



APLICAÇÃO DE BIG DATA ANALYTICS PARA TOMADA DE DECISÃO NA ÁREA DE MANUTENÇÃO

Felipe José dos Santos Bastos

Fundação Educacional Inaciana "PE. Sabóia de Medeiros- FEI

<https://orcid.org/0009-0004-8752-3193>

felipejsbastos@gmail.com

Romulo Rieder

Fundação Educacional Inaciana "PE. Sabóia de Medeiros- FEI

<https://orcid.org/0009-0007-6571-0694>

romulorieder@gmail.com

Marcos Geraldo Gomes

Fundação Educacional Inaciana "PE. Sabóia de Medeiros- FEI

<http://lattes.cnpq.br/7689589378187051>

mggomes1970@gmail.com

Geraldo Cardoso de Oliveira Neto

Universidade Federal do ABC - UFABC

<https://orcid.org/0000-0002-4744-3963>

geraldo.prod@gmail.com

Resumo

Este estudo investiga a aplicação de Big Data Analytics (BDA) na tomada de decisão na área de manutenção, explorando como essa tecnologia pode melhorar a eficiência operacional. O método de pesquisa adotado foi aplicação de técnica computacional em python via google colab para analisar um conjunto de dados sintéticos obtidos de máquinas de um processo de produção, representando condições reais de manutenção. Os resultados indicaram três primeiros componentes principais, relacionados às temperaturas, potência da máquina e desgaste da ferramenta



que explicam a maior parte da variância dos dados. Constataram-se que as falhas de dissipação de calor ocorrem em rotações por minuto (RPM) mais baixos e torque alto, enquanto falhas de desgaste da ferramenta são independentes de rpm e torque, mas relacionadas ao tempo de desgaste. Também, as falhas de sobrecarga ocorreram em torques altos e rpm baixos, e falhas de energia em torque alto e rpm baixos ou vice-versa. Com isso, o BDA aplicado na manutenção preventiva é essencial para antecipar problemas, otimizar processos e melhorar a competitividade. A integração dessas abordagens permite uma visão holística da operação, decisões baseadas em dados e melhorias contínuas nos processos de manutenção, resultando em economia de custos e maior eficiência.

Palavras-chave: Big Data Analytics, Manutenção Preventiva, Tomada de Decisão.

Abstract

This study investigates the application of Big Data Analytics (BDA) in decision-making in the area of maintenance, exploring how this technology can improve operational efficiency. The research method adopted was the application of a computational technique in Python via Google Colab to analyze a set of synthetic data obtained from machines in a production process, representing real maintenance conditions. The results indicated three main components related to temperatures, machine power, and tool wear that explain most of the data variance. It was found that heat dissipation failures occur at lower revolutions per minute (RPM) and high torque, while tool wear failures are independent of RPM and torque but related to wear time. Also, overload failures occurred at high torque and low RPM, and power failures at high torque and low RPM or vice versa. Thus, BDA applied to preventive maintenance is essential for anticipating problems, optimizing processes, and improving competitiveness. The integration of these approaches allows for a holistic view of the operation, data-driven decisions, and continuous improvements in maintenance processes, resulting in cost savings and greater efficiency.

Keywords: Big Data Analytics, Preventive Maintenance, Decision Making.

1 Introdução

Nas últimas décadas, o volume de dados disponíveis para as empresas tem aumentado de maneira significativa, resultando em desafios cada vez mais complexos no processo de tomada de decisão. Além do crescimento exponencial do volume de dados, há também uma multiplicação das fontes e de formatos dessas informações, caracterizando o fenômeno conhecido como Big Data. Segundo Baú, [2019](#), a tomada de decisão baseada em dados possui o potencial de transformar estratégias empresariais. Por outro lado, esse cenário exige uma preparação mais robusta dos profissionais responsáveis pela gestão estratégica e operacional nas organizações.



No contexto atual, o Big Data desempenha um papel importante na tomada de decisões estratégicas. As organizações enfrentam um ambiente empresarial cada vez mais complexo e competitivo, demandando informações precisas e em tempo real para fundamentar decisões bem embasadas. Conforme mencionado por Wamba et al., [2015](#), por meio da análise de Big Data, é possível extrair informações valiosas e basear decisões em evidências quantitativas.

Uma das principais vantagens do Big Data reside em sua capacidade de lidar com a variedade, velocidade e volume de dados, como destacado por Kwon e Sim, [2013](#); McAfee e Brynjolfsson, [2012](#); Russom, [2011](#). Ao combinar diferentes fontes de dados, como dados operacionais coletados por sensores, dados transacionais, dados de redes entre outros, é possível obter uma visão completa e abrangente de um problema ou situação específica. Além disso, as tecnologias relacionadas ao Big Data possibilitam o processamento desses dados em tempo real ou quase real, permitindo uma resposta ágil diante de mudanças e oportunidades, conforme discutido por Jeble et al., [2018](#).

No setor produtivo, a área de manutenção desempenha um papel crítico na garantia do desempenho, disponibilidade e confiabilidade dos ativos de uma organização. No entanto, essa área frequentemente enfrenta desafios na coleta, análise e interpretação de um grande volume de dados gerados por máquinas, sensores e sistemas de monitoramento (Souza et al., [2020](#)). A aplicação de Big Data Analytics (BDA) pode proporcionar indicações valiosas a partir desses dados, permitindo uma tomada de decisão mais informada e baseada em evidências. Dessa forma, este estudo tem o objetivo de investigar a aplicação de Big Data Analytics (BDA) na tomada de decisão na área de manutenção, explorando como essa tecnologia pode melhorar a eficiência operacional.

2 Revisão da Literatura sobre Evolução da Gestão de Manutenção

A gestão da manutenção tem evoluído significativamente ao longo da história, passando de uma abordagem reativa para modelos mais proativos e estratégicos. Inicialmente, a manutenção era realizada apenas quando ocorriam falhas, resultando em paradas inesperadas e



custos elevados. Com o tempo, surgiram práticas preventivas e preditivas, impulsionadas pelo desenvolvimento tecnológico e pela necessidade de aumentar a eficiência operacional (Roda e Macchi, 2021). Hoje, a manutenção é vista como um fator crítico para o sucesso organizacional, contribuindo para a redução de custos, aumento da produtividade e garantia da qualidade. A integração de tecnologias avançadas, como o Big Data, tem potencializado essa área, permitindo uma análise mais precisa e tomada de decisões baseadas em dados (Lee et al., 2020).

A manutenção preditiva é uma abordagem que utiliza dados em tempo real e técnicas avançadas de análise para prever quando equipamentos e sistemas podem falhar, permitindo intervenções antes que ocorram problemas. Suas características incluem o uso de sensores, monitoramento contínuo e algoritmos de análise preditiva (Zonta et al., 2020). As principais vantagens são a redução de tempo de inatividade não planejado, maior vida útil dos ativos e otimização dos recursos de manutenção. No entanto, a implementação dessa estratégia apresenta desafios, como o alto custo inicial de instalação de sensores e sistemas de monitoramento, a necessidade de habilidades especializadas para interpretar os dados e a integração com sistemas existentes (Nunes et al., 2023).

A Indústria 4.0 representa a quarta revolução industrial, caracterizada pela integração de tecnologias avançadas, como Internet das Coisas (IoT), inteligência artificial (IA), big data e automação, nos processos de manufatura e produção (Martinelli et al., 2021). No contexto da gestão da manutenção, a Indústria 4.0 oferece oportunidades significativas, como a implementação de manutenção preditiva, otimização de processos e maior eficiência operacional (Zonta et al., 2020). Contudo, esses avanços também trazem desafios, incluindo a necessidade de investimentos substanciais em tecnologia, treinamento de pessoal para lidar com novas ferramentas e a complexidade da integração de sistemas antigos com novas soluções digitais (Silvestri et al., 2020). A adaptação bem-sucedida a esses desafios pode resultar em maior competitividade e sustentabilidade para as organizações.

No contexto da Indústria 4.0, Big Data refere-se ao vasto volume de dados gerados e coletados em velocidades sem precedentes, provenientes de diversas fontes, como diferentes sensores instalados nas cadeias de produção (Wang et al., 2022). Para as empresas, Big Data oferece informações valiosas para tomar decisões informadas, identificar tendências de mer-



cado e melhorar a eficiência operacional (Azeem et al., 2022). Especificamente na gestão da manutenção, Big Data possibilita a análise preditiva de falhas, a otimização de cronogramas de manutenção e a redução de custos operacionais (Bousdekis et al., 2021). Ao processar e analisar grandes quantidades de dados em tempo real, as empresas podem antecipar problemas antes que estes ocorram, prolongando a vida útil dos equipamentos e minimizando paradas não planejadas, resultando em operações mais seguras e produtivas.

Business Intelligence (BI) é o processo de coleta, análise e transformação de dados em informações acionáveis para suportar decisões estratégicas e operacionais nas empresas. Em geral, BI permite às organizações monitorar desempenho, identificar oportunidades de melhoria e tomar decisões baseadas em dados concretos (Tavera Romero et al., 2021). No contexto da gestão da manutenção, BI, combinado com Big Data, oferece ferramentas poderosas para análise preditiva e prescritiva, ajudando a prever falhas e otimizar atividades de manutenção (Picozzi et al., 2024). Através de dashboards interativos e relatórios detalhados, gestores podem visualizar padrões, monitorar indicadores-chave de desempenho (KPIs) e implementar estratégias mais eficientes, resultando em maior confiabilidade e eficiência dos equipamentos e operações.

3 Metodologia

Esta seção apresenta a metodologia empregada nesta pesquisa, a partir da coleta e análise dos dados, conforme descrito nos itens a seguir.

3.1 Coleta de dados

Considerando que conjuntos de dados reais de manutenção são geralmente difíceis de obter e, em particular, difíceis de serem publicados (Braun e Ong, 2018), os dados foram obtidos de máquinas de um processo de produção, Kaggle: Your Machine Learning and Data Science Community, que disponibiliza uma base de dados aberta do UCI Machine Learning Repository, 2020.

O repositório de aprendizado de máquina da UCI (Universidade da Califórnia Irvine) mantém atualmente 488 conjuntos de dados de diversas características como um serviço para a



comunidade de machine learning (Chang et al., 2020). O conjunto sintético de dados utilizado reflete a manutenção real encontrada na indústria e é composto por 10.000 pontos de dados armazenados como linhas (Mota et al., 2023), com as seguintes características em colunas:

- a) UID: Identificador único variando de 1 a 10.000;
- b) ID Produto: Número de série específico para cada variante de Tipo;
- c) Tipo: Composto por uma letra L, M ou H para representar variantes de qualidade do produto. Baixa qualidade (L), representando 60% dos produtos; média qualidade (M) representando 30% dos produtos; e alta qualidade (H), representando 10% dos produtos;
- d) Temperatura do ar [K]: Gerada usando um processo de caminhada aleatória e posteriormente normalizada para um desvio padrão de 2 K em torno de 300 K;
- e) Temperatura do processo [K]: Gerada usando um processo de caminhada aleatória, normalizada para um desvio padrão de 1 K e adicionada à temperatura do ar mais 10 K;
- f) Velocidade de rotação [rpm]: Calculada a partir de uma potência de 2860 W, sobreposta com ruído distribuído normalmente;
- g) Torque [Nm]: Valores de torque distribuídos normalmente em torno de 40 N.m, com um desvio padrão de 10 N.m e sem valores negativos;
- h) Desgaste da ferramenta [min]: As variantes de qualidade H/M/L adicionam 5/3/2 minutos de desgaste à ferramenta usada no processo;
- i) Alvo: Indicador de falha;
- j) Tipo da Falha: Rótulo que indica se a máquina falhou neste ponto de dados específico, para qualquer um dos seguintes modos de falha, conforme Patil et al., 2023:
 - Falha de desgaste da ferramenta (TWF): A ferramenta será substituída ou falhará em um momento de desgaste da ferramenta selecionado aleatoriamente entre 200 e 240 minutos;

- Falha de dissipação de calor (HDF): A dissipação de calor causa uma falha no processo se a diferença entre a temperatura do ar e a temperatura do processo for inferior a 8,6 K e a velocidade de rotação da ferramenta for inferior a 1380 rpm;
- Falha de energia (PWF): O produto do torque e da velocidade de rotação (em rad/s) corresponde à potência necessária para o processo. Se essa potência estiver abaixo de 3500 W ou acima de 9000 W, o processo falha;
- Falha de sobrecarga (OSF): Se o produto do desgaste da ferramenta e do torque exceder 11.000 min.N.m para a variante de produto L (12.000 para M, e 13.000 para H), o processo falha devido à sobrecarga;
- Falhas aleatórias (RNF): Cada processo tem uma chance de 0,1% de falhar, independentemente de seus parâmetros de processo. Se pelo menos um dos modos de falha mencionados acima for verdadeiro, o processo falhará e o rótulo será “falha da máquina”.

3.2 Procedimento de Análise de Dados

Os dados coletados serão analisados de forma exploratória, com análise descritiva para identificar padrões e informações relevantes (Myatt, 2007). Para a análise de Big Data, foi utilizada a simulação computacional por meio da linguagem de programação Python, através do uso da ferramenta on-line Google Colab (Carneiro et al., 2018). Os resultados são apresentados a seguir.

4 Resultados

A análise dos dados iniciou-se com uma exploração da distribuição da coluna “Tipo de falha”, revelando um desequilíbrio significativo na distribuição das classes. Notavelmente, aproximadamente 96,69% das observações foram categorizadas como “Sem Falha”, o que poderia ser uma indicação de um problema de classificação. Um exame mais aprofundado foi desenvolvido nos 3,31% restantes do conjunto de dados em que ocorreu algum tipo de falha de máquina. Outra análise importante diz respeito à ocorrência de falhas de máquina em todo o

conjunto de dados, cuja percentagem foi de 3,31%. Para demonstrar a distribuição das falhas, gerou-se um novo gráfico no modelo "pie chart", como mostra a Figura 1, detalhando as causas envolvidas em cada falha, revelando um grau significativo de desequilíbrio.

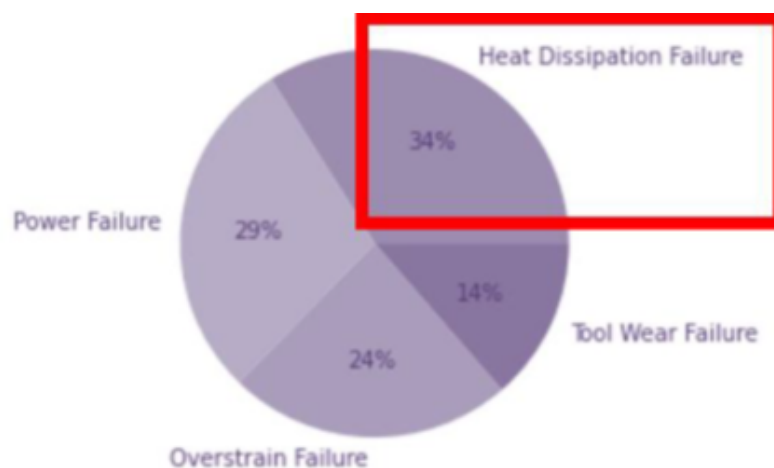


Figura 1: Ocorrência das Causas de Falhas

Fonte: Elaborado pelo Autor.

Neste primeiro gráfico, nota-se que o percentual de observações classificadas como "Falha de Dissipação de Calor" representou o maior foco de falhas, com 34% do total observado, destacado em vermelho. Essa análise pode indicar falhas relacionadas à temperatura no processo, porém, é necessário analisar a relação entre as variáveis para confirmação.

Diante das análises acima, foram gerados novos gráficos de dispersão para explorar as relações entre as variáveis, conforme pode ser observado na Figura 2. Os gráficos de dispersão, também conhecidos como gráficos de pares, ilustram a relação entre duas variáveis. No contexto de um conjunto de dados de manutenção, esses gráficos podem ser úteis para examinar as relações entre diferentes recursos e a probabilidade de falha de uma máquina.

Através do gráfico de dispersão (Figura 2), identificou-se que as falhas de energia tendem a ser correlacionadas com valores de torque próximos aos limites máximo ou mínimo, conforme destacado em vermelho no gráfico. Essa análise sugere que a falha de dissipação de calor, notada no Gráfico 2, pode estar correlacionada com as falhas de energia.

Seguindo as investigações, a partir dos gráficos de densidade, foi possível observar que os casos de falha de produto estão principalmente relacionados ao "Desgaste de Máquina" de qualidade "L", seguidos pela qualidade "M" e, em poucas ocasiões, pela qualidade "H". No

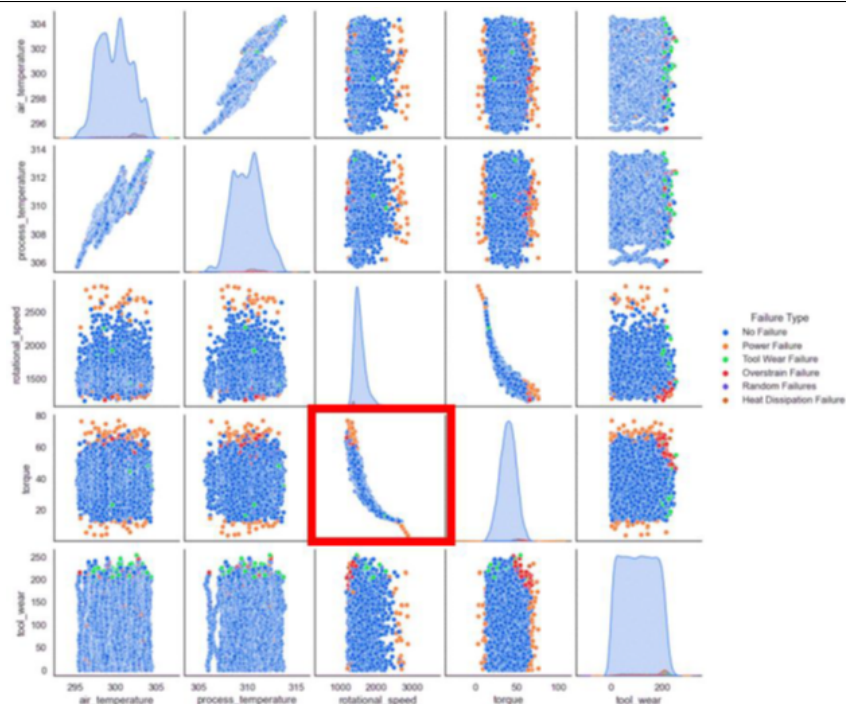


Figura 2: Relações entre Diferentes Recursos e Variáveis

Fonte: Elaborado pelo Autor.

entanto, isso não está amplamente correlacionado com as características, pois, ao diferenciar a qualidade, a distribuição das características não apresenta grandes diferenças, exceto pelo pico lateral no Desgaste da Ferramenta, conforme ressaltado no gráfico apresentado na Figura 3.

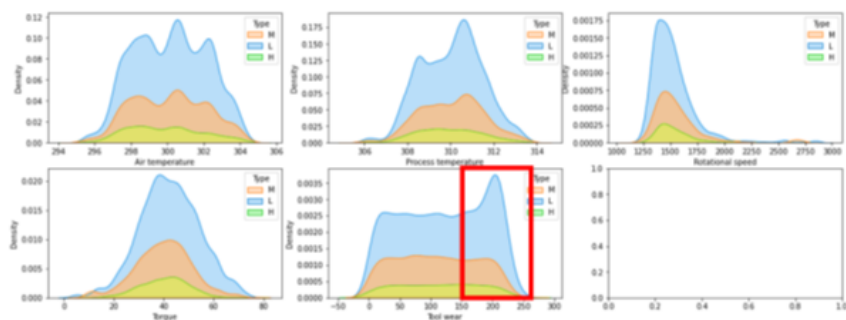


Figura 3: Relações entre Diferentes Recursos e Variáveis

Fonte: Elaborado pelo Autor.

Estes gráfico da Figura 3 foram gerados com os comandos abaixo:

Essa interpretação sugere que a maioria das falhas que ocorre em produtos do tipo L se deve à maior presença desse tipo no conjunto de dados, e a correlação com a falha da máquina se deve a razões estatísticas. Em seguida, foi aplicada a Análise de Componentes Principais (PCA) para proporcionar uma visualização adicional dos dados em vez de realizar a seleção de

```
fig, axs = plt.subplots(nrows=2, ncols=3, figsize=(19,7))
custom_palette = {'L':'tab:blue', 'M':'tab:orange', 'H':'tab:green'}
for j, feature in enumerate(num_features):
    sns.kdeplot(ax=axs[j//3, j-3*(j//3)], data=df_res, x=feature,
               hue='Type', fill=True, palette=custom_palette)
plt.show()
```

Figura 4: Comandos para Geração dos Gráficos da Figura 3
Fonte: Elaborado pelo Autor.

recursos manualmente. Através dessa análise, foi calculada a representatividade percentual de cada um dos principais componentes.

```
Percentual de variância explicada por componentes:
PC1    37.69
PC2    36.81
PC3    19.84
PC4     3.08
PC5     2.58
Tipo de dados: float64
Percentual de Variância explicada com 3 componentes: 94.34
```

Figura 5: Resultados Gerados no Console
Fonte: Elaborado pelo Autor.

Nota-se que no gráfico da Figura 6, um "barplot" dos principais componentes, a compreensão da representação da variância dos três principais componentes foi facilitada.

- PC1 está intimamente relacionado aos dados de temperatura;
- PC2 está associado à potência da máquina, que é o produto da Velocidade de Rotação e do Torque;
- PC3 está relacionado ao Desgaste da Ferramenta.

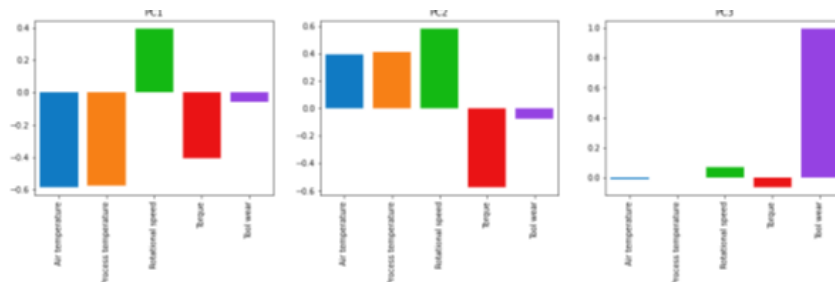


Figura 6: Principais Componentes
Fonte: Elaborado pelo Autor.

Visto que os três primeiros componentes são capazes de representar quase completamente a variância dos dados, foi gerada a projeção desses componentes em um espaço tridimensional, conforme mostrado na Figura 7.

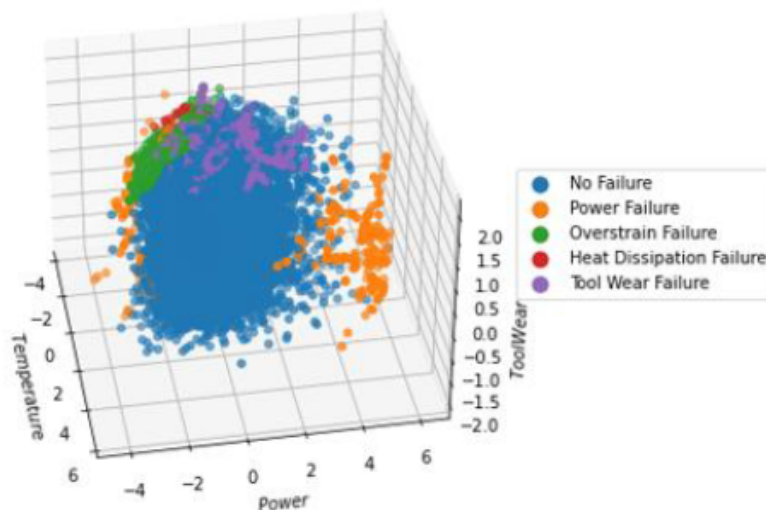


Figura 7: Visão Tridimensional dos Principais Componentes

Fonte: Elaborado pelo Autor.

Com a renderização do gráfico apresentado na Figura 7, identificou-se que a inclusão das ocorrências classificadas como "No Failure" dificultava a interpretação das demais ocorrências em relação às variáveis de temperatura e energia. Assim, gerou-se um novo gráfico, exibido abaixo na Figura 8, desconsiderando essas observações para uma análise mais clara.

Este gráfico foi gerado por meio dos comandos mostrados na Figura 9:

A partir da projeção no espaço gerado por esses três eixos, pode-se destacar o seguinte:

- A falha de desgaste da ferramenta (TWF) é a classe de falhas mais bem separada de todas as outras.
- A falha de energia (PWF) ocupa duas faixas extremas ao longo do PC2 (Potência) e é independente dos outros dois componentes.
- As classes Falha de sobrecarga (OSF) e Falha de dissipação de calor (HDF) estão menos separadas do que as outras, embora a primeira seja caracterizada por um alto Desgaste da Ferramenta e baixa potência, enquanto a segunda é caracterizada por uma alta temperatura e baixa potência.

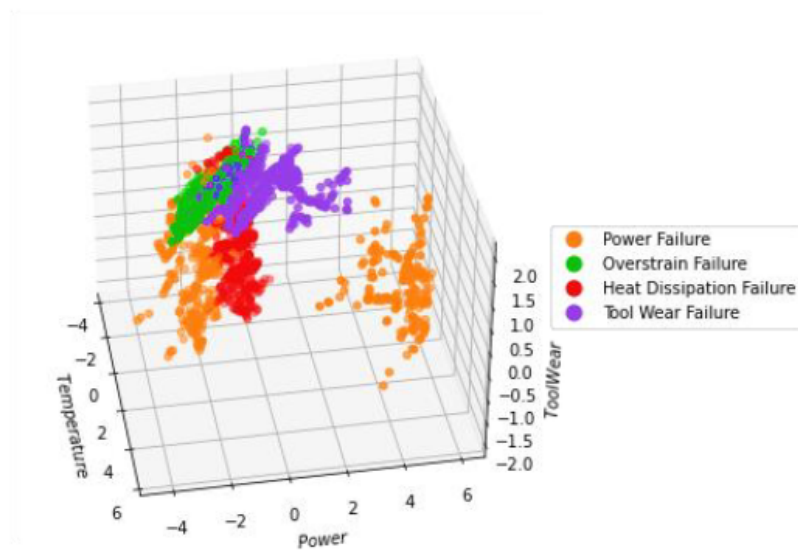


Figura 8: Visão Tridimensional dos Principais Componentes

Fonte: Elaborado pelo Autor.

Com essas informações e para uma análise de correlação mais profunda, gerou-se um mapa de calor, conforme o gráfico da Figura 10, que permite observar que há uma forte correlação positiva entre "Temperatura do Ar" e "Temperatura de Processo" (0,88). Isso se explica pelo fato de que uma alta temperatura de processo naturalmente aquece o ar ao redor da máquina. Além disso, existe uma forte correlação negativa entre "Velocidade de Rotação" e "Torque" (-0,88). As demais correlações são mais fracas, próximas de 0, indicando relações menos significativas.

Inferir-se do gráfico da Figura 10 que os tipos de falha ocorrem na correlação entre temperatura do ar e temperatura do processo de maneira mais atenuada, e na velocidade de rotação e torque. Compreender as correlações entre diferentes variáveis em um conjunto de dados pode ser útil para construir modelos preventivos, pois pode indicar quais recursos são mais importantes para prever um determinado alvo.

5 Conclusão

De acordo com as análises realizadas e os resultados obtidos, é possível fazer algumas considerações conclusivas. Após realizar as análises através dos gráficos apresentados, foi identificado que os três primeiros componentes explicam a maior parte da variância dos dados. Esses

```
X_pca3.rename(mapper={'PC1':'Temperature',
                    'PC2':'Power',
                    'PC3':'Tool Wear'}, axis=1, inplace=True)

# PCA plot
color = []
col = df_pre['Failure Type'].map({0:'tab:blue',1:'tab:orange',2:'tab:green',
3:'tab:red',4:'tab:purple'})
color.append(col)
idx_w = col[col == 'tab:blue'].index
color.append(col.drop(idx_w))
colors = ['tab:blue', 'tab:orange', 'tab:green', 'tab:red', 'tab:purple']
labelTups = [('No Failure', 'tab:blue'),
             ('Power Failure', 'tab:orange'),
             ('Overstrain Failure', 'tab:green'),
             ('Heat Dissipation Failure', 'tab:red'),
             ('Tool Wear Failure', 'tab:purple')]

fig = plt.figure(figsize=(18,6))
fig.suptitle('Data in 3D PCA space')
full_idx = X_pca3.index

for j, idx in enumerate([full_idx,idx_fail_res]):
    ax = fig.add_subplot(1, 2, j+1, projection='3d')

    lg = ax.scatter(X_pca3.loc[idx, 'Temperature'],
                   X_pca3.loc[idx, 'Power'],
                   X_pca3.loc[idx, 'Tool Wear'],
                   c=color[j])

    ax.set_xlabel('$Temperature$')
    ax.set_ylabel('$Power$')
    ax.set_zlabel('$Tool Wear$')
    ax.title.set_text('With'+str(j*'out')+' "No Failure" class')
    ax.view_init(35, -10)
    custom_lines = [plt.Line2D([],[], ls="", marker='.',
                                mec='k', mfc=c, mew=.1, ms=20) for c in colo
rs[j:]]
    ax.legend(custom_lines, [lt[0] for lt in labelTups[j:]],
              loc='center left', bbox_to_anchor=(1.0, .5))

plt.show()
```

Figura 9: Comandos para Geração do Gráfico da Figura 8

Fonte: Elaborado pelo Autor.

componentes podem ser representados pelas seguintes características: a combinação das duas temperaturas (ar e processo), a potência da máquina (produto da velocidade de rotação e torque) e o desgaste da ferramenta. Com base nisso, constatou-se que essas são as características que mais influenciam no percentual de variância.

Adicionalmente, foi possível constatar que:

- A falha de dissipação de calor ocorre em rotações por minuto (RPM) mais baixas e torque mais alto quando a diferença de temperatura é menor, indicando uma maior frequência quando a máquina está iniciando.
- A falha de desgaste da ferramenta ocorre principalmente quando o tempo de desgaste da ferramenta é elevado, sendo independente de RPM e torque.

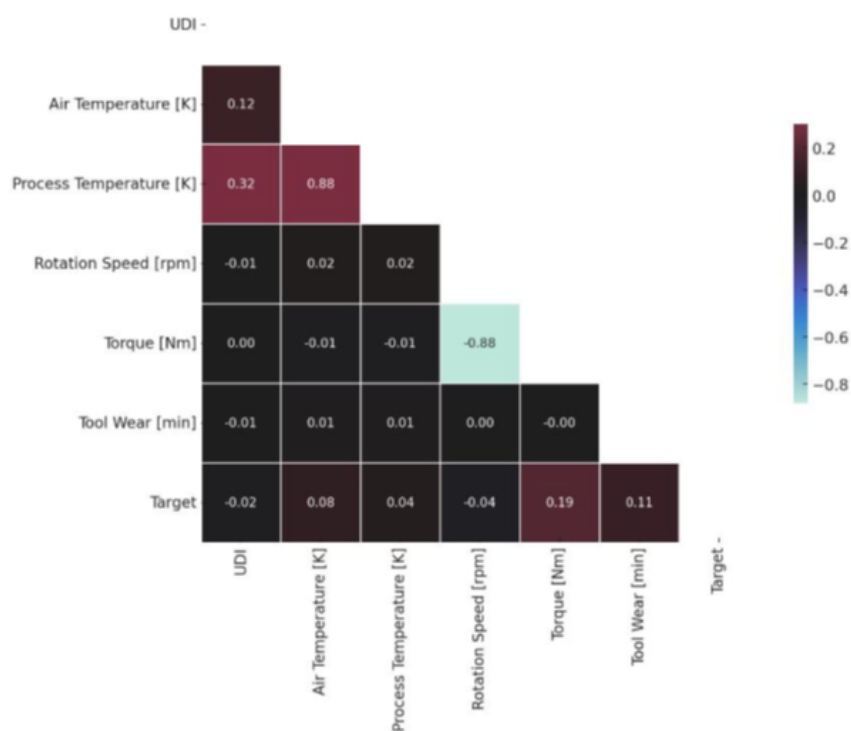


Figura 10: Mapa de Calor
Fonte: Elaborado pelo Autor.

- A falha de sobrecarga ocorre em torques mais altos e RPM mais baixas, mas, ao contrário da falha de dissipação de calor, é independente da diferença de temperatura. No entanto, requer um tempo de desgaste da ferramenta elevado, similar à falha por desgaste da ferramenta.
- As falhas de energia ocorrem quando há torque alto e RPM baixas ou RPM altas e torque baixo.
- As principais falhas identificadas nas máquinas revelam uma possível deficiência de manutenção, resultando em prejuízos no andamento dos processos ao interferir na eficiência da produção.

A manutenção preditiva e o Business Intelligence desempenham papéis fundamentais na eficiência e no sucesso dos negócios. A manutenção preditiva, por meio da análise avançada de dados, permite identificar falhas futuras e problemas potenciais, evitando interrupções e reduzindo custos de manutenção. O Business Intelligence, por sua vez, coleta e analisa dados de várias fontes, fornecendo informações valiosas sobre o desempenho operacional e oportuna-



des de melhoria. A combinação dessas abordagens possibilita uma visão holística da operação, além de tomada de decisões embasadas em dados e aprimoramento contínuo dos processos e estratégias de manutenção.

A integração eficaz da manutenção preditiva e do Business Intelligence pode resultar em melhorias significativas na eficiência operacional. A capacidade de identificar padrões de falha, relacionar dados de manutenção com dados operacionais e tomar medidas proativas permite antecipar problemas e realizar ajustes precisos nos processos. Além disso, a análise avançada de dados proporciona uma compreensão mais profunda dos impactos financeiros e das oportunidades de otimização, gerando economia de custos e maior competitividade.

Ao adotar a manutenção preditiva e o Business Intelligence, as empresas alcançam excelência operacional, impulsionam a inovação e se destacam no mercado. Essas práticas combinadas permitem tomar decisões embasadas em dados, melhorar a qualidade dos produtos e serviços, e obter um retorno mais alto sobre os investimentos em ativos e infraestrutura. Com o uso inteligente dessas abordagens, as empresas podem obter uma vantagem competitiva duradoura, oferecendo serviços superiores e garantindo sucesso a longo prazo.

Como demonstrado nesse relatório, a quantidade de falhas geradas pode ser prevista com a utilização de sensores instalados nas máquinas e equipamentos utilizados pela empresa. Há um momento claro na produção, por exemplo, entre 4000 e 5000 unidades onde diversas falhas ocorrem, como em um efeito em cadeia. Essa informação demonstra quão assertiva a incorporação dessas tecnologias pode ser. Conforme apresentado neste relatório, a utilização do Big Data Analytics e da ferramenta Python podem influenciar positivamente o gerenciamento de dados e criação de relatórios gerenciais, fazendo deste um instrumento importante no contexto da manutenção preditiva, imprescindível para redução de desperdícios.

Com o advento da Inteligência Artificial, é possível que em breve as estratégias relacionadas ao Big Data Analytics e à manutenção preditiva estejam mais integradas, com dados reais sendo processados no momento da manufatura dos produtos, o que levará a menor desperdício e melhor controle de toda a cadeia produtiva. Como sugestão para um relatório futuro, novos gráficos de correlação, específicos para cada tipo de erro podem auxiliar no processo de tomada de decisão.



Referências Bibliográficas

Baú, J. M. (2019). Implicações do fenômeno Big Data na tomada de decisão baseada em dados em uma cooperativa de crédito [Acesso em: 21 maio 2024].

Wamba, S. F., et al. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234–246.

Kwon, O., & Sim, J. M. (2013). Effects of data set features on the performances of classification algorithms. *Expert Systems with Applications*, 40(5), 1847–1857.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard Business Review*, 90(10), 60–68.

Russom, P. (2011). Big data analytics. *TDWI Best Practices Report*, 19(4), 1–34.

Jeble, S., Kumari, S., & Patil, Y. (2018). Role of Big Data in Decision Making. *Operations And Supply Chain Management*, 11(1), 36–44.

Souza, M., da Costa, C., de Oliveira Ramos, G., & da Rosa Righi, R. (2020). A survey on decision-making based on system reliability in the context of Industry 4.0. *Journal of Manufacturing Systems*, 56, 133–156.

Roda, I., & Macchi, M. (2021). Maintenance concepts evolution: A comparative review towards advanced maintenance conceptualization. *Computers in Industry*, 133, 103531.

Lee, J., Ni, J., Singh, J., Jiang, B., Azamfar, M., & Feng, J. (2020). Intelligent maintenance systems and predictive manufacturing. *Journal of Manufacturing Science and Engineering*, 142(11), 110805.



Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers & Industrial Engineering, 150*, 106889.

Nunes, P., Santos, J., & Rocha, E. (2023). Challenges in predictive maintenance: A review. *CIRP Journal of Manufacturing Science and Technology, 40*, 53–67.

Martinelli, A., Mina, A., & Moggi, M. (2021). The enabling technologies of industry 4.0: examining the seeds of the fourth industrial revolution. *Industrial and Corporate Change, 30*(1), 161–188.

Silvestri, L., Forcina, A., Introna, V., Santolamazza, A., & Cesarotti, V. (2020). Maintenance transformation through Industry 4.0 technologies: A systematic literature review. *Computers in Industry, 123*, 103335.

Wang, J., Xu, C., Zhang, J., & Zhong, R. (2022). Big data analytics for intelligent manufacturing systems: A review. *Journal of Manufacturing Systems, 62*, 738–752.

Azeem, M., Haleem, A., Bahl, S., Javaid, M., Suman, R., & Nandan, D. (2022). Big data applications to take up major challenges across manufacturing industries: A brief review. *Materials Today: Proceedings, 49*, 339–348.

Bousdekis, A., Lepenioti, K., Apostolou, D., & Mentzas, G. (2021). A review of data-driven decision-making methods for industry 4.0 maintenance applications. *Electronics, 10*(7), 828.

Tavera Romero, C., Ortiz, J., Khalaf, O., & Ríos Prado, A. (2021). Business intelligence: business evolution after industry 4.0. *Sustainability, 13*(18), 10026.

Picozzi, P., Nocco, U., Pezzillo, A., De Cosmo, A., & Cimolin, V. (2024). The Use of Business Intelligence Software to Monitor Key Performance Indicators (KPIs) for the Evaluation of a Computerized Maintenance Management System (CMMS). *Electronics, 13*(12), 2286.

Braun, M. L., & Ong, C. S. (2018). Open science in machine learning. Em *Implementing Reproducible Research* (pp. 343–365). Chapman; Hall/CRC.



UCI Machine Learning Repository. (2020). AI4I 2020 Predictive Maintenance Dataset [Acesso em: 21 maio 2024].

Chang, S., Shihong, Y., & Qi, L. (2020). Clustering characteristics of UCI dataset. *2020 39th Chinese Control Conference (CCC)*, 6301–6306.

Mota, B., Faria, P., & Ramos, C. (2023). Machine Learning Applied to Industrial Machines for an Efficient Maintenance Strategy: A Predictive Maintenance Approach. *Energy Informatics Academy Conference*, 289–299.

Patil, C., Jadhav, S., Bardiya, A., Davande, A., & Raverkar, M. (2023). Machine Learning-Based Predictive Maintenance of Industrial Machines. *International Journal of Computer Trends and Technology*, 71(3), 50–56.

Myatt, G. J. (2007). *Making sense of data: a practical guide to exploratory data analysis and data mining*. John Wiley & Sons.

Carneiro, T., da Nóbrega, R., Nepomuceno, T., Bian, G., de Albuquerque, V., & Reboucás Filho, P. (2018). Performance analysis of google colab as a tool for accelerating deep learning applications. *IEEE Access*, 6, 61677–61685.