



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

RAPHAEL ARANTES E SILVA

**MONITORAMENTO DA QUALIDADE DE EMBALAGENS DE PET COM O USO DE
FERRAMENTAS DE APRENDIZADO DE MÁQUINA**

Recife
2024

RAPHAEL ARANTES E SILVA

**MONITORAMENTO DA QUALIDADE DE EMBALAGENS DE PET COM O USO DE
FERRAMENTAS DE APRENDIZADO DE MÁQUINA**

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

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

Recife
2024

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

Silva, Raphael Arantes.

Monitoramento da qualidade de embalagens de pet com o uso de
ferramentas de aprendizado de máquina / Raphael Arantes Silva. - Recife,
2024.

55p : 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, 2024.

1. Aprendizado de máquina. 2. Indústria. 3. Resina PET. 4. Parâmetros . 5.
Físico-químico. I. da Costa Gomes, Felipe Pedro. (Orientação). II. Título.

660 CDD (22.ed.)

RAPHAEL ARANTES E SILVA

MONITORAMENTO DA QUALIDADE DE EMBALAGENS DE PET COM O USO DE FERRAMENTAS DE APRENDIZADO DE MÁQUINA

Trabalho de conclusão de curso apresentado à Coordenação do Curso de Graduação em Engenharia Química da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de Engenheiro Químico.

Aprovado em: 25/03/2024.

BANCA EXAMINADORA

Documento assinado digitalmente
 **FELIPE PEDRO DA COSTA GOMES**
Data: 08/04/2024 14:06:07-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Felipe Pedro da Costa Gomes (Orientador) Universidade Federal de Pernambuco - UFPE

Documento assinado digitalmente
 **LEANDRO DANIELSKI**
Data: 08/04/2024 15:40:32-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Leandro Danielski (1º Examinador (a)) Universidade Federal de Pernambuco - UFPE

Documento assinado digitalmente
 **GRAZIELLE ROZENDO DE CERQUEIRA**
Data: 08/04/2024 16:16:03-0300
Verifique em <https://validar.iti.gov.br>

Dra. Grazielle Rozendo de Cerqueira (2º Examinador (a)) Universidade Federal de Pernambuco - UFPE

DEDICATÓRIA

A minha família.

AGRADECIMENTOS

Primeiramente gostaria de agradecer a minha vó, minha tia e meu pai que me deram todo suporte durante a graduação.

Gostaria de agradecer aos meus amigos do curso de engenharia química que me apoiaram e fizeram a minha jornada ser incrível.

Gostaria de agradecer ao meu orientador, o professor Felipe Pedro da Costa Gomes, por me orientar em todo o processo do TCC.

Agradecer, ainda, à Universidade Federal de Pernambuco que me permitiu conhecer o mundo.

RESUMO

Com o desenvolvimento recente das tecnologias digitais e a necessidade de melhorar processos, decisões e recursos, as indústrias estão cada vez mais investindo em sistema de informações para reduzir erros e perdas em seus processos internos. Dessa maneira, a produção de resina PET tem o desafio de atender todos os parâmetros dentro do estabelecido pelo produtor de embalagens, onde qualquer variabilidade pode afetar o seu processamento e a qualidade dos recipientes, tendo a necessidade de investimentos de melhorias e oportunidades para evitar falhas e prejuízos. A obtenção de informações confiáveis e precisas é crucial para tomadas de decisões eficazes e o uso de ferramentas de aprendizado de máquina ou *machine learning* (ML) tem sido fundamental para contribuir com o ganho de resultados e evitar prejuízos. Nesse contexto, o presente trabalho tem como objetivo aplicar técnicas de ML com o uso da linguagem de programação Python e algoritmo de predição a partir dos dados dos parâmetros físico-químicos obtidos de um banco de dados histórico de produção da resina PET, buscando prever o risco de ocorrência de falha gerada por alteração nas variáveis na produção de pré-formas e garrafas. Diante disso, foram utilizadas 3 técnicas de aprendizado de máquina, k-vizinhos mais próximos (KNN), máquina de vetor de suporte (SVM) e árvore de decisão, com a aplicação das métricas de desempenho através da comparação dos modelos pela acurácia, precisão e revocação. Apesar da análise preliminar demonstrar uma baixa correlação entre as variáveis, os resultados dos modelos de predição foram satisfatórios com destaque para o modelo de árvore de decisão que obteve resultado de 99,81% para a métrica da acurácia e 100% para precisão e revocação.

Palavras-chave: Aprendizado de Máquina, Industria, resina PET, Parâmetro físico-químico

ABSTRACT

With the recent development of digital technologies and the need to improve processes, decisions and resources, industries are increasingly investing in information systems to reduce errors and losses in their internal processes. Therefore, the production of PET resin has the challenge of meeting all configurations within what is foreseen by the packaging producer, where any variability can affect its processing and the quality of the containers, requiring investment in improvements and opportunities to avoid failures. and losses. Obtaining reliable and accurate information is crucial for effective decision-making and the use of machine learning (ML) tools has been essential to contribute to gaining results and avoiding losses. In this context, the present work aims to apply ML techniques using the Python programming language and prediction algorithm based on data on physical-chemical parameters obtained from a historical database of PET resin production, seeking to predict the risk of failures caused by changes in variations in the production of preforms and bottles. Therefore, 3 machine learning techniques were used, k-nearest neighbors (KNN), support vector machine (SVM) and decision tree, with the application of performance metrics through the comparison of models by accuracy, precision and revocation. Despite the preliminary analysis demonstrating a low observation among the variations, the results of the prediction models were impressive, with emphasis on the decision tree model, which obtained results of 99.81% for the accuracy metric and 100% for results and revocation.

Key-words: Machine Learning, Industry, PET resin, Physical-Chemical Parameters.

LISTA DE FIGURAS

Figura 1 -	Estrutura molecular do polietileno tereftalato.	17
Figura 2 -	Reações de Pré-Polimerização do PET.	19
Figura 3 -	Resina PET amorfa.	20
Figura 4 -	Resina PET cristalizada.	21
Figura 5 -	Reação do processo termodegradativo do PET em acetaldeído.	22
Figura 6 -	Máquina injetora usada para o processamento da resina PET.	24
Figura 7 -	Matriz de confusão.	26
Figura 8 -	Exemplo de classificação por meio do KNN.	28
Figura 9 -	Exemplo de SVM para Classificação.	29
Figura 10 -	Distribuição dos Registros de Notificações.	34
Figura 11 -	Distribuição dos dados por aplicação de resina.	35
Figura 12 -	Distribuição dos dados por Linhas de Produções.	36
Figura 13 -	Mapa de calor das variáveis.	37
Figura 14 -	Gráfico da correlação entre a IV e CL com base na Notificação.	38
Figura 15 -	Gráfico da correlação entre a IV e CB com base na Notificação.	38
Figura 16 -	Gráfico da correlação entre a IV e AD com base na Notificação.	39
Figura 17 -	Gráfico da correlação entre a CL e CB com base na Notificação.	39
Figura 18 -	Gráfico da correlação entre a CL e AD com base na Notificação.	40
Figura 19 -	Gráfico da correlação entre a CB e AD com base na Notificação.	40
Figura 20 -	Gráfico da correlação entre a IV e CL com base no Tipo da resina.	41
Figura 21 -	Gráfico da correlação entre a IV e CB com base no Tipo da resina.	42
Figura 22 -	Gráfico da correlação entre a IV e AD com base no Tipo da resina.	42
Figura 23 -	Gráfico da correlação entre a CL e CB com base no Tipo da resina.	43
Figura 24 -	Gráfico da correlação entre a CL e AD com base no Tipo da resina.	43
Figura 25 -	Gráfico da correlação entre a CB e AD com base no Tipo da resina.	44
Figura 26 -	Gráfico da correlação entre a IV e CL com base na Linha de produção.	45
Figura 27 -	Gráfico da correlação entre a IV e CL com base na Linha de produção.	45
Figura 28 -	Gráfico da correlação entre a IV e CL com base na Linha de produção.	46
Figura 29 -	Gráfico da correlação entre a IV e CL com base na Linha de produção.	46
Figura 30 -	Gráfico da correlação entre a IV e CL com base na Linha de produção.	47
Figura 31 -	Gráfico da correlação entre a IV e CL com base na Linha de produção.	47

LISTA DE TABELAS

Tabela 1 -	Quantidade de atributo por categoria.	32
Tabela 2 -	Resultados do modelo KNN com base na IV e a Notificação	48
Tabela 3 -	Resultados do modelo KNN com base na IV e AD e a Notificação	48
Tabela 4 -	Resultados do modelo KNN com base na IV, CL, CB e AD e a Notificação	48
Tabela 5 -	Resultados do modelo SVM com base na IV e a Notificação	49
Tabela 6 -	Resultados do modelo SVM com base na IV e AD e a Notificação	49
Tabela 7 -	Resultados do modelo SVM com base na IV, CL, CB e AD e a Notificação	49
Tabela 8 -	Resultados do modelo árvore de decisão com base na IV e a Notificação	50
Tabela 9 -	Resultados do modelo árvore de decisão com base na IV e AD e a Notificação	50
Tabela 10 -	Resultados do modelo árvore de decisão com base na IV, CL, CB e AD e a Notificação	50

LISTA DE ABREVIATURAS E SIGLAS

ABIPET	Associação Brasileira da Indústria de PET
ABIPLAST	Associação Brasileira da Indústria de Plástico
AD	Acetaldeído
BHET	bis-2-hidroxietil-tereftalato
CL	Cor L*
CB	Cor b*
DP	<i>Deep Learning</i>
IA	Inteligência Artificial
ISBM	<i>Injection Stretch Blow Molding</i> - Moldagem por Sopro Estiramento por Injeção
IV	Viscosidade Intrínseca
KNN	K-Vizinhos Mais Próximos
ML	<i>Machine Learning</i>
PET	Polietileno Tereftalato
PLN	Processamento de Linguagem Natural
SVM	Máquina de vetores de suporte
T _c	temperatura de cristalização
T _g	temperatura de transição vítrea
T _m	temperatura de transição vítrea

SUMÁRIO

1.	INTRODUÇÃO.....	14
1.1	OBJETIVO GERAL.....	16
1.1.1	Objetivos específicos.....	16
2.	FUNDAMENTAÇÃO TEÓRICA.....	17
2.1	POLI TERFETALATO DE ETILENO (PET).....	17
2.2	PROCESSO DE PRODUÇÃO DE RESINA PET.....	19
2.3	PRINCIPAIS PARÂMETROS DA RESINA PET.....	21
2.3.1.	Viscosidade.....	21
2.3.2.	Acetaldeído.....	22
2.3.3.	Colorimetria.....	22
2.3.4.	Temperatura de Transição.....	23
2.3.5.	Umidade.....	23
2.4	PROCESSAMENTO DA RESINA PET.....	24
2.5	MACHINE LEARNING.....	24
2.6	TÉCNICAS DE MACHINE LEARNING.....	28
2.6.1.	K-vizinhos mais próximos (KNN).....	28
2.6.2.	Máquinas de vetor de suporte.....	29
2.6.3.	Árvore de decisão.....	30
3.	METODOLOGIA.....	31
3.1	CARACTERIZAÇÃO DO PROCESSO.....	31
3.2	BASE DE DADOS.....	31
3.3	APRENDIZADO DE MÁQUINA.....	32
4.	RESULTADOS E DISCUSSÕES.....	34
4.1	ANÁLISE DETALHADA DA BASE DE DADOS.....	34
4.2	ANÁLISE DOS RESULTADOS DOS MODELOS PREDITIVOS.....	38
5.	CONCLUSÃO.....	52
6.	REFERÊNCIAS.....	53

1. INTRODUÇÃO

No mundo altamente competitivo onde as empresas estão incessantemente buscando por aperfeiçoamento, as indústrias têm direcionado investimentos para implementar melhorias em seus processos, visando aprimorar a qualidade de seus produtos e serviços. Com o desenvolvimento recente das tecnologias digitais e eletrônicas, juntamente com métodos automáticos de trabalhos avançados, aliados a modelos matemáticos e à inteligência computacional, a indústria 4.0 se mostra bastante eficaz a melhorar seus modelos de negócios e otimizar suas produções (BOMFIM, 2021).

Nesse contexto, as organizações que adotam o uso de Inteligência Artificial (IA) podem notar que há um poder benéfico, independentemente do setor que estão inseridas. Isso porque a utilização da IA além de gerar automação mecânica, abrange processos cognitivos que são capazes de gerar aprendizado para executar atividades repetitivas, numerosas e manuais, podendo também ter a responsabilidade de analisar e tomar decisões. (GOMES, 2010).

Dentre os atributos fundamentais da IA utilizados, destacam-se o ML *machine learning*, que emprega métodos de análise de dados para automatizar a identificação de padrões em desenvolvimento, o *deep learning* (DL), que enfatiza o uso de redes neurais artificiais com múltiplas camadas de abstração, especialmente aplicado no reconhecimento de padrões e o Processamento de Linguagem Natural (PLN), que se dedica a estudar maneiras de replicar os processos de desenvolvimento associados ao funcionamento da linguagem humana (COSSETTI, 2019).

Desse modo, o uso dos recursos fundamentais da IA, como as técnicas de ML, tem sido aplicado em diversos setores das indústrias para auxiliar na análise de dados e tomada de decisão, para isso, saber qual método utilizar é essencial para que a implementação dessas tecnologias seja totalmente efetiva. Essas aplicações são encontradas na área de extração, transformação, energética, construção civil e de embalagens. Com isso, uso dessas ferramentas no âmbito de embalagens é de extrema importância, com destaque na produção de resinas e garrafas de poli tereftalato de etileno (PET), por se tratar de segurança alimentar que essa embalagem fornece para a humanidade (BOMFIM, 2021).

O polietileno tereftalato ou o poli (tereftalato de etileno) ou pela sua forma amplamente reconhecida, PET, é um polímero que pertence à categoria dos poliésteres termoplásticos por possuir um grupo funcional éster na sua cadeia principal e que pode ser reprocessado diversas vezes pelo mesmo ou por outro processo de transformação. Com o destaque de suas

propriedades, como, excelente em barreiras a gases e odores, baixo peso, maleabilidade, cristalinidade, transparência e resistência mecânica, possibilitou que esse material fosse aplicado para a fabricação de fibras têxteis, embalagens e filmes, garantindo assim uma utilização significativa, sobretudo, na indústria alimentícia (STRONG, 2006).

Dado que o PET é um polímero cujo a matéria-prima principal é derivado do petróleo que passa por processos de polimerização, copolimerização, extrusão, injeção e sopro para a produção de garrafas a serem envasadas, principalmente com refrigerantes e água mineral, existem diversas normas sanitárias que buscam controlar e estabelecer limites de migração de substâncias nas bebidas. Diante disso, a resina PET, além de suprir as necessidades técnicas do processamento de injeção e sopro que submete o material a condições extremas de temperatura e cisalhamento, deve garantir também uma excelente aparência visual (SCOTT, 2000).

Em vista disso, os parâmetros da resina PET devem atender a todas as especificações dos parâmetros técnicos pré-estabelecidos para que o processamento seja realizado pelo produtor de garrafas de forma eficiente, assegurando também a saúde física do consumidor final. Por se tratar de uma longa cadeia, o não atendimento dos parâmetros da resina pode ocasionar elevados prejuízos financeiros à indústria de produção, assim como, o comprometimento do negócio mediante o dano na integridade física do cliente final (SILVA, 2022).

Por este motivo, esse trabalho discute sobre os principais parâmetros técnicos da resina PET conforme a sua aplicação e como eles podem influenciar no processamento de injeção e sopro na formação de garrafas para envase de água e refrigerantes. A partir dos dados dos parâmetros será analisado por meios de técnicas de ML com o uso de linguagem de programação e algoritmo de predição a possibilidade de ocorrer falha gerada por alterações nas variáveis no processo de produção das garrafas.

1.1. OBJETIVO GERAL

Correlacionar, a partir de técnicas de ML, os parâmetros técnicos da produção da resina PET com as principais implicações do seu uso e do possível impacto do processamento da resina pelo produtor de garrafas.

1.1.1 Objetivos específicos

Além disso, o trabalho visou atingir os seguintes objetivos específicos:

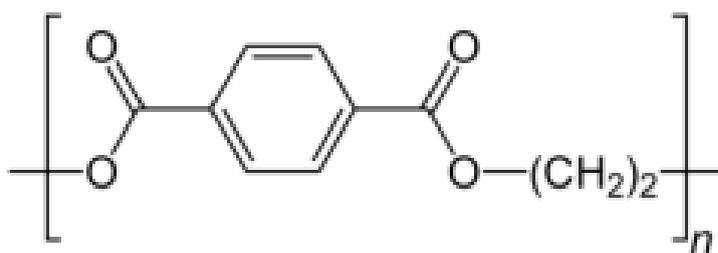
- Obter e preparar a base de dados a partir do histórico de uma produção de resina PET para aplicação ML;
- Implementar técnicas de aprendizado de máquina a partir de linguagem Python capazes de prever os impactos dos parâmetros da resina no processo de injeção e sopro;
- Comparar a eficiência dos modelos aplicados a partir dos resultados obtidos.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. POLI TEREFTALATO DE ETILENO (PET)

Os materiais poliméricos são utilizados desde a antiguidade, no entanto, a história do PET se inicia em 1941 em contexto da Segunda Guerra Mundial, em que o conflito gerou desabastecimentos em diversos setores, incluindo o têxtil, afetando as principais matérias-primas da época, como algodão e lã. Em 1946 os químicos britânicos John Rex Whinfield e James Tennant Dickson, sintetizaram pela primeira vez o polietileno tereftalato, representado na Figura 1, o produto da reação entre o ácido tereftálico e o etileno glicol, dando origem ao polímero termoplástico que inicialmente fora usado para suprir a crise têxtil (COUTINHO et al, 2003).

Figura 1 - Estrutura molecular do polietileno tereftalato.



Fonte: Wikipedia, 2015.

A partir da década de 1960 o poli tereftalato de etileno passa a ser usado como material de embalagem na forma de filme biorientado nos Estados Unidos e depois na Europa. Após o desenvolvimento da tecnologia do processo de injeção e sopro pela DuPont no ano em 1973 foi desenvolvida pela primeira vez a garrafa de poliéster empregado no envase de refrigerantes e águas minerais. O crescimento global do PET fez com que a tecnologia chegasse ao Brasil em 1998, sendo atualmente um dos três maiores produtores de garrafas PET do mundo, ficando só atrás do Estados Unidos e México (TOLEDO et al, 2008).

Além disso, segundo a Associação Brasileira da Indústria de PET (ABIPET), em 2020 a produção de resina virgem do material chegou a 596 mil toneladas no ano, como também, segundo a Associação Brasileira da Indústria de Plástico (ABIPLAST) o PET foi o 6º plástico

mais consumido dentro do país no ano de 2019, correspondendo cerca de 8,5% da produção, onde o setor de alimentos e bebidas correspondeu por mais de um quarto de todo o consumo desse material no Brasil no mesmo ano.

O sucesso e a substituição das garrafas de vidro estão relacionados às propriedades do PET e ao custo-benefício apresentado pelo material. Antigamente, apenas o vidro tinha a capacidade de barrar gases e vapores e manter propriedades ópticas. Com o advento do PET e a descoberta de suas propriedades, uma grande revolução no mercado de embalagens surgiu (UCHIDA, 2022).

Entender as propriedades do PET é fundamental para entender com mais consistência porque esse material se mostrou bastante atrativo no mercado de embalagens para alimentos e bebidas. As importantes características, como elevada resistência ao calor, leveza, baixa permeabilidade aos gases CO₂ e O₂, boa estabilidade química e bom isolante elétrico, além de, uma alta temperatura de fusão (275°C) e transição vítrea (70°C), permite que esse material possua alta rigidez e baixa mobilidade e consiga manter suas propriedades mecânicas mesmos em altas temperaturas, faz com que o uso desse material seja bastante aplicado na indústria de embalagens (HARADA et al, 2005).

Suas cadeias podem se organizar na forma cristalina ou amorfas, dependendo dessas organizações, podemos ter diferentes tipos de aplicações e propriedades. Desse modo, o PET na forma cristalina, forma barreiras que reduzem a permeabilidade de umidade e a difusão de gases. Essa propriedade explica o porquê do PET ser uma excelente alternativa para o armazenamento de bebidas. Além disso, a baixa permeabilidade permite que os gases fiquem presos nas garrafas e que as bebidas não percam o efeito desejado (PREVILATE, 2022).

O PET também apresenta a vantagem de ser um polímero que pode receber pigmentos de diferentes cores e tons. É por isso que, no mercado, são encontradas embalagens PET transparentes, azuladas e verdes. Outra característica importante do material é sua baixa densidade. Quando comparado com o vidro, o PET é bastante vantajoso para o uso em embalagens por conta desta característica também. Uma carreta comum pode transportar 60% mais produto se sua carga for feita com garrafas PET no lugar do vidro. A densidade do PET está entre 1,37 g/mL e 1,45 g/mL enquanto a do vidro para embalagem é de 2,5 g/mL (PREVILATE, 2022).

2.2. PROCESSO DE PRODUÇÃO DE RESINA PET

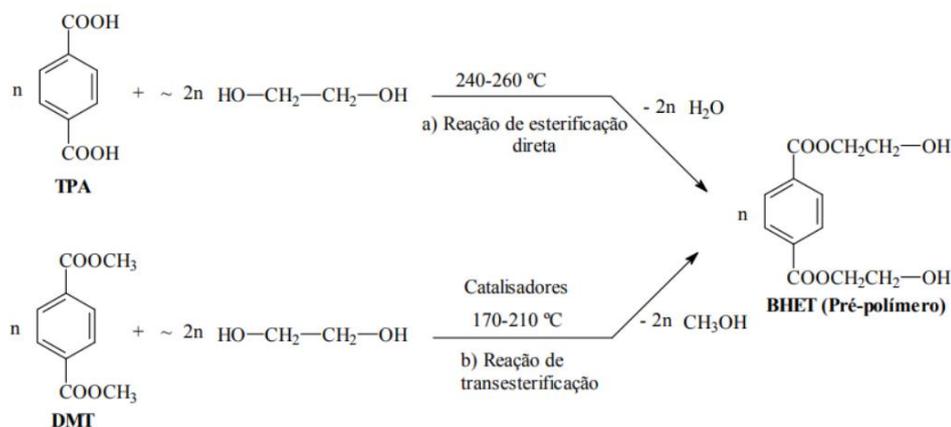
O polietileno tereftalato é um material polimérico que se origina principalmente, pela reação entre um ácido dicarboxílico, o ácido tereftálico (C₈H₆O₄), e um glicol, o etilenoglicol (C₂H₆O₂). Para a formação dos chips ou pellets, como é conhecido o produto da resina PET, é necessário passar por uma série de etapas, como, pré-polimerização, poli condensação e a poli condensação no estado sólido. Durante a etapa da pré-polimerização, duas rotas distintas podem ser tomadas para a produção do monômero do BHET [tereftalato de bis(2-hidroxi-etileno)] (PREVILATE, 2022).

a) Esterificação direta: Esta reação é obtida através do ácido tereftálico (TPA) com o etileno glicol (EG). A produção do TPA é dada a partir da oxidação do p-xileno junto ao ácido acético. Esta rota, é uma das mais atrativas no mercado, sendo atualmente utilizada pelas maiores empresas fabricantes de garrafas PET (PREVILATE, 2022).

b) Transesterificação: É um tipo de reação antiga e que por um tempo foi a técnica preferida pelas empresas pela possibilidade de obtenção do monômero DMT (com maior grau de pureza). Neste processo, o reator é carregado com DMT:EG e com catalisadores. Por meio de um receptor, todo metanol liberado é coletado. A reação é considerada finalizada quando se acaba a destilação do metanol e o BHET é obtido (PREVILATE, 2022).

Os dois métodos de obtenção do pré polímero do PET, podem ser observados através das reações ilustradas na Figura 2.

Figura 2 - Reações de Pré-Polimerização do PET.

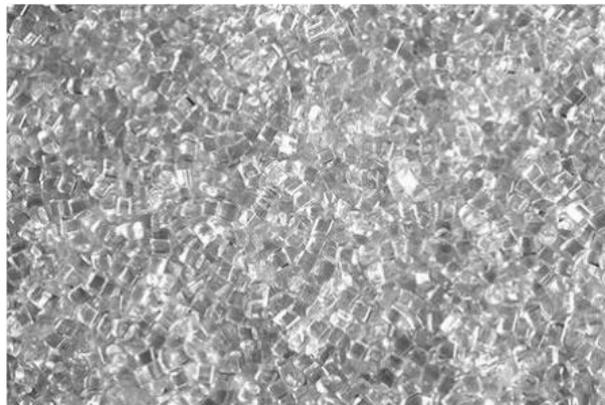


Fonte: ROMÃO et al (2009).

Na esterificação direta, a reação acontece entre 240 e 260 °C e por ser uma reação heterogênea e autocatalítica, para que ela aconteça, não é necessária a adição de catalisadores. Como produto da reação de pré-polimerização, tem-se além do pré polímero de PET, a água ou o metanol e os monômeros excessivos. Todas essas substâncias são removidas da atmosfera reacional através de vácuo, que deve ser aplicado ao processo. A massa molecular do pré polímero de PET, obtido nesta etapa é, no entanto, bastante baixa e deve ser aumentada através de uma reação de policondensação, conhecida como “melt phase condensation”, que acontece em temperaturas acima de 280°C (KEMKES, 1970).

Durante o aquecimento que acontece nesse processo, a pressão interna do reator é reduzida para valores abaixo de 130 Pa e, com isso, o grau de polimerização é elevado para 100. O grau de polimerização representa a quantidade média de monômeros existentes em uma molécula, e portanto, o tamanho médio da cadeia polimérica. Após esta etapa, são obtidos polímeros com massa molar em torno de 33000 g/mol, com isso, a massa resultante é fundida e cortada na forma de pellets de PET amorfo, representado na Figura 3 (KEMKES, 1970).

Figura 3 - Resina PET amorfa.



Fonte: O Autor (2023).

A massa molecular do material amorfo é ainda considerada bastante baixa para a aplicação do PET na forma de embalagens, por exemplo. Por isso, para materiais que necessitam de maiores massas molares (>30000 g/mol), outra reação de condensação adicional pode ser realizada, desta vez em estado sólido. Conhecida como “solid-state polycondensation”, essa reação de policondensação em estado sólido, acontece em temperaturas

entre 180 e 230 °C, em atmosfera inerte ou vácuo, sob agitação. Durante essa reação, que acontece em estado sólido, sem que haja fusão do PET, há eliminação de produtos voláteis e o aumento da massa molar da resina, resultando assim nas resinas cristalinas de PET, representada na Figura 4 (ROMÃO, 2009).

Figura 4 - Resina PET cristalizada.



Fonte: O Autor (2023).

2.3. PRINCIPAIS PARÂMETROS DA RESINA PET

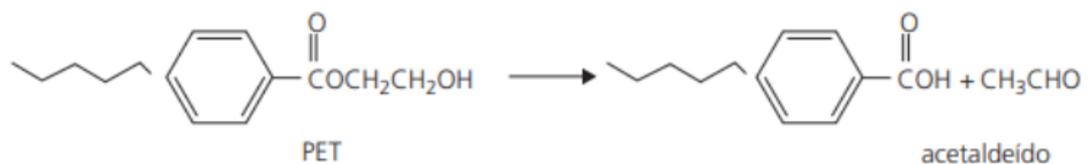
2.3.1. Viscosidade

O índice de fluidez representa um parâmetro técnico crucial no controle de qualidade da resina PET, sendo sua magnitude inversamente proporcional à viscosidade do polímero fundido e sua massa molar. Entretanto, é importante observar que esse polímero é suscetível à hidrólise, especialmente em temperaturas próximas às do processo de fabricação e do teste de índice de fluidez. Isso pode resultar em degradação durante o ensaio, introduzindo desafios experimentais e possíveis erros nas medições. Embora a viscosidade intrínseca possa ser convertida em massa molar viscosimétrica ou média numérica por meio de equações apropriadas, é notável que essa conversão é frequentemente realizada na prática industrial. Na maioria dos casos, a preferência da indústria é confiar diretamente nos valores de viscosidade intrínseca para avaliações de desempenho e controle de qualidade (MANCINI et al, 2004).

2.3.2. Acetaldeído

O acetaldeído pode ter sua formação através do processo termodegradativo natural do PET, podendo migrar da embalagem para alimentos e bebidas. Com isso, durante a síntese do polietileno tereftalato ou no seu processamento pode haver a degradação do material produzindo o composto volátil mais comum nesse processo, o acetaldeído. A reação referida é ilustrada na Figura 5. (TOLEDO et al., 2008).

Figura 5 - Reação do processo termodegradativo do PET em acetaldeído.



Fonte: TOLEDO (2008).

Por não apresentar riscos toxicológicos comprovados à saúde, as normas sanitárias regulamentadoras não estabelecem limites específicos de migração deste composto para produtos alimentícios. Entretanto, pelo fato da sua concentração apresentar a capacidade de alterar o sabor de alimentos e bebidas, as empresas produtoras de resinas estabelecem limites internos na ordem de partes por milhão (ppm), para que assim não haja variação das características sensoriais pelo contato do polímero e o alimento ou bebida (TOLEDO, et al., 2008).

2.3.3. Colorimetria

As propriedades ópticas de um material plástico estão relacionadas à sua interação com a luz. Essas interações podem ser categorizadas em três tipos distintos: a passagem da luz ou difusão através do polímero, a reflexão da luz na superfície do material e a manifestação da cor específica do plástico. De maneira geral, os cristais presentes nos polímeros têm um tamanho que propicia o espalhamento da luz, tornando os polímeros transparentes não cristalinos, ao

contrário dos translúcidos, que se encontram na faixa de comprimento de onda da luz visível (CHEREMISINOFF,1997).

Os três eixos responsáveis pela medição dos níveis de cores são compostos pelo eixo vertical, denominado de cor L^* - que varia do preto (0) ao branco (100), o eixo horizontal denominado de cor a^* - que varia do verde (parte negativa) ao vermelho (parte positiva) e por fim o eixo horizontal denominado de cor b^* - que varia do azul (parte negativa) ao amarelo (parte positiva) (BUENO, 2016).

2.3.4. Temperaturas de Transição

Os importantes três intervalos de temperatura a serem observados no PET, são: a temperatura de transição vítrea (T_g), que se trata do intervalo de temperatura onde as moléculas do polímero amorfo adquirem mobilidade por ganhar energia na forma de calor; a temperatura de fusão cristalina (T_m), que consiste da faixa de temperatura em que há o desaparecimento das regiões cristalinas devido à fusão dos cristalitos por aquecimento; e por fim, a temperatura de cristalização (T_c), que é o intervalo de temperatura onde ocorre a formação de cristais quando submete-se o polímero ao aquecimento ou resfriamento a partir de seu estado fundido (MC KEEN, 2014).

2.3.5. Umidade

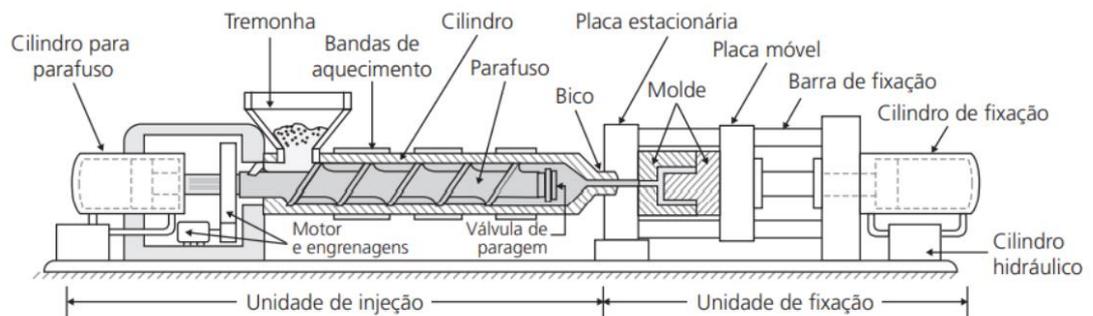
Devido à significativa higroscopicidade do polímero semi-cristalino PET, a determinação do teor de umidade desse material é conduzida por meio de um teste gravimétrico de grande importância. Nesse processo, uma alíquota da resina é pesada em uma termobalança, sendo que a quantidade de umidade absorvida até o momento depende das condições ambientais de temperatura e umidade durante o armazenamento do material. Posteriormente, essa alíquota é exposta a condições de temperatura superiores ao ponto de ebulição da água por um período específico. Após o aquecimento, a amostra é novamente pesada, permitindo quantificar a quantidade de água perdida e, conseqüentemente, determinar o teor de absorção de umidade do material sob as condições de armazenamento existentes (SILVA, 2023).

2.4. PROCESSAMENTO DE RESINA PET

O processamento da resina PET para a formação de embalagem, como garrafas e potes pode ser desenvolvida por diferentes métodos; porém é comumente obtida por meio da injeção e sopro. O processo de injeção pode ser seguido pelo processo de sopro, podendo ou não ser complementado pelo processo de estiramento. Na moldagem por injeção e sopro, o polímero fundido é injetado em uma matriz moldada com o formato desejado da embalagem. No caso específico da produção de garrafas, utiliza-se o modelo de pré-formas, onde a garrafa é moldada previamente antes do processo de sopro (WHELAN, 1989). A pré-forma é posteriormente encaminhada para a fase de sopro, na qual o formato final da garrafa é alcançado. No caso do processo de moldagem por injeção e sopro com estiramento, também conhecido pela sigla ISBM (Injection Stretch Blow Molding - Moldagem por Sopro Estiramento por Injeção), a pré-forma é inicialmente produzida, subsequente reauecida, es tirada e moldada na configuração final da embalagem (TOLEDO et al., 2008).

A Figura 6, representa os principais componentes da máquina de injeção para o processamento da resina PET para a formação das embalagens como garrafas.

Figura 6 - Máquina injetora usada para o processamento da resina PET.



Fonte: SILVA (2022).

2.5. MACHINE LEARNING

Machine Learning, traduzido como "aprendizado de máquina" em português, foi inicialmente introduzido por Arthur Lee Samuel, um pioneiro em jogos de computador e inteligência artificial. Ele descreve a ciência que capacita computadores a operar de forma

autônoma, sem a necessidade de programação explícita. De forma resumida, o aprendizado de máquina como um campo de estudo no qual programas de computador podem aprender e aprimorar o desempenho em tarefas específicas por meio do treinamento com dados históricos (SYAM; SHARMA, 2018).

As técnicas de ML são identificadas como uma subdivisão da IA, que são projetadas por meio de algoritmos que aprendem com dados, utilizando-os como modelos. Os princípios do ML são aplicados em diversas áreas da IA, sendo sua utilização observada em vários setores, contribuindo tanto para a resolução de problemas quanto para o aprimoramento de atividades. Para ser considerado inteligente, um sistema em um ambiente dinâmico deve ter a capacidade de aprender. Se o sistema pode se adaptar a mudanças, o designer do sistema não precisa antecipar e fornecer soluções para todas as situações possíveis (SCHÜSSLER; BASTIANI; BUSSLER, 2018).

Existe uma diversidade de métodos e ferramentas de aprendizado de máquina disponíveis para estudo. O primeiro passo crucial ao desenvolver qualquer novo aplicativo é a seleção do método mais apropriado, uma vez que essa escolha está intrinsecamente ligada ao tipo de dados, à quantidade de pontos de dados disponíveis para treinamento do sistema e à saída desejada. Portanto, tomar decisões equivocadas nesse processo pode resultar na criação de falsas correlações durante o treinamento e na construção de modelos preditivos ineficazes (BOMFIM, 2021).

Os métodos de aprendizado de máquina podem ser classificados de acordo com o tipo de aprendizado. Assim, esses métodos podem ser agrupados em quatro categorias distintas: supervisionado, não supervisionado, semisupervisionado e por reforço (BAHVSAR et al, 2017).

Com isso, no aprendizado supervisionado, é analisado que o ganho estatístico ocorre através da supervisão, que envolve a construção de um modelo estatístico para prever ou estimar uma saída com base em uma ou mais entradas. Em outras palavras, na aprendizagem supervisionada, para um conjunto de variáveis preditoras no conjunto de dados (input), o valor correto da variável dependente é fornecido ao algoritmo, ou seja, existe uma variável de saída correspondente a um conjunto de variáveis de entrada (SYAM; SHARMA, 2018).

As ferramentas de aprendizado de máquina supervisionado podem ser conceituadas como um processo que utiliza a experiência para adquirir especialização ou conhecimento. Nesse contexto, a aprendizagem supervisionada se desenvolve em um cenário em que a

experiência contém informações significativas ausentes nos exemplos de teste e invisíveis para os quais a experiência aprendida deve ser aplicada. Nesse cenário, a experiência adquirida é destinada a prever as informações ausentes nos dados separados para teste. Analogamente, pode-se entender este método como um ambiente em que um professor supervisiona o aluno, fornecendo informações extras ou necessárias (SHALEV-SHWARTZ; BEN-DAVID, 2014).

A partir disso, para abordar problemas de classificação, o algoritmo de aprendizado de máquina deve visar categorizar ou classificar as entradas fornecidas com base no conjunto de dados de treinamento. Os dados empregados podem se manifestar sob a forma de texto, números ou outros tipos de valores nominais. O conjunto de dados de treinamento em um problema de classificação é composto por uma série de pares de entrada e saída categorizados em classes, sendo que muitos desses problemas são binários, como exemplificado pela representação da resposta em verdadeiro e falso (BHAVSAR et al., 2017).

As métricas desempenham um papel crucial na avaliação do modelo. Nesse sentido, a matriz de confusão é empregada para fornecer uma visão mais abrangente e completa ao avaliar o desempenho de um modelo. Sua definição é ilustrada na Figura 7:

Figura 7 - Matriz de confusão.

		Classe Prevista	
		+	-
Classe Real	+	VP Verdadeiro Positivo	FN Falso Negativo
	-	FP Falso Positivo	VN Verdadeiro Negativo

Fonte: Adaptado de Amidi e Amidi (2018).

A métrica empregada na avaliação do desempenho das técnicas de classificação do aprendizado de máquina é denominada acurácia, representada pela Equação 1:

$$A = \frac{V_p + V_n}{V_p + F_p + V_n + F_n} \quad (1)$$

Para estabelecer a métrica de precisão, a fim de quantificar o número de exemplos classificados como pertencentes a uma classe, que de fato são dessa classe (VP), utiliza-se a Equação 2:

$$P = \frac{V_p}{V_p + F_p} \quad (2)$$

A especificidade avalia a capacidade do método de detectar resultados negativos, e é definida pela Equação 3:

$$E = \frac{V_n}{V_n + F_p} \quad (3)$$

A revocação avalia a habilidade do método de identificar com sucesso resultados classificados como positivos e é expressa pela Equação 4:

$$S = \frac{V_p}{V_p + F_n} \quad (4)$$

O score F1 é um indicador que representa a média harmônica entre precisão e revocação, pois ambos avaliam VP, mas com base em referências distintas, conforme expresso na Equação 5:

$$F1 = \frac{2V_p}{2V_p + F_p + F_n} \text{ OU } F1 = 2 * \frac{S * P}{S + P} \quad (5)$$

2.6. TÉCNICAS DE MACHINE LEARNING

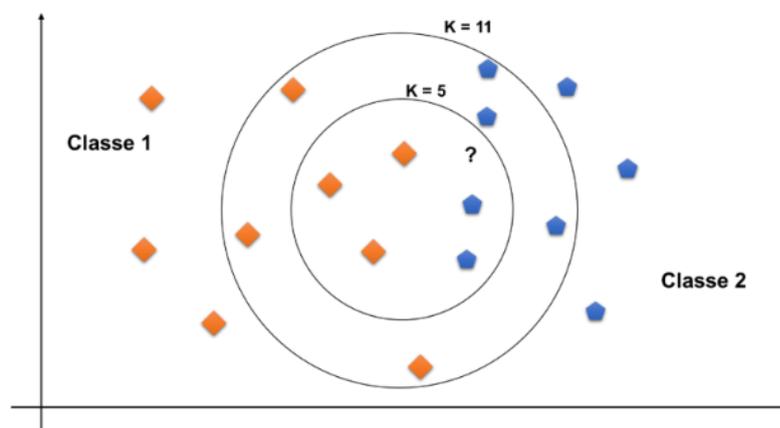
2.6.1. K-vizinhos mais próximos (KNN)

O algoritmo dos k-vizinhos mais próximos (KNN) do inglês k-Nearest Neighbors é uma abordagem simples para classificar um ponto de dados, determinando sua classe com base na maioria dos K-vizinhos mais próximos. Os dados utilizados nesse método abrangem diversos atributos multivariados, empregados para a classificação. O método de KNN é uma abordagem de ML antiga e intuitiva. Ele utiliza observações dos dados de treinamento para identificar os KNN em relação ao dado que se deseja prever, seguindo uma métrica de proximidade (LOPES 2018).

A estrutura do algoritmo visa buscar os registros classificados mais próximos e, com base nessa informação, classificar o novo registro. O número de vizinhos mais próximos a serem comparados é definido pelo valor "K". A aplicação do algoritmo KNN utiliza a distância euclidiana como medida, fundamentada no teorema de Pitágoras. A parametrização do número de vizinhos considerados na análise é representada por K, sendo geralmente muito menor que N, o tamanho da amostra (ALPAYDIN, 2010).

Para ilustração do método, tem-se a Figura 8, tendo duas classes, uma de pentágonos azuis e a outra de losangos laranjas, na qual a chegada de uma nova variável é simbolizada pela interrogação, e a mesma precisa ser classificada.

Figura 8 - Exemplo de classificação por meio do KNN.



Fonte: BOMFIM (2021).

No exemplo de classificação da Figura 8, com $k=5$, a nova instância é classificada como laranja, pois três vizinhos pertencem a essa classe. Essa classificação se mantém quando $k=11$, já que seis vizinhos são da classe laranja, enquanto cinco são azuis.

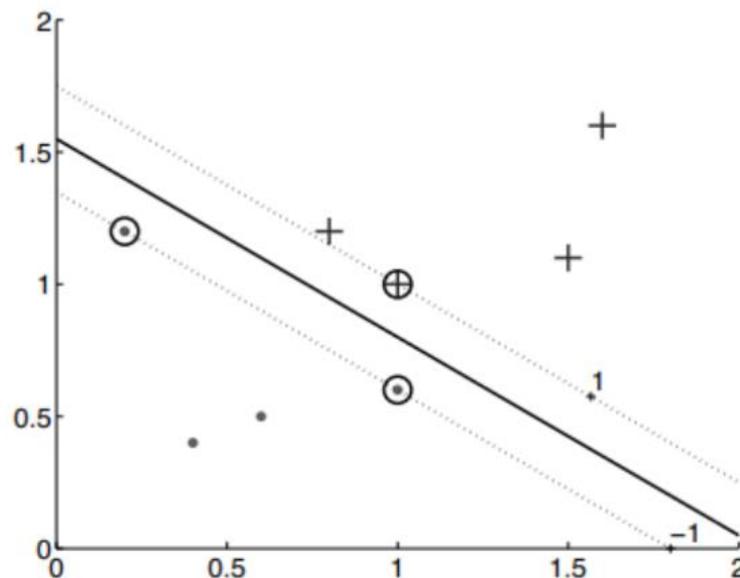
2.6.2. Máquinas de vetor de suporte

Máquinas de vetor de suporte (do inglês Support Vector Machine - SVM) são algoritmos de classificação que visam aproximar as margens de um dado a ser classificado com os dados mais próximos. O algoritmo apresenta grande capacidade de generalização e robustez, possibilitando sua aplicação em vetores de grandes dimensões (AMARAL, 2016).

SVM é um método de aprendizado supervisionado que classifica os pontos de dados, segregando-os usando um hiperplano N -dimensional, onde N é o número de atributos que caracterizam um ponto de dados. A ideia geral é encontrar uma função com uma margem de erro definida que mapeie as variáveis de entrada para a variável de saída de modo que a saída prevista não se desvie da saída real mais do que a margem de erro definida (BOMFIM, 2021).

Na Figura 9, temos um problema de duas classes, onde as instâncias de cada classe são representadas por sinais de mais e pontos. A linha espessa representa o limite de decisão, enquanto as linhas tracejadas definem as margens em ambos os lados.

Figura 9 – Exemplo de SVM para Classificação.



Fonte: Alpaydin (2010).

Por ser um método baseado em discriminante, o SVM se preocupa apenas com as instâncias próximas ao limite e descarta aquelas que estão no interior (ALPAYDIN, 2010).

2.6.3. Árvore de Decisão

A árvore de decisão é um método não paramétrico que apresenta uma estrutura semelhante à de uma árvore ou fluxograma, sendo aplicável tanto a problemas de classificação quanto de regressão. Muitos pesquisadores têm usado o método de previsão de árvore de decisão por sua clareza e facilidade em expor conjuntos de dados pequenos e grandes e prever o valor. A lógica na aplicação de técnicas de árvore de decisão equivale a uma série de declarações IF-THEN, que podem ajudar a simplificar o entendimento deste método (ZEINEDDINE et al. 2020).

No método de árvore de decisão, para um problema de classificação, o processo segue a estrutura de uma árvore onde um atributo mais informativo para classificar os dados é dividido em ramos hierárquicos de forma que a próxima pergunta a ser feita dependa da resposta da pergunta atual. Para uma árvore de decisão de classificação, o nível de impureza é medido para avaliar o desempenho da árvore. Caso a árvore de decisão classifique todos os padrões de dados em classes às quais eles realmente pertencem, as divisões entre a classe e as ramificações são consideradas puras (BHAVSAR et al., 2017).

3. METODOLOGIA

3.1. CARACTERIZAÇÃO DO PROCESSO

Qualquer manifestação por parte do cliente, isto é, do fabricante de garrafas responsável pela operação de injeção e sopro, em relação à empresa mencionada, necessita de análises detalhadas e demoradas acerca dos parâmetros técnicos do produto vendido e investigação da ocorrência para conseguir proceder ou não a reclamação do cliente. Com isso, serão listados e identificados os principais parâmetros físico-químicos que a resina pode gerar problema na produção das garrafas e potes por parte do cliente. Atualmente, a execução dessa atividade depende da experiência do especialista, no entanto, a aplicação de técnicas de aprendizado de máquina pode facilitar a assimilação do conhecimento especializado e contribuir para a definição de dados em futuros lotes de produção por parte de outros colaboradores.

3.2. BASE DE DADOS

A base de dados obtida para o estudo é construída por meio de planilhas com os principais parâmetros técnicos da resina PET, identificando lotes que podem produzir algum tipo de notificação por parte do cliente e lotes sem notificação para resinas específicas para o processamento de garrafas voltadas ao envase de bebidas carbonatadas e resinas específicas para o processamento de garrafas voltadas ao envase de água mineral.

O banco de dados é composto por 31612 registros de lotes produzidos das resinas para as diferentes aplicações, em que todos os dados já estão tratados e não foi necessário realizar a retirada de campos vazios, ou informações duplicadas ou incorretas.

A planilha contendo a base de dados é constituída por sete colunas, cada uma correspondendo uma característica diferente, como exposto abaixo:

- **Tipo de resina:** Corresponde ao tipo de resina utilizada para aplicação desejada.
- **Linha:** Corresponde ao tipo de linha de produção utilizada para fabricar o lote dos chips.
- **IV:** Corresponde à viscosidade intrínseca do produto em dL/g.
- **CL:** Corresponde a Cor L*.

- **CB:** Corresponde a Cor b*.
- **AD:** Corresponde a concentração de acetaldeído em partes por milhão (ppm).
- **Notificação:** Corresponde à possibilidade do lote ter ou não notificação por parte do cliente.

A Tabela 1 abaixo, mostra a relação de categorias por atributos de acordo com a base de dados trabalhada.

Tabela 1 – Quantidade de atributo por categoria.

Categorias	Atributos
Tipo	Água mineral, Carbonatada
Linha	B, C
IV	0,79 - 0,88 dL/g.
CL	76,30-85,40
CB	-4,40 - 0,11
AD	0,36 - 1,20 ppm
Notificação	(0) Não - (1) Sim

Fonte: O Autor (2024).

3.3. APRENDIZADO DE MÁQUINA

A construção da ferramenta de aprendizado de máquina foi realizada por meio da plataforma computacional que emprega a linguagem de programação Python, como a plataforma do Google Colab em suas versões de código aberto. Para o desenvolvimento de algoritmos e a aplicação das técnicas de aprendizado de máquina, a principal biblioteca utilizada foi Scikit-learn, dada sua ampla aceitação na comunidade de Ciência de Dados, facilidade de aprendizado, acessibilidade abrangente, simplicidade e eficiência. Diferentes técnicas foram comparadas e sua eficiência avaliada a partir dos resultados de classificação dos eventos conhecidos da base de dados.

A plataforma citada é capaz de ler várias extensões de arquivos de dados, como “json”, “csv”, “txt”, entre outras, e realizar integrações com bancos de dados relacionais e não relacionais. No entanto, o formato do arquivo utilizado está em “xlsx”, uma vez que a empresa

possui um banco de dados capaz de exportar os dados somente para o Excel para a manipulação dos dados com maior facilidade.

Com isso, as técnicas de ML aplicadas foram KNN, SVM e árvore de decisão analisando como performance dos modelos as métricas acurácias, precisão e revocação.

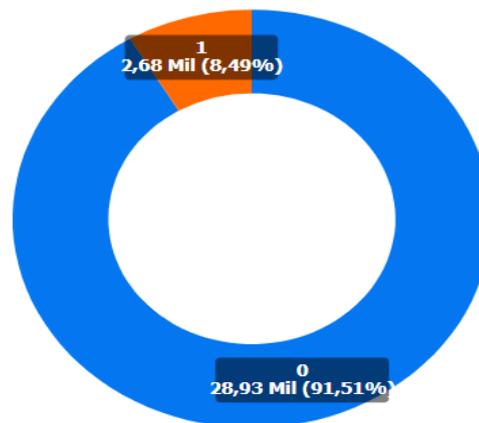
4. RESULTADOS E DISCUSSÕES

4.1. ANÁLISE DETALHADA DA BASE DE DADOS

Inicialmente, conduziu-se uma análise exploratória dos dados com o objetivo de interpretar e compreender seus padrões e tendências, visando extrair informações pertinentes para a elaboração dos modelos de implementação. Nesse processo, empregaram-se principalmente técnicas visuais, tais como mapas de calor, gráficos de pizza, de barras e de rosca.

De acordo com a base de dados, tem-se no total 31.612 registros para a realização das análises detalhadas com os seus respectivos parâmetros. Primeiramente foi realizada a análise quantos registros podem gerar ou gerou notificação por parte do produtor de pré-formas e garrafas. A Figura 10 mostra as quantidades dos quais não podem gerar notificação, representada pelo 0, e dos quais podem gerar notificação, representado pelo 1.

Figura 10 – Distribuição dos Registros de Notificações.

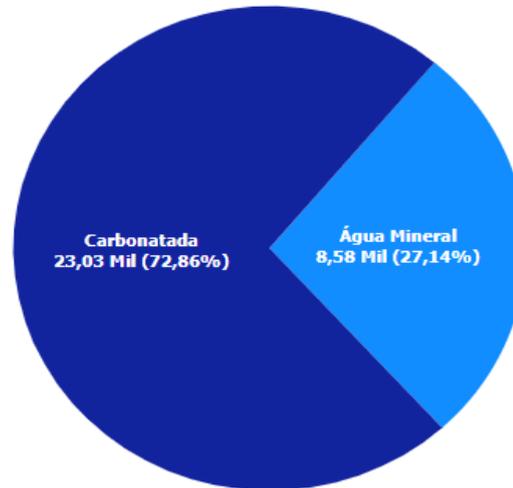


Fonte: O Autor (2024).

De acordo com a Figura 10, pode se perceber que a maioria dos dados não correspondem a registros que podem gerar notificações, sendo representado por 28900 mil dados, cerca de 91,51%, já os registros que podem gerar algum tipo de notificação correspondem a cerca de 2680 mil registros, representando cerca de 8,49%. Seguindo para a próxima análise, foi elaborado um gráfico de pizza para representar os registros por aplicação de resina, dividida em

resina voltada para o processamento de garrafas no envase de bebidas carbonatadas e resina voltada para o processamento de garrafas no envase de água mineral, ilustrado na Figura 11.

Figura 11 - Distribuição dos dados por aplicação de resina.

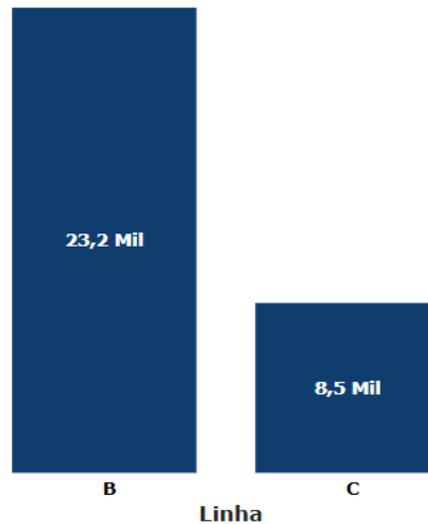


Fonte: O Autor (2024).

Conforme representado na Figura 11, a maioria dos dados corresponde a resina voltada para aplicação de garrafas carbonatadas, com cerca de 23030 mil registros e 77,86%, já a resina voltada para aplicação de garrafas de água mineral equivale a cerca de 9580 mil registros, representando 27,14% dos dados.

Outra análise realizada, foi identificar os registros separados por linhas de produção diferentes, separadas em B e C, representado na Figura 12.

Figura 12 - Distribuição dos dados por Linhas de Produções.



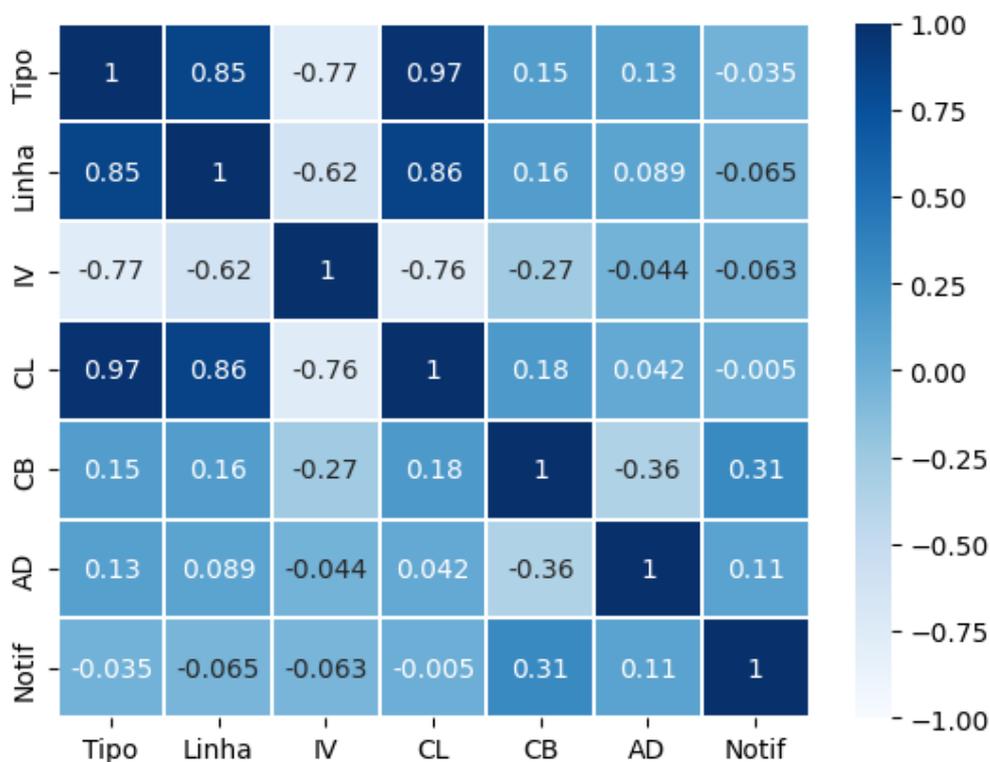
Fonte: O Autor (2024).

Segundo a Figura 12, pode-se perceber que a maioria das resinas produzidas são na Linha B, representando cerca de 23,2 mil registros, totalizando cerca de 73,4% e a Linha C, representa 8,5 mil dos registros, totalizando cerca de 26,6% dos registros.

Com base nas Figuras 10,11 e 12, é evidente que os parâmetros que mais influenciaram nas análises de correlação e predição subsequentes são aqueles associados à ausência de notificações (0), à resina utilizada para o envase de bebidas carbonatadas e aos registros da linha de produção da Linha B.

Para investigar as relações entre as variáveis, foi realizada a computação de uma matriz de correlação dos dados amostrais, seguida pela geração de um mapa de calor dividido em clusters, com o objetivo de fornecer uma representação visual. Cada célula é apresentada em uma cor distinta, proporcional à sua posição em um gradiente de cores. Importante mencionar que esta análise gerou exclusivamente o mapa de calor para variáveis do tipo inteiro e float. A ordenação das linhas foi determinada por meio de análises de clusters hierárquicos, conforme ilustrado na Figura 13.

Figura 13 - Mapa de calor das variáveis.

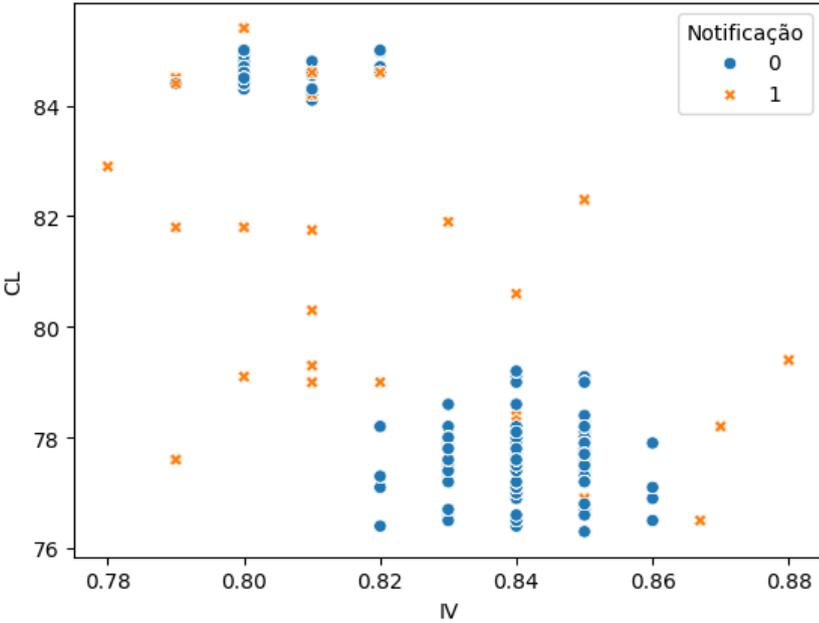


Fonte: O Autor (2024).

Observando o mapa de calor das variáveis na Figura 13, foi possível identificar que há uma fraca correlação das variáveis entre si, com destaque para o parâmetro de IV que tem a pior correlação entre as variáveis e à do tipo de resina que tem a melhor correlação entre as variáveis se comparado com as outras correlações. Isso porque a IV têm pouca variabilidade nos dados, variando entre 0,78 e 0,86 tornando difícil a correlação com as outras variáveis, e a do Tipo de resina, devido aos seus limites de especificação e os resultados das variáveis em faixa diferentes torna a correlação com os parâmetros CL, CB e AD melhores.

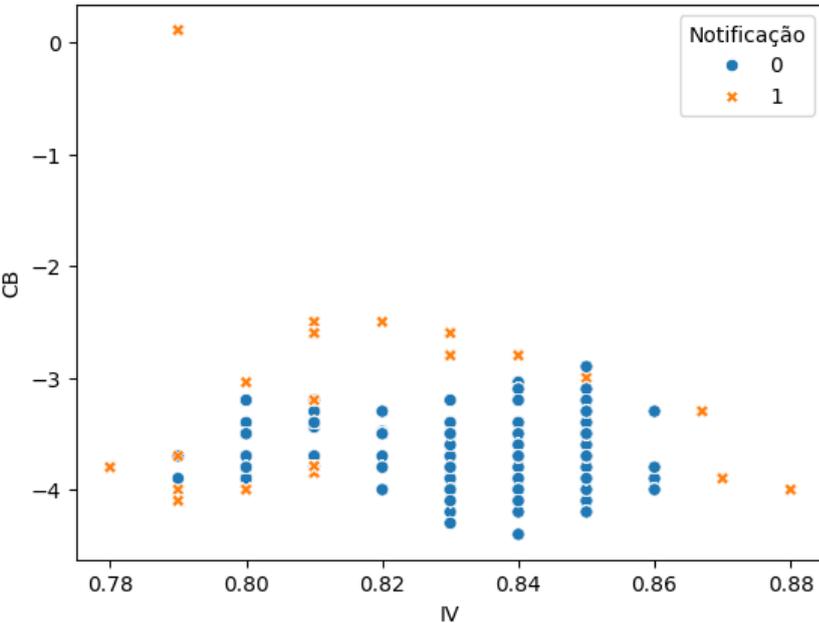
A partir desse mapeamento inicial, é importante ter também o conhecimento das correlações entre as variáveis físico-químicas, IV, CL, CB e do AD. Com isso, foram criados gráficos de dispersão entre duas variáveis a fim de identificar e analisar as correlações, com base na variável de notificação, tipo de resina e por linha de produção. Primeiro, foram representados os gráficos das correlações dos parâmetros com base nas notificações, ilustrados nas Figuras 14 a 19.

Figura 14 – Gráfico da correlação entre a IV e CL com base na Notificação.



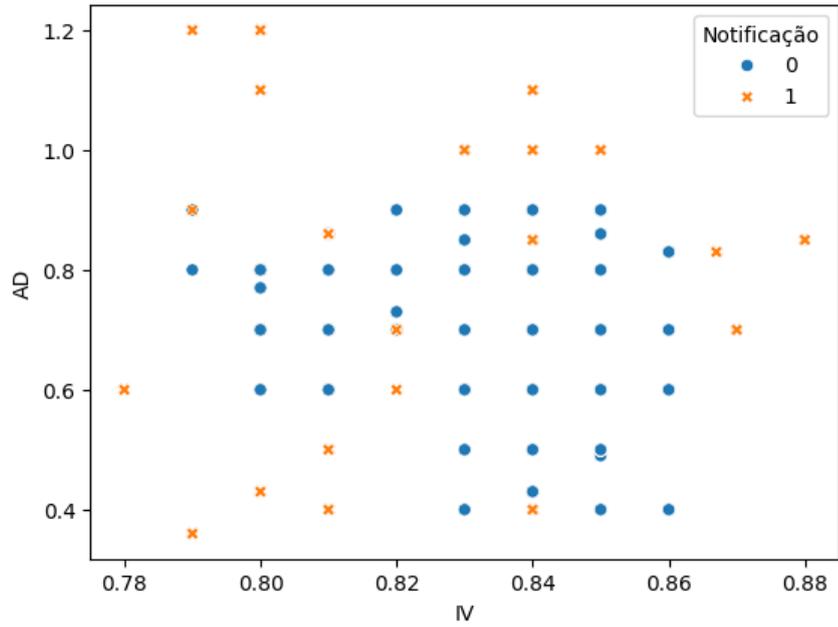
Fonte: O Autor (2024).

Figura 15 – Gráfico da correlação entre a IV e CB com base na Notificação.



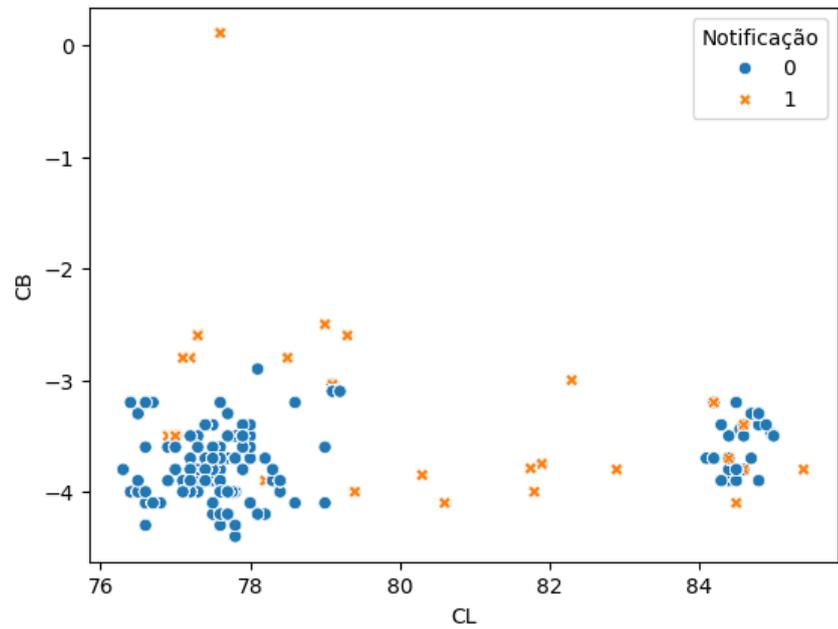
Fonte: O Autor (2024).

Figura 16 – Gráfico da correlação entre a IV e AD com base na Notificação.



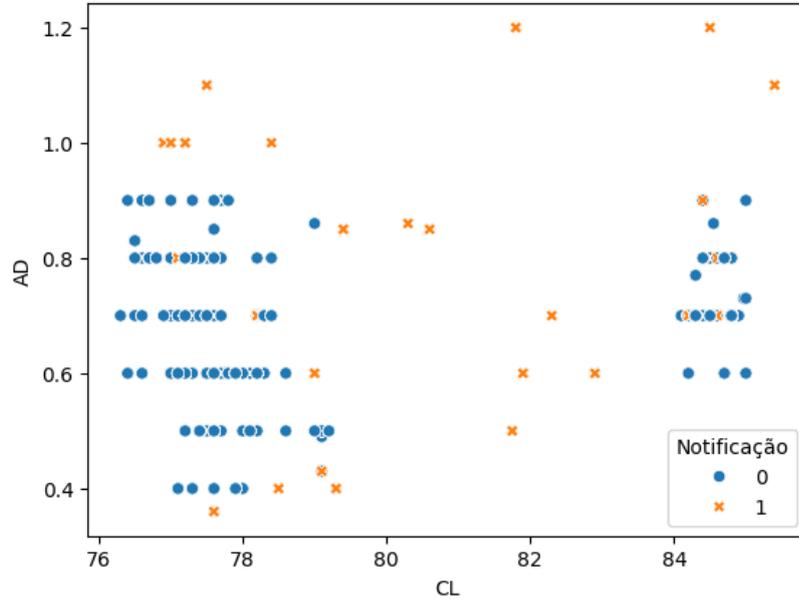
Fonte: O Autor (2024).

Figura 17 – Gráfico da correlação entre a CL e CB com base na Notificação.



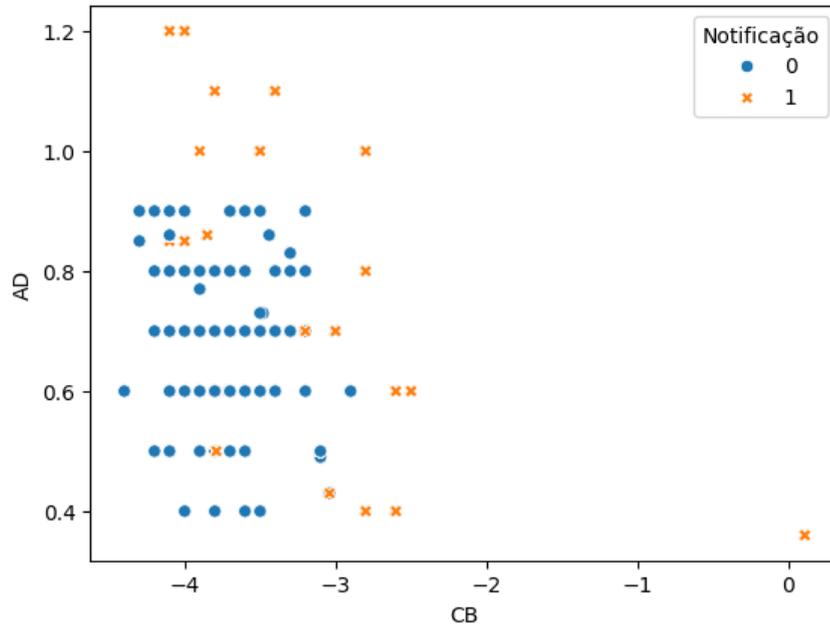
Fonte: O Autor (2024).

Figura 18 – Gráfico da correlação entre a CL e AD com base na Notificação.



Fonte: O Autor (2024).

Figura 19 – Gráfico da correlação entre a CB e AD com base na Notificação.



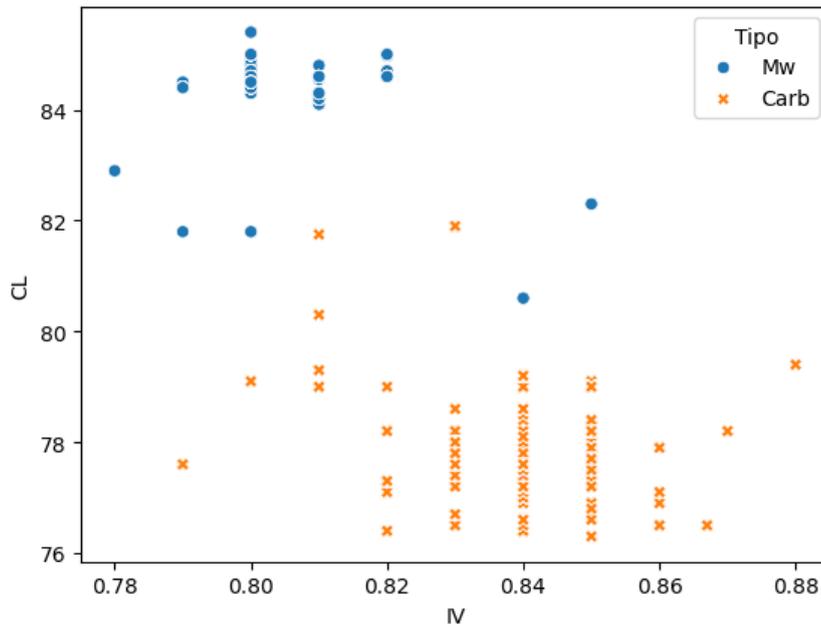
Fonte: O Autor (2024).

Conforme apresentado nas Figuras acima, é perceptível que as correlações entre a variáveis são fracas e não houve boa separação entre os grupos da variável de Notificação,

exceto na dispersão entre IV e AD, mostrado na Figura 14, em que nota-se uma separação entre uma área de não notificação, localizados de forma central, e possíveis notificações, localizadas ao redor.

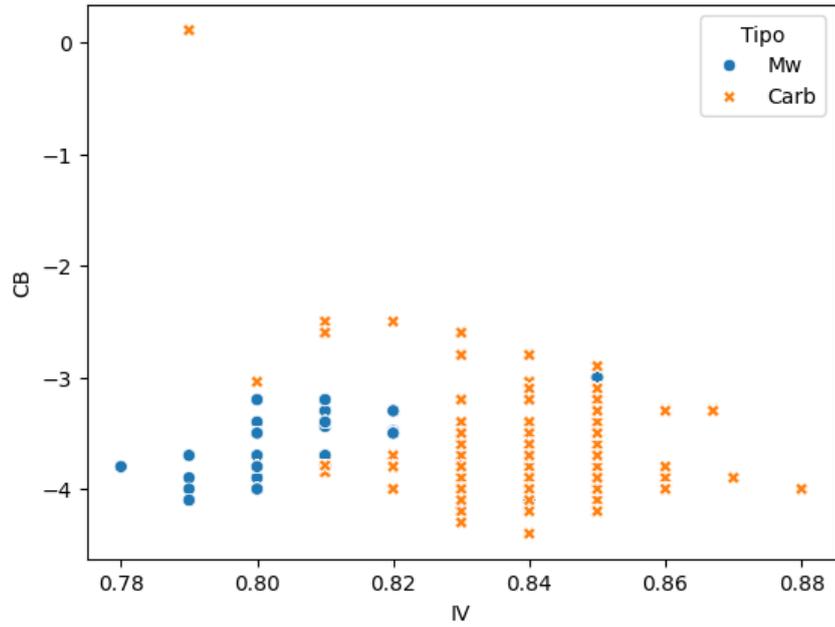
Com as correlações acima escritas, outro ponto que foi verificado foi a correlação entre as variáveis físico-químicas com base no tipo de aplicação de resina, expostos pelas Figuras 20 a 25.

Figura 20 – Gráfico da correlação entre a IV e CL com base no Tipo da resina.



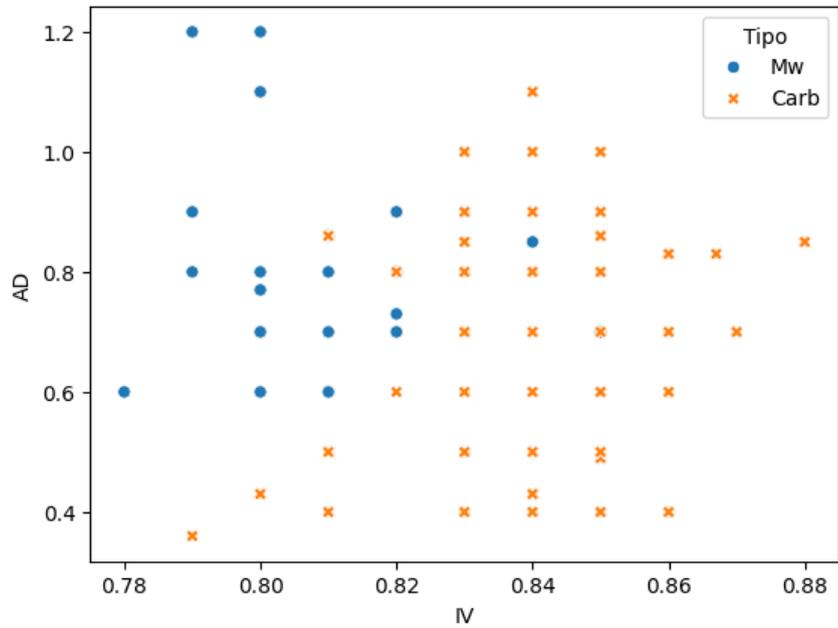
Fonte: O Autor (2024).

Figura 21 – Gráfico da correlação entre a IV e CB com base no Tipo da resina.



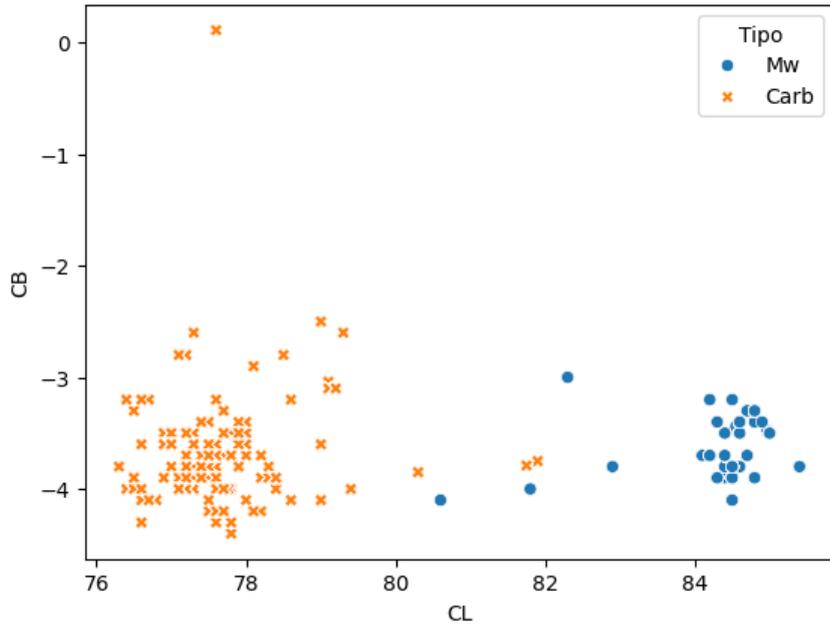
Fonte: O Autor (2024).

Figura 22 – Gráfico da correlação entre a IV e AD com base no Tipo da resina.



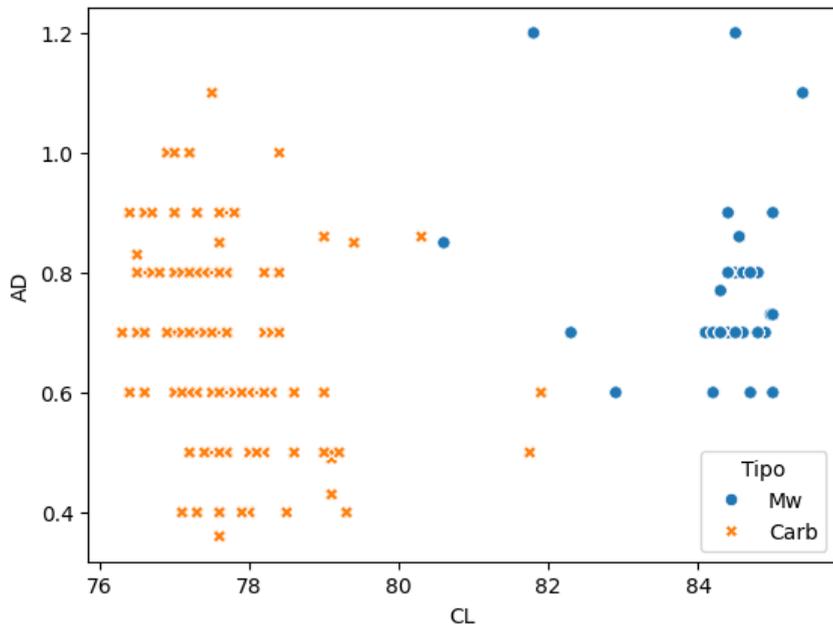
Fonte: O Autor (2024).

Figura 23 – Gráfico da correlação entre a CL e CB com base no Tipo da resina.



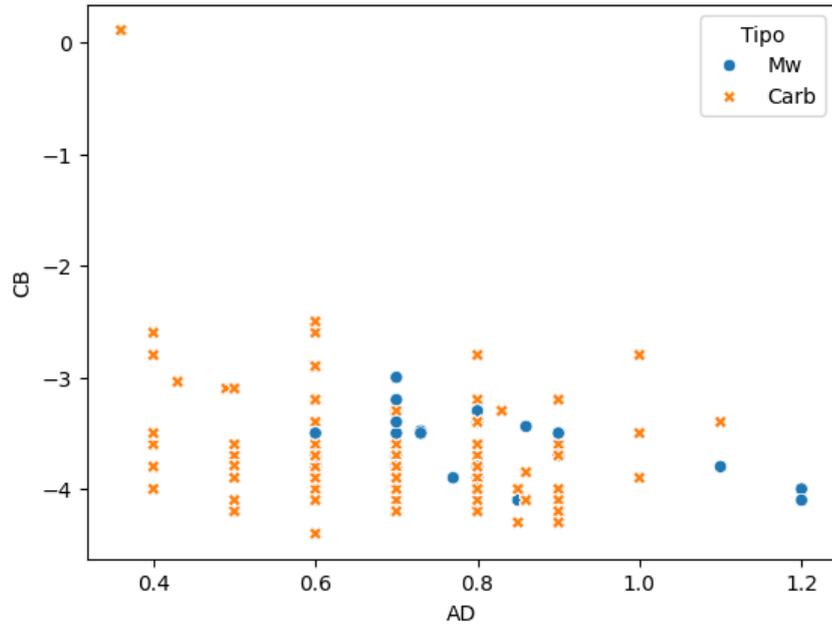
Fonte: O Autor (2024).

Figura 24 – Gráfico da correlação entre a CL e AD com base no Tipo da resina.



Fonte: O Autor (2024).

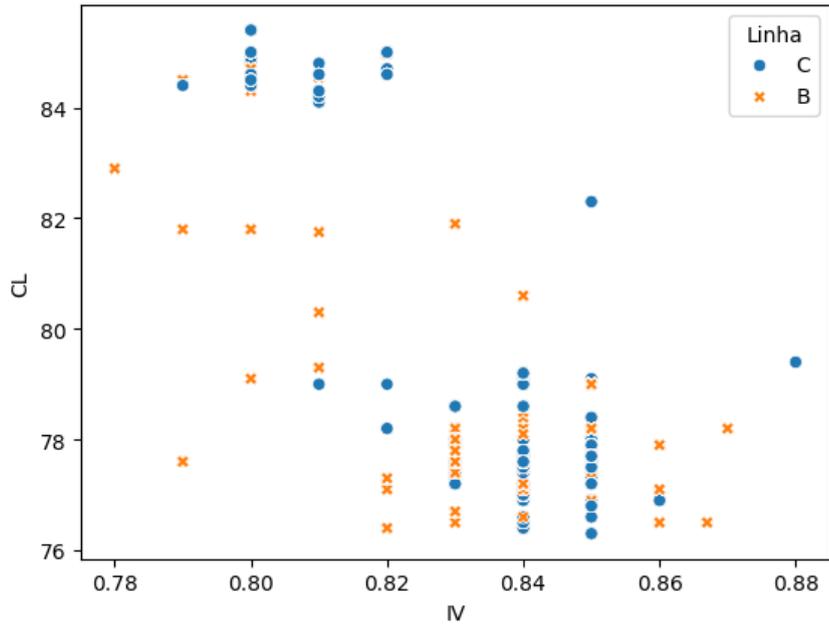
Figura 25 – Gráfico da correlação entre a CB e AD com base no Tipo da resina.



Com base nas Figuras de correlações acima, em relação ao tipo de aplicação da resina, é evidente que tem-se uma correlação bem mais definida do que em relação ao de notificações, isso porque, os limites dessas variáveis físico-químicas são bem definidos e diferentes para o tipo de aplicação. Nota-se que a variável de IV e CL contribui de forma significativa para as correlações entre as variáveis, ficando entre 76 a 80 para aplicação de garrafas de água mineral e 81 a 85 para aplicação de garrafa de bebidas carbonatadas para CL, e o IV, ficando 0,78 a 0,82 para aplicação de garrafas de água mineral e 0,82 a 0,88 para aplicação de garrafa de bebidas carbonatadas.

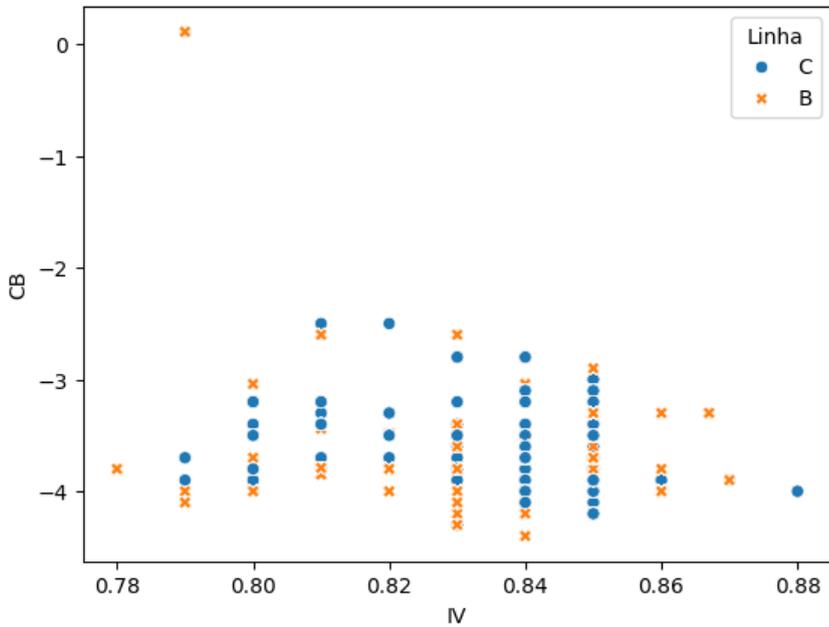
Além disso, foram analisadas também as correlações das variáveis com base na Linha de produção, ilustrado pelas Figuras 26 a 31.

Figura 26 – Gráfico da correlação entre a IV e CL com base na Linha de produção.



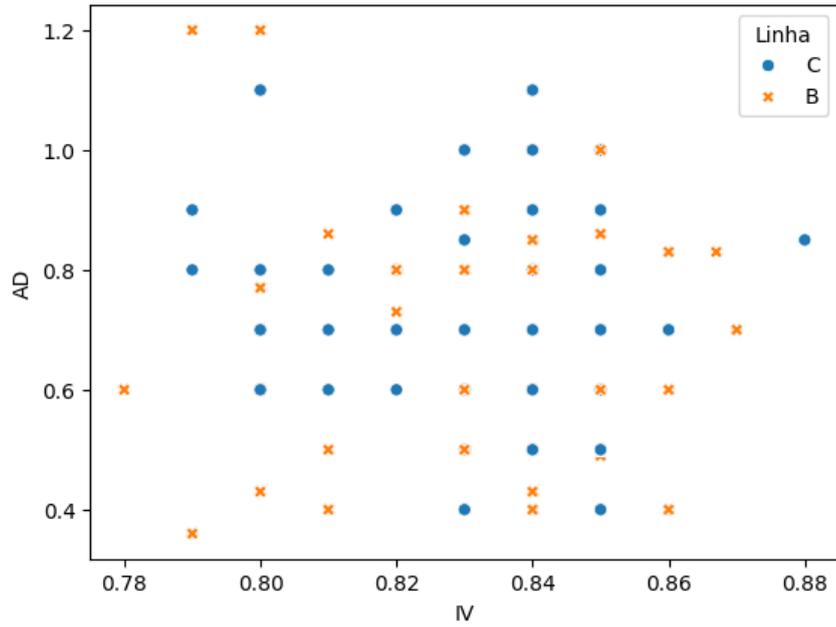
Fonte: O Autor (2024).

Figura 27 – Gráfico da correlação entre a IV e CB com base na Linha de produção.



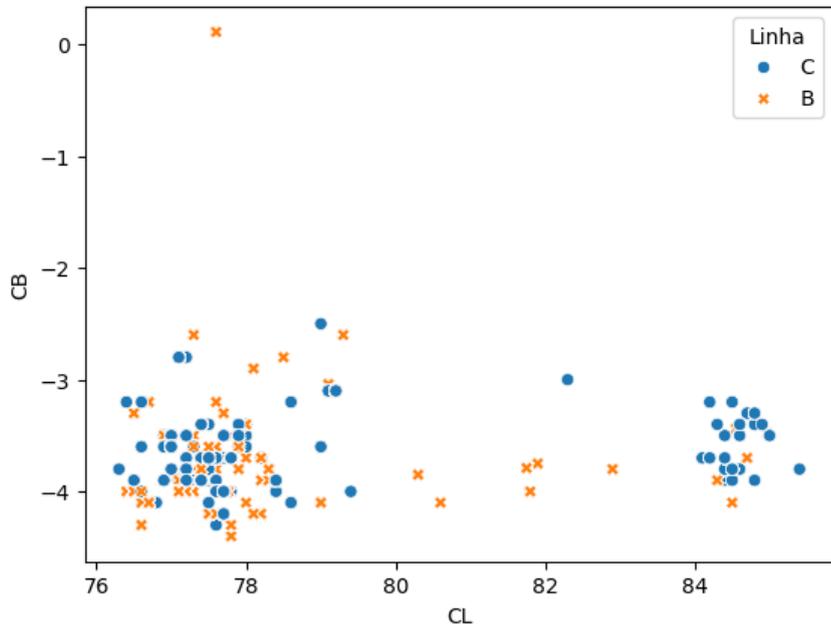
Fonte: O Autor (2024).

Figura 28 – Gráfico da correlação entre a IV e AD com base na Linha de produção.



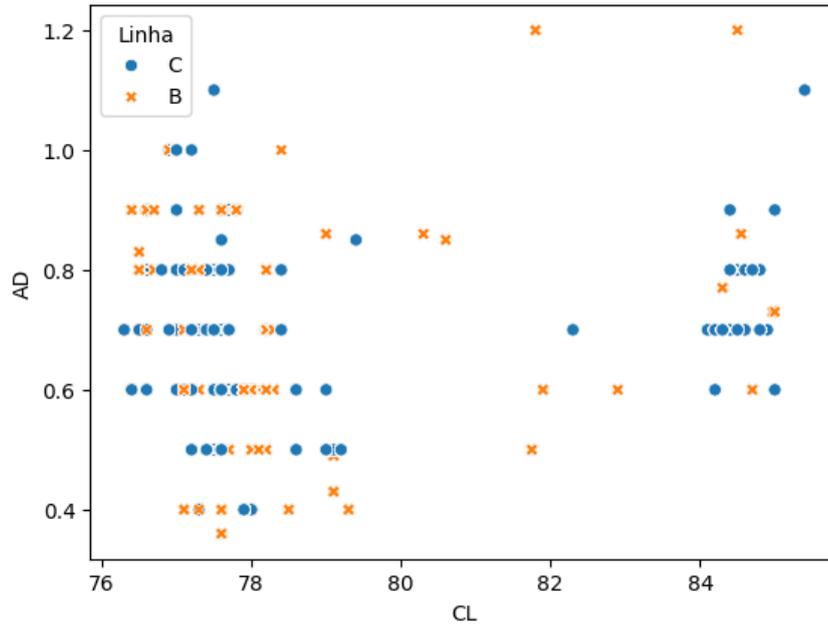
Fonte: O Autor (2024).

Figura 29 – Gráfico da correlação entre a CL e CB com base na Linha de produção.



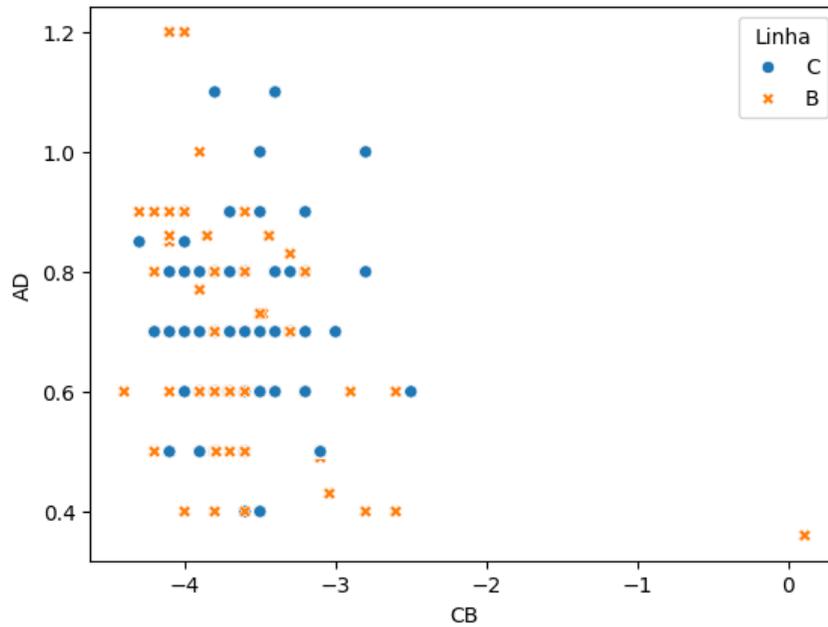
Fonte: O Autor (2024).

Figura 30 – Gráfico da correlação entre a CL e AD com base na Linha de produção.



Fonte: O Autor (2024).

Figura 31 – Gráfico da correlação entre a CB e AD com base na Linha de produção.



Fonte: O Autor (2024).

Com as Figuras acima representadas, é visível como as correlações são fracas entre si de modo que os pontos referentes aos dados se misturam mutuamente nos gráficos.

4.2. ANÁLISE DOS RESULTADOS DOS MODELOS PREDITIVOS

Depois de implementadas as técnicas de aprendizado de máquina, foram analisados os resultados tanto da acurácia, quanto da precisão e revocação utilizando as métricas descritas pelas equações 1, 2 e 4 para técnicas de classificação. As avaliações seguiram os modelos de análise e suas respectivas configurações, conforme mencionado na seção 2.5 e 2.6.

Foram iniciadas as análises pelo método KNN, avaliando a acurácia, precisão e revocação com base em 3 atributos, IV, IV e AD e IV, CL, CB e AD para predição das possíveis notificações ou não. As Tabelas 2, 3 e 4 mostram os resultados dos parâmetros a serem avaliados.

Tabela 2 – Resultados do modelo KNN com base na IV e a Notificação.

	Acurácia	Precisão	Revocação
0	0,3745	0,99	0,32
1		0,12	0,96

Fonte: O Autor (2024).

Tabela 3 – Resultados do modelo KNN com base na IV e AD e a Notificação.

	Acurácia	Precisão	Revocação
0	0,9886	0,99	1
1		1	0,87

Fonte: O Autor (2024).

Tabela 4 – Resultados do modelo KNN com base na IV, CL, CB e AD e a Notificação.

	Acurácia	Precisão	Revocação
0	0,9978	0,99	1
1		1	0,87

Fonte: O Autor (2024).

A partir dos dados apresentados nas Tabelas acima, pode-se perceber que os resultados do modelo KNN que teve o melhor resultado no geral foi com base em todas as variáveis para a predição das notificações ou não. Além disso, destaca-se o resultado da acurácia, ou seja, a

frequência com que o modelo de aprendizado de máquina prevê corretamente os resultados, com o valor de 0,9978, superando os outros testes com os resultados de 0,9886 e o 0,3745. As métricas da precisão e revocação também apresentaram bons resultados no teste, demonstrando que o modelo pode ser usado para prever as notificações.

Outro ponto de destaque, é os resultados do teste com base na variável de IV para a predição da notificação, tendo uma frequência de acerto corretamente bem baixo, com cerca de 0,3745, além da precisão e revocação, demonstrando assim que o modelo KNN não consegue prever bem a relação com a IV e a notificação. Os resultados do teste com base nas variáveis de IV e AD para predição da notificação mostram-se resultados similares comparado com os de todas as variáveis, demonstrando também que o modelo consegue prever bem um possível problema.

Outra análise feita, foi com modelo SVM avaliando também a acurácia, precisão e revocação, usando as mesmas variáveis e o parâmetro de previsão, ilustradas nas Tabelas 5, 6 e 7.

Tabela 5 – Resultados do modelo SVM com base na IV e a Notificação.

	Acurácia	Precisão	Revocação
0	0,9151	0,92	1
1		0	0

Fonte: O Autor (2024).

Tabela 6 – Resultados do modelo SVM com base na IV e AD e a Notificação.

	Acurácia	Precisão	Revocação
0	0,9155	0,92	1
1		0	0

Fonte: O Autor (2024).

Tabela 7 – Resultados do modelo SVM com base na IV, CL, CB e AD e a Notificação.

	Acurácia	Precisão	Revocação
0	0,922	0,92	1
1		1	0,07

Fonte: O Autor (2024).

Com base nos dados obtidos pelas Tabelas, é notável que os resultados apresentaram uma boa acurácia em todos os testes, no entanto a performance da precisão e do revocação relacionado aos testes com a predição das possíveis notificações (1) foram extremamente ruins, demonstrando que o modelo não consegue prever os dados das variáveis com as possíveis notificações para o teste com base na IV e IV e AD. No entanto, a performance da precisão e do revocação em geral para as não notificações, foram extremamente satisfatório demonstrando que o modelo consegue prever com boa porcentagem.

A última análise foi feita com base no modelo de árvore de decisão, com as mesmas variáveis e parâmetros dos modelos anteriores, representadas nas Tabelas 8, 9 e 10.

Tabela 8 – Resultados do modelo árvore de decisão com base na IV e a Notificação.

	Acurácia	Precisão	Revocação
0	0,93	0,93	1
1		1	0,18

Fonte: O Autor (2024).

Tabela 9 – Resultados do modelo árvore de decisão com base na IV e AD e a Notificação.

	Acurácia	Precisão	Revocação
0	0,98987	0,99	1
1		0,97	0,91

Fonte: O Autor (2024).

Tabela 10 – Resultados do modelo árvore de decisão com base na IV, CL, CB e AD e a Notificação.

	Acurácia	Precisão	Revocação
0	0,9981	1	1
1		1	0,98

Fonte: O Autor (2024).

Conforme mostrado nas Tabelas acima, a acurácia apresentou uma performance bastante satisfatória nos testes em geral, destacando o teste com todas as variáveis com resultado de 0,9981, demonstrando assim que o modelo consegue prever com frequência corretamente as possíveis notificações ou não. Outro ponto é que a precisão e a revocação

apresentaram também performance em todos os testes, conseguindo prever bem os dados que podem ter notificações e os dados que não tiveram notificações. Comparando os resultados do modelo de árvore de decisão com os modelos KNN e SVM, pode se afirmar que esse modelo consegue prever e ter a melhor performance para mostrar a possibilidade de ter problemas no processo de produção de injeção e sopro.

5. CONCLUSÃO

Diante do trabalho exposto, a ferramenta de aprendizado de máquinas proposta buscou prever e definir a possibilidade de haver notificações ou não através da base de dados contendo as informações das variáveis físico-químicas, como, IV, CL, CB e AD a partir dos limites de especificação. Para isso, foram aplicadas técnicas de classificação, obtendo os índices da acurácia, precisão e revocação para avaliação e análise dos modelos aplicados aos testes.

Entre os modelos KNN, SVM e árvore de decisão apresentados, o modelo de árvore de decisão atingiu os melhores índices de acurácia, precisão e revocação em todos os testes com as diferentes variáveis, além disso, apesar de um teste do modelo KNN ter apresentado um valor de acurácia muito baixo, a performance do modelo SVM foi extremamente ruim para prever os valores das possíveis notificações, apresentando assim a pior técnica com a base de dados utilizada. Com isso, a expectativa com os resultados satisfatórios foi superada devido à complexidade da problemática em estudo e a baixa correlação dos dados obtidos, ressaltando assim a relevância do trabalho, uma vez que as previsões possam vir a auxiliar na identificação das possíveis notificações dos produtores de injeção e sopro.

A aplicação de técnicas de aprendizado de máquina pode oferecer previsões precisas para desafios encontrados na indústria, e presume-se que essas técnicas tenham potencial para serem aplicadas em diferentes fases do processo produtivo de qualquer empresa. Além disso, uma das principais contribuições deste estudo é a viabilização da aplicação de Machine Learning e Inteligência Artificial na indústria, demonstrando como isso impulsiona a adoção da Indústria 4.0 através da automatização de tarefas e aprendizado de processos repetitivos ou carentes de procedimentos claros para sua execução.

Para pesquisas futuras, sugere-se que sejam utilizados outras técnicas e outros ajustes de parametrização nos algoritmos na construção da ferramenta de aprendizado de máquina para a predição das possíveis notificações dos produtores de garrafas devido a problemas no processo de injeção e sopro. Outra sugestão é a utilização de parâmetros de processo que tenha correlação direta com os possíveis impactos no processamento da resina PET para a predição das notificações.

6. REFERÊNCIAS

- ALPAYDIN, E. **Introduction to machine learning**. 2a. ed. Londres: [s.n.], 2010.
- AMARAL, F. **Aprenda mineração de dados: Teoria e Prática**. São Paulo: Alta Books, 2016.
- AMIDI, A.; AMIDI, S. **Dicas e truques de aprendizado de máquina**. Stanford University CS 229, p. 2–4, 2018. Disponível em: <<https://stanford.edu/~shervine/l/pt/teaching/cs-229/dicas-truques-aprendizado-maquina>>. Acesso em: 02 Fevereiro 2024.
- BHAVSAR, P. et al. **Machine learning in transportation data analytics**. In: CHOWDHURY, M.; APON, A.; DEY, K. *Data Analytics for Intelligent Transportation Systems*. [S.l.]: Elsevier, 2017. p. 283-307.
- BOMFIM, M A. **Aplicação de técnicas de aprendizado de máquina para determinar a velocidade de produção em máquinas coextrusoras**. 2021. p76. Engenharia de Produção - Universidade Tecnológica Federal do Paraná, Paraná, 2021.
- BUENO, E. P. **Avaliação do uso de aditivos no PET reciclado grau alimentício**. Trabalho de Conclusão de Curso. Universidade Federal do Rio Grande do Sul, 2016.
- CHEREMISINOFF, N. P. **Handbook of Engineering Polymeric Materials**. 1. ed. New York: Marcel Dekker, Inc. 1997.
- COSSETTI, M. C. **O que é inteligência artificial? TECNOBLOG**, 23 Novembro 2019. Disponível em: <<https://tecnoblog.net/responde/o-que-e-inteligencia-artificial/>>
- COUTINHO, Fernanda M.B.; MELLO, Ivana L.; MARIA, Luiz C. de Santa. **Polietileno: Principais Tipos, Propriedades e Aplicações**. *Polímeros: Ciência e Tecnologia*, vol. 13, n. 1, págs. 1-13, 2003.
- GOMES, D. D. S. **Inteligência artificial. Olhar Científico**, p. 234-246, 2010.
- HARADA, Júlio; WIEBECK, Hélio. **Plásticos de Engenharia - Tecnologia e Aplicações**. São Paulo: Artliber, 2005. 350 p.

Kemkes, J. F. - **US Patent 3497473**, 1970.

LOPES, H. F. **Aprendizado de máquina aplicado a previsão de desempenho de jogadores de futebol**. São Carlos: [s.n.], 2018.

Mancini, S. D; Matos I. G; Almeida R F. **Polímeros: Ciência e Tecnologia**, vol. 14, nº 2, p. 69-73, 2004.

MC KEEN, W. **Fatigue and Tribological Properties of Plastics and Elastomers**. 2 ed. New York: Elsevier Inc. 2010.

PREVILATE, M. F. **Análise do processo de pós condensação da resina de pet pós consumo reciclada**. Trabalho de Conclusão de Curso. Universidade Federal de São Carlos, 2022.

ROMÃO, Wanderson; SPINACÉ, Márcia; DE PAOLI, Marco. **Poli(tereftalato de etileno), PET: uma revisão sobre os processos de síntese, mecanismos de degradação e sua reciclagem**. *Polímeros*, São Carlos, v. 19, p. 121-132, jun 2009.

SCHÜSSLER, P. J.; BASTIANI, E.; BUSSLER, N. R. C. **Inteligência artificial e aprendizado de máquina: Utilizando o entendimento da inteligência humana para reprodução na computação**. *Salão do Conhecimento: Ciência para a redução das desigualdades - XXVI Seminário de Iniciação Científica*, 04 Outubro 2018.

SCOTT, D.V. **Advanced Materials for Water Handling**. 1 ed. New York: Elsevier Ltd., 2000.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding machine learning: From theory to algorithms**. Nova iorque: Cambridge University Press, 2014.

SILVA, L. A. **Otimização do Processamento da Resina PET Virgem**. Trabalho de Conclusão de Curso. Universidade Federal de Pernambuco, Pernambuco. 2023.

SILVA, L. **Estudo das variáveis envolvidas no controle de qualidade da resina PET, seu impacto, problemas e soluções para o processamento de injeção e sopro de garrafas**. Relatório de Estágio. Recife, PE: UFPE, 2022.

STRONG, A. Brent. **Plastics: materials and processing**. 3. ed. New Jersey: Pearson Prentice Hall, 2006.

SYAM, N.; SHARMA, A. **Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice**. *Industrial Marketing Management*, p. 135-146, Dezembro 2018.

TOLEDO, E.; OLIVEIRA, E.; FORTES, C.; MORAES, T; ARENA, M. **Manual Técnico PET**. 1. ed. São Paulo: M&G Polímeros S.A., 2008.

UCHIDA, C. N. **Os desafios do uso da resina de pet reciclado na indústria de alimentos e bebidas**. Trabalho de Conclusão de Curso. Universidade Federal de São Carlos, 2022.

WHELAN, T. **The Bekum Blow Moulding Handbook**. 1. ed. London: Bekum Maschinenfabriken GmbH, 1989.

ZEINEDDINE, H.; BRAENDLE, U.; FARAH, A. **Enhancing prediction of student success: Automated machine learning approach**. *Computers and Electrical Engineering*, 24 Novembro 2020.