



ORANGE DATA MINING NO RECRUTAMENTO: UMA ANÁLISE PREDITIVA PARA OTIMIZAR PROCESSOS

Arthur Quadros Pegoretti¹ - <https://orcid.org/0009-0003-3996-7630>

Alexandre Meira de Vasconcelos² - <https://orcid.org/0000-0003-0824-9495>

Gabriella Zanoto Botton^{3*} - <https://orcid.org/0000-0001-9589-1202>

^{1 2 3} FAENG, Universidade Federal de Mato Grosso do Sul, Campo Grande, Brasil

¹ Email: arthurpegoretti@gmail.com

² Email: meira1970@gmail.com

³ Email: gabizanotobotton@gmail.com

* Corresponding author

RESEARCH ARTICLE

Annals of Data Analysis

<https://periodicos.ufpe.br/revistas/ADA/>

Submitted on: 12.12.2025.

Accepted on: 01.08.2026.

Published on: 01.08.2026.

Copyright © 2026 by author(s).

This work is licensed under the Creative Commons Attribution International License CC BY-NC-ND 4.0

<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>



Resumo

As empresas de Recursos Humanos têm passado por transformações devido à globalização e avanços tecnológicos, como Inteligência Artificial (IA) e Machine Learning (ML). A seleção de candidatos para entrevistas ainda enfrenta desafios, como processos manuais e demorados, destacando a necessidade de soluções automatizadas para melhorar eficiência e assertividade. O estudo visa aplicar algoritmos de ML, utilizando o software Orange Data Mining, para criar um modelo preditivo que auxilie na seleção de candidatos em uma empresa de RH, reduzindo custos e tempo no processo de recrutamento. Utilizou-se um estudo quantitativo com dados da empresa. Quatro algoritmos (Random Forest, Neural Network, KNN e Árvore de Decisão) foram testados para analisar 300 candidatos, com variáveis como experiência, escolaridade e habilidades. O método DMAIC (Define, Measure, Analyse, Improve, Control) estruturou a abordagem. Verificou-se que o algoritmo Neural Network apresenta os melhores resultados de predição para candidatos reprovados, enquanto o KNN para candidatos aprovados. Dificuldades na coleta automatizada de dados e a necessidade de avaliar manualmente os aspectos qualitativos (como habilidades interpessoais) limitaram a generalização dos resultados. Sugere-se a inclusão de mais variáveis qualitativas e a atualização contínua da base de dados, além da integração de sistemas para automatizar a coleta de informações, melhorando a precisão dos modelos.

Palavras-chave

Algoritmos; Análise Preditiva; Mapeamento de processos; Modelo de aprendizagem de máquina; Orange Data Mining.

1. Introdução

O setor de Recursos Humanos (RH) tem passado por rápidas transformações impulsionadas pela globalização e pelo avanço tecnológico (Artar, Balcioglu & Erdil, 2024). Nesse contexto, o recrutamento e a seleção de profissionais qualificados tornaram-se funções estratégicas para o crescimento organizacional. Diante disso, no momento de procurar por novos funcionários, as empresas se deparam com um paradigma: fazer a contratação utilizando os seus próprios colaboradores, despendendo hora/trabalho dos mesmos, para selecionar seus futuros colegas ou terceirizar esse serviço para que outras companhias, especializadas em recrutamento e seleção, busquem os melhores profissionais disponíveis no mercado (Bosmans, 2016).

Com a chegada da Revolução 4.0, o RH passa a incorporar ferramentas tecnológicas que automatizam tarefas operacionais e tornam a gestão de pessoas mais eficiente. Entre essas inovações, destacam-se a Inteligência Artificial e o Machine Learning, que permitem a análise de grandes volumes de dados (Silveira, 2021) e a geração de insights preditivos (Furuya, 2020). Softwares como o Orange Data Mining têm ganhado espaço por facilitar o processamento e a visualização de dados, contribuindo para a tomada de decisão em diferentes áreas do conhecimento (Naik & Samant, 2016), podendo ser aplicado em: pesquisas no ramo da medicina (Kodati, 2019), estudos desenvolvidos sobre botânica e ecologia (Madhumathi, 2017), análises e previsões da performance das notas escolares de alunos (Adekitan, 2019).

Esses avanços evidenciam a necessidade de constante atualização dos processos organizacionais, especialmente frente ao crescimento de empresas com estruturas mais ágeis e tecnológicas. Segundo dados do IBGE (2019), já existiam 961 empresas no Brasil especializadas na gestão terceirizada de Recursos Humanos, onde na região Centro-Oeste são mais de 75 empresas que também estão oferecendo soluções estratégicas na contratação de novos colaboradores. Diante disso, a busca por inovações tecnológicas e refinamento dos serviços oferecidos se torna cada vez mais importante e presente no dia a dia das companhias de terceirização de R&S.

A empresa a ser estudada é uma consultoria de Recursos Humanos, presente há mais de 30 anos em Campo Grande/MS, onde atualmente encontra-se sua única sede física. Porém, ao longo dos anos já possuiu mais de 15 filiais espalhadas pelas quatro regiões do Brasil, e hoje, pelo “know how” constituído, oferece suas soluções em todo o território nacional. Seus serviços ofertados vão desde o recrutamento e seleção de novos colaboradores, sendo eles de carteira assinada ou não, terceirização de temporários, avaliações psicológicas até consultorias personalizadas para desenvolvimento organizacional.

Ao longo dessas três décadas de operação a companhia atendeu mais de treze mil clientes, de todos os portes e tamanhos, e em todos os estados do território nacional, com ênfase na região Centro-Oeste e empresas de médio porte. Além disso, também já exerceu suas atividades até em outros países, como: Paraguai, Bolívia e Angola. Em 2019, adquiriu um software de recrutamento e seleção online com o objetivo de integrar os setores Comercial e de Recrutamento, centralizar os processos seletivos em uma única plataforma e reduzir tarefas manuais. A decisão foi impulsionada pelo aumento expressivo no número de candidatos: apenas até julho de 2021, mais de 3.100 novos usuários se cadastraram no site da consultoria.

No intuito de auxiliar na gestão das atividades do projeto, foi escolhido um método de resolução de problemas: DMAIC, composto pelas etapas: define (definir), measure (medir), analyse (analisar), improve (melhorar) e control (controlar), que foi escolhido pois, ultimamente, é o mais utilizado pelos autores e profissionais da área, já que por contar com cinco etapas, possibilita uma adequada organização da sua implantação, desenvolvimento e medição (Andritetta & Miguel, 2007). Assim, o objetivo geral do trabalho consiste em utilizar o banco de dados da empresa, para com o auxílio do software Orange Data Mining, criar uma análise preditiva dos futuros candidatos a serem selecionados pela empresa de RH.

2. Procedimentos metodológicos

Trata-se de um estudo de caso com natureza mista e aplicada, com fonte de dados mistos e com instrumentos de coleta de dados questionários, entrevistas e dados disponibilizados pela empresa. O primeiro procedimento do estudo foi uma entrevista não estruturada com os dois proprietários da organização, que consistiu basicamente na identificação das possíveis problemáticas de cada um dos serviços ofertados e possíveis melhorias para o processo como um todo. Dentre os problema mapeados, verificou-se que a etapa de filtragem dos candidatos a serem selecionados e as entrevistas individuais, pessoalmente ou por alguma plataforma online, era o mais impactante e

se concentrava no departamento de seleção da empresa dado são trabalhos manuais que exigem aplicação prática e não contam com apoio tecnológico para a execução.

Para auxiliar na etapa de análise dos dados, utilizou-se o software Orange Data Mining, uma ferramenta baseada em linguagem de programação utilizada para análises estatísticas, principalmente para modelagens visuais, devido ao seu ótimo layout de sistema (Isahk et al, 2020). Assim, utilizou-se quatro algoritmos de aprendizado de máquina, sendo eles os seguintes: Random Forest (Floresta Aleatória) na qual o algoritmo de florestas aleatórias combina previsões de várias árvores desenvolvidas isoladamente, oferecendo processamento rápido e desempenho superior a outros baseados em árvores de decisão (KODATI, 2018); Neural Network (Rede Neural) que o algoritmo de redes neurais identifica padrões e tendências complexas em dados, aprendendo a modelar a relação entre entradas e saídas para gerar os resultados esperados (TURRA, 2017).

Ademais, também se utilizou o Knn (K-Nearest Neighbor Ou K-Vizinhos Mais Próximos) no qual O KNN, baseado em aprendizado por analogia, classifica novos itens considerando os “k” vizinhos mais próximos e atribuindo-os à classe mais frequente (Caetano, 2015; Kodati, 2018). Esse método utiliza comparações com grupos já definidos para determinar a classificação de um novo elemento. E por fim, a TREE (Árvore De Decisão) em que o método funciona como uma árvore, onde decisões sucessivas baseadas em testes dividem os dados em subconjuntos, levando à classificação final representada pelas folhas (Caetano, 2015; Mohi, 2020).

3. Referencial teórico

A evolução das tecnologias aplicadas ao processo de recrutamento e seleção de pessoas dentro das organizações vêm potencializando o volume de análise dos candidatos, como também, otimizando as etapas de identificação dos perfis (Nain & Shyam, 2024). Assim, a tomada de decisão dentro das organizações que utilizam IA na seleção de pessoas deve ser aprimorada além disso, alinhada com a cultura organizacional da corporação.

Dentre os benefícios percebidos, pode-se citar a diminuição da rotatividade de funcionários em uma organização, dado que com uma análise preditiva amparada pelo uso de tecnologias, é possível detectar padrões de sucesso ou fracasso de determinado perfil dentro da empresa (Nain & Shyam, 2024). Ademais, a redução dos custos também é apontada como fator de impacto para adoção da metodologia (Odili et al., 2024), dado que otimizam o processo de recrutamento, ampliam o alcance das vagas, facilitam as etapas de candidatura (Meshram, 2023) e melhoram a performance dos candidatos dentro das corporações (Kshetri, 2021).

Na etapa de “define” verificou-se qual a problemática a ser solucionada. Para isso, as etapas do serviço e as atividades exercidas por cada departamento foram identificadas e mapeadas. O cronograma de Análise Preditiva Candidatos é uma ferramenta que auxilia às organizações por meio das avaliações de desempenho e dimensões estruturais de um fluxo contínuo de trabalho, sendo considerada um instrumento da área de qualidade, pois proporciona mudanças na estrutura organizacional da empresa, como: eliminação de vícios, redução de falhas, integração de sistemas, simplificação de atividades e principalmente, redução de custos (Longaray et al, 2017). Partindo da premissa que o serviço mais comercializado pela companhia é o de Recrutamento e Seleção, em que outras empresas contratam a consultoria para buscar, filtrar e selecionar os melhores perfis disponíveis do mercado. Foi identificado pelo Departamento de Seleção que o processo mais trabalhoso dessa solução ofertada é o recrutamento dos possíveis candidatos aptos a participarem do processo seletivo e as entrevistas individuais com cada um dos aprovados na etapa de filtragem, pois o número de inscritos por vagas é relativamente alto, uma média superior a 50 candidatos por vaga ofertada.

A etapa de medição tem como objetivo identificar e analisar o processo como um todo, reunindo informações e dados essenciais sobre o problema estudado, avaliando o desempenho dos procedimentos e a qualidade frente às expectativas da organização e dos clientes (Gonçalves et al., 2016). Para participar de um processo seletivo, o candidato deve cadastrar seu currículo no site da empresa com o máximo de informações e contatos possíveis. A partir desse cadastro, o Departamento de Seleção inicia a filtragem manual dos candidatos que possuem as habilidades desejadas pelo cliente. Após a identificação dos aprovados, verifica-se se o número é suficiente para avançar; caso contrário, o processo retorna ao Departamento de Marketing para novas divulgações, enquanto as entrevistas com os aptos já podem ocorrer.

Para auxiliar na etapa de análise dos dados com auxílio do software Orange Data Mining, que com auxílio de algoritmos de Machine Learning, sendo esse um processo computacional, o qual busca deduzir automaticamente um modelo de aprendizado, fundamentado no conjunto de dados, finito e amostral, que foi adicionado anteriormente ao sistema. De uma maneira mais descomplicada, esse processo permite o Software aprender, sem necessariamente ser programado para isso, otimizando sua análise e desempenho com base nos resultados de cada um dos modelos de aprendizado (Gomes, 2019).

4. Análise e discussão dos resultados

4.1. Definição

Para melhorar a visualização e o entendimento dos processos de triagem e seleção dos candidatos, utilizou-se uma das principais ferramentas de *Lean Six Sigmas*, o SIPOC. Esta, por sua vez, tem seu nome derivado das palavras em inglês, *Supply* (Fornecedor), *Input* (Entrada), *Process* (Processo), *Output* (Saída) e *Customer* (Cliente) e tem como propósito mapear de forma macro todas as informações pertinentes, do início ao fim, de um projeto ou processo (Mattos, 2021). Além disso, identificar elementos básicos de um conjunto de processos (Baldam et al, 2014), como demonstrado na Figura 1.

Quadro 1: SIPOC da filtragem e entrevista dos candidatos.

S	I	P	O	C
Fornecedores	Entradas	Processos	Saídas	Clientes
Clientes Candidatos	Equipamentos eletrônicos (notebook e celular) Recursos telefônicos Informações da vaga e perfil solicitado pelo cliente Dados cadastrais informados pelos candidatos Treinamento Acesso ao sistema da empresa	Fazer a busca dos candidatos no sistema Realizar a triagem manual dos participantes	Possíveis candidatos a serem contratados pelos clientes. Boa experiência do cliente e do candidato	Selecionadores Candidatos

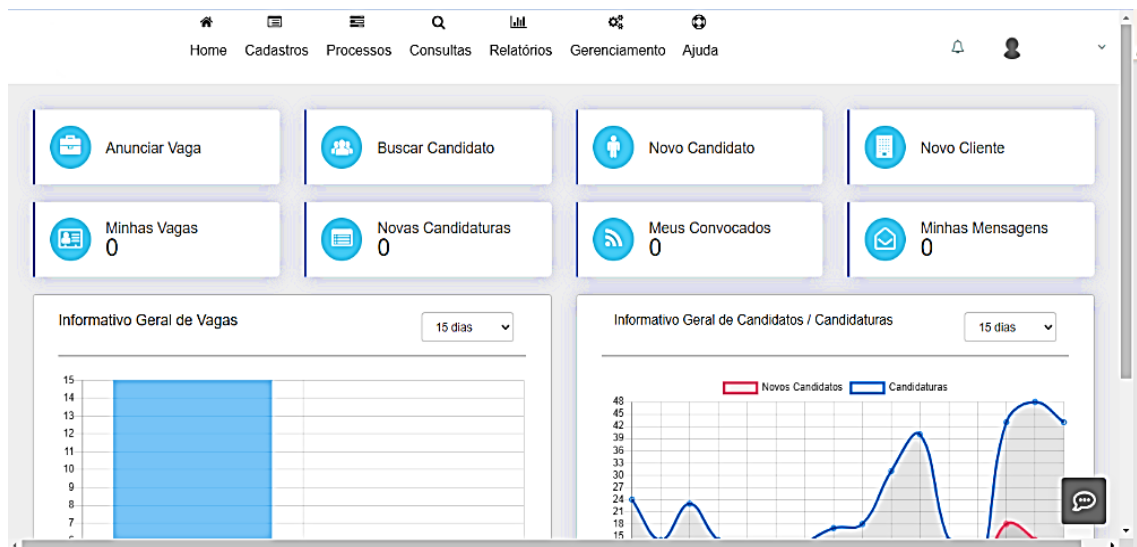
Fonte: Elaborado pelos autores, 2025.

O processo de análise dos dados curriculares de cada perfil é feito de forma manual e individual, tornando o processo trabalhoso e de alta demanda por atenção aos mínimos detalhes. Já a etapa de entrevista com cada participante filtrado também utiliza uma metodologia de aplicação manual e por isso exige total empenho do colaborador. Sendo então os dois processos executados sem nenhum apoio de tecnologia na seleção de cada perfil. A atividade de recrutamento exige uma série de esforços para ser concluída com êxito, demandando empenho de alguns departamentos como: Comercial, Marketing e por último, o de Seleção. Por isso, vários processos precisam estar sempre bem alinhados, possibilitando que o sistema tenha um funcionamento eficiente e sem comprometimentos. Enquanto a execução das entrevistas com os candidatos demanda esforços únicos e exclusivamente do departamento de Seleção, visto que somente esses profissionais estão aptos a desempenhar o papel de Entrevistadores, onde analisam diversas informações, desde técnicas quanto comportamentais, de cada candidato, para definir sua aprovação para a próxima etapa. Por ser um papel único, de alta complexidade e muitas vezes personalizado para cada tipo de vaga, demanda muitos esforços dos colaboradores que o executam, além disso, frequentemente se faz necessário o retrabalho da etapa devido ao não atingimento das metas estabelecidas. Para o êxito do projeto, foi necessário a colaboração do departamento de Seleção, Marketing, Comercial e da empresa terceirizada que fornece o sistema para a consultoria. Aquém disso, baseado no Gráfico de GANTT, uma ferramenta de qualidade voltada para o planejamento e programação de atividades, houve a demarcação do início e o fim das atividades (Junqueira et al, 2015).

4.2 Medição

O processo de entrevista é conduzido exclusivamente pelo Departamento de Seleção, formado por Psicólogas Organizacionais, e é considerado a etapa de maior valor para o cliente, pois nela são escolhidos os futuros profissionais. Essas entrevistas são personalizadas, considerando especificações dos clientes, dificuldade de seleção do perfil, recursos necessários e urgência no fechamento da vaga. A Figura 2 apresenta o sistema utilizado pela instituição, de onde foram extraídos relatórios com dados para este estudo.

Figura 1: Print de tela do Dashboard da empresa.



Fonte: Elaborado pelos autores, 2025.

Os documentos extraídos continham informações como descrição detalhada das vagas, dados de candidatos e informações de clientes, que foram compilados em planilhas no Excel. Em etapas seguintes, foi necessário extrair manualmente dados adicionais do Sistema para completar as planilhas, que serviram de base para a análise preditiva no software Orange. Por razões éticas, de sigilo e em conformidade com a Lei Geral de Proteção de Dados (Lei 13.709/2018), todos os dados sensíveis que poderiam identificar os participantes, como nome, CPF e endereço, foram desconsiderados (Oliveira et al., 2021).

Para compor as planilhas base, foram selecionadas todas as vagas abertas em 2020 e parte de 2021, totalizando cem posições de diferentes níveis de escolaridade. Definidas as vagas, foram escolhidos os cem profissionais contratados para essas posições, acrescidos de dois candidatos selecionados aleatoriamente, totalizando trezentos participantes. Cada perfil foi registrado em planilha para futura manipulação e análise dos dados, conforme descrito na Figura 3.

Figura 2: Print de tela da planilha com os dados compilados.

	A	B	C	D	E	F	G	H	I
1	CA_AP	CA_ID	CA_IDA	CA_EXP	CA_END	CA_SEX	CA_EC	CA_CNH	CA_TRA
2	Apr	74989	25 - 35	Muita Experiência	Não	Feminino	Solteiro	Sim	Sim
3	Rep	47711	25 - 35	Média Experiência	Sim	Masculino	Solteiro	Sim	Sim
4	Rep	23578	25 - 35	Muita Experiência	Sim	Masculino	Solteiro	Sim	Sim
5	Apr	71252	25 - 35	Muita Experiência	Sim	Masculino	Casado	Não	Sim
6	Rep	66271	25 - 35	Muita Experiência	Não	Masculino	Solteiro	Sim	Não
7	Rep	76222	35 - 45	Média Experiência	Sim	Masculino	Solteiro	Sim	Não
8	Apr	75897	25 - 35	Média Experiência	Sim	Feminino	Solteiro	Não	Não
9	Rep	69030	25 - 35	Média Experiência	Sim	Feminino	Solteiro	Não	Não
10	Rep	626	25 - 35	Muita Experiência	Sim	Feminino	Casado	Sim	Não
11	Apr	15610	25 - 35	Muita Experiência	Sim	Masculino	Solteiro	Sim	Sim
12	Rep	74892	25 - 35	Pouca Experiência	Sim	Masculino	Solteiro	Não	Não
13	Rep	33241	35 - 45	Muita Experiência	Sim	Masculino	Solteiro	Sim	Sim
14	Apr	74943	25 - 35	Pouca Experiência	Sim	Feminino	Solteiro	Sim	Sim
15	Rep	75101	25 - 35	Média Experiência	Sim	Feminino	Solteiro	Não	Sim
16	Rep	40607	25 - 35	Muita Experiência	Sim	Masculino	Solteiro	Sim	Sim
17	Apr	16298	25 - 35	Média Experiência	Sim	Masculino	Solteiro	Sim	Não
18	Rep	74474	45 - 60	Muita Experiência	Sim	Masculino	Solteiro	Sim	Não
19	Rep	72195	35 - 45	Muita Experiência	Sim	Feminino	Casado	Sim	Não

Fonte: Elaborado pelos autores, 2025.

As variáveis para preenchimento da planilha foram definidas em reunião com as Psicólogas e a proprietária da empresa, utilizando o método de Brainstorming, que consiste no compartilhamento de ideias sem críticas (Buchele

et al., 2017). Nesse debate, foram selecionadas apenas características disponíveis no banco de dados da empresa, oriundas do cadastro realizado pelo candidato no site, desconsiderando aspectos interpessoais, de personalidade e de alinhamento de valores com a companhia, pois essas informações são avaliadas apenas nas entrevistas e podem ser influenciadas por outros fatores, conforme Tabela 1.

Tabela 1: Variáveis da planilha definida no brainstorming.

Variáveis	Legenda	Valores
CA_AP	Se o candidato foi aprovado na vaga	(Apr) Aprovado, (Rep) Reprovado.
CA_ID	Número identificador no Sistema	Valor Min: 82, Valor Max: 77556
CA_IDA	Faixa-etária	18 - 25, 25 - 35, 35 - 45, 45 - 60, "+" 60
CA_EXP	Experiência comprovada	Sem Experiência, Pouca Experiência, Média Experiência, Muita experiência
CA_END	Se reside na mesma cidade da vaga	Sim, Não
CA_SEX	Sexo	Feminino, Masculino
CA_EC	Estado civil	Solteiro, Casado
CA_CNH	Se possui CNH	Sim, Não
CA_TRA	Se está trabalhando atualmente	Sim, Não
CA_ESC	Escolaridade do candidato	Ensino Médio Completo, Ensino Técnico Completo, Ensino Superior Incompleto, Ensino Superior Completo, Pós-graduação
CA_CEX	Se possui cursos extracurriculares	Sim, Não
CA_PCO	Se possui pacote office	Sim, Não
CA_IDM	Se fala um segundo idioma	Sim, Não

Fonte: Elaborado pelos autores, 2025.

4.3 Analisar

Após a definição dos algoritmos, realizou-se o relacionamento individual com o banco de dados e o teste de precisão de cada um dos modelos. Sendo estes calculados pela proporção de verdadeiros positivos entre as instâncias classificadas como positivas (Ajda, 2017). Conforme feita a constatação pelo software, os algoritmos com as melhores precisões para a análise preditiva desse objeto de estudo, em ordem decrescente, foram: *Neural Network* (Rede Neural), *Random Forest* (Floresta Aleatória), KNN (k-Nearest Neighbor ou k-vizinhos mais próximos) e por último *Tree* (Árvore de decisão).

Tabela 2 - Precisão dos algoritmos.

Modelo	Precisão
Neural Network	0.714
KNN	0.675
Random Forest	0.667
Tree	0.601

Fonte: Elaborado pelos autores, 2025.

Para dar apoio às análises, elaborou-se, por meio do software Orange, matrizes de confusão para cada um dos algoritmos de aprendizado de máquina utilizados no estudo.

Tabela 2a – Matriz de Confusão – Random Forest

	Apr	Rep	Σ
Apr	37	63	100
Rep	31	169	200
Σ	68	232	300

Tabela 2b – Matriz de Confusão – Neural Network

	Apr	Rep	Σ
Apr	21	79	100
Rep	8	192	200
Σ	29	271	300

Tabela 2c – Matriz de Confusão – KNN

Tabela 2d – Matriz de Confusão – Tree

	Apr	Rep	Σ
Apr	42	58	100
Rep	35	165	200
Σ	77	223	300

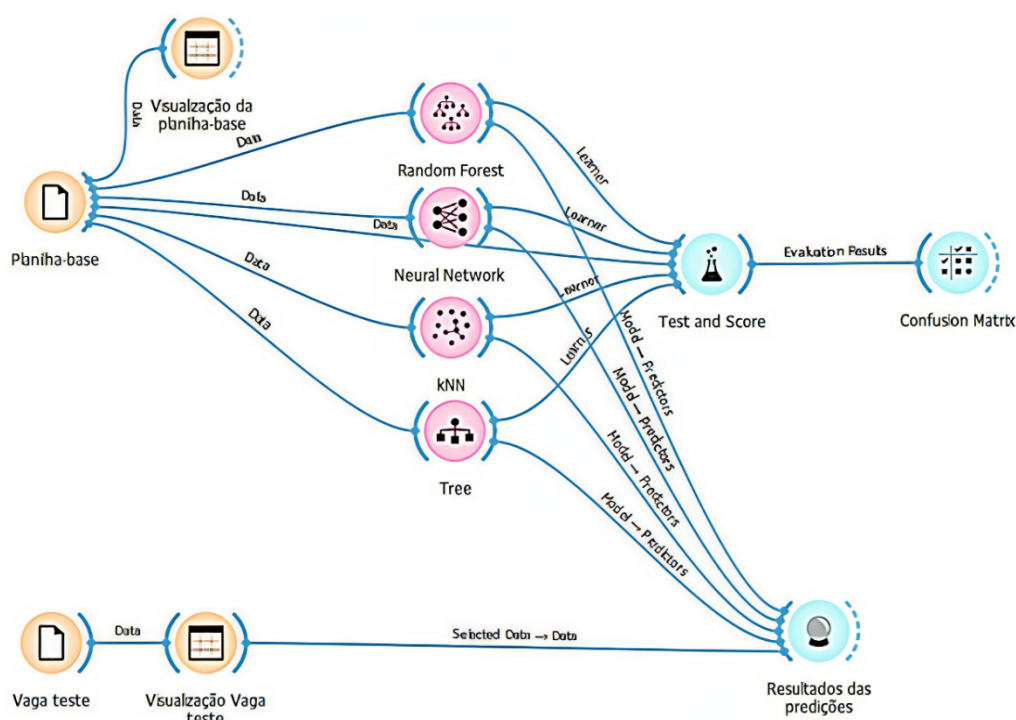
	Apr	Rep	Σ
Apr	33	67	100
Rep	47	153	200
Σ	80	220	300

Fonte: Elaborado pelos autores, 2025.

Essa matriz é utilizada para oferecer medidas efetivas do modelo de classificação do método aplicado, mostrando o número de classificações corretas e incorretas para cada classe do estudo (Matos et al, 2021). Há duas classes sendo consideradas na classificação dos dados: Apr e Rep. Neste caso, a classe Apr corresponde aos candidatos da base de dados que foram aprovados nos processos seletivos; e a classe Rep corresponde aos candidatos que foram reprovados. A matriz mostra o número dos dados classificados como uma classe e aqueles que realmente são dessa classe, ou seja, também ilustra os falsos positivos e os verdadeiros positivos. A utilização dessa ferramenta se dá por meio da ligação de células entre uma categoria e outra, por exemplo: no algoritmo Random Forest o número 37 é referente ao encontro da coluna “Apr” com a linha “Apr”, sendo então o número de acertos desse modelo; já a célula com o numeral 63 é a junção da coluna “Rep” com a linha “Apr”, sendo esses a quantidade de acertos que o modelo deixou de afirmar.

Feita a análise dos algoritmos com os valores de precisão e a matriz de confusão, realizou-se a aplicação dos algoritmos em vagas recém-fechadas pela empresa, onde analisaram-se somente os profissionais que participaram da última etapa de seleção, a qual é pertinente a entrevista feita do candidato direto com a empresa contratante. Para esse teste foram selecionadas as últimas cinco posições que a empresa havia concluído, sendo divididas da seguinte forma: quatro vagas de nível médio ou superior incompleto (dos códigos de sistema: 21826, 21857, 21863 e 21873) e uma vaga de nível superior completo (do código de sistema: 21878). Foi criada a visualização da análise preditiva dentro do software Orange para a aplicação de cada modelo de aprendizado de máquina estabelecido, conforme Figura 5.

Figura 3: Modelo de predição no software Orange.



Fonte: Elaborado pelos autores, 2025.

A aplicação do algoritmo foi processada individualmente por cada vaga, sem que houvesse a influência de uma nova base de dados para a posição seguinte. Considerando que o processo já havia sido finalizado e um dos profissionais já havia sido contratado, o escolhido possui indicado as letras “AP” na sua variável “CA_ID” da planilha, correspondente ao seu ID de sistema. Os valores da probabilidade de aprovação de cada um dos candidatos em cada uma das vagas estão apresentados por meio das tabelas 3, 4, 5, 6 e 7.

Para a vaga 21826, de nível superior incompleto ou médio, obtiveram-se os resultados da Tabela. O candidato da linha oito foi aprovado pela empresa e possui os seguintes resultados preditivos de aprovação: 74% de chance de ser aprovado pelo algoritmo *Random Forest*, 33% pelo *Neural Network*, 57% pelo KNN e indeterminado pelo *Tree*.

Tabela 3 - Resultado da predição dos candidatos da vaga 21826.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
73273	21826	21% - Rep	33% - Rep	39% - Rep	0% - Rep
77229	21826	30% - Rep	33% - Rep	0% - Rep	0% - Rep
77233	21826	48% - Rep	34% - Rep	60% - Apr	0% - Rep
23141	21826	28% - Rep	23% - Rep	39% - Rep	1% - Apr
76929	21826	14% - Rep	24% - Rep	0% - Rep	0% - Rep
76944	21826	39% - Rep	24% - Rep	43% - Rep	1% - Apr
77181	21826	5% - Rep	16% - Rep	30% - Rep	0% - Rep
AP-77808	21826	74% - Apr	33% - Rep	57% - Apr	0% - Rep

A vaga de código 21857, também de nível superior incompleto ou médio, resultou na solução representada na Tabela 4. Nesta posição, o candidato da linha dois foi o selecionado e possui as seguintes probabilidades de aprovação: 51% pelo *Random Forest*, 25% pelo *Neural Network*, 59% pelo KNN e 24% pelo *Tree*.

Tabela 4: Resultado da predição dos candidatos da vaga 21857.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
67777	21857	69% - Apr	43% - Rep	62% - Apr	83% - Apr
AP-77542	21857	51% - Apr	25% - Rep	59% - Apr	24% - Rep
77514	21857	47% - Rep	34% - Rep	41% - Rep	75% - Apr
63754	21857	31% - Rep	27% - Rep	37% - Rep	24% - Rep
77472	21857	17% - Rep	30% - Rep	28% - Rep	25% - Rep

Fonte: Elaborado pelos autores, 2025.

A vaga de número 21863, para candidatos de nível médio ou superior incompleto, obteve as respostas da Tabela 5. Nessa vaga, o candidato da linha dois também foi o selecionado, tendo ele as seguintes probabilidades de aprovação: 45% pelo *Random Forest*, 50% pelo *Neural Network*, 61% pelo KNN e 100% pelo *Tree*.

Tabela 5: Resultado da predição dos candidatos da vaga 21863.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
77493	21863	47% - Rep	20% - Rep	0% - Rep	33% - Rep
AP-77498	21863	45% - Rep	50% - Rep	61% - Apr	1% - Apr
77627	21863	46% - Rep	37% - Rep	29% - Rep	14% - Rep
77570	21863	47% - Rep	20% - Rep	0% - Rep	33% - Rep

Fonte: Elaborado pelos autores, 2025.

Para a vaga de código 21873, a última para profissionais de nível médio ou superior incompleto, os valores obtidos estão na Tabela 6. Na vaga 21873 os valores de predição para o candidato selecionado, da linha dois, foram: 44% de probabilidade de aprovação pelo algoritmo *Random Forest*, 47% pelo algoritmo *Neural Network*, 70% pelo KNN e indeterminado pelo *Tree*.

Tabela 6: Resultado da predição dos candidatos da vaga 21873.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
77598	21873	35% - Rep	24% - Rep	58% - Apr	0% - Rep
AP-77553	21873	44% - Rep	47% - Rep	70% - Apr	0% - Rep
77638	21873	36% - Rep	37% - Rep	0% - Rep	0% - Rep
74022	21873	29% - Rep	25% - Rep	0% - Rep	17% - Rep
76851	21873	35% - Rep	28% - Rep	61% - Apr	100% - Apr

Fonte: Elaborado pelos autores, 2025.

Para a única vaga destinada a profissionais de nível superior, de código 21878, encontrou-se as predições da Tabela 7. A vaga 21878 resultou em uma particularidade em suas predições, pois foi a única em que as porcentagens do candidato selecionado, da linha um, *Random Forest* = 89%, *Neural Network* = 59%, KNN = 81%

e *Tree* = 50%, foram acima ou igual a 50%, sendo então considerado aprovado por todos os algoritmos de aprendizados de máquina testados neste estudo.

Tabela 7: Resultado da predição dos candidatos da vaga 21878.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
AP-30694	21878	89% - Apr	59% - Apr	81% - Apr	50% - Apr
51717	21878	44% - Rep	38% - Rep	34% - Rep	50% - Apr
36374	21878	56% - Apr	39% - Rep	89% - Apr	50% - Apr
1421	21878	14% - Rep	30% - Rep	37% - Rep	0% - Rep

Fonte: Elaborado pelos autores, 2025.

A singularidade obtida nos resultados da vaga de nível superior pode ter diversas explicações, como: maior homogeneidade dos dados; o candidato aprovado, em específico pode ter dados mais relevantes para a precisão dos modelos ou até melhor aplicabilidade da análise preditiva baseada no banco de dados elaborado para as vagas destinadas a candidatos de nível superior.

Para aperfeiçoar a aplicabilidade dos modelos de aprendizado de máquina, verificou-se com a empresa que havia outras quatro vagas (dos códigos de sistema: 21886, 21805, 21652 e 21328), destinadas a profissionais com escolaridade de nível superior completo, que iriam iniciar a etapa de “Entrevista com os candidatos selecionados”, conforme o mapeamento de processos descrito no Apêndice B, que poderiam então ter os candidatos testados pelos algoritmos do software. Para isso, foram seguidos os mesmos passos que as aplicações apresentadas anteriormente, porém, nesse caso, ainda não havia nenhum candidato selecionado, então na coluna “CA_ID” não possui nenhuma linha com os caracteres AP antecedendo a algum código. Os resultados para as vagas estão sendo apresentados por meio das Tabelas 8, 9, 10 e 11.

Tabela 8: Resultado da predição dos candidatos da vaga 21886.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
66990	21886	85% - Apr	55% - Apr	64% - Apr	100% - Apr
7589	21886	46% - Rep	28% - Rep	16% - Rep	0% - Rep
34156	21886	38% - Rep	36% - Rep	33% - Rep	33% - Rep
62400	21886	58% - Apr	40% - Rep	79% - Apr	100% - Apr
25474	21886	51% - Apr	25% - Rep	50% - Apr	20% - Rep

Fonte: Elaborado pelos autores, 2025.

Conforme verificado pela Tabela 8, somente o candidato da linha um foi considerado aprovado por todos os algoritmos no teste, os da linha dois e três foram reprovados pelos quatro algoritmos, o da linha quatro foi reprovado somente pelo *Neural Network* e o da linha cinco foi aprovado pelos algoritmos *Random Forest* e KNN.

Tabela 9: Resultado da predição dos candidatos da vaga 21805.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
77403	21805	64% - Apr	53% - Apr	60% - Apr	83% - Apr
48156	21805	45% - Rep	42% - Rep	0% - Rep	78% - Apr
43111	21805	23% - Rep	21% - Rep	33% - Rep	20% - Rep
75505	21805	59% - Apr	36% - Rep	50% - Apr	50% - Apr
51717	21805	31% - Rep	38% - Rep	40% - Rep	100% - Apr
36374	21805	26% - Rep	42% - Rep	22% - Rep	100% - Apr

Fonte: Elaborado pelos autores, 2025.

Nesta predição o candidato da linha um foi aprovado por todos os algoritmos; os candidatos das linhas dois, cinco e seis foram considerados reprovados por todos os algoritmos, menos o *Tree*, onde foram avaliados como aprovados; candidato da linha três foi reprovado por todos os algoritmos e o candidato da linha quatro foi reprovado somente pelo algoritmo *Neural Network*, pelos demais, ele foi aprovado.

Tabela 10: Resultado da predição dos candidatos da vaga 21652.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
74466	21652	88% - Apr	54% - Apr	100% - Apr	100% - Apr
69030	21652	28% - Rep	21% - Rep	33% - Rep	20% - Rep
16298	21652	45% - Rep	36% - Rep	33% - Rep	33% - Rep
5359	21652	69% - Apr	62% - Apr	50% - Apr	83% - Apr

74795	21652	51% - Apr	25% - Rep	50% - Apr	20% - Rep
-------	-------	-----------	-----------	-----------	-----------

Fonte: Elaborado pelos autores, 2025.

Os resultados da vaga 21652 foram: os candidatos da linha um e quatro foram aprovados pelos quatro algoritmos aplicados; candidatos dois e três foram reprovados por todos os algoritmos e o candidato cinco foi aprovado pelos algoritmos *Random Forest* e KNN.

Tabela 11: Resultado da predição dos candidatos da vaga 21328.

CA ID	Cód. Vaga	Random Forest	Neural Network	KNN	Tree
75000	21328	55% - Apr	25% - Rep	50% - Apr	20% - Rep
75199	21328	65% - Apr	55% - Apr	50% - Apr	50% - Apr
68548	21328	65% - Apr	56% - Apr	100% - Apr	83% - Apr
4635	21328	9% - Rep	25% - Rep	0% - Rep	14% - Rep
23578	21328	50% - Rep	33% - Rep	50% - Apr	20% - Rep
5359	21328	62% - Apr	62% - Apr	50% - Apr	83% - Apr

Fonte: Elaborado pelos autores, 2025.

Na tabela 11 observou-se que o candidato da linha um foi aprovado pelos algoritmos *Random Forest* e KNN e reprovado pelo *Neural Network* e pelo *Tree*, os candidatos das linhas dois, três e seis foram aprovados por todos os quatro algoritmos, o da linha quatro foi reprovado em todas as aplicações e por último, o da linha cinco, foi aprovado somente pelo algoritmo KNN.

Após a aplicação dos modelos de aprendizado de máquina, descritos anteriormente, verificou-se com a empresa quais candidatos haviam sido aprovados para a próxima etapa: entrevista com a empresa contratante.

Tabela 12: Candidatos aprovados para a etapa de entrevista com o cliente.

Vaga	Candidatos aprovados
21886	66990, 62400 e 25474.
21805	77403, 75505 e 51717
21652	74466, 5359 e 74795
21328	75000, 75199, 68548 e 5359

Fonte: Elaborado pelos autores, 2025.

Deste modo, pela tabela 12, todos os candidatos que passaram para a próxima etapa do processo seletivo, foram considerados como aprovados, ao menos, por um dos algoritmos utilizados na aplicação dos modelos apresentados no estudo em questão. Para ilustrar quais participantes foram aprovados ao menos em uma das aplicações dos algoritmos, de cada uma das vagas, além disso, quais desses candidatos aprovados não foram selecionados para a próxima etapa, foi elaborada a tabela 13.

Tabela 13: Manipulação candidatos aprovados.

Vaga	Candidatos aprovados por pelo menos um algoritmo	Candidatos não aprovados para a próxima etapa
21886	66990, 62400 e 25474.	
21805	77403, 75505, 51717, 48156 e 36374	48516 e 36374
21652	74466, 5359 e 74795	
21328	75000, 75199, 68548, 5359 e 23578	23578

Fonte: Elaborado pelos autores, 2025.

Constatou-se, por meio da tabela 13, que em 13 das 16 vezes, ou 81% das ocasiões, em que pelo menos um dos quatro algoritmos aplicados apontou a aprovação do candidato estudado, ele foi aprovado para a próxima etapa do processo seletivo. Além disso, para a análise da assertividade dos modelos em relação ao candidato contratado, no final do processo seletivo verificou-se, junto às empresas contratantes, qual candidato foi admitido pelas companhias, estando estes apresentados na tabela 14.

Tabela 14: Candidatos aprovados no final do processo seletivo.

Vaga	Candidato contratado
21886	66990
21805	77403
21652	74466
21328	68548

Fonte: Elaborado pelos autores, 2025.

Conforme identificado as predições individuais de cada participante nas tabelas 8, 9, 10 e 11 os candidatos contratados no final do processo seletivo, descritos na tabela 04, foram os com as melhores possibilidades de efetivação após a etapa de “Filtragem dos candidatos”. Isso devido ao fato de terem apresentado percentuais de aprovação maiores ou iguais a 50% em todas as quatro aplicações dos algoritmos e além disso, quando havia outros participantes com a mesma particularidade, a maior parte de suas porcentagens eram superiores, fazendo com que a média geral dos quatro algoritmos de probabilidade de aprovação, fosse a maior do que os outros candidatos.

Tabela 15: Médias dos valores de predição de cada candidato selecionado.

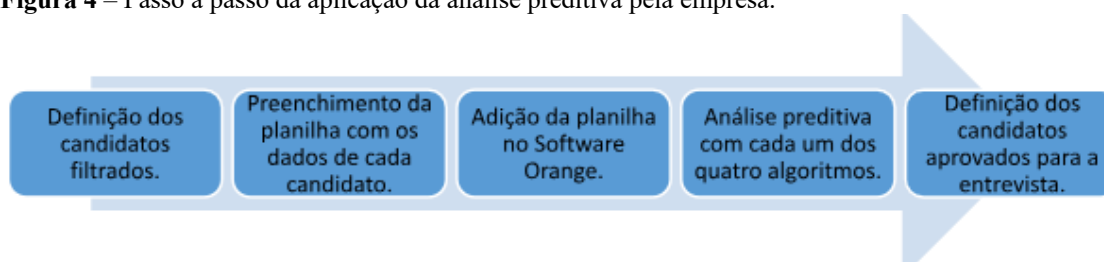
Vaga	Candidato	<i>Random Forest</i>	<i>Neural Network</i>	KNN	<i>Tree</i>	Média
21886	66990	85%	55%	64%	100%	76%
21886	62400	58%	40%	79%	100%	69%
21886	25474	51%	25%	50%	20%	37%
21886	34156	38%	36%	33%	33%	35%
21886	75897	46%	28%	16%	0%	23%
21805	77403	64%	53%	60%	83%	65%
21805	51717	31%	38%	40%	100%	52%
21805	75505	59%	36%	50%	50%	49%
21805	36374	26%	42%	22%	100%	48%
21805	48156	45%	42%	0%	78%	41%
21805	43111	23%	21%	33%	20%	24%
21652	74466	88%	54%	100%	100%	86%
21652	5359	69%	62%	50%	83%	66%
21652	16298	45%	36%	33%	33%	37%
21652	74795	51%	25%	50%	20%	37%
21652	69030	28%	21%	33%	20%	26%
21328	68548	65%	56%	100%	83%	76%
21328	5359	62%	62%	50%	83%	64%
21328	75199	65%	55%	50%	50%	55%
21328	23578	50%	33%	50%	20%	38%
21328	75000	55%	25%	50%	20%	38%
21328	4635	9%	25%	0%	14%	12%

Fonte: Elaborado pelos autores, 2025.

Conforme a tabela 15, os candidatos selecionados são os que possuem as maiores médias de predição para cada uma das vagas estudadas. Dessa forma, é reforçado a assertividade dos algoritmos aplicados não somente a etapa de pré-entrevista, mas também para a seleção dos candidatos.

Portanto, verificou-se que a aplicação dos modelos de aprendizado de máquina para as vagas destinadas a candidatos com escolaridade de ensino superior completo se mostrou mais eficiente do que para a análise preditiva das vagas com perfil de candidatos de ensino médio ou superior incompleto. Alguns fatores que podem ter auxiliado na melhor aplicabilidade dos modelos são as vagas destinadas a perfis de nível superior completo possuem candidatos com maior homogeneidade das variáveis estudadas, facilitando o reconhecimento de padrões pelos algoritmos; e o banco de dados elaborado para o estudo pode ter contemplado melhor a predição para esse tipo de vaga. Assim, a aplicação do processo sugerido seguirá os cinco passos contidos na Figura 4:

Figura 4 – Passo a passo da aplicação da análise preditiva pela empresa.



Fonte: Esta pesquisa (2021).

A definição dos candidatos filtrados inicia o processo, com a Psicóloga selecionando os perfis mais adequados para cada vaga. Em seguida, os dados são preenchidos na planilha (conforme figura 12), inseridos no software Orange Data Mining e submetidos a quatro modelos de aprendizado de máquina. Os resultados das previsões auxiliam as Psicólogas a decidirem quais candidatos avançam para a etapa de entrevista.

5. Considerações finais

Após aplicar os algoritmos às vagas e analisar os resultados, constatou-se a eficácia na filtragem de candidatos, levando à proposta de incorporar modelos de aprendizado de máquina no processo de recrutamento, usando o software Orange, sob responsabilidade do Departamento de Seleção. Para implementar essa etapa, é necessário instalar o Orange Data Mining nos computadores da equipe, software gratuito e compatível com diversos sistemas operacionais, além de capacitar os colaboradores por meio de treinamentos especializados ou recursos online disponíveis.

A nova etapa foi inserida entre a “Filtragem dos candidatos” e a “Entrevista com os candidatos selecionados” porque o sistema da empresa não gera relatórios personalizados para alimentar as planilhas da análise preditiva, exigindo coleta manual dos dados. Como a média de inscritos por vaga é superior a 50, aplicar os algoritmos antes da filtragem tornaria o processo ainda mais complexo e trabalhoso, já que o colaborador precisaria acessar cada cadastro, preencher planilhas individualmente e rodar o modelo para definir os aprovados.

Além disso, a sugestão da aplicação dos algoritmos pelo software encontra-se anteriormente a etapa de “Entrevista com os candidatos selecionados”, pois as avaliações para a aprovação dos participantes nessa etapa são baseadas em uma pesquisa qualitativa, a qual tem como condição indispensável para êxito de desempenho: a interrelação dos dados, o apontamento das experiências cotidianas dos participantes e a linguagem do senso comum (Gaskel, 2014).

Como as Psicólogas consideram, além dos dados das planilhas, aspectos como experiências, carreira, habilidades interpessoais e percepção profissional, a aplicação do algoritmo foi posicionada antes dessa etapa. Além disso, essa análise preditiva visa reduzir o número de candidatos a serem entrevistados, tornando o processo mais eficiente e menos custoso para a empresa, que envolve recursos como espaço físico, contatos telefônicos e tempo dos colaboradores.

6. Conclusão

O mercado de Recursos Humanos progrediu muito nos últimos anos, a tecnologia passou a fazer parte do dia a dia dos processos e a manipulação de dados se tornou um dos grandes diferenciais competitivos ao oferecer soluções com menos burocracia e mais assertividade. As empresas que compreendem o novo cenário e promovem atualizações constantes, acabam se tornando mais desenvolvidas e dispostas a oferecer soluções com os melhores métodos de resoluções disponíveis no mercado.

Desse modo, a consultoria abordada neste estudo possui muitos processos executados de forma manual e sem apoio de ferramentas tecnológicas no desenvolvimento das atividades, consequentemente, tomando muitos esforços de seus colaboradores e muitos recursos não necessários. Assim, com as matrizes de confusão apresentadas, verificou-se que o algoritmo *Neural Network* apresenta os melhores resultados de previsão para candidatos reprovados, já que este foi capaz de prever em 192 das 200 vezes ou 96% dos casos os profissionais que não foram contratados. Porém, mesmo sendo o melhor algoritmo de valor de precisão e para a classe de reprovação, na aprovação, o modelo foi preciso somente em 21% das vezes. Conforme podemos verificar pela figura 16, para os candidatos aprovados, o algoritmo KNN foi responsável pela precisão de acerto em 42 vezes ou 42% dos profissionais contratados, portanto, sendo esse o mais indicado para análises preditivas de candidatos a serem aprovados.

Portanto, após a implantação da nova etapa é necessário o acompanhamento e medição dos resultados. Após e é fundamental acompanhar e medir os resultados para avaliar a eficácia dos algoritmos e propor atualizações conforme a filosofia Kaizen, que valoriza melhorias contínuas por meio de pequenos ajustes sem grandes investimentos (Santos et al., 2021). Para garantir o sucesso do processo, é importante o alinhamento constante entre colaboradores e proprietários, por meio de reuniões diárias que permitam compartilhar progressos e dificuldades, possibilitando uma visão integrada do andamento da nova etapa e da empresa como um todo (Soares et al., 2007).

Além da avaliação interna com os colaboradores, foi proposta a aplicação do NPS (Net Promoter Score) para medir a satisfação dos clientes finais por meio da pergunta “Quão provável é que você nos recomendaria?”, utilizando uma escala de 1 a 10 (Mizutani, 2016). Essa métrica permite avaliar não só a nova etapa do processo, mas toda a solução de recrutamento oferecida pela empresa. Para aprimorar as análises preditivas, recomenda-se a atualização periódica da planilha-base com novos candidatos, garantindo dados atuais e consistentes para futuras avaliações.

O estudo teve dificuldades no levantamento dos dados necessários por encontrar entraves ao gerar os relatórios gerenciais, mas ainda assim foi feita uma reunião com a empresa terceira responsável pelo software e foi sugerido mudanças futuras para os sistemas. Dessa forma, possibilitando uma maior facilidade para a implantação do processo na companhia, tornando a etapa mais automatizada. Além disso, sugere-se analisar a otimização do tempo proporcionada por essa nova etapa, reduzindo as horas gastas pelos colaboradores. A aplicação desses modelos representa um avanço importante na digitalização e automação dos processos da empresa, acompanhando a crescente presença da tecnologia no ambiente empresarial.

7. Referências

1. Abiodun, O. I., et al. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11).
2. Adekitan, A. I., & Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, 24(2), 1527-1543.
3. Ajda. (2017, November 29). How to properly test models. *Orange Blog*. <https://orangedatamining.com/blog/2017/11/29/how-to-properly-test-models/>
4. Andrietta, J. M., & Miguel, P. A. C. (2007). Aplicação do programa seis sigma no Brasil: resultados de um levantamento tipo survey exploratório-descritivo e perspectivas para pesquisas futuras. *Gestão da Produção*, 14(2), 203-219.
5. Artar, M., Balcioglu, Y. S., & Erdil, O. (2024). Improving the quality of hires via the use of machine learning and an expansion of the person–environment fit theory. *Management Decision*.
6. Baldam, R., Valle, R., & Rozenfeld, H. (2014). Gerenciamento de processos de negócios – BPM: uma referência para implantação prática (1ª ed.). Elsevier.
7. Blaz. (2016). The beauty of a random forest. *Orange Blog*. <https://orangedatamining.com/blog/2016/12/22/the-beauty-of-random-forest/>
8. Bosmans, K., Hardonk, S., De Cuyper, N., & Vanroelen, C. (2016). Explaining the relation between precarious employment and mental well-being. A qualitative study among temporary agency workers. *Work*, 53(2), 249-264.
9. Buchele, G. T., et al. (2017). Métodos, técnicas e ferramentas para inovação: O uso do brainstorming no processo de design contribuindo para a inovação. *Pensamento & Realidade*, 32(1), 61-81.
10. Caetano, M. (2015). Modelos de classificação: aplicações no setor bancário (Master's thesis). Universidade Estadual de Campinas, Instituto de Matemática Estatística e Computação Científica, Campinas, SP. <http://www.repositorio.unicamp.br/handle/REPOSIP/306286>
11. De Oliveira, C., Machado, A. F., & Watanabe, C. Y. V. (2021). Utilização de dados pessoais pelas empresas: LGPD e o comportamento do consumidor com o macro modelo APCO. *Brazilian Journal of Development*, 7(6), 63580-63591.
12. Gaskell, G. (2014). Entrevistas individuais e de grupos. In M. W. Bauer & G. Gaskell (Orgs.), *Pesquisa qualitativa com texto, imagem, e som: um manual prático* (pp. 64-89). Vozes.
13. Gonçalves, C., et al. (2016). Utilização do método DMAIC para melhoria dos processos de fabricação e do produto: Estudo de caso em uma empresa do ramo metalúrgico. In VI Congresso Brasileiro de Engenharia de Produção, 2016.
14. IBGE – Instituto Brasileiro de Geografia e Estatística. (2019). Estatísticas do Cadastro Central de Empresa – CEMPRE. <https://www.ibge.gov.br/estatisticas/economicas/comercio/9016-estatisticas-do-cadastro-central-de-empresas.html?=&t=downloads>. Acesso em ago. 2025.

15. Ishak, A., Siregar, K., Ginting, R., & Afif, M. (2020). Orange software usage in data mining classification method on the dataset lenses. In IOP Conference series: materials science and engineering (Vol. 1003, No. 1, p. 012113). IOP Publishing.
16. Junqueira, M. N., et al. (2015). Utilização da ferramenta gráfico de GANTT no processo produtivo de uma empresa de equipamentos médicos de Franca-SP. Fortaleza: Enegep.
17. Kodati, S., & Vivekanandam, R. (2018). Analysis of heart disease using in data mining tool Orange and Weka. Global Journal of Computer Science and Technology.
18. Kshetri, N. (2021). Evolving uses of artificial intelligence in human resource management in emerging economies in the global South: some preliminary evidence. Management Research Review, 44(7), 970-990.
19. Longaray, A. A., et al. (2017). Proposta de mapeamento de processos usando a BPMN: estudo de caso em uma indústria da construção naval brasileira. Revista Eletrônica de Estratégia & Negócios, 10, 247-275.
20. Madhumathi, R., Dharshana, R., & Harini, P. (2017). Stimulation of plant growth in hydroponics using nutrient film technology and report generation using Orange data mining tool. Biometrics and Bioinformatics, 9(1), 1-2.
21. Matos, P. F. et al., (2009). Relatório técnico “métricas de avaliação”. Universidade Federal de Sao Carlos. Disponível em: https://www.academia.edu/download/50713968/Relatrio_Tcnico_Mtricas_de_Avaliao20161204-24828-grj1rd.pdf. Acesso em: ago. 2025.
22. Meshram, R. (2023). The role of artificial intelligence (AI) in recruitment and selection of employees in the organisation. Russian Law Journal, 11(9S), 322-333.
23. MIZUTANI, B. (2016). Implementação do net promoter score em uma empresa de varejo online. São Paulo: POLI-USP, 2016. 127f. Monografia (Graduação em Engenharia de Produção). Escola Politécnica da Universidade de São Paulo, São Paulo.
24. Mohi, Z. R. (2020). Orange data mining as a tool to compare classification algorithms. Dijlah Journal, 3(3).
25. Naik, A., & Samant, L. (2016). Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. Procedia Computer Science, 85, 662-668.
26. Nain, V., & Shyam, H. S. (2024). Empirical analysis of the role of artificial intelligence in human resources recruitment and selection. Proceedings on Engineering Sciences, 6(2), 817-826.
27. Odili, P. O., et al. (2024). The impact of artificial intelligence on recruitment and selection processes in the oil and gas industry: a review. Engineering Science & Technology Journal, 5(2), 612-638.
28. Santos, A., et al. (2021). Lean management: melhoria contínua, kaizen-muda, mura e muri, combate ao desperdício numa empresa do setor automóvel: estudo de caso. In Atas do VII Encontro Científico da UI&D (ecUI&D 21) (p. 8).
29. Silveira, R. F. (2021). Análise preditiva sobre o consumidor litigante: A inteligência artificial pode prever o litígio entre o passageiro e a companhia aérea? (Dissertação de mestrado). Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, Brasília.
30. Soares, F., et al. (2007). Adoção de Scrum em uma fábrica de desenvolvimento distribuído de software. In I Workshop de Desenvolvimento Distribuído de Software. João Pessoa.
31. Turra, C., et al. (2017). Multielement determination in orange juice by ICP-MS associated with data mining for the classification of organic samples. Information Processing in Agriculture, 4(3), 199-205.
32. Teffé, C. S., & Viola, M. (2020). Tratamento de dados pessoais na LGPD: estudo sobre as bases legais. Civilistica, 9(1), 1-38.