



ESCOLA DE
NEGÓCIOS



MARCUS VINICIUS DE CAMPOS

**MODELOS DE *MACHINE LEARNING* NA GESTÃO DE PESSOAS POR
COMPETÊNCIAS DE UMA COOPERATIVA DE TRABALHO MÉDICO**

CURITIBA

2023

MARCUS VINICIUS DE CAMPOS

**MODELOS DE *MACHINE LEARNING* NA GESTÃO DE PESSOAS POR
COMPETÊNCIA DE UMA COOPERATIVA DE TRABALHO MÉDICO**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão em Cooperativas. Mestrado Profissional em Gestão de Cooperativas da Pontifícia Universidade Católica do Paraná, como requisito parcial à obtenção do título de mestre em Gestão de Cooperativas.

Orientador: Prof. Dr. Alex Sandro Quadros Weymer.

CURITIBA

2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.
Catalogação na publicação Serviço de Biblioteca e Documentação Pontifícia Universidade Católica do Paraná – PUC-PR

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Edilene de Oliveira dos Santos CRB-9/1636

C198m
2023

Campos, Marcus Vinicius de
Modelos de *machine learning* na gestão de pessoas por competências de uma cooperativa de trabalho médico / Marcus Vinicius de Campos ; orientador: Alex Sandro Quadros Weymer. – 2023
70 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2023.
Bibliografia: p. 66-70

1. Cooperativas - Administração. 2. Eficiência organizacional. 3. 4. Administração de pessoal. 5. Aprendizado do computador. I. Weymer, Alex Sandro Quadros. II. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Gestão de Cooperativas. III. Título

CDD. 20. ed. – 658.047

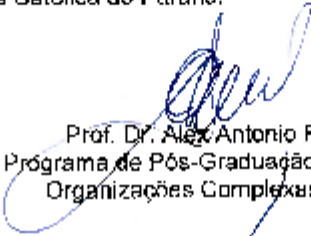
TERMO DE APROVAÇÃO

**MODELOS DE MACHINE LEARNING NA GESTÃO DE PESSOAS POR
COMPETÊNCIA DE UMA COOPERATIVA DE TRABALHO MÉDICO**

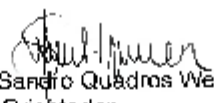
Por

Marcus Vinícius de Campos

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Gestão de Cooperativas e Organizações Complexas e Plurais, área de concentração em Gestão de Cooperativas, da Escola de Negócios da Pontifícia Universidade Católica do Paraná.



Prof. Dr. Alex Antonio Ferraresi
Coordenador do Programa de Pós-Graduação em Gestão de Cooperativas e
Organizações Complexas e Plurais



Prof. Dr. Alex Sandro Quadros Weymer
Orientador



Profa. Dra. Leila Andressa Dissanha
Examinadora



Profa. Dra. Simone Cristina Ramos
Examinadora

AGRADECIMENTOS

A Deus, por estar sempre presente em minha vida, me dando sabedoria a cada decisão tomada.

A minha esposa Michelle, meus filhos Felipe e Lívia pela compreensão e suporte nas horas dedicadas ao mestrado.

Aos meus pais, Jovis e Neuza, por sempre se preocuparem em ter como prioridade a nossa educação, minha e da minha irmã Alessandra.

Ao time da área de tecnologia da Unimed Londrina, por dar continuidade as atividades durante minhas ausências.

Ao Ricardo Cassio Pinelli, Superintendente Administrativo Financeiro da Unimed Londrina por permitir que me ausentasse da cooperativa para dedicar-me ao mestrado.

A Lucia Maria Almeida Baum, Gerente de Desenvolvimento Humano e sua equipe por disponibilizarem os dados utilizados na pesquisa e pelo compromisso de estarem sempre à disposição quando necessário.

Ao meu orientador, professor Dr. Alex Sandro Quadros Weymer, pelo incentivo e disponibilidade, mas principalmente pela compreensão que todos temos limitações.

RESUMO

Objetivo: O objetivo deste estudo foi identificar de que maneira a utilização de modelos de Aprendizagem de máquina pode contribuir na gestão de pessoas da Unimed Londrina a tomar decisões, especialmente no que diz respeito aos treinamentos ofertados pela cooperativa aos seus empregados. **Contexto:** A relevância do tema se apoia no potencial do uso de *Machine Learning* (ML) no ambiente de trabalho, à medida que proporcionam dados sistematizados capazes de reduzir ou otimizar o esforço humano em tarefas que podem ser realizadas por máquinas. **Fundamentação teórica:** Como principal pressuposto, assume-se que a literatura sobre *Machine Learning* (ML) apresenta uma ampla variedade de sugestões sobre como técnicas de análise de dados poderiam ser implementadas para finalidades específicas de Recursos Humanos. Nessa linha de pensamento, a base teórica deste estudo contempla a seguinte estrutura: inteligência artificial, aprendizagem de máquinas, gestão por competências com foco na avaliação, treinamento e carreira. **Método:** A estratégia de pesquisa do presente trabalho caracteriza-se como um estudo de caso, no qual o fenômeno estudado é a utilização de modelos de *Machine Learning* na gestão de pessoas de uma cooperativa de trabalho médico. Os dados foram coletados, considerando um intervalo de tempo entre 2017 e 2021, a partir do registro das avaliações de competências realizadas, bem como a quantidade de avaliadores, histórico de cargos, absenteísmo, tempo de treinamento, tempo de casa e idade, e que foram utilizados para identificar os atributos que tiveram influência sobre o resultado final das avaliações. **Resultados:** A partir da análise realizada, foi possível constatar que a utilização de modelos de *Machine Learning* apresentou acurácia elevada, chegando a 87% de acerto de um modelo específico. Assim, é possível afirmar que a utilização destes modelos na Gestão de Recursos Humanos para avaliação de competências pode ser benéfica, pois tais modelos podem direcionar as atividades realizadas. Além disso, os modelos Sintéticos e Analíticos demonstraram ser eficazes ao identificar a relevância dos atributos presentes no modelo, o que pode ser utilizado pela Gestão de Recursos Humanos para direcionar as ações necessárias para corrigir as lacunas existentes em um determinado grupo de empregados. **Implicações Gerenciais:** Os resultados obtidos proporcionam uma avaliação multinível, sendo identificadas a acurácia em relação a utilização de *Machine Learning* na Gestão de Recursos Humanos na tomada de decisões relacionadas a avaliação de competências realizada pela Unimed Londrina.

Palavras-chave: Cooperativas. *Machine Learning*. Gestão de Pessoas por Competência.

ABSTRACT

Objective: The objective of this study was to identify how the use of Machine Learning models can contribute to the personnel management of Unimed Londrina in making decisions, especially regarding the training offered by the cooperative to its employees. **Brief context:** The relevance of this topic is based on the potential of using Machine Learning (ML) in the workplace, as it provides systematic data that can reduce or optimize human effort in tasks that can be done by machines. **Theoretical foundation:** As the main assumption, it is assumed that the literature on Machine Learning (ML) presents a wide variety of suggestions on how specific techniques could be implemented for specific Human Resources purposes. In this line of thought, the theoretical basis of this study includes the following structure: Artificial Intelligence, Machine Learning, Competency Management with a focus on assessment, training and career. **Method:** The research strategy of this work is characterized as a case study, in which the studied phenomenon is the use of Machine Learning models in the management of people of a medical work cooperative. The data was collected, considering a time frame between 2017 and 2021, from the record of the competence evaluations carried out by the employees, as well as the amount of evaluators, job history, absenteeism, training time, home time and age, and which were used to identify the attributes that had an influence on the final result of the evaluations. **Results:** From the analysis carried out, it was possible to ascertain that the use of Machine Learning models was highly assertive, reaching 87% accuracy of a specific model. Thus, it can be stated that the use of these models in Human Resource Management for competence evaluation can be beneficial, as these models can guide the activities carried out. In addition, the Synthetic and Analytical models have proven to be effective in identifying the relevance of the attributes present in the model, which can be used by Human Resource Management to direct the necessary actions to correct the gaps existing in a given group of employees. **Managerial Implications:** The results obtained provide a multi-level assessment, identifying the accuracy in the use of Machine Learning in Human Resource Management in decisions related to the evaluation of skills carried out by Unimed Londrina.

Keywords: Cooperatives. Machine Learning. Competency-based People Management.

SUMÁRIO

1 INTRODUÇÃO	8
1.1 APRESENTAÇÃO DO PROBLEMA DE PESQUISA	12
1.2 OBJETIVOS DA PESQUISA.....	12
1.2.1 Objetivos geral e específicos.....	12
1.3 JUSTIFICATIVAS TEÓRICA E PRÁTICA.....	12
1.4 ESTRUTURA DA DISSERTAÇÃO	14
2 REFERÊNCIAL TEÓRICO	15
2.1 <i>MACHINE LEARNING</i>	15
2.2 GESTÃO DE PESSOAS POR COMPETÊNCIAS	18
2.3 ALINHANDO <i>MACHINE LEARNING</i> E GESTÃO DE PESSOAS POR COMPETÊNCIAS ..	22
3 PROCEDIMENTOS METODOLÓGICOS	25
3.1 ESTRATÉGIA DA PESQUISA	26
3.2 COLETA E ANÁLISE DE DADOS	26
3.3 DEFINIÇÕES CONSTITUTIVAS E DEFINIÇÕES OPERACIONAIS	26
3.3.1 <i>Machine Learning</i>	26
3.3.2 Gestão por Competências	35
4 ANÁLISE DE DADOS.....	36
4.1 CONTEXTO DE ANÁLISE.....	36
4.2 ORIGEM DOS DADOS	36
4.3 ESTRUTURAÇÃO DA BASE DE DADOS.....	36
4.3.1 Estrutura Dados Sintéticos	39
4.3.2 Estrutura Dados Analíticos.....	41
4.4 MODELOS DE ANALISE DA BASE DE DADOS	42
4.4.1 Matriz de Confusão	44
4.4.2 Dados Sintéticos Sem Matrícula	46
4.4.3 Dados Sintéticos Com Matrícula.....	48
4.4.4 Dados Analíticos Sem Matrícula.....	51
4.4.5 Dados Analíticos Com Matrícula	54
5 DISCUSSÃO DOS RESULTADOS.....	60
6 CONSIDERAÇÕES FINAIS	63
7 IMPACTOS GERENCIAIS.....	65
REFERÊNCIAS	67

1 INTRODUÇÃO

Para uma melhor compreensão a respeito dos conceitos de cooperativa, é importante o entendimento a respeito da definição de alguns elementos, tal como a Sociedade de Pessoas, que diz respeito a atividade econômica estar centralizada na pessoa do cooperado. Diferentemente das sociedades de capital, as cooperativas têm em sua natureza a “sociedade de uso”, ou seja, tem como propósito conectar o cooperado ao mercado, assim se destacando pela ausência de finalidade de lucro.

A Cooperativa tem o propósito de conectar o cooperado ao mercado. Esta instituição não tem fins lucrativos, o que a caracteriza.

A razão da existência da cooperativa é a prestação de serviços ao sócio cooperado, o que reafirma que a cooperativa é a ponte entre o cooperado e o mercado, tendo como papel prover o crescimento econômico e status social de seus cooperados, uma vez que estes são a sua principal atividade econômica.

Diferentemente de outras sociedades, as cooperativas recebem a atribuição de “natureza sui generis”, uma vez que não se assemelham com outras empresas, pois, ao contrário do restante das empresas, cujo maior interesse é o lucro, a cooperativa tem como principal tarefa a promoção dos interesses de seus sócios, os cooperados.

As cooperativas se diferem das empresas capitalistas por possuírem características peculiares a esse tipo de sociedade, como, por exemplo, a Adesão Voluntária, fundamentada no princípio cooperativista da Aliança Cooperativa Internacional da Adesão Livre Voluntária, que dá direito a qualquer pessoa que atenda às regras da cooperativa de se associar ou sair da sociedade a qualquer momento. Isto significa que o sócio pode, a qualquer tempo, romper a relação jurídica com a cooperativa.

A exemplo de um melhor entendimento sobre cooperativismo e seus princípios, é importante trazer a luz o princípio da identidade, onde a sociedade cooperativa tem caráter instrumental ou auxiliar, pois o seu fim é amparar e melhorar a situação econômica dos cooperados-clientes, mediante os serviços que lhes presta (Franke 1978).

Nas cooperativas, o fim que o empreendimento visa identifica-se com o da clientela-associada, ou seja, de seus membros. Assim, as relações entre cliente e empreendimento desenvolvem-se de acordo com o “princípio da identidade”, o que significa que o fim da cooperativa identifica-se com o de sua clientela, funcionando a sociedade como um instrumento para satisfazer as necessidades dos cooperados, sejam elas sociais ou econômicas.

Diferentemente das organizações empresariais, as sociedades cooperativas não

investem na capacitação de seus empregados visando à obtenção de lucros, mas sim à satisfação de seus associados, seja para o crescimento econômico, social ou simplesmente para que eles possam receber um atendimento de qualidade por parte dos empregados da cooperativa, dos quais são sócios.

A popularização da utilização de modelos de *Machine Learning* teve início há mais de uma década, quando as empresas começaram a utilizar algoritmos para tornar processos relacionados a negócios mais eficientes. A utilização de modelos de *Machine Learning* permitiu às empresas automatizar tarefas, processar grandes quantidades de dados e tomar decisões inteligentes, baseadas nos dados históricos.

Desde então, a utilização de modelos de *Machine Learning* evoluiu muito, tornando-se ainda mais acessível para pessoas que não possuem grandes conhecimentos em programação. Grandes corporações como a *Google*, *Microsoft* e *Amazon* desenvolveram ferramentas e serviços que permitem que qualquer pessoa possa ter acesso à essa tecnologia com mais facilidade.

Mesmo com a popularização dessa tecnologia, é importante ressaltar que a utilização de modelos de *Machine Learning Open Source*, ou seja, tecnologia com código aberto, é uma tarefa desafiadora, considerando que além da necessidade de possuir conhecimentos em programação, matemática e lógica, é preciso conhecer os frameworks e bibliotecas disponíveis no mercado.

A utilização dos modelos de *Machine Learning* na Unimed Londrina só foi possível devido a investimentos em tecnologia e capacitação de empregados. Os modelos de *Machine Learning* criados durante este estudo utilizaram tecnologia embarcada no banco de dados *Oracle Enterprise* utilizado pela cooperativa, ou seja, utilizaram uma única tecnologia, não necessitando a busca de frameworks e bibliotecas de mercado. Se faz importante ressaltar que a criação e utilização dos modelos apenas foi possível devido a existência de uma base de dados estruturada, do investimento em tecnologia de banco de dados e pelo fato da cooperativa possuir um time de profissionais seniores com experiência em banco de dados e processos.

Considerando a importância do capital humano nas organizações, o papel estratégico da Gestão de Recursos Humanos tem sua relevância cada vez mais reconhecida. No cenário atual, poucas pessoas sustentariam argumentos contra a importância do treinamento e desenvolvimento dos empregados, uma vez que essa é a principal força para sucesso da organização.

Segundo Ostroff e Bowen (2000), a atenção em relação aos recursos humanos e a adesão as práticas de recursos humanos tecnicamente superiores têm relação direta com empregados

mais produtivos, motivados, satisfeitos e comprometidos, promovendo assim organizações mais eficazes.

A criação de vantagem competitiva por meio de capital humano requer uma atenção cada vez mais cuidadosa em relação as práticas eficazes que melhor alavancam esses ativos. Para Delery e Doty (1996), as práticas de gestão de recursos humanos são um conjunto de políticas e práticas consistentes que são projetadas e implementadas com a finalidade de garantir que o capital humano de uma organização contribua na realização dos objetivos estratégicos da mesma.

Nas cooperativas, o Fundo de Assistência Técnica Educacional e Social (FATES) destina-se à prestação de assistência aos associados, aos seus familiares e, quando previsto nos estatutos, aos empregados da cooperativa. Este fundo deverá ser constituído, no mínimo, de 5% (cinco por cento) das sobras líquidas apuradas no exercício. Vale a pena descartar a utilização desse fundo na Unimed Londrina, cooperativa de trabalho médico onde o presente estudo será realizado, uma vez que o mesmo foi utilizado para criar um fundo de previdência social, atendendo assim uma das finalidades para as quais as sociedades cooperativas se propõem, que é trazer bem-estar financeiro e social aos seus associados.

No que se refere a educação, capacitação e formação, a Unimed Londrina, local onde presente estudo será realizado, investiu 131.464,69 mil reais em cursos superiores no ano de 2020, beneficiando assim cinquenta empregados. Ainda no referido ano, foram investimentos 163.626,65 mil reais na capacitação profissional de cooperados e empregados, totalizando 262 cooperados treinados e 665 empregados treinados (Unimed Londrina, 2021).

O desenvolvimento e o desempenho dos empregados estão relacionados a diferentes variáveis, mas neste estudo são abordadas especificamente três: treinamento, eficácia do treinamento e gestão por competências.

Nas cooperativas, assim como nas organizações convencionais, uma das formas mais clássicas de aprendizagem é o treinamento, sendo as ações referentes a treinamento e desenvolvimento, T&D, consideradas estratégicas, uma vez que são capazes de influenciar ou mesmo de serem influenciadas pelas diretrizes estratégicas das organizações.

Borges-Andrade (2002) definem treinamento como sendo uma ação da organização planejada sistematicamente, tendo como objetivo desenvolver as mais diversas habilidades e estratégias cognitivas, possibilitando assim que o empregado possa melhorar sua eficiência atual ou mesmo futura. Ainda para Hanashiro, Teixeira e Zaccarelli (2007), treinamento é definido como sendo o aprimoramento do desempenho dos empregados de forma que se possa aumentar a produtividade dos recursos que estão a sua disposição durante a realização do

trabalho, desta forma, o treinamento tem como objetivo proporcionar ao treinando o aperfeiçoamento das habilidades existentes e novas habilidades que serão usadas em prol da organização.

Sitzmann e Weinhardt (2017) concluíram em seu estudo que os empregados adquirem mais conhecimentos e habilidades a partir das ações de treinamento, e utilizam essas competências adquiridas em seu ambiente de trabalho. Eles garantem que a organização possa alcançar as competências necessárias para superar seus concorrentes.

Referente a eficácia treinamento, a mesma se encontra na dependência de um processo que envolve desde a concepção dos objetivos instrucionais que estão diretamente alinhados aos objetivos estratégicos, das metodologias e do papel dos indivíduos no que diz respeito ao desenvolvimento e predisposição para a aprendizagem. Considerando que todas as etapas desse processo são importantes, uma das maiores dificuldades relacionadas ao tema é certificar que o investimento realizado com treinamento retornará aos empregados em forma de melhoria no desempenho, e conseqüentemente para organização.

Sitzmann e Weinhardt (2019) propõem uma avaliação de eficácia de treinamento de estrutura multinível, que examina os critérios nos níveis pessoal, interpessoal e macro, utilizando as variáveis: utilização, afeto, desempenho e impacto financeiro do treinamento.

Assim, cada vez mais as organizações têm aumentado seu nível de exigência em relação ao capital humano, principalmente no que diz respeito a entregas, estando esse aumento no nível de exigência relacionado ao ambiente organizacional que vem se tornando cada vez mais competitivo e tecnológico (Bowen e Ostroff, 2004). Tem-se observado uma forte tendência das organizações em destinarem mais recursos na capacitação de seu capital humano, pois as mesmas entendem que esses recursos podem ser o diferencial em relação as organizações concorrentes.

Nesse sentido, a utilização de um modelo de gestão por competência pode auxiliar a organização de maneira que tais recursos sejam melhor aproveitados e sejam realmente considerados um diferencial em relação a concorrência.

A fim de prover o crescimento econômico e status social de seus cooperados, a cooperativa no qual o estudo está sendo realizado investe em programas de treinamento para seus empregados como parte da estratégia da gestão por competência, sendo tais treinamentos utilizados para preencher lacunas referentes a conhecimentos técnicos e comportamentais identificados durante as avaliações de desempenho, assim a área responsável pelos treinamentos pode focar em treinamentos que sejam mais eficientes e que estejam alinhados com o atingimento dos objetivos estratégicos da cooperativa.

1.1 APRESENTAÇÃO DO PROBLEMA DE PESQUISA

A partir das considerações precedentes, pretende-se analisar neste estudo o seguinte problema:

De que maneira a utilização de modelos de *Machine Learning* pode contribuir na gestão de pessoas por competências de uma cooperativa de trabalho médico?

1.2 OBJETIVOS DA PESQUISA

1.2.1 Objetivos geral e específicos

Objetivo geral:

Esta pesquisa tem como objetivo identificar modelos de *Machine Learning* que possam contribuir na gestão de pessoas por competências de uma cooperativa de trabalho médico.

Objetivos específicos:

1. Identificar quais modelos de *Machine Learning* podem aplicados na gestão de pessoas por competência
2. Comparar os resultados obtidos por meio da utilização de modelos de *Machine Learning* na gestão de pessoas por competências com os resultados obtidos com métodos tradicionais
3. Analisar o impacto da utilização de modelos de *Machine Learning* na eficiência dos processos de gestão de pessoas por competências

1.3 JUSTIFICATIVAS TEÓRICA E PRÁTICA

Segundo o relatório emitido pela Organização das Cooperativas do Estado do Paraná, Sistema Ocepar, em dezembro de 2021, o ano encerrou-se com um total de 216 cooperativas registradas no Estado do Paraná, fazendo as cooperativas paranaenses ocuparem cada vez mais

uma posição de destaque no cenário paranaense. Dentre as 216 cooperativas paraenses, 36 delas são do ramo de atividade relacionados a saúde, onde 21 dessas cooperativas fazem parte do Sistema Unimed, sendo uma delas a Unimed Londrina, cooperativa essa no qual este estudo está sendo realizado.

Em 2021, as cooperativas paranaenses atingiram um faturamento histórico e expressivo de 153,7 bilhões de reais, tendo um crescimento de 32,8% em relação ao faturamento do ano anterior. Neste período, o número de associados também aumentou, chegando a 2.741.270, um crescimento de 10,5% em relação ao ano de 2020. Além disso, as cooperativas paranaenses do ramo de saúde contaram com 16.774 associados.

As cooperativas paraenses geram mais de 129 mil empregos diretos em todo estado, onde 7.558 desses empregos são em cooperativas do ramo de saúde, demonstrando desta forma a importância que as cooperativas desempenham para o desenvolvimento econômico e social do estado do Paraná, tornando assim cada vez maior o desafio de manter a competitividade por meio da manutenção de seus indicadores de forma ascendente, tarefa essa que a cada ano se torna mais significativa se considerarmos os números expressivos que já foram atingidos até o momento.

Segundo o último relatório de gestão emitido pelo Sescoop, em 2021 foram realizados 8.078 eventos de formação profissional, contando com a participação de 206.782 pessoas, onde foram investidos 33,2 milhões de reais, voltados exclusivamente para profissionalização, sendo destinados a dirigentes, gestores, empregados e cooperados das cooperativas paranaenses. Assim, o Sescoop/PR atua com educação corporativa, viabilizando treinamentos para os colaboradores, cursos técnicos e de nível superior: graduação, pós-graduação, especialização, mestrado e doutorado.

No que diz respeito a educação cooperativista, é importante trazer à luz a importância que o Serviço Nacional de Aprendizagem do Cooperativismo (SESCOOP) tem nesse processo. Fundado em 1988, o SESCOOP é responsável pelo ensino, formação, profissional, organização e promoção social dos trabalhadores, associados e empregados das cooperativas brasileiras, atuando desta forma no desenvolvimento organizacional das cooperativas paranaenses. Ainda sobre o Serviço Nacional de Aprendizagem do Cooperativismo (SESCOOP), o mesmo recebe uma contribuição de 2,5% (dois e meio por cento) sobre a folha de pagamento dos empregados, sendo essa contribuição realizada por força legal, sendo assim fiscalizada diretamente pela Receita Federal do Brasil.

As descobertas científicas a respeito da eficácia de treinamentos podem ser utilizadas pelas organizações para realizar bons investimentos e tomar decisões políticas mais assertivas.

No entanto, as organizações não devem promover ou investir esforços em treinamentos que não incorporem os princípios de um bom treinamento (Salas et al., 2012).

Na visão prática, o vínculo existente entre eventos de treinamento e características organizacionais se destaca como um dos avanços mais notáveis no que diz respeito a ciência do treinamento, ou seja, o desenvolvimento de modelos de eficácia de treinamento. Contudo, pesquisas e teorias modernas consideram muitos fatores como sendo responsáveis por afetar a eficácia do treinamento, como por exemplo pesquisas sobre motivação de treinamento (Colquitt, LePine, & Noe, 2000) e medição de desempenho (Cannon-Bowers & Salas, 1997).

A automação de tarefas da Gestão de Recursos Humanos busca transferir parcialmente a execução da tarefa humana para máquinas, substituindo a qualificação humana e esforço por processos automatizados que podem ser realizados de forma mais rápida e a um custo reduzido (Zuboff 1985). Dessa maneira, tem-se desenvolvido esforços significativos para aprimorar a automação de tarefas da Gestão de Recursos Humanos. A informação gerada a partir desta automação é valiosa e pode servir como suporte para a tomada de decisões, aumentando a qualidade global das decisões (Zuboff 1985). Por consequência, há também esforços para que se possa aproveitar o potencial das informações desta automação como apoio as decisões da Gestão de Recursos Humanos.

Podemos considerar que as pesquisas científicas na área de gestão de pessoas em sociedades cooperativas são cada vez mais necessárias, especialmente pelo campo promissor para a construção do conhecimento científico (Furlanetto, Weymer & Matos, 2023; Christofoli & Weymer, 2022; Pellin et al., 2021; Goulart, Weymer, & Moreira, 2022). Nesse sentido, principal justificativa para que o estudo seja realizado na Unimed Londrina é trazer a luz um estudo que possa relacionar a eficácia dos treinamentos realizados por seus empregados com o seu crescimento profissional no que diz respeito as competências adquiridas nos treinamentos e aplicadas no ambiente de trabalho.

1.4 ESTRUTURA DA DISSERTAÇÃO

Na primeira seção foi apresentada a introdução, com uma contextualização a respeito do tema proposto. Na sequência, foram apresentados o problema de pesquisa, juntamente com os objetivos gerais, as justificativas teórica e prática, assim como a estrutura da dissertação.

Na segunda seção, são apresentados os elementos conceituais da pesquisa, que são responsáveis por sustentar a demarcação teórica do presente trabalho. Inicialmente é apresentado um contexto a respeito de *Machine Learning*, a seguir, Gestão de Pessoas por

Competência, e por fim, Alinhando *Machine Learning* e Gestão de Pessoas por Competência.

Os procedimentos metodológicos são apresentados na terceira seção, sendo os mesmos responsáveis por nortear a pesquisa e as questões de pesquisa, bem como a delimitação do campo de pesquisa, as variáveis, as definições constitutivas e operacionais. Nesta seção também é apresentada a definição e amostra do presente estudo, assim como as fontes e o tratamento dos dados coletados. Por fim, são apresentados os modelos propostos, implicações gerenciais e considerações finais.

2 REFERÊNCIAL TEÓRICO

Nesta seção do estudo a respeito da utilização de modelos de *Machine Learning*, são apresentados os elementos conceituais da pesquisa, que compõem o referencial teórico do trabalho, na seguinte sequência: *Machine Learning*, Gestão de Pessoas por Competências e Alinhamento entre *Machine Learning* e Gestão de Pessoas.

2.1 MACHINE LEARNING

A Inteligência Artificial (IA) é uma área diversificada e desmembrada da ciência da computação, pois existe heterogeneidade e até certo grau de confusão em termos de definição e classificação adequadas de técnicas de IA (Duch 2007; Kahraman et al. 2010; Wang 2008). Segundo Wang (2008), as definições restritas se concentram nas analogias estruturais ou comportamentais com a inteligência natural, ou seja, uma técnica de inteligência artificial é construída e/ou age como um sistema de linguagem natural. Isso possibilita uma determinação clara das técnicas relevantes, bem como uma separação da ciência da computação em geral, enquanto essa compreensão estreita é cada vez mais chamada de "Inteligência Computacional" (Duch 2007; Kahraman et al. 2010).

Para Wang (2008), definições mais amplas da Inteligência Artificial focam nas capacidades e funções semelhantes aos sistemas inteligentes naturais, uma vez que esta técnica tem como objetivo desempenhar funções ou possuir capacidades, assim como os sistemas inteligentes naturais.

Uma abordagem ampla que define as técnicas de Inteligência Artificial (IA) como instruções processáveis por máquina para resolver tarefas que exigem capacidades cognitivas claras se forem realizadas por humanos, permite considerar técnicas “clássicas” de IA, tais como a representação de conhecimento. No entanto, isso também dificulta a separação clara de

técnicas de IA de outras técnicas de computação, resultando em uma demarcação inadequada da IA da ciência da computação como um todo. Baseada nessa definição, é possível classificar as técnicas de Inteligência Artificial (IA) de acordo com o seu nível cognitivo. Dessa forma, conhecimento, pensamento e linguagem são consideradas funções cognitivas principais que abrangem diferentes categorias de técnicas de IA.

Segundo o autor, a geração, a preservação e o processamento do conhecimento são considerados capacidades cognitivas claras, a partir da compreensão dos fatos relevantes e da consciência. As principais técnicas de Inteligência Artificial relacionadas a esta área podem ser agrupadas em três categorias: descoberta do conhecimento, representação do conhecimento e processamento do conhecimento. A descoberta do conhecimento, também conhecida como *Machine Learning* reconhecimento de padrões ou mineração de dados, é definida como o processo de identificação de informações novas, potencialmente úteis e válidas nos dados (Fayyad et al., 1996).

Uma ampla variedade de técnicas de descoberta de conhecimento, tais como classificação, associação, segmentação e previsão, tornam-se categorias proeminentes (Wu et al., 2008). A representação de conhecimento refere-se à transcrição de proposições relevantes (“conhecimento”) para símbolos formais, de forma a permitir aos computadores o uso desses símbolos para resolver tarefas (Brachman e Levesque, 2004; Davis et al., 1993). Entre as principais formas de representação de conhecimento estão a representação declarativa (representação de fatos) e a representação procedimental (representação de procedimentos para uso de conhecimentos).

Segundo Tanwar et al. (2010), para a representação de conhecimento, existem diversas técnicas, como quadros, redes semânticas e ontologias, sendo estas últimas consideradas exemplos proeminentes de categorias. Para Brachman & Levesque (2004), o processamento de conhecimento, ou seja, a utilização de conhecimentos representados em computadores para gerar novo conhecimento, depende das representações existentes como base e fonte para a racionalização. No que se refere ao raciocínio, as principais categorias são o dedutivo, indutivo e abdutivo (Brachman & Levesque, 2004).

O processamento do conhecimento pode ser entendido como o processamento cognitivo intencional de conhecimentos já existentes para gerar novo conhecimento e solucionar problemas. Desse modo, o processamento do conhecimento representa uma capacidade cognitiva relevante para a Inteligência Artificial, pois conecta-se ao conhecimento de maneiras distintas, ao usar conhecimento existente como entrada e buscar desenvolver novos conhecimentos como saída. Por conseguinte, técnicas relacionadas ao processamento do

conhecimento devem ser consideradas como decorrentes do pensamento, pois técnicas utilizadas para a procura de soluções formam uma segunda categoria importante de técnicas relacionadas ao pensamento. Essencialmente, essas técnicas têm por objetivo formalizar "problemas desafiadores" e solucioná-los mediante uma busca inteligente de um espaço de busca para obter uma solução ótima ou, pelo menos, viável (Kahraman et al. 2010).

O *Machine Learning* caracteriza-se como um campo da computação que dá aos computadores a capacidade de aprender sem que sejam explicitamente programados. Ele permite que os computadores sejam treinados para realizar tarefas específicas usando algoritmos que aprendem com dados, sem que esse processo seja explícito.

O *Machine Learning* refere-se à capacidade de uma máquina de aprender e executar um processo de acordo com um objetivo e etapas definidas para treinamento. Um exemplo genérico poderia ser uma possível recomendação de promoções de empregados, onde o objetivo nesse caso seria recomendar uma promoção (a tarefa) e os passos para chegar lá seriam qualificações e critérios ponderados contra candidatos (a métrica de desempenho). Em um cenário ideal, os dados de treinamento consistiriam em uma lista de empregados recomendados anteriormente para promoção e seus correspondentes registros, contribuindo para a tomada de decisão. A máquina examinaria então estas dezenas, centenas ou milhares de registros e identificaria padrões para recomendar um conjunto de novos dados (o estímulo/experiência) (Mitchell, 2006).

O *Machine Learning* se beneficia das técnicas de aprendizado profundo, que usam redes neurais para aprofundar ainda mais suas análises e realizar tarefas complexas de recomendação. As redes neurais podem avaliar grandes quantidades de dados de forma rápida e ajudar as pessoas a tomarem decisões empresariais com base na análise realizada por meio desses algoritmos (Khan et al., 2020). Elas também podem ser utilizadas para identificar as correlações entre eventos reais e percepções de sentimento a partir de milhares de tweets (do Twitter) e identificar os tópicos que influenciaram a percepção do público em relação a esses eventos (Ridhwan e Hargreaves, 2021).

O *Machine Learning* constitui-se, atualmente, como um dos tópicos de maior relevância, ao permitir que as máquinas possam aprender a partir de dados e prever resultados futuros sem necessidade de programação explícita.

O aprendizado supervisionado é uma das principais ramificações da aprendizagem de máquina, que permite ao modelo prever resultados futuros após o treinamento baseado em dados anteriores. Esta abordagem de aprendizado envolve o uso de pares de entrada/saída ou dados rotulados para treinar um modelo com o objetivo de produzir uma função que seja

suficientemente aproximada para prever saídas para novas entradas quando elas são apresentadas. Os problemas de aprendizado supervisionado podem ser divididos em problemas de regressão e problemas de classificação, onde os problemas de regressão envolvem saídas contínuas, enquanto os problemas de classificação lidam com saídas categorizadas (Sen et al., 2020).

Um algoritmo de classificação é um processo que aprende com dados selecionados para treinamento, em seguida, atribui novas entradas de dados a uma determinada classe e constrói um modelo preditivo baseado em atributos ou características específicas encontradas nos dados. Os modelos de classificação são usados em diferentes tipos de tarefas para prever a classe correta para novas entradas de dados. Um atributo é um parâmetro importante em tarefas de classificação, eles são usados na criação dos modelos preditivos. Os atributos são informações relevantes presentes no conjunto de dados usadas para informar a decisão de classificação.

A classificação é uma função de mineração preditiva, onde o modelo de classificação utiliza dados históricos para prever um alvo específico. Os modelos de classificação podem usar os algoritmos *Naïve Bayes*, *Árvore de Decisão*, *Regressão Logística* ou *Máquina de Vetor de Suporte*. O *Naïve Bayes* é um algoritmo de aprendizado supervisionado de classificação probabilística, ou seja, baseado em estatísticas, ele se baseia no Teorema de *Bayes* para calcular a probabilidade de ocorrências em um evento a partir do conhecimento de outros eventos. *Naïve Bayes* é usado para classificar dados usando uma combinação de diversas probabilidades condicionais, assumindo que cada atributo usado para classificar os dados é independente dos outros, o que torna o algoritmo mais simples.

Considerando o potencial da utilização de modelos de *Machine Learning* para otimização dos dados de determinadas organizações, como suporte estratégico para tomada de decisões, a próxima seção trata da gestão de pessoas por competências e seus respectivos desdobramentos, notadamente, os processos de avaliação, treinamento e carreira.

2.2 GESTÃO DE PESSOAS POR COMPETÊNCIAS

Para Fleury e Fleury (2001), o tema competência entrou na pauta das discussões acadêmicas e empresariais, associado as instâncias de compreensão como, nível da pessoa, das organizações e dos países. Ainda para os autores o conceito de competência é pensado como sendo um conjunto de conhecimentos, habilidades e atitudes que venham a justificar um alto desempenho, acreditando-se assim que os melhores desempenhos estão fundamentados na inteligência e personalidade das pessoas, desta forma, competência é percebida como estoque

de recursos que o indivíduo possui.

Ainda segundo Fleury e Fleury (2004), competência individual ou humana é definida como “um saber agir responsável e reconhecido, que implica mobilizar, integrar, transferir conhecimentos, recursos, habilidades, que agreguem valor econômico à organização e valor social ao indivíduo”. Baseados na mesma perspectiva, Carbone *et al.*, (2009) reforçam que as competências humanas são reveladas quando as pessoas se deparam diante de situações profissionais que servem de ligação entre os atributos individuais e a estratégia definida pela organização.

Considerando que as competências são reveladas quando as pessoas estão diante de situações profissionais com as quais elas se deparam, podemos considerar que a perspectiva da aplicação de conhecimentos, habilidade e atitudes no ambiente de trabalho gera um desempenho profissional, sendo esse desempenho demonstrado por meio de comportamentos manifestados nesse ambiente e pelas consequências desses comportamentos no que diz respeito a realizações e resultados (Gilbert, 1978).

Segundo Zangiski, Lima e Costa (2013), o termo competências organizacional é geralmente utilizado em um contexto estratégico, podendo associar-se à obtenção de vantagem competitiva, sendo tais competências essenciais para sobrevivência e diferenciação da organização frente a seus concorrentes. Para Souza e Vasconcelos (2020) enfatizam que é possível que o modelo de gestão de pessoas por competências possa ser mais meritocrático, tendo em vista que a competência é uma maneira de se avaliar o empregado.

Para Hanashiro (2007), as competências organizacionais são o resultado da integração dos recursos, habilidades e processos organizacionais que interagem com as estratégias da organização e com a cultura organizacional, contribuindo assim para a geração de valor e de vantagem competitiva de maneira sustentável às organizações.

Carbone *et al.* (2009) classificam as competências organizacionais como básicas e essenciais para o atingimento das estratégias definidas por uma organização. O autor define as competências organizacionais básicas como sendo os atributos necessários ao funcionamento da organização, entretanto eles não os diferem da concorrência. Para Prahalad e Hamel (1990), as competências organizacionais essenciais são definidas como os atributos da organização responsáveis por proporcionar vantagem competitiva e geração de valor particular, sendo esse percebido pelo cliente e difícil de ser imitado pelas organizações concorrentes.

Segundo Dutra (2001), o alinhamento entre as competências individuais e as competências organizacionais ocorre durante o processo de troca entre a organização e seus empregados, enriquecendo assim as competências de ambas as partes.

O processo de mapeamento de competências da cooperativa é subsidiado por meio da identificação das competências necessárias para execução das atividades de cada função exercida por seus empregados, essas competências podem ser organizacionais ou humanas, sendo elas responsáveis para o atingimento dos objetivos estratégicos da mesma. O procedimento para identificação de lacunas existentes nas competências ocorre durante a avaliação de desempenho realizada pelos empregados, sendo parte do processo da gestão por competências.

O modelo de gestão por competências ganhou notoriedade no final da década de 80 e início da década de 90, período em que foi possível perceber maior disposição por parte das organizações em evidenciar os trabalhadores que mais agregavam valor para que a organização prosperasse. Assim, as organizações passaram a explicitar suas estratégias em relação aos empregados por meio da definição das competências desejadas para eles (Dutra, 2017).

A gestão por competências é um conceito transversal, podendo ser estudado associado a diferentes correntes de pensamento (Weymer & Silva, 2014). Além disso, é considerado pelas organizações como um modelo alternativo aos instrumentos tradicionais, pois, baseando-se no pressuposto de que o domínio de certos recursos é determinante para um desempenho superior de pessoas e organizações, esse modelo de gestão propõe-se a gerenciar as lacunas referentes às competências, reduzindo assim ao máximo a discrepância existente entre as competências necessárias para o alcance dos objetivos da organização e aquelas que já estão disponíveis. Contudo, a redução ou até mesmo a eliminação destas lacunas depende do mapeamento das competências necessárias para o alcance dos objetivos da organização.

Segundo Brandão e Guimarães (2001), a gestão por competências se propõe a direcionar esforços no planejamento, captação, desenvolvimento e avaliação das competências necessárias à execução de seus objetivos nos diferentes níveis da organização, seja de forma individual, grupal ou mesmo organizacional. Contudo, para Carbone (2005), uma vez que a organização tenha sua estratégia definida, torna-se possível a realização do mapeamento de competências, que objetiva identificar as lacunas existentes, ou seja, a diferença entre as competências necessárias para o atingimento da estratégia e as competências já existentes na organização.

Para Carbone *et. Al.* (2008), a Gestão por Competências permite aos líderes que seja possível alinhar os esforços no desenvolvimento e ampliação das habilidades e capacidades de seus empregados, sendo essa uma forma de lidar com competências relacionadas ao próprio indivíduo. Quando aliadas a recursos e processos, dão sustentação as competências organizacionais, podendo assim suprir as necessidades inerentes ao atingimento dos objetivos e metas da organização.

O modelo de Gestão por Competências faz parte da estratégia que a cooperativa tem para que seus empregados atuais e aqueles que possam vir a ser admitidos tenham um critério de qualificação profissional específico de acordo com a função que desempenham em cada momento da sua trajetória de carreira. Por meio da avaliação de desempenho, isso permite aos líderes identificarem as lacunas técnicas e comportamentais que precisam ser desenvolvidas em cada empregado.

Para Mendes (2017), a Gestão por Competência exige para o recrutamento interno do empregado que o mesmo tenha capacidade de entrega, realização e agregação de valor relacionado a suas atividades, estando essas relacionadas à compreensão da missão da organização a qual faz parte e dos objetivos da mesma para com a comunidade, processos e equipes de trabalho inseridas no ambiente político, social e organizacional.

Baseado na perspectiva de Ruzzarin e Simionovschi (2010), a cooperativa adota a postura de olhar primeiramente para os recursos existentes internamente nos momentos de decisão para alavancar novos negócios ou mesmos em promoções para posições de lideranças, analisando dessa forma a situação por meio de uma perspectiva relacionada a competência de seus recursos humanos, ou seja, de seus empregados, fazendo desta forma o uso da avaliação de desempenho realizada pelos empregados.

Para Mendes (2017) a competência técnica é responsável por avaliar “conhecimentos, habilidades e experiências necessárias e desenvolvidas para garantir o nível de *performance* e qualidade esperados e adequados para realização e agregação de valor ao trabalho”, o que significa que os empregados devem ser avaliados de acordo com a formação técnica que possuem.

Os treinamentos de formação que são realizados na cooperativa devem atender as especificidades das funções dos empregados participantes, os mesmos visam preencher as lacunas técnicas e comportamentais existentes, pois por mais que se evite contratar empregados com lacunas de competências técnicas, o que reduz custos com treinamentos, algumas vezes isso não é totalmente possível, muitas vezes em função da especificidade de conhecimentos e habilidades exigidas para a função que será exercida na cooperativa.

O processo de avaliação de desempenho por competências é, de acordo com Snell e Bohlander (2009), o mecanismo utilizado pelas organizações para auxiliar os empregados a entenderem suas funções, objetivos, expectativas e o sucesso de seu desempenho dentro da organização. De acordo com Lucena (1992), este processo contribui para o gerenciamento dos empregados no âmbito organizacional e pode verificar o real desempenho do empregado em seu ambiente de trabalho.

Segundo os autores, a avaliação de desempenho por competências tem sido descrita como uma abordagem para avaliar o quanto os trabalhadores alcançam os objetivos estabelecidos para suas respectivas funções (Kirkpatrick & Locke, 1991). Para John B. Miner (2005), a avaliação de desempenho por competências é uma abordagem que vincula as habilidades e conhecimentos necessários para alcançar resultados específicos. Ela envolve a medição do desempenho do empregado em relação a padrões de desempenho específicos, definidos em termos de competências.

Steven E. Egglund e Kenneth G. Brown (2010) abordam o uso da avaliação de desempenho por competências como meio para identificar o potencial de desenvolvimento dos empregados e avaliar os resultados de seu desempenho. Esta abordagem é considerada extremamente benéfica para a formulação de planos de desenvolvimento individual apropriados a cada função, contribuindo para que os empregados alcancem seus objetivos profissionais.

Para William J. Rothwell (2012), a avaliação de desempenho por competências proporciona às organizações uma abordagem sistemática para avaliar a capacidade de seus empregados em desempenhar seu trabalho de acordo com os padrões estabelecidos. Ele afirma que a avaliação de desempenho por competências fornece às organizações informações importantes sobre os resultados de desempenho de seus empregados e também sobre como melhorar a produtividade e o desempenho no futuro.

O relacionamento dessas dimensões permite que as lideranças tenham o entendimento da situação em que o desempenho do empregado está enquadrado, e que cada situação requer um tipo de atitude específica para lidar com o mesmo e avaliar seu desempenho. Na cooperativa, a avaliação de desempenho tem a função de permitir ao empregado compreender o seu papel no atingimento da estratégia, suas potencialidades e a motivação empregada nas atividades desempenhadas, o que possibilita o planejamento de trajetórias de desenvolvimento específicas para cada um.

2.3 ALINHANDO *MACHINE LEARNING* E GESTÃO DE PESSOAS POR COMPETÊNCIAS

Os Recursos Humanos são reconhecidos como um dos principais ativos de qualquer organização e o gerenciamento bem-sucedido deste ativo é visto como um dever gerencial essencial (Devanna et al., 1984). Para uma gestão eficaz destes, técnicas adequadas se tornam fundamentais, pois são elas que proporcionam aos seres humanos a capacidade de enfrentar com sucesso as exigências de um determinado domínio.

Além das técnicas específicas do domínio, também se faz necessário considerar técnicas "importadas" de outros domínios, como os testes psicométricos da psicologia organizacional (Wolf & Jenkins, 2006), a otimização da pesquisa operacional (Ernst et al., 2004) ou o processamento analítico online do sistema de informação (Burgard & Piazza, 2009). Uma outra área de destaque que oferece técnicas que podem ser aplicadas na área de Recursos Humanos é a Inteligência Artificial (Jantan et al., 2010).

A utilização de modelos de *Machine Learning* tem um grande potencial para melhorar a gestão de pessoas por competências, pois pode ser utilizado para auxiliar na identificação dos melhores candidatos, avaliação e monitoramento do desempenho dos empregados e avaliação mais precisa das competências necessárias para cada cargo. Os algoritmos de *Machine Learning* podem analisar os dados de desempenho de cada empregado e identificar quais competências são necessárias para cada cargo, o que auxilia no desenvolvimento de estratégias de recrutamento mais eficazes. Desta forma, a utilização de modelos de *Machine Learning* oferece aos gestores de pessoas uma grande quantidade de dados que podem ser usados para avaliar, treinar e recrutar os melhores candidatos, o que melhora significativamente a gestão de pessoas por competências.

O Gerenciamento de Recursos Humanos possui importância proeminente como domínio gerencial, sendo definido e categorizado distintamente de acordo com a literatura. Compreender os empregados como principal fonte de eficiência institucional e vantagem competitiva, além do alinhamento sistemático de todas as ações relacionadas ao empregado de acordo com a estratégia de negócios, são características comuns do conceito desde seu surgimento (Devanna et al. 1984; Jackson et al. 2014). Estas tarefas de responsabilidade da Gestão de Recursos Humanos podem cada vez mais ser apoiadas por técnicas inteligentes de duas formas interligadas, que são a automação e a informação (Zuboff 1985).

Atualmente, com a pressão dos negócios aumentando cada vez mais, os gerentes têm compreendido a relevância do uso de Inteligência Artificial no ambiente de trabalho, assim a Inteligência Artificial cada vez mais vem sendo inserida nos sistemas das organizações, inclusive no departamento de Recursos Humanos. Neste sentido, o uso da Inteligência Artificial tem proporcionado a substituição ou redução do esforço humano em tarefas como triagem de candidatos, recrutamento, alinhamento de atividades de recursos humanos e gerenciamento de desempenho, entre outras, para que desta forma as organizações possam monitorar seu desempenho e suas funções diárias de forma mais eficiente (Yawalkar, M. V. V., 2019).

Na última década do século passado, a combinação entre técnicas de representação e processamento de conhecimento gerou expectativas quanto ao desenvolvimento de sistemas

"especialistas" baseados em conhecimento em Recursos Humanos (Inoue, 1993; Lawler e Elliot, 1996). Contudo, a tecnologia de sistema especialista não se mostrou amplamente desenvolvida e implementada, o que impossibilitou o cumprimento das expectativas estabelecidas.

Na área da Inteligência Artificial, a literatura apresenta uma ampla variedade de sugestões sobre como técnicas específicas de Inteligência Artificial poderiam ser implementadas para finalidades específicas de Recursos Humanos. Por exemplo, Chien e Chen (2008) propuseram a aplicação de técnicas de mineração de dados na seleção dos empregados; Giotopoulos et al. (2007) ofereceram a utilização de agentes inteligentes para o desenvolvimento dos empregados; e Kaczmarek et al. (2005) sugeriram o emprego de técnicas de extração de informações na contratação dos empregados.

O setor de Recursos Humanos tem direcionado a sua atividade para uma revolução digital, utilizando diferentes métodos para simplificar os processos, como análise de grandes conjuntos de dados, inteligência artificial e computação em nuvem (Amla & Malhotra, 2017). As organizações estão implementando tecnologias digitais no âmbito de recursos humanos, como *chatbots*, *Machine Learning* e automação de processos robóticos, que auxiliam o processo de recrutamento, triagem, admissão e entrevistas, entre outros. Assim, a inteligência artificial tem um papel fundamental na gestão de recursos humanos, permitindo a execução de tarefas de forma ágil e eficiente.

A automação de tarefas da Gestão de Recursos Humanos busca transferir parcialmente a execução da tarefa humana para máquinas, substituindo a qualificação humana e esforço por processos automatizados que podem ser realizados de forma mais rápida e a um custo reduzido (Zuboff 1985). Dessa maneira, tem-se desenvolvido esforços significativos para aprimorar a automação de tarefas da Gestão de Recursos Humanos. A informação gerada a partir desta automação é valiosa e pode servir como suporte para a tomada de decisões, aumentando a qualidade global das decisões (Zuboff 1985). Por consequência, há também esforços para que se possa aproveitar o potencial das informações desta automação como apoio as decisões da Gestão de Recursos Humanos.

Em relação ao ajuste tarefa-tecnologia, a aplicação bem-sucedida das técnicas de Inteligência Artificial está diretamente relacionada às necessidades e requisitos das tarefas de Gestão de Recursos Humanos. Assim, é possível encontrar potenciais combinações eficazes para a utilização de técnicas de Inteligência Artificial em tarefas de Recursos Humanos. No entanto, esses potenciais cenários ainda não foram devidamente explorados, pois nem toda técnica de Inteligência Artificial pode ser aplicada a tarefas de Recursos Humanos, nem toda

tarefa de Recursos Humanos pode ser solucionada por uma técnica de Inteligência Artificial. Consequentemente, o ajuste entre a técnica e a tarefa deve ser meticulosamente elaborado em nível individual (Strohmeier et al., 2015).

Qualquer técnica de Inteligência Artificial precisa ser mais eficaz (resultados melhorados) e/ou mais eficiente (menos esforço de implementação) do que as técnicas de Recursos Humanos já estabelecidas, pois caso isso não ocorra, a aplicação desse recurso se torna inútil. Dessa forma, oferecer soluções “personalizadas” para Gestão de Recursos Humanos constitui um método proeminente para o melhor aproveitamento dos potenciais que ainda não foram explorados por meio da aplicação de Inteligência Artificial, de tal forma que a mesma gere valor para o Gerenciamento de Recursos Humanos.

Atualmente, dado o envolvimento da Inteligência Artificial no setor de Recursos Humanos, é necessário que cada vez mais os empregados possuam um conjunto de habilidades para poder operar de maneira eficiente esses recursos, porém, o processo de aquisição de conhecimento a respeito das ferramentas de Inteligência Artificial e das tecnologias digitais, em geral, é dificultoso para os empregados (Jain S., 2017).

Implementar um sistema de Inteligência Artificial em um negócio tem implicações significativas na gestão dos seus recursos humanos, uma que essa mudança pode criar um sentimento de medo e incerteza entre os empregados, resultando em um desafio para o setor de Recursos Humanos em encontrar o profissional adequado para trabalhar com essas tecnologias. Uma outra limitação e desafio para os setores de Recursos Humanos é restringir a capacidade de tomar decisões no dia-a-dia, uma vez que a tecnologia tem sido usada para substituir a autoridade e o papel de Recursos Humanos na tomada de decisões dentro de uma organização.

3 PROCEDIMENTOS METODOLÓGICOS

A ciência é definida, de acordo com Stake (2016), como uma compilação de ótimas explicações sobre coisas físicas, biológicas e sociológicas. Ela é, de maneira geral, a explicação para o funcionamento de todas as coisas, desde a química até as culturas.

Se os pesquisadores decidem coletar dados experimentais e não medidas, sua pesquisa é chamada de “qualitativa”, mas, mesmo assim, podem focar ou o individual ou o geral. Se as descobertas forem baseadas principalmente na combinação de muitas observações individuais, o estudo é chamado de “quantitativo”, mas o pesquisador pode optar por focar tanto o particular quanto o geral (Stake, 2016).

Para Silverman (2000), “a pesquisa qualitativa removeu a pesquisa social da ênfase na explicação de causa e efeito e a colocou no caminho da interpretação pessoal. A pesquisa qualitativa é conhecida por sua ênfase no tratamento holístico dos fenômenos”.

O método qualitativo é adequado aos estudos da história, das representações e crenças, das relações, das percepções e opiniões, ou seja, dos produtos das interpretações que os humanos fazem durante suas vidas, da forma como constroem seus artefatos materiais e a si mesmos, sentem e pensam (Minayo, 2008, p.57).

3.1 ESTRATÉGIA DA PESQUISA

A estratégia de pesquisa do presente trabalho caracteriza-se como sendo um estudo de caso, no qual o fenômeno a ser estudado é a utilização de modelos de *Machine Learning* na gestão de pessoas por competências de uma cooperativa de trabalho médico. O público e os instrumentos utilizados são descritos respeitando-se as etapas referentes à coleta e análise de dados.

3.2 COLETA E ANÁLISE DE DADOS

Este estudo está delimitado a uma cooperativa de trabalho médico do Estado do Paraná que contribui com o Serviço Nacional de Aprendizagem do Cooperativismo – SESCOOP/PR.

O estudo foi realizado a partir da coleta de dados secundários dos empregados, referentes a treinamentos realizados, avaliações de desempenho, planos de desenvolvimento individual e valor de investimento em treinamentos.

Foi considerado para o presente estudo um intervalo de tempo entre os anos de 2017 e 2021, com uma amostra de 2823 avaliações de competências realizadas no período proposto, com uma média de 560 respondentes por ano, considerando apenas empregados ativos com mais de seis meses de trabalho no período da avaliação.

3.3 DEFINIÇÕES CONSTITUTIVAS E DEFINIÇÕES OPERACIONAIS

A seguir são apresentadas as definições constitutivas (DC) e as definições operacionais (DO) do presente estudo.

3.3.1 *Machine Learning*

DC: Um dos principais componentes da Inteligência Artificial moderna, e que foi utilizado maciçamente no estudo de caso do presente trabalho, é o *Machine Learning*, sendo este um campo da computação que dá a computadores a capacidade de aprender sem que sejam explicitamente programados, ou seja, ele permite que os computadores sejam treinados para realizar tarefas específicas usando algoritmos que “aprendem” com dados, sem que esse processo seja explícito.

DO: A operacionalização dessa categoria considerou os registros das avaliações de competências realizadas pelos empregados como base, assim como a quantidade de avaliadores, histórico de cargos, absenteísmo, tempo de treinamento, tempo de casa e idade, que serão utilizados para identificar quais os atributos tem influência sobre o resultado final da avaliação do empregado.

Para a análise dos dados utilizados neste trabalho, foram criadas tabelas no banco de dados, a partir de fontes diversificadas para cada tabela, sendo que todas elas foram carregadas com dados referentes ao período entre 2017 e 2021. Cabe ressaltar que os dados dos empregados foram preservados, onde o número de matrícula dos mesmos foi anonimizado antes da importação para a tabela do banco de dados, para que não fosse possível a identificação dos empregados. A única exceção é o código de matrícula do próprio pesquisador, para que ele possa avaliar e validar os dados referentes às suas avaliações de competência.

A tabela "Avaliação de Competências" foi criada com base nos dados obtidos a partir dos sistemas de avaliação de competências utilizados pela Unimed Londrina. Durante o período, houve um processo de substituição do sistema, bem como uma reestruturação das dimensões avaliadas, o que foi considerado e tratado na planilha Excel fornecida pela área de Desenvolvimento Humano para carregamento dos dados em banco de dados.

A Tabela 1 serve como exemplificação da estrutura da tabela "Avaliação de Competências", na qual os dados relacionados às avaliações de competência estão armazenados para análise posterior.

Tabela 1

Tabela de Banco de Dados “Avaliação de Competências”

Atributo	Descrição do Atributo
Código da Avaliação	Código único utilizado como chave primário da tabela
Número da Matrícula	Campo utilizado para armazenar de maneira anonimizada o número de registro do empregado

Ano	Campo utilizado para armazenar o ano em que a avaliação de competência foi realizada
Descrição da Competência	Campo utilizado para armazenar a competência avaliada
Nota	Campo utilizada para armazenar a nota referente a competência avaliada
Escala	Campo utilizada para armazenar a escala da nota referente a competência avaliada

Os atributos Nota e Escala criados a partir dos dados da avaliação de competência estão exemplificados conforme Tabela 2.

Tabela 2

Escala Avaliação de Competência

Escala	Nota
Não atende	1
Atende parcialmente	2
Atende	3
Supera	4

A Tabela 3 é utilizada para ilustrar de modo quantitativo os dados armazenados na tabela "Avaliação de Competências" referentes às avaliações de competência. Ela agrupa o número de participantes da avaliação de desempenho em cada ano, e consequentemente a quantidade de registros dentro desses anos, que representam as perguntas referentes às dimensões que foram respondidas por eles.

Tabela 3

Participação nas Avaliações de Competências

Ano	Participantes	Registros
2017	531	4026
2018	525	4112
2019	481	10485
2020	608	26686
2021	633	56954

A tabela "Absentéismo" foi criada com base nos dados obtidos a partir do sistema de folha de pagamento utilizado pela Unimed Londrina. Os dados foram gerados a partir de consulta em banco de dados e salvos em planilha Excel para anonimização do código de matrícula e posterior importação para a estrutura de banco de dados criada para essa finalidade.

A Tabela 4 serve como exemplificação da estrutura da tabela "Absentéismo", na qual os dados relacionados ao absentéismo estão armazenados para análise posterior.

Tabela 4

Tabela de Banco de Dados “Absentéismo”

Atributo	Descrição do Atributo
Código Absenteísmo	Código único utilizado como chave primário da tabela
Número de Matrícula	Campo utilizado para armazenar o número de registro do empregado de forma anonimizada
Ano	Campo utilizado para armazenar o ano do absentéismo
Índice Absenteísmo	Campo utilizado para armazenar o índice de absentéismo do empregado durante o referido ano
Faixa Absenteísmo	Campo utilizado para armazenar a faixa de absentéismo do empregado durante o referido ano

O atributo “Faixa Absenteísmo” foi criado a partir dos dados do atributo “Índice Absenteísmo”, sendo esse distribuído em faixas conforme exemplificado na Tabela 5.

Tabela 5

Distribuição Faixa Absenteísmo

Critério	Faixa Absenteísmo
Sem absentéismo	0
Menos de 3 %	3
Entre 3 e 6 %	6
Entre 6 e 9 %	9
Entre 9 e 12 %	12
Entre 12 e 15 %	15
Entre 15 e 18 %	18
Entre 18 e 21 %	21
Entre 21 e 24 %	24
Acima de 24 %	30

A Tabela 6 é utilizada para ilustrar de modo quantitativo os dados armazenados na tabela "Absentéismo" referentes aos registros de percentuais de absentéismo dos empregados no período.

Tabela 6

Registro de Percentual Absenteísmo dos Empregados no período

Referência	Empregados
2017	609
2018	659
2019	737
2020	745
2021	821

A tabela "Histórico de Cargos" foi criada com base nos dados obtidos a partir do sistema de folha de pagamento utilizado pela Unimed Londrina. Os dados foram gerados a partir de consultas em banco de dados e salvos em planilha Excel para anonimizar o código de matrícula e para posterior importação para a estrutura de banco de dados criada para essa finalidade.

A Tabela 7 serve como exemplo da estrutura da tabela "Histórico de Cargos", na qual os dados relacionados às promoções estão armazenados para análise posterior.

Tabela 7

Tabela de Banco de Dados “Histórico de Cargos”

Atributo	Descrição do Atributo
Código do Histórico	Código único utilizado como chave primária da tabela
Número de Matricula	Campo utilizado para armazenar o número de registro do empregado de forma anonimizada
Ano	Campo utilizado para armazenar o ano referente a alteração do cargo
Código do Cargo	Campo utilizado para armazenar o código do cargo de maneira anonimizada
Descrição do Cargo	Descrição Anonimizada do Cargo
Mudança de Cargo	Campo utilizado para armazenar se houve ou não mudança de cargo no mês e ano
Último Cargo	Campo utilizado para armazenar se o cargo é o cargo atual do empregado

A Tabela 8 é utilizada para ilustrar de modo quantitativo os dados armazenados na tabela "Histórico de Cargos", referentes aos registros de promoções dos empregados no período.

Tabela 8

Registro de Promoções dos Empregados no período

Referência	Promoções
2017	30
2018	140
2019	56
2020	28
2021	27

A tabela "Treinamentos" foi criada com base nos dados obtidos a partir dos registros referentes aos treinamentos realizados pelos empregados da Unimed Londrina. A importação para o banco de dados foi realizada a partir da planilha Excel fornecida pela área de Desenvolvimento Humano.

A Tabela 9 serve como exemplo da estrutura da tabela "Treinamentos", na qual os dados relacionados aos treinamentos realizados pelos empregados estão armazenados para análise posterior.

Tabela 9

Tabela de Banco de Dados “Treinamentos”

Atributo	Descrição do Atributo
Código Treinamento	Código único utilizado como chave primária da tabela
Número Matricula	Campo utilizado para armazenar o número de registro do empregado de forma anonimizada
Data Inicial	Campo utilizado para armazenar a Data de Início do treinamento
Código Curso	Campo utilizado para armazenar o código do treinamento
Descrição Curso	Campo utilizado para armazenar a descrição do treinamento
Carga Horária	Campo utilizado para armazenar a carga horário do treinamento realizado

A Tabela 10 é utilizada para ilustrar de modo quantitativo os dados armazenados na tabela "Treinamentos" referentes aos registros de horas de treinamento dos empregados no período.

Tabela 10

Registro de Horas de Treinamentos dos Empregados no período

Referência	Horas de Treinamento
2017	13.156
2018	3.981
2019	18.232
2020	20.763
2021	18.097

A tabela "Faixa Etária" foi criada com base nos dados dos registros dos empregados da Unimed Londrina armazenados no sistema de folha de pagamento utilizado pela gestão de Recursos Humanos. A importação para o banco de dados foi realizada a partir de uma planilha Excel fornecida pela área de Recursos Humanos. O atributo "Idade" foi armazenado em faixas de idade a partir da idade do empregado no ano de apuração, isso foi feito para garantir o anonimato do empregado.

A Tabela 11 serve como exemplo da estrutura da tabela "Faixa Etária", na qual os dados relacionados à Faixa Etária dos empregados estão armazenados para análise posterior.

Tabela 11

Tabela de Banco de Dados "Faixa Etária"

Atributo	Descrição do Atributo
Código Faixa Etária	Código único utilizado como chave primária da tabela
Número Matrícula	Campo utilizado para armazenar o número de registro do empregado de forma anonimizada
Ano	Campo utilizado para armazenar o ano de apuração da faixa etária
Situação	Campo utilizado para armazenar a situação do empregado no período, podendo o valor ser Ativo ou Inativo
Idade	Campo utilizado para armazenar a faixa etária a qual o empregado pertencia no ano de apuração

O atributo "Idade" é utilizado para armazenar a faixa etária do empregado no ano de apuração, criado a partir de critérios conforme exemplificado na Tabela 12.

Tabela 12

Distribuição Faixa Etária

Critério	Faixa Etária
Menos de 20 anos	20 anos
Entre 20 e 25 anos	25 anos
Entre 25 e 30 anos	30 anos
Entre 30 e 35 anos	35 anos
Entre 35 e 40 anos	40 anos
Entre 40 e 45 anos	45 anos
Entre 45 e 50 anos	50 anos
Entre 50 e 55 anos	55 anos
Entre 55 e 60 anos	60 anos
Acima de 60 anos	65 anos

A Tabela 13 é utilizada para ilustrar de modo quantitativo os dados armazenados na tabela "Faixa Etária" referentes aos registros de faixa de idade empregados no período.

Tabela 13

Registro de Faixa de Idade dos Empregados no período

Faixa Etária	Empregados
20 anos	36
25 anos	127
30 anos	201
35 anos	237
40 anos	216
45 anos	170
50 anos	101
55 anos	54
60 anos	28
65 anos	11

A tabela "Tempo de Casa" foi criada com base nos dados dos registros dos empregados da Unimed Londrina armazenados no sistema de folha de pagamento utilizado pela gestão de Recursos Humanos. A importação para o banco de dados foi realizada a partir de uma planilha

Excel fornecida pela área de Recursos Humanos. O atributo “Tempo de Casa” foi armazenado em faixas de anos a partir dos anos de cooperação do empregado no ano de apuração, a fim de garantir o anonimato do empregado.

A Tabela 14 serve como exemplo da estrutura da tabela "Tempo de Casa", na qual os dados relacionados ao tempo de casa dos empregados estão armazenados para análise posterior.

Tabela 14

Tabela de Banco de Dados “Tempo de Casa”

Atributo	Descrição do Atributo
Código Tempo de Casa	Código único utilizado como chave primária da tabela
Número Matrícula	Campo utilizado para armazenar o número de registro do empregado de forma anonimizada
Ano	Campo utilizado para armazenar o ano de apuração do tempo de casa
Situação	Campo utilizado para armazenar a situação do empregado no período, podendo o valor ser Ativo ou Inativo
Tempo de Casa	Campo utilizado para armazenar anos de tempo de casa ao qual o empregado pertencia no ano da apuração

O atributo "Tempo de Casa" é utilizado para armazenar os anos de casa do empregado no ano de apuração, criado a partir de critérios conforme exemplificado na Tabela 15.

Tabela 15

Distribuição Tempo de Casa

Critério	Tempo de Casa
Menos de 5 anos	5 anos
Entre 5 e 10 anos	10 anos
Entre 10 e 15 anos	15 anos
Entre 15 e 20 anos	20 anos
Entre 20 e 25 anos	25 anos
Acima de 25 anos	30 anos

A Tabela 15 é utilizada para ilustrar de modo quantitativo os dados armazenados na tabela "Tempo de Casa" referentes aos registros de anos de tempo de casa empregados no período.

Tabela 15

Registro de Anos de Tempo de Casa dos Empregados no período

Tempo de Casa	Quantidade Empregados
5 anos	495
10 anos	226
15 anos	85
20 anos	62
25 anos	43
30 anos	10

3.3.2 Gestão por Competências

DC: Para Fleury e Fleury (2013), competência se define como “um saber agir responsável e reconhecido, que implica mobilizar, integrar, transferir conhecimentos, recursos, habilidades, agregando valor econômico à organização. Todos esses saberes e valor social ao indivíduo”, assim, a gestão por competência tem como foco principal os processos que venham a incentivar as pessoas envolvidas em uma organização a criarem, compartilharem, disponibilizarem e trocarem conhecimentos objetivando o andamento dos negócios da organização de maneira que o conhecimento venha a se tornar um recurso que deve ser gerenciado com foco no melhor desempenho da própria organização.

DO: A operacionalização dessa categoria considerou as avaliações de desempenho por competências realizadas pelos empregados para avaliar a evolução profissional e comportamental dos mesmos dentro do recorte de tempo proposto no estudo, procurando assim identificar a influência dos treinamentos e consequentemente da gestão por competências na melhora do desempenho da organização.

4 ANÁLISE DE DADOS

4.1 CONTEXTO DE ANÁLISE

Anualmente por meio da Gestão de Recursos Humanos a Unimed Londrina aplica para seus empregados uma avaliação de desempenho, avaliação essa que está dividida em competências avaliada de acordo com a função de cada empregado. A avaliação de competência é utilizada pela Gestão de Recursos Humanos da Unimed Londrina para direcionar os treinamentos realizados pelo empregado no ano consequente a avaliação, para que assim no próximo possa ser verificado as variações entre as avaliações realizadas pelo empregado, principalmente no que diz respeito a competência a qual o mesmo recebeu treinamento.

Esse estudo tem como objetivo auxiliar a Gestão de Recursos Humanos da Unimed Londrina a tomar decisões mais assertivas utilizando os resultados dos modelos gerados por meio da utilização de *Machine Learning*, principalmente no que diz respeito aos treinamentos ofertados pela cooperativa aos seus empregados.

4.2 ORIGEM DOS DADOS

Para realizar a análise contemplada nesse estudo foi necessário estruturar os dados oriundos de diferentes fontes de dados em tabelas de banco de dados Oracle, banco de dados esse que possui um recurso de *Machine Learning* que permite a seus desenvolvedores e administradores de banco de dados criarem modelos de *Machine Learning* a partir de dados armazenados no mesmo, podendo esses modelos serem usados para prever resultados, descobrir padrões ou identificar tendências.

O banco de dados Oracle possui de maneira nativa uma variedade de funções de *Machine Learning*, o *Oracle Machine Learning*, que permite a seus usuários criarem, testarem e executarem modelos de *Machine Learning* a partir dos dados armazenados, que é o caso dessa pesquisa.

4.3 ESTRUTURAÇÃO DA BASE DE DADOS

Para que fosse possível fazer uma análise dos dados utilizando-se da tecnologia de *Machine Learning*, foi necessário criar as tabelas “Dados Sintéticos” e “Dados Analíticos”, sendo estas a representação da junção entre os dados existentes nas tabelas “Avaliação de Competências”, “Absentéismo”, “Histórico de Cargos”, “Treinamentos”, “Faixa Etária” e

“Tempo de Casa”. A relação entre os dados dessas tabelas é feita pelos atributos “Ano” e “Número da Matrícula”.

Os dados existentes nas "Dados Sintéticos" e "Dados Analíticos" são os mesmos, a diferença se faz na forma como estão agrupados. Na tabela "Dados Sintéticos", os dados do empregado estão reunidos em um único registro, pois o atributo "Descrição da Competência" foi desconsiderado. Já na tabela "Dados Analíticos", os dados do empregado estão agrupados considerando a "Descrição da Competência", o que significa que, para um único ano, o empregado terá a quantidade de registros de acordo com a quantidade de competências avaliadas naquele ano.

O atributo "Faixa de Respondentes" foi criado a partir dos dados do atributo "Respondentes", sendo esse distribuído em faixas, conforme exemplificado na Tabela 16.

Tabela 16

Distribuição Faixa Respondentes

Critério	Faixa de Respondentes
Menos de 5 respondentes	5
Entre 5 e 10 respondentes	10
Entre 10 e 15 respondentes	15
Entre 15 e 20 respondentes	20
Entre 20 e 25 respondentes	25
Acima de 25 respondentes	30

O atributo "Faixa de Horas de Treinamento" foi criado a partir dos dados do atributo "Horas de Treinamento", sendo distribuído em faixas conforme exemplificado na Tabela 17.

Tabela 17

Distribuição Faixa Horas de Treinamento

Critério	Faixa de Horas de Treinamento
Não realizou treinamento	0
Entre 1 e 30 horas	30
Entre 30 e 60 horas	60
Entre 60 e 90 horas	90
Entre 90 e 120 horas	120
Entre 120 e 150 horas	150
Entre 150 e 180 horas	180
Entre 180 e 210 horas	210

Entre 210 e 240 horas	240
Acima de 240 horas	270

As Tabelas 18 e 19 são utilizadas para ilustrar de maneira simples a diferença entre as tabelas “Dados Sintéticos” e “Dados Analíticos”.

Tabela 18

Agrupamento Tabela “Dados Sintéticos”

Número	Matrícula	Ano	Nota	Escala
	1	2017	3	ATENDE
	1	2018	3	ATENDE
	1	2019	3	ATENDE
	1	2020	3,29	ATENDE
	1	2021	3,31	ATENDE
	4	2017	2	ATENDE PARCIALMENTE
	4	2018	2	ATENDE PARCIALMENTE
	4	2019	2,36	ATENDE PARCIALMENTE
	4	2020	2,57	ATENDE PARCIALMENTE
	4	2021	2,95	ATENDE PARCIALMENTE

Tabela 19

Agrupamento Tabela “Dados Analíticos”

Matrícula	Ano	Competência	Nota	Escala
1	2020	GESTÃO DE PESSOAS	3,39	ATENDE
1	2020	FOCO EM RESULTADOS E MELHORIA CONTÍNUA	3,28	ATENDE
1	2020	FOCO NO CLIENTE	3,22	ATENDE
1	2020	FLEXIBILIDADE E CAPACIDADE DE ADAPTAÇÃO	3,22	ATENDE
1	2020	INTERCOOPERAÇÃO E ATUAÇÃO ESTRATÉGICA	3,56	ATENDE
1	2020	INTELIGÊNCIA EMOCIONAL	3	ATENDE
1	2020	COMUNICAÇÃO	3,33	ATENDE
1	2021	INTERCOOPERAÇÃO E ATUAÇÃO ESTRATÉGICA	3,44	ATENDE
1	2021	GESTÃO DE PESSOAS	3,48	ATENDE
1	2021	FOCO EM RESULTADOS E MELHORIA CONTÍNUA	3,2	ATENDE
1	2021	FOCO NO CLIENTE	3,32	ATENDE
1	2021	INTELIGÊNCIA EMOCIONAL	3,2	ATENDE
1	2021	FLEXIBILIDADE E CAPACIDADE DE ADAPTAÇÃO	3,28	ATENDE

1	2021	COMUNICAÇÃO	3,28	ATENDE
4	2020	COMUNICAÇÃO	3	ATENDE
4	2020	INTELIGÊNCIA EMOCIONAL	2	ATENDE PARCIALMENTE
4	2020	FLEXIBILIDADE E CAPACIDADE DE ADAPTAÇÃO	3	ATENDE
4	2020	INTERCOOPERAÇÃO E ATUAÇÃO SISTÊMICA	2	ATENDE PARCIALMENTE
4	2020	FOCO EM RESULTADOS E MELHORIA CONTÍNUA	2	ATENDE PARCIALMENTE
4	2020	NEGOCIAÇÃO	3	ATENDE
4	2020	FOCO NO CLIENTE	3	ATENDE
4	2021	INTELIGÊNCIA EMOCIONAL	2,83	ATENDE PARCIALMENTE
4	2021	FLEXIBILIDADE E CAPACIDADE DE ADAPTAÇÃO	3	ATENDE
4	2021	NEGOCIAÇÃO	3	ATENDE
4	2021	INTERCOOPERAÇÃO E ATUAÇÃO SISTÊMICA	3	ATENDE
4	2021	COMUNICAÇÃO	3	ATENDE
4	2021	FOCO EM RESULTADOS E MELHORIA CONTÍNUA	2,67	ATENDE PARCIALMENTE
4	2021	FOCO NO CLIENTE	3,17	ATENDE

4.3.1 Estrutura Dados Sintéticos

A Tabela 20 serve como exemplificação da estrutura da tabela de banco de dados “Dados Sintéticos”, na qual os dados foram utilizados na criação dos modelos “Dados Sintéticos com Matricula” e “Dados Sintéticos sem Matricula”.

Tabela 20

Agrupamento Tabela “Dados Sintéticos”

Atributo	Descrição do Atributo
Número Sequência	Código único utilizado como chave primária da tabela
Número de Matricula	Campo utilizado para armazenar o número de registro do empregado
Ano	Campo utilizado para armazenar o ano de apuração dos dados

Respondentes	Campo utilizado para armazenar a quantidade de respondentes da avaliação de competências do empregado no ano de apuração
Faixa Respondentes	Campo utilizado para armazenar a Faixa de Respondentes
Nota	Campo utilizado para armazenar a média das notas recebidas na avaliação de competências do empregado no ano da apuração
Escala	Campo utilizado para armazenar a escala referente a nota recebida durante a avaliação de competências no ano da apuração
Horas Treinamento	Campo utilizado para armazenar as horas de treinamento realizadas pelo empregado no ano da apuração
Faixa Horas Treinamento	Campo utilizado para armazenar a faixa referente as horas de treinamento realizadas pelo empregado no ano da apuração
Absenteísmo	Campo utilizado para armazenar a informação referente percentual de absenteísmo do empregado no ano da apuração
Faixa Absenteísmo	Campo utilizado para armazenar a informação referente Faixa de absenteísmo do empregado no ano da apuração
Faixa Etária	Campo utilizado para armazenar a idade do empregado referente a faixa etária no ano da apuração
Tempo de Casa	Campo utilizado para armazenar o tempo de casa do empregado no ano da apuração
Promoção	Campo utilizado para armazenar a informação referente ao empregado ter recebido ou não uma promoção no ano da apuração

A Tabela 21 é utilizada para ilustrar de forma quantitativa os dados armazenados na tabela "Dados Sintéticos".

Tabela 21

Avaliações Dados Sintéticos

Ano	Avaliações
------------	-------------------

2017	531
2018	525
2019	481
2020	608
2021	633

4.3.2 Estrutura Dados Analíticos

A estrutura da tabela de banco de dados “Dados Analíticos” foi exemplificada na Tabela 22, a partir dos dados utilizados na criação dos modelos “Dados Analíticos com Matrícula” e “Dados Analíticos sem Matrícula”.

Tabela 22

Agrupamento Tabela “Dados Analíticos”

Atributo	Descrição do Atributo
Número Sequência	Código único utilizado como chave primária da tabela
Número de Matrícula	Campo utilizado para armazenar o número de registro do empregado
Ano	Campo utilizado para armazenar o ano de apuração dos dados
Competência	Campo utilizado para armazenar a descrição da competência à qual o empregado foi avaliado no ano de apuração
Respondentes	Campo utilizado para armazenar a quantidade de respondentes da avaliação de competências do empregado no ano de apuração
Faixa Respondentes	Campo utilizado para armazenar a Faixa de Respondentes
Nota	Campo utilizado para armazenar a média das notas recebidas na avaliação de competências do empregado no ano da apuração
Escala	Campo utilizado para armazenar a escala referente a nota recebida durante a avaliação de competências no ano da apuração

Horas Treinamento	Campo utilizado para armazenar as horas de treinamento realizadas pelo empregado no ano da apuração
Faixa Horas Treinamento	Campo utilizado para armazenar a faixa referente as horas de treinamento realizadas pelo empregado no ano da apuração
Absenteísmo	Campo utilizado para armazenar a informação referente percentual de absenteísmo do empregado no ano da apuração
Faixa Absenteísmo	Campo utilizado para armazenar a informação referente Faixa de absenteísmo do empregado no ano da apuração
Faixa Etária	Campo utilizado para armazenar a idade do empregado referente a faixa etária no ano da apuração
Tempo de Casa	Campo utilizado para armazenar o tempo de casa do empregado no ano da apuração
Promoção	Campo utilizado para armazenar a informação referente ao empregado ter recebido ou não uma promoção no ano da apuração

A Tabela 23 é utilizada para ilustrar de forma quantitativa os dados armazenados na tabela "Dados Analíticos".

Tabela 23

Avaliações Dados Analíticos

Ano	Avaliações
2017	2560
2018	2970
2019	2722
2020	4307
2021	4499

4.4 MODELOS DE ANALISE DA BASE DE DADOS

Os modelos de *Machine Learning* são algoritmos que permitem que uma máquina “aprenda” com dados fornecidos, sem que eles sejam explicitamente programados. Estes

modelos são capazes de prever com precisão, identificar padrões e obter *insights* de dados complexos, utilizando técnicas avançadas de processamento de dados e otimização. Os modelos podem ser treinados usando técnicas de aprendizado supervisionado, semi-supervisionado ou não-supervisionado.

A aprendizagem supervisionada, que foi utilizado nesse estudo, é um tipo de *Machine Learning* em que o modelo é treinado usando dados que contêm rótulos ou respostas esperadas. Os modelos aprendem a associar entradas (variáveis independentes) com saídas (variáveis dependentes) usando os dados rotulados. O objetivo geral é prever a saída para um novo conjunto de dados de entrada. Os modelos de aprendizado supervisionado são usados para classificação, regressão e outros tipos de previsão.

A criação dos modelos de utilizados nesse estudo se fez por meio dos dados existentes nas tabelas de banco de dados denominadas “Dados Sintéticos” e “Dados Analíticos”. A partir de dados históricos existentes nessas tabelas se fez possível a criação de modelos para serem utilizados na realização de previsões a partir do mesmo contexto, para isso, os modelos foram criados a partir de amostras dos dados existentes considerando o período entre 2017 e 2020, para que dessa forma, os dados referentes a 2021 fossem utilizados na validação do modelo.

Os dados existentes no período entre 2017 e 2020 foram utilizados para "treinar" o modelo de Machine Learning. Assim, os dados referentes a 2021 foram utilizados para "testar" o mesmo, ou seja, para validar o modelo.

Para criar os modelos, foi necessário informar um atributo "Número Sequência" como chave, devendo o mesmo ser um valor único, que não se repita dentro da amostra de dados usada para treinamento e criação do modelo. Além disso, foi necessário informar um atributo como sendo o "Alvo", sendo ele o resultado que se deseja alcançar com o *Machine Learning* durante a criação dos modelos e aplicação futura para fazer previsões com base nos atributos usados na criação do modelo.

É importante destacar que as tabelas utilizadas neste estudo estão armazenadas em um Banco de Dados *Oracle* na versão *Enterprise* 19c, que permite aos usuários da versão a utilização do pacote *Data Mining* para criação dos modelos sem custo adicional. Assim, todos os modelos de *Machine Learning* utilizados para o estudo fizeram uso do procedimento *Create Model* existente no pacote *Data Mining*.

Vale também destacar que a função de mineração utilizada foi a de Classificação, que é uma função preditiva que se utiliza de dados históricos para prever respostas futuras. Os modelos de classificação podem utilizar as técnicas *Naive Bayes*, *Decision Tree*, *Logistic*

Regression ou *Support Vector Machine*, sendo a *Naive Bayes* a técnica padrão dessa função e que foi a utilizada neste estudo.

4.4.1 Matriz de Confusão

Para analisar os resultados obtidos a partir dos modelos gerados neste estudo, foi utilizada uma Matriz de Confusão, também conhecida como *Confusion Matrix*, que é uma forma de avaliar o desempenho do modelo de classificação utilizado. Esta matriz é uma comparação entre os valores reais e os valores preditivos resultantes do modelo em relação ao alvo, que é o valor desejado.

A matriz de confusão é uma importante ferramenta utilizada para avaliar a performance de um modelo de *Machine Learning*. É uma representação gráfica e numérica que compara os resultados previstos de um modelo de *Machine Learning* com os resultados reais obtidos. É usada para avaliar a precisão de um modelo de classificação.

A matriz de confusão fornece informações importantes sobre a performance do modelo. Por exemplo, a taxa de acerto e a taxa de erro podem ser calculadas a partir dos valores contidos nela. Além disso, ela pode ser usada para medir a precisão e o *recall* do modelo. A precisão é uma medida da capacidade do modelo de prever corretamente os dados, enquanto o *recall* é uma medida da capacidade do modelo de identificar os dados que devem ser previstos corretamente.

A matriz de confusão é útil no *Machine Learning* supervisionado quando se é utilizado um conjunto de dados previamente identificados. Um exemplo de matriz de confusão é demonstrado na imagem abaixo, onde um modelo de classificação binário é utilizado: 0 para resultados negativos e 1 para os positivos.

Tabela 24

Matriz de Confusão

		Valor Preditivo	
		Sim	Não
Real			

	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

A matriz de confusão é representada por duas classes: positiva e negativa. A classe positiva representa o comportamento classificado como anormal, sendo geralmente menos representada do que a classe oposta. Por sua vez, a classe negativa é representada pela normalidade, ou seja, um comportamento normal.

A construção de modelos tem como finalidade prever o que não é considerado como normal, ou seja, detectar anomalias com a maior precisão possível, como, por exemplo, prever possíveis clientes inadimplentes ou mesmo clientes em potencial que podem contrair determinada doença.

A matriz de confusão é dividida em quatro quadrantes: verdadeiros positivos (TP), verdadeiros negativos (TN), falsos positivos (FP) e falsos negativos (FN), como exemplificado na Tabela 25.

Tabela 25

Definição dos Quadrantes da Matriz de Confusão

Quadrantes	Definição
Verdadeiro Positivo (TP)	O quadrante de resultados verdadeiros positivos da Matriz de Confusão é o quadrante superior direito. Esse quadrante contém os resultados verdadeiros positivos, ou seja, aqueles em que o modelo previu corretamente a classe positiva. Ele é usado para calcular a precisão do modelo.
Verdadeiro Negativo (TN)	O quadrante de resultados verdadeiro negativo da Matriz de Confusão é o quadrante localizado na parte inferior esquerda, ele representa o número de vezes que um modelo classificou corretamente como negativo e geralmente é usado para medir a precisão de um modelo de classificação.
Falso Positivo (FP)	O quadrante de resultados falso positivo na Matriz de Confusão é o quadrante superior direito. Ele é o quadrante que registra os casos em que o classificador previu que o resultado era positivo, mas, na realidade, era negativo.
Falso Negativo (FN)	O quadrante de resultados falso negativo da Matriz de Confusão é o quadrante da parte inferior esquerda da matriz, que representa o número de vezes em que o classificador previu incorretamente que um item não pertence a uma determinada classe, quando na verdade ele realmente pertence. Isso significa

que o modelo errou ao rotular um item como pertencente a uma classe, quando na verdade ele não era.

4.4.2 Dados Sintéticos Sem Matrícula

Para criar o modelo “Dados Sintéticos Sem Matrícula”, foi utilizada a amostra de dados existente na tabela “Dados Sintéticos”, considerando-se como atributos o “Código da Sequência”, “Faixa de Respondentes”, “Faixa de Horas de Treinamento”, “Faixa de Absenteísmo”, “Faixa de Idade”, “Faixa de Tempo de Casa” e “Promoção”.

O atributo “Código da Sequência” foi utilizado como chave única para criar o modelo, uma vez que o valor desse atributo não se repete.

O atributo “Atende” foi utilizado como “Alvo” no aprendizado, ou seja, ele consiste no resultado a ser buscado com a utilização do *Machine Learning*, que neste caso se trata de o empregado ter uma nota igual ou acima de três na avaliação de competência no ano de apuração. Para os empregados com nota igual ou acima de três, tem-se o entendimento de que eles têm as habilidades desenvolvidas referentes ao item avaliado. Já para notas inferiores a três, tem-se o entendimento de que existem lacunas a serem preenchidas referentes ao item avaliado.

Após a aplicação do modelo na amostra de dados utilizada para teste, onde foram consideradas as 633 avaliações de competência realizadas em 2021, os seguintes resultados foram obtidos:

Foram encontrados 525 resultados Verdadeiros Positivos, ou seja, resultados em que o modelo conseguiu prever corretamente que o empregado ficaria com nota igual ou superior a três no ano de 2021.

Foram encontrados 108 resultados Falso Positivo, ou seja, resultados em que o modelo prévio previa que os empregados ficariam com uma nota superior ou igual a três no ano de 2021, mas na verdade obtiveram um resultado inferior a três.

Para os quadrantes Falso Negativo e Verdadeiro Negativo não foram encontrados resultados.

A fim de representar o resultado do modelo de maneira gráfica e numérica, foi criada a Tabela 26 utilizando a Matriz de Confusão.

Tabela 26

Matriz de Confusão do modelo “Dados Sintéticos Sem Matrícula”

		Valor Preditivo	
		Sim	Não
Real	Sim	Verdadeiro Positivo (525)	Falso Negativo (0)
	Não	Falso Positivo (108)	Verdadeiro Negativo (0)

O modelo conta com uma precisão geral de aproximadamente 83%, assim tendo 17% de erros de classificação, ou seja, a porcentagem de resultados que foram classificados de forma incorreta.

O modelo apresentou uma precisão de 82,94%, ou seja, esse percentual representa o número de empregados que tinham uma nota igual ou superior a três e que o modelo conseguiu prever a nota corretamente.

O modelo apresentou um percentual de sensibilidade igual a 100 %, ou seja, conseguiu acertar todos os empregados que obtiveram nota igual ou superior a três no ano de 2021.

A Tabela 27 foi criada para apresentar os dados da tabela de observação referente aos resultados obtidos na aplicação do modelo.

Tabela 27

Tabela de Observação do modelo “Dados Sintéticos Sem Matrícula”

Conjunto de Treinamento (Total de Registros = 2145)	
Métricas	Pontuação
Precisão	0,6648 (66,48%)
Erro de Classificação Incorreta	0,3352 (33,52%)
Precisão (Atende)	0,6648 (66,48%)
Recall (Atende)	1 (100%)
Pontuação F1	0,7987 (79,87%)
Conjunto de Testes (Total de Registros = 633)	

Métricas	Pontuação
Precisão	0,8294 (82,94%)
Erro de Classificação Incorreta	0,1706 (17,06%)
Precisão (Atende)	0,8294 (82,94%)
Recall (Atende)	1 (100%)
Pontuação F1	0,9067 (90,67%)

A fim de entender a relevância de cada atributo do modelo "Dados Sintéticos Sem Matrícula", foi possível realizar uma análise dos atributos de maior importância na criação do modelo. Portanto, os atributos foram organizados por grau de importância e foram utilizados durante o processo de aprendizagem de máquina que resultou no modelo.

Para criar o modelo, a aprendizagem de máquina utilizou apenas dois atributos existentes na tabela: a "Faixa de Tempo de Casa" e a "Faixa de Respondentes", variando o grau de importância dentro dos registros da amostra utilizada para testes.

A tabela 28 foi criada para exemplificar a distribuição da importância dos atributos.

Tabela 28

Grau de Importância dos atributos do modelo "Dados Sintéticos Sem Matrícula"

Primeiro	Segundo	Alvo	Real	Avaliações
Tempo de Casa	Respondentes	Sim	Sim	331
Respondentes	Tempo de Casa	Sim	Sim	194
Tempo de Casa	Respondentes	Não	Sim	70
Respondentes	Tempo de Casa	Não	Sim	38

Por meio da Tabela 28, é possível observar que, para os casos em que o empregado obteve nota igual ou superior a três, a aprendizagem de máquina utilizou a "Faixa de Tempo de Casa" como atributo principal na criação do modelo. Já para os casos em que o empregado obteve nota inferior a três, o atributo principal para a aprendizagem de máquina utilizado na criação do modelo foi a "Faixa de Respondentes".

4.4.3 Dados Sintéticos Com Matrícula

Para criar o modelo "Dados Sintéticos Com Matrícula", foi utilizada para treino a amostra de dados existente na tabela "Dados Sintéticos", considerando como atributos o "Código da Sequência", "Número de Matrícula", "Faixa de Respondentes", "Faixa de Horas de Treinamento", "Faixa de Absenteísmo", "Faixa de Idade", "Faixa de Tempo de Casa" e "Promoção".

O atributo "Código da Sequência" foi utilizado como chave única para criar o modelo, uma vez que o valor deste atributo não se repete.

O atributo "Atende" foi utilizado como "Alvo" no aprendizado, ou seja, ele consiste no resultado a ser buscado com a utilização do *Machine Learning*, que neste caso se trata de que o empregado tenha uma nota igual ou acima de três na avaliação de competência no ano de apuração.

Após a aplicação do modelo na amostra de dados utilizada para teste, onde foram consideradas as 663 avaliações de competência realizadas em 2021, os seguintes resultados foram obtidos:

Foram encontrados 525 resultados verdadeiros positivos, ou seja, resultados em que o modelo conseguiu prever corretamente que o empregado ficaria com nota igual ou superior a três no ano de 2021.

Foram encontrados 108 resultados Falsos Positivos, ou seja, resultados em que o modelo prévio previu que os empregados teriam uma nota superior ou igual a três no ano de 2021, mas na verdade eles obtiveram um resultado inferior a três.

Para os quadrantes Falso Positivo e Verdadeiro Negativo não foram encontrados resultados.

Afim de representar o resultado do modelo de maneira gráfica e numérica utilizando a Matriz de Confusão foi criada a Tabela 29.

Tabela 29

Matriz de Confusão do modelo “Dados Sintéticos Com Matrícula”

		Valor Preditivo	
		Sim	Não
Real	Sim	Verdadeiro Positivo (525)	Falso Negativo (0)

	Não	Falso Positivo (108)	Verdadeiro Negativo (0)
--	------------	----------------------	-------------------------

O modelo conta com uma Precisão Geral de aproximadamente 83%, assim temos 17% de Erros de Classificação, ou seja, a porcentagem de resultados que podem ter sido classificados de maneira incorreta.

O modelo apresentou uma precisão de 82,94%, o que significa que, entre todos os empregados com nota igual ou superior a três, o modelo conseguiu prever corretamente a nota deles.

O modelo apresentou um percentual de sensibilidade igual a 100%, ou seja, ele conseguiu acertar todos os empregados que obtiveram nota igual ou superior a três no ano de 2021.

A Tabela 30 foi criada para apresentar os dados da tabela de observação referentes aos resultados apresentados na aplicação do modelo.

Tabela 30

Tabela de Observação do modelo “Dados Sintéticos Com Matrícula”

Conjunto de Treinamento (Total de Registros = 2145)	
Métricas	Pontuação
Precisão	0,6648 (66,48%)
Erro de Classificação Incorreta	0,3352 (33,52%)
Precisão (Atende)	0,6648 (66,48%)
Recall (Atende)	1 (100%)
Pontuação F1	0,7987 (79,87%)
Conjunto de Testes (Total de Registros = 663)	
Métricas	Pontuação
Precisão	0,8294 (82,94%)
Erro de Classificação Incorreta	0,1706 (17,06%)
Precisão (Atende)	0,8294 (82,94%)
Recall (Atende)	1 (100%)
Pontuação F1	0,9067 (90,67%)

A fim de entender a relevância de cada atributo do modelo "Dados Sintéticos com Matrícula", foi possível realizar uma análise dos atributos de maior importância na criação do modelo, ou seja, aqueles organizados por grau de importância que foram utilizados durante o

processo de aprendizagem de máquina e que resultaram no modelo.

Para criar o modelo, a aprendizagem de máquina utilizou apenas dois atributos presentes na tabela: a “Faixa de Tempo de Casa” e a “Faixa de Respondentes”, variando o grau de importância dentro dos registros da amostra utilizada para teste.

Para exemplificar essa distribuição, foi criada a Tabela 31 para listar a importância dos atributos.

Tabela 31

Grau de Importância dos atributos do modelo “Dados Sintéticos Com Matrícula”

Primeiro	Segundo	Alvo	Real	Avaliações
Tempo de Casa	Respondentes	Sim	Sim	331
Respondentes	Tempo de Casa	Sim	Sim	194
Tempo de Casa	Respondentes	Não	Sim	70
Respondentes	Tempo de Casa	Não	Sim	38

Por meio da Tabela 31, é possível observar que, para os casos em que o empregado obteve nota igual ou superior a três, a aprendizagem de máquina utilizou a “Faixa de Tempo de Casa” como atributo principal na criação do modelo. Por outro lado, nos casos onde a nota foi inferior a três, a “Faixa de Respondentes” foi o atributo principal para a aprendizagem de máquina na criação do modelo.

4.4.4 Dados Analíticos Sem Matrícula

Para criar o modelo “Dados Analíticos Sem Matrícula”, foi utilizada a amostra de dados existente na tabela “Dados Analíticos”, considerando como atributos o “Número da Sequência”, o “Código da Sequência”, a “Descrição da Competência”, a “Faixa de Respondentes”, a “Faixa de Horas de Treinamento”, a “Faixa de Absenteísmo”, a “Faixa de Idade”, a “Faixa de Tempo de Casa” e a “Promoção”.

O atributo “Código da Sequência” foi utilizado como chave única para a criação do modelo, uma vez que o seu valor não se repete.

O atributo “Atende” foi utilizado como “Alvo” no aprendizado. Ele consiste no resultado a ser buscado com a utilização do *Machine Learning*, que, neste caso, se trata de o empregado ter um resultado igual ou superior a três em determinada competência da avaliação de competência no ano de apuração.

Após a aplicação do modelo na amostra de dados utilizada para teste, onde foram considerados os 4499 registros distribuídos entre as 633 avaliações de competência realizadas em 2021, os seguintes resultados foram obtidos:

Foram encontrados 3.933 resultados verdadeiros positivos, ou seja, resultados em que o modelo conseguiu prever corretamente que o empregado teria uma nota igual ou superior a três no ano de 2021.

Foram encontrados 566 resultados Falsos Positivos, ou seja, resultados em que o modelo prévio previa que os empregados obteriam uma nota superior ou igual a três no ano de 2021, mas na verdade os mesmos obtiveram um resultado inferior a três.

Para esse modelo não foram encontrados resultados Verdadeiro Negativo, ou seja, resultados em que o modelo conseguiu prever corretamente a classe negativa, nem Falso Negativo, ou seja, resultados em que o modelo previu incorretamente a classe negativa.

A fim de representar o resultado do modelo de maneira gráfica e numérica, utilizou-se a Matriz de Confusão, conforme mostrada na Tabela 32.

Tabela 32

Matriz de Confusão do modelo “Dados Analíticos Sem Matrícula”

		Valor Preditivo	
		Sim	Não
Real	Sim	Verdadeiro Positivo (3933)	Falso Negativo (0)
	Não	Falso Positivo (566)	Verdadeiro Negativo (0)

O modelo conta com uma precisão geral de 87,42%, o que significa que 12,58% dos resultados podem ter sido classificados incorretamente.

O modelo apresentou uma precisão de 87,42%, o que significa que, entre todos os empregados que tinham uma nota igual ou superior a três, o modelo conseguiu prever a nota corretamente para o percentual representado.

O modelo apresentou um percentual de sensibilidade igual a 100%, ou seja, ele conseguiu acertar todos os empregados que obtiveram nota igual ou superior a três no ano de 2021.

A Tabela 33 foi criada para apresentar os dados da tabela de observação referentes aos resultados obtidos na aplicação do modelo.

Tabela 33

Tabela de Observação do modelo “Dados Analíticos Sem Matrícula”

Conjunto de Treinamento (Total de Registros = 12559)	
Métricas	Pontuação
Precisão	0,7629 (76,29%)
Erro de Classificação Incorreta	0,2371 (23,71%)
Precisão (Atende)	0,7629 (76,29%)
Recall (Atende)	1 (100%)
Pontuação F1	0,8655 (86,55%)
Conjunto de Testes (Total de Registros = 4499)	
Métricas	Pontuação
Precisão	0,8742 (87,42%)
Erro de Classificação Incorreta	0,1258 (12,58%)
Precisão (Atende)	0,8742 (87,42%)
Recall (Atende)	1 (100%)
Pontuação F1	0,9329 (93,29%)

A fim de entender a relevância de cada atributo do modelo “Dados Analíticos Sem Matrícula”, foi possível realizar uma análise dos atributos de maior importância na criação do modelo. Ou seja, os atributos organizados por grau de importância que foram utilizados durante o processo de aprendizagem de máquina e que resultaram no modelo.

Para criar o modelo, a aprendizagem de máquina utilizou quatro atributos: a “Descrição da Competência”, a “Faixa de Tempo de Casa”, a “Faixa de Horas de Treinamento” e a “Faixa de Idade”, variando o grau de importância dentro dos registros existentes na amostra utilizada para testes.

A Tabela 34 lista a importância dos atributos para exemplificar a distribuição.

Tabela 34

Grau de Importância dos atributos do modelo “Dados Analíticos Sem Matrícula”

Primeiro	Segundo	Terceiro	Quarto	Alvo	Real	Qtde.
-----------------	----------------	-----------------	---------------	-------------	-------------	--------------

Tempo de Casa	Competência	Treinamento	Idade	Sim	Sim	1581
Idade	Tempo de Casa	Treinamento	Idade	Sim	Sim	607
Competência	Tempo de Casa	Treinamento	Idade	Sim	Sim	603
Idade	Tempo de Casa	Competência	Treinamento	Sim	Sim	479
Competência	Idade	Tempo de Casa	Treinamento	Sim	Sim	205
Idade	Idade	Tempo de Casa	Treinamento	Sim	Sim	179
Tempo de Casa	Idade	Treinamento	Idade	Sim	Sim	136
Idade	Tempo de Casa	Idade	Treinamento	Sim	Sim	56
Idade	Tempo de Casa	Treinamento	Competência	Sim	Sim	21
Idade	Competência	Tempo de Casa	Treinamento	Sim	Sim	14
Treinamento	Tempo de Casa	Competência	Idade	Sim	Sim	9
Treinamento	Competência	Tempo de Casa	Idade	Sim	Sim	6
Idade	Treinamento	Tempo de Casa	Competência	Sim	Sim	6
Tempo de Casa	Idade	Competência	Treinamento	Sim	Sim	6
Idade	Treinamento	Tempo de Casa	Idade	Sim	Sim	4
Treinamento	Idade	Tempo de Casa	Competência	Sim	Sim	4
Idade	Treinamento	Competência	Tempo de Casa	Sim	Sim	4
Treinamento	Tempo de Casa	Idade	Idade	Sim	Sim	3
Tempo de Casa	Competência	Idade	Treinamento	Sim	Sim	3
Idade	Idade	Treinamento	Tempo de Casa	Sim	Sim	3
Treinamento	Competência	Idade	Tempo de Casa	Sim	Sim	2
Idade	Treinamento	Idade	Tempo de Casa	Sim	Sim	1
Treinamento	Idade	Tempo de Casa	Idade	Sim	Sim	1
Tempo de Casa	Competência	Treinamento	Idade	Não	Sim	175
Competência	Tempo de Casa	Treinamento	Idade	Não	Sim	131
Idade	Tempo de Casa	Competência	Treinamento	Não	Sim	74
Idade	Tempo de Casa	Treinamento	Idade	Não	Sim	67
Competência	Idade	Tempo de Casa	Treinamento	Não	Sim	56
Idade	Idade	Tempo de Casa	Treinamento	Não	Sim	27
Tempo de Casa	Idade	Treinamento	Idade	Não	Sim	9
Idade	Competência	Tempo de Casa	Treinamento	Não	Sim	7
Idade	Tempo de Casa	Idade	Treinamento	Não	Sim	5
Tempo de Casa	Idade	Competência	Treinamento	Não	Sim	5
Idade	Tempo de Casa	Treinamento	Competência	Não	Sim	4
Treinamento	Competência	Idade	Tempo de Casa	Não	Sim	2
Treinamento	Idade	Tempo de Casa	Competência	Não	Sim	2
Idade	Treinamento	Idade	Tempo de Casa	Não	Sim	1
Treinamento	Idade	Tempo de Casa	Idade	Não	Sim	1

Por meio da Tabela 34, é possível observar que, mesmo para os casos em que o empregado teve nota igual ou superior a três ou inferior a três, a aprendizagem de máquina, que utilizou várias combinações de atributos, aponta que os atributos “Competência”, “Tempo de Casa” e “Idade” são os mais relevantes para os casos onde a maior quantidade de registros está concentrada.

4.4.5 Dados Analíticos Com Matrícula

Para criar o modelo "Dados Analíticos com Matrícula", foi utilizada a amostra de dados existente na tabela "Dados Analíticos", considerando os seguintes atributos: “Código da Sequência”, “Número de Matrícula”, “Descrição da Competência”, “Faixa de Respondentes”,

“Faixa de Horas de Treinamento”, “Faixa de Absenteísmo”, “Faixa de Idade”, “Faixa de Tempo de Casa” e “Promoção”.

O atributo “Código da Sequência” foi utilizado como chave única para criar o modelo, uma vez que o valor desse atributo não se repete.

O atributo "Atende" foi utilizado como "Alvo" no aprendizado, ou seja, ele consiste no resultado a ser buscado com a utilização de *Machine Learning*, que, neste caso, se trata de que o empregado tenha um resultado igual ou superior a três em determinada competência da avaliação de competências no ano de apuração.

Após a aplicação do segundo modelo na amostra de dados utilizada para teste, onde foram considerados os 4499 registros distribuídos entre as 663 avaliações de competência realizadas em 2021, foram obtidos os seguintes resultados:

Foram encontrados 3840 resultados Verdadeiros Positivos, ou seja, resultados em que o modelo conseguiu prever corretamente que o empregado ficaria com nota igual ou superior a três no ano de 2021.

Foram encontrados 18 resultados Verdadeiro Negativo, ou seja, resultados em que o modelo conseguiu prever corretamente a classe negativa.

Foram encontrados 548 resultados Falsos Positivos, ou seja, resultados em que o modelo prévio previa que os empregados teriam uma nota superior ou igual a três no ano de 2021, mas na verdade obtiveram resultados inferiores a três.

Foram encontrados 93 resultados Falsos Negativos, ou seja, resultados em que o modelo previu incorretamente a classe negativa.

A fim de representar o resultado do modelo de maneira gráfica e numérica, foi criada a Matriz de Confusão. A Tabela 35 foi utilizada para tal.

Tabela 35

Matriz de Confusão do modelo “Dados Analíticos Com Matrícula”

		Valor Preditivo	
		Sim	Não
Real			

	Sim	Verdadeiro Positivo (3840)	Falso Negativo (93)
	Não	Falso Positivo (548)	Verdadeiro Negativo (18)

O modelo conta com uma Precisão Geral de 85,75%, assim temos 14,25% de Erros de Classificação, ou seja, a porcentagem de resultados que podem ter sido classificados de maneira incorreta.

O modelo apresentou uma precisão de 87,51%, ou seja, esse percentual representa o quanto o modelo conseguiu prever corretamente a nota de todos os empregados que tinham uma nota igual ou superior a três.

O Modelo apresentou um percentual de sensibilidade igual a 97,64%, ou seja, o modelo conseguiu acertar todos os empregados que ficaram com nota igual ou superior a três no ano de 2021.

A Tabela 36 foi criada para apresentar os dados da tabela de observação referentes aos resultados obtidos na aplicação do modelo.

Tabela 36

Tabela de Observação do modelo “Dados Analíticos Com Matrícula”

Conjunto de Treinamento (Total de Registros = 12559)	
Métricas	Pontuação
Precisão	0,7708 (77,08%)
Erro de Classificação Incorreta	0,2292 (22,92%)
Precisão (Atende)	0,7766 (77,66%)
Recall (Atende)	0,982 (98,20%)
Pontuação F1	0,8673 (86,73%)
Conjunto de Testes (Total de Registros = 4499)	
Métricas	Pontuação
Precisão	0,8575 (85,75%)
Erro de Classificação Incorreta	0,1425 (14,25%)
Precisão (Atende)	0,8751 (87,51%)
Recall (Atende)	0,9764 (97,64%)
Pontuação F1	0,9230 (92,30%)

A fim de compreender a relevância de cada atributo no modelo “Dados Analíticos com Matrícula”, foi possível realizar uma análise dos atributos de maior importância na criação deste. Assim, os atributos foram organizados de acordo com o seu grau de importância e utilizados no processo de aprendizagem de máquina que resultou no modelo.

Para criar o modelo, a aprendizagem de máquina utilizou cinco atributos: “Descrição da Competência”, “Número de Matrícula”, “Faixa de Tempo de Casa”, “Faixa de Horas de Treinamento” e “Faixa de Idade”. Os níveis de importância desses atributos foram variados dentro dos registros da amostra usada para a realização dos testes.

A Tabela 37 foi criada para exemplificar a distribuição de importância dos atributos.

Tabela 37

Grau de Importância dos atributos do modelo “Dados Analíticos Com Matrícula”

Primeiro	Segundo	Terceiro	Quarto	Quinto	Alvo	Real	Qtde.
Tempo de Casa	Competência	Treinamento	Idade	Matricula	Sim	Sim	624
Matricula	Tempo de Casa	Competência	Treinamento	Idade	Sim	Sim	402
Tempo de Casa	Competência	Matricula	Treinamento	Idade	Sim	Sim	302
Idade	Tempo de Casa	Treinamento	Idade	Matricula	Sim	Sim	245
Competência	Tempo de Casa	Treinamento	Idade	Matricula	Sim	Sim	243
Tempo de Casa	Matricula	Competência	Treinamento	Idade	Sim	Sim	240
Idade	Tempo de Casa	Competência	Treinamento	Matricula	Sim	Sim	197
Idade	Tempo de Casa	Matricula	Treinamento	Idade	Sim	Sim	184
Competência	Tempo de Casa	Matricula	Treinamento	Idade	Sim	Sim	150
Idade	Matricula	Tempo de Casa	Treinamento	Idade	Sim	Sim	120
Competência	Matricula	Tempo de Casa	Treinamento	Idade	Sim	Sim	115
Competência	Idade	Tempo de Casa	Treinamento	Matricula	Sim	Sim	81
Matricula	Competência	Tempo de Casa	Treinamento	Idade	Sim	Sim	77
Idade	Idade	Tempo de Casa	Treinamento	Matricula	Sim	Sim	76
Idade	Tempo de Casa	Competência	Matricula	Treinamento	Sim	Sim	72
Idade	Matricula	Tempo de Casa	Competência	Treinamento	Sim	Sim	71
Idade	Tempo de Casa	Matricula	Competência	Treinamento	Sim	Sim	64
Tempo de Casa	Idade	Treinamento	Idade	Matricula	Sim	Sim	57
Matricula	Idade	Tempo de Casa	Treinamento	Idade	Sim	Sim	55
Competência	Idade	Tempo de Casa	Matricula	Treinamento	Sim	Sim	55
Matricula	Idade	Tempo de Casa	Competência	Treinamento	Sim	Sim	51
Idade	Idade	Tempo de Casa	Matricula	Treinamento	Sim	Sim	46
Matricula	Tempo de Casa	Idade	Treinamento	Idade	Sim	Sim	34
Tempo de Casa	Matricula	Idade	Treinamento	Idade	Sim	Sim	32
Competência	Idade	Matricula	Tempo de Casa	Treinamento	Sim	Sim	32
Idade	Tempo de Casa	Idade	Treinamento	Matricula	Sim	Sim	23
Idade	Idade	Matricula	Tempo de Casa	Treinamento	Sim	Sim	23
Matricula	Idade	Idade	Tempo de Casa	Treinamento	Sim	Sim	21
Tempo de Casa	Idade	Matricula	Treinamento	Idade	Sim	Sim	13
Idade	Matricula	Idade	Tempo de Casa	Treinamento	Sim	Sim	11
Idade	Tempo de Casa	Treinamento	Competência	Matricula	Sim	Sim	11
Idade	Tempo de Casa	Matricula	Idade	Treinamento	Sim	Sim	10
Idade	Tempo de Casa	Idade	Matricula	Treinamento	Sim	Sim	9
Matricula	Competência	Idade	Tempo de Casa	Treinamento	Sim	Sim	8
Matricula	Idade	Tempo de Casa	Treinamento	Competência	Sim	Sim	7

Idade	Competência	Tempo de Casa	Treinamento	Matricula	Sim	Sim	7
Matricula	Idade	Tempo de Casa	Idade	Treinamento	Sim	Sim	7
Treinamento	Idade	Tempo de Casa	Competência	Matricula	Sim	Sim	4
Idade	Competência	Tempo de Casa	Matricula	Treinamento	Sim	Sim	4
Tempo de Casa	Idade	Competência	Matricula	Treinamento	Sim	Sim	4
Idade	Matricula	Tempo de Casa	Idade	Treinamento	Sim	Sim	4
Idade	Treinamento	Tempo de Casa	Competência	Matricula	Sim	Sim	3
Idade	Tempo de Casa	Treinamento	Matricula	Competência	Sim	Sim	3
Treinamento	Tempo de Casa	Competência	Idade	Matricula	Sim	Sim	3
Tempo de Casa	Competência	Idade	Matricula	Treinamento	Sim	Sim	3
Matricula	Idade	Treinamento	Tempo de Casa	Competência	Sim	Sim	3
Treinamento	Matricula	Tempo de Casa	Competência	Idade	Sim	Sim	3
Treinamento	Tempo de Casa	Competência	Matricula	Idade	Sim	Sim	3
Tempo de Casa	Idade	Matricula	Competência	Treinamento	Sim	Sim	2
Treinamento	Competência	Tempo de Casa	Idade	Matricula	Sim	Sim	2
Idade	Treinamento	Tempo de Casa	Idade	Matricula	Sim	Sim	2
Treinamento	Competência	Tempo de Casa	Matricula	Idade	Sim	Sim	2
Matricula	Idade	Treinamento	Competência	Tempo de Casa	Sim	Sim	2
Treinamento	Competência	Matricula	Tempo de Casa	Idade	Sim	Sim	2
Idade	Treinamento	Competência	Tempo de Casa	Matricula	Sim	Sim	2
Matricula	Idade	Idade	Treinamento	Tempo de Casa	Sim	Sim	2
Idade	Competência	Matricula	Tempo de Casa	Treinamento	Sim	Sim	2
Idade	Tempo de Casa	Treinamento	Matricula	Idade	Sim	Sim	1
Idade	Treinamento	Matricula	Tempo de Casa	Idade	Sim	Sim	1
Matricula	Idade	Competência	Tempo de Casa	Treinamento	Sim	Sim	1
Idade	Idade	Treinamento	Tempo de Casa	Matricula	Sim	Sim	1
Treinamento	Tempo de Casa	Idade	Matricula	Idade	Sim	Sim	1
Treinamento	Idade	Tempo de Casa	Idade	Matricula	Sim	Sim	1
Idade	Treinamento	Idade	Tempo de Casa	Matricula	Sim	Sim	1
Treinamento	Tempo de Casa	Idade	Idade	Matricula	Sim	Sim	1
Idade	Treinamento	Tempo de Casa	Matricula	Idade	Sim	Sim	1
Treinamento	Matricula	Tempo de Casa	Idade	Idade	Sim	Sim	1
Matricula	Competência	Idade	Tempo de Casa	Treinamento	Sim	Não	29
Matricula	Idade	Tempo de Casa	Competência	Treinamento	Sim	Não	24
Matricula	Competência	Tempo de Casa	Treinamento	Idade	Sim	Não	18
Matricula	Tempo de Casa	Competência	Treinamento	Idade	Sim	Não	13
Matricula	Idade	Tempo de Casa	Idade	Treinamento	Sim	Não	3
Matricula	Idade	Tempo de Casa	Treinamento	Idade	Sim	Não	2
Matricula	Idade	Idade	Tempo de Casa	Treinamento	Sim	Não	2
Treinamento	Competência	Idade	Tempo de Casa	Matricula	Sim	Não	2
Tempo de Casa	Competência	Treinamento	Idade	Matricula	Não	Sim	84
Competência	Tempo de Casa	Treinamento	Idade	Matricula	Não	Sim	52
Competência	Tempo de Casa	Matricula	Treinamento	Idade	Não	Sim	35
Matricula	Tempo de Casa	Competência	Treinamento	Idade	Não	Sim	35
Tempo de Casa	Competência	Matricula	Treinamento	Idade	Não	Sim	32
Idade	Tempo de Casa	Competência	Treinamento	Matricula	Não	Sim	27
Idade	Tempo de Casa	Treinamento	Idade	Matricula	Não	Sim	26
Competência	Matricula	Tempo de Casa	Treinamento	Idade	Não	Sim	24
Idade	Tempo de Casa	Matricula	Treinamento	Idade	Não	Sim	23
Tempo de Casa	Matricula	Competência	Treinamento	Idade	Não	Sim	22
Competência	Idade	Tempo de Casa	Matricula	Treinamento	Não	Sim	19
Competência	Idade	Tempo de Casa	Treinamento	Matricula	Não	Sim	16
Matricula	Competência	Tempo de Casa	Treinamento	Idade	Não	Sim	15
Idade	Tempo de Casa	Competência	Matricula	Treinamento	Não	Sim	15
Idade	Tempo de Casa	Matricula	Competência	Treinamento	Não	Sim	14

Competência	Idade	Matricula	Tempo de Casa	Treinamento	Não	Sim	11
Idade	Matricula	Tempo de Casa	Treinamento	Idade	Não	Sim	11
Idade	Idade	Tempo de Casa	Treinamento	Matricula	Não	Sim	10
Idade	Idade	Tempo de Casa	Matricula	Treinamento	Não	Sim	9
Idade	Matricula	Tempo de Casa	Competência	Treinamento	Não	Sim	9
Matricula	Idade	Tempo de Casa	Treinamento	Idade	Não	Sim	7
Matricula	Idade	Tempo de Casa	Competência	Treinamento	Não	Sim	6
Idade	Idade	Matricula	Tempo de Casa	Treinamento	Não	Sim	6
Tempo de Casa	Idade	Treinamento	Idade	Matricula	Não	Sim	4
Tempo de Casa	Idade	Competência	Matricula	Treinamento	Não	Sim	3
Matricula	Tempo de Casa	Idade	Treinamento	Idade	Não	Sim	3
Idade	Competência	Tempo de Casa	Treinamento	Matricula	Não	Sim	3
Matricula	Idade	Tempo de Casa	Treinamento	Competência	Não	Sim	3
Tempo de Casa	Matricula	Idade	Treinamento	Idade	Não	Sim	2
Treinamento	Competência	Idade	Tempo de Casa	Matricula	Não	Sim	2
Treinamento	Idade	Tempo de Casa	Competência	Matricula	Não	Sim	2
Idade	Matricula	Tempo de Casa	Idade	Treinamento	Não	Sim	2
Tempo de Casa	Idade	Matricula	Competência	Treinamento	Não	Sim	2
Matricula	Competência	Idade	Tempo de Casa	Treinamento	Não	Sim	2
Idade	Competência	Matricula	Tempo de Casa	Treinamento	Não	Sim	2
Idade	Tempo de Casa	Treinamento	Competência	Matricula	Não	Sim	1
Matricula	Idade	Idade	Tempo de Casa	Treinamento	Não	Sim	1
Idade	Treinamento	Idade	Tempo de Casa	Matricula	Não	Sim	1
Idade	Matricula	Idade	Tempo de Casa	Treinamento	Não	Sim	1
Idade	Tempo de Casa	Idade	Treinamento	Matricula	Não	Sim	1
Idade	Competência	Tempo de Casa	Matricula	Treinamento	Não	Sim	1
Matricula	Idade	Competência	Tempo de Casa	Treinamento	Não	Sim	1
Idade	Tempo de Casa	Idade	Matricula	Treinamento	Não	Sim	1
Idade	Tempo de Casa	Matricula	Idade	Treinamento	Não	Sim	1
Treinamento	Idade	Tempo de Casa	Idade	Matricula	Não	Sim	1
Matricula	Competência	Idade	Tempo de Casa	Treinamento	Não	Não	8
Matricula	Competência	Tempo de Casa	Treinamento	Idade	Não	Não	5
Matricula	Idade	Tempo de Casa	Competência	Treinamento	Não	Não	3
Matricula	Tempo de Casa	Competência	Treinamento	Idade	Não	Não	2

É possível observar, por meio da Tabela 37, que mesmo para os casos onde o empregado teve nota igual ou superior a três ou inferior a três, mesmo a aprendizagem de máquina tendo utilizado várias combinações de atributos, os atributos “Competência”, “Tempo de Casa” e “Idade” são os atributos mais relevantes para os casos onde a maior quantidade de registros está concentrada.

5 DISCUSSÃO DOS RESULTADOS

É importante ressaltarmos que a única diferença na criação dos modelos “Dados Sintéticos Sem Matrícula” e “Dados Sintéticos Com Matrícula” está no atributo “Número da Matrícula”, o que para fins de resultado não teve influência nenhuma no resultado dos modelos, uma vez que todos os resultados apresentados nas Tabelas de Observação desses modelos são idênticos. A utilização desses modelos implica em uma acurácia de aproximadamente 83%, contra aproximadamente 17% de classificações incorretas. Ambos os modelos consideram apenas os atributos “Tempo de Casa” e “Faixa de Respondentes” para classificações dos dados, ou seja, para prever os resultados.

No que diz respeito aos resultados após a aplicação dos modelos aos “Dados Sintéticos Sem Matrícula” e “Dados Sintéticos Com Matrícula” os mesmos são idênticos, ou seja, ambos trazer os mesmos valores tanto em relação aos resultados quanto em relação a importância de seus atributos, dessa maneira se considerarmos de forma isolado o atributo de maior relevância para os resultados verdadeiros positivo, a “Faixa de Tempo de Casa” é o principal atributo a ser considerado, pois entre as 525 ocorrências referentes a essa situação, 331 dessas ocorrências, o que representa 63,05% de todas as ocorrências tem esse atributo como sendo o mais importante para decisão tomada pela máquina. Em segundo e último lugar, vem a “Faixa de Respondentes”, com 194 ocorrências, o que representa 36,95% de todas as ocorrências.

Em relação aos modelos sintéticos, quando consideramos os resultados onde o atributo principal é a “Faixa de Respondentes”, independentemente do resultado ser Verdadeiro Positivo ou Falso Positivo, o atributo não sofre nenhuma variação em relação ao valor da faixa, assim não é possível afirmar que quanto maior o valor da “Faixa de Respondentes” pior será o resultado. Ou seja, através do resultado dos modelos não é possível afirmar que quanto mais empregados avaliarem um outro empregado, pior será o resultado do empregado avaliado.

Em relação aos "Dados Sintéticos Sem Matrícula" e "Dados Sintéticos Com Matrícula", é possível observar que os Atributos “Tempo de Casa” e “Faixa de Respondentes” influenciam diretamente o resultado da avaliação de competências, uma vez que são estes dois atributos que o modelo utiliza para fazer previsões sobre os resultados futuros.

Em relação aos modelos “Dados Analíticos Sem Matrícula” e “Dados Analíticos Com Matrícula”, assim como nos modelos sintéticos, a diferença existente na criação dos mesmos se concentra apenas no atributo “Número da Matrícula”, o que traz uma leve diferença em relação à acurácia dos modelos.

Na aplicação do modelo “Dados Analíticos Sem Matrícula” os resultados ficaram

concentrados apenas nos quadrantes Verdadeiro Positivo e Falso Positivo, por sua vez, quando o modelo “Dados Analíticos Com Matrícula” é utilizado, os dados são distribuídos entre os quatro quadrantes, mas assim como no modelo “Dados Analíticos Sem Matrícula”, a maior concentração de resultados está presente nos quadrantes Verdadeiro Positivo e Verdadeiro Negativo. Para fins de avaliação do resultado final em relação a aplicação desses modelos, ambos tiveram uma de acurácia de aproximadamente 87%, e aproximadamente 13% de classificações incorretas.

Considerando de maneira isolada os resultados verdadeiros positivo do modelo “Dados Analíticos Sem Matrícula”, ou seja, onde o resultado do modelo foi igual ao resultado real, é possível concluir que o atributo “Tempo de Casa” foi o atributo mais relevante, pois entre as 3933 ocorrências que o modelo teve sucesso, o mesmo foi o principal atributo em 1726 dessas ocorrências, o que representa 43,89% de todas as ocorrências. A “Faixa de Idade” vem em segundo lugar, aparecendo como principal atributo em 1374 ocorrências, ou seja, em 34,94% das ocorrências. A “Competência” ocupa a terceira posição, com 808 ocorrências, o que representa 20,54% das ocorrências. O quarto e último atributo é a “Faixa de Treinamento”, com apenas 25 ocorrências, que representa apenas 0,64% das ocorrências. A partir dessa análise, podemos concluir que a “Faixa de Tempo de Casa” e a “Faixa de Idade” tem grande influência quando aplicamos esse modelo, pois juntos representam 78,83% das ocorrências, assim como a “Faixa de Treinamento” não tem influência nos resultados do modelo, uma vez esse atributo aparece como atributo principal em menos de 1% dos casos.

Assim como no modelo anterior, se considerarmos de maneira isolada os resultados verdadeiros positivo do modelo “Dados Analíticos Com Matrícula”, ou seja, onde o resultado do modelo foi igual ao resultado real, é possível concluir que o atributo “Tempo de Casa” novamente foi o atributo mais relevante, pois entre as 3840 ocorrências que o modelo teve sucesso, o mesmo foi o principal atributo em 1279 dessas ocorrências, o que representa 33,26% de todas as ocorrências. A “Faixa de Idade” novamente vem em segundo lugar, aparecendo como principal atributo em 1194 ocorrências, ou seja, em 31,09% das ocorrências. A “Competência” ocupa a terceira posição, com 676 ocorrências, o que representa 20,54% das ocorrências. O quarto atributo é a “Matrícula”, com 670 ocorrências, representando 17,45% das ocorrências. O quinto e último atributo é a “Faixa de Treinamento”, com apenas 23 ocorrências, que representa apenas 0,60% das ocorrências. A partir dessa análise, podemos concluir que a “Faixa de Tempo de Casa” e a “Faixa de Idade novamente” são os atributos de maior influência quando aplicamos esse modelo, pois juntos representam 64,35% das ocorrências, assim como a “Faixa de Treinamento” não tem influência nos resultados do modelo, uma vez esse atributo

aparece como atributo principal em menos de 1% dos casos.

A partir da aplicação dos modelos analíticos, conclui-se que os principais atributos são a “Faixa de Tempo de Casa” e a “Faixa de Idade”, assim como a “Faixa de Horas de Treinamento”, que não tem influência expressiva no resultado dos modelos.

Em relação aos modelos criados, sejam eles sintéticos ou analíticos, após a aplicação do modelo, a “Faixa de Tempo de Casa” tem se mostrado um dos principais atributos. Isso pode indicar que o tempo de cooperativa e os vínculos gerados durante esse período podem influenciar diretamente no resultado da avaliação de competência.

Para fins de utilização pela Gestão de Recursos Humanos, mesmo os modelos sintéticos tendo um percentual de acurácia tão alto quanto o dos modelos analíticos, os modelos analíticos fazem mais sentido, uma vez que esses modelos trazem a competência a qual o empregado foi avaliado, podendo assim direcionar esforços para corrigir uma lacuna específica para um grupo de empregados.

É possível observar que, em todos os modelos criados utilizando *Machine Learning*, a acurácia ficou acima de 80%. Dessa forma, podemos considerar que a utilização desses modelos pela Gestão de Recursos Humanos no que diz respeito à avaliação de competências pode ser benéfica, uma vez que podem direcionar as atividades.

6 CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo identificar modelos de *Machine Learning* que possam contribuir na gestão de pessoas por competências de uma cooperativa de trabalho médico. Para tanto foi adotada uma metodologia quantitativa com um intervalo de tempo entre os anos de 2017 e 2021, com uma amostra de 2823 avaliações de competências realizadas no período proposto, com uma média de 560 respondentes por ano, considerando apenas empregados ativos com mais de seis meses de trabalho no período da avaliação.

Considerando os resultados obtidos a partir da aplicação dos modelos de *Machine Learning* criados para o presente estudo, sejam eles modelos Sintéticos ou Analíticos, ambos podem ser utilizados na gestão de pessoas por competência, principalmente para auxiliar no direcionamento de estratégias da Gestão de Desenvolvimento Humano.

Os modelos de *Machine Learning* podem ser utilizados para oferecer *Insights* no que diz respeito a decisões futuras, uma vez que os modelos de *Machine Learning* utilizados durante esse estudo trabalham com predição, eles olham para situações futuras baseadas nos dados do passado, algo que não é possível de ser feito usando métodos tradicionais.

Os modelos de *Machine Learning* utilizados nesse estudo podem ser utilizados para prever o desempenho dos empregados. Assim, a cooperativa pode utilizar o modelo para prever o desempenho futuro do empregado em uma avaliação de competência baseada em seu histórico anterior. A partir da previsão dos resultados das avaliações de competência, também é possível planejar treinamentos de maneira mais assertiva, a fim de preencher lacunas existentes nas competências dos empregados.

A partir dos resultados obtidos, conclui-se que os modelos de Sintéticos fornecem resultados semelhantes e que os Atributos “Tempo de Casa” e “Faixa de Respondentes” são os principais fatores que contribuem para a classificação correta dos dados, sendo assim, ambos os modelos Sintéticos apresentam uma acurácia de aproximadamente 83% contra 17% de classificações incorretas.

A partir da análise dos modelos Analíticos, conclui-se que a utilização de modelos de *Machine Learning* para avaliação de competências pode ser benéfica para a Gestão de Recursos Humanos, pois os modelos proporcionaram uma acurácia superior a 80%. Os resultados também indicam que os principais atributos para a avaliação de competência são a “Faixa de Tempo de Casa” e a “Faixa de Idade”.

Em suma, os modelos de dados sintéticos e analíticos mostraram-se eficazes para avaliação de competências, podendo ser utilizados de forma eficiente pela Gestão de Recursos

Humanos para direcionar esforços e corrigir lacunas específicas.

Alguns pontos podem ser observados como limitadores do estudo, dentre eles a quantidade de atributos utilizados, pois no que diz respeito a modelos de Machine Learning, quanto maior a quantidade de atributos disponíveis, melhor é possível fazer simulações utilizando-os, a fim de identificar a acurácia do modelo, com ou sem a utilização dos mesmos.

Visualizando a continuidade do estudo, é possível ampliar a quantidade de atributos dos modelos, afim torna-los ainda mais precisos e com uma amplitude maior em relação a sua utilização, como por exemplo para Prever Retenções, onde o modelo é capaz de prever a tendência ou não do empregado deixar a cooperativa. Os modelos também podem vir a ser utilizados na Avaliação de Promoções e Planejamento de Carreiras, baseado na previsão de resultados futuros de desempenho, habilidades e competências dos empregados.

7 IMPACTOS GERENCIAIS

Os resultados apresentados neste trabalho podem colaborar com processos gerenciais relacionados à gestão de pessoas nas sociedades cooperativas, independentemente de serem cooperativas na área de saúde ou não.

A modernização das cooperativas passa necessariamente pelo investimento em tecnologias e na capacitação dos empregados, sendo a realização dessas ações fundamentais para que as mesmas consigam se destacar no mercado competitivo.

O investimento em tecnologia deve ser priorizado para que as cooperativas possam modernizar sua gestão, oferecendo serviços de qualidade aos seus associados. Os avanços tecnológicos possibilitam melhorias na gestão dos processos, com menor custo e maior eficiência.

A capacitação dos empregados é imprescindível para que sejam criadas equipes altamente qualificadas. Atualizar o conhecimento dos profissionais é fundamental para que as cooperativas consigam oferecer serviços de qualidade e se destacar no mercado.

A utilização de modelos de *Machine Learning* na Gestão de Pessoas tem se tornado cada vez mais importante no ambiente corporativo devido aos benefícios que ela pode trazer. A utilização destes modelos permite às empresas se adaptarem mais rapidamente às mudanças no mercado e aproveitarem melhor as oportunidades. Elas também podem ajudar as empresas a melhorarem o desempenho de seus empregados, otimizando os processos de recrutamento e seleção.

A utilização de modelos de *Machine Learning* na Gestão de Pessoas tem se mostrado de grande valia no que diz respeito à otimização de processos de tomada de decisão, pois por meio desses modelos é possível avaliar um grande volume de dados em tempo real e gerar *insights* estratégicos, além de auxiliar na identificação de oportunidades de melhoria.

Um grande ponto favorável relacionado à utilização de modelos de *Machine Learning* na Gestão de Pessoas, assim como em outras gestões, é sua capacidade de fornecer análise preditiva, permitindo a tomada de decisões baseadas em dados, ao invés de se basear somente na intuição. Dessa forma, a tomada de decisões sobre recursos humanos se torna mais assertiva, aumentando a produtividade e gerando melhores resultados para a empresa.

Podemos concluir que a utilização de modelos de *Machine Learning* é extremamente importante para o crescimento e o sucesso das empresas, pois elas possibilitam que as empresas tomem decisões mais assertivas com base em dados e *insights* acurados, otimizando o processo de tomada de decisão e aumentando a produtividade dos empregados. Para que isso se torne

uma realidade, é essencial que as empresas possam se adaptar rapidamente às mudanças do mercado e aproveitar melhor as oportunidades.

REFERÊNCIAS

- Amla, M., & Malhotra, P. M. (2017). Digital transformation in HR. *International Journal of Interdisciplinary and Multidisciplinary Studies (IJIMS)*, 4(3), 536-544.
- Borges-Andrade, J. E. (2002). *Desenvolvimento de medidas em avaliação de treinamento*. (Dissertação de Mestrado). Universidade Federal do Rio Grande do Norte, Rio Grande do Norte, Brasil.
- Bowen, D. E., & Ostroff, C. 2004. Understanding HRM–firm performance linkages: The role of the “strength” of the HRM system. *Academy of Management Review*, 29: 203-221.
- Burgard, M., Piazza, F.: Data warehouse and business intelligence systems in the context of e-HRM. In: Torres-Coronas, T., Arias-Oliva, M. (eds.) *Encyclopedia of HRIS: Challenges in e-HRM*, pp. 223–229. Hershey, PA (2009).
- Brachman, R.J., Levesque, H.J.: *Knowledge Representation and Reasoning*. Morgan Kaufmann, Burlington (2004).
- Brandão, Hugo Pena; Guimarães, Tomás de Aquino. Gestão de competências e gestão de desempenho: tecnologias distintas ou instrumentos de um mesmo constructo? *Revista de Administração de Empresas*, São Paulo, v. 41, n.1, p. 8-15, jan./mar. 2001.
- Cannon-Bowers, J. A., & Salas, E. (1997). A framework for developing team performance measures in training. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 45–62).
- Carbone, Pedro Paulo; Brandão, Hugo Pena; Leite, João Batista Diniz. *Gestão por competências e gestão do conhecimento*. Rio de Janeiro: Fundação Getúlio Vargas, 2005.
- Carbone, P. P. et al. *Gestão por competências e gestão do conhecimento*. São Paulo: FGV, 2008.
- Carbone, P. P.; Brandão, H. P.; Leite, J. B. D. ; Vilhena, R. M. de P.; *Gestão por competências e Gestão do conhecimento*. Rio de Janeiro: Editora FGV, 2009.
- Chien, C.F., Chen, L.F.: Data mining to improve personnel selection and enhance human capital: a case study in high-technology industry. *Expert Syst. Appl.* 34(1), 280–290 (2008).
- Christofoli, V., & Weymer, A. S. Q. (2023). A relação entre autoeficácia e reputação organizacional em organizações cooperativas. *Cadernos EBAPE. BR*, 21, e2022-0015. <https://doi.org/10.1590/1679-395120220015>.
- Christofoli, V., & Weymer, A. S. Q. (2023). The relationship between self-efficacy and organizational reputation in cooperative organizations. *Cadernos EBAPE. BR*, 21, e2022-0015. <https://doi.org/10.1590/1679-395120220015x>.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, 85(3), 678–707.

- De Souza Minayo, M. C. (2008). O desafio do conhecimento. *Pesquisa qualitativa em saúde*, 13.
- Delery, J.E. and Doty, D.H. (1996), "Modes of theorizing in strategic human resource management: tests of universalistic, contingency, and configurational performance predictions", *Academy of Management Journal*, Vol. 39 No. 4, pp. 802-835.
- Devanna, M.A., Fombrun, C.J., Tichy, N.: A framework for strategic human resource management. In: Fombrun, C.J., Tichy, N., Devanna, M.A. (eds.) *Strategic Human Resource Management*, pp. 33–51. Wiley, New York (1984).
- Duch, W.: What is Computational Intelligence and where is it going? Challenges for Computational Intelligence, pp. 1–13. Springer, Berlin (2007).
- Dutra, J. S. *Gestão por Competências*. São Paulo: Gente, 2001.
- Dutra, J. S. *Competências: conceitos, instrumentos e experiências*. 2. ed. São Paulo: Atlas, 2017.
- Eggland, S. E., & Brown, K. G. (2010). Evaluating performance by competencies: An approach to identify employee development potential and performance outcomes. *Journal of Management*, 9(1), 1-5.
- Ernst, A.T., Jiang, H., Krishnamoorthy, M., Sier, D.: Staff scheduling and rostering: a review of applications, methods and models. *Eur. J. Oper. Res.* 153(1), 3–27 (2004).
- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. *AI Mag.* 17(3), 36–53 (1996).
- Fleury, M. T. L. & Fleury, A. (2001). Construindo o conceito de competência. especial. ed. [S.l.]: RAC.
- Fleury, A. C. C.; Fleury, M. T. *Estratégias empresariais e formação de competências*. 3 ed. São Paulo: Atlas, 2004.
- Fleury, M. T. L. *As pessoas na organização*. São Paulo: Atlas, 2013.
- Franke, Walmor. *O Direito das Sociedades Cooperativas*. São Paulo: Saraiva/Editora da Universidade de São Paulo, 1978.
- Furlanetto, C. D. D. M., Weymer, A. S. Q., & Matos, R. D. (2023). Capitalismo Consciente e Construção de Relações Humanizadas: Um Estudo em uma Cooperativa de Crédito sob a Perspectiva do Sensemaking. *Revista de Administração Contemporânea*, 27. <https://doi.org/10.1590/1982-7849rac2022210251.por>.
- Furlanetto, C. D. D. M., Weymer, A. S. Q., & Matos, R. D. (2023). Conscious Capitalism and Construction of Humanized Relationships: A Study in a Credit Cooperative from the Sensemaking Perspective. *Revista de Administração Contemporânea*, 27. <https://doi.org/10.1590/1982-7849rac2022210251.en>.

Gilbert, Thomas F. *Human competence: engineering worthy performance*. New York: McGraw-Hill Book Company, 1978.

Giotopoulos, K.C., Alexakos, C.E., Beligiannis, G.N., Likothanassis, S.D.: Integrating agents and computational intelligence techniques in E-learning environments. *Int. J. Soc. Hum. Sci. Eng.* 1 (7), 412–419 (2007).

Goulart, E. L., Weymer, A. S. Q., & Moreira, V. R. (2022). The influence of self-efficacy on training effectiveness in cooperative organizations. *Revista de Administração da UFSM*, 15, 331-353. <https://doi.org/10.5902/1983465967213>.

Hanashiro, D. M. M., Teixeira, M. L. M., & Zaccarelli, L. M. others.(2007). *Gestão do fator humano: uma visão baseada em stakeholders*. São Paulo: Saraiva.

Jain, S. (2017). IS ARTIFICIAL INTELLIGENCE –THE NEXT BIG THING IN HR International Conference on Innovation Research in Scinece, Techonology and Management (pp. 220-224). Rajasthan : Modi Intitute of Management & Technology.

Kaczmarek, T., Kowalkiewicz, M., Pikorski, J: Information Extraction from CV. In: *Proceedings of the 8th International Conference on Business Information Systems*, pp. 3–7 (2005).

Kahraman, C., Kaya, I., Çinar, D.: Computational intelligence: past, today, and future. In: Ruan, D. (ed.) *Computational Intelligence in Complex Decision Systems*, pp. 1–46. Atlantis Press, Paris (2010).

Khan, W. A., Chung, S. H., Awan, M. U., & Wen, X. (2020). Machine learning facilitated business intelligence (Part I). *Industrial Management & Data Systems*, 120(1), 164–195. <https://doi.org/10.1108/watIMDS-07-2019-0361>.

Kirkpatrick, S. A., & Locke, E. A. (1991). *Measurement of performance in business: Evaluation and planning for performance improvement*. New York, NY: Free Press.

Lucena, M. D. S.; *Avaliação de Desempenho*. São Paulo: Atlas, 1992.

Mendes, W. *Gestão estratégica de pessoas e competências*. Curitiba: FGV, 2017.

Miner, J. B. (2005). Competency-based performance evaluation: A strategic approach to performance management. *Journal of Management*, 3(4), 359-374.

Mitchell, T. M. (2006). *The discipline of machine learning* (Vol. 9). Pittsburgh: Carnegie Mellon University, School of Computer Science, Machine Learning Department.

Ostroff, C. and Bowen, A.E. (2000), “Moving HR to a higher level: HR practice and organizational effectiveness”, in Klein, K.J. and Koslowski, S.W.J. (Eds), *Multilevel Theory, Research, and Methods in Organizations, Foundations, Extensions, and New Directions*, Jossey-Bass, San Francisco, pp. 211-266.

Pellin, N. P., Weymer, A., Dissenha, L. A., & Bauer, M. A. L. (2021). Organizational links and

sensemaking in a medical work cooperative system. *Qualitative Research in Organizations and Management: An International Journal*, 16(2), 388-408. <https://doi.org/10.1108/QROM-08-2020-2015>.

- Prahalad, C. K.; Hamel, G. The Core Competence of the Corporation. *Harvard Business Review*, 1990.
- Ridhwan, K. M. , & Hargreaves, C. A. (2021). Leveraging Twitter Data to Understand Public Sentiment for the COVID-19 Outbreak in Singapore. *International Journal of Information Management Data Insights* , Article 100021.
- Rothwell, W. J. (2012). Performance appraisal by competencies: A guide to improving performance and productivity. *Journal of Management*, 11(2), 123-139.
- Ruzzarin, T.; Simionovschi, M. Competências: uma base para a governança corporativa. Porto Alegre: AGE, 2010.
- Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological science in the public interest*, 13(2), 74-101.
- Sen, P. C., Hajra, M., & Ghosh, M. (2020). Supervised classification algorithms in machine learning: A survey and review. In *Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018* (pp. 99-111). Springer Singapore.
- Silverman, D., & Marvasti, A. (2000). *Doing qualitative research*. London. Sage.
- Sitzmann, T., & Weinhardt, J. M. (2017). Approaching evaluation from a multilevel perspective: A comprehensive analysis of the indicators of training effectiveness. *Human Resource Management Review*.
- Sitzmann, T., & Weinhardt, J. M. (2019). Approaching evaluation from a multilevel perspective: A comprehensive analysis of the indicators of training effectiveness. *Human Resource Management Review*, 29(2), 253-269.
- Souza, A. R. S.; Vasconcelos, I. F. F. G. Meritocracia e gestão de pessoas por competências: tema utópico ou realidade organizacional? *Cadernos EBAPE.BR*, p. 1-17, 2020.
- Stake, R. E. (2016). *Pesquisa qualitativa: estudando como as coisas funcionam*. Penso Editora.
- Strohmeier, S., & Piazza, F. (2015). Artificial intelligence techniques in human resource management—a conceptual exploration. *Intelligent Techniques in Engineering Management: Theory and Applications*, 149-172.
- Tanwar, P., Prasad, T.V., Aswal, M.S.: Comparative study of three declarative knowledge representation techniques. *Int. J. Comput. Sci. Eng.* 2(07), 2274–2281 (2010).
- Unimed. Unimed de Londrina – Quem Somos – Governança. Unimed de Londrina, 2021. Disponível em: <<https://www.unimedlondrina.com.br/quem-somos/#governanca>>. Acesso em: out. 2021.

- Wang, P.: What do you mean by “AI”? In: Artificial General Intelligence, 2008: Proceedings of the First AGI Conference, vol. 171, p. 362. IOS Press (2008).
- Weymer, A. S. Q., & Silva, E. D. D. (2014). A contribuição do conceito de competências nos campos científico e gerencial: um tour sobre pesquisas empíricas e suas relações epistemológicas. <http://repositorio.furg.br/handle/1/6937>.
- Wolf, A., Jenkins, A.: Explaining greater test use for selection: The role of HR professionals in a world of expanding regulation. *Hum. Resour. Manage. J.* 16(2), 193–213 (2006).
- Yawalkar, M. V. V. (2019). a Study of Artificial Intelligence and its role in Human Resource Management. *International Journal of Research and Analytical Reviews (IJRAR)*, 6(1), 20-24.
- Zangiski, M. A. S. G.; lima, E. P.; Costa, S. E. G. Organizational competence building and development: Contributions to operations management. *International Journal of Production Economics*, v. 144, n. 1, p. 76-89, 2013.
- Zuboff, S.: Automate/informate: the two faces of intelligent technology. *Org. Dyn.* 14(2), 5–18 (1985).