresentando as conclusões e apontando caminhos para possíveis aplicações futuras no âmbito da Marinha do Brasil.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os referenciais teóricos que sustentam a presente pesquisa. O primeiro traz os conceitos da GC, destacando sua relevância para a criação, compartilhamento e aplicação do conhecimento organizacional. Na

sequência, o capítulo explora as tarefas e aplicações da MD, demonstrando como essas ferramentas contribuem para a descoberta de padrões e a geração de conhecimento a partir de grandes volumes de dados. Em última análise, discute-se a integração entre a GC e a MD, denotando como essa convergência pode potencializar os processos decisórios, e apoiar a gestão estratégica.

2.1 GESTÃO DO CONHECIMENTO: CONCEITOS E ABORDAGENS

Conforme Davenport e Prusak (1998), a GC envolve processos integrados de criação, captura e disseminação do saber, de modo a transformar dados e informações em conhecimento organizacional. Quando associada às tecnologias da informação, essa prática fortalece a inteligência institucional e possibilita converter informações dispersas, sejam de natureza tácita, como experiências e habilidades individuais, ou explícita, como documentos e processos formalizados, em conhecimento estruturado e acessível às pessoas certas, no momento adequado.

No setor público, a GC assume papel ainda mais relevante, pois contribui para promover maior transparência, eficiência nos processos e continuidade nas ações administrativas. Como destacam Siuko, Myllärniemi e Hellsten (2023), ela atua como impulsionadora da transformação digital nas instituições públicas, apoiando a modernização e permitindo que decisões sejam tomadas com base em dados acessíveis e atualizados em tempo real.

De acordo com Nonaka e Takeuchi (2008), o conhecimento organizacional provém da interação contínua entre conhecimento tácito e explícito, sendo esse processo representado no modelo *SECI* (Socialização, Externalização, Combinação e Internalização). Para que essa dinâmica ocorra de forma eficiente, é essencial que as organizações desenvolvam mecanismos eficazes para capturar, estruturar e reutilizar esse conhecimento. Esse ciclo pode ser compreendido em quatro etapas principais: criação do conhecimento, armazenamento e organização, compartilhamento e aplicação.

As etapas do ciclo de GC podem ser implementadas por distintos enfoques metodológicos, conforme descrito por diversos autores na literatura. A abordagem processual enfatiza a definição e a padronização de fluxos, métodos e sistemas para

gerenciar o conhecimento em todas as etapas do seu ciclo de vida (Davenport; Prusak, 1998). Já a abordagem estratégica compreende a GC como um pilar para vantagem competitiva e alinhamento das competências organizacionais aos objetivos institucionais (Probst; Raub; Romhardt, 2002).

A abordagem sociotécnica integra dimensões humanas, culturais e tecnológicas, reconhecendo que a efetividade da GC depende da harmonização entre sistemas e ferramentas, práticas colaborativas e uma cultura organizacional voltada ao compartilhamento do conhecimento. Essa perspectiva, discutida por Bostrom e Heinen (1977), sustenta que o desempenho organizacional resulta do equilíbrio entre fatores sociais e tecnológicos, o que é particularmente relevante em contextos onde decisões estratégicas demandam infraestrutura robusta e participação ativa das pessoas.

Por sua vez, a abordagem orientada a dados, destacada por Chen, Chiang e Storey (2012), associa a GC a métodos de análise de dados e inteligência artificial, permitindo transformar grandes volumes de informações em conhecimento aplicável. Essa perspectiva amplia a capacidade de identificar padrões, tendências e correlações relevantes, fortalecendo o apoio à decisão organizacional em contextos complexos e dinâmicos.

Entre as abordagens mencionadas, destacam-se, para esta pesquisa, a orientada a dados e a sociotécnica. A primeira, conforme proposta por Chen, Chiang e Storey (2012), apresenta maior aderência ao objetivo do estudo por considerar a análise de dados, incluindo métodos de MD, como elemento central para a criação e aplicação do conhecimento organizacional. Essa perspectiva aproxima a GC de métodos quantitativos capazes de extrair valor informacional a partir de grandes volumes de dados, favorecendo a geração de previsões e a tomada de decisões baseadas em evidências.

Já a abordagem sociotécnica, conforme Bostrom e Heinen (1977), reforça que a efetividade da GC depende de uma integração equilibrada entre dimensões humanas, culturais e tecnológicas, o que, no contexto da MB, é especialmente relevante para processos decisórios estratégicos que exigem tanto infraestrutura tecnológica sólida quanto práticas colaborativas e uma cultura voltada ao compartilhamento do conhecimento.

2.2 MINERAÇÃO DE DADOS: FUNDAMENTOS, TÉCNICAS E APLICAÇÕES

A mineração de dados (*data mining*), segundo Han, Pei e Tong (2023), é a extração automática de informações relevantes e previamente desconhecidas a partir de grandes volumes de dados, por meio da aplicação de algoritmos estatísticos e computacionais. Integra uma das etapas centrais do modelo Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Databases* – KDD), proposto por Fayyad *et al.* (1996). Esse modelo compreende cinco fases: seleção de dados relevantes; pré-processamento para remoção de ruídos e inconsistências; transformação dos dados para formatos adequados; mineração propriamente dita, com uso de algoritmos apropriados para extrair padrões; e interpretação e avaliação, que convertem os padrões úteis em conhecimento compreensível (Tan, Steinbach e Kumar, 2016).

Transformar grandes volumes de dados em conhecimento estratégico exige a aplicação de tarefas analíticas robustas e metodologicamente estruturadas, conforme previsto no campo da MD (Han; Pei; Tong, 2023). Essas tarefas, também denominadas funções ou operações analíticas, constituem a essência dos processos de extração de padrões e geração de *insights* a partir de bases de dados extensas e heterogêneas.

Entre as mais relevantes, a classificação consiste em construir modelos capazes de descrever e diferenciar classes ou conceitos previamente definidos, utilizando algoritmos supervisionados treinados com conjuntos de dados cujos resultados são conhecidos. O modelo obtido pode atribuir novos registros à classe correta, produzindo resultados interpretáveis e aplicáveis em diferentes contextos, motivo pelo qual é amplamente empregada em áreas que demandam decisões baseadas em dados confiáveis.

A regressão mapeia itens de dados para variáveis de valor real, permitindo prever valores contínuos. É especialmente útil em projeções numéricas, como previsões de vendas, estimativas financeiras e cálculos de demanda futura, utilizando métodos como regressão linear, regressão múltipla e determinadas configurações de redes neurais.

O agrupamento (*clustering*) é uma tarefa não supervisionada que busca identificar grupos naturais nos dados, reunindo registros com alta similaridade interna e baixa similaridade em relação a outros grupos. Essa abordagem é utilizada, por exemplo, na segmentação de mercado, na análise de padrões de comportamento e na organização de grandes acervos de informações não rotuladas.

A modelagem de dependência, ou aprendizado de regras de associação, identifica relações frequentes e significativas entre variáveis ou atributos. Um exemplo clássico é a análise de cestas de compras, que revela itens com maior probabilidade de serem adquiridos em conjunto, fornecendo subsídios para estratégias de marketing e sistemas de recomendação.

A detecção de desvios (ou anomalias) concentra-se na identificação de mudanças relevantes ou valores fora do padrão esperado, sendo fundamental para áreas como segurança da informação, prevenção de fraudes financeiras e monitoramento de equipamentos, por permitir identificar eventos raros que indicam falhas ou comportamentos suspeitos.

Por fim, a tarefa resumo envolve a elaboração de descrições condensadas de um conjunto de dados, que podem incluir visualizações, relatórios agregados ou generalizações estatísticas. Essa tarefa favorece a compreensão global do conjunto analisado e fornece suporte à tomada de decisão ao apresentar a informação de maneira sintética, porém representativa.

Assim, conforme sintetizado por Han, Pei e Tong (2023), essas seis tarefas constituem o núcleo da MD, permitindo que grandes volumes de informação sejam transformados em conhecimento útil para diferentes contextos organizacionais e científicos.

De forma geral, a classificação destaca-se por sua ampla aplicabilidade em diferentes domínios, especialmente em contextos que exigem a previsão de categorias a partir de dados históricos. Conforme salientam Silwattananusarn e Tuamsuk (2012), essa tarefa ocupa papel central em iniciativas de GC apoiadas por tecnologias analíticas, ao viabilizar a identificação de padrões relevantes e recorrentes nos dados institucionais.

As tarefas analíticas apresentadas podem ser visualizadas de forma consolidada na Figura 1, que sintetiza as principais funções aplicadas na MD e

evidencia suas contribuições para a extração de padrões e a geração de conhecimento estratégico. Essa compreensão conceitual é fundamental para discutir, a seguir, como a MD pode integrar-se aos processos de GC, ampliando a capacidade institucional de transformar dados em informações úteis e conhecimento acionável.

Figura 1 – Resumo das principais tarefas analíticas na Mineração de Dados

Tarefa Analítica	Descrição	Objetivo Principal
Classificação	Atribui registros a categorias pré-definidas com base em dados históricos.	Predizer categorias discretas.
Regressão	Estima valores contínuos a partir de variáveis preditoras.	Projetar valores futuros.
Agrupamento (<i>Clustering</i>)	Agrupa registros semelhantes sem categorias pré-definidas.	Descobrir padrões ou perfis ocultos.
Mineração de Padrões Frequentes / Associações	Identifica relações recorrentes entre variáveis.	Descobrir coocorrências relevantes.
Deteção de <i>Outliers</i>	Identifica registros com comportamento atípico.	Detectar anomalias ou casos extremos.
Resumo Multidimensional de Dados	Consolida e hierarquiza dados em múltiplas dimensões.	Permitir análises exploratórias dinâmicas.

Fonte: Adaptado de Han; Pei; Tong (2023).

2.3 A Integração entre Mineração de Dados e Gestão do Conhecimento

A gestão da carreira no âmbito do Setor de Pessoal da MB, sob responsabilidade da DPM, envolve um elevado grau de complexidade. O processo demanda a análise de um grande volume de informações institucionais, distribuídas em diferentes sistemas corporativos. Nesse contexto, a integração entre práticas de GC e técnicas de MD apresenta-se como uma abordagem particularmente promissora, ao possibilitar a transformação de dados dispersos em insumos analíticos capazes de qualificar o processo decisório.

O constante avanço das tecnologias da informação, aliado à crescente disponibilidade de dados gerados em diversos níveis organizacionais, amplia tanto as oportunidades quanto os desafios para as instituições. A capacidade de

transformar esses dados em conhecimento aplicável tornou-se um diferencial competitivo, especialmente para organizações que precisam tomar decisões fundamentadas e precisas. A GC desempenha papel essencial nesse processo ao sistematizar, compartilhar e aplicar o conhecimento acumulado nos processos organizacionais, favorecendo a inovação e a melhoria contínua.

Entretanto, para que a GC atinja todo o seu potencial, é indispensável contar com ferramentas capazes de lidar de forma eficaz com o grande volume de dados produzidos. Nessa conjuntura, a MD desponta como um recurso capaz de potencializar a GC, pois torna visível aquilo que está oculto nas bases de dados e revela padrões de valor estratégico. O alinhamento entre essas duas perspectivas, análise de dados e GC, favorece o planejamento, antecipa cenários e contribui para uma atuação mais assertiva e responsiva (Silwattananusarn e Tuamsuk, 2012).

No campo da literatura contemporânea, destaca-se a abordagem orientada a dados, que associa a GC a métodos de análise de dados e inteligência artificial, ampliando a capacidade de identificar padrões e apoiar decisões com base em evidências (Chen; Chiang; Storey, 2012). Essa perspectiva está alinhada à tendência de utilização de modelos preditivos e técnicas avançadas de MD para apoiar decisões estratégicas, reforçando o potencial dessa integração no contexto institucional.

Como destacam Fayyad, Piatetsky-Shapiro e Smyth (1996), nos ambientes institucionais, os bancos de dados concentram informações de alto valor estratégico, e a MD é capaz de convertê-las em conhecimento aplicável, impulsionando o aprendizado organizacional e superando o papel restrito de ferramenta analítica. No contexto da DPM, essa capacidade assume importância ainda maior, pois a gestão de carreira demanda análises abrangentes e integradas que permitam não apenas compreender o cenário atual, mas também antecipar tendências e necessidades futuras. A aplicação de técnicas de MD, nesse sentido, torna-se um instrumento decisivo para extrair, consolidar e interpretar informações de diferentes fontes, subsidiando a formulação de políticas e práticas de gestão de pessoal mais eficazes, proativas e alinhadas aos objetivos institucionais.

Ao concluir este capítulo, observa-se que a integração entre GC e MD forma um referencial teórico e prático capaz de ampliar o aproveitamento estratégico das

informações institucionais. No âmbito da DPM, essa convergência favorece a transformação de dados históricos em conhecimento aplicável, apoiando a tomada de decisões e servindo de base para o desenvolvimento da proposta que será detalhada nos capítulos seguintes.

3 CONTEXTO INSTITUCIONAL E APLICAÇÃO NA MB

Considerando o potencial da MD como suporte à GC, este capítulo apresenta o contexto institucional da MB, com foco na estrutura de planejamento da carreira dos militares e nos desafios relacionados à evasão de pessoal. Adicionalmente, são discutidas possibilidades de integração de técnicas de MD aos processos de GC no âmbito da MB, de modo a ampliar a capacidade analítica e fornecer subsídios consistentes ao processo decisório.

3.1 PLANEJAMENTO DA CARREIRA DOS MILITARES

O Sistema de Planejamento de Pessoal (SPP) é um conjunto de elementos da estrutura administrativa do Comando da Marinha destinado a orientar o cumprimento das atividades inerentes à logística e à otimização dos recursos humanos empregados na MB. Por meio do SPP, é possível determinar as necessidades de pessoal e planejar as atividades decorrentes dessa demanda (Brasil, 2016).

O planejamento de carreira busca atender às necessidades de pessoal de cada Corpo ou Quadro, distribuídas por postos e graduações, de acordo com a quantidade e a qualificação exigidas, além de orientar os militares quanto às etapas a serem cumpridas ao longo de suas trajetórias profissionais. Nesse processo, as Organizações Militares atuam como agentes responsáveis pelas atividades logísticas de recursos humanos, que envolvem a procura e admissão, destinada à captação de recursos humanos na sociedade para o serviço na Marinha; a preparação, que compreende a formação básica, especialização e aperfeiçoamento do efetivo, assegurando sua prontidão e capacidade para o exercício de cargos; e a administração, voltada ao preenchimento dos cargos e funções previstos nas Tabelas de Lotação da Marinha (Brasil, 2016).

A estruturação da carreira visa estimular o desenvolvimento pessoal e a realização profissional, assegurando proficiência no exercício dos diversos cargos da Marinha. O detalhamento encontra-se nos Planos de Carreira, que compõem o SPP e contêm orientações sobre o emprego dos militares de carreira, dividindo-se em Plano de Carreira de Oficiais da Marinha (PCOM) e Plano de Carreira de Praças da Marinha (PCPM) (Brasil, 2016).

O fluxo de carreira corresponde à ascensão seletiva, gradual e sucessiva na hierarquia militar, concretizada por meio de promoções e balizada por critérios como o processo de escolha, o reajustamento dos interstícios, a aplicação da quota compulsória, a transferência entre Corpos e Quadros e a distribuição de efetivos. O interstício é a condição de acesso representada pelo tempo mínimo de permanência, em efetivo serviço, em cada posto ou graduação dos diversos Corpos e Quadros da Marinha. Os interstícios de planejamento são definidos considerando o emprego do militar no posto ou graduação e o tempo necessário para a obtenção dos requisitos de acesso ao nível superior, enquanto os interstícios para acesso são estabelecidos a partir dos de planejamento, com o objetivo de regularizar o fluxo de carreira de cada Corpo ou Quadro (Brasil, 2016).

Os fluxos ideais de carreira são dimensionados com base em dois instrumentos: as curvas-padrão, que apresentam o perfil de carreira do Corpo ou Quadro a partir das necessidades de militares em cada posto ou graduação, dos Planos de Carreira específicos, das taxas de administração e das evasões previsíveis; e o Índice de Possibilidade de Acesso (IPA), que indica a variação percentual planejada para a distribuição do efetivo entre postos ou graduações consecutivos (Brasil, 2016).

Nessa perspectiva, a previsão de evasão desempenha papel relevante, sendo calculada a partir de dados históricos que representam a média de militares que deixam determinado posto ou graduação, por diferentes motivos, ao longo dos anos. Esses valores são incorporados às fórmulas de dimensionamento para ajustar o efetivo médio em cada posto ou graduação, contribuindo para maior precisão no planejamento. Assim, compreender e projetar o impacto da evasão torna-se uma estratégia importante para manter o equilíbrio do fluxo de carreira e a disponibilidade de pessoal qualificado, tema que será aprofundado na seção seguinte.

3.2 EVASÃO DE MILITARES

Considerando a diversidade de formas de desligamento existentes, este trabalho adota o termo *evasão* como categoria geral para compreender diferentes situações que resultam na saída do militar do serviço ativo da MB. Para fins analíticos, destacam-se três modalidades: licenciamento a pedido, referente aos oficiais ainda sem estabilidade; demissão a pedido, aplicável aos oficiais de carreira que solicitam a saída do quadro permanente; e demissão *ex officio*, que ocorre quando o desligamento decorre da posse em cargo público.

O quadro a seguir apresenta os dados consolidados do período de 2019 a 2023, contemplando o total de exclusões de oficiais nas três modalidades citadas.

Tabela 1 – Exclusão de Oficiais da Marinha do Brasil por tipo e ano (2019–2023)

Ano	Demissão a pedido	Demissão <i>ex officio</i>	Licenciamento a pedido
2019	19	6	82
2020	9	5	77
2021	33	7	77
2022	50	25	104
2023	48	22	79

Fonte: Adaptado do Anuário Estatístico da Marinha do Brasil.

Observa-se que, no período analisado, o licenciamento a pedido apresentou, de forma geral, os maiores quantitativos anuais. Já as demissões a pedido e as demissões *ex officio* variaram em menor escala, registrando picos mais expressivos em 2022 e 2023. Ainda que, à primeira vista, esses números representem apenas registros anuais, eles concentram informações valiosas que, quando submetidas a métodos analíticos e de mineração de dados, permitem identificar relações entre variáveis e compreender fatores subjacentes que influenciam a decisão de desligamento.

Embora os dados apresentados neste estudo se refiram exclusivamente à MB, a compreensão do fenômeno da evasão é igualmente relevante para as demais Forças Armadas, uma vez que estudos acadêmicos apontam fatores estruturais

comuns à carreira militar no país. Conforme analisam Vito *et al.* (2022), a saída precoce de militares pode estar relacionada a aspectos como transferências para localidades indesejadas, escalas de serviço que resultam em jornadas excessivas e não remuneradas, disponibilidade integral, lentidão nas promoções e atratividade do mercado civil para os perfis formados nas Forças Armadas.

De forma convergente, Raimundo e Correia Neto (2017) destacam elementos como a falta de perspectivas no plano de carreira, questões salariais e transferências não desejadas como contribuintes para a decisão de desligamento. Essas constatações reforçam a importância de um monitoramento contínuo e sistemático da evasão, permitindo identificar causas recorrentes, antecipar tendências e orientar políticas de retenção mais eficazes.

Ao compreender os perfis mais suscetíveis ao desligamento precoce, a MB pode alinhar seu planejamento de carreira às necessidades estratégicas da organização, reduzindo impactos sobre a continuidade das atividades e preservando o retorno do investimento realizado na formação de seus oficiais. Nesse cenário, a aplicação de técnicas analíticas avançadas, como a MD integrada à GC, desponta como um recurso valioso para identificar padrões, antecipar tendências e apoiar decisões que fortaleçam a retenção de pessoal.

3.3 POTENCIAL DA MINERAÇÃO DE DADOS NA GESTÃO DE CARREIRA

A gestão da carreira dos militares envolve decisões estratégicas relacionadas à movimentação, promoções, capacitações, requerimentos e outros elementos da trajetória profissional. Esse processo demanda análise criteriosa de grandes volumes de dados acumulados ao longo do percurso do militar na Instituição, contemplando informações sobre formação, desempenho em avaliações, tempo de serviço, funções exercidas, histórico de movimentações e cumprimento de requisitos de carreira. Nesse cenário, a MD configura-se como ferramenta capaz de transformar dados brutos em conhecimento estratégico, oferecendo suporte mais preciso e proativo às decisões que impactam diretamente a carreira. Conforme Provost e Fawcett (2013), a aplicação da ciência de dados na gestão institucional possibilita decisões mais rápidas, precisas e alinhadas aos objetivos estratégicos da

organização.

Considerando essa abordagem, as pessoas constituem ativos estratégicos cujo desenvolvimento, retenção e alocação adequada influenciam diretamente o desempenho organizacional. Schultz (2020) observa que, em um ambiente cada vez mais dinâmico, a MD viabiliza análises qualificadas de fenômenos como a rotatividade interna, subsidiando ações de retenção e alocação inteligente. As decisões que antes se apoiavam majoritariamente na experiência e intuição dos gestores passam, gradativamente, a fundamentar-se em informações derivadas de dados estruturados e analisados de forma sistemática, aumentando a assertividade e reduzindo riscos.

No âmbito da DPM, destaca-se a existência de um extenso acervo de dados históricos relativos à trajetória dos militares, reunidos ao longo de décadas por meio de diferentes sistemas informatizados. Esse patrimônio informacional, quando processado com o apoio de técnicas analíticas, transforma-se em um recurso estratégico para investigações aprofundadas, dentre as quais se sobressai a análise da evasão, fenômeno que afeta diretamente o fluxo de carreira e influencia desde a captação e distribuição de efetivos até a definição de critérios como o interstício.

Atualmente, a previsão de evasão tem se apoiado principalmente em médias históricas de desligamentos anuais, constituindo-se em uma referência inicial relevante. Esse procedimento pode ser complementado com a incorporação de variáveis contextuais e institucionais, de modo a ampliar a compreensão dos fatores que influenciam a decisão do militar de permanecer ou deixar a Força.

Dessa maneira, as técnicas de MD permitem integrar diferentes variáveis e identificar padrões que oferecem suporte consistente às decisões estratégicas. Conforme destacam Silwattananusarn e Tuamsuk (2012), essa capacidade de revelar informações ocultas confere à MD um papel central na GC, fornecendo subsídios sólidos para estratégias de retenção e para o aprimoramento do planejamento de carreira.

Em síntese, este capítulo demonstrou que o planejamento de carreira na MB pode ser fortalecido pela integração entre GC e MD, ampliando a capacidade institucional de analisar e direcionar a trajetória profissional dos militares ao longo de toda a carreira. Entre os diversos fenômenos que podem ser explorados por meio

dessa integração, a evasão se destaca como um ponto crítico, cuja análise aprofundada pode gerar subsídios valiosos para antecipar cenários e orientar decisões estratégicas, estabelecendo a base para a proposta apresentada no próximo capítulo.

4 PROPOSTA DE APLICAÇÃO E RESULTADOS ESPERADOS

Considerando o potencial da MD como suporte à GC, esta seção apresenta uma abordagem analítica voltada ao planejamento da carreira dos militares, com foco nos desafios relacionados à evasão de pessoal. Busca-se demonstrar como a exploração estruturada de informações institucionais pode contribuir para a geração de conhecimento estratégico, subsidiando decisões gerenciais no âmbito da Administração Naval.

4.1 MODELO PREDITIVO PARA IDENTIFICAÇÃO DE PERFIS DE EVASÃO

Um modelo preditivo é uma estrutura matemática ou computacional desenvolvida com o objetivo de estimar a probabilidade de ocorrência de eventos futuros com base em dados históricos. Ele aprende padrões e relações entre variáveis (entradas) e um resultado esperado (saída), aplicando esse conhecimento para prever novos casos. Segundo Han, Pei e Tong (2023), os modelos preditivos integram as tarefas supervisionadas de MD, notadamente a classificação e a regressão, sendo construídos a partir de conjuntos de treinamento cujos resultados são previamente conhecidos. A lógica aprendida pelo modelo é, então, aplicada a novos dados com o objetivo de inferir comportamentos ou categorias futuras.

A classificação é empregada quando a variável-alvo é categórica e não ordenada, como no caso da evasão, que pode ser representada por rótulos como “evadiu” ou “não evadiu”. A principal vantagem está em sua capacidade de identificar padrões recorrentes em atributos históricos que favorecem a previsão de eventos futuros, auxiliando na definição de políticas mais direcionadas. Já a regressão é utilizada quando o objetivo é prever valores numéricos contínuos, como, por exemplo, o tempo estimado até a evasão ou a probabilidade percentual de desligamento de um militar em determinado intervalo de tempo.

De acordo com Schultz (2020), a aplicação de modelos preditivos na gestão de pessoas possibilita transformar grandes volumes de dados funcionais em conhecimento útil para embasar decisões estratégicas. No contexto da evasão de militares, a escolha do método de MD mais apropriado dependerá da natureza e da qualidade dos dados disponíveis, do volume histórico armazenado, da granularidade das variáveis analisadas e, sobretudo, dos objetivos analíticos definidos. É necessário avaliar se o foco está na explicação dos fatores associados ao fenômeno ou na obtenção do melhor desempenho preditivo.

Considerando que a evasão configura um evento binário ou categórico, os métodos de classificação preditiva se mostram particularmente indicados. Essas técnicas permitem não apenas estimar a probabilidade de ocorrência do evento com base em características históricas e contextuais da carreira militar, mas também revelar padrões e perfis de risco, contribuindo de forma estratégica para o planejamento de efetivos e para a formulação de políticas de retenção mais eficazes. Para que tais modelos alcancem desempenho e confiabilidade adequados, contudo, torna-se imprescindível realizar um rigoroso processo de pré-processamento dos dados, garantindo a qualidade e a relevância das variáveis a serem analisadas.

O pré-processamento dos dados representa uma etapa fundamental na construção de modelos preditivos, uma vez que garante a qualidade, a consistência e a relevância das informações analisadas. Esse processo abrange a identificação e o tratamento de registros incompletos, inconsistentes ou duplicados, a normalização de atributos com escalas distintas e a seleção das variáveis mais significativas para o problema em estudo (Han; Pei; Tong, 2023; Shmueli; Bruce; Patel, 2020). Tais procedimentos contribuem para a redução da dimensionalidade dos dados, o aumento da eficiência computacional e a melhoria da interpretabilidade dos resultados (Witten; Frank; Hall; Pal, 2017), assegurando que o modelo seja construído sobre uma base informacional sólida. Dessa forma, favorecem-se análises mais precisas e alinhadas às necessidades do planejamento de carreira e da formulação de políticas de retenção.

Dentre as técnicas de classificação mais conhecidas, destacam-se: árvores de decisão, redes neurais artificiais, máquinas de vetores de suporte (SVM),

classificador *Naïve Bayes* e k-vizinhos mais próximos (k-NN). Cada uma apresenta características próprias, vantagens, limitações e diferentes graus de interpretabilidade, fatores que devem ser considerados na escolha da técnica mais adequada ao problema.

As árvores de decisão organizam o processo de classificação em uma estrutura hierárquica composta por nós de decisão, ramos e folhas. Cada nó de decisão representa um teste lógico baseado em um atributo dos dados; cada ramo indica o resultado possível desse teste; e cada folha corresponde a uma classe final ou categoria de previsão (Han; Pei; Tong, 2023). A construção da árvore ocorre por meio da divisão sucessiva do conjunto de dados em subconjuntos cada vez mais homogêneos em relação à variável-alvo, buscando maximizar a pureza de cada grupo resultante. Uma das principais vantagens dessa técnica é sua alta interpretabilidade, pois o resultado final pode ser expresso em regras claras do tipo “SE... ENTÃO...”, permitindo que os critérios de classificação sejam facilmente compreendidos e auditados (Dreyer Galvão; Marin, 2009).

No contexto do planejamento de carreira, essa interpretabilidade torna-se especialmente relevante, pois possibilita que gestores compreendam de forma transparente quais combinações de variáveis estão associadas a diferentes probabilidades de evasão. Essa clareza não apenas facilita a comunicação dos resultados em diferentes níveis hierárquicos, como também fortalece a confiabilidade das decisões baseadas no modelo.

As redes neurais artificiais são modelos inspirados na estrutura do cérebro humano, compostos por unidades de processamento interligadas, organizadas em camadas de entrada, ocultas e de saída. Essas redes possuem alta capacidade de modelar relações complexas e não lineares entre atributos e classes (Han; Pei; Tong, 2023), sendo particularmente eficientes em cenários com grandes volumes de dados e variáveis interdependentes. No entanto, apresentam baixa interpretabilidade, funcionando como “caixas-pretas” que dificultam a compreensão dos critérios utilizados para chegar a determinada previsão (Carvalho; Pereira; Cardoso, 2019). Isso representa um desafio quando se busca justificar as decisões a partir do modelo, como no caso de gestão de pessoal.

As máquinas de vetores de suporte atuam encontrando o hiperplano que melhor separa as classes, maximizando a margem entre elas (Han; Pei; Tong, 2023). Quando os dados não são linearmente separáveis, utilizam funções de *kernel* para projetar as informações em espaços de maior dimensão, em que a separação linear seja viável. Essa técnica apresenta boa capacidade de generalização e, em muitos casos, acurácia elevada, especialmente em problemas binários. Contudo, sua interpretabilidade é limitada, principalmente quando se empregam *kernels* não lineares (Alejandrino, 2021). Além disso, o processo de treinamento do modelo pode demandar elevado custo computacional em bases de dados muito extensas. No caso específico da evasão de militares, essa técnica é capaz de produzir previsões com elevado grau de precisão. No entanto, sua utilização como principal ferramenta para apoiar decisões estratégicas apresenta uma limitação, pois o modelo não oferece uma explicação intuitiva sobre como chegou às conclusões, o que pode dificultar a interpretação e o uso dos resultados por gestores que não sejam especialistas na área.

O *Naïve Bayes* é um classificador probabilístico baseado no teorema de Bayes, que assume independência condicional entre os atributos dado a classe (Han; Pei; Tong, 2023). Apesar de tal suposição raramente ocorrer de forma plena no mundo real, o método mantém bom desempenho em diversos cenários devido à sua simplicidade e eficiência. Sua capacidade de fornecer probabilidades associadas às classes permite priorizar casos de maior risco com base na probabilidade estimada de evasão, apoiando ações preventivas direcionadas. No entanto, a limitação em capturar interações complexas entre atributos pode comprometer a acurácia em situações em que essas interdependências sejam relevantes para o fenômeno analisado.

O método dos k-vizinhos mais próximos (k-NN) classifica uma instância com base nas classes dos k exemplos mais próximos no conjunto de treinamento, medindo-se a proximidade por meio de métricas de distância (Han; Pei; Tong, 2023). Trata-se de uma abordagem simples, que não exige treinamento complexo, mas que apresenta baixa interpretabilidade e pode ter desempenho afetado por atributos irrelevantes e pela alta dimensionalidade dos dados (Dreyer Galvão; Marin, 2009).

Essas técnicas, embora distintas em concepção e aplicação, apresentam pontos fortes e limitações que precisam ser ponderados à luz do contexto institucional e dos objetivos estratégicos da análise. Para sintetizar essas informações, a Figura 2 apresenta uma comparação das principais características dessas técnicas de classificação, evidenciando seu potencial de aplicação, vantagens e restrições no contexto de modelagem preditiva.

Figura 2 – Comparação de técnicas de classificação utilizadas na modelagem preditiva

Técnica de Classificação	Descrição resumida	Benefícios relevantes	Limitações
Árvore de decisão	Estrutura hierárquica que divide os dados em testes lógicos, gerando regras do tipo "SE... ENTÃO...".	Alta interpretabilidade; identifica variáveis mais relevantes; comunicação clara dos resultados; decisões transparentes.	Pode sofrer <i>overfitting</i> ; árvores muito complexas requerem ajuste de parâmetros e poda para manter generalização.
Redes neurais artificiais	Modelo inspirado no funcionamento cerebral; captura padrões complexos e não lineares.	Alta capacidade preditiva; detecta padrões sutis; adequada para grandes volumes de dados.	Baixa interpretabilidade ("caixa-preta"); exige grande volume de dados para bom desempenho.
Máquinas de vetores de suporte (SVM)	Define hiperplano ótimo para separar classes, maximizando margens.	Alta acurácia em problemas binários; boa generalização.	Difícil interpretação; custo computacional elevado em grandes bases.
Naïve Bayes	Classificador probabilístico baseado no teorema de Bayes; assume independência entre atributos.	Simple e rápido; fornece probabilidades por classe; útil para priorizar casos de risco.	Suposição de independência raramente atendida; menor desempenho em dados correlacionados.
k-Vizinhos mais próximos (k-NN)	Classifica instâncias com base nos exemplos mais próximos no treinamento.	Fácil implementação; útil para identificar casos semelhantes; não exige treinamento complexo.	Baixa interpretabilidade; sensível a atributos irrelevantes e à alta dimensionalidade.

Fonte: Adaptado de Han; Pei; Tong (2023), Dreyer Galvão; Marin (2009), Chen; Chiang; Storey (2012).

4.2 ESCOLHA DA TÉCNICA DE CLASSIFICAÇÃO PARA MODELOS PREDITIVOS

A definição da técnica de classificação mais adequada para prever evasões exige uma análise cuidadosa dos objetivos do estudo e das características dos dados disponíveis. Conforme discutido, a evasão de militares configura-se como um evento categórico (evadiu ou não evadiu), o que direciona naturalmente a tarefa para algoritmos de classificação supervisionada em vez de regressão. A decisão deve equilibrar desempenho preditivo e transparência dos resultados, atendendo às necessidades estratégicas da gestão de pessoal da DPM.

A escolha da técnica analítica mais adequada depende de diversos fatores, como a quantidade de dados históricos disponíveis, a qualidade e o nível de detalhamento (granularidade) das variáveis coletadas, bem como o objetivo da análise, se voltado à explicação de fenômenos ou à previsão de ocorrências futuras. Esses aspectos, conforme discutido anteriormente, influenciam diretamente o desempenho e a aplicabilidade dos modelos construídos.

Em resumo, o método de MD adotado deve não apenas prever com precisão quais militares apresentam maior probabilidade de evasão, mas também identificar os fatores que contribuem para esse comportamento. Dessa forma, os resultados gerados poderão subsidiar ações alinhadas aos objetivos institucionais de retenção e valorização do efetivo.

Diante das alternativas analisadas, as árvores de decisão configuram-se, neste estudo, como a escolha mais adequada para o modelo preditivo de evasão. A técnica alia desempenho satisfatório e clareza interpretativa, atributos essenciais para que as previsões possam subsidiar decisões estratégicas de gestão de pessoal de forma transparente e justificável. Além disso, possibilita identificar variáveis e combinações de fatores que mais influenciam a evasão, fornecendo subsídios objetivos para a formulação de políticas de retenção.

Na pesquisa de Mercês (2023), por exemplo, técnicas de *gradient boosting* (*XGBoost* e *CatBoost*) figuraram entre as de melhor desempenho na tarefa de identificar Oficiais propensos a deixarem o Exército. Contudo, essas abordagens sofisticadas operam como verdadeiras caixas-pretas, dificultando a extração de justificativas simples para as predições realizadas. Em um cenário de gestão de pessoal, tal opacidade pode representar um entrave, uma vez que recomendações de políticas ou intervenções precisam estar embasadas em motivos compreensíveis.

De acordo com Han, Pei e Tong (2023), a interpretabilidade constitui elemento essencial para a adoção de modelos preditivos em contextos que demandam decisões estratégicas, pois garante que os resultados possam ser compreendidos e auditados por profissionais sem formação especializada em ciência de dados. De forma convergente, Carvalho, Pereira e Cardoso (2019) ressaltam que, embora modelos mais complexos, como aqueles baseados em *ensemble*, que combinam múltiplos algoritmos ou redes neurais profundas, possam apresentar ganhos

pontuais de acurácia, a falta de transparência nos critérios utilizados para gerar as previsões pode restringir sua utilização como ferramenta principal de apoio à decisão.

Por esse motivo, optou-se por utilizar modelos baseados em árvore de decisão, cuja estrutura lógica e transparente favorece a compreensão por parte dos gestores, sem comprometer significativamente a acurácia preditiva. Sua capacidade de gerar regras explícitas e de fácil validação facilita a incorporação dos resultados às decisões institucionais, atendendo aos critérios de interpretabilidade exigidos em ambientes organizacionais. Assim, as árvores de decisão contribuem para aproximar o domínio técnico da análise de dados da realidade prática da gestão, funcionando como uma ponte entre o conhecimento especializado e a tomada de decisão estratégica.

A adoção dessa técnica implica considerar mecanismos que assegurem equilíbrio entre precisão e capacidade de generalização. Nesse sentido, destacam-se estratégias como a poda (*pruning*), que simplifica a árvore eliminando ramos pouco relevantes, e a validação cruzada, que testa o desempenho do modelo em diferentes subconjuntos de dados. Esses procedimentos evitam o sobreajuste (*overfitting*) e contribuem para que a árvore mantenha sua aplicabilidade em cenários inéditos, ampliando a confiabilidade dos resultados obtidos.

Em síntese, a adoção de árvores de decisão oferece uma solução equilibrada entre precisão e interpretabilidade. Essa escolha está alinhada às recomendações da literatura em MD e às particularidades do problema estudado, que requer informações claras e acionáveis sobre os perfis de risco. Com o modelo arbóreo, será possível identificar militares com maior probabilidade de evadir e simultaneamente elucidar quais fatores, isolados ou combinados, contribuem para decisão, subsidiando assim a elaboração de estratégias de intervenção e políticas de retenção mais eficazes e focalizadas.

Assim, a técnica escolhida cumpre simultaneamente o papel de previsão e de prevenção, ao possibilitar a identificação antecipada de potenciais desligamentos e oferecer subsídios objetivos para a adoção de medidas proativas que favoreçam a retenção da força de trabalho. Além disso, contribui para fortalecer a capacidade

institucional de planejar cenários futuros com maior segurança, reduzindo incertezas e ampliando a eficácia das políticas de gestão de pessoal.

4.3 TOMADA DE DECISÃO APOIADA POR MINERAÇÃO DE DADOS

A tomada de decisão é reconhecida como elemento essencial para o alcance de resultados organizacionais sustentáveis, pois orienta a formulação de estratégias, a alocação de recursos e a execução de ações alinhadas aos objetivos institucionais. Conforme Schultz (2020), decisões baseadas em análises consistentes de dados permitem direcionar recursos de forma mais precisa, alinhar ações aos objetivos organizacionais e responder com agilidade a mudanças no contexto.

Nesse sentido, como destacam Chen, Chiang e Storey (2012), a adoção de abordagens orientadas por evidências, apoiadas por ferramentas analíticas robustas, contribui para reduzir a subjetividade, mitigar riscos e ampliar a eficácia das políticas e práticas organizacionais. Assim, a qualificação contínua do processo decisório configura-se como requisito indispensável à gestão organizacional.

A retenção de pessoal qualificado constitui aspecto estratégico para organizações, incluindo as Forças Armadas, dada a necessidade de preservar o investimento contínuo realizado na formação e no treinamento ao longo da carreira dos militares (Raimundo; Correia Neto, 2017). A perda desses profissionais implica não apenas desperdício de recursos, mas também riscos à eficiência operacional e à manutenção do conhecimento organizacional acumulado.

Estudos apontam que a decisão de permanência ou desligamento é influenciada por múltiplos fatores, como remuneração, oportunidades de progressão, condições de trabalho, localização de lotação e aproveitamento de especializações (Raimundo; Correia Neto, 2017; Vito *et al.*, 2022). A compreensão integrada desses elementos é essencial para subsidiar decisões estratégicas voltadas à retenção de efetivos e, conseqüentemente, à sustentação da capacidade operacional e administrativa da MB.

Diante do exposto na seção anterior, a árvore de decisão mantém-se como abordagem preditiva adequada, por conciliar exatidão e transparência nos resultados. Sua estrutura interpretável facilita a conversão de padrões identificados

em diretrizes operacionais, contribuindo para a antecipação de cenários críticos, a mitigação de impactos institucionais e a preservação do capital humano e do conhecimento organizacional.

Considerando os princípios da GC, a utilização da MD para a retenção de efetivos caracteriza-se como um processo sistemático e contínuo de aprimoramento. A cada ciclo, dados sobre permanências e desligamentos enriquecem a base institucional, possibilitando o refinamento dos modelos, a atualização de parâmetros analíticos e a avaliação objetiva da eficácia das políticas adotadas. Consolida-se, assim, uma cultura de decisão orientada por dados, em que inteligência analítica e experiência administrativa se complementam. Nessa perspectiva, verifica-se que o uso de árvores de decisão apresenta potencial para oferecer benefícios relevantes, tais como:

- definição da captação de pessoal, apoiando a identificação de quantitativos e perfis de militares a serem incorporados, alinhados às demandas administrativas;
- otimização da distribuição de efetivos, subsidiando a alocação equilibrada entre organizações militares, evitando carências ou excedentes;
- aprimoramento do planejamento do interstício, ajustando prazos de permanência em postos ou graduações conforme necessidades estratégicas;
- identificação de padrões críticos de evasão, detectando variáveis ou combinações de fatores que elevam o risco de desligamento;
- classificação de perfis de risco, agrupando militares com maior propensão à evasão e possibilitando ações preventivas específicas;
- fortalecimento da comunicação decisória, sustentando decisões com evidências objetivas e facilitando a compreensão em diferentes níveis de comando;
- consolidação do conhecimento organizacional, transformando dados dispersos em conhecimento estruturado e institucionalizado, favorecendo o aprendizado organizacional.

Assim, o uso de árvores de decisão, aliado à análise sistemática de dados, apresenta-se como um recurso promissor para apoiar decisões estratégicas

alinhadas às demandas e aos objetivos da instituição. Essa técnica contribui para a definição de políticas de gestão mais assertivas e o fortalecimento da capacidade institucional de responder de maneira ágil e fundamentada a contextos de elevada complexidade, como o planejamento de efetivos.

4.4 ASPECTOS DE IMPLEMENTAÇÃO: DESAFIOS E OPORTUNIDADES

A aplicação de modelos preditivos na gestão de carreira dos militares da MB, embora promissora, demanda atenção a aspectos técnicos, organizacionais e culturais que podem influenciar diretamente sua efetividade. A implementação deve considerar as particularidades da estrutura administrativa da instituição, garantindo alinhamento com as diretrizes institucionais e com a missão da DPM.

Entre os principais desafios identificados estão:

- integrar o modelo preditivo aos sistemas existentes na DPM, assegurando interoperabilidade e consistência de dados entre plataformas legadas e heterogêneas, o que pode demandar atualização de sistemas e padronização de formatos;
- capacitar gestores para compreender e interpretar os resultados, limitações e pressupostos do modelo preditivo, permitindo o uso integrado de indicadores de risco, probabilidades e fatores explicativos na tomada de decisão. Na DPM, essa competência possibilita transformar previsões em ações concretas, abrangendo desde a definição da captação e distribuição de efetivos até a retenção, o aprimoramento do planejamento do interstício e a consolidação do conhecimento organizacional, fortalecendo de forma abrangente a gestão de carreira dos militares da MB.

Principais oportunidades de aplicação:

- integrar com sistemas de *Business Intelligence* (BI), possibilitando que o modelo preditivo seja incorporado a plataformas que ofereçam visualizações dinâmicas dos resultados, como *dashboards*, indicadores e painéis interativos. Essa integração facilitaria o monitoramento contínuo de tendências e a identificação rápida de áreas críticas, permitindo que

decisões estratégicas sejam tomadas de forma mais ágil, fundamentada e transparente;

- realizar análises preditivas em tempo real, aproveitando o potencial do processamento de dados em fluxo contínuo (*streaming data*). Essa capacidade permitiria atualizar de forma instantânea as estimativas de risco de evasão e outros indicadores relevantes para a gestão de carreira, incorporando dados provenientes de diferentes sistemas corporativos da MB. Tal atualização contínua ampliaria significativamente a capacidade da instituição de responder proativamente a mudanças no contexto operacional e administrativo, assegurando decisões mais rápidas e precisas;
- expandir para previsão de outros eventos de carreira, estendendo a metodologia inicialmente voltada à evasão para contemplar fenômenos correlatos, igualmente estratégicos para o planejamento e gestão da carreira militar. Essa ampliação tornaria o modelo uma ferramenta abrangente de apoio à gestão de pessoal, fortalecendo a capacidade institucional de monitorar padrões, antecipar cenários e adotar medidas proativas em diferentes etapas da trajetória profissional dos militares.

Em síntese, a proposta apresentada neste capítulo demonstra que a aplicação estruturada de técnicas de MD, em especial as árvores de decisão, pode fortalecer significativamente o planejamento e a gestão da carreira dos militares da MB. Ao oferecer o potencial de gerar previsões mais precisas e interpretáveis sobre a probabilidade de evasão, o modelo pode subsidiar ações direcionadas à retenção e contribuir para decisões mais qualificadas ao longo de todas as etapas da trajetória militar.

5 CONSIDERAÇÕES FINAIS

A presente monografia teve como propósito investigar o potencial da integração entre a gestão do conhecimento e a mineração de dados no apoio à gestão da carreira dos militares da Marinha do Brasil, com ênfase no planejamento de pessoal. Considerando o contexto contemporâneo, em que as organizações

enfrentam transformações aceleradas impulsionadas pelo avanço das tecnologias da informação e pela crescente geração de dados, a capacidade de processar grandes volumes de informações e convertê-las em conhecimento estratégico tornou-se um diferencial essencial para garantir o desempenho e a continuidade das instituições.

Ao longo do estudo, foram revisados os conceitos e abordagens centrais da gestão do conhecimento e mapeadas as principais tarefas da mineração de dados. Essa fundamentação teórica evidenciou que técnicas analíticas avançadas permitem extrair padrões relevantes e transformar dados em conhecimento aplicável, ampliando a precisão e a agilidade das decisões organizacionais. Em síntese, constatou-se que a convergência entre gestão do conhecimento e mineração de dados potencializa os processos decisórios e favorece um planejamento de pessoal mais proativo, fundamentado em evidências e alinhado aos objetivos estratégicos.

Na Marinha do Brasil, a evasão de militares representa um ponto de atenção relevante, pois influencia o planejamento de efetivos, a continuidade dos serviços, o retorno sobre o investimento em formação e capacitação, bem como a preservação do conhecimento organizacional. Compreender os fatores que levam à evasão e desenvolver mecanismos preditivos que permitam antecipá-la configura-se, portanto, como uma estratégia essencial para a gestão de pessoal. Nesse cenário, destaca-se a Diretoria do Pessoal da Marinha, responsável pela administração de uma extensa base histórica de dados de pessoal militar. Esse repositório pode oferecer subsídios relevantes para decisões mais precisas e assertivas.

Diante desse panorama, a pesquisa analisou, sob a perspectiva da gestão do conhecimento e da mineração de dados, o potencial de um modelo preditivo para apoiar a identificação de perfis de militares com maior propensão à evasão. O uso de algoritmos de classificação aplicados a dados históricos mostrou-se capaz de revelar padrões relevantes, permitindo uma atuação proativa na formulação de estratégias de retenção, no aprimoramento do planejamento de carreira e na consolidação de uma cultura organizacional orientada pela tomada de decisão baseada em evidências.

Foram avaliados diferentes métodos de classificação, considerando critérios como acurácia, interpretabilidade, facilidade de implementação e adequação ao

contexto da gestão de pessoal. Entre eles, a árvore de decisão destacou-se por oferecer um equilíbrio adequado entre desempenho preditivo e clareza na apresentação dos resultados, permitindo identificar variáveis-chave e gerar regras compreensíveis para gestores. Essa característica mostrou-se fundamental para garantir que os resultados pudessem ser diretamente aplicados nas decisões estratégicas, sem necessidade de conhecimentos técnicos avançados em ciência de dados.

Assim, a construção de um modelo preditivo baseado em árvore de decisão demonstrou ser uma solução eficaz para aliar desempenho preditivo e interpretabilidade, tornando seus resultados claros e aplicáveis no processo decisório. Este trabalho evidencia o potencial de consolidar uma cultura de tomada de decisão fundamentada em evidências, especialmente na gestão de pessoal da Marinha do Brasil, e aponta a possibilidade de adaptação futura do modelo para outros eventos relevantes da carreira militar.

Conclui-se, portanto, que a integração entre a mineração de dados, com ênfase na aplicação de algoritmos de classificação como a árvore de decisão, e a gestão do conhecimento apresenta-se como uma estratégia consistente e alinhada aos desafios contemporâneos, oferecendo recursos para aprimorar o planejamento de pessoal e subsidiar decisões mais assertivas. Considera-se, assim, alcançado o objetivo geral desta pesquisa, que consistiu em analisar o potencial de aplicação dessa abordagem integrada como suporte à tomada de decisão no âmbito da gestão de pessoal, evidenciando sua contribuição efetiva para a modernização, a eficiência institucional e o fortalecimento da capacidade estratégica da MB.

REFERÊNCIAS

ALEJANDRINO, G. C. **Máquinas de vetores de suporte (SVM) e suas aplicações**. São Paulo: Editora Técnica, 2021.

BOSTROM, R. P.; HEINEN, J. S. **MIS problems and failures: a sociotechnical perspective**. *MIS Quarterly*, v. 1, n. 3, p. 17–32, 1977. Disponível em: <https://www.jstor.org/stable/248710>. Acesso em: 20 jul. 2025.

BRASIL (2016). Diretoria-Geral do Pessoal da Marinha. **Normas para o Sistema de Pessoal da Marinha**. 4ªRev, Rio de Janeiro, RJ, 2016. Disponível em: <https://www.dgpm.mb.dgpm/documentos/secretaria/normas/normas.pdf/> . Acesso em: 10 jun. 2025.

CARVALHO, L. F.; PEREIRA, M. B.; CARDOSO, T. R. **Redes neurais artificiais: fundamentos e aplicações**. Rio de Janeiro: Elsevier, 2019.

CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. **Business intelligence and analytics: from big data to big impact**. *MIS Quarterly*, v. 36, n. 4, p. 1165–1188, 2012. Disponível em: <https://www.jstor.org/stable/41703503>. Acesso em: 21 jul. 2025.

DAVENPORT, T. H.; PRUSAK, L. **Working Knowledge: How Organizations Manage What They Know**. Boston: Harvard Business School Press, 1998.

DREYER GALVÃO, A. F.; MARIN, H. F. **Data Mining: Conceitos, Técnicas e Aplicações**. São Paulo: Elsevier, 2009.

FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. **From data mining to knowledge discovery in databases**. *AI Magazine*, v. 17, n. 3, p. 37-54, 1996. Disponível em: <https://ojs.aaai.org/index.php/aimagazine/article/view/1230> Acesso em: 02 mai. 2025.

HAN, Jiawei; PEI, Jian; TONG, Hanghang. **Data mining: concepts and techniques**. [S.l.]: Morgan kaufmann, 2023. Disponível em: https://books.google.com.br/books/about/Data_Mining_Concepts_and_Techniques.html?id=pQws07tdpjoC&redir_esc=y. Acesso em: 05 abr. 2025.

MERCÊS, Marcella Guarnieri. **Predição de evasão de militares do Exército Brasileiro utilizando técnicas de machine learning**. 2023. 155 f. Dissertação (Mestrado em Ciência da Computação) – Universidade de Brasília, Brasília, 2023. Disponível em: <https://educapes.capes.gov.br/handle/capes/952193?mode=full>. Acesso em: 14 jul. 2025.

NONAKA, I.; TAKEUCHI, H. **The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation**. New York: Oxford University Press, 2008.

PROBST, G.; RAUB, S.; ROMHARDT, K. **Managing Knowledge: Building Blocks for Success**. 2. ed. Wiley, 2002.

PROVOST, Foster; FAWCETT, Tom. **Data Science for Business: What you need to know about data mining and data-analytic thinking**. [S.l.]: O'Reilly Media, 2013. Disponível em:
https://www.researchgate.net/publication/256438799_Data_Science_for_Business.
 Acesso em: 22 jun. 2025.

RAIMUNDO, J. N.; CORREIA NETO, J. S. **Motivação e evasão na carreira militar: um estudo de caso no Exército Brasileiro**. *Cadernos da Fucamp*, 19(41), p. 1-13, 2017.
 Disponível em: <https://idonline.emnuvens.com.br/id/article/view/870/1240>
 Acesso em: 14 jun. 2025.

SCHULTZ, Valéria Simas. **Mineração de dados aplicada à gestão de pessoas: revisão e análise bibliométrica**. *Revista Gestão do Conhecimento e Tecnologia da Informação*, v. 4, n. 2, p. 19–24, 2020. Disponível em:
<https://portalrevistas.ucb.br/index.php/rgcti/article/download/13942/11261/63257>.
 Acesso em: 05 mai. 2025.

SILWATTANANUSARN, Tipawan; TUAMSUK, Kulthida. **Data Mining and Its Applications for Knowledge Management**. ArXiv, 2012.

SIUKO, Viivi; MYLLÄRNIEMI, Jussi; HELLSTEN, Pasi. **Towards digital transformation: knowledge management as an enabler in a public sector asset lifecycle**. In: International Joint Conference On Knowledge Discovery, Knowledge Engineering And Knowledge Management – KMIS, 15., 2023, Roma. Disponível em:
https://trepo.tuni.fi/bitstream/handle/10024/207345/Towards_Digital_Transformation.pdf?sequence=-1. Acesso em: 15 jul. 2025.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introduction to data mining**. [S. l.]: Pearson Education India, 2016.

VITO, Luis Gustavo Ronsani; MELO, Thiago Silva; MESQUITA, Alessandro Pimentel; BELDERRAIN, Mischel Carmen Neyra; GARCIA, Solange. **Evasão na carreira das Forças Armadas dos militares do Corpo Técnico – aplicação de Soft OR utilizando SSM**. In: Simpósio Brasileiro de Pesquisa Operacional, 54., 2022, Juiz de Fora. Anais... Campinas: Galoá, 2022. Disponível em:
<https://proceedings.science/sbpo/sbpo-2022/trabalhos/evasao-na-carreira-das-forcas-armadas-dos-militares-do-corpo-tecnico-aplicacao-d?lang=pt-br>. Acesso em: 14 jun. 2025.

WITTEN, I.H.; FRANK, E.; HALL, M.A.; PAL, C.J. **Data Mining: Practical Machine Learning Tools and Techniques**. 4. ed. Burlington, MA: Morgan Kaufmann, 2017.