



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA QUÍMICA
CURSO DE GRADUAÇÃO EM ENGENHARIA QUÍMICA

JOSÉ CLAUDIO VILAR GOMES JUNIOR

**MODELO DE INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAÇÃO DO REFUGO
EM BOBINAS DE PAPEL RECICLADO A PARTIR DA VARIAÇÃO DE
GRAMATURA E UMIDADE**

Recife
2025

JOSÉ CLAUDIO VILAR GOMES JUNIOR

**MODELO DE INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAÇÃO DO REFUGO
EM BOBINAS DE PAPEL RECICLADO A PARTIR DA VARIAÇÃO DE
GRAMATURA E UMIDADE**

Trabalho de Conclusão de Curso
apresentado à Coordenação do Curso de
Graduação em Engenharia Química da
Universidade Federal de Pernambuco,
como requisito parcial para obtenção do
título de Bacharel em Engenharia Química.

Orientador(a): Prof. Dr. Felipe Pedro da Costa Gomes

Recife

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Gomes Júnior, José Cladio Vilar.

Modelo de inteligência artificial para identificação do refugo em bobinas de papel reciclado a partir da variação de gramatura e umidade / José Cladio Vilar Gomes Júnior. - Recife, 2025.

65 p : il., tab.

Orientador(a): Felipe Pedro da Costa Gomes

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Engenharia Química - Bacharelado, 2025.

Inclui referências.

1. Predição de refugo. 2. Gradient Boosting. 3. Random Forest. 4. MLPClassifier. 5. Controle de qualidade. I. Gomes, Felipe Pedro da Costa. (Orientação). II. Título.

620 CDD (22.ed.)

JOSÉ CLAUDIO VILAR GOMES JUNIOR

Modelo de inteligência artificial para identificação do refugo em bobinas de papel reciclado a partir da variação de gramatura e umidade

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Graduação em Engenharia Química da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Engenharia Química.

Aprovado em: 09/04/2025.

BANCA EXAMINADORA

Profº. Dr. Felipe Pedro da Costa Gomes (Orientador)
Universidade Federal de Pernambuco

Profº. Dr. Fábio Machado Cavalcanti (Examinador Interno)
Universidade Federal de Pernambuco

Dr. Alan Gomes da Câmara (Examinador Interno)
Universidade Federal de Pernambuco

Dedico este trabalho à minha família.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder força, sabedoria e saúde ao longo desta jornada.

Agradeço à minha esposa, pelo apoio incondicional, paciência e incentivo em cada etapa desta jornada. Sua compreensão, força e exemplo de dedicação foram fundamentais para que eu pudesse seguir adiante, mesmo diante dos desafios.

Aos meus pais e à minha irmã, que me ensinaram desde cedo a importância da coragem, da resiliência e de enfrentar a vida com determinação e dignidade. Sem o exemplo e o apoio de vocês, esta conquista não seria possível.

Ao meu orientador, pela orientação e pelos ensinamentos que contribuíram diretamente para a realização deste Trabalho de Conclusão de Curso.

À Universidade Federal de Pernambuco, pela oportunidade de aprendizado e crescimento pessoal e profissional ao longo desta formação.

Não poderia deixar de agradecer ao meu fiel companheiro Alfredo, que me ofereceu suporte emocional nos momentos mais desafiadores desta caminhada. Por todas as vezes em que deitou a cabeça nos meus pés enquanto eu redigia este trabalho, trazendo-me calma e conforto, e por cada momento em que trouxe um brinquedo, como se soubesse exatamente que eu precisava de uma pausa para reencontrar a leveza e a esperança. Sua companhia fez toda a diferença para que este TCC fosse concluído.

Inteligência é a habilidade de se adaptar à
mudança (Hawking, 1996).

RESUMO

A rápida identificação de falhas é essencial para reduzir perdas e aumentar a eficiência nos processos industriais. Nesse contexto, ferramentas baseadas em Inteligência Artificial têm se destacado como alternativas promissoras para o controle de qualidade em tempo real. Este trabalho teve como objetivo desenvolver e validar modelos de classificação binária baseados em algoritmos de aprendizado supervisionado para prever a tendência de refugo em bobinas de papel reciclado, utilizando dados de gramatura e umidade coletados ao longo do processo produtivo. A pesquisa foi conduzida com base em um banco de dados histórico composto por mais de seis mil lotes produzidos. Foram aplicados e comparados três algoritmos: *Gradient Boosting*, *Random Forest* e *MLPClassifier*. Todos os modelos apresentaram acurácia superior a 97%, com destaque para o *MLPClassifier*, que obteve o melhor desempenho global ao combinar elevados níveis de acurácia, precisão, revocação e maior sensibilidade para rever lotes desclassificados. A análise das matrizes de confusão revelou, entretanto, que os modelos apresentaram baixa taxa de identificação dos lotes refugados, o que indica limitações na detecção da classe minoritária. Essa dificuldade pode estar relacionada ao desbalanceamento da base de dados e à perda de detalhamento provocada pelo uso de estatísticas agregadas. Apesar das limitações observadas, os resultados demonstram o potencial dos modelos de aprendizado de máquina como apoio ao controle de qualidade, contribuindo para o avanço da automação, a redução de desperdícios e o aumento da confiabilidade nos processos industriais.

Palavras-chave: predição de refugo; *Gradient Boosting*; *Random Forest*; *MLPClassifier*; controle de qualidade.

ABSTRACT

The rapid identification of failures is essential to reduce losses and increase efficiency in industrial processes. In this context, Artificial Intelligence-based tools have emerged as promising alternatives for real-time quality control. This study aimed to develop and validate binary classification models based on supervised learning algorithms to predict the scrap tendency in recycled paper rolls, using grammage and moisture data collected throughout the production process. The research was conducted using a historical database comprising over six thousand produced batches. Three algorithms were applied and compared: Gradient Boosting, Random Forest, and MLPClassifier. All models achieved accuracy above 97%, with the MLPClassifier standing out for its overall performance, combining high levels of accuracy, precision, recall, and greater sensitivity in identifying rejected batches. However, the analysis of the confusion matrices revealed a low detection rate of scrap batches, indicating limitations in identifying the minority class. This challenge may be related to data imbalance and the loss of detail caused by the use of aggregated statistics. Despite the observed limitations, the results demonstrate the potential of machine learning models to support quality control, contributing to the advancement of automation, reduction of waste, and increased reliability in industrial processes.

Keywords: scrap prediction; Gradient Boosting; Random Forest; MLPClassifier; quality control.

LISTA DE FIGURAS

Figura 1 -	Planta de descontaminação do preparo de massa para produção de papel reciclado	18
Figura 2 -	Etapas de depuração grossa do preparo de massa para produção de papel reciclado	19
Figura 3 -	Etapas da máquina de papel	21
Figura 4 -	Máquina de teste de compressão (RCT)	26
Figura 5 -	Mullen Tester	27
Figura 6 -	Esquema do funcionamento do algoritmo <i>Gradient Boosting</i>	33
Figura 7 -	Esquema do funcionamento do algoritmo <i>Random Forest</i>	34
Figura 8 -	Esquema de uma unidade básica de processamento de uma rede neural artificial	35
Figura 9 -	Modelo de matriz de confusão	37
Figura 10 -	Distribuição da classificação final dos lotes avaliados	45
Figura 11 -	Distribuição dos lotes desclassificados segundo critérios de não conformidade nos ensaios de RCT e Mullen	46
Figura 12 -	Comparação da distribuição de gramatura – Lote aprovado vs. refugado	47
Figura 13 -	Comparação da distribuição de umidade – Lote aprovado vs. refugado	48
Figura 14 -	Matriz de correlação dos parâmetros analisados das bobinas de papel reciclado	49
Figura 15 -	Matriz de confusão do modelo <i>Gradient Boosting</i>	52
Figura 16 -	Matriz de confusão do modelo <i>MLPClassifier</i>	52
Figura 17 -	Matriz de confusão do modelo <i>Random Forest</i>	52
Figura 18 -	Importância das variáveis no modelo <i>Gradient Boosting Classifier</i> aplicado à classificação dos lotes de papel reciclado	54
Figura 19 -	Importância das variáveis no modelo <i>Random Forest Classifier</i> aplicado à classificação dos lotes de papel reciclado	55

LISTA DE ABREVIATÖES

AED	Análise exploratória de dados
DP	Desvio padrão
IA	Inteligência Artificial
IoT	<i>Internet of things</i>
MLP	<i>Multilayer Perceptron</i>
RCT	<i>Ring crush test</i>
RF	<i>Random forest</i>
RNA	Rede neural artificial

SUMÁRIO

1. INTRODUÇÃO	13
1.1.1 Objetivo Geral	15
1.1.2 Objetivos Específicos	15
2 REVISÃO DA LITERATURA	16
2.1 PRODUÇÃO DE PAPEL RECICLADO	16
2.1.1 Preparo de massa	16
2.1.2 Refinação	20
2.1.3 Etapas na máquina de papel.....	20
2.2 CONTROLE DE QUALIDADE EM PAPEL RECICLADO	23
2.2.1 Parâmetros de qualidade no papel reciclado	24
2.3 APLICAÇÕES DE INTELIGÊNCIA ARTIFICIAL NA INDÚSTRIA 4.0	28
2.4 MODELOS DE CLASSIFICAÇÃO PARA PREDIÇÃO DE REFUGO	30
2.4.1 Tipos de aprendizado de máquina	30
2.4.2 Modelos supervisionados de classificação.....	32
2.4.3 Avaliação de modelos de classificação supervisionada	36
3 METODOLOGIA	40
3.1 TIPO DE PESQUISA.....	40
3.2 COLETA E PRÉ-PROCESSAMENTO DOS DADOS	40
3.3 DESENVOLVIMENTO E TREINAMENTO DOS MODELOS.....	42
3.4 VALIDAÇÃO E TESTE	43
4 RESULTADOS.....	44
4.2 ANÁLISE EXPLORATÓRIA DOS DADOS (AED)	44
4.2.1 Distribuição das variáveis de gramatura e umidade	46
4.3 ANÁLISE DE CORRELAÇÃO	48
4.4 AVALIAÇÃO DOS MODELOS DE CLASSIFICAÇÃO	50

4.5 ANÁLISE DA IMPORTÂNCIA DAS VARIÁVEIS NOS MODELOS DE CLASSIFICAÇÃO.....	54
4.5.1 <i>Gradient Boosting Classifier</i>	54
4.5.2 <i>Random Forest Classifier</i>	55
4.5.4 Comparação das importâncias das variáveis	56
5 CONCLUSÃO	57
REFERÊNCIAS.....	59

1. INTRODUÇÃO

A indústria de papel e celulose passou por profundas transformações nas últimas décadas, impulsionada por avanços tecnológicos voltados à automação, à eficiência produtiva e à sustentabilidade. O Brasil se destaca globalmente nesse setor, sendo o maior exportador mundial de celulose e o segundo maior produtor, atrás apenas dos Estados Unidos (IBÁ, 2024). Apesar do desempenho robusto, o setor enfrenta desafios relacionados a altos custos operacionais, variabilidade na qualidade do produto final e limitações no controle em tempo real de variáveis críticas, como gramatura e umidade (Merlin; Caleman; Padgett, 2024; He *et al.*, 2022).

A produção de papel reciclado, em particular, apresenta instabilidades decorrentes da falta de controle sobre a matéria-prima, o papel reciclado coletado. Variações na gramatura e umidade são responsáveis por aproximadamente 1,7% do refugo total, afetando diretamente os custos e a produtividade (Devi *et al.*, 2023). A imprevisibilidade desses parâmetros compromete a eficiência operacional e a sustentabilidade do processo produtivo.

A crescente digitalização dos processos produtivos em diferentes segmentos industriais tem impulsionado o uso de sistemas inteligentes baseados em inteligência artificial (IA), especialmente no contexto da chamada Indústria 4.0. Em áreas como manufatura, agronegócio e química de processos, a IA tem sido aplicada para prever falhas, ajustar parâmetros em tempo real e apoiar decisões com base em grandes volumes de dados operacionais (Raschka; Mirjalili, 2019). Esse movimento também alcança o setor papeleiro, onde modelos preditivos podem antecipar variações de qualidade e minimizar perdas antes que comprometam os resultados finais da produção.

Entre as diversas abordagens da IA, o aprendizado de máquina supervisionado tem se destacado como uma estratégia eficaz para problemas em que há disponibilidade de dados históricos rotulados, como testes físicos de qualidade. Essa classe de modelos permite construir algoritmos capazes de aprender padrões complexos a partir de variáveis de entrada e prever resultados futuros com elevada precisão. Modelos como *Random Forest*, *Gradient Boosting* e redes neurais artificiais (*MLPClassifier*) vêm sendo amplamente utilizados na indústria por sua robustez, capacidade de generalização e flexibilidade para lidar com conjuntos de dados estruturados (Han *et al.*, 2011; James *et al.*, 2021).

A utilização de modelos supervisionados se justifica especialmente em ambientes nos quais os resultados dos testes de qualidade como RCT e Mullen são previamente conhecidos, possibilitando que algoritmos aprendam a mapear variáveis operacionais, como gramatura e umidade, com os resultados finais do produto. Essa abordagem está alinhada aos princípios da Indústria 4.0, que defende o uso de tecnologias inteligentes para aumentar a eficiência, reduzir perdas e fortalecer a competitividade do setor (Srinivasan *et al.*, 2024).

Apesar dos avanços, ainda são escassos os estudos voltados especificamente à aplicação de modelos de aprendizado supervisionado na predição de testes físicos em papel reciclado, sobretudo considerando variáveis captadas em tempo real nas linhas de produção. Nesse sentido, este trabalho se insere como uma contribuição prática e acadêmica, ao propor a integração entre dados operacionais e inteligência artificial com foco na redução do refugo e na otimização da qualidade industrial.

1.1 OBJETIVOS

Nesta seção são apresentados o objetivo geral e os objetivos específicos que nortearam o desenvolvimento da pesquisa, alinhados à proposta de aplicar técnicas de inteligência artificial para aprimorar o controle de qualidade no processo produtivo de papel reciclado.

1.1.1 Objetivo Geral

Desenvolver e validar um modelo de classificação, baseado em algoritmos de Inteligência Artificial, capaz de prever com alta acurácia a tendência de refugo em bobinas de papel reciclado, a partir da análise das variáveis gramatura e umidade durante o processo produtivo.

1.1.2 Objetivos Específicos

- Coletar e tratar dados históricos de gramatura e umidade de bobinas de papel reciclado, consolidando estatísticas descritivas para cada lote produzido.
- Treinar e validar modelos de classificação supervisionada, utilizando algoritmos como *Gradient Boosting*, *Random Forest* e *MLPClassifier*, para prever a tendência de refugo das bobinas.
- Avaliar o desempenho dos modelos de classificação, com base em métricas como acurácia, precisão, revocação e matriz de confusão, selecionando o mais adequado para a aplicação.
- Analisar a importância das variáveis nos modelos de classificação para identificar quais parâmetros possuem maior influência na predição do refugo.

2 REVISÃO DA LITERATURA

Esta seção apresenta conceitos, teorias e estudos relacionados ao desenvolvimento de modelos de Inteligência Artificial aplicados ao controle de qualidade na indústria de papel reciclado. A revisão da literatura foi estruturada para fornecer uma base teórica sólida, abordando desde o processo de produção do papel reciclado e os desafios relacionados ao controle de qualidade até a aplicação de técnicas de aprendizado de máquina, com foco nos modelos de classificação supervisionada utilizados para a predição de refugo.

2.1 PRODUÇÃO DE PAPEL RECICLADO

A reciclagem de papel consiste na produção de novas folhas a partir de papéis recicláveis, como rebarbas, aparas e artefatos descartados durante ou após o consumo. Esses materiais, provenientes do processo industrial ou do descarte final pelos usuários, são reaproveitados, minimizando o impacto ambiental e reduzindo a demanda por matéria-prima virgem (Moraes *et al.*, 2015).

Presente em diversos aspectos do cotidiano, o papel reciclado tem um destaque na indústria de base, e sua fabricação abrange um conjunto integrado de etapas. A produção de papel reciclado inicia-se com a recepção e triagem das aparas, seguida por processos de limpeza, peneiramento, refino, lavagem e desaguamento. Essas etapas têm como objetivo remover impurezas e preparar adequadamente a polpa (Bajpai, 2015; Kühn *et al.*, 2024).

Durante o ciclo produtivo, o monitoramento constante de variáveis operacionais como gramatura e umidade é fundamental para assegurar a qualidade do produto final. A coordenação eficiente das etapas de formação, prensagem, secagem e acabamento não apenas otimiza o controle sobre as características do papel, mas também contribui para a redução de resíduos e o aumento da eficiência energética (He *et al.*, 2022).

2.1.1 Preparo de massa

A produção de papel reciclado tem início na central de aparas, onde a matéria-prima é recebida e submetida a processos de amostragem e classificação. Amostras

representativas são coletadas e analisadas visualmente para identificação dos tipos de papel e a presença de contaminantes, como plásticos e metais. Também é avaliado o teor de umidade, fator essencial para determinar a qualidade e adequação do material. Com base nesses critérios, as aparas são classificadas conforme as diretrizes da NBR 15483, apresentadas no Quadro 1.

Quadro 1 – Classificação de aparas de papel e papelão ondulado

Material	Descrição	Especificação
Aparas de refil de papelão ondulado	Aparas de produtos de papelão ondulado de fibra virgem ou reciclada, resultantes dos processos de produção do papelão ondulado e caixas não utilizadas no mercado, sem cola insolúvel em água e sem grampo, podendo apresentar refil de capa branca.	Teor máximo de umidade: 13% Teor máximo de impurezas: 0% Teor máximo de materiais proibitivos: 0%
Aparas de papelão ondulado I	Aparas de produtos de papelão ondulado de fibra virgem ou reciclada marrom e sem outros papéis que não sejam papelão ondulado. É permitida a presença de fita adesiva, <i>hot melt</i> , grampo e etiquetas provenientes da própria embalagem.	Teor máximo de umidade: 15% Teor máximo de impurezas: 3% Teor máximo de materiais proibitivos: 0%
Aparas de papelão ondulado II	Aparas de produtos de papelão ondulado de fibra virgem ou reciclada, com até 5% (em massa) de outros papéis que não sejam papelão ondulado. É permitida a presença de fita adesiva, <i>hot melt</i> , grampo e etiquetas provenientes da própria embalagem, capa branca.	Teor máximo de umidade: 15% Teor máximo de impurezas: 3% Teor máximo de materiais proibitivos: 1%
Aparas de papelão ondulado III	Aparas de produtos de papelão ondulado de fibra virgem ou reciclada, com até 20% (em massa) de outros papéis que não sejam papelão ondulado. É permitida a presença de fita adesiva, <i>hot melt</i> , grampo e etiquetas provenientes da própria embalagem, capa branca.	Teor máximo de umidade: 15% Teor máximo de impurezas: 5% Teor máximo de materiais proibitivos: 3%

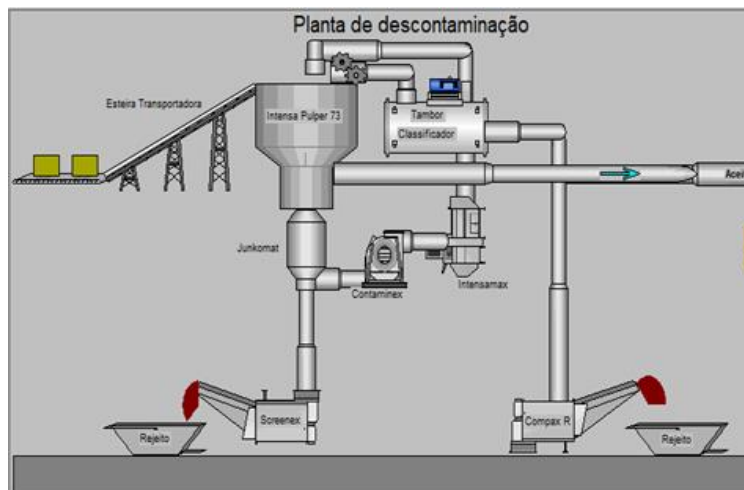
Fonte: NBR 15483, 2009.

Os fardos de aparas que não se enquadram nas especificações estabelecidas no Quadro 1 são rejeitados. Esses resíduos, originados de embalagens e caixas, são coletados de diferentes fontes, como indústrias e cooperativas e também do refugo gerado internamente no processo produtivo. As aparas podem ser classificadas como pré-consumo, quando recolhidas diretamente nas fábricas antes de serem utilizadas pelos consumidores finais, ou pós-consumo, quando provenientes de materiais descartados pelos usuários finais (NBR ISO 14021, 2004; Alvarenga; Reis, 2018).

O objetivo principal da central de aparas é desagregar, depurar e selecionar fibras, de modo a garantir a produção de papel reciclado com a qualidade desejada e

o menor nível possível de impurezas (NBR 15483, 2009). Na Figura 1 estão representados os principais processos realizados na Central de Aparas.

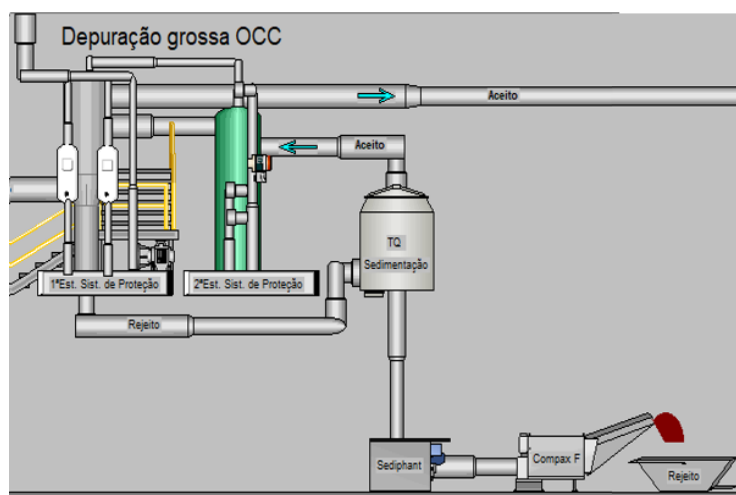
Figura 1 - Planta de descontaminação do preparo de massa para produção de papel reciclado



Fonte: O autor (2025).

A primeira etapa na central de aparas consiste no processo de desagregação e descontaminação das fibras, conforme ilustrado na Figura 1. As aparas, no equipamento *intensa pulper*, são misturadas com água e agitadas mecanicamente por um rotor que modifica a estrutura do papel, desagregando as fibras e formando uma massa homogênea. Durante essa etapa, materiais indesejados, como grampos, plásticos e outras impurezas maiores, começam a ser separados. A retirada de contaminantes, como fios de aço e náilon, protege os equipamentos subsequentes, garantindo maior eficiência e durabilidade dos equipamentos (Alvarenga; Reis, 2018). Contaminantes de tamanho menor são retirados na etapa de depuração grossa, apresentada na Figura 2.

Figura 2 – Etapa de depuração grossa do preparo de massa para produção de papel reciclado



Fonte: O autor (2025).

Na etapa de depuração grossa, Figura 2, a massa de papel é processada para a remoção de impurezas de maior tamanho, que poderiam comprometer a qualidade do papel reciclado. O funcionamento dessa etapa baseia-se em princípios mecânicos de separação, utilizando equipamentos como depuradores de alta consistência, peneiras vibratórias ou rotativas, ciclones e removedores de contaminantes pesados. Os depuradores de alta consistência operam com massas de papel de maior densidade, empregando rotores ou peneiras para reter partículas maiores. As peneiras vibratórias ou rotativas realizam a separação por meio de movimentos mecânicos, enquanto os ciclones promovem a separação de partículas densas, como areia e metais, por força centrífuga. Já os removedores de contaminantes pesados atuam com fluxo descendente ou também por força centrífuga, isolando materiais indesejados de maior densidade presentes na pasta (Campos 2011).

A última etapa na Central de Aparas envolve a classificação das fibras longas e curtas. Esse processo é realizado por fracionadores que utilizam peneiras e fendas para separar as fibras conforme seu comprimento. As fibras longas são essenciais para a resistência do papel, enquanto as fibras curtas contribuem para a formação de uma superfície lisa. Após a primeira separação, as fibras longas passam por uma nova etapa destinada à recuperação das fibras curtas que ainda permanecem misturadas. Em seguida, o material segue para tanques de estocagem, onde aguarda o próximo passo na produção do papel (Campos 2011).

2.1.2 Refinação

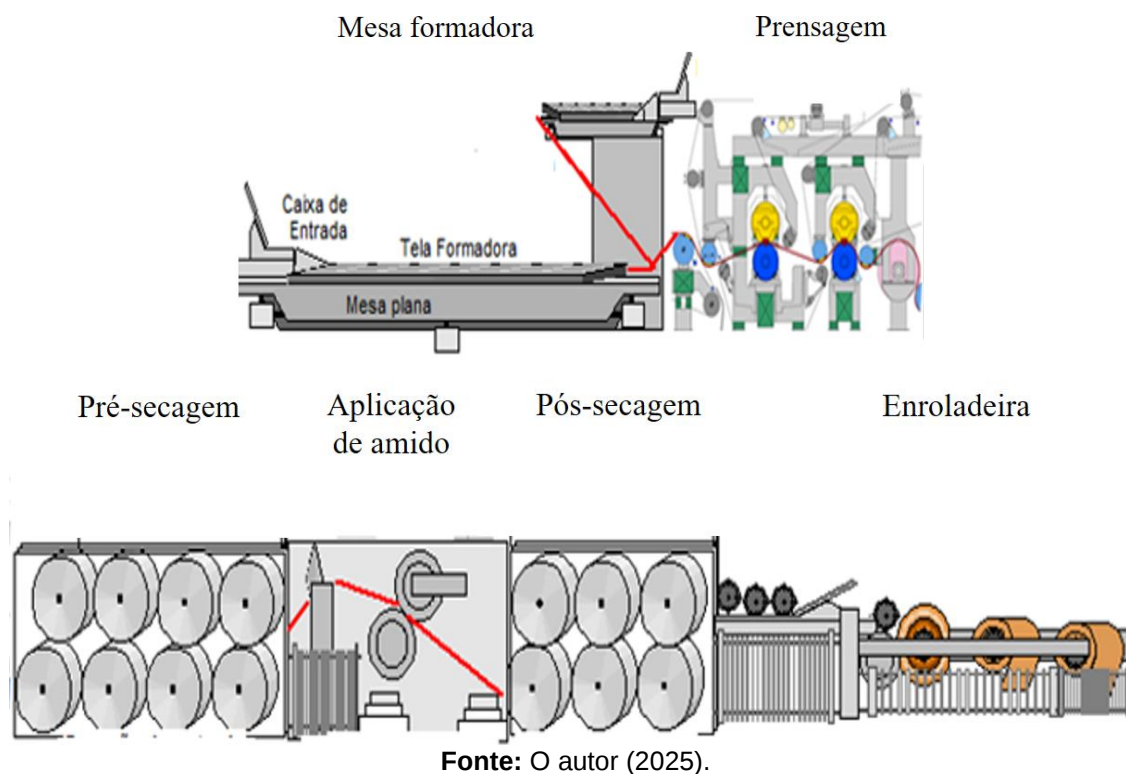
A refinação da massa é um tratamento mecânico essencial, sendo seu principal objetivo o aprimoramento da capacidade das fibras de se entrelaçarem, possibilitando a formação de uma folha de papel homogênea e resistente a esforços mecânicos. Esses esforços podem ocorrer durante a produção da folha de papel no decorrer da máquina, nas etapas de acabamento, ou ainda durante a aplicação final do produto (Alvarenga; Reis, 2018).

Além de melhorar a união das fibras, a refinação também permite modificar suas características para atender às propriedades desejadas no papel final. Durante esse processo, as fibras podem ser encurtadas, hidratadas ou fibriladas, o que contribui para o desenvolvimento de propriedades como resistência física, absorvência, porosidade e atributos ópticos (Campos 2011).

2.1.3 Etapas na máquina de papel

A máquina de papel é um equipamento essencial na indústria de papel e celulose, sendo responsável pela transformação da polpa/massa de papel em folhas contínuas de papel. Este processo ocorre em diversas etapas, desde a formação da folha até a rebobinagem final. As etapas principais incluem a formação da folha, prensagem, secagem, adição de amido (a depender do tipo de papel) e enrolamento. A eficiência e qualidade do papel final dependem diretamente do controle e execução dessas fases (Campos, 2011; Alvarenga; Reis, 2018).

Abaixo, na Figura 3, estão ilustrados os principais processos relacionados a cada etapa da máquina de papel.

Figura 3 - Etapas da máquina de papel

Fonte: O autor (2025).

2.1.3.1 Mesa formadora

A formação da folha de papel é uma das etapas mais importantes do processo produtivo, pois define as características estruturais do produto final. O processo se inicia com a chegada da massa de papel à caixa de entrada da máquina de papel. A polpa é distribuída uniformemente sobre a tela formadora, onde ocorre a drenagem inicial da água, formando uma folha úmida e ainda frágil (Alvarenga; Reis, 2018).

Nesta fase, a consistência da massa de papel varia entre 1,5% e 2%, o que permite que a fibra se espalhe de maneira homogênea. A água excedente é removida por gravidade e por sucção, enquanto a folha se move em alta velocidade sobre a tela formadora. A uniformidade da distribuição das fibras busca garantir as propriedades físicas e mecânicas do papel, como resistência e uniformidade superficial (Campos, 2011).

2.1.3.2 Prensagem

Após a formação inicial, a folha passa pela seção de prensagem, cujo objetivo é remover o excesso de água e aumentar a densidade e resistência do papel. A

prensagem ocorre entre cilindros de alta pressão, comprimindo a folha e forçando a saída da água através dos espaços entre as fibras (Campos, 2011; Alvarenga; Reis, 2018).

Esse processo, além de reduzir o teor de umidade, melhora a ligação entre as fibras, conferindo maior integridade à estrutura do papel. São utilizadas, ainda, prensas de sucção, que combinam vácuo e pressão mecânica para otimizar a remoção de água, reduzindo o consumo de energia térmica nas etapas subsequentes de secagem (Campos, 2011; Alvarenga; Reis, 2018).

2.1.3.3 Pré-secagem

Na fase de pré-secagem, a folha prensada passa por uma série de cilindros aquecidos que promovem a evaporação gradual da umidade, ainda em excesso. Esta etapa objetiva estabilizar a estrutura do papel e evitar deformações. A temperatura e a velocidade dos cilindros são controladas para evitar o ressecamento excessivo ou a quebra da folha (Campos, 2011; Alvarenga, 2018).

2.1.3.4 Aplicação de amido

A adição de amido ocorre após a etapa de pré-secagem e tem como objetivo melhorar as propriedades mecânicas e superficiais do papel. Embora essa etapa não esteja presente em todas as máquinas de papel, na unidade industrial analisada neste estudo, a aplicação de amido é realizada em todos os tipos de papel produzidos, por meio do processo de colagem superficial (*size press*). Essa aplicação confere maior resistência ao papel, reduz a penetração de líquidos e melhora sua capacidade de impressão (Campos, 2011).

O amido atua como agente de reforço, promovendo melhor ligação entre as fibras e aumentando a rigidez do material. Além disso, esse tratamento contribui para a redução da formação de poeira durante as etapas de impressão e corte, favorecendo o desempenho e a qualidade do produto final (Campos, 2011).

2.1.3.5 Pós-secagem

Após a adição de amido, o papel passa por uma nova série de cilindros secadores, onde ocorre a remoção da umidade residual, reduzindo-a para níveis entre 7% e 9%. Esse teor de umidade é considerado ideal para garantir a estabilidade dimensional do papel e evitar problemas de deformação durante o armazenamento e uso (Campos, 2011; Alvarenga 2018).

Durante a pós-secagem, o papel adquire suas propriedades finais de resistência, flexibilidade e brilho. O controle assertivo da temperatura e do tempo de secagem é importante para evitar o endurecimento excessivo do papel, o que poderia comprometer sua qualidade (Alvarenga; Reis, 2018).

2.1.3.6 Enroladeira

A seção de enrolamento transforma a folha em unidades finitas para facilitar o processamento e uso. A enroladeira produz bobinas de grande diâmetro e com densidade uniforme, garantindo a integridade do papel até o corte ou acabamento. Durante essa fase, o papel seco é enrolado em rolo jumbo, que podem alcançar centenas de metros, sendo preparados para transporte ou armazenamento sem danos ou deformações (Campos, 2011).

Na fase final, a bobina jumbo passa pela rebobinadeira, que refila e corta o papel de acordo com as especificações do cliente. Esse equipamento ajusta a largura e o diâmetro das bobinas, eliminando bordas irregulares e garantindo uniformidade no produto final (Campos, 2011; Alvarenga; Reis, 2018)

2.2 CONTROLE DE QUALIDADE EM PAPEL RECICLADO

A indústria papeleira desempenha um papel fundamental na economia global, mas enfrenta desafios significativos relacionados ao controle de qualidade, especialmente na produção de papel reciclado. A crescente demanda por produtos sustentáveis tem impulsionado o reaproveitamento de fibras recicladas, no entanto, a variabilidade da matéria-prima e das condições de produção impactam diretamente a qualidade final do produto (Righetti *et al.*, 2005; Schroth *et al.*, 2021).

O processo produtivo de papel reciclado envolve diversas etapas, desde a triagem da matéria-prima até o acabamento do produto final. Fatores como a heterogeneidade das aparas, a presença de contaminantes e a instabilidade nos processos de refino e secagem dificultam o alcance de padrões de qualidade consistentes (Sousa *et al.*, 2016).

O controle de qualidade é realizado tradicionalmente por meio de ensaios laboratoriais e inspeções visuais, com foco na avaliação de propriedades físicas e mecânicas essenciais para a utilização do papel reciclado em diferentes aplicações (Borba, 2012). Porém, essas práticas apresentam limitações quanto à detecção precoce de desvios durante o processo produtivo, o que pode gerar aumento no índice de refugo e elevação dos custos operacionais (Pereira, 2023).

Dentre as principais causas de refugo na produção de papel reciclado, destacam-se as variações excessivas de gramatura e umidade, que comprometem a resistência física das bobinas e resultam na reprovação nos testes laboratoriais de RCT (*Ring Crush Test*) e Mullen (He *et al.*, 2022; Devi *et al.*, 2023).

Com o avanço das tecnologias de automação industrial e Inteligência Artificial, novas abordagens têm sido aplicadas ao controle de qualidade, promovendo monitoramento em tempo real, análise preditiva e intervenções preventivas que minimizam o risco de refugo e aumentam a eficiência produtiva (Zhang *et al.*, 2023; Srinivasan *et al.*, 2024).

2.2.1 Parâmetros de qualidade no papel reciclado

O desempenho e a aceitação do papel reciclado no mercado dependem diretamente de alguns parâmetros de qualidade críticos, que refletem a integridade estrutural e a funcionalidade do produto final. No presente estudo, os principais parâmetros monitorados foram gramatura, umidade, *Ring Crush Test* (RCT) e Mullen. Esses indicadores são essenciais para garantir a conformidade técnica das bobinas de papel reciclado e foram considerados variáveis-chave na modelagem preditiva realizada.

2.2.1.1 Gramatura

A gramatura refere-se ao peso do papel por unidade de área, geralmente expresso em gramas por metro quadrado (g/m²). É um parâmetro fundamental que influencia as propriedades físicas e mecânicas do papel, como rigidez, resistência mecânica e capacidade de impressão (Silva, 2023).

Variações significativas de gramatura ao longo da bobina podem gerar zonas de fraqueza ou excesso de material, resultando em falhas durante o uso final ou no processo de conversão (Cardoso *et al.*, 2012). Conforme Devi *et al.* (2023), manter a gramatura dentro dos limites especificados é essencial para garantir homogeneidade e desempenho mecânico adequado.

2.2.1.2 Umidade

A umidade do papel representa o conteúdo de água presente no material, normalmente expresso em percentual (%) e pode ser determinada pela Equação 1:

$$Umidade = \frac{(P_{inicial} - P_{seco})}{P_{inicial}} * 100 \quad (1)$$

Onde $P_{inicial}$ é o peso da amostra antes da secagem e o P_{seco} é o peso da amostra depois da secagem.

Esse parâmetro afeta diretamente a estabilidade dimensional, resistência mecânica e propriedades de impressão do papel reciclado (Borba, 2012; Pereira, 2023). Níveis elevados ou flutuações excessivas de umidade comprometem a resistência à compressão e a integridade do papel durante o rebobinamento e armazenamento, além de aumentar o risco de deformações e falhas nos ensaios de qualidade (Sousa *et al.*, 2016).

2.2.1.3 Ring Crush Test

O Ring Crush Test (RCT) mede a resistência do papel à compressão anular, ou seja, sua capacidade de suportar forças de esmagamento ao longo da direção da fibra (Cardoso *et al.*, 2012) e pode ser determinado pela Equação 2, onde F é força

máxima aplicada até a ruptura dada em Newtons N e L a largura da amostra dada em metros (m).

$$RCT = \frac{F}{L} \quad (2)$$

Este teste é fundamental para avaliar a resistência estrutural de papéis reciclados utilizados em embalagens, especialmente aquelas submetidas a empilhamento e armazenamento em ambientes com variações de umidade (Devi *et al.*, 2023). Valores abaixo do limite estabelecido nos ensaios de RCT indicam fragilidade e são causa recorrente de reprovação de bobinas (Borba, 2012).

O equipamento que mede o RCT do papel reciclado é a máquina de teste de compressão, apresentada na Figura 4.

Figura 4 – Máquina de teste de compressão (RCT)



Fonte: Arquivo pessoal (2025).

O método de ensaio segue as diretrizes da ABNT NBR NM-ISO 12192:2013, que estabelece que a amostra de papel deve ser cortada em tiras e moldada em formato cilíndrico, por meio de um dispositivo metálico denominado anel de

compressão. Posteriormente, o cilindro é posicionado na máquina de compressão vertical que aplica uma força crescente até que ocorra a deformação ou o colapso da amostra (NBR 12192, 2013).

2.2.1.4 Teste de Mullen

O Teste de Mullen avalia a resistência à ruptura do papel quando submetido a pressão hidrostática. Esse ensaio determina a pressão máxima que o papel pode suportar antes de romper, sendo expresso em kilopascal (kPa) ou libra-força por polegada quadrada (psi) (Borba, 2012).

O Mullen é especialmente relevante para verificar a resistência à perfuração e impactos durante o transporte e manuseio do papel reciclado (Sousa *et al.*, 2016). A não conformidade nesse teste indica baixa resistência à pressão e é um critério de refugo na avaliação da qualidade das bobinas (Devi *et al.*, 2023).

O equipamento que mede o Mullen do papel reciclado é o Mullen Tester, apresentada na Figura 5.

Figura 5 – Mullen Tester



Fonte: Arquivo pessoal (2025).

O método de ensaio segue as diretrizes estabelecidas na ABNT NBR 7755:1983, que define que a amostra de papel deve ser fixada no Mullen Tester que é equipado com um diafragma de borracha. A pressão é aplicada hidraulicamente através do diafragma até que ocorra a ruptura do papel. O valor da pressão no momento da ruptura é registrado como o resultado do teste (NBR 7755, 1983).

2.3 APLICAÇÕES DE INTELIGÊNCIA ARTIFICIAL NA INDÚSTRIA 4.0

A Indústria 4.0 representa um marco na evolução da manufatura, caracterizada pela fusão de tecnologias digitais com processos industriais, resultando em fábricas inteligentes e altamente automatizadas. Esse conceito integra novas abordagens como a aplicação de sistemas ciberfísicos, Internet das Coisas (IoT) e Inteligência Artificial (IA), permitindo que máquinas, sistemas e operadores trabalhem de forma integrada e interativa, otimizando o desempenho e reduzindo desperdícios (Santos, 2021). A introdução dessas tecnologias na cadeia produtiva tem impulsionado a automação e facilitado o monitoramento em tempo real, proporcionando respostas rápidas a variações de processos e melhorando a competitividade das indústrias (Souza; Silva, 2024).

O conceito de Indústria 4.0 abrange a digitalização completa das operações industriais, promovendo uma transformação que vai além da automação tradicional. A manufatura inteligente utiliza sensores e redes interconectadas para coletar dados, analisar tendências e prever falhas antes que elas ocorram (Kagermann, Wahlster e Helbig, 2013). Dessa forma, as indústrias são capazes de se adaptar rapidamente às demandas do mercado, garantindo maior eficiência e redução de custos operacionais. Estudos demonstram que a integração de IA em processos industriais tem viabilizado ganhos expressivos de produtividade, ao mesmo tempo em que promove práticas sustentáveis e otimizadas (Zhang *et al.*, 2023).

A aplicação de IA na cadeia produtiva tornou-se um dos pilares fundamentais da Indústria 4.0. Em setores como o de papel e celulose, essa tecnologia tem sido utilizada para melhorar o controle de qualidade e prever falhas operacionais. Modelos de aprendizado de máquina e redes neurais são empregados para monitorar variáveis críticas, como gramatura e umidade, reduzindo o refugo e garantindo uma produção mais eficiente (Wan *et al.*, 2022). Liu *et al.* (2020) apresentou em seus resultados que a aplicação de algoritmos de aprendizado extremo (*Extreme Learning Machines* –

ELM) aumentou a precisão na previsão de qualidade de efluentes em processos industriais, resultando em uma melhora de 27,5% em comparação com métodos tradicionais. Essa inovação foi fundamental para otimizar o uso de recursos e minimizar desperdícios.

Além disso, a utilização de gêmeos digitais (Digital Twins) na indústria de papel permite a simulação de diferentes cenários produtivos, facilitando a detecção de anomalias e possibilitando ajustes operacionais em tempo real. Trata-se de uma tecnologia que cria uma réplica virtual do processo físico, capaz de espelhar em tempo real os dados da operação, permitindo testes, previsões e intervenções com base em condições reais (Zhang *et al.*, 2023). Essa abordagem oferece uma visão detalhada de cada etapa do processo, possibilitando que variáveis de produção sejam ajustadas de forma dinâmica, o que contribui para a redução de perdas e o aumento da eficiência produtiva. A capacidade de antecipar falhas e corrigir desvios antes que afetem a qualidade do produto final representa um avanço significativo na busca por processos industriais mais sustentáveis e competitivos (Liu *et al.*, 2020).

A IA também desempenha um papel crucial na redução de refugos industriais, promovendo uma abordagem preditiva que substitui métodos tradicionais de controle de qualidade. He *et al.* (2022) destacam que a introdução de sensores inteligentes e modelos preditivos na indústria de papel reciclado possibilitou a identificação precoce de variações durante o processo produtivo, resultando em uma redução de refugo de até 17,3%. Esse tipo de tecnologia tem sido essencial para minimizar perdas de material e evitar desperdícios, garantindo maior eficiência em toda a cadeia produtiva.

O monitoramento contínuo proporcionado pela IA permite que as indústrias identifiquem padrões complexos que muitas vezes passam despercebidos na análise manual. Modelos de *Machine Learning* são treinados com dados históricos, permitindo que o sistema reconheça condições favoráveis à ocorrência de defeitos e atue proativamente para evitá-los (Wan *et al.*, 2022). Essa capacidade de aprendizado contínuo é um dos diferenciais da IA tornando-a uma ferramenta essencial para o sucesso das indústrias modernas.

A incorporação de IA nos processos industriais não apenas melhora a qualidade do produto final, mas também contribui para a sustentabilidade ao reduzir o consumo de matéria-prima e energia. A automação e a digitalização proporcionam um controle mais rigoroso sobre cada etapa da produção, permitindo ajustes precisos e imediatos, o que resulta em uma operação mais limpa e eficiente (Dias *et al.*, 2024).

Empresas que adotam essas tecnologias se tornam mais competitivas, conseguindo oferecer produtos de maior qualidade com menor impacto ambiental, consolidando-se no mercado global como referências em inovação e eficiência.

2.4 MODELOS DE CLASSIFICAÇÃO PARA PREDIÇÃO DE REFUGO

A predição de refugo no ambiente industrial pode ser tratada como um problema de classificação binária, no qual o objetivo é identificar, com base em variáveis de processo, se determinado lote será aprovado ou refugado. Para isso, são aplicados algoritmos de aprendizado de máquina supervisionado, capazes de aprender padrões a partir de dados rotulados, gerando modelos que classificam novas observações com base em suas características (Han; Kamber; Pei, 2011; Murphy, 2012). A construção de modelos eficazes depende, entre outros fatores, da escolha do tipo de aprendizado adequado à natureza do problema, conforme discutido a seguir.

2.4.1 Tipos de aprendizado de máquina

Segundo Murphy (2012) e Alpaydin (2020), os algoritmos de aprendizado de máquina podem ser classificados de acordo com a natureza do problema abordado e do conjunto de dados disponíveis.

Além da distinção entre os tipos de aprendizado de máquina, é fundamental considerar fatores que afetam o desempenho dos modelos. Um dos desafios recorrentes é o desbalanceamento entre as classes presentes no conjunto de dados. Em problemas de classificação binária, quando uma classe está significativamente sub-representada em relação à outra, os algoritmos podem apresentar viés em favor da classe majoritária, comprometendo a capacidade de identificação da classe minoritária. Segundo Lemaître, Nogueira e Aridas (2017), esse tipo de assimetria tende a reduzir a sensibilidade do modelo, sendo necessária a adoção de técnicas específicas de reamostragem ou ajustes nos algoritmos para mitigar esse efeito.

Outro aspecto relevante refere-se à forma de representação dos dados de entrada. O uso de estatísticas agregadas, como médias ou desvios-padrão, pode ser adequado para simplificar o processamento, mas também pode limitar a capacidade dos modelos de captar padrões sutis presentes nos dados. Géron (2019) destaca que,

ao reduzir dados temporais ou espaciais a resumos estatísticos, perde-se granularidade isto é, o nível de detalhe ou resolução com que os dados descrevem o fenômeno observado e, com isso, a possibilidade de identificar variações significativas ao longo do tempo. O autor recomenda, sempre que viável, a utilização de dados brutos ou sequenciais, especialmente em contextos industriais, onde pequenas flutuações podem indicar desvios relevantes no processo produtivo.

A efetividade dessa análise, contudo, também depende da escolha da abordagem de aprendizado mais compatível com a estrutura dos dados disponíveis, o que envolve compreender como diferentes algoritmos tratam a relação entre entradas e saídas (Alpaydin, 2020).

2.4.1.1 Aprendizado supervisionado

O aprendizado supervisionado consiste no treinamento de algoritmos a partir de dados rotulados, ou seja, entradas associadas às suas saídas conhecidas. O objetivo é encontrar uma função $f: X \rightarrow Y$ que mapeie características de entrada (X) para classes ou valores contínuos (Y) (James *et al.*, 2021).

Os problemas supervisionados podem ser divididos em duas categorias principais: classificação, quando a variável de saída é categórica, como no caso da distinção entre lotes "Aprovado" ou "Refugado" e regressão, quando a variável de saída é contínua, como na previsão de valores de resistência física, a exemplo do RCT.

2.4.1.2 Aprendizado não supervisionado

O aprendizado não supervisionado é uma abordagem dentro do aprendizado de máquina que lida com conjuntos de dados não rotulados, ou seja, em que não há valores de saída conhecidos para orientar o treinamento do modelo (Murphy, 2012; Alpaydin, 2020). O principal objetivo dessa técnica é identificar estruturas ocultas, relações ou agrupamentos naturais nos dados, sem a supervisão externa que ocorre no aprendizado supervisionado (James *et al.*, 2021).

Dentre as técnicas mais empregadas, destacam-se os algoritmos de agrupamento (clustering), como o K-means, utilizados para segmentar dados com base em similaridade; as regras de associação, como o algoritmo Apriori, voltadas à

descoberta de relações frequentes entre variáveis; e os métodos de redução de dimensionalidade, como a Análise de Componentes Principais (PCA), que possibilitam a simplificação de conjuntos de dados complexos sem perda significativa de informação. Essas técnicas são amplamente utilizadas em aplicações industriais e científicas para a exploração de dados, identificação de padrões anômalos e suporte à tomada de decisão em contextos com alta variabilidade e ausência de rotulagem prévia (DATACAMP, 2024).

De acordo com Goodfellow, Bengio e Courville (2016), enquanto o aprendizado supervisionado busca minimizar a diferença entre previsões e respostas conhecidas, o aprendizado não supervisionado tenta aprender a estrutura subjacente dos dados com base em padrões estatísticos internos.

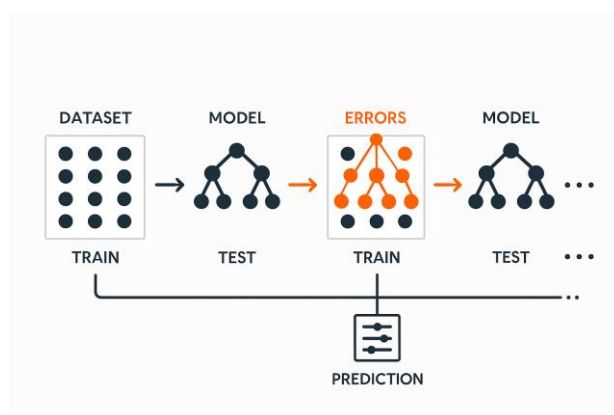
2.4.2 Modelos supervisionados de classificação

Entre os modelos de classificação aplicados na predição, podem ser destacados o *Gradient Boosting*, o *Random Forest* e o *MLPClassifier*. Estes algoritmos de aprendizado supervisionado têm se mostrado eficazes na análise de variáveis críticas, como a gramatura e a umidade das bobinas, proporcionando elevada capacidade preditiva e contribuindo para a melhoria do controle dos processos produtivos (Grinsztajn; Oyallon; Varoquaux, 2022).

2.4.2.1 Gradient Boosting

O *Gradient Boosting* é um algoritmo de aprendizado supervisionado que pertence à classe dos métodos *ensemble*, ou seja, combina vários modelos simples para formar um modelo mais robusto e preciso. Nesse caso, os modelos simples (chamados de modelos fracos) são geralmente árvores de decisão rasas. O funcionamento do *Gradient Boosting* é baseado na construção sequencial desses modelos, em que cada nova árvore é treinada para corrigir os erros (ou resíduos) cometidos pelos modelos anteriores (Friedman, 2001). A Figura 6 apresenta um esquema visual simplificado que ilustra o funcionamento sequencial do algoritmo Gradient Boosting.

Figura 6 - Esquema do funcionamento do algoritmo Gradient Boosting



Fonte: Adaptado de UC-R (2024)

O esquema evidencia o princípio central do Gradient Boosting, a construção sucessiva de modelos fracos, cada um ajustado para reduzir os erros do modelo anterior. A partir de um conjunto inicial de dados, o primeiro modelo é treinado e os erros resultantes são destacados. Esses erros passam a orientar o treinamento do modelo seguinte, e assim sucessivamente. O processo continua até que um número predefinido de modelos seja construído, sendo a predição final resultante da combinação ponderada de todas as predições anteriores, otimizando o desempenho do modelo global (FRIEDMAN, 2001; UC-R, 2024).

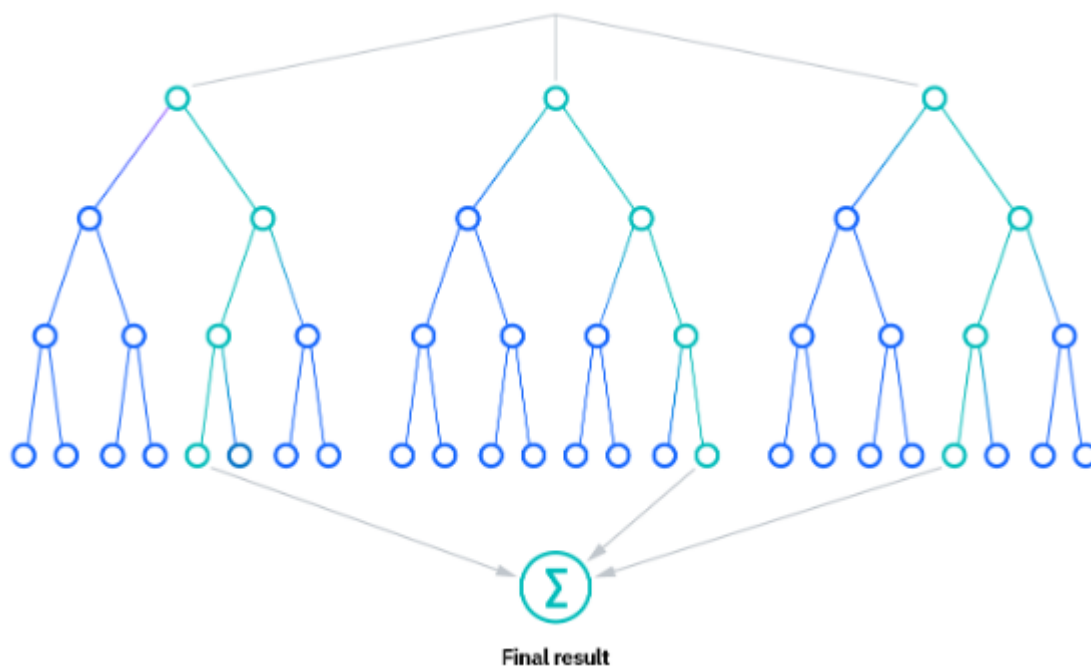
Na indústria papeleira, o *Gradient Boosting* tem sido empregado para a predição de parâmetros críticos de qualidade, como gramatura e resistência à tração, integrando sistemas de gêmeo digital para monitoramento preditivo da produção. Zhang, Yu e Li (2023) desenvolveram um sistema de gêmeo digital que utilizou o *Gradient Boosting* para prever falhas em bobinas de papel antes que ocorressem.

2.4.2.2 Random Forest

O Random Forest (RF) é um algoritmo baseado no conceito de Ensemble Learning, no qual múltiplas árvores de decisão são construídas a partir de subconjuntos aleatórios do conjunto de dados original e, posteriormente, combinadas para gerar uma predição mais robusta. Proposto por Breiman (2001), esse método utiliza o *bootstrap aggregating (bagging)*, técnica que consiste em gerar diversas amostras aleatórias com reposição do conjunto de dados original, sobre as quais diferentes modelos são treinados. Além disso, emprega a seleção aleatória de

atributos em cada divisão de nó, promovendo diversidade entre as árvores e reduzindo a variância do modelo. A Figura 7 ilustra a arquitetura conceitual do algoritmo Random Forest.

Figura 7 – Esquema do funcionamento do algoritmo *Random Forest*



Fonte: IMB (2024).

O diagrama representa múltiplas árvores de decisão, cada uma treinada de forma independente com subconjuntos aleatórios do conjunto de dados original, utilizando a técnica de bootstrap (amostragem com reposição). Cada árvore realiza uma predição individual, por exemplo, classificando se um lote é "conforme" ou "refugado". As decisões de todas as árvores são então combinadas em uma etapa final de agregação, representada no centro inferior do diagrama pelo símbolo Σ . No caso de problemas de classificação, a Random Forest realiza a votação da maioria das árvores para determinar o resultado final. (BREIMAN, 2001; JAMES *et al.*, 2021; IBM, 2024).

Suas principais características incluem a robustez a dados ruidosos e a superajuste, a capacidade de lidar com dados de alta dimensionalidade e variáveis faltantes, além de fornecer estimativas da importância das variáveis, o que contribui para a interpretabilidade dos resultados (Grinsztajn; Oyallon; Varoquaux, 2022). Contudo, apesar de sua eficiência, o *Random Forest* pode apresentar limitações na identificação de relações complexas entre variáveis quando comparado a métodos

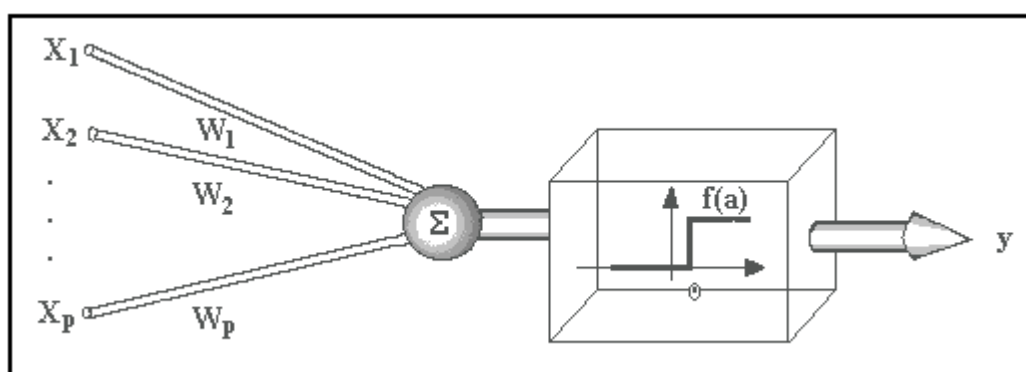
mais sofisticados, como o *Gradient Boosting*, além de exigir maior capacidade computacional para realizar previsões em tempo hábil (Grinsztajn; Oyallon; Varoquaux, 2022).

Devi, Sakthivel e Priya (2023) desenvolveram um modelo de *Random Forest* que alcançou mais de 90% de precisão na classificação de bobinas como "conforme" ou "não conforme". O modelo ajudou a reduzir o tempo de inspeção manual e forneceu *insights* sobre as variáveis mais relevantes para a decisão de qualidade, como a gramatura e a resistência mecânica.

2.4.2.3 MLPClassifier

O *MLPClassifier*, sigla para *Multilayer Perceptron Classifier*, é um tipo de Rede Neural Artificial (RNA) composta por múltiplas camadas de neurônios. É um modelo de aprendizado supervisionado que realiza classificação não linear, sendo capaz de aprender padrões complexos em bases de dados diversas. Sua estrutura básica inclui uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída, e seu treinamento é realizado com o algoritmo de retropropagação, utilizando gradiente descendente para ajuste dos pesos (Goodfellow; Bengio; Courville, 2016). A Figura 8 ilustra o funcionamento de uma unidade fundamental de redes neurais artificiais.

Figura 8 – Esquema de uma unidade básica de processamento de uma rede neural artificial



Fonte: Adaptado de Braga; Carvalho; Ludermir (2007).

O esquema representa a unidade básica de processamento de uma rede neural artificial, fundamental para o funcionamento do *MLPClassifier*. Cada entrada x_i do sistema é associada a um peso w_i , refletindo sua importância relativa. Essas entradas ponderadas são somadas por uma unidade agregadora (Σ), que gera um valor de

ativação a . Esse valor é então processado por uma função de ativação não linear, $f(a)$, responsável por introduzir complexidade e capacidade de aprendizado não linear ao modelo. A saída y corresponde à resposta da unidade frente aos estímulos recebidos. Quando organizadas em camadas (entrada, ocultas e saída), essas unidades formam as bases do *MLPClassifier*, permitindo a modelagem de padrões complexos em tarefas de classificação (Haykin, 2009; Braga; Carvalho; Ludermir, 2007).

Uma das principais vantagens do *MLPClassifier* é sua capacidade de capturar relações não lineares entre as variáveis de entrada e a variável alvo. Por outro lado, o algoritmo pode ser computacionalmente intensivo e requer ajustes de parâmetros, como o número de camadas, neurônios, função de ativação e taxa de aprendizado, para alcançar bons resultados (Haykin, 2009).

Na indústria papaleira, o *MLPClassifier* tem sido utilizado para prever a ocorrência de defeitos em bobinas de papel, com base em dados de sensores integrados ao processo produtivo. He, Zhang e Li (2022) implementaram uma Rede Neural Multicamadas para monitorar a produção em tempo real e prever a conformidade das bobinas antes da etapa final de acabamento. O sistema contribuiu para reduzir o índice de falhas e aumentar a eficiência ao permitir ajustes automáticos dos parâmetros de operação das máquinas.

2.4.3 Avaliação de modelos de classificação supervisionada

A avaliação de modelos de aprendizado supervisionado, especialmente em problemas de classificação binária, requer o uso de métricas estatísticas específicas que permitem medir a eficácia preditiva dos algoritmos. Essas métricas são fundamentais para a validação dos modelos, possibilitando a comparação objetiva entre diferentes técnicas de aprendizado de máquina (James *et al.*, 2021).

Entre as principais métricas adotadas na literatura para avaliação de classificadores estão a acurácia, a precisão e a revocação, além da matriz de confusão, que fornece uma visão detalhada sobre o desempenho dos modelos em cada classe (Sarwar, 2020).

2.4.3.2 Matriz de confusão

A matriz de confusão é uma ferramenta fundamental na avaliação de algoritmos de classificação supervisionada, permitindo uma análise detalhada do desempenho do modelo em relação às classes reais e previstas e seu modelo está apresentado na Figura 9.

Figura 9 – Modelo de matriz de confusão

Real	Positivo	Negativo
	VP	FN
Negativo	FP	VN
		Predição
		Positivo Negativo

Fonte: Adaptado de IBM (2024).

Em problemas binários, como o abordado neste estudo, a matriz organiza os resultados em quatro categorias principais: Verdadeiros Positivos (VP), que representam os casos corretamente classificados como positivos; Verdadeiros Negativos (VN), que indicam as classificações corretas da classe negativa; Falsos Positivos (FP), que ocorrem quando o modelo classifica erroneamente um caso negativo como positivo; e Falsos Negativos (FN), que refletem a falha do modelo em identificar corretamente a classe positiva (SARWAR, 2020; JAMES *et al.*, 2021).

Essa representação tabular permite visualizar diretamente quais erros o modelo mais comete, sendo particularmente útil em contextos com desequilíbrio de classes, como é o caso da identificação de refugos, onde a classe minoritária (refugado) tende a ser preterida pelas métricas globais. A matriz de confusão também é a base para o cálculo de outras métricas, como acurácia, precisão e revocação, sendo, portanto, essencial para uma avaliação crítica e contextualizada da performance dos modelos (HAN *et al.*, 2011).

2.4.3.2 Acurácia

A acurácia é uma das métricas mais utilizadas em problemas de classificação e representa a proporção de previsões corretas em relação ao total de observações avaliadas (Trevor; Hastie; Friedman, 2009). É calculada pela razão entre os verdadeiros positivos e verdadeiros negativos sobre o total de amostras, que incluem os falsos positivos e falsos negativos, Equação 3:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (3)$$

Embora seja uma métrica de fácil interpretação, sua eficácia diminui em cenários de classes desbalanceadas, onde outras métricas tornam-se mais apropriadas (James *et al.*, 2021).

2.4.3.3 Precisão

A precisão refere-se à capacidade do modelo de classificar corretamente os exemplos positivos, ou seja, mede a proporção de verdadeiros positivos entre todas as predições classificadas como positivas, conforme Equação 4:

$$Precisão = \frac{VP}{VP + FP} \quad (4)$$

A literatura destaca a importância da precisão em contextos onde falsos positivos devem ser minimizados, como em sistemas de controle de qualidade que buscam evitar a aprovação indevida de produtos defeituosos (James *et al.*, 2021).

2.4.3.4 Revocação

A revocação, também chamada de sensibilidade, avalia a capacidade do modelo em identificar corretamente todos os exemplos positivos, sendo crucial em contextos onde a falha em detectar casos positivos pode resultar em graves consequências (Trevor; Hastie; Friedman, 2009), é determinada pela Equação 5:

$$Revocação = \frac{VP}{VP + FN} \quad (5)$$

Em problemas industriais, a revocação é uma métrica amplamente utilizada para garantir a detecção completa de lotes defeituosos, reduzindo o risco de falhas no produto final (Sarwar, 2020).

3 METODOLOGIA

Este capítulo apresenta a descrição dos procedimentos metodológicos adotados para o desenvolvimento deste estudo, que teve como objetivo aplicar técnicas de aprendizado de máquina na predição da conformidade de lotes de papel reciclado, com base em parâmetros de controle de qualidade industrial. A metodologia foi estruturada de modo a garantir a transparência das etapas realizadas, abrangendo desde a coleta e tratamento dos dados até o desenvolvimento, treinamento e avaliação dos modelos de classificação supervisionada.

3.1 TIPO DE PESQUISA

Esta pesquisa é caracterizada como quantitativa e aplicada, com abordagem explicativa, uma vez que objetiva compreender a relação entre variáveis de processo (gramatura, umidade, resistência RCT e Mullen) e a classificação da qualidade dos lotes de papel reciclado. A pesquisa aplicada busca solucionar um problema prático identificado na indústria papeleira, enquanto a abordagem quantitativa se baseia na análise de dados numéricos e na utilização de algoritmos de aprendizado de máquina para classificação preditiva (Gil, 2010; Marconi; Lakatos, 2017).

3.2 COLETA E PRÉ-PROCESSAMENTO DOS DADOS

Os dados utilizados neste estudo foram extraídos da base de dados de uma fábrica de papel reciclado localizada no estado de Pernambuco, abrangendo um período contínuo de 11 meses. Foram coletadas informações referentes a 6.613 lotes de produção, sendo cada lote composto por 254 medições pontuais de gramatura e umidade obtidas transversalmente ao longo da bobina, a cada 105 segundos de processo.

Inicialmente, os dados foram organizados e tratados no software Microsoft Excel, com o objetivo de eliminar registros inconsistentes, valores ausentes e outliers, conforme os critérios estatísticos recomendados por Han *et al.* (2011).

Devido ao elevado volume de registros e às limitações de processamento local, optou-se por utilizar estatísticas descritivas consolidadas para representar as principais variáveis de entrada. Esses parâmetros foram calculados com base nas

leituras ponto a ponto de gramatura e umidade realizadas ao longo da extensão de cada bobina, sendo os resultados agregados por lote para compor a base de modelagem. As estatísticas utilizadas incluíram:

- Média: calculada pela soma de todas as leituras individuais da variável, dividida pelo número total de medições em cada bobina;
- Mediana: valor central da distribuição das leituras, obtido após ordenar os dados em ordem crescente;
- Desvio padrão: medida da dispersão dos dados em relação à média, calculado a partir da raiz quadrada da variância das leituras de cada bobina;
- Quartis (Q1, Q2, Q3): valores que dividem a distribuição ordenada dos dados em quatro partes iguais;
- Valor máximo e mínimo: respectivamente, os maiores e menores valores observados nas leituras da variável em cada bobina.

Em seguida, foi conduzida uma Análise Exploratória dos Dados (AED), com o intuito de identificar padrões estatísticos, avaliar a distribuição das variáveis e caracterizar diferenças entre lotes classificados como “aprovado” e “refugado”. Essa etapa incluiu a construção de gráficos de distribuição e boxplots comparativos, viabilizando a identificação de dispersões elevadas associadas a não conformidades nos testes físicos.

Também foi realizada uma análise de correlação entre os parâmetros estatísticos consolidados e os testes laboratoriais de qualidade (Ring Crush Test – RCT e Mullen). Utilizou-se o coeficiente de correlação de Pearson para medir o grau de associação linear entre as variáveis de entrada e os atributos de desempenho físico, auxiliando na interpretação e na seleção das variáveis mais relevantes para a modelagem.

Ao final desta etapa, os dados estruturados foram integrados em um único banco de dados consolidado, sendo posteriormente importados para o ambiente de desenvolvimento Python, onde foram aplicados os modelos de aprendizado de máquina.

3.3 DESENVOLVIMENTO E TREINAMENTO DOS MODELOS

Na etapa de modelagem, foram aplicados três algoritmos de aprendizado supervisionado de classificação binária, com o objetivo de prever a conformidade dos lotes de papel reciclado: *Gradient Boosting Classifier*, *Random Forest Classifier* e *MLPClassifier* (Multilayer Perceptron).

O processo de desenvolvimento foi realizado na linguagem Python, no ambiente Google Colab, com suporte das bibliotecas *scikit-learn*, *pandas*, *numpy*, *matplotlib* e *seaborn*, que forneceram as ferramentas necessárias para a construção, treinamento, avaliação e visualização dos modelos.

Os dados consolidados foram importados para o ambiente Colab e divididos em 80% para o conjunto de treinamento e 20% para o conjunto de teste, utilizando a técnica de *Holdout*, que consiste em separar o conjunto de dados em dois subconjuntos mutuamente exclusivos: um para treinar o modelo e outro para avaliá-lo. Foi aplicada a estratificação durante a divisão, com o objetivo de preservar a proporção das classes nos dois subconjuntos (Han *et al.*, 2011). A escolha da proporção 80/20 é amplamente utilizada na literatura por representar um equilíbrio entre a quantidade de dados disponível para o aprendizado do modelo e um volume suficiente de dados inéditos para a validação, garantindo maior confiabilidade na avaliação do desempenho preditivo (JAMES *et al.*, 2021).

Os hiperparâmetros de cada modelo foram definidos conforme as recomendações da biblioteca *scikit-learn* e ajustados empiricamente com base em validações sucessivas. No *Gradient Boosting Classifier*, utilizou-se *n_estimators*=100, que define o número de árvores sequenciais construídas no processo; *learning_rate*=0.1, responsável por controlar o peso de cada árvore adicionada ao modelo, afetando diretamente a taxa de aprendizado; *max_depth*=3, que limita a profundidade máxima de cada árvore base para evitar sobreajuste; e *random_state*=42, utilizado para garantir a reprodutibilidade dos resultados (Scikit-learn, 2025).

No *Random Forest Classifier*, adotou-se *n_estimators*=100, indicando a quantidade de árvores independentes geradas; *max_depth*=None, permitindo que as árvores cresçam até o seu limite natural, maximizando a profundidade conforme necessário; e *random_state*=42, assegurando consistência nos experimentos (Scikit-learn, 2025).

Já no *MLPClassifier*, os hiperparâmetros selecionados foram *hidden_layer_sizes*=(50, 50), o que define duas camadas ocultas com 50 neurônios cada; *activation*='relu', função de ativação que introduz não linearidade e melhora o desempenho computacional; *solver*='adam', um otimizador que ajusta os pesos da rede neural de forma adaptativa a cada iteração, promovendo maior eficiência e estabilidade em problemas com grandes conjuntos de dados; *max_iter*=500, determinando o número máximo de iterações para o treinamento da rede; e *random_state*=42, para manter a reprodutibilidade (Scikit-learn, 2025).

3.4 VALIDAÇÃO E TESTE

A avaliação dos modelos de classificação desenvolvidos foi realizada com base em métricas de desempenho amplamente utilizadas para problemas de classificação binária, conforme discutido na revisão de literatura (James *et al.*, 2021). As métricas aplicadas foram: Acurácia (Equação 3), Precisão (Equação 4), Revocação (Equação 5) e Matriz de Confusão.

Para cada modelo treinado, os resultados das métricas foram analisados de forma comparativa, com o objetivo de determinar qual algoritmo apresentou o melhor desempenho global na classificação dos lotes de papel reciclado como “Classificado” ou “Desclassificado”.

Além da análise quantitativa das métricas, foram geradas e avaliadas as matrizes de confusão, que permitiram verificar a distribuição dos acertos e erros de cada modelo, auxiliando na interpretação dos resultados e na validação prática dos classificadores no contexto do processo produtivo estudado.

4 RESULTADOS

O conjunto de informações coletadas totalizou milhares de dados, o que, embora proporcione uma visão detalhada das variabilidades locais, impôs desafios significativos em termos de capacidade de processamento computacional e eficiência do tempo de treinamento dos modelos de aprendizado de máquina. Esses desafios foram potencializados pelas limitações da máquina utilizada para análise, equipada com processador Intel Core i5-10210U 1.60GHz, 8 GB de memória RAM e sistema operacional Windows 10 Home (64 bits), o que demandou estratégias de consolidação estatística e uso de plataformas como o Google Colab para complementar a execução dos experimentos.

Diante dessa limitação prática, foi realizada uma Análise Exploratória dos Dados (AED) que demonstrou que as estatísticas descritivas (média, desvio padrão, valores mínimos e máximos) de cada lote são representativas dos padrões de variação observados nas séries de dados brutos. Como destacado por Han *et al.* (2011), a redução dimensional dos dados por meio de técnicas de sumarização estatística é uma estratégia recomendada para aumentar a eficiência computacional sem comprometer a qualidade preditiva dos modelos.

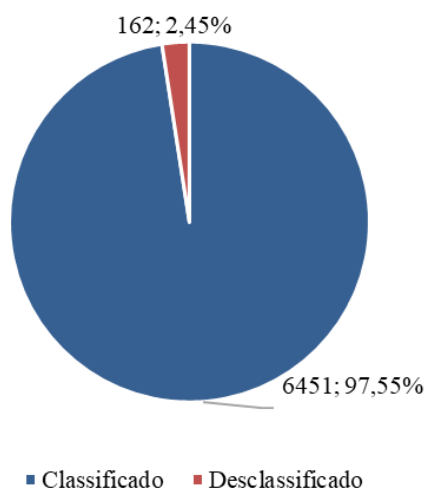
A decisão de trabalhar com atributos estatísticos resumidos se alinha ainda à prática descrita por Devi *et al.* (2023), que destaca a importância de simplificar as variáveis de entrada em processos industriais de larga escala, garantindo maior agilidade nos ciclos de treinamento e inferência dos modelos, além de proporcionar facilidade de interpretação dos resultados preditivos. Portanto, a abordagem adotada, utilizando estatísticas descritivas dos lotes, visa não apenas contornar a limitação de recursos computacionais, mas também potencializar a aplicabilidade prática do modelo de Inteligência Artificial proposto, garantindo eficiência e precisão no monitoramento preditivo da qualidade das bobinas de papel reciclado.

4.2 ANÁLISE EXPLORATÓRIA DOS DADOS (AED)

A análise dos dados coletados ao longo do processo produtivo de bobinas de papel reciclado permitiu a identificação de padrões e variáveis críticas que influenciam diretamente a qualidade final do produto. Dos 6.613 lotes analisados, 6.451 lotes eram classificados como aprovados e 162 lotes desclassificados (refugados), o que

corresponde a uma taxa de desclassificação de 2,45%, conforme está apresentado no gráfico da Figura 10.

Figura 10 – Distribuição da classificação final dos lotes avaliados

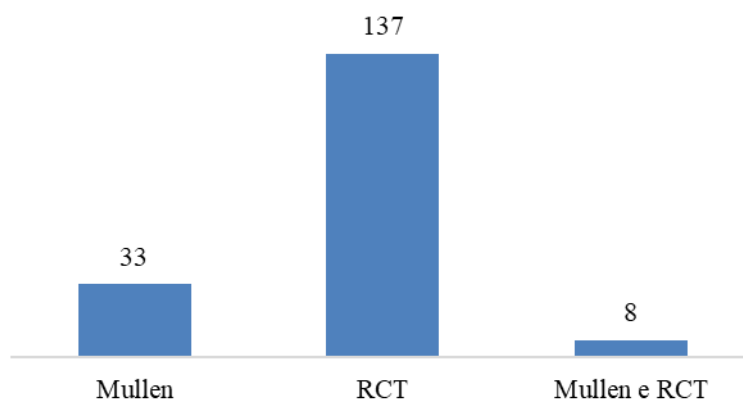


Fonte: Dados da pesquisa (2025).

Este índice de refugo está alinhado com a média reportada pela literatura para processos de reciclagem de papel, que apontam taxas variando de 1,7% a 5% em contextos industriais semelhantes (Devi *et al.*, 2023; He *et al.*, 2022).

A reprovação dos lotes está associada ao não atendimento aos critérios mínimos estabelecidos para o Ring Crush Test (RCT) e para o teste Mullen, ambos fundamentais para assegurar a resistência estrutural e a qualidade do papel reciclado conforme ilustrado na Figura 11.

Figura 11 - Distribuição dos lotes desclassificados segundo critérios de não conformidade nos ensaios de RCT e Mullen

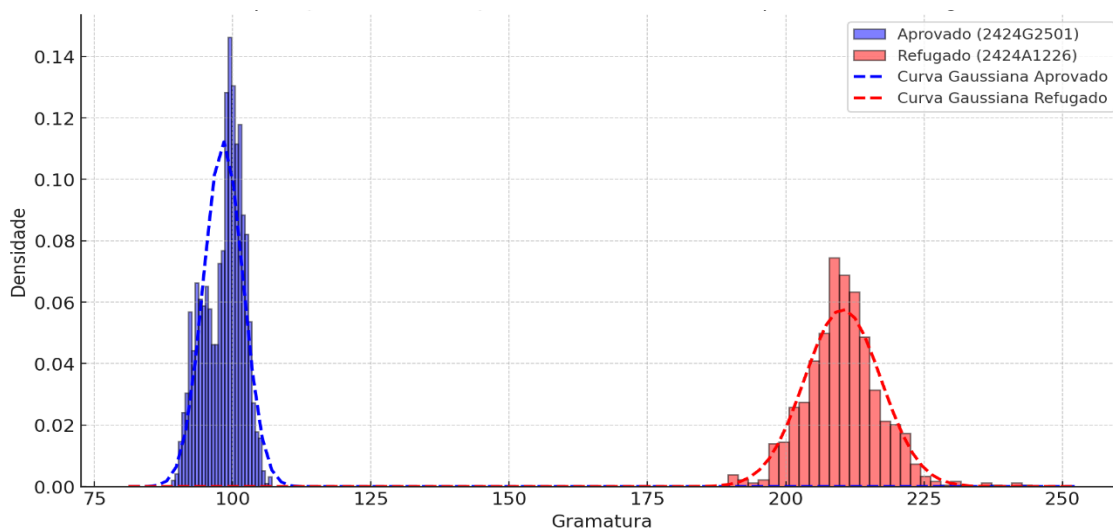


Fonte: Dados da pesquisa (2025).

Dos 162 lotes desclassificados durante o período analisado, a maior parte, 137 lotes, foi reprovada por não atingir o limite de especificação mínimo exigido pela fábrica para o RCT de seu respectivo produto. Este resultado evidencia o RCT como o principal fator de desclassificação na linha de produção avaliada. Adicionalmente, 33 lotes foram desclassificados por não conformidade no teste Mullen. Por fim, 8 lotes apresentaram não conformidade simultânea em ambos os testes (RCT e Mullen), o que demonstra que a falha conjunta nos critérios é menos frequente, mas compromete severamente a qualidade final do produto. Esses dados reforçam a importância do controle rigoroso das variáveis que impactam diretamente no desempenho do RCT, como a uniformidade do perfil do papel (variação de gramatura e umidade), uma vez que desvios nesses parâmetros comprometem a resistência à compressão anular (He *et al.*, 2022). Além disso, a não conformidade no teste Mullen, apesar de menos frequente, também representa um ponto de atenção, uma vez que está relacionada à resistência à ruptura e integridade do papel sob pressão, aspectos cruciais para o desempenho em aplicações finais (Borba, 2012).

4.2.1 Distribuição das variáveis de gramatura e umidade

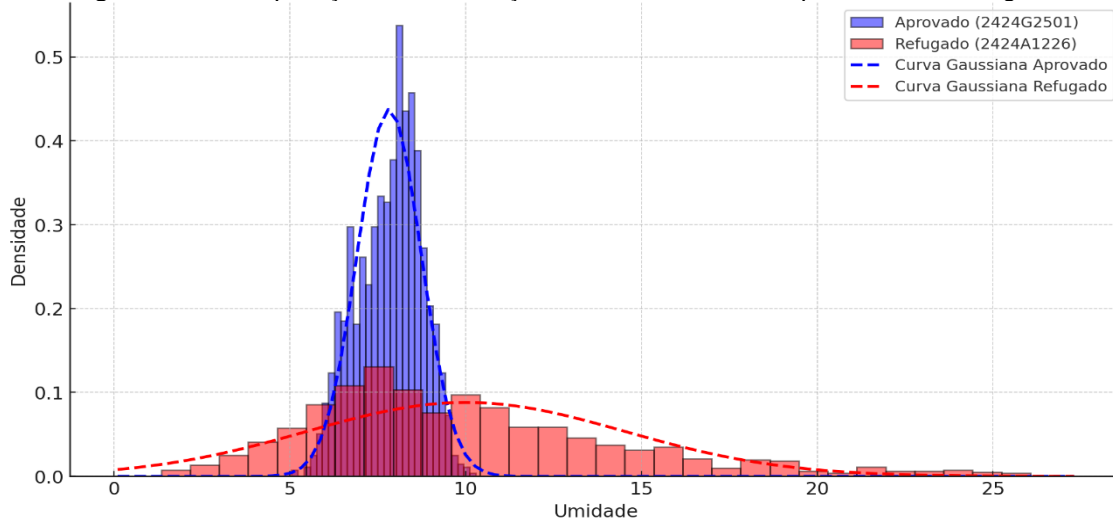
Para compreender melhor a relação entre variabilidade dos parâmetros produtivos e a classificação dos lotes, foram analisadas as distribuições estatísticas da gramatura e da umidade em um lote classificado e outro desclassificado. A Figura 12 evidencia a distribuição da gramatura de dois lotes selecionados a partir das suas classificações.

Figura 12 – Comparação da distribuição de gramatura – Lote aprovado vs. refugado

Fonte: Dados da pesquisa (2025).

O lote aprovado dentre das especificações apresentou uma distribuição mais concentrada, com média de 98,31 g/m² e desvio padrão de 3,55 g/m², indicando uma variação mais controlada ao longo do processo produtivo. Por outro lado, o lote refugado apresentou uma dispersão significativamente maior, com média de 210,29 g/m² e desvio padrão de 6,92 g/m², evidenciando oscilações frequentes na gramatura do papel. Conforme discutido por Devi *et al.* (2023) e Borba (2012), essa alta variação compromete a integridade estrutural do papel, elevando a probabilidade de falhas mecânicas e reprovação nos ensaios físicos.

A Figura 13 apresenta a distribuição da umidade para os mesmos lotes.

Figura 13 – Comparação da distribuição de umidade – Lote aprovado vs. Refugado

Fonte: Dados da pesquisa (2025).

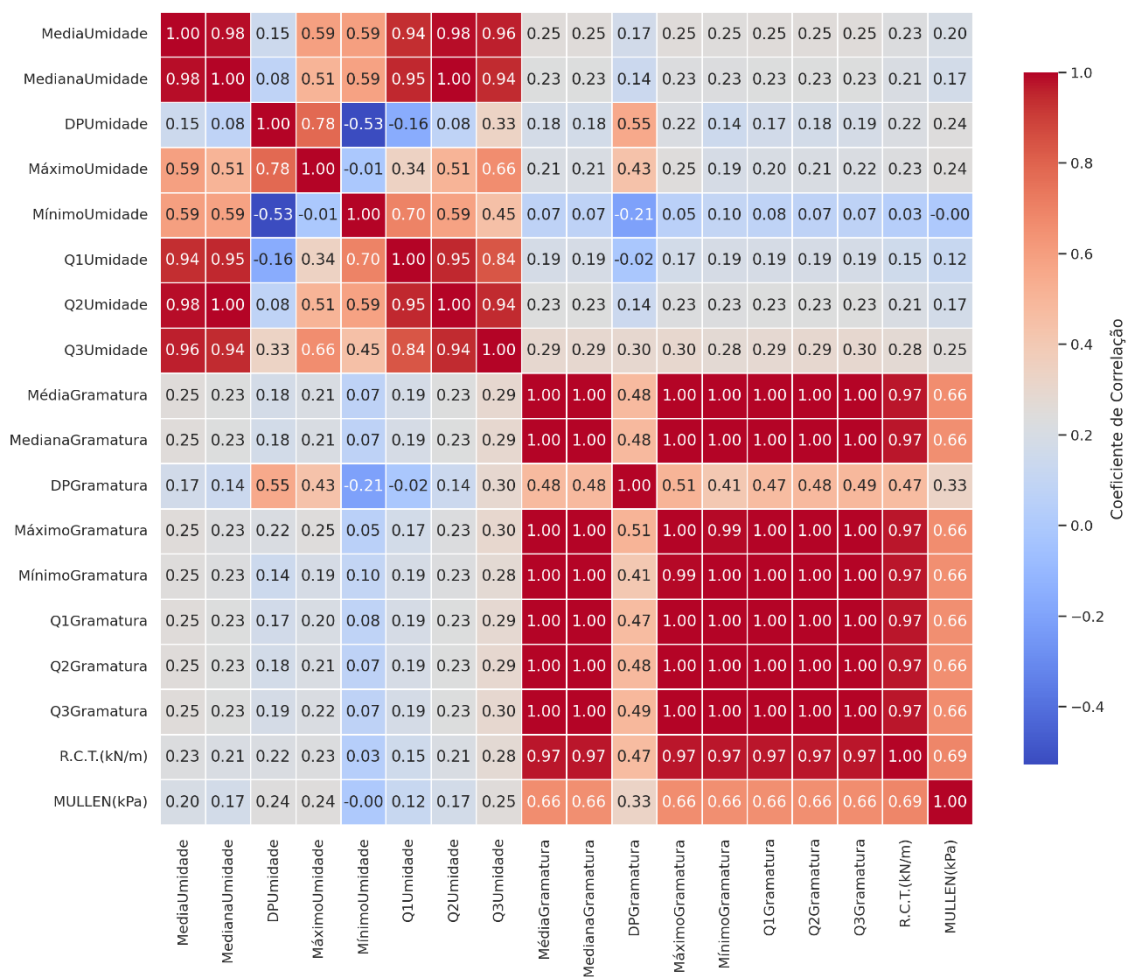
A dispersão dos valores de umidade no lote refugado, com desvio padrão de 4,53%, foi cinco vezes maior do que a registrada no lote aprovado, cujo desvio foi de 0,91%. Essa alta variabilidade na umidade é um fator crítico, pois oscilações excessivas impactam diretamente o desempenho no teste Mullen, reduzindo a resistência do papel à ruptura sob pressão (He *et al.*, 2022; Cardoso *et al.*, 2012).

A análise confirma que a maior variabilidade na gramatura e umidade dos lotes refugados está diretamente associada às não conformidades nos testes físicos. Segundo Zhang *et al.* (2023), a manutenção do desvio padrão abaixo de 5% é um fator determinante para garantir a estabilidade estrutural do papel reciclado.

A alta dispersão observada no lote refugado compromete a uniformidade do perfil do papel, tornando-o mais suscetível a falhas mecânicas durante a conversão e aplicação final. Conforme destacado por Srinivasan *et al.* (2024), processos industriais que operam com baixa variabilidade apresentam menores taxas de refugo e maior previsibilidade na qualidade do produto final. Essas constatações reforçam a necessidade de métodos preditivos baseados em Inteligência Artificial que integrem variáveis estatísticas, visando antecipar o risco de refugo e garantir a consistência do processo produtivo (Srinivasan *et al.*, 2024; Zhang *et al.*, 2023).

4.3 ANÁLISE DE CORRELAÇÃO

A análise de correlação foi conduzida com o objetivo de investigar as relações lineares existentes entre as variáveis de entrada do processo produtivo e os principais parâmetros de qualidade das bobinas de papel reciclado. Por meio do coeficiente de correlação de Pearson, buscou-se identificar quais variáveis possuem maior influência sobre os testes laboratoriais de resistência, em especial o Ring Crush Test (RCT) e o teste Mullen, ambos essenciais para a avaliação da integridade estrutural do produto final. Os resultados estão apresentados no mapa de calor da matriz de correlação, conforme ilustrado na Figura 14.

Figura 14 – Matriz de correlação dos parâmetros analisados das bobinas de papel reciclado

Fonte: Dados da pesquisa (2025).

Observou-se, pela análise da Figura 14, que a variável *MédiaGramatura* apresenta forte correlação positiva com o RCT ($r = 0,97$), indicando que o aumento da gramatura média do papel reciclado está diretamente associado à elevação de sua resistência à compressão anular. Esse achado está de acordo com os resultados de Zhang *et al.* (2023), que destacam a gramatura como um dos principais determinantes da resistência mecânica de papéis reciclados. Além disso, a média de gramatura apresentou correlação positiva com o Mullen ($r = 0,66$), sugerindo que papéis com maior massa por área tendem a ter melhor desempenho também no teste de resistência à ruptura. Essa relação é consistente com os estudos de Borba (2012), que apontam a gramatura como uma variável crítica para a integridade do papel, não apenas em condições de compressão, mas também de pressão interna.

As variáveis *MáximoGramatura* e *MínimoGramatura* exibiram correlações muito altas com o RCT ($r = 0,97$), reforçando a ideia de que o controle dos limites

superiores e inferiores da gramatura ao longo da produção é fundamental para assegurar a qualidade do produto final. A variável de desvio padrão da gramatura (DPGramatura), que representa a variabilidade no processo de formação do papel, apresentou correlação moderada com o RCT ($r = 0,47$) e fraca com o Mullen ($r = 0,33$). Embora positivamente associadas, essas correlações são estatisticamente baixas a moderadas, indicando que a instabilidade da gramatura não foi um fator fortemente determinante no desempenho mecânico das bobinas. Ainda assim, os resultados sugerem uma tendência de que variações maiores possam impactar parcialmente os testes físicos, o que reforça, mesmo que de forma limitada, as observações de Devi *et al.* (2023) sobre a importância da uniformidade da gramatura para a integridade estrutural do papel.

No que se refere às variáveis relacionadas à umidade, a Média de Umidade apresentou correlação baixa com o RCT ($r = 0,23$) e com o Mullen ($r = 0,20$). Embora tais correlações sejam menores em relação às observadas para a gramatura, ainda é possível inferir que níveis elevados de umidade podem comprometer a integridade mecânica do papel, especialmente no teste Mullen, que mede a resistência à ruptura sob pressão. Isso está alinhado com as considerações de Borba (2012), que ressaltam a influência da umidade no comportamento do papel frente a cargas de pressão.

De modo geral, os resultados da matriz de correlação indicam que a gramatura é a variável de maior influência nos parâmetros de qualidade, particularmente no RCT. A umidade, embora menos expressiva em termos de correlação, ainda se configura como um fator relevante no desempenho do teste Mullen e deve ser considerada no monitoramento dos processos. Essa análise reforça a importância de priorizar o controle da gramatura para assegurar a qualidade das bobinas de papel reciclado, sem negligenciar o impacto potencial das flutuações na umidade.

4.4 AVALIAÇÃO DOS MODELOS DE CLASSIFICAÇÃO

Após a realização das análises preditivas, foram implementados e avaliados três modelos de classificação supervisionada com o objetivo de prever a classificação final dos lotes, categorizados como "Classificado" ou "Desclassificado". Os algoritmos utilizados foram: *Gradient Boosting Classifier*, *Random Forest Classifier* e *MLPClassifier*.

Para garantir a representatividade de ambas as classes, lotes classificados e desclassificados, os dados foram divididos em dois subconjuntos, com 80% destinados ao treinamento e 20% ao teste. Utilizou-se a técnica de estratificação, que assegura que a proporção entre as classes seja mantida em ambos os conjuntos. Isso contribuiu para evitar vieses na aprendizagem dos modelos e garantir maior confiabilidade na avaliação dos resultados.

A Tabela 1 apresenta os principais resultados obtidos para cada modelo, considerando as métricas de acurácia, precisão e revocação, obtidos, respectivamente, pelas Equações 3, 4 e 5.

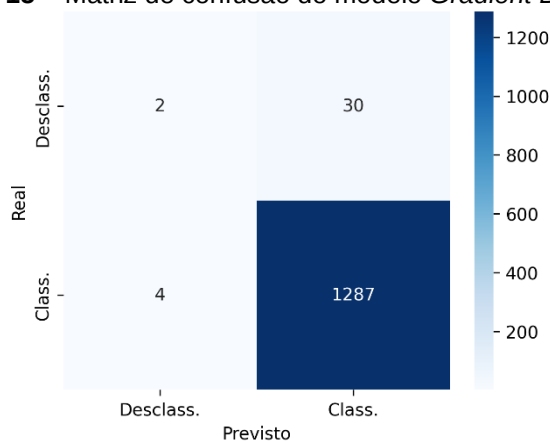
Tabela 1 – Resultados das Métricas de Desempenho dos Modelos de Classificação para Previsão da Classificação Final dos Lotes

Modelo	Acurácia	Precisão	Revocação
Gradient Boosting	0,9743	0,9772	0,9969
Random Forest	0,9751	0,9758	0,9992
MLPClassifier	0,9766	0,9773	0,9992

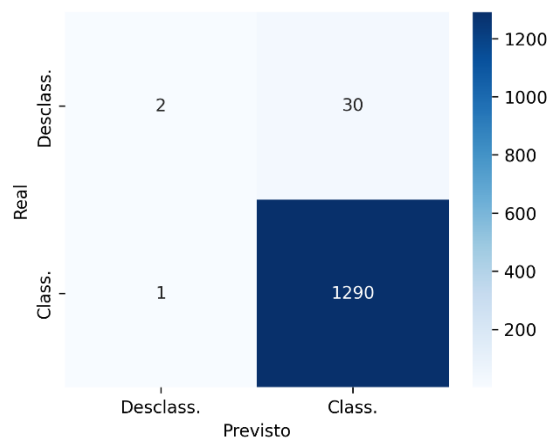
Fonte: Dados da pesquisa (2025).

Os três modelos avaliados apresentaram desempenho elevado e consistente, com acurácia superior a 97%, demonstrando capacidade de prever corretamente a classificação da grande maioria dos lotes analisados. O *MLPClassifier* obteve ligeira vantagem nos indicadores de desempenho, com acurácia de 97,66%, precisão de 97,73% e revocação de 99,92%, seguido de perto pelo *Gradient Boosting*, cuja precisão foi de 97,72% e revocação de 99,69%. O *Random Forest*, por sua vez, obteve a mesma revocação do *MLPClassifier* (99,92%), com precisão levemente inferior (97,58%). Diante da pequena variação entre os resultados, os três algoritmos podem ser considerados eficazes na tarefa de classificação dos lotes. Embora o *MLPClassifier* tenha apresentado o melhor desempenho numérico em todos os indicadores, a diferença entre os modelos é estatisticamente pequena, o que não permite afirmar com segurança a superioridade absoluta de um sobre os demais. A escolha do modelo ideal, portanto, pode considerar também critérios secundários, como tempo de treinamento, interpretabilidade e capacidade de atualização.

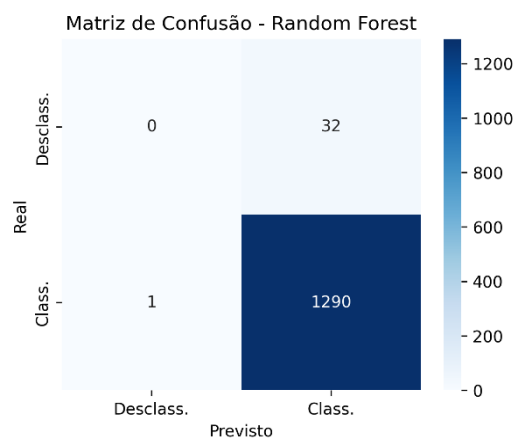
Ao final da avaliação dos modelos de classificação, foram geradas as matrizes de confusão correspondentes aos três classificadores analisados. As Figuras 15, 16 e 17 apresentam, respectivamente, as matrizes de confusão dos modelos *Gradient Boosting*, *MLPClassifier* e *Random Forest*.

Figura 15 – Matriz de confusão do modelo *Gradient Boosting*

Fonte: Dados da pesquisa (2025).

Figura 16 – Matriz de confusão do modelo *MLPClassifier*

Fonte: Dados da pesquisa (2025).

Figura 17 – Matriz de confusão do modelo *Random Forest*

Fonte: Dados da pesquisa (2025).

Analisando as matrizes de confusão, observa-se que todos os modelos apresentaram elevada capacidade de identificar corretamente os lotes classificados como "Classificado", com destaque para o *MLPClassifier*, que obteve o melhor desempenho geral. Esse modelo registrou apenas 1 falso negativo e 30 falsos positivos. O *Gradient Boosting* e o *Random Forest* também apresentaram desempenho robusto, porém com maior número de erros: o primeiro com 4 falsos negativos, e o segundo com 32 falsos positivos e nenhum verdadeiro negativo identificado.

Apesar das métricas globais indicarem desempenho satisfatório, os resultados revelam uma limitação importante no contexto do objetivo central deste trabalho, que é a previsão do refugio. Todos os modelos apresentaram baixa taxa de identificação correta dos lotes desclassificados (verdadeiros negativos), com apenas dois acertos no caso do *MLPClassifier* e do *Gradient Boosting*, e nenhum acerto no *Random Forest*. Esse resultado sugere que, embora os modelos tenham aprendido com eficiência os padrões relacionados à classe majoritária (lotes classificados como bons), ainda enfrentam dificuldades significativas em reconhecer padrões associados à não conformidade.

Essa limitação pode estar relacionada a dois fatores centrais. O primeiro diz respeito ao desbalanceamento da base de dados, em que os casos positivos (lotes bons) são muito mais numerosos que os negativos (lotes refugados), o que pode levar os classificadores a um viés em favor da classe majoritária. Segundo Lemaître, Nogueira e Aridas (2017), classificadores convencionais tendem a apresentar desempenho insatisfatório na classe minoritária em tarefas de classificação binária com dados desbalanceados, sendo necessário o uso de técnicas específicas de reamostragem ou ajustes nos algoritmos para melhorar a sensibilidade.

O segundo fator refere-se à natureza dos dados utilizados no treinamento, baseados em estatísticas descritivas agregadas por lote, como médias, desvios padrão e quartis. Embora essa abordagem tenha sido necessária devido a limitações de processamento computacional, ela pode ocultar variações temporais ou padrões sutis críticos para a correta identificação dos lotes refugados. Nesse sentido, Géron (2019) destaca que a agregação de dados temporais em estatísticas resumidas pode comprometer significativamente a capacidade dos modelos de detectar padrões relevantes, sendo mais recomendável, sempre que possível, a utilização de dados brutos ou sequenciais, especialmente em aplicações industriais.

4.5 ANÁLISE DA IMPORTÂNCIA DAS VARIÁVEIS NOS MODELOS DE CLASSIFICAÇÃO

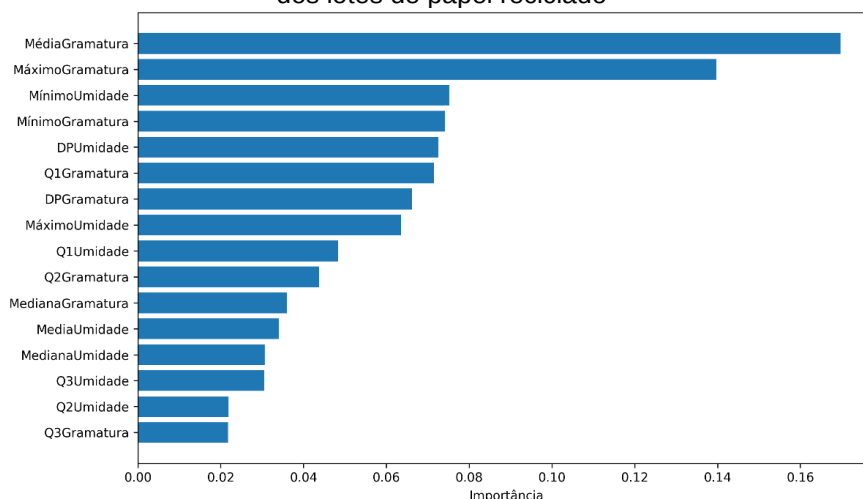
A análise de importância das variáveis desempenha um papel fundamental na interpretação dos modelos preditivos, permitindo identificar quais atributos mais influenciaram a classificação dos lotes como “Classificado” ou “Desclassificado”.

Conforme apresentado na Matriz de Correlação, Figura 14, a variável MédiaGramatura apresentou alta correlação com a resistência RCT, sugerindo que lotes com maior gramatura média tendem a possuir maior resistência mecânica. Esse padrão foi confirmado pelos modelos baseados em árvores, que atribuíram elevada importância a essa variável. Segundo Cardoso *et al.* (2012), a gramatura influencia diretamente a resistência estrutural dos papéis reciclados, o que justifica a preponderância dessa variável no processo decisório.

4.5.1 Gradient Boosting Classifier

A Figura 18 apresenta a importância das variáveis no modelo *Gradient Boosting Classifier*.

Figura 18 - Importância das variáveis no modelo *Gradient Boosting Classifier* aplicado à classificação dos lotes de papel reciclado



Fonte: Dados da pesquisa (2025).

O modelo destacou a MédiaGramatura como a variável de maior relevância, seguida pelo MáximoGramatura e MínimoUmidade. Essas variáveis representam,

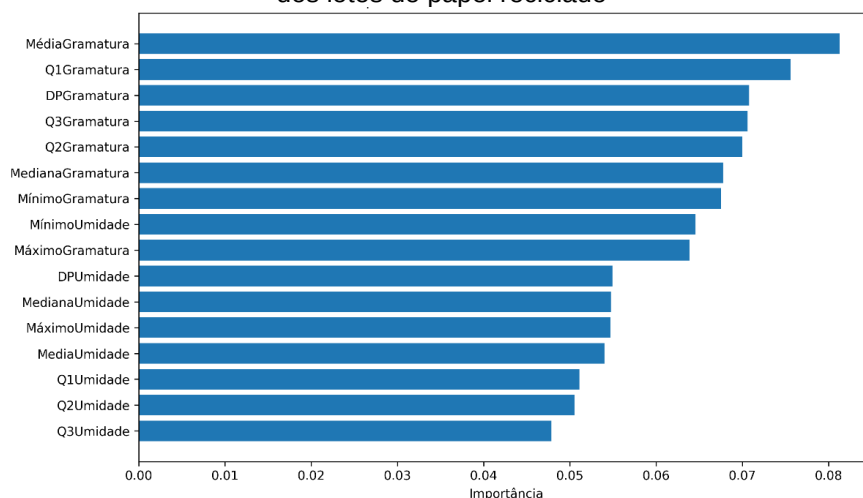
respectivamente, a tendência central e a dispersão das características físicas do papel reciclado. A importância da MáximoGramatura indica que picos de espessura ao longo da produção também impactam na qualidade final, podendo gerar zonas críticas de fragilidade ou rigidez excessiva (Devi; Priyanka; Sakthivel, 2023).

Outro ponto relevante é o destaque dado ao MínimoUmidade e DPUmidade, variáveis que refletem a homogeneidade do processo de secagem. Conforme He *et al.* (2022), a estabilidade da umidade durante a fabricação é um dos principais desafios para a manutenção da qualidade do papel reciclado.

4.5.2 Random Forest Classifier

A Figura 19 apresenta a importância das variáveis no modelo *Random Forest Classifier*.

Figura 19 - Importância das variáveis no modelo *Random Forest Classifier* aplicado à classificação dos lotes de papel reciclado



Fonte: Dados da pesquisa (2025).

O modelo *Random Forest* apresentou um perfil de importância semelhante ao *Gradient Boosting*, com MédiaGramatura, Q1Gramatura e DPGramatura figurando como as principais variáveis. A maior diversidade na seleção das estatísticas descritivas de gramatura (quartis e desvio padrão) sugere que o modelo valoriza a uniformidade do papel, corroborando o impacto da variabilidade estrutural na resistência mecânica, conforme argumentado por Oliveira *et al.* (2015).

A importância atribuída às variáveis relacionadas à umidade, como MínimoUmidade e MáximoUmidade, reflete a influência das flutuações no teor de umidade no desempenho dos produtos, como também observado na literatura (ZHANG *et al.*, 2023).

4.5.4 Comparação das importâncias das variáveis

De modo geral, a variável *MédiaGramatura*, juntamente com suas medidas de dispersão (como quartis e desvio padrão), foi destacada como a mais relevante para a predição da classificação dos lotes nos modelos baseados em árvores, *Gradient Boosting* e *Random Forest*. Esses algoritmos permitem, de forma nativa, a geração de medidas de importância das variáveis a partir da frequência e do ganho de informação obtido em cada divisão de nó. O modelo *MLPClassifier*, por sua vez, não fornece diretamente uma medida de importância das variáveis, devido à sua natureza baseada em redes neurais com múltiplas camadas ocultas e conexões não lineares entre os dados. Por esse motivo, a análise de importância foi realizada apenas para os modelos baseados em árvores.

Os resultados obtidos corroboram as análises da matriz de correlação apresentadas na Figura 14, reforçando a forte associação entre a gramatura e a resistência física das bobinas.

5 CONCLUSÃO

O presente trabalho propôs o desenvolvimento e a validação de modelos de classificação baseados em algoritmos de Inteligência Artificial para a previsão de refugo em bobinas de papel reciclado, utilizando como variáveis de entrada os valores de gramatura e umidade obtidos durante o processo produtivo.

Os resultados demonstraram que os três modelos supervisionados avaliados *Gradient Boosting*, *Random Forest* e *MLPClassifier* apresentaram desempenho semelhante, com acurácia superior a 97% em todos os casos, o que evidencia sua capacidade de classificar corretamente a grande maioria dos lotes. Dada a proximidade entre os valores de acurácia, precisão e revocação, não foi possível afirmar, com base apenas nessas métricas, a superioridade absoluta de um modelo em relação aos demais. No entanto, observou-se que o *MLPClassifier* apresentou maior sensibilidade na identificação de verdadeiros negativos, ou seja, maior capacidade de reconhecer corretamente os lotes refugados, o que pode ser particularmente relevante em contextos industriais onde a falha em detectar um produto não conforme acarreta perdas significativas. Esse comportamento reforça o potencial da rede neural para lidar com padrões complexos, mesmo que sua interpretabilidade seja mais limitada em relação aos modelos baseados em árvores.

Entretanto, uma análise mais aprofundada das matrizes de confusão revelou limitações significativas na capacidade dos modelos de identificar corretamente os lotes refugados. O número reduzido de verdadeiros negativos e a proporção expressiva de falsos positivos indicam uma dificuldade dos algoritmos em reconhecer padrões associados à classe minoritária. Essa limitação pode estar relacionada tanto ao desequilíbrio na distribuição das classes, quanto ao uso de variáveis estatísticas agregadas por lote, que, embora eficientes do ponto de vista computacional, podem ter comprometido a capacidade dos modelos de capturar variações sutis e relevantes ao longo do processo.

Apesar dessas restrições, os resultados obtidos são promissores e demonstram o potencial da aplicação de técnicas de aprendizado de máquina no controle de qualidade industrial. A automatização da análise de dados operacionais, aliada à capacidade preditiva dos modelos, representa um avanço importante no sentido de reduzir desperdícios, antecipar falhas e aumentar a confiabilidade do processo produtivo. Além disso, o estudo abre caminho para a construção de soluções

mais robustas e sensíveis, capazes de operar em tempo real e de forma integrada à linha de produção.

Como recomendações para pesquisas futuras, destaca-se a adoção de dados brutos coletados ponto a ponto, substituindo as estatísticas consolidadas por informações mais detalhadas e contínuas. Além disso, recomenda-se a implementação de técnicas específicas para lidar com o desbalanceamento das classes, ou ajustes nos algoritmos de aprendizado que favoreçam a correta identificação da classe minoritária. Tais estratégias visam aumentar a sensibilidade dos modelos e, conseqüentemente, sua eficácia na detecção de lotes com não conformidades.

Dessa forma, este trabalho contribui para o avanço das aplicações da Inteligência Artificial no setor industrial, mais especificamente no segmento de papel reciclado, demonstrando a viabilidade e os benefícios do uso de modelos preditivos no apoio à tomada de decisão, à melhoria da qualidade e ao aumento da competitividade.

REFERÊNCIAS

ALPAYDIN, Ethem. **Introduction to Machine Learning**. 4. ed. Cambridge: MIT Press, 2020.

ALVARENGA, Daví Lopes; REIS, Rodrigo Eustáquio dos. **Análise do processo de fabricação do papel utilizando matéria-prima reciclada**. 2018. 34 f. Monografia (Graduação em Engenharia Química) – Universidade Uberaba, Uberaba, 2018. Disponível em: <<https://dspace.uniube.br:8443/bitstream/123456789/1092/1/ANALISE%20DO%20PROCESSO%20DE%20FABRICA%C3%87%C3%83O%20DO%20PAPEL%20UTILIZANDO%20MAT%C3%89RIA%20PRIMA%20REICLADA%20-%20revisado.pdf>>. Acesso em 10 de fev de 2025.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 14021: **Rótulos e declarações ambientais – Autodeclarações ambientais** (Rotulagem do tipo II). Rio de Janeiro, 2004. 26 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 15483 /**Emenda 1:2009 – Aparas de papel e papelão ondulado** – Classificação. Rio de Janeiro, 2009.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 7755: **Papel — Determinação da resistência à ruptura por pressão (Mullen)**. Rio de Janeiro, 1983.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR NM-ISO 12192: **Papel e papelão — Determinação da resistência ao esmagamento em anel (RCT)**. Rio de Janeiro, 2013.

BAJPAI, P. Basic Overview of Pulp and Paper Manufacturing Process. In: BAJPAI, P. (Org.). **Green Chemistry and Sustainability in Pulp and Paper Industry**. Cham: Springer International Publishing, 2015. p. 11-39. Disponível em: <https://doi.org/10.1007/978-3-319-18744-0_2>. Acesso em 10 de fev de 2025.

BORBA, Carla Izidoro de. **Os reflexos da implantação do sistema de gestão da qualidade baseado na NBR ISO 9001:2008 no setor de produção de embalagens de papelão ondulado**. 2012. Trabalho de Conclusão de Curso (Especialização em Celulose e Papel) – Universidade Presbiteriana Mackenzie, São Paulo, 2012. Disponível em: <<https://adelpha-api.mackenzie.br/server/api/core/bitstreams/a6e24cd4-76e4-428b-b1c7-41051c3d13d7/content>>. Acesso em 10 de fev de 2025.

BRAGA, A. P.; CARVALHO, A. P. L. F. de; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2. ed. Rio de Janeiro: LTC, 2007.

BREIMAN, Leo. **Random Forests**. *Machine Learning*, v. 45, p. 5–32, 2001. Disponível em: <<https://link.springer.com/content/pdf/10.1023/a:1010933404324.pdf>>. Acesso em 20 de mar de 2025.

CAMPOS, E. S. **Curso básico de fabricação de celulose e papel**. Eucalyptus Online Book & Newsletter, 2011.

CARDOSO, Marco Túlio *et al.* **Propriedades físicas e mecânicas de papéis reciclados utilizados para fabricação de tubetes**. Ciência Florestal, v. 22, n. 2, p. 403-411, 2012. Disponível em: <<https://www.scielo.br/j/cflo/a/Rfp3kFzj64HfYLrv8X6ZYfp/?format=pdf&lang=pt>>. Acesso em 7 de Jan de 2025.

DATA CAMP. **Introdução ao aprendizado não supervisionado**. 2024. Disponível em: <<https://www.datacamp.com/pt/blog/introduction-to-unsupervised-learning>>. Acesso em: 12 abr. 2025.

DEVI, T. K.; PRIYANKA, E. B.; SAKTHIVEL, P. **Paper quality improvement and prediction models using machine learning techniques**. Results in Engineering, v. 17, 2023. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2590123023000774>>. Acesso em 07 de jan de 2025.

DIAS FLAMIA, R.; HOFFMANN, K.; JANUÁRIO, M.; PATRICK PRIGOL, J. **Determinação da Qualidade de Formação do Papel Reciclado utilizando Redes Neurais Convolucionais**. Disponível em: <[https://periodicos.unoesc.edu.br/siepe/areminário de Iniciação Científica e Seminário Integrado de Ensino, Pesquisa e Extensão \(SIEPE\)](https://periodicos.unoesc.edu.br/siepe/areminário%20de%20Iniciac%C3%A3o%20Cient%C3%ADfica%20e%20Semin%C3%A1rio%20Integrado%20de%20Ensino%20Pesquisa%20e%20Extens%C3%A3o%20(SIEPE)), [S. l.], p. e28505, 2021. Disponível em: <[ticle/view/28505](https://periodicos.unoesc.edu.br/siepe/areminário%20de%20Iniciac%C3%A3o%20Cient%C3%ADfica%20e%20Semin%C3%A1rio%20Integrado%20de%20Ensino%20Pesquisa%20e%20Extens%C3%A3o%20(SIEPE))>. Acesso em: 26 dez. 2024.

DIAS, Henrique; ARAÚJO, Fernando Rafael da C.; MOREIRA, Kleber da Silveira; DIAS, Marcio José. **A influência da inteligência artificial (IA) nos processos de automação industrial**. Revista Aracê, v. 6, n. 4, p. 15373-15392, 2024. Disponível em: <https://editoraes.com.br/wp-content/uploads/2024/05/A-INFLUENCIA-DA-INTELIGENCIA-ARTIFICIAL-IA-NOS-PROCESSOS-DE-AUTOMACAO-INDUSTRIAL1.pdf?utm_source=chatgpt.com>. Acesso em: 22 dez. 2024.

ELE, Z.; QIAN, J.; LI, J.; HONG, M.; MAN, Y. **Data-oriented soft sensors for paper manufacturing process and their application to cleaner production with multi-objective optimization**. Journal of Cleaner Production, v. 372, 2022. Disponível em: <https://doi.org/10.1016/j.jclepro.2022.133803>. Acesso em: 15 dez. 2024.

FORUGHI, Amir Farzad; GREEN, Sheldon; STOEBER, Boris. **Effect of dryer fabric structure on the performance of contact paper drying**. Drying Technology, 2018. DOI: 10.1080/07373937.2018.1469141. Acesso em: 22 dez. 2024.

FRIEDMAN, J. H. **Greedy function approximation: A gradient boosting machine**. Annals of Statistics, 2001.

GALLON, Alessandra Vasconcelos; SALAMONI, Franciane Luiza; BEUREN, Ilse Maria. **O processo de fabricação de papel reciclado e as ações associadas aos custos ambientais em indústria de Santa Catarina**. ABCustos, São Leopoldo: Associação Brasileira de Custos, v. 3, n. 1, p. 53-80, jan./abr. 2008. Disponível em: <

<https://revista.abcustos.org.br/abcustos/article/view/28>>. Acesso em 08 de fev de 2025.

GÉRON, Aurélien. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. 2. ed. Sebastopol: O'Reilly Media, 2019.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2010

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. Cambridge: MIT press, 2016.

GRINSZTAJN, Léo; OYALLON, Edouard; VAROQUAUX, Gaël. **Why do tree-based models still outperform deep learning on tabular data?** Journal of Machine Learning Research, v. 23, p. 1-27, 2022. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2022/file/0378c7692da36807bdec87ab043cdadc-Paper-Datasets_and_Benchmarks.pdf>. Acesso em 23 de mar de 2025.

HAN, Jiawei; KAMBER, Micheline; PEI, Jian. **Data mining: concepts and techniques**. 3rd ed. San Francisco: Morgan Kaufmann, 2011.

HAYKIN, Simon. **Redes neurais: princípios e prática**. 3. ed. Porto Alegre: Bookman, 2009

HE, Y.; LIU, W.; DENG, Y.; LI, J. **Quality improvement in recycled paper production: A review of challenges and strategies**. Journal of Cleaner Production, [S. l.], v. 365, 2022. Disponível em: <https://doi.org/10.1016/j.jclepro.2022.132634>. Acesso em: 15 dez. 2024

HEO, Chang Hoe; CHO, Hyunjun; YEO, Yeong-Koo. **Dynamic modeling of paper drying processes**. Korean Journal of Chemical Engineering, v. 28, n. 8, p. 1651-1657, 2011. DOI: 10.1007/s11814-011-0046-0. Acesso em: 22 dez. 2024.

HOMEM, Y.; HU, S.; GAO, J.; LI, J.; HONG, M. **Combustão química integrada em fábrica de celulose para alta eficiência energética e baixa emissão de carbono**. Jornal de Produção Mais Limpa, v. 275, 2020. Disponível em: <https://doi.org/10.1016/j.jclepro.2020.122979>. Acesso em: 15 dez. 2024.

IBM. **O que é matriz de confusão?**. 2024. Disponível em: <<https://www.ibm.com/br-pt/think/topics/confusion-matrix>>. Acesso em: 12 abr. 2025.

IBM. **What is Random Forest?**. 2024. Disponível em: <https://www.ibm.com/topics/random-forest>. Acesso em: 12 abr. 2025.

INDÚSTRIA BRASILEIRA DE ÁRVORES (IBÁ). **Relatório Anual 2024**. São Paulo, 2024. Disponível em: <https://www.iba.org>. Acesso em: 15 dez. 2024.

JAMES, Gareth *et al.* **An Introduction to Statistical Learning: with Applications in R**. 2. ed. New York: Springer, 2021.

KAGERMANN, H.; WAHLSTER, W.; HELBIG, J. **Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry**. Final report of the Industrie 4.0 Working Group. Frankfurt: National Academy of Science and Engineering, 2013. Disponível em: <https://www.din.org/en/innovation-and-research/industrie-4-0>. Acesso em: 22 dez. 2024.

KÜHL, Diumara; REIS, Diogo; SILVA, Anaisa Lourenço da; SANDRINO, Bianca. **Resíduos da Indústria Papeleira - Tratamentos e Aplicações**. Faculdade Sant'Ana em Revista, [S. l.], v. 8, n. 2, p. p. 469 – 483, 2024. Disponível em: <https://iessa.edu.br/revista/index.php/fsr/article/view/2540>. Acesso em: 29 dez. 2024.

LEMAÎTRE, Guillaume; NOGUEIRA, Fernando; ARIDAS, Christos K. **Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning**. Journal of Machine Learning Research, v. 18, n. 17, p. 1–5, 2017. Disponível em: <https://www.jmlr.org/papers/volume18/16-365/16-365.pdf>. Acesso em 04 de abr de 2025.

LIU, Hongbin; ZHANG, Yuchen; ZHANG, Hao. **Prediction of effluent quality in papermaking wastewater treatment processes using dynamic kernel-based extreme learning machine**. Process Biochemistry, v. 97, p. 72–79, 2020. Disponível em: <https://doi.org/10.1016/j.procbio.2020.06.020>. Acesso em: 22 dez. 2024.

MACEDO, Angela Regina Pires; VALENÇA, Antonio Carlos de Vasconcelos. **A indústria de papel no Brasil e no mundo: uma visão geral**. Rio de Janeiro: BNDES, 1995.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 8. ed. São Paulo: Atlas, 2017.

MERLIN, H. A.; CALEMAN, S. M. de Q.; PADGETT, R. C. M. L. **Organização da indústria brasileira de papel e celulose: reflexões à luz da nova economia institucional**. Caderno de Administração, Maringá, v. 32, n. 1, p. 120-145, jan./jun. 2024. DOI: <https://doi.org/10.4025/cadadm.v32i1.67201>. Disponível em: <https://www.redalyc.org/articulo.oa?id=733878577006>. Acesso em: 15 dez. 2024.

MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS (MDIC). **Indústria de papel e celulose anuncia investimentos de R\$ 105 bilhões**. Brasília, 2024. Disponível em: <https://www.gov.br/mdic/pt-br/assuntos/noticias/2024/agosto/industria-de-papel-e-celulose-anuncia-investimentos-de-r-105-bilhoes>. Acesso em: 15 dez. 2024.

MORAES, M. D. S.; SANTOS, M. A. W. D.; GUZMAN, E. F.; SANTOS, S. D.; MORO, A. R. P. **Um estudo sobre a reciclagem de papel: Um panorama desta atividade no Brasil**. Artigo. Semana Acadêmica. Revista Científica. ISSN 2236-6717. Fortaleza. Ceará, 2015. Disponível em: <https://semanaacademica.org.br/system/files/artigos/reciclagem_de_papel_um_pano_rama_dest_a_atividade_no_brasil.pdf>. Acesso em 14 de fev de 2025.

MURPHY, Kevin P. **Machine Learning: A Probabilistic Perspective**. Cambridge: MIT Press, 2012.

OLIVEIRA, C. F. *et al.* **Propriedades físicas e mecânicas de papéis reciclados utilizados para fabricação de tubetes**. ResearchGate, 2015. Disponível em: < <https://periodicos.ufsm.br/cienciaflorestal/article/view/5748/3389> >. Acesso em: 07 mar. 2025.

PEREIRA, Lucas Lucena de Moraes. **Soft sensor empregando rede neural recorrente LSTM para estimação da gramatura numa máquina de papel**. 2023. 59 f. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Química) – Centro de Tecnologia, Universidade Federal da Paraíba, João Pessoa, 2023. Disponível em: < https://repositorio.ufpb.br/jspui/bitstream/123456789/27905/1/TCC_LUCAS%20LUCENA%20DE%20MORAIS%20PEREIRA.pdf >. Acesso em 8 de jan de 2025.

PESCIO, Paulo Henrique. **Análise da integração da tecnologia da automação aplicada ao processo de fabricação de papel**. 2016. 112 f. Dissertação (Mestrado em Sistemas de Potência) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2016. Disponível em: <https://www.teses.usp.br/teses/disponiveis/3/3143/tde-29082016-154458/publico/PauloHenriquePescioOrig16.pdf>. Acesso em: [data de acesso].Potência). Acesso em 20 de jan de 2025.

RASCHKA, Sebastian; MIRJALILI, Vahid. **Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2**. 3. ed. Birmingham: Packt Publishing, 2019.

RIGHETTI, Carlos Cesar; RODRIGUES, Ivete; FACÓ, Júlio F. B.; SAKURAMOTO, Carlos Y.; BARBIERI, José Carlos. **Estratégias de gestão ambiental nas empresas: um estudo de caso sobre o papel reciclado**. In: ENCONTRO DA ANPAD (ENANPAD), 29., 2005, Brasília. Anais [...]. Brasília: ANPAD, 2005. Disponível em: < https://arquivo.anpad.org.br/diversos/down_zips/9/enanpad2005-apsc-2227.pdf >. Acesso em 22 de fev de 2025.

SANTOS, Frederico de Oliveira. **Inovação, tecnologia e o uso potencial do aprendizado de máquina na fabricação inteligente para detecção de anomalias nos processos de produção**. 2021. 158 f. Dissertação (Mestrado em Desenvolvimento e Sociedade) – Universidade do Alto Vale do Rio do Peixe – União, Caçador, 2021.

SARWAR, Muhammad. **Performance evaluation of machine learning classifiers in quality prediction for the paper manufacturing industry**. International Journal of Advanced Computer Science and Applications, v. 11, n. 8, 2020. DOI: 10.14569/IJACSA.2020.0110872.

SCHROTH, Moritz *et al.* **Optimization of paper production through digitalization by developing an assistance system for machine operators including quality forecast: a concept**. 2022. Disponível em: <https://arxiv.org/abs/2206.11581>. Acesso em: 15 dez. 2024.

SILVA, Gislaine Izidio da Guia. **Produção e caracterização de papel reciclado obtido com aproveitamento de resíduo de jardim, material foliar da espécie Terminalia catappa**. 49 f. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Ambiental) – Universidade Estadual Paulista (UNESP), Instituto de Ciência e Tecnologia, Sorocaba, 2023. Disponível em: <
<https://repositorio.unesp.br/server/api/core/bitstreams/4239dec2-b519-4689-8d67-3057c513c6ba/content>>. Acesso em 22 de abr de 2025.

SOUSA, Derlicio Carlos Goes; MATOS, Leandro Lisboa; ARAUJO, Myllane Kelry Sa; LIMA, Elon Vieira. **A importância da reciclagem do papel na melhoria da qualidade do meio ambiente**. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO (ENEGEP), 36., 2016, João Pessoa. Anais [...]. João Pessoa: ABEPRO, 2016. Disponível em: <
https://d1wqtxts1xzle7.cloudfront.net/54495043/2016_-_ENEGEP_A_importancia_da_reciclagem_do_papel-libre.pdf?1505997756=&response-content-disposition=inline%3B+filename%3DA_IMPORTANCIA_DA_RECICLAGEM_DO_PAPEL_NA.pdf&Expires=1744590618&Signature=cZfdFshdi5fCRx7RDvwGaNrxNWyeRSDqByW2hxVbQxc~-7G~ToShZsZ60cxCNcS3Jn~F~MNMt~brKMrU4Bmyl9j23-MP8zQJ~D8T5e7x9dS~Bo9Nfk181WgJjslqC6RYeKsbuUBOX2UuawOi3GWPz--ZJgfWrr~LCDMVuzFypaHk2iwTCw5BGgnQlbwAhAu~YeLLXrebBwgXUD0JBisM8ItT81goZGQWYu9YGrM9qhePriEP1Iri10qMUp1cS5Bt~qSYpreOr1H4KYBJn0CTTFTpilud6tfO-qcJZ7Q2a20HoL8dxSbKTKQGaiuzfKaSAIvsZH9xFDDHLKfyIGcUw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA>. Acesso em 25 de jan de 2025.

SOUZA, Alan Christian Almeida; SILVA, Paula Renatha Nunes da. **Aplicação da inteligência artificial na engenharia de confiabilidade e manutenção preditiva: um estudo de caso na indústria de mineração**. Revista Ibero-Americana de Humanidades, Ciências e Educação, São Paulo, v. 10, n. 10, p. 3646-3658, out. 2024. ISSN 2675-3375. Disponível em: <https://doi.org/10.51891/rease.v10i10.16252>. Acesso em: 22 dez. 2024.

SRINIVASAN, S.; MURUGAN, S.; SWAGTHY, R.; OLIVEIRA, C.; SRIVIDHYA, V.; MUTHULEKSHMI, M. **Data-driven recycling transformation to enhance the efficiency of paper and cardboard bins using IoT and Random Forest**. In: AMATHE 2024: International Conference on Advances in Modern Technology for Health Sciences and Engineering, Anais [...], 2024. DOI: 10.1109/AMATHE61652.2024.10582250. Disponível em: <https://ieeexplore.ieee.org/document/10582250>. Acesso em: 15 dez. 2024.

TEIXEIRA, M. B. D.; OLIVEIRA, R. A.; GATTI, T. H.; SUAREZ, P. A. Z. **O Papel: Uma Breve Revisão Histórica, Descrição da Tecnologia Industrial de Produção e Experimentos para Obtenção de Folhas Artesanais**. Revista Virtual de Química, v. 9, n. 3, p. 1364-1380, 2017. Disponível em: <
https://d1wqtxts1xzle7.cloudfront.net/65980332/1984-6835-libre.pdf?1615528209=&response-content-disposition=inline%3B+filename%3DThe_Paper_A_Brief_Historical_Review_a_De.pdf&Expires=1744590710&Signature=TzUs6AW-qniOfjO6RsSkaiVPqCdMJK2S3jMj1JuAqw5J2igvF5766xADS6aydWfo7bj0e1dOZuu6XT2DZIMEwlOcRzah1MCDFOHrxwFvLp4jbQluhUxoDgORfD-

9EyxLOteUZ0TmQjrNygkfn2BH5YMTL6HzFz7DjsiT4piYzND-2z1L45vw-u0a1imWx000sgTuoovURSk6h0W~WLH6LiHsnE5nzzeXrWSIOMVfUgEwsXZ2JcpPInAChxVdQXPFLuoylxfQSVob6S6nfxDFPyu7AFgYAYi4b3BKHKGFXZ7Rf9fracq33S9m8FQkOoeN9IAH4CXrFqYH~KZXVvRGheQ__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA>. Acesso em 14 de fev de 2025.

TREVOR, Hastie; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning**. 2. ed. Springer, 2009.

UC-R. **Gradient Boosting Regression**. UC Business Analytics R Programming Guide, 2024. Disponível em: <https://uc-r.github.io/gbm_regression>. Acesso em: 12 abr. 2025.

WAN, Xin *et al.* **Water quality prediction model using Gaussian process regression based on deep learning for carbon neutrality in papermaking wastewater treatment system**. Environmental Research, v. 211, p. 112942, 2022. Disponível em: <https://doi.org/10.1016/j.envres.2022.112942>. Acesso em: 22 dez. 2024.

ZHANG, Jiwei *et al.* **An intelligent digital twin system for paper manufacturing in the paper industry**. Expert Systems with Applications, v. 230, p. 120614, 2023. Disponível em: <https://doi.org/10.1016/j.eswa.2023.120614>. Acesso em: 22 dez. 2024.