

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA
E TECNOLOGIA FLUMINENSE**

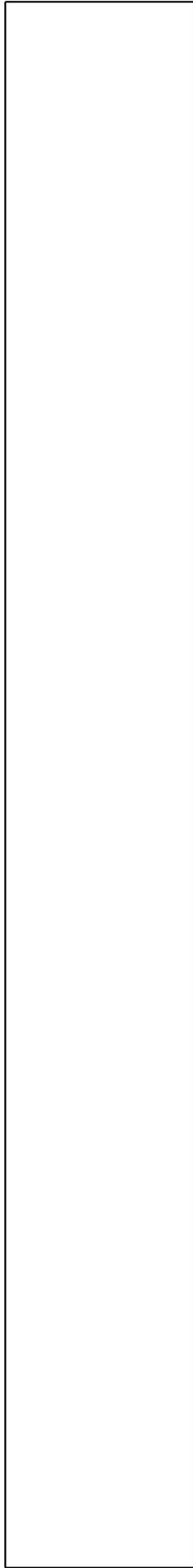
**PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS APLICADOS À
ENGENHARIA E GESTÃO**

Higor Medina Velloso

**PREVISÃO DE SURGIMENTO DE *SLUGS* EM SISTEMAS DE
BOMBEIO DE ÓLEO E GÁS DE UM CAMPO DE PRODUÇÃO
SUBMARINO ATRAVÉS DE FERRAMENTAS DE INTELIGÊNCIA
ARTIFICIAL**

Campos dos Goytacazes/RJ

2019



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA
E TECNOLOGIA FLUMINENSE**

**PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS APLICADOS À
ENGENHARIA E GESTÃO**

HIGOR MEDINA VELLOSO

**PREVISÃO DE FALHAS EM ÁRVORES DE NATAL SUBMARINAS DO PRÉ-SAL
UTILIZANDO FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL**

**Henrique Rego Monteiro Da Hora
(Orientador)**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação do Instituto Federal de Educação, Ciência e Tecnologia Fluminense, no Curso de Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão (MPSAEG), como parte dos requisitos necessários à obtenção do título de Mestre em Sistemas Aplicados à Engenharia e Gestão.

**Campos dos Goytacazes/RJ
2019**

Biblioteca Anton Dakitsch

CIP – Catalogação na Publicação

S586pSiSilva, Sergio Tobias da

S586p

Propagação do Som: Conceitos e Experimentos / Sergio
Tobias da Silva - Rio de Janeiro: UFRJ / IF, 2011. viii, 77 f.: il.;30cm.

Orientador: Carlos Eduardo Aguiar

Dissertação (mestrado) – UFRJ / Instituto de Física / Programa de
Pós-Graduação em Ensino de Física, 2011. Referências Bibliográficas: f. 74-
77.

1. Ensino de Física. 2. Som. 3. Velocidade do som. I. Aguiar, Carlos
Eduardo. II. Universidade Federal do Rio de Janeiro, Instituto de Física,
Programa de Pós-Graduação em Ensino de Física. III. Propagação do Som:
Conceitos e Experimentos.

Elaborada pelo Sistema de Geração Automática de Ficha catalográfica da Biblioteca Anton
Dakitsch do IFF com os dados fornecidos pelo(a) autor(a).

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
FLUMINENSE

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS APLICADOS À
ENGENHARIA E GESTÃO

Higor Medina Velloso

PREVISÃO DE FALHAS EM ÁRVORES DE NATAL SUBMARINAS DO PRÉ-SAL
UTILIZANDO FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação do Instituto Federal de Educação, Ciência e Tecnologia Fluminense, no Curso de Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão (MPSAEG), como parte dos requisitos necessários à obtenção do título de Mestre em Sistemas Aplicados à Engenharia e Gestão.

Aprovado(a) em xx de xxxx de 20**.

Banca Examinadora:

Henrique Rego Monteiro da Hora, D.Sc
Instituto Federal de Educação, Ciência e Tecnologia Fluminense
(Orientador)

Nome, Título
Instituição
(Coorientador)

Nome, Título
Instituição

Nome, Título
Instituição

Dedico este trabalho para toda a minha família que me apoiou e incentivou a estudar, em especial à minha futura esposa Natália que me colocou para a frente todas as vezes em que quis desistir do programa de mestrado. Sem sua ajuda, não teria chegado até aqui.

Dedico também este trabalho à comunidade Python que é absurdamente colaborativa e, sempre que tem uma oportunidade, cria uma nova biblioteca para facilitar a vida de outros desenvolvedores. Se não fossem os exemplos e bibliotecas que vocês criam, eu com certeza teria que pedir a postergação do prazo de apresentação.

Por fim, dedico este trabalho aos meus gerentes e coordenadores da TechnipFMC, em especial a Gustavo Oliveira e Ranieri Santos, que me permitiram sair todas as sextas ao meio-dia para assistir às aulas do mestrado, ou quando era necessário ir a congressos e conferências. A visão de que esse conhecimento seria um investimento e não um perda de tempo para a empresa é fundamental para o crescimento do negócio e desenvolvimento de novas mentes.

AGRADECIMENTOS

Agradeço ao meu orientador por toda a paciência e chamadas que foram dadas para que eu voltasse para os trilhos e não vagasse por diversas ideias. Sua dedicação e disponibilidade fizeram toda a diferença para o acontecimento deste trabalho.

Agradeço à Daniela Carvalho Martins por ter me ajudado na organização e estruturação dos dados para a confecção do artigo C. Seu empenho poupou muitas horas de trabalho e fez com que o trabalho saísse mais rápido.

RESUMO

A indústria de Óleo & Gás é uma das maiores, mais complexas e mais importantes indústrias a nível global, influenciando a vida de todos as pessoas com diversos produtos como combustíveis, asfalto e outros diversos produtos de tapetes a óculos a roupas. Na economia o panorama não é diferente, em 2015, a indústria de óleo e gás gerou um valor de 1,3 trilhões de dólares nos Estados Unidos, o que representa mais de 7% do PIB deste mesmo ano do país. Nos campos *offshore*, diversos equipamentos críticos compõem o arranjo submarino, como Árvores de Natal Molhadas (ANMs) e bombas elétricas submersíveis (ESPs, em inglês). Neste contexto, saber dizer qual componente irá falhar e quando se torna fundamental para evitar intervenções desnecessárias. Visando melhorar a detecção e reconhecimento de padrões de falha, ferramentas de inteligência artificial (IA) tem sido aplicadas e estudadas tanto da comunidade científica quanto das indústrias. Este trabalho vem dar sua contribuição com o objetivo de aplicar a técnica de *Support vector Machine* (SVM) para a criação de um modelo a fim de prever *slugs* em um sistema submarino através de dados de campo. Este trabalho está estruturado em três artigos sendo o primeiro focado em bibliometria (artigo A), o segundo focado em revisão sistemática (artigo B) e o terceiro um estudo de caso (artigo C). Como resultado, foi possível gerar modelos de predição com acurácia superior a 90%.

Palavras-chave: previsão de falhas, ESP, SVM.

ABSTRACT

The oil and gas industry is one of the biggest, more complex and important industry in a global level, influencing everybody's life with many products such as fuel, carpets, glasses and clothes. In economy there is no difference. In 2015, the oil and gas industry generated a total of 1.3 trillion of dollars in United States. This same value represents 7% of the GDP of that year in the country. In offshore fields, many critical equipment compound the subsea layout, such as Christmas Trees (XTs) and Electrical Submersible Pumps (ESPs). In this context, to know when and which component will fail becomes crucial to avoid unnecessary interventions. Aiming to improve the detection and recognition of failure patterns, artificial intelligence tools have been used and studied by the scientific community and companies. This work come to use the Support Vector Machine (SVM) technique to create a model able to predict slugs in an subsea system through field data. This work is structured in three papers: the first one is focused on a bibliometry (paper A), the second one on a systematic revision (paper B) and the last one on a case study (paper C). As result, prediction models were generated and achieved a precision over 90%.

Keywords: failure prediction, ESP, SVM.

LISTAS

Lista de ilustrações

Figura 1-1 – Parcela de energia global por fonte.	12
Figura 1-2 – Parcela de energia por fonte no Brasil.	12
Figura 1-3 – Histórico de produção de petróleo no Brasil por origem.	13
Figura 3-1 – Palavras-chave, tesouros e termos correspondentes.	19
Figura 3-2 – Pesquisa realizada no Scopus®.	20
Figura 3-3 – Diagrama de Venn das pesquisas realizadas.	21
Figura 3-4 – Diagrama de Pareto dos veículos com o maior número de publicações para os conceitos AB.	22
Figura 3-5 – Diagrama de Pareto dos veículos com o maior número de publicações para os conceitos AC.	22
Figura 3-6 – Levantamento cronológico das publicações.	24
Figura 3-7 – Diagrama de Pareto dos artigos com o maior número de citações para os conceitos AB.	25
Figura 3-8 – Diagrama de Pareto dos artigos com o maior número de citações para os conceitos AC.	25
Figura 3-9 – Diagrama de Pareto dos autores com o maior número de publicações para os conceitos AB.	29
Figura 3-10 – Diagrama de Pareto dos autores com o maior número de publicações para os conceitos AC.	30
Figura 3-11 - Setença de busca no Scopus®.	31
Figura 3-12 - Gráfico de Sankey do fluxo de conhecimento.	32
Figura 3-13 - Fontes mais relevantes de publicação de artigos na área.	33
Figura 3-14 - Autores com maior número de produções e citações.	33
Figura 3-15 - Evolução da produção científica por ano.	34
Figura 3-16 - Distribuição da produção científica por país.	35
Figura 3-17 - Termos mais utilizados por ano.	35
Figura 3-18 - Mapa de árvore dos termos mais citados.	36
Figura 4-1 - Palavras-chave e tesouros utilizados na busca.	43
Figura 4-2 - Fluxograma do procedimento de pesquisa.	45
Figura 4-3 - Segmentação dos artigos de previsão de falhas.	46
Figura 5-1 - Esquema de arranjo submarino.	60

Lista de quadros

Quadro 1-1 – Subdivisão dos objetivos do trabalho	15
Quadro 3-1 – Veículos com maior número de publicações sobre o tema	22
Quadro 3-2 – Artigos mais citados por pesquisa.....	25
Quadro 4-1 - Sentenças de busca.....	43
Quadro 5-1 - Variáveis coletadas.....	63
Quadro 5-2 - Resumo dos valores da coluna (v)	63
Quadro 5-3 - Resumo das colunas do banco de dados	64
Quadro 5-4 - Distâncias e vazão dos poços	64
Quadro 5-5 - Dados dos poços	65
Quadro 5-6 - Quantidade de amostras corrigidas para cada poço e bomba.....	65
Quadro 5-7 - Resultado das classificações	66
Quadro 5-8 - Matriz de confusão para o <i>kernel</i> linear.....	67
Quadro 5-9 - Matriz de confusão para o <i>kernel</i> rbf.....	67
Quadro 5-10 - Matriz de confusão para o <i>kernel</i> polinomial	67

Lista de tabelas

Tabela 2-1 – Lista de artigos que compoem o trabalho	16
Tabela 4-1 - Técnica PICOS para elaboração da pergunta de pesquisa	42
Tabela 4-2 - Desempenho dos modelos	50
Tabela 4-3 - Valores de RMSE para o experimento (ii)	51
Tabela 4-4 - Desempenho das ferramentas de classificação	52

Lista de siglas

ANM – Árvore de Natal Molhada

BAP – Base Adaptadora de Produção

boe – Barris de Óleo Equivalente

EPE – Empresa de Pesquisa Energética

IEA – *International Energy Agency*

MCV-A – Módulo de Conexão Vertical de Anular

MCV-P – Módulo de Conexão Vertical de Produção

MCV-U – Módulo de Conexão Vertical de Umbilical

MEC – Ministério da Educação

SCM – *Subsea Control Module*

TCAP – *Tree Cap*

TH – *Tubing Hanger*

SUMÁRIO

1 Apresentação	12
2 Estrutura	16
3 Artigo A - Inteligência Artificial para Detecção de Falhas em Equipamentos: Um Estudo Bibliométrico no Setor de Óleo e Gás	17
3.1 Resumo	17
3.2 Abstract.....	17
3.3 Introdução	18
3.4 Metodologia	18
3.5 Resultados	20
3.5.1 Pesquisa na amostra com o uso das palavras-chave e tesouros	20
3.5.2 Identificações dos veículos com maior número de artigos publicados sobre o tema	21
3.5.3 Levantamento da cronologia da produção	24
3.5.4 Seleção dos artigos com o maior número de citações	24
3.5.5 Identificação dos autores com maior número de publicações	29
3.6 Considerações Finais	30
3.7 Apêndice	31
3.8 Referências.....	36
4 Artigo B - Utilização de Técnicas de MACHINE LEARNING na Previsão de Falhas Em Equipamentos: Uma Revisão Sistemática	39
4.1 Resumo	39
4.2 <i>Abstract</i>	39
4.3 Introdução	40
4.4 Metodologia	41
4.4.1 Critérios de elegibilidade	41
4.4.2 Fontes de informação	42
4.4.3 Estratégia de busca	42

4.4.4 Seleção dos estudos	43
4.5 Resultados	44
4.5.1 Técnicas de Classificação	46
4.5.1.1 Equipamentos de manufatura	46
4.5.1.2 Equipamentos de óleo em gás	47
4.5.1.3 Outras aplicações.....	48
4.5.2 Técnicas de regressão	48
4.5.2.1 Equipamentos de manufatura	48
4.5.2.2 Equipamentos de óleo e gás	49
4.5.3 Híbridos	50
4.6 Discussão e Conclusão	52
4.7 Referências.....	53
5 Artigo C - Previsão de <i>slugs</i> em pipelines de produção de óleo e gás - Um modelo utilizando SVM	57
5.1 Resumo	58
5.2 <i>Abstract</i>	58
5.3 Introdução	59
5.4 Metodologia	62
5.4.1 Coleta de dados	62
5.4.2 Tratamento de dados	63
5.4.3 Aplicação da ferramenta de Inteligência Artificial	66
5.5 Resultados e Discussões.....	66
5.6 Conclusão.....	68
5.7 Referências.....	68
Referências	74

1 APRESENTAÇÃO

A indústria de Óleo & Gás é uma das maiores, mais complexas e mais importantes indústrias a nível global, influenciando a vida de todos as pessoas com diversos produtos como combustíveis, asfalto e outros diversos produtos de tapetes a óculos a roupas (INKPEN; MOFFETT, 2011). Segundo IEA (2018), em 2016, 54,0% de toda a energia primária fornecida globalmente era proveniente do petróleo (32%) ou do gás natural (22%), conforme Figura 1-1. No Brasil, em 2017, essa porcentagem foi de 49,4%, sendo 36,4% proveniente do petróleo e seus derivados e 13% do gás natural (EPE, 2018), conforme Figura 1-2.

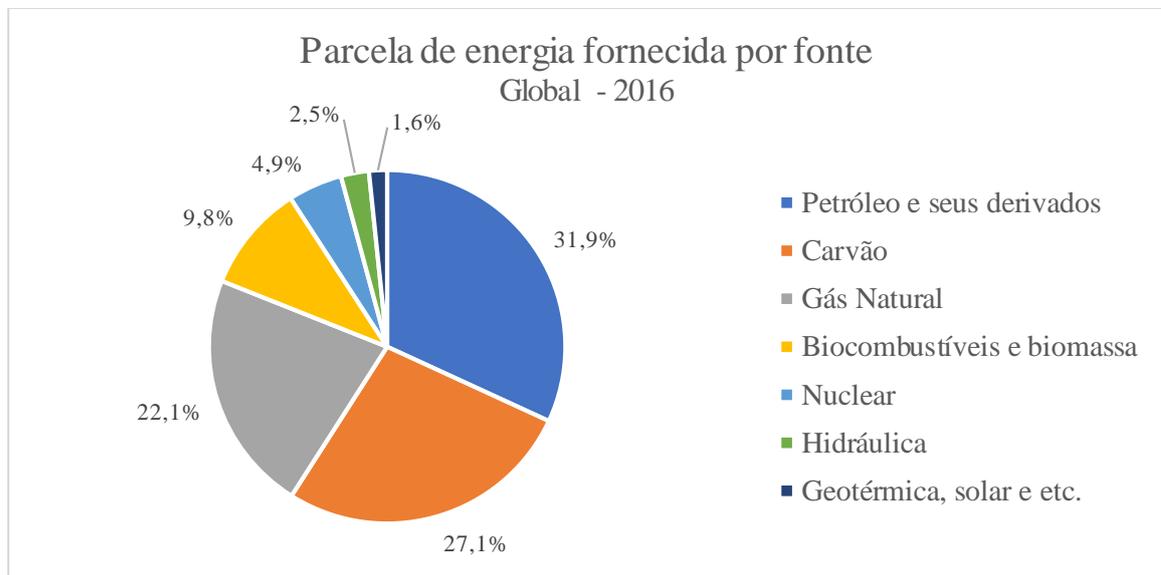


Figura 1-1 – Parcela de energia global por fonte.
Fonte: Adaptado de IEA (2018).

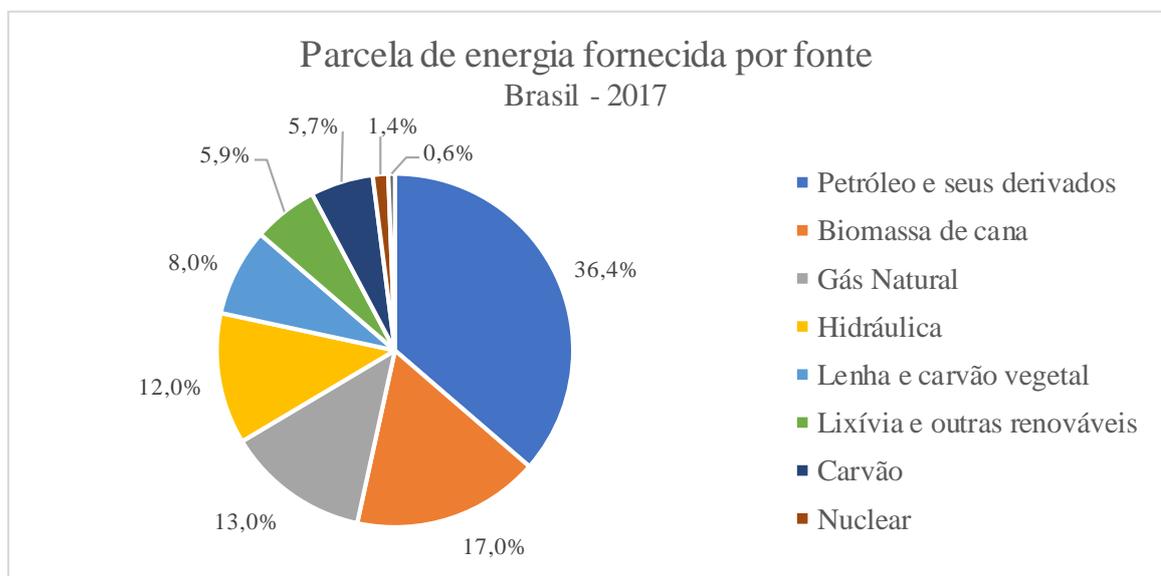


Figura 1-2 – Parcela de energia por fonte no Brasil.
Fonte: Adaptado de EPE (2018).

Na economia o panorama não é diferente, em 2015, a indústria de óleo e gás gerou um valor de 1,3 trilhões de dólares nos Estados Unidos, o que representa mais de 7% do PIB deste mesmo ano do país (API, 2017). No Brasil, a relevância é ainda maior: em 2014, este setor representou 13% do PIB e um aumento de dez pontos percentuais se comparado com a parcela do ano 2000 (BRASIL, 2014).

Um dos maiores responsáveis pelo crescimento da relevância da indústria de óleo e gás na composição do PIB foi a descoberta das reservas do Pré-Sal em 2006, uma nova realidade geológica para o Brasil e para a indústria petrolífera mundial (SCHUTTE, 2012). Estes reservatórios receberam esse nome por estarem situados sob a extensa camada de sal que se estende na região costa-afora entre os estados do Espírito Santo e Santa Catarina, numa faixa com cerca de 800 km de comprimento por 200 km de largura (RICCOMINI; SANT'ANNA; TASSINARI, 2012). A capacidade de produção dos poços do Pré-Sal é tamanha que, em abril de 2019, estes poços produziram 1,98 MMboe/d. Esse valor representa 59,7% do total de óleo produzido em todo território brasileiro nesse mesmo mês, apesar de representar apenas 12,4% do total de poços submarinos produtores (ANP, 2019b). Na Figura 1-3 é possível ver o histórico de produção entre os poços do Pré-Sal, Pós-Sal e os poços terrestres.

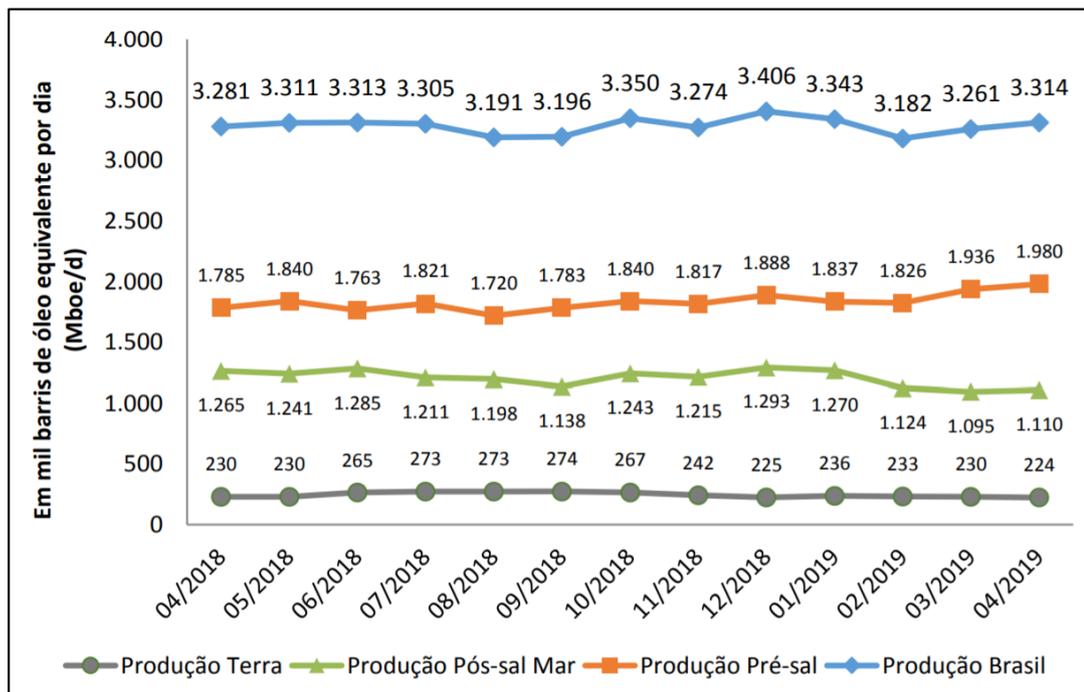


Figura 1-3 – Histórico de produção de petróleo no Brasil por origem.
Fonte: (ANP, 2019b).

Junto com a relevância dos poços Pré-Sal, cresce também a criticidade de trabalho e operação dos equipamentos. Cada poço é completado com um conjunto de equipamentos submarinos composto por até 8 equipamentos (BAP, TH, ANM, TCAP, MCV-P, MCV-A,

MCV-U e SCM) que devem suportar a pressão hidrostática de profundidades que chegam a quase 3000m (ANP, 2019a) e apresentar alta confiabilidade, uma vez que a intervenção nestes equipamentos envolvem embarcações com custos diretos diários que podem variar de 50 a 500 mil dólares e custos indiretos como custos de oportunidade e lucro cessante de um poço fechado (SILVA; HORA; FERNANDES, 2017). Além destes equipamentos, outros compõem o arranjo submarino, como ESPs, *manifolds* e linhas de fluxo para o escoamento da produção e injeção de fluidos nos poços.

Neste contexto, saber dizer qual componente irá falhar e quando se torna fundamental para evitar intervenções desnecessárias. Com estas premissas em mente, a manutenção preditiva, definida por Scheffer e Girdhar (2004) como uma manutenção realizada apenas quando é necessário e com poder de aumentar a capacidade produtiva da máquina, se encaixa perfeitamente por eliminar os tempos de parada, aumentar eficiência e segurança e minimizar o número de manutenções (ATS, 2018).

Visando melhorar a detecção e reconhecimento de padrões de falha, Liu *et al.* (2018a) propõem a utilização de ferramentas de inteligência artificial (IA). Esta área de monitoramento do estado das máquinas é foco de um interesse crescente, tanto da comunidade científica quanto das indústrias, assim como a utilização de IA como ferramenta de previsão de falhas (SUGUMARAN; MURALIDHARAN; RAMACHANDRAN, 2007; VELLOSO; HORA; CARVALHO, 2018).

Mohammadzaheri *et al.* (2016) aplicam ferramentas de IA (redes neurais artificiais e algoritmos evolutivos) para maximizar o lucro de campos submarinos de petróleo com elevação artificial por ESP (abreviação para bombas elétricas submersíveis, em inglês). A aplicação destes algoritmos se mostrou bem sucedida, sendo possível encontrar valores otimizados para as variáveis operacionais, como taxa de fluxo e velocidade de rotação do motor da bomba, para um sistema submarino composto por vários poços com sistemas de elevação por ESPs.

A técnica de rede neurais artificiais também foi implementada por El-Abbasy *et al.* (2014), porém sua aplicação foi voltada para a prever a condição de tubulações *offshore* de petróleo e gás baseada em dados históricos de inspeções coletados em três destas tubulações no Catar. Os modelos encontrados foram capazes de avaliar e prever a condição dos dutos com uma acurácia superior a 97%.

Já em Silva, da Hora e Fernandes (2017), foi realizado um estudo sobre a confiabilidade de sistemas de controle submarinos aplicados a produção de óleo e gás a fim de identificar os pontos críticos e modos de falha mais prováveis de ocorrerem durante a fase de operação. Foram

utilizados os dados de confiabilidade do OREDA e aplicadas técnicas de diagrama de blocos de confiabilidade e árvore de análise de falha. O estudo mostrou que os equipamentos de superfície são os que tem maior impacto da não confiabilidade do sistema, seguido pelo bloco de válvulas direcionais que atuam os equipamentos.

Uma vez consolidado o arcabouço científico sobre as aplicações de ferramentas de IA na predição de falhas, verifica-se que apesar de os artigos abordarem diversas aplicações práticas, há uma lacuna na predição de falhas de equipamentos submarinos, em especial na predição de falhas de ESPs. Uma vez que um dos modos de falha de ESPs é o surgimento de bolsões de gás conhecidos como *slugs*, este trabalho vem dar sua contribuição com o objetivo geral de aplicar a técnica de *Support Vector Machine* (SVM) para a criação de um modelo a fim de prever *slugs* em uma sistema submarino através de dados de campo.

O objetivo supracitado pode ser subdividido em três objetivos específicos conforme Quadro 1-1, sendo todos eles voltados para os assuntos objetos deste trabalho: Inteligência artificial, Análise de falha, e Óleo e gás.

Quadro 1-1 – Subdivisão dos objetivos do trabalho

Nº do objetivo	Objetivos específicos
I	Identificar as tendências históricas da produção científica, períodos de maior produção e artigos mais relevantes
II	Identificar conceitos importantes, realizar análises confrontando as produções da literatura, concluir sobre o que os estudos apresentam e apontar os problemas/questões que necessitam de novos estudos
III	Realizar um estudo de caso com dados históricos de surgimento de <i>slugs</i> em uma ESP e gerar um modelo de predição para <i>slugs</i> .

Fonte: Dos autores (2018).

2 ESTRUTURA

Este trabalho é composto por três artigos, conforme Tabela 2-1. Cada um dos artigos é identificado por um código e irá atender um dos objetivos listados no Quadro 1-1. A partir deste ponto do trabalho, cada artigo será referenciado apenas utilizando seu código de identificação.

Tabela 2-1 – Lista de artigos que compoemo trabalho

Tipo de artigo	Código de identificação
Bibliometria	A
Revisão sistemática	B
Estudo de caso	C

Fonte: Dos autores (2018).

O artigo A é uma bibliometria realizada na base de dados Scopus® a fim de atender ao objetivo I. Foram procuradas produções que citassem os termos de busca Inteligência artificial, Análise de falha, Óleo e gás e seus *thesaurus*. Os artigos selecionados serviram como base para o desenvolvimento das partes seguintes do trabalho. Este artigo foi publicado no XV Simpósio de Excelência em Gestão em 2018, portanto foi realizada uma atualização buscando realizar as avaliações bibliométricas em trabalhos mais atualizados.

O artigo B é uma revisão sistemática dos principais artigos encontrados no artigo A. A revisão sistemática é capaz de atender ao objetivo II, como descrito em Sampaio e Mancini (2007). Este artigo apresenta as melhores práticas e as ferramentas de IA mais indicadas para o diagnóstico de falhas de acordo com a comunidade científica.

O artigo C é um estudo de caso a fim de atender aos objetivo III. Nele, é utilizada a ferramenta SVM. Uma vez definida a técnica de IA a ser utilizada, o *dataset* foi tratado, separado na proporção 70/30, sendo a maior parte para treino e a menor para teste, e submetido à geração do modelo utilizando SVM. Este método de validação foi escolhido por demandar um menor custo computacional dentre os outros métodos. O modelo obtido neste artigo foi capaz de prever o acontecimento de *slugs*, objeto deste trabalho.

3 ARTIGO A - INTELIGÊNCIA ARTIFICIAL PARA DETECÇÃO DE FALHAS EM EQUIPAMENTOS: UM ESTUDO BIBLIOMÉTRICO NO SETOR DE ÓLEO E GÁS

3.1 Resumo

Contexto: O setor de óleo e gás gera um grande volume de dados das suas operações e maquinário que necessita ser analisado com o objetivo de estratificar informações para agregar valor à gestão da operação e manutenção. Os métodos de inteligência artificial são ferramentas poderosas para realizar esta análise dada sua capacidade de reconhecimento de padrões.

Objetivo: O objetivo deste trabalho é realizar uma bibliometria da área de métodos de inteligência artificial para análise de falhas em equipamentos do setor de óleo e gás.

Metodologia: É realizado um estudo bibliométrico na base de dados Scopus utilizando conceitos chave a fim de analisar as publicações na área.

Resultados: A partir das pesquisas realizadas utilizando os conceitos chaves e seus tesouros, são analisados os principais autores e periódicos do tema, além de uma análise cronológica de produção.

Conclusões: Foram encontrados poucos trabalhos que englobem todos os conceitos estabelecidos, o que demonstra que esta é uma área ainda a ser explorada mais a fundo.

Palavras-chave: Inteligência artificial; Análise de falha; Óleo e gás.

3.2 Abstract

Context: *The oil and gas industry generate a great amount of operation and machinery data that needs to be analyzed to get valuable information to operational and maintenance management. The artificial intelligence methods are powerful tools to make this analysis given its capability to recognize patterns.*

Objective: *This research objective is to make a bibliometric study with the published works in artificial intelligence methods for failure analysis in oil and gas equipment area.*

Methodology: *Make a bibliometric study in Scopus database using key concepts to analyze the published works in this area.*

Results: *From the made researches using the key concepts and their thesaurus, the theme main authors and journals as well as a publishing chronology.*

Conclusions: *Just a few works were found covering all the established concepts, which shows that this is an area yet to be deeply explored.*

Keywords: Artificial intelligence; Failure analysis; Oil and gas.

3.3 Introdução

Nos diversos segmentos da indústria é comum encontrar diversas máquinas e equipamentos dotados de sensores capazes de gerar informações sobre o funcionamento daquele item. Para Rengaswamy e Venkatasubramanian (1995), os operadores se deparam com uma multidão de dados do processo, porém pode ser difícil diferenciar o comportamento normal do anormal apenas olhando os dados brutos dos sensores.

Levando em consideração o aumento da capacidade produtiva de cada uma das máquinas, para Nandi e Toliyat (1999), se tornou muito importante diagnosticar falhas assim que elas ocorrem, uma vez que paradas de máquina não agendadas podem atrasar deadlines e causar grande perdas financeiras. Já para Chen e Patton (2012), existe um aumento na demanda de para sistemas dinâmicos se tornarem mais seguros e confiáveis. Para solucionar este problema, Liu *et al.* (2018) propõem a utilização de métodos de inteligência artificial como ferramenta para reconhecimento de padrões.

Segundo Nandi e Toliyat (1999), sistemas inteligentes que utilizam metodologias de inteligência artificial como redes neurais e lógica *fuzzy* estão substituindo o envolvimento humano para tomadas decisão em detecção de falhas, uma vez que os humanos apresentam uma resposta muito mais lenta quando comparados com esses sistemas. A detecção e diagnóstico de falhas com base em dados são amplamente aceitas na prática da indústria graças às características baseadas em dados e multivariáveis (QIN, 2012).

Neuroth, MacConnell, Stronach e Vamplew (2000) definem que a situação da utilização dessas metodologias no setor de óleo e gás é surpreendente. Mesmo que, em diversas regiões do mundo, esta indústria tenha chegado ao mais alto nível de maturidade com experiência em exploração, produção e processamento, o potencial do uso da inteligência artificial para organizar e processar a informação mais eficientemente ainda não foi completamente percebido.

O objetivo desta pesquisa é realizar uma bibliometria da área de detecção e diagnóstico de falhas utilizando inteligência artificial. Busca-se definir os veículos e autores mais relevantes para o assunto, assim como verificar o avanço na produção de trabalhos ao longo dos anos e tendência para os anos futuros.

3.4 Metodologia

Com o objetivo de identificar os trabalhos que envolvam a análise de falhas utilizando inteligência artificial no setor de óleo e gás é realizado um estudo bibliométrico. Para Hood &

Wilson (2001), a bibliometria é o estudo das dinâmicas das disciplinas como refletidos na produção da literatura.

O modelo desta pesquisa é baseado no proposto por Costa (2010), com as seguintes etapas:

1. Definição da amostra da pesquisa;
2. Pesquisa na amostra, com as palavras-chave;
3. Identificações dos veículos com maior número de artigos publicados sobre o tema;
4. Levantamento da cronologia da produção;
5. Seleção dos artigos com o maior número de citações;
6. Identificação dos autores com maior número de publicações.

Na definição de amostra (etapa 1), serão utilizados os artigos indexados na base de conhecimento Scopus®. Esta escolha é justificada pelo Scopus® ser o maior banco de dados de resumos e citações de literaturas revisadas, como periódicos científicos e livros. O conjunto de trabalhos utilizado é composto por todas as publicações na base até o ano de 2017.

As palavras-chaves escolhidas para a etapa de pesquisa na amostra (etapa 2) são: *artificial intelligence*, *failure analysis* e *oil and gas*. A partir destas palavras-chaves, foram definidos os tesouros a serem utilizados. A partir destes termos, foram extraídos os tesouros e outros termos utilizados na literatura, conforme Figura 3-1.

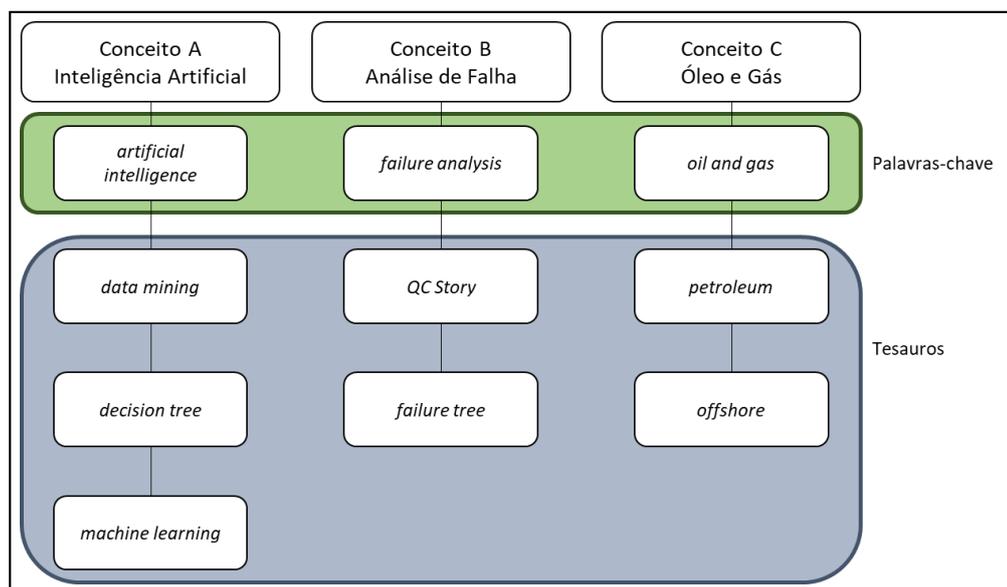


Figura 3-1 – Palavras-chave, tesouros e termos correspondentes.
Fonte: Dos autores (2018).

Conforme apresentado na Figura 3-1, os conceitos, palavras-chave e tesouros receberam as identificações A, B e C. Essa identificação continuará a ser utilizada até o final deste trabalho.

Uma vez definidos todos os termos, foi formulada e realizada a pesquisa no Scopus®. A pesquisa possui um corte temporal que exclui trabalhos publicados em 2018, uma vez que o ano não terminou na data de finalização deste trabalho. A consulta realizada está representada na Figura 3-2.

(TITLE-ABS-KEY	
("data mining" OR "decision tree" OR "artificial intelligence" OR "machine learning")	# Tesouros de A
AND TITLE-ABS-KEY	
("QC Stor*" OR "failure*tree" OR "failure* analysis")	# Tesouros de B
AND TITLE-ABS-KEY	
("oil and gas" OR "petroleum" OR "offshore")	# Tesouros de C
) AND (EXCLUDE	
(PUBYEAR, 2018)	# Corte temporal
)	

Figura 3-2 – Pesquisa realizada no Scopus®.
Fonte: Dos autores (2018).

Além da consulta supracitada, foram feitas outras 3 consultas na base de dados Scopus® combinando os conceitos dois a dois (AB, AC e BC).

As análises deste trabalho irão se concentrar nas pesquisas que envolvem o conceito Inteligência Artificial (AB e AC), uma vez que este é o método, sendo imprescindível para a pesquisa.

Para as etapas de 3 a 6, serão levantadas as informações de veículos com maior número de publicações sobre o tema, autores com maior de publicações, cronologia da produção e publicações com maior número de citações conforme resultado da pesquisa na amostra. Os resultados serão apresentados de forma gráfica e em tabelas.

Para as etapas 3, 4 e 6, serão selecionados os veículos, autores e artigos utilizando o corte de Pareto. Para a etapa 6, serão considerados apenas os artigos que possuem 3 ou mais citações.

3.5 Resultados

3.5.1 PESQUISA NA AMOSTRA COM O USO DAS PALAVRAS-CHAVE E TESAUROS

Este trabalho compõe uma pesquisa sobre métodos de inteligência artificial para análise de falha em equipamentos do setor de óleo e gás. Para realizar a pesquisa bibliométrica, foram utilizados 3 conceitos que definem o objetivo da pesquisa, o método e o corte setorial: inteligência artificial (método), análise de falha (objetivo da pesquisa) e óleo e gás (corte setorial).

Utilizando todas as consultas realizadas (AB, AC, BC e ABC), é possível montar o diagrama de Venn ilustrado na Figura 3-3 para representar os resultados encontrados.

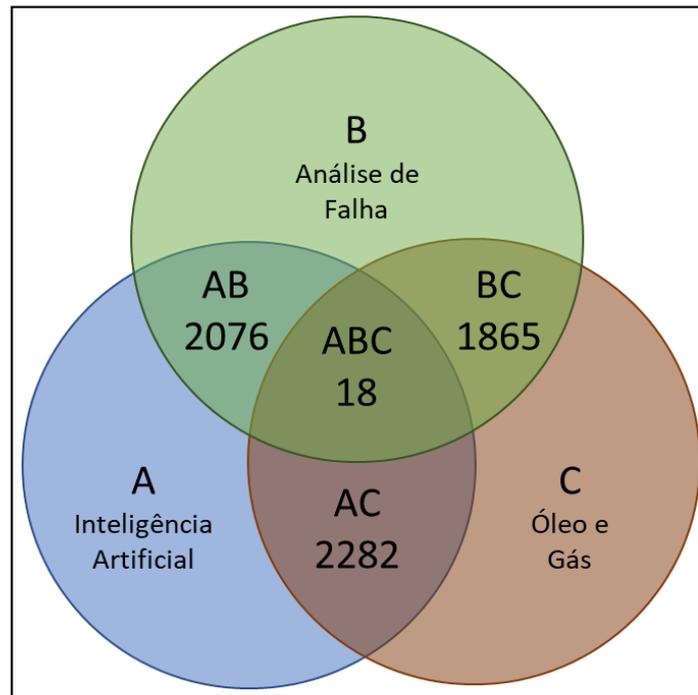


Figura 3-3 – Diagrama de Venn das pesquisas realizadas.
Fonte: Dos autores (2018).

Através do diagrama de Venn é possível perceber que, embora a pesquisa envolvendo os três conceitos tenha retornado 18 resultados, este é considerado um número baixo se comparado às outras pesquisas que utilizaram apenas 2 conceitos.

3.5.2 IDENTIFICAÇÕES DOS VEÍCULOS COM MAIOR NÚMERO DE ARTIGOS PUBLICADOS SOBRE O TEMA

Nesta seção, serão listados os veículos com mais publicações em cada uma das pesquisas analisadas. Para a pesquisa com os conceitos AB, foram encontrados 124 veículos com publicações sobre o tema. Já para a pesquisa com os conceitos AC, 86 veículos foram encontrados. Os diagramas de Pareto dispostos abaixo na Figura 3-4 e na Figura 3-5 representam os resultados das pesquisas com os conceitos AB e com os conceitos AC, respectivamente.

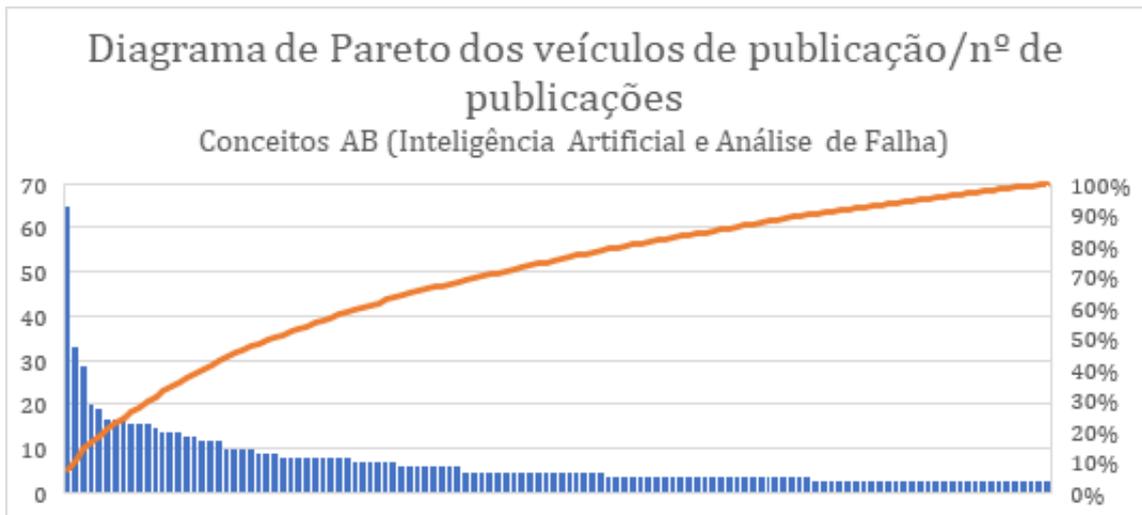


Figura 3-4 – Diagrama de Pareto dos veículos como maior número de publicações para os conceitos AB.
Fonte: Dos autores (2018).

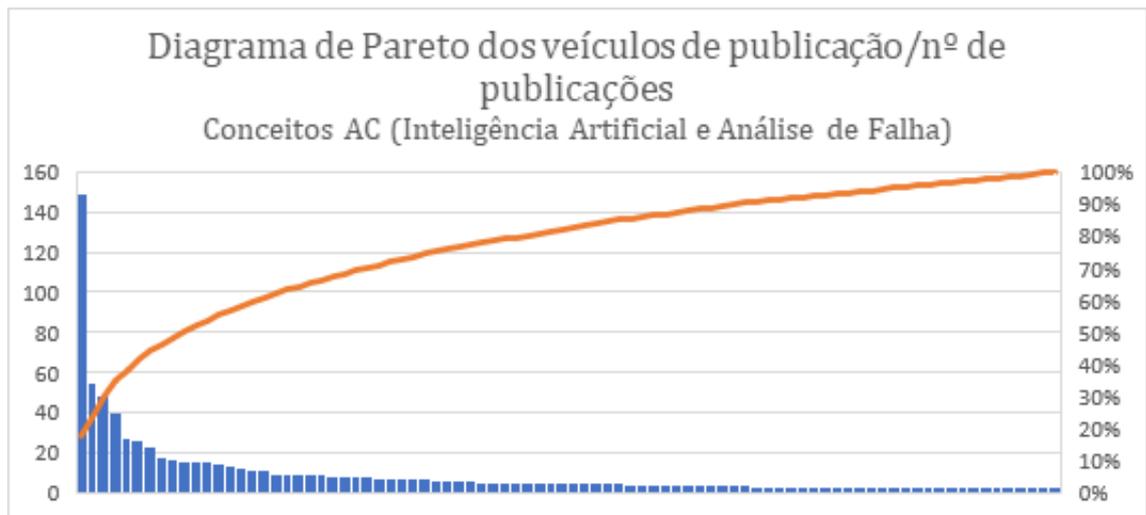


Figura 3-5 – Diagrama de Pareto dos veículos como maior número de publicações para os conceitos AC.
Fonte: Dos autores (2018).

O Quadro 3-1 mostra os veículos com maior produção nas pesquisas com os conceitos AB e AC.

Quadro 3-1 – Veículos com maior número de publicações sobre o tema

	Conceitos AB (<i>Artificial Intelligence e Failure Analysis</i>)	Conceitos AC (<i>Artificial Intelligence e Oil and Gas</i>)
Applied Mechanics And Materials / ISSN: 1662-7482	28	-
Arabian Journal For Science And Engineering / ISSN: 2193-567X, 2191-4281	-	40
Autotestcon Proceedings / ISSN: 0734-7510	20	-
Bioinspiration And Biomimetics / ISSN: 1748-3182	16	-

Conference Proceedings Annual International Conference Of The IEEE Engineering In Medicine And Biology Society IEEE Engineering In Medicine And Biology Society Conference / ISSN: 1557-170X	16	-
Expert Systems With Applications / ISSN: 0957-4174	-	17
IEEE Transactions On Biomedical Engineering / ISSN: 0018-9294	17	-
IEEE Transactions On Neural Networks / ISSN: 2162-237X	17	-
IEEE Transactions On Neural Systems And Rehabilitation Engineering / ISSN: 1534-4320	17	-
IEEE Transactions On Systems Man And Cybernetics Part B Cybernetics / ISSN: 1083-4419	33	-
Journal Of Natural Gas Science And Engineering / ISSN: 1875-5100	-	27
Journal Of Petroleum Science And Engineering / ISSN: 0920-4105	-	48
Lecture Notes In Artificial Intelligence Subseries Of Lecture Notes In Computer Science	15	-
Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics / ISSN: 0302-9743	65	55
Neurocomputing / ISSN: 0925-2312	19	-
Oil And Gas Journal / ISSN: 0030-1388	-	18
Proceedings Of SPIE The International Society For Optical Engineering	16	-
Proceedings Of The Annual Offshore Technology Conference / ISSN: 0160-3663	-	23
Proceedings Of The International Offshore And Polar Engineering Conference	-	26
Proceedings SPE Annual Technical Conference And Exhibition	-	149

Fonte: Dos autores (2018).

No total, o Quadro 3-1 apresenta 21 veículos, sendo que apenas o periódico de ISSN 0302-9743 aparece em ambas as pesquisas. Além disto, o periódico é o primeiro em número de publicações na pesquisa com os conceitos AB e segundo na pesquisa com os conceitos AC, fato que demonstra que este é um periódico a ser acompanhado por ter uma maior atuação na área pesquisada.

3.5.3 LEVANTAMENTO DA CRONOLOGIA DA PRODUÇÃO

Para o levantamento cronológico relevante ao tema, foram consideradas as publicações que constam na base de conhecimento Scopus® contendo os conceitos AB e AC. A representação gráfica deste levantamento consta na Figura 3-6.

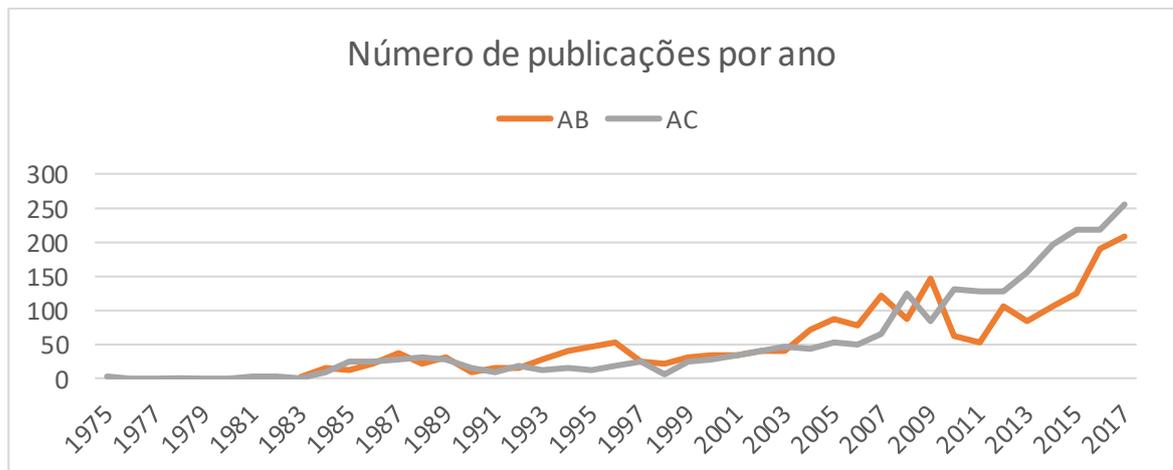


Figura 3-6 – Levantamento cronológico das publicações.
Fonte: Dos autores (2018).

Com os dados apresentados na Figura 3-6, pode-se observar que ambos os assuntos estão quase que em uma crescente constante desde o início da década de 2010. Nota-se também que os temas abordando Inteligência Artificial e Análise de Falha e abordando Inteligência Artificial e Óleo e Gás passaram por um ciclo de maior pesquisa entre os anos de 2007 a 2009.

3.5.4 SELEÇÃO DOS ARTIGOS COM O MAIOR NÚMERO DE CITAÇÕES

Nesta última etapa, foram selecionados os artigos mais citados. Estes artigos servirão de possíveis referências para pesquisas futuras. Abaixo, estão representados os diagramas de Pareto da pesquisa com os conceitos AB na Figura 3-7 e com os conceitos AC na Figura 3-8.

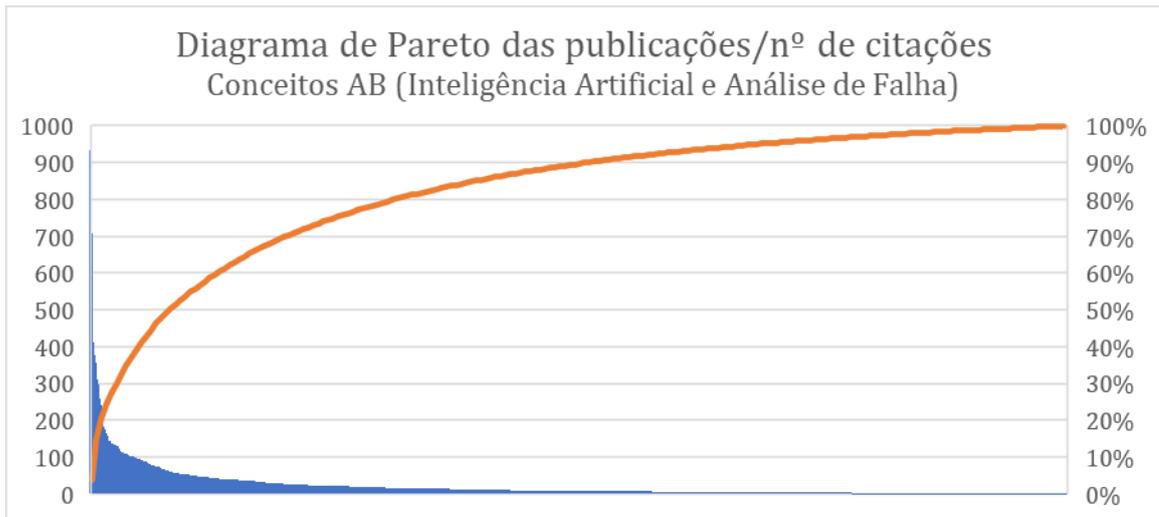


Figura 3-7 – Diagrama de Pareto dos artigos como maior número de citações para os conceitos AB.
Fonte: Dos autores (2018).

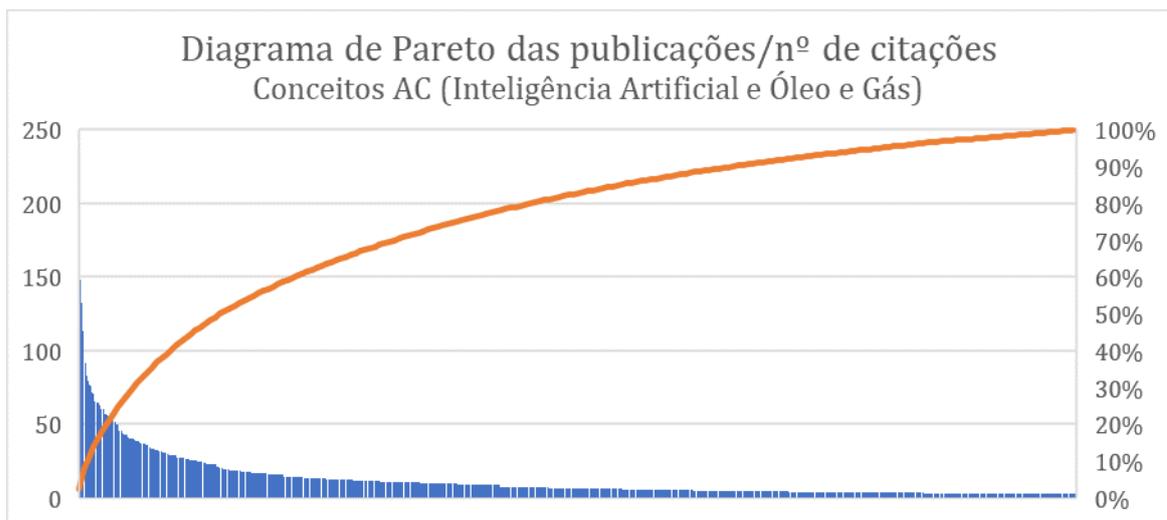


Figura 3-8 – Diagrama de Pareto dos artigos como maior número de citações para os conceitos AC.
Fonte: Dos autores (2018).

Quadro 3-2 – Artigos mais citados por pesquisa

Título da publicação	Autor(es)	Título do periódico ou conferência	Número de Citações
Conceitos AB (Inteligência Artificial e Análise de Falha)			
Survey of robust residual generation and evaluation methods in observer-based fault detection systems	Frank, P.M., Ding, X.	Journal of Process Control/ ISSN: 0959-1524	934

From computing with numbers to computing with words - From manipulation of measurements to manipulation of perceptions	Zadeh, L.A.	IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications / ISSN: 1057-7122	772
Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring	Karantonis, D.M., Narayanan, M.R., Mathie, M., Lovell, N.H., Celler, B.G.	IEEE Transactions on Information Technology in Biomedicine / ISSN: 1089-7771	710
Activity classification using realistic data from wearable sensors	Pärkkä, J., Ermes, M., Korpipää, P., Mäntyjärvi, J., Peltola, J., Korhonen, I.	IEEE Transactions on Information Technology in Biomedicine / ISSN: 1089-7771	413
Survey on data-driven industrial process monitoring and diagnosis	Qin, S.J.	Annual Reviews in Control / ISSN: 1367-5788	379
Condition monitoring and fault diagnosis of electrical machines - a review	Nandi, Subhasis, Toliyat, Hamid A.	Conference Record - IAS Annual Meeting (IEEE Industry Applications Society) / ISSN: 0160-8592, 0197-2618	357
Bankruptcy prediction using neural networks	Wilson, R.L., Sharda, R.	Decision Support Systems / ISSN: 13675788	345
Using health care Failure Mode and Effect Analysis: the VA National Center for Patient Safety's prospective risk analysis system.	DeRosier, J., Stalhandske, E., Bagian, J.P., Nudell, T.	The Joint Commission journal on quality improvement / ISSN: 1070-3241	311
Development of a biomimetic robotic fish and its control algorithm	Yu, J., Tan, M., Wang, S., Chen, E.	IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics / ISSN: 1083-4419	297
The use of design descriptions in automated diagnosis	Genesereth, M.R.	Artificial Intelligence / ISSN: 0004-3702	261
Control of an electrical prosthesis with an SSVEP-based BCI	Müller-Putz, G.R., Pfurtscheller, G.	IEEE Transactions on Biomedical Engineering / ISSN: 0018-9294	257
Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS	Kourti, T., Nomikos, P., MacGregor, J.F.	Journal of Process Control / ISSN: 0959-1524	242
The prediction and diagnosis of wind turbine faults	Kusiak, A., Li, W.	Renewable Energy / ISSN: 0960-1481	194
Takagi-Sugeno fuzzy-model-based fault detection for networked control systems with Markov delays	Zheng, Y., Fang, H., Wang, H.O.	IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics / ISSN: 1083-4419	184
Design of a high-resolution optoelectronic retinal prosthesis	Palanker, D., Vankov, A., Huie, P., Baccus, S.	Journal of Neural Engineering / ISSN: 1741-2560	176

Conceitos AC (Inteligência Artificial e Óleo e Gás)			
A neural network methodology for process fault diagnosis	Venkatasubramanian, V., Chan, K.	AIChE Journal / ISSN: 1547-5905	235
Solar Energy Engineering	Kalogirou, S.	Solar Energy Engineering / ISBN: 978-0-12397-256-9, 978-0-12397-270-5	193
Fuel cell high-power applications	Thounthong, P., Davat, B., Raël, S., Sethakul, P.	IEEE Industrial Electronics Magazine / ISSN: 1932-4529	149
Support vector machine regression (SVR/LS-SVM) - An alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data	Balabin, R.M., Lomakina, E.I.	Analyst / ISSN: 0003-2654	133
A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting	Cheng, C.-H., Chen, T.-L., Wei, L.-Y.	Information Sciences / ISSN: 0020-0255	114
Practical application of fuzzy logic and neural networks to fractured reservoir characterization	Ouenes, A.	Computers and Geosciences / ISSN: 0098-3004	92
Biodiesel classification by base stock type (vegetable oil) using near infrared spectroscopy data	Balabin, R.M., Safieva, R.Z.	Analytica Chimica Acta / ISSN: 0003-2670	86
First harmful dinophysis (dinophyceae, dinophysiales) bloom in the U.S. is revealed by automated imaging flow cytometry	Campbell, L., Olson, R.J., Sosik, H.M., Abraham, A., Henrichs, D.W., Hyatt, C.J., Buskey, E.J.	Journal of Phycology / ISSN: 1529-8817	83
Internet-scale collection of human-reviewed data	Su, Q., Pavlov, D., Chow, J.-H., Baker, W.C.	16th International World Wide Web Conference, WWW2007 / ISBN: 978-1-59593-654-7	80
Decision support system for risk management: A case study	Dey, P.K.	Management Decision / ISSN: 0025-1747	77
Bridging between SeaWiFS and MODIS for continuity of chlorophyll-a concentration assessments off Southeastern China	Zhang, C., Hu, C., Shang, S., Müller-Karger, F.E., Li, Y., Dai, M., Huang, B., Ning, X., Hong, H.	Remote Sensing of Environment / ISSN: 0034-4257	76
Dynamic spatial and temporal prediction of changes in depositional processes on clastic shorelines: Toward improved subsurface uncertainty reduction and management	Ainsworth, R.B., Vakarelov, B.K., Nanson, R.A.	AAPG Bulletin / ISSN: 0149-1423	72

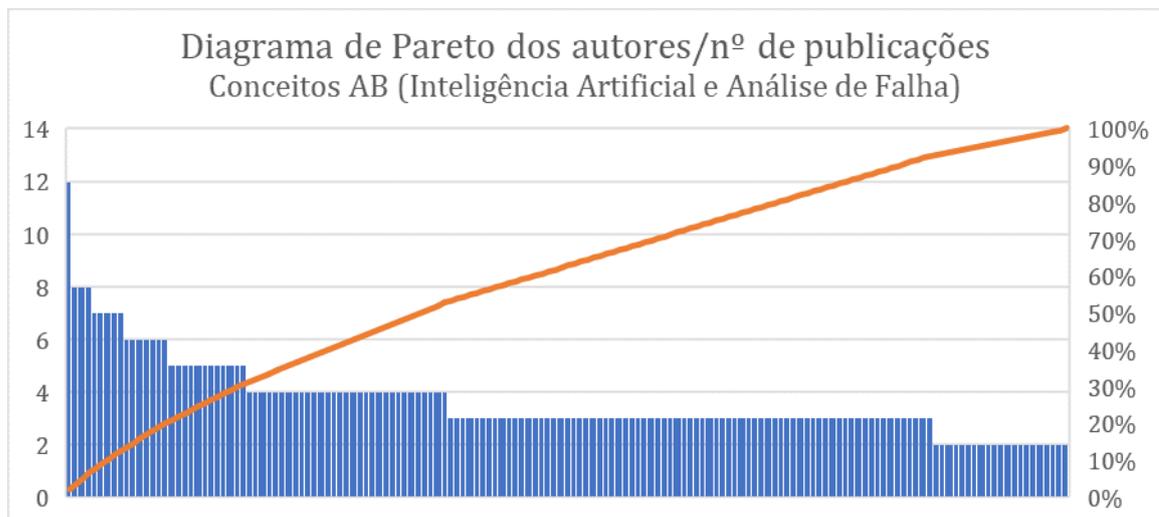
Hindcasting of wave parameters using different soft computing methods	Mahjoobi, J., Etemad-Shahidi, A., Kazeminezhad, M.H.	Applied Ocean Research / ISSN: 0141-1187	71
Implementing Radial Basis Function Networks for modeling CO ₂ reservoir oil minimum miscibility pressure	Tatar, A., Shokrollahi, A., Mesbah, M., Rashid, S., Arabloo, M., Bahadori, A.	Journal of Natural Gas Science and Engineering / ISSN: 1875-5100	66
Laser scanning and digital outcrop geology in the petroleum industry: A review	Hodgetts, D.	Marine and Petroleum Geology / ISSN: 0264-8172	66
Comparison of stochastic sampling algorithms for uncertainty quantification	Mohamed, L., Christie, M., Demyanov, V.	SPE Journal / ISSN: 1086-055X, 1930-0220	65
Using genetic algorithms for reservoir characterization	Romero, C.E, Carter, J.N	Journal of Petroleum Science and Engineering / ISSN: 0920-4105	65
Forecasting PVT properties of crude oil systems based on support vector machines modeling scheme	El-Sebakhy, E.A.	Journal of Petroleum Science and Engineering / ISSN: 0920-4105	63
Removal of sulfur compounds from utility pipelined synthetic natural gas using modified activated carbons	Cui, H., Turn, S.Q., Reese, M.A.	Catalysis Today / ISSN: 0920-5861	62
Design a new mixed expert decision aiding system using fuzzy ELECTRE III method for vendor selection	Montazer, G.A., Saremi, H.Q., Ramezani, M.	Expert Systems with Applications / ISSN: 0957-4174	61
Optimum rehabilitation strategy of water distribution systems using the HBMO algorithm	Haddad, O.B., Adams, B.J., Mariño, M.A.	Journal of Water Supply: Research and Technology – AQUA / ISSN: 1606-9935	61
Prediction of significant wave height using regressive support vector machines	Mahjoobi, J., Adeli Mosabbeh, E.	Ocean Engineering / ISSN: 0029-8018	60
A completions guidebook to shale-play development: A review of successful approaches towards shale-play stimulation in the last two decades	Chong, K.K., Grieser, W.V., Passman, A., Tamayo, C.H., Modeland, N., Burke, B.	Society of Petroleum Engineers - Canadian Unconventional Resources and International Petroleum Conference 2010 / ISBN: 978-1-55563-312-7	57
Project risk management: A combined analytic hierarchy process and decision tree approach	Dey, P.K.	Cost Engineering (Morgantown, West Virginia) / ISSN: 0274-9696	57
Uncertainty quantification for porous media flows	Christie, M., Demyanov, V., Erbas, D.	Journal of Computational Physics / ISSN: 0021-9991	56

Fonte: Dos autores (2018).

Analisando o Quadro 3-2, percebe-se que as publicações da pesquisa AB possuem um número muito superior às publicações da pesquisa AC. Para ilustrar este fato, considera-se como comparação o artigo mais citado na pesquisa com os conceitos AC (235 citações) e o 12º artigo mais citado na pesquisa com os conceitos AB (242 citações). Observa-se então um maior interesse das pesquisas relacionadas aos conceitos AB, uma vez que, conforme mostrado na Figura 3-6, a primeira pesquisa destas áreas foi realizada em 1983, oito anos após a primeira pesquisa correlata aos conceitos AC.

3.5.5 IDENTIFICAÇÃO DOS AUTORES COM MAIOR NÚMERO DE PUBLICAÇÕES

Na quarta etapa desta pesquisa, são identificados os dez autores com o maior número de publicações. A distribuição da quantidade de artigos publicados por cada autor está disposta na Figura 3-9 para os conceitos AB e na Figura 3-10 para os conceitos AC.



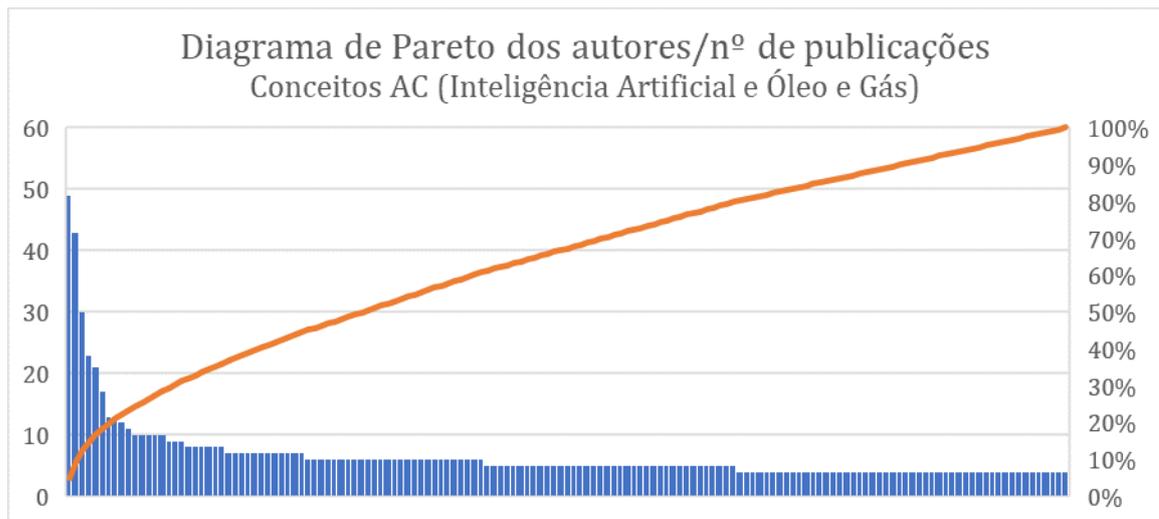


Figura 3-10 – Diagrama de Pareto dos autores com o maior número de publicações para os conceitos AC.
Fonte: Dos autores (2018).

Nota-se, nos gráficos acima, que a pesquisa com os conceitos AB possui uma distribuição mais uniforme, onde cada autor fica na mesma faixa de quantidades de publicação, enquanto que o da pesquisa AC possui uma concentração de produções em alguns nomes.

Dentre os nomes encontrados, vale ressaltar os autores Frank, P.M., Lovell, N.H., Liu, C.-C., Jung, J., Ouenes, A., Dey, P.K., J. e Zhang, C. que possuem artigos listados no Quadro 3-2 entre os mais citados de suas respectivas áreas.

3.6 Considerações Finais

Com todas as análises realizadas, pode-se perceber que o tema proposto de métodos de inteligência artificial para análise de falha em equipamentos do setor de óleo e gás ainda é uma área com poucas publicações e com grande potencial para pesquisas futuras.

As buscas com apenas dois conceitos combinados retornaram uma quantidade significativa de publicações cada, sendo a pesquisa envolvendo os conceitos Inteligência Artificial e Óleo e Gás a com mais publicações (2282 publicações). Quanto à cronologia de publicação dos conceitos Inteligência Artificial com Análise de Falha e Inteligência Artificial com Offshore, é perceptível que estes temas estão em ascensão.

Entre os artigos mais citados estão aqueles que envolvem os conceitos Inteligência Artificial e Análise de Falha. O grupo de publicações apresentado neste trabalho irá servir como base de conhecimento para pesquisas futuras.

Em relação aos veículos com maior número de publicações, temos o periódico ISSN 0302-9743 como principal periódico a ser monitorado, uma vez que, além de apresentar publicações

tanto nas pesquisas com os conceitos AB e AC, possui uma quantidade satisfatória de publicações em ambas as pesquisas.

Já entre os autores com maior número de publicações, temos Frank, P.M., Lovell, N.H., Liu, C.-C., Jung, J., Ouenes, A., Dey, P.K., J. e Zhang, C. como sendo os principais autores dos temas pesquisados. Todos os autores possuem publicações entre as mais citadas de cada área e serão monitorados para publicações futuras.

Para trabalhos futuros, propõem-se a realização da pesquisa em outras bases de dados e realizar um comparativo entre os resultados ou uma revisão sistemática dos artigos encontrados, assim como realizar outras análises.

3.7 Apêndice

Este artigo já se encontra publicado no XV Simpósio de Excelência em Gestão e Tecnologia e possui um corte temporal até 2017. A fim de manter o artigo original, mas também trazer dados atualizados, essa seção dispõe de uma nova busca pelos termos e de novas análises sobre os resultados encontrados.

Foi realizada uma nova pesquisa utilizando a sentença de busca apresentada na Figura 3-11. Esta busca retornou 37 resultados, quantidade suficiente para executar uma análise nestes trabalhos.

```
( TITLE-ABS-KEY ( "Machine Learn*" OR "Data Mining" OR "Data-mining" OR "Deep Learn*" OR "Decision Tree" OR c45 OR "Random Forest" OR "Artificial Intel*" OR "Artificial Learn*" OR "Decision tree" OR "Data mini*" OR "Data-mining" OR "datamining" OR c45 OR "random forest" ) AND TITLE-ABS-KEY ( "QC stor*" OR "QCstor*" OR "QCstor*" OR "fail* tree" OR "fail* Analysis" ) AND TITLE-ABS-KEY ( "Oil and gas" OR "Petroleum" OR "Oil & Gas" OR offshore OR off-shore OR "off shore" ) )
```

Figura 3-11 - Setença de busca no Scopus®.

Fonte: Dos autores (2020).

Para realizar a análise bibliométrica, a ferramenta *bibliometrix* foi utilizada. Esta ferramenta auxilia na execução do fluxo de trabalho de uma análise bibliométrica, além de gerar análises que mapeiam o conhecimento científico de forma compreensiva (ARIA; CUCCURULLO, 2017).

Na Figura 3-12 está disposto um gráfico de Sankey que representa o fluxo de conhecimento com base em três atributos: palavra-chave (coluna à esquerda), filiação (coluna central) e país (coluna à direita). Nele, pode-se perceber que as palavras-chave *failure analysis*, *fault detection* e *artificial intelligence* são as mais frequentes nos trabalhos encontrados e que esse

conhecimento está difundido entre várias universidades. Na parte dos países, nota-se uma predominância da China, seguida de Estados Unidos e Itália. No Brasil, a Universidade Federal do Espírito Santo é a única representante, estando ligada com os três termos principais além de *fault diagnosis system* e *decision trees*.

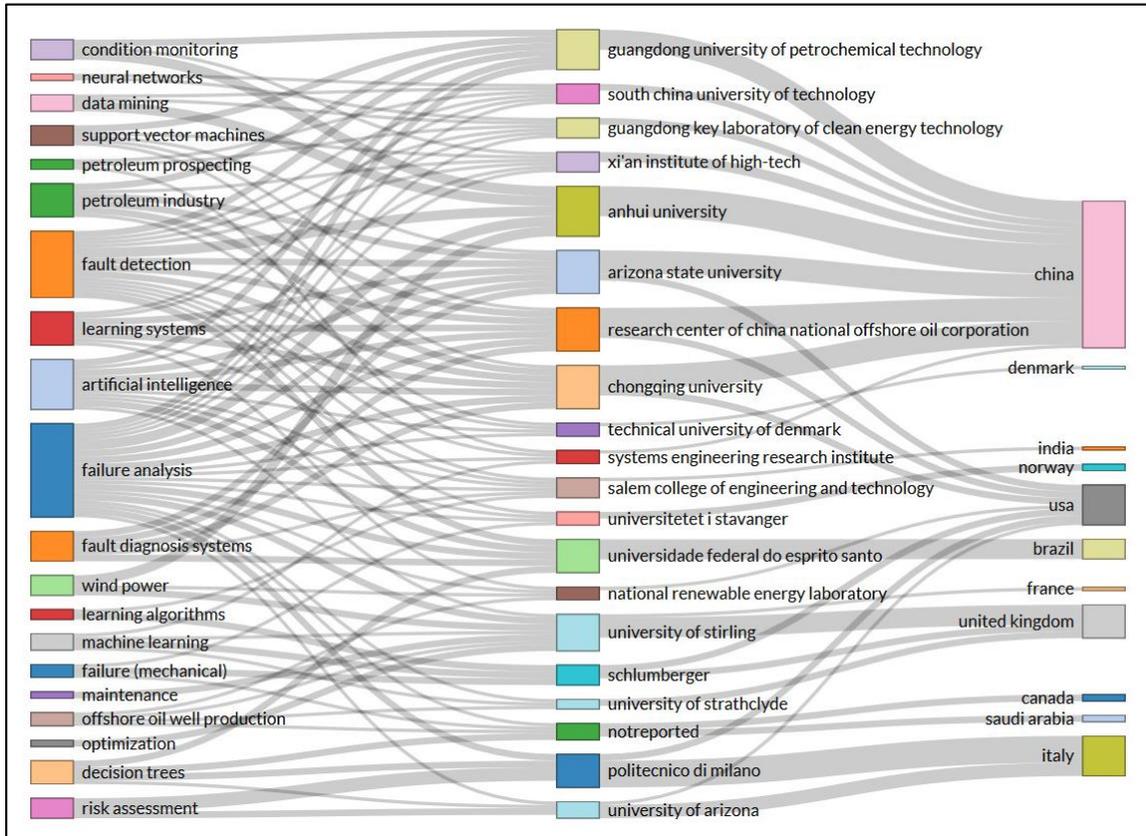


Figura 3-12 - Gráfico de Sankey do fluxo de conhecimento

Vale ressaltar ainda que dentre as filiações presentes na Figura 3-12, a China National Offshore Oil Corporation (CNOOC) e a Schlumberger são as únicas empresas a aparecer, sendo a CNOOC um grande gerador de conhecimento científico na área.

A Figura 3-13 apresenta um gráfico com as fontes que mais publicaram, dentre os artigos encontrados. Nele, é possível notar que dentre os primeiros colocados, apenas o Lecture Notes in Computer Science é um periódico, sendo então uma área que, em sua maioria, se comunica por conferências. Além disto, nenhuma das fontes, seja periódico ou conferência, desponta como principal fonte de artigos. Assim, essa área se mostra como emergente e que ainda demanda pesquisa para se consolidar.

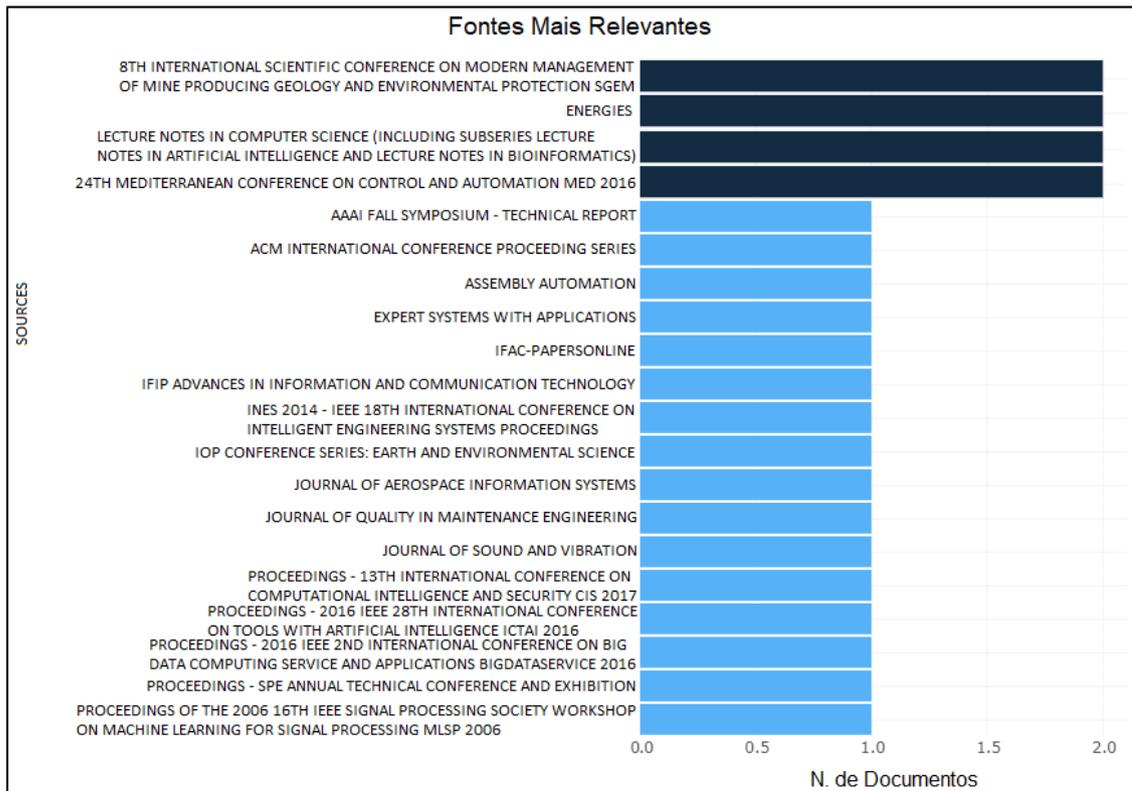


Figura 3-13 - Fontes mais relevantes de publicação de artigos na área

Na Figura 3-14, o gráfico apresentado mostra os autores com maior produção de artigos ao longo dos anos. A quantidade de artigos publicados é representada pelo diâmetro do círculo, sendo que quanto mais escura sua cor de preenchimento, mais esse artigo foi citado.

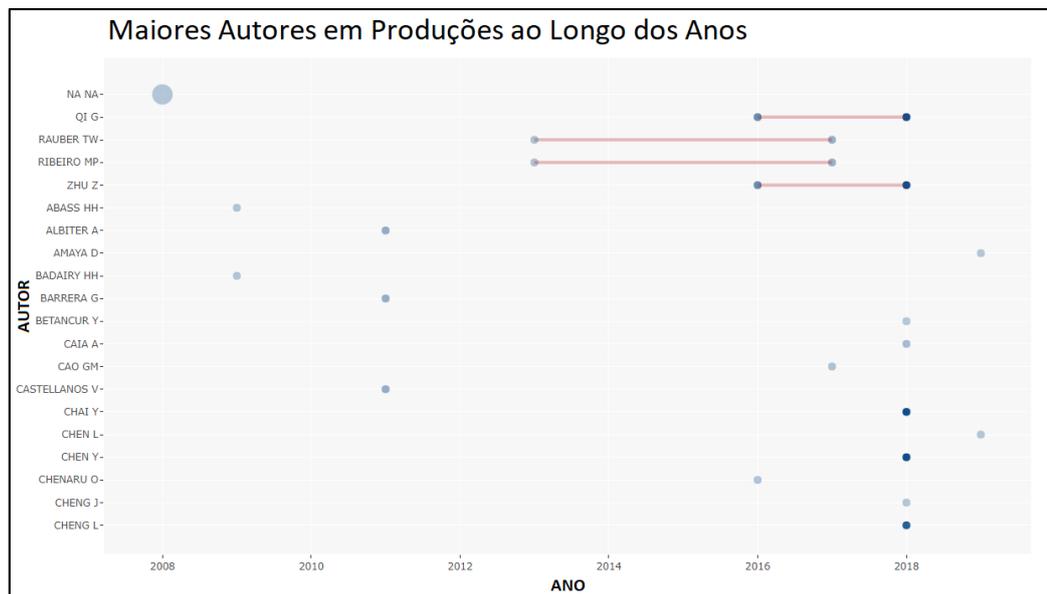


Figura 3-14 - Autores com maior número de produções e citações

Nele, é possível notar que apenas o autor Na Na realizou mais de uma produção em um mesmo ano, porém Qi G, Rauber TW, Ribeiro MP e Zhu Z produziram mais de um artigo entre 2013 e 2018. Outro ponto importante que deve ser notado é que os artigos mais citados são

recentes (2018), corroborando assim com as análises já realizadas neste trabalho de que esta área de pesquisa ainda está em seu período de consolidação.

Na Figura 3-15, está representada a evolução cronológica da produção científica. Neste gráfico é possível verificar que, apesar de um artigo ter sido publicado em 1997, a área iniciou sua produção científica apenas em 2005, se mantendo com, no máximo, três artigos por ano até 2017. Já em 2018, houve uma produção maior de artigos que, apesar de não ter se mantido no mesmo valor em 2019, se manteve acima dos anos anteriores.

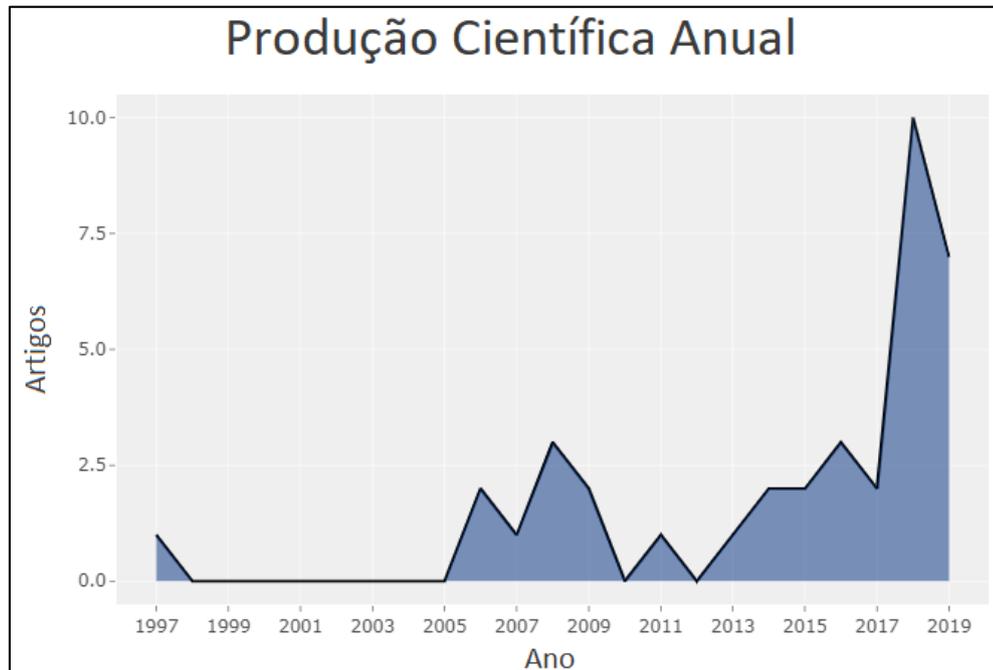


Figura 3-15 - Evolução da produção científica por ano.
Fonte: Dos autores (2020).

Já na Figura 3-16, são apresentados os países que mais produzem nesta área. Conforme visto na Figura 3-12, China e Estados Unidos são os países com maior quantidade de produções. Outra informação que pode ser retirada desta imagem é que a maioria dos países que estão produzindo conhecimento são países produtores de petróleo, porém há uma relação de cooperação apenas entre China e Estados Unidos. Percebe-se então que a área ainda necessita de uma maior pulverização e colaboração entre outros países para um melhor desenvolvimento.

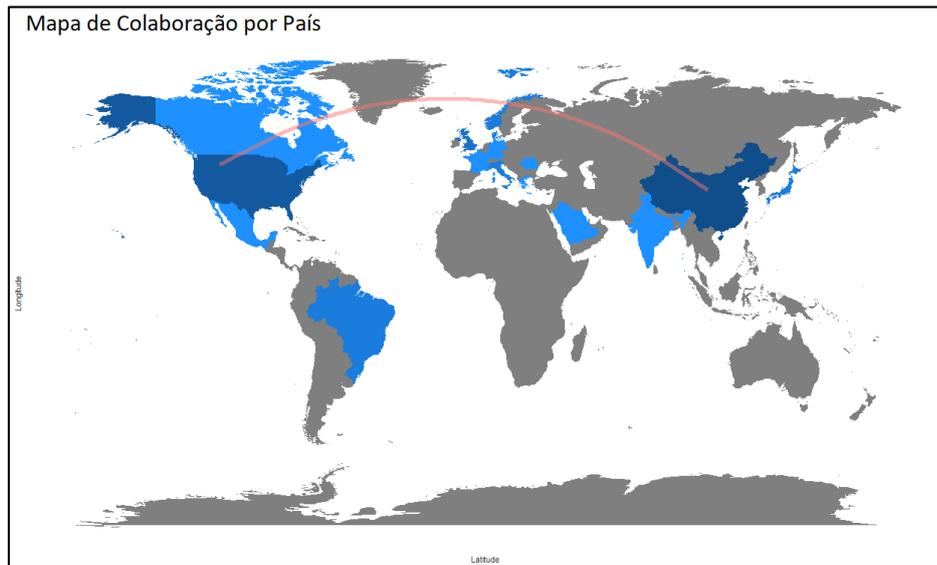


Figura 3-16 - Distribuição da produção científica por país.
Fonte: Dos autores (2020).

Na Figura 3-17 está representado um gráfico que mostra os termos mais utilizados ao longo dos anos. Percebe-se que todos os termos ficam bem próximos até 2013, onde *failure analysis*, *fault detection* e *artificial intelligence* avançam na quantidade de utilizações e figuram nas três primeiras posições, assim como mostrado na Figura 3-12.

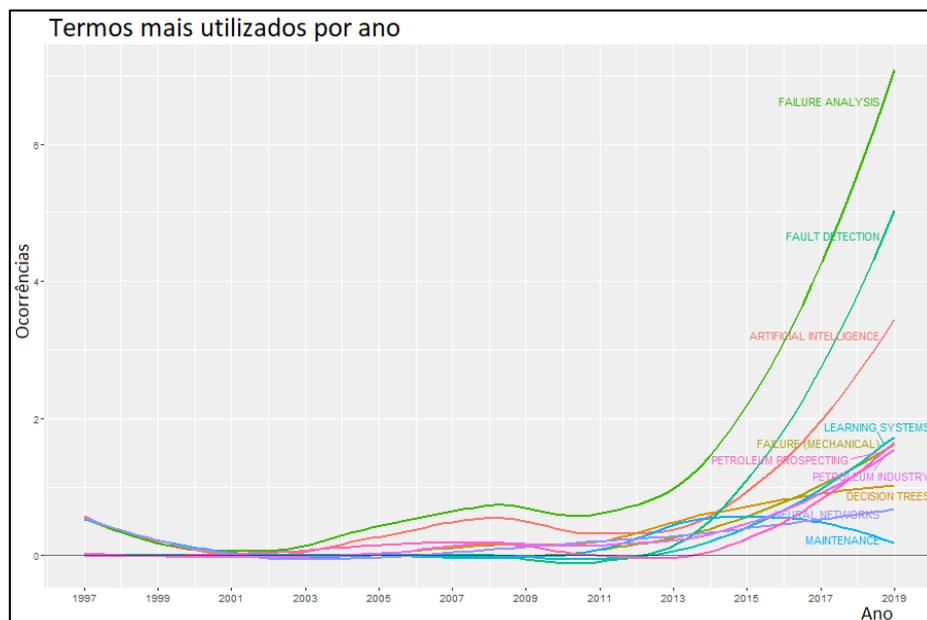


Figura 3-17 - Termos mais utilizados por ano.
Fonte: Dos autores (2020).

Na Figura 3-18, está representado um mapa de árvore com os termos mais frequentes nos artigos encontrados. Ao contrário do gráfico presente na Figura 3-17, o mapa de árvore é estático e apresenta um resultado geral. Dentre os termos de maior frequência, *failure analysis*, *artificial intelligence*, *falt detection*, *decision trees*, *petroleum industry* e *learning systems*

estavam na sentença de busca, sendo assim termos que possuem maior probabilidade de atingir um alto valor de frequência.

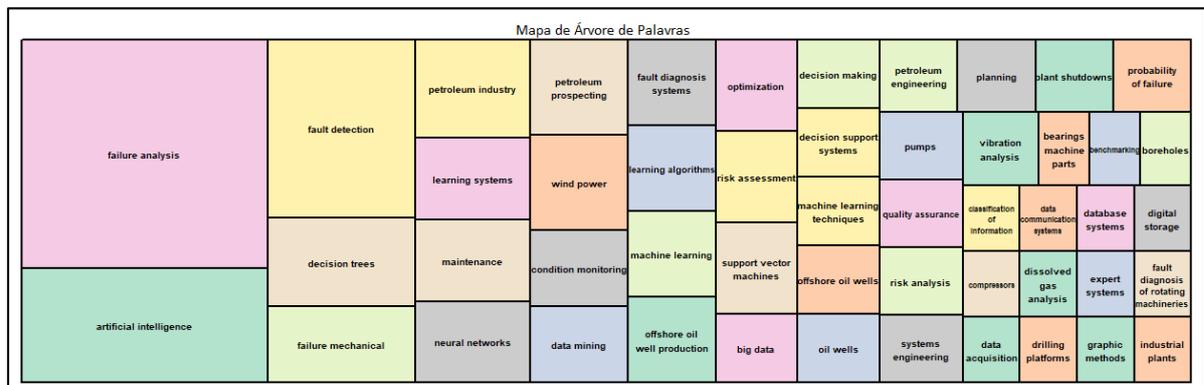


Figura 3-18 - Mapa de árvore dos termos mais citados

Entretanto, *failure mechanical*, *maintenance* e *risk assessment* figuram entre os termos com maior frequência mesmo sem serem buscados, mostrando assim que as pesquisas estão voltadas para a predição de falhas visando a manutenção e componentes mecânicos. O mesmo acontece com *wind power*, o que demonstra um crescente interesse em energias renováveis, e com *support vector machines*, técnica de predição que mesmo sem ter sido diretamente pesquisada aparece devido à sua relevância para a academia.

Dadas as atualizações, foi possível realizar uma análise voltada para os artigos que atendessem à estratégia de busca. Os novos artigos passaram por um processo de análise mais criterioso e profundo, explorando fatores não explorados anteriormente.

3.8 Referências

- ARIA, M.; CUCCURULLO, C. bibliometrix : An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959–975, nov. 2017.
- CHEN, J.; PATTON, R. J. Robust Model-Based Fault Diagnosis for Dynamic Systems. [s.l.] Springer Science & Business Media, 2012.
- COSTA, H. G. Modelo para webibliomining: proposta e caso de aplicação. *Revista da FAE*, v. 13, p. 115–126, 2010.
- HOOD, W. W.; WILSON, C. S. The Literature of Bibliometrics, Scientometrics, and Informetrics. *Scientometrics*, v. 52, n. 2, p. 291, 1 out. 2001.
- LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, v. 108, p. 33–47, 2018.

NANDI, S.; TOLIYAT, H. A. Condition monitoring and fault diagnosis of electrical machines - a review. . In: CONFERENCE RECORD - IAS ANNUAL MEETING (IEEE INDUSTRY APPLICATIONS SOCIETY). 1999

NEUROTH, M. et al. Improved modelling and control of oil and gas transport facility operations using artificial intelligence. Knowledge-Based Systems, v. 13, n. 2, p. 81–92, 1 abr. 2000.

QIN, S. J. Survey on data-driven industrial process monitoring and diagnosis. Annual Reviews in Control, v. 36, n. 2, p. 220–234, 2012.

RENGASWAMY, R.; VENKATASUBRAMANIAN, V. A syntactic pattern-recognition approach for process monitoring and fault diagnosis. Engineering Applications of Artificial Intelligence, v. 8, n. 1, p. 35–51, 1 fev. 1995.

CHEN, J.; PATTON, R. J. Robust Model-Based Fault Diagnosis for Dynamic Systems. [s.l.] Springer Science & Business Media, 2012.

COSTA, H. G. Modelo para webibliomining: proposta e caso de aplicação. Revista da FAE, v. 13, p. 115–126, 2010.

HOOD, W. W.; WILSON, C. S. The Literature of Bibliometrics, Scientometrics, and Informetrics. Scientometrics, v. 52, n. 2, p. 291, 1 out. 2001.

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. Mechanical Systems and Signal Processing, v. 108, p. 33–47, 2018.

NANDI, S.; TOLIYAT, H. A. Condition monitoring and fault diagnosis of electrical machines - a review. . In: CONFERENCE RECORD - IAS ANNUAL MEETING (IEEE INDUSTRY APPLICATIONS SOCIETY). 1999

NEUROTH, M. et al. Improved modelling and control of oil and gas transport facility operations using artificial intelligence. Knowledge-Based Systems, v. 13, n. 2, p. 81–92, 1 abr. 2000.

QIN, S. J. Survey on data-driven industrial process monitoring and diagnosis. Annual Reviews in Control, v. 36, n. 2, p. 220–234, 2012.

RENGASWAMY, R.; VENKATASUBRAMANIAN, V. A syntactic pattern-recognition approach for process monitoring and fault diagnosis. Engineering Applications of Artificial Intelligence, v. 8, n. 1, p. 35–51, 1 fev. 1995.

4 ARTIGO B - UTILIZAÇÃO DE TÉCNICAS DE MACHINE LEARNING NA PREVISÃO DE FALHAS EM EQUIPAMENTOS: UMA REVISÃO SISTEMÁTICA

4.1 Resumo

A quantidade total de dados no mundo tem crescido de forma exponencial nos últimos anos. Porém, este crescimento não é uma exclusividade das empresas de tecnologia, tendo atingido diversos ramos da indústria. Dentro do cenário competitivo que as indústrias estão envolvidas, a manutenção se tornou protagonista para que a qualidade e a produtividade empresarial mantenha altos padrões. Com a capacidade de captação de dados através de sensores, a manutenção preditiva apresenta um melhor desempenho para evitar falhas. Apesar dessa disponibilidade de dados, é necessário extrair informações e gerar conhecimento para conseguir atingir o verdadeiro potencial. Desta forma, as técnicas de inteligência artificial se mostram promissoras ferramentas de detecção e reconhecimento de padrões de falhas, tendo a produção científica voltada para essa área crescido nos últimos anos. Assim, este trabalho propõe uma revisão sistemática com a finalidade de evidenciar os estudos de maior qualidade. Foi utilizado o método PRISMA para a preparação da revisão sistemática. Foram buscados artigos nas bases de dados Scopus e IEEE Xplore e os artigos selecionados passaram por etapas de filtragem baseadas no período de publicação e área de interesse e por etapas de avaliação baseadas nos títulos, resumos e no texto completo do artigo. A estratégia de busca retornou 2.963 artigos de onde, após a etapa de análise, foram retirados 12 artigos. Os artigos foram agrupados quanto ao tipo de técnica utilizada e quanto sua área de aplicação. As técnicas de classificação e regressão utilizando SVM foram as mais encontradas, além da utilização de árvores de decisão e redes neurais, por exemplo. Assim, este trabalho obteve sucesso em demonstrar a utilização e aplicabilidade das ferramentas de inteligência artificial, assim como a variedade de técnicas contidas nos trabalhos.

Palavras-chave: revisão sistemática, mineração de dados, previsão de falhas.

4.2 Abstract

The total amount of data in the world has been grow in an exponential way in the last years. But this growth is not an exclusivity of technology companies, reaching industry branch in general. The industries are involved in a competitive scenario, where maintenance took the main role so quality and corporate productivity keep high standards. With the ability of data capture through sensors, predictive maintenance presents a better performance to avoid failures. Even with this data availability, this is not enough to reach the full potential of this data and info extraction and knowledge generation are needed. Therefore, the artificial

intelligence techniques show themselves as promising tools to detect and recognize failure patterns, so the scientific community has been putting effort to study this area in the last years. Thereby, this work aims to do a systematic revision to spotlight the studies with greater quality. The PRISMA method was used to prepare the systematic revision. The papers were searched in Scopus and IEEE Xplorer databases and the selected ones passed through filtering and a nalysis steps, based on publishing year, area of interest, title, abstract and the complete paper text. The search strategy returned 2,963 papers from where, after analysis step, 12 papers were selected. They were grouped by the used technique type and by the application area. Classification and regression using SVM were found with greater frequency, followed by other techniques such as decision trees and neural networks, for example. So, this paper achieved its goal by demonstrating the use and applicability of artificial intelligence tools, as the variety of other techniques.

Keywords: systematic revision, data mining, failure prediction.

4.3 Introdução

A quantidade total de dados no mundo aumentou nove vezes entre os anos de 2006 e 2011 (GANTZ; REINSEL, 2011) e dobrando a cada dois anos a partir de 2012 até 2020, chegando a um valor 300 vezes maior se comparado com 2005 (GANTZ; REINSEL, 2012). Todo esse crescimento de dados não se restringiu ao uso de empresas de tecnologia, chegando até às indústrias. Nos últimos anos, quase todos os ramos da indústria tem experimentado um crescimento exponencial de dados (MONDAL; BARUA, 2019).

Dado o cenário competitivo em que as indústrias estão envolvidas, a manutenção se tornou protagonista para manter a qualidade e a produtividade empresarial (ALI, 2000; MARCORIN; LIMA, 2003; SWANSON, 2001). Dentre todos os tipos de manutenção, a manutenção preditiva apresenta um melhor desempenho para evitar falhas (MUNION, 2017) e se aproveita da quantidade de dados gerados por sistemas instrumentados. Uma única máquina, por exemplo, é capaz de gerar 5.000 amostras a cada 33ms, resultando em um total de 4 trilhões de amostras em um ano (GE AUTOMATION, 2015).

Mesmo com as atuais soluções existentes para aquisição de dados simplificarem a captura dos dados de campo, é necessário criar conhecimento a partir desses dados utilizando outras ferramentas e modelos (ACCORSI *et al.*, 2017). Uma das formas de geração de conhecimento utilizando esse grande volume de dados disponível são aplicações de ferramentas de inteligência artificial (IA) que tem se mostrado promissoras na detecção e reconhecimento de padrões de falha (LIU *et al.*, 2018a). Velloso, Hora e Atem (2018) demonstraram que a produção científica

sobre a utilização de inteligência artificial para a detecção de falhas vem aumentando nos últimos anos.

As ferramentas de IA podem ser usadas tanto em modelos *physics-based*, modelos baseados nas propriedades físicas do equipamento que levam em consideração o tipo de material do qual a máquina é feita, método de funcionamento e etc., quanto em modelos *data-driven*, modelos baseados apenas nos dados capturados da máquina. Zhou *et al.* (2017) mostram que apesar dos modelos *physics-based* serem altamente acurados, eles, geralmente, possuem alta dimensionalidade e são computacionalmente custosos e, em contrapartida, os modelos *data-driven* visam baixa dimensionalidade e diminuir a complexidade computacional. Zhou *et al.* (2017) mostraram ainda que os modelos *data-driven* podem ser utilizados como substitutos para os *physics-based* com uma perda insignificante de acurácia de predição em aplicações que possuem dados operacionais disponíveis.

Dado o crescente desenvolvimento da área, diversos métodos foram surgindo, sendo aplicados em vários cenários e os modelos *data-driven* foram mostrando sua aplicabilidade. Quando há uma diversidade de estudos que podem ser até controversos, a revisão sistemática surge como o caminho mais coerente para evidenciar os estudos de melhor qualidade (GALVÃO; PEREIRA, 2014). Assim, este trabalho se propõe a realizar uma revisão sistemática a fim de evidenciar as principais técnicas de *machine learning* para previsão de falhas em equipamentos e máquinas que estão sendo utilizadas pela academia.

4.4 Metodologia

Este trabalho segue a metodologia PRISMA (Principais Itens para Relatar Revisões sistemáticas e Meta-Análises) de preparação de uma revisão sistemática como apresentado por Galvão e Pansani (2015). Esta metodologia é uma evolução da metodologia que seguia a recomendação QUORUM (Qualidade dos Relatos de Meta-análises) por abranger tanto as revisões sistemáticas quanto as meta-análises.

Baseando-se na metodologia PRISMA, esta seção será dividida em: (1) Critérios de elegibilidade, (2) Fontes de informação, (3) Busca e (4) Seleção dos estudos. Após essas etapas, os resultados serão apresentados no capítulo.

4.4.1 CRITÉRIOS DE ELEGIBILIDADE

Os trabalhos selecionados devem responder a seguinte pergunta de pesquisa: Qual é o desempenho das técnicas de *machine learning* na previsão de falhas em equipamentos? Esta pergunta foi elaborada utilizando a metodologia PICOS (anagrama em inglês para os termos

população, intervenção, comparação, desfecho e estudo), como apresentado por Galvão e Pereira (2014) e listado na Tabela 4-1.

Tabela 4-1 - Técnica PICOS para elaboração da pergunta de pesquisa

Descrição	Abreviação	Componentes da pergunta
População	P	Equipamentos
Intervenção	I	Técnicas de <i>machine learning</i>
Comparação	C	Desempenho
Desfecho	O	Previsão de falhas
Tipo de estudo	S	Estudos de caso

Fonte: Dos autores (2020).

Foram considerados artigos completos, a fim de verificar estudos com tenham apresentado uma conclusão, publicados dentro dos últimos 10 anos (2009 - 2019), para identificar os estudos mais atuais, e no idioma inglês e português, sendo o inglês o principal idioma de publicação de artigos (DI BITETTI; FERRERAS, 2017) e o português a língua nativa dos autores. Foram excluídos artigos de outras áreas que não engenharia e ciência da computação.

4.4.2 FONTES DE INFORMAÇÃO

Foram buscados artigos em duas bases de dados, Scopus® e IEEE Xplore. O Scopus® foi selecionado por ser considerado um banco de dados de fonte neutra curado por especialistas independentes no assunto (SCOPUS, 2019). Já o IEEE Xplore foi selecionado por ser um banco de publicações que engloba conteúdos técnicos e científicos divulgados pelo Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE) (IEEE, 2019).

4.4.3 ESTRATÉGIA DE BUSCA

A fim de filtrar os trabalhos que não atendessem aos critérios elencados na seção 4.4.1, foram aplicados filtros na busca dos artigos. Foram utilizados também tesouros que pudessem abranger um número maior de trabalhos sem que o objeto de pesquisa se alterasse.

Foram utilizados os termos de pesquisa e tesouros conforme Figura 4-1. Todos esses termos foram utilizados na busca realizada na base Scopus®, porém a base IEEE Xplore não possibilita a combinação das palavras-chave com tesouros na sentença de busca. Além destes termos, foi realizado um recorte quanto à área de estudo na base Scopus®. Entretanto, este recorte não se aplica à base IEEE Explore uma vez que todos os trabalhos ali publicados estão correlacionados apenas às áreas de engenharia, ciência da computação e eletrônica.

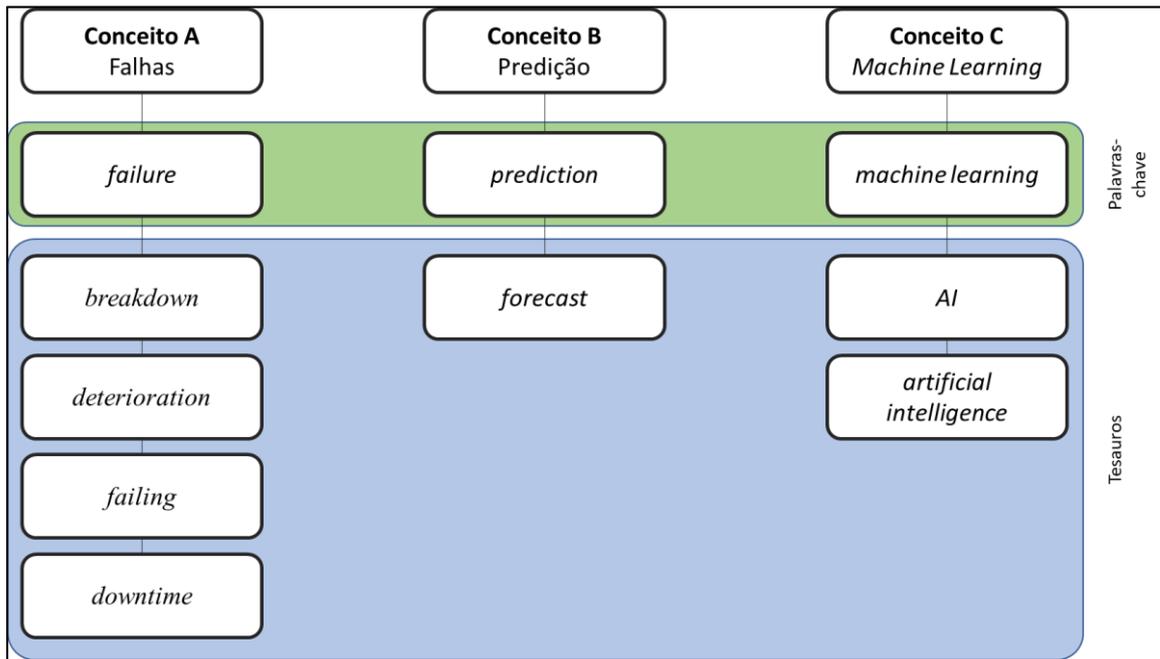


Figura 4-1 - Palavras-chave e tesauros utilizados na busca.
 Fonte: Dos autores (2020).

Quadro 4-1 - Sentenças de busca

Base de dados	Scopus®	IEEE Xplore
Sentença de Busca	(TITLE-ABS-KEY ("failure" OR "breakdown" OR "deterioration" OR "failing" OR "downtime") AND TITLE-ABS-KEY ("prediction" OR "forecast*") AND TITLE-ABS-KEY ("machine learning" OR "AI" OR "artificial intelligence"))	((("All Metadata":failure) AND "All Metadata":prediction) AND "All Metadata":machine learning)

Fonte: Dos autores (2020).

No Quadro 4-1 estão dispostas as sentenças de busca utilizadas nesta etapa para cada uma das base de dados.

4.4.4 SELEÇÃO DOS ESTUDOS

Além dos filtros já aplicados para a busca do artigo, alguns requisitos foram elencados para dar maior objetividade à revisão. Os artigos devem citar a utilização de técnicas de *machine learning* e descrever qual ou quais foram as técnicas utilizadas. Além dos requisitos já citados, os trabalhos devem utilizar modelos de predição puramente *data-driven* e devem ser aplicados a previsão de falhas em equipamentos/máquinas, uma vez que esses itens compõem o objetivo do trabalho.

O método de avaliação seguirá o seguinte fluxo: (1) o título de todos os artigos foram lidos, (2) o resumo dos artigos selecionados e (3) leitura do artigo completo. Em cada um dos passos serão selecionados aqueles que apresentarem maior afinidade com o tema deste trabalho e servirão de base de avaliação para o passo seguinte.

4.5 Resultados

As pesquisas foram realizadas nos dias 11/02/2020, retornando um total de 2.963 artigos, sendo 2.462 da base Scopus® e 501 da base IEEE Xplorer.

Inicialmente, os artigos foram submetidos a uma verificação de trabalhos duplicados, porém não foi encontrado nenhum. Por causa da grande quantidade de resultados da busca, foi necessário realizar uma filtragem por área de conhecimento, levando a uma exclusão de 1.330 artigos. Seguinte a esta etapa, 262 documentos foram removidos após a aplicação do corte temporal. Após esses filtros de pré-análise, foram realizadas as análises de título, resumo e texto completo, resultando em uma seleção de 13 trabalhos. O fluxograma mostrado na Figura 4-2 apresenta a estratégia de pesquisa.

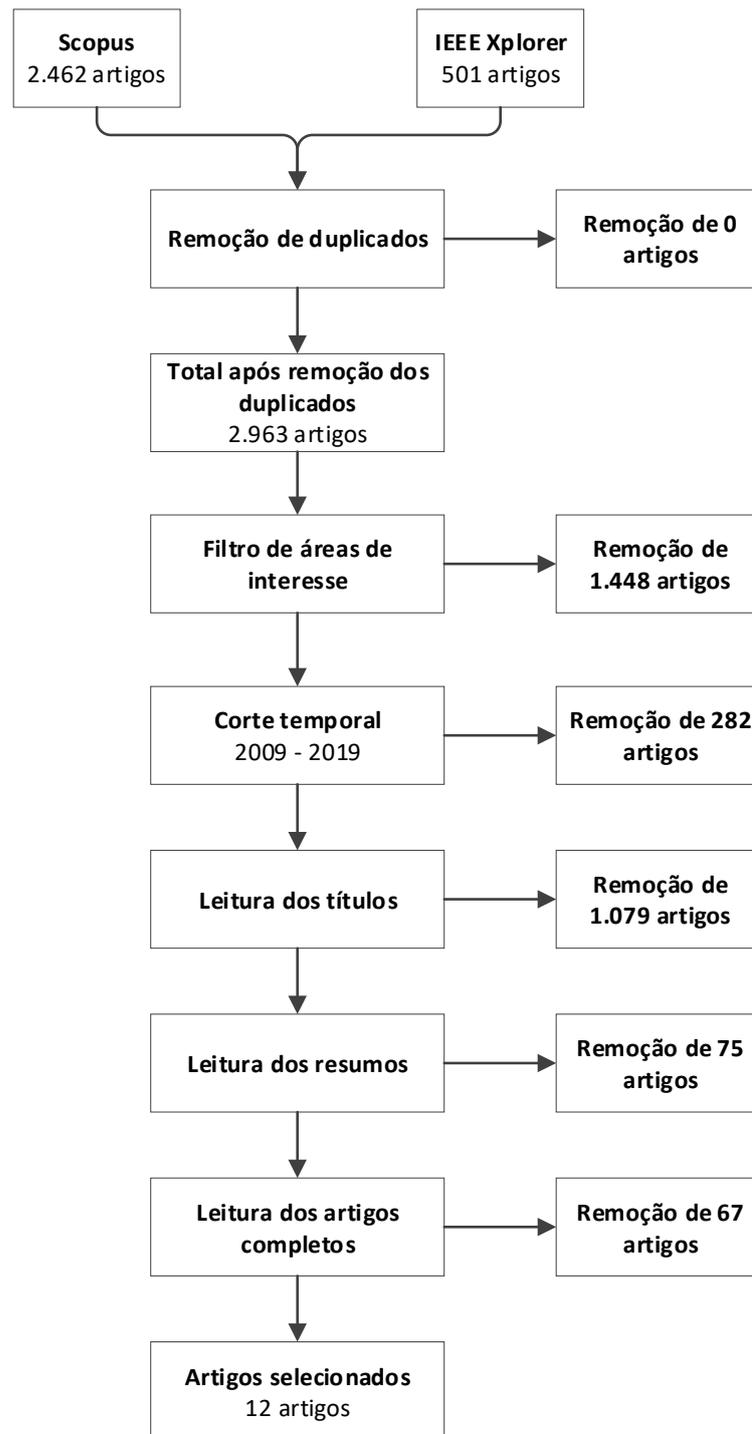


Figura 4-2 - Fluxograma do procedimento de pesquisa.

Fonte: Dos autores (2020).

Apesar de todos os trabalhos selecionados terem como tema a previsão de falhas, eles se diferenciam de acordo com a técnica de *machine learning* utilizada e qual tipo de equipamento no qual a técnica foi utilizada.

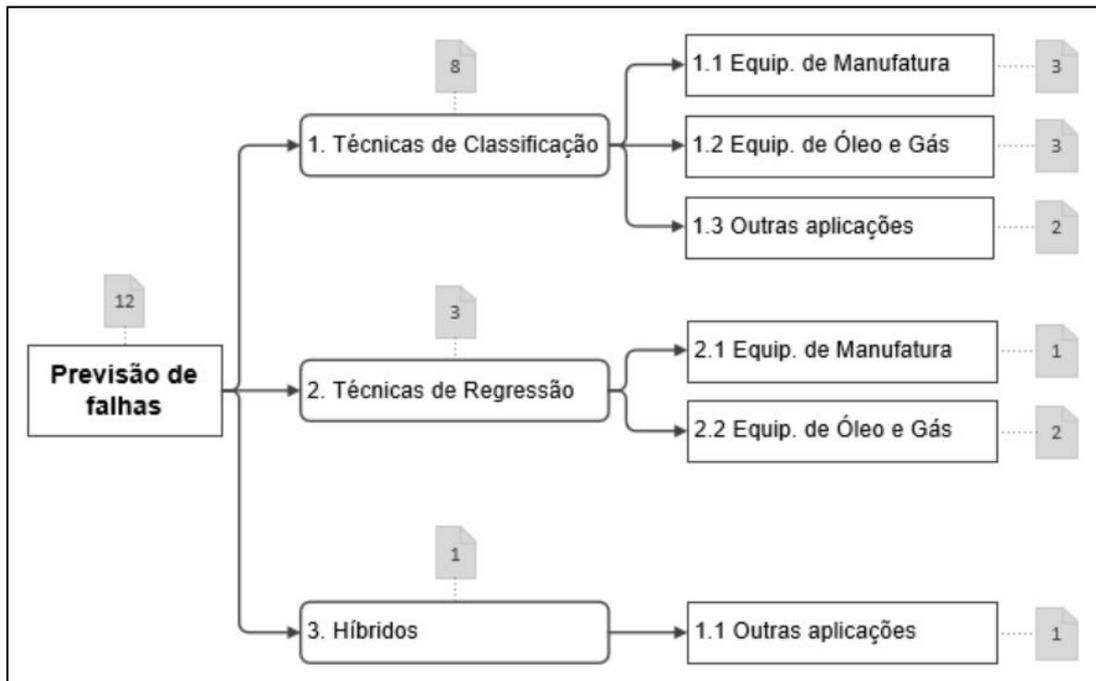


Figura 4-3 - Segmentação dos artigos de previsão de falhas.

Fonte: Dos autores (2020).

A Figura 4-3 mostra os dois principais tipos de técnica utilizadas para previsão de falhas: (1), com 9 artigos e (2) com 4 artigos. Na categoria (3), ainda é possível verificar artigos híbridos que utilizam as técnicas propostas em (1) e (2). Dentre os equipamentos utilizados como base para o estudo, os equipamentos de manufatura e óleo e gás demonstram prevalência sobre os outros equipamentos que foram agrupados na categoria Outras aplicações.

4.5.1 TÉCNICAS DE CLASSIFICAÇÃO

4.5.1.1 Equipamentos de manufatura

Accorsi *et al.* (2017) fizeram um estudo de comparação de técnicas de classificação para a previsão de falhas nos equipamentos de uma linha de montagem de maquinário de manufatura utilizando 116 atributos de diferentes sensores. As técnicas utilizadas foram árvore de decisão, *random forests* e redes neurais. Chegou-se à conclusão que para a linha de montagem estudada o algoritmo de *random forest* apresenta melhor resultado, com uma acurácia (71%) maior se comparado a árvore de decisão (63%) e redes neurais (64%), mesmo apresentando uma predição de valores negativos muito baixa (38%) o que pode acarretar em paradas inesperadas por alarmes falsos.

A utilização dos dados provenientes dos sensores das máquinas se apresenta como um meio de alimentar modelos também em Qiang, Cheng e Jing (2010). Neste trabalho, é criado um modelo de predição de falha que utiliza o sinal do sensor de força de corte de uma máquina de

usinagem e os algoritmos de *Support Vector Machine* (SVM) e *Hidden Markov Models* (HMM). Os autores demonstram que, se essas duas técnicas forem usadas na forma híbrida de SVM/HMM, elas tem um desempenho maior (95%) se comparado a utilização do algoritmo HMM de forma independente, que apresentou um desempenho de 88,9%.

Enquanto Accorsi *et al.* (2017) e Qiang, Cheng e Jing (2010) utilizaram atributos provenientes de diferentes sensores já presentes nos equipamentos de manufatura, Han *et al.* (2017) pesquisaram um modelo de redes neurais para previsão de falhas em equipamentos industriais com dados provenientes de um sensor de som externo ao equipamento. Neste, foram utilizadas duas configurações desta técnica: redes neurais convolucionais e redes neurais com retropropagação, onde a primeira apresentou uma acurácia ligeiramente maior que a segunda (94,6% contra 93,3%), mas muito superior a utilização de *randomforest* proposta por Accorsi *et al.* (2017).

4.5.1.2 Equipamentos de óleo em gás

Alguns equipamentos de óleo e gás apresentam uma maior dificuldade pois os equipamentos que necessitam ser monitorados estão, geralmente, em operação no fundo do mar. Em Okoli *et al.* (2019), por exemplo, são apresentadas formas para prever o nível de vibração da coluna de perfuração de um poço de petróleo, uma vez que essas vibrações podem causar desgaste e diminuir sua vida útil. Foram utilizados parâmetros como a taxa de penetração e torque na broca como dados de entrada de forma independente e cinco diferentes técnicas de *machine learning*: regressão logística multinomial, análise discriminante linear, *k-Nearest Neighbour* (KNN), árvore de decisão e Naïve Bayes Gaussiano. As previsões utilizando a variável Torque apresentaram maior acurácia, tendo as técnicas de regressão logística multinomial, análise discriminante linear e Naïve Bayes Gaussiano apresentado os melhores resultados (76%).

Guo *et al.* (2015) também utilizaram variáveis dinâmicas que dependem de como os equipamentos estão sendo utilizados. Em seu trabalho, foi aplicada a técnica de SVM para prever desgastes em uma bomba elétrica submarina (ESP, em inglês) utilizando os dados de corrente, tensão, frequência de operação e quantidade de trabalho diária do equipamento em dois campos diferentes (FMT1 e FMT2). A performance do algoritmo foi diferente entre os campos FMT1 e FMT2, dadas as particularidades de cada campo. Os autores encontraram uma precisão geral de 61,8% e uma sensibilidade de 61,5%. Apesar do modelo apresentar um desempenho razoavelmente bom, ele perde com falta de atributos e exemplos de treino para a geração de um modelo de previsão mais acurado e generalizado.

Diferente de Okoli *et al.* (2019) e Guo *et al.* (2015), Ouadah (2019) demonstra a utilização das ferramentas de previsão de falhas em um sistema estático, *pipelines*. Nos dados captados por *pigs* instrumentados (sensores que captam os dados conforme passam por dentro dos dutos) foi utilizada a técnica de *clustering K-means* combinada com as técnicas de classificação árvore de decisão e redes neurais. Os resultados evidenciam que as duas técnicas de classificação apresentam um percentual de instâncias classificadas corretamente acima de 98%, mas com uma ligeira superioridade da técnica de árvore de decisão (99,80% de precisão) sobre a técnica de redes neurais (98,85% de precisão).

4.5.1.3 Outras aplicações

Foram encontradas aplicações para previsão de falhas também em ferrovias no trabalho de Wang *et al.* (2017) e em turbinas eólicas no trabalho de Leahy *et al.* (2016). Apesar de áreas diferentes, ambos utilizaram a técnica de SVM para a predição das falhas. Em Wang *et al.* (2017), foram coletados os registros de falhas relacionadas a condições climáticas em sistemas de desvio ferroviário de uma empresa chinesa em um período de cinco anos e aplicadas as técnicas de AdaBoost.LS e AdaBoost.RT para encontrar os melhores parâmetros a fim de encontrar a melhor performance do algoritmo de classificação. Os resultados apresentaram um melhor desempenho da técnica de AdaBoost.LS se comparada com AdaBoost.RT para aplicações onde o modelo é treinado com pelo menos 60% do *dataset*.

Já em Leahy *et al.* (2016), foram utilizados dados de três turbinas eólicas de geração de energia coletados por onze meses entre os anos de 2014 e 2015. Os autores aplicaram a técnica SVM para predição de falhas em geral e para predição de modos de falha específicos. Apesar de apresentar alta sensibilidade (90%) e especificidade (83%) para falhas no geral, o modelo teve uma precisão muito baixa (8%) e, por isso, muitas entradas foram classificadas como falha de forma incorreta. Este resultado é condizente com outros trabalhos que apresentam *datasets* desbalanceados como o utilizado nessa pesquisa. Mesmo com o baixo desempenho para a classificação geral de falhas, os resultados foram mais promissores para a classificação de falhas provenientes do aquecimento do gerador, com uma precisão de 56%, uma sensibilidade de 100% e especificidade de 99%.

4.5.2 TÉCNICAS DE REGRESSÃO

4.5.2.1 Equipamentos de manufatura

Em Yao *et al.* (2019), a técnica de *Support Vector Regression* (SVR) foi aplicada, juntamente ao algoritmo de *Particle Swam Optimazation* (PSO), em uma planta química da empresa Tennessee Eastman. Foram coletadas 980 amostras para treinamento do modelo, onde

500 representam o funcionamento normal e 480 representam 21 tipos de falha. Para teste do modelo, foram coletadas 960 amostras contendo os mesmos tipos de falha. A coleta foi realizada a cada 3 minutos, resultando em um tempo amostral de 49h para treino e 48h para teste. Os resultados apresentaram um erro máximo de, aproximadamente, 13% e um erro médio de 0,0468%, demonstrando alta acurácia nesta aplicação.

Ainda em Yao *et al.* (2019), o algoritmo de SVR é aplicado a um *dataset* da NASA com dados de quatro rolamentos submetidos a testes de vibração. Foram selecionadas 1.850 amostras de treino e quinze amostras de teste. Os resultados demonstraram desempenho semelhante à aplicação anterior.

4.5.2.2 Equipamentos de óleo e gás

Assim como nos trabalhos que utilizam algoritmos de classificação, Shirzad, Tabesh e Farmani (2014) estudaram a predição de falhas em *pipelines* e Sneed (2017) em ESPs. Shirzad, Tabesh e Farmani (2014) realizaram uma comparação entre as técnicas de redes neurais e SVR para prever falhas em dois *pipelines* de transporte de água do Irã, um em Mashhad e outro em Mahabad. Diâmetro, idade, comprimento, profundidade de instalação e pressão hidráulica (média e máxima) foram as características consideradas para o desenvolvimento dos modelos, sendo a última coletada durante o intervalo de um ano. O desempenho dos resultados foi classificado quanto às medidas de erro médio quadrático (RMSE), erro médio quadrático normalizado (NMSE), erro médio de *bias* normalizado (NMBE), erro absoluto médio (MAE), percentual de erro absoluto médio (MAPE), índice de concordância (IOA) e coeficiente de determinação (R^2).

No modelo do *pipeline* de Mashhad, foram considerados dutos de 80 a 300mm de diâmetro e 337 casos de falha reportados durante o período de coleta. Após serem aplicadas as ferramentas de regressão, a regressão SVR apresentou melhor desempenho em seis das sete métricas de avaliação, como pode ser visto abaixo na Tabela 4-2. Apesar de performar melhor em apenas uma das métricas, a técnica de redes neurais apresentou desempenho muito semelhante à técnica de SVR.

Tabela 4-2 - Desempenho dos modelos

Métricas de performance	Mashhad Model		Mahabad Model	
	SVR	ANN	SVR	ANN
RMSE	0,028967	0,037251	0,919037	0,045931
NMSE	0,002547	0,004212	0,228903	0,000572
NMBE	0,000697	-0,001556	-0,036415	-0,000917
MAE	0,019655	0,016823	0,440707	0,026051
MAPE	0,015719	0,019864	0,197710	0,016886
IOA	0,999998	0,999991	0,998956	0,999999
R2	0,997465	0,995755	0,770797	0,999426

Fonte: Adaptado de Shirzad, Tabesh e Farmani (2014).

No modelo proposto para o *pipeline* de Mahabad, foram considerados dutos de 32 a 160mm de diâmetro e 349 casos de falha reportados. Diferente do modelo anterior, a técnica de redes neurais obteve um desempenho superior à técnica de SVR em todas as métricas de performance, como apresentado também na Tabela 4-2.

Os autores apresentam como conclusão que a técnica de redes neurais é, de forma geral, um preditor melhor do que o SVR, porém não pode ser facilmente utilizada para propósitos de generalização.

Já em Sneed (2017), o objetivo é prever a quantidade de vida útil remanescente de uma ESP, ou seja, deseja-se prever quando irá ocorrer a falha que ocasionará na parada do equipamento. Foram gerados modelos utilizando diferentes técnicas para verificar qual delas tem a melhor performance. As técnicas regressão logística múltipla, árvore de decisão e *random forest* de alta performance (HP Forest) foram aplicadas em um *dataset* composto por 51 falhas de ESPs entre 37 poços ocorridas entre janeiro de 2015 a julho de 2016.

O modelo criado pelo HP Forest obteve o menor erro quadrático médio de todos os modelos considerados e considerado o mais apto para a predição das falhas desta aplicação. Os resultados apresentados mostram que esta técnica conseguiu prever, em média, a vida útil remanescente do equipamento dentro de um range de 5 dias de diferença do tempo real até a falha, aproximadamente. Apesar do bom desempenho, o modelo possui a desvantagem de que seu resultado não é facilmente explicado uma vez que as relações entre as variáveis de entrada e de saída não ficam claras.

4.5.3 HÍBRIDOS

Jalali *et al.* (2019) propõem a utilização de modelos de predição de falhas baseado em dados para equipamentos de gravação por plasma em uma indústria de semicondutores, uma

vez que esse maquinário executa procedimentos complexos e é difícil de capturar todas as propriedades, variáveis e comportamentos relevantes para um único modelo físico. Foram propostas três diferentes tarefas de predição: (i) predição do Tempo-Até-Falha (TTF, em inglês), (ii) predição do estado de saúde, e (iii) predição dos intervalos do TTF de um equipamento.

As propostas de predição foram modeladas como problemas de regressão (i e ii) e classificação (iii). Para os problemas de regressão foram utilizadas as técnicas de regressão linear, SVM, árvores de decisão, *random forest* e redes neurais. Já para o problema de classificação foram utilizadas as mesmas técnicas de regressão (excluindo regressão linear) e *Gradient Boosting Classifier* (GBC), SVM com *Stochastic Gradient Descent* (SGD) e *k-Nearest Neighbor* (kNN).

Para identificar o melhor conjunto de características, cada modelo foi testado com sete diferentes *datasets* como variáveis de entrada, onde cada *dataset* é uma combinação de características operacionais, como medidas de sensores do equipamento, limites e alarmes e medidas de engenharia calculadas. Esses *datasets* foram identificados como FS1, FS2, FS3, FS4, FS5, FS6 e FS7.

Nos modelos de regressão, o resultado para a tarefa (i) demonstrou que a técnica de regressão linear obteve o melhor desempenho, apresentando o menor erro médio quadrático para modelos treinados com o conjunto FS1. Na tarefa (ii), a variável objetivo TTF teve seu valor normalizado para o intervalo de zero a um, onde um seria um equipamento com todos os seus processos funcionando perfeitamente e zero o outro extremo. Os resultados mostraram que o algoritmo de redes neurais apresentou os melhores resultados para todos os conjuntos em que foi treinado, conforme Tabela 4-3.

Tabela 4-3 - Valores de RMSE para o experimento (ii)

Conjunto de dados	Regressão Linear	SVM	Redes Neurais	Árvore de Decisão	<i>Random Forest</i>
FS1	0,14	0,17	0,14	0,33	0,17
FS2	0,19	0,17	0,08	0,33	0,17
FS3	0,25	0,17	0,11	0,47	0,19
FS4	0,25	0,17	0,11	0,47	0,17
FS5	0,25	0,17	0,11	0,47	0,19
FS6	0,22	0,14	0,11	0,44	0,17
FS7	0,22	0,14	0,11	0,42	0,17

Fonte: Adaptado de Jalali *et al.* (2019)

Na tarefa (iii), o problema de classificação foi dividido em períodos predefinidos de 0 a x , onde $x \in \{8h, 16h, 24h, 48h, 72h, 96h, 120h, 144h, 168h, 336h\}$. Para determinado intervalo, o desfecho é considerado verdadeiro se o TTF estiver compreendido no intervalo objeto.

Tabela 4-4 - Desempenho das ferramentas de classificação

Desfecho	Conjunto	Melhor modelo	Precisão	Sensibilidade	F1
0-8 h	FS4	Árvore de decisão	0,25	0,36	0,29
0-16 h	FS5	Árvore de decisão	0,36	0,53	0,42
0-24 h	FS7	SVM	0,39	0,75	0,51
2 dias	FS7	SVM	0,59	0,76	0,66
3 dias	FS1	GBC	0,68	0,91	0,78
4 dias	FS1	GBC	0,74	0,96	0,83
5 dias	FS1	GBC	0,79	0,98	0,88
6 dias	FS1	GBC	0,84	0,99	0,91
7 dias	FS1	GBC	0,88	0,99	0,93
14 dias	FS6	Redes neurais	1	1	1

Fonte: Adaptado de Jalali *et al.* (2019)

Na Tabela 4-4 é apresentado o melhor resultado para cada desfecho, assim como suas métricas de precisão, sensibilidade e F1-score. Ou seja, para o desfecho de previsão de dois dias, o melhor resultado é encontrado quando se utiliza o conjunto FS7 e aplica-se a técnica de SVM, encontrando uma precisão de 59% de precisão, sensibilidade de 76% e um F1-score de 66%. O mesmo princípio se aplica a todos os outros desfechos.

4.6 Discussão e Conclusão

Neste artigo, foram apresentados 12 trabalhos que evidenciam a aplicação de algoritmos de *machine learning* para previsão de falhas em equipamentos de diferentes áreas e sua aplicabilidade. De forma geral, todos os modelos descritos apresentaram assertividade nas predições e se mostraram aptos para executar a tarefa de predição de falhas.

Para os problemas de classificação a técnica SVM foi a mais presente com cinco aplicações, seguida pelas técnicas de árvore de decisão e redes neurais que foram aplicadas em quatro trabalhos. Apesar de possuir uma menor frequência, as técnicas de árvore de decisão e redes neurais apresentaram bom desempenho em aplicações de diferentes áreas, como em Accorsi *et al.* (2017), Han *et al.* (2017) e Ouadah (2019). Já o SVM se destacou nos modelos propostos por Qiang, Cheng e Jing (2010), Guo *et al.* (2015), Wang *et al.* (2017) e Leahy *et al.* (2016). Quando tratadas juntas, como em Jalali *et al.* (2019), os três modelos apareceram como a técnica de melhor desempenho em metade dos conjuntos de dados testados.

Assim como nas técnicas de classificação, a técnica de regressão utilizando SVM (também conhecida como SVR) foi a mais citada nos trabalhos avaliados, seguida novamente de árvore de decisão e redes neurais. Em Shirzad, Tabesh e Farmani (2014), foi apresentada uma comparação entre SVR e redes neurais, com um melhor desempenho da primeira na aplicação de Mashhad (RMSE de 2,90%) e da segunda na aplicação de Mahabad (RMSE de 4,59%). Em Yao *et al.* (2019), o modelo gerado pelo SVR representa ainda mais o funcionamento real do equipamento, uma vez que obteve o RMSE de 0,0468%. Enquanto que em Jalali *et al.* (2019), o modelo de redes neurais apresentou melhor resultado com o menor RMSE para todos os conjuntos de teste.

Este trabalho tinha como objetivo apresentar uma revisão sistemática de artigos sobre a utilização de *machine learning* na previsão de falhas de equipamentos. Os resultados apresentados evidenciaram a utilização e aplicabilidade das ferramentas, assim como a variedade de técnicas contidas nos trabalhos. Três técnicas principais foram evidenciadas: SVM, árvore de decisão e redes neurais, estando presentes na maioria dos trabalhos avaliados e conseguindo prever os eventos com bom desempenho. Como trabalho futuro, propõe-se realizar uma revisão sistemática em outras bases de artigos.

4.7 Referências

- ACCORSI, R. et al. Data Mining and Machine Learning for Condition-based Maintenance. **Procedia Manufacturing**, v. 11, p. 1153–1161, 2017.
- ALI, A. The Impact of Innovativeness and Development Time on New Product Performance for Small Firms. **Marketing Letters**, v. 11, n. 2, p. 151–163, 1 maio 2000.
- AL-SAFRAN, E. Investigation and prediction of slug frequency in gas/liquid horizontal pipe flow. **Journal of Petroleum Science and Engineering**, v. 69, n. 1–2, p. 143–155, nov. 2009.
- AMARI, S.; WU, S. Improving support vector machine classifiers by modifying kernel functions. **Neural Networks**, v. 12, n. 6, p. 783–789, 1 jul. 1999.
- ANP. **Tabela de Poços 2019**, 6 fev. 2019a. Disponível em: <<http://www.anp.gov.br/arquivos/dados-tecnicos/acd/2019-tabela-de-pocos.xlsx>>. Acesso em: 15 jun. 2019
- ANP. **Boletim da Produção de Petróleo e Gás Natural**. [s.l.] ANP, 31 maio 2019b.
- API. **Impacts Of The Natural Gas And Oil Industry On The Us Economy In 2015**. EUA: API, 2017.
- ATS. **Four Ways Predictive Maintenance Can Cut Costs and Improve Productivity**, 2018. Disponível em: <<https://www.advancedtech.com/wp-content/uploads/2019/06/Four-Ways-Predictive-Maintenance-Can-Cut-Costs-and-Improve-Productivity.pdf>>. Acesso em: 15 jun. 2019

- BAI, Y.; BAI, Q. **Subsea Engineering Handbook**. [s.l.] Gulf Professional Publishing, 2018.
- BANSAL, S. K.; KAGEMANN, S. Integrating Big Data: A Semantic Extract-Transform-Load Framework. **Computer**, v. 48, n. 3, p. 42–50, mar. 2015.
- BASTOS, I. C. Eficiência Energética em Sistemas de Elevação Artificial – Bombeio Mecânico. **HOLOS**, v. 1, n. 0, p. 255–263, 18 jul. 2017.
- BASU, J. K.; BHATTACHARYYA, D.; KIM, T. Use of artificial neural network in pattern recognition. **International journal of software engineering and its applications**, v. 4, n. 2, 2010.
- BELGACEM, I. et al. **Experimental investigation of slug pattern in a horizontal two-phase flow**. In: MULTIPHASE FLOW 2013. A Coruña, Spain: 3 jul. 2013 Disponível em: <<http://library.witpress.com/viewpaper.asp?pcode=MPF13-035-1>>. Acesso em: 10 dez. 2019
- BROWN, K. E. Technology of artificial lift methods. Volume 1. Inflow performance, multiphase flow in pipes, the flowing well. 1 jan. 1977.
- DI BITETTI, M. S.; FERRERAS, J. A. Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. **Ambio**, v. 46, n. 1, p. 121–127, 1 fev. 2017.
- EL-ABBASY, M. S. et al. Artificial neural network models for predicting condition of offshore oil and gas pipelines. **Automation in Construction**, v. 45, p. 50–65, set. 2014.
- EPE. **Balanco Energético Nacional 2018: Ano base 2017**. Rio de Janeiro: EPE, 2018.
- GALVÃO, T. F.; PANSANI, T. DE S. A. Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA. **Epidemiologia e Serviços de Saúde**, v. 24, n. 2, p. 335–342, jun. 2015.
- GALVÃO, T. F.; PEREIRA, M. G. Revisões sistemáticas da literatura: passos para sua elaboração. **Epidemiologia e Serviços de Saúde**, v. 23, n. 1, p. 183–184, mar. 2014.
- GANTZ, J.; REINSEL, D. Extracting Value from Chaos. p. 12, 2011.
- GANTZ, J.; REINSEL, D. THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. p. 16, 2012.
- GE AUTOMATION. **The Rise of Industrial Big Data**. Disponível em: <<http://www.geautomation.com/download/rise-industrial-big-data>>. Acesso em: 8 nov. 2019.
- GEREVINI, G. et al. **Increase of Oil Production in Offshore Production Systems through NMPC Control**. 8 out. 2016
- GOUVEIA, F. Tecnologia nacional para extrair petróleo e gás do pré-sal. **Conhecimento & Inovação**, v. 6, n. 1, p. 30–35, 2010.
- GOVERNO DO BRASIL. **Setor de petróleo e gás chega a 13% do PIB brasileiro**. Disponível em: <<http://www.brasil.gov.br/economia-e-emprego/2014/06/setor-de-petroleo-e-gas-chega-a-13-do-pib-brasileiro>>. Acesso em: 14 jun. 2019.

GUO, D. et al. **Data driven approach to failure prediction for electrical submersible pump systems**. SPE West. Reg. Meet.: Old Horizons, New Horizons Enabling Technol. **Anais...**Society of Petroleum Engineers, 2015Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84960337869&partnerID=40&md5=0d287ddaaf9a34f98e0f02727fdbbfe6>>

HAN, Q. et al. **On fault prediction based on industrial big data**. 2017 36th Chinese Control Conference (CCC). **Anais...** In: 2017 36TH CHINESE CONTROL CONFERENCE (CCC). jul. 2017

IEA. **World Energy Balances 2019**. Disponível em: <<https://webstore.iea.org/world-energy-balances-2019>>. Acesso em: 20 jan. 2020.

IEEE. **About IEEE Xplore**. Disponível em: <<https://ieeexplore.ieee.org/Xplorehelp/#/overview-of-ieee-xplore/about-ieee-xplore>>. Acesso em: 12 nov. 2019.

INKPEN, A. C.; MOFFETT, M. H. **The Global Oil & Gas Industry: Management, Strategy & Finance**. [s.l.] PennWell Books, 2011.

INTERNATIONAL ENERGY AGENCY. **WORLD ENERGY BALANCES 2018**. Place of publication not identified: ORGANIZATION FOR ECONOMIC, 2018.

JALALI, A. et al. **Predicting time-to-failure of plasma etching equipment using machine learning**. IEEE Int. Conf. Progn. Health Manag., ICPHM. **Anais...**Institute of Electrical and Electronics Engineers Inc., 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072772020&doi=10.1109%2fICPHM.2019.8819404&partnerID=40&md5=badc7477077c62e549bb902519aaf8b0>>

LEAHY, K. et al. **Diagnosing wind turbine faults using machine learning techniques applied to operational data**. 2016 IEEE International Conference on Prognostics and Health Management (ICPHM). **Anais...** In: 2016 IEEE INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT (ICPHM). jun. 2016

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33–47, 2018a.

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33–47, 1 ago. 2018b.

MARCORIN, W. R.; LIMA, C. R. C. Análise dos Custos de Manutenção e de Não-manutenção de Equipamentos Produtivos. v. 11, p. 8, 2003.

MOHAMMADZAHERI, M. et al. An intelligent approach to optimize multiphase subsea oil fields lifted by electrical submersible pumps. **Journal of Computational Science**, v. 15, p. 50–59, jul. 2016.

MONDAL, K. C.; BARUA, H. B. **Fault Analysis and Trend Prediction in Telecommunication Using Pattern Detection: Architecture, Case Study and Experimentation**. [s.l.] Springer Verlag, 2019. v. 1030

MUNION, M. C. **Predictive vs. preventive: The debate—and the future**. Disponível em: <<https://www.plantengineering.com/articles/predictive-vs-preventive-the-debate-and-the-future/>>. Acesso em: 27 dez. 2018.

NASCIMENTO, J. M. A. DO. Simulador computacional para poços de petróleo com método de elevação artificial por bombeio mecânico. 11 mar. 2005.

OKOLI, P.; VEGA, J. C.; SHOR, R. **Estimating downhole vibration via machine learning techniques using only surface drilling parameters**. SPE West. Reg. Meet. Proc. **Anais...Society of Petroleum Engineers (SPE)**, 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066314507&partnerID=40&md5=a61e71ee673b18e1b2481072ee69fbdb>>

OUADAH, A. **Pipeline Defects Risk Assessment Using Machine Learning and Analytical Hierarchy Process**. Proc. Int. Conf. Appl. Smart Syst., ICASS. **Anais...Institute of Electrical and Electronics Engineers Inc.**, 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063503902&doi=10.1109%2fICASS.2018.8651970&partnerID=40&md5=e7cf9cf9b2d5a24347cc1d7526929c0d>>

QIANG, S.; CHENG, S.; JING, K. **Hybrid SVM/HMM Method for Tool Wear Intelligence Measure in Cutting Process**. 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering. **Anais... In: 2010 3RD INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING**. nov. 2010

RICCOMINI, C.; SANT'ANNA, L. G.; TASSINARI, C. C. G. Pré-sal: geologia e exploração. **Revista USP**, v. 0, n. 95, p. 33, 30 nov. 2012.

SAMPAIO, R.; MANCINI, M. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Revista Brasileira de Fisioterapia**, v. 11, n. 1, p. 83–89, fev. 2007.

SCHEFFER, C.; GIRDHAR, P. **Practical Machinery Vibration Analysis and Predictive Maintenance**. [s.l.] Elsevier, 2004.

SCHUTTE, G. R. PANORAMA DO PRÉ-SAL: DESAFIOS E OPORTUNIDADES. p. 79, 2012.

SCOPUS. **The largest database of peer-reviewed literature - Scopus | Elsevier Solutions**. Disponível em: <<https://www.elsevier.com/solutions/scopus>>. Acesso em: 12 nov. 2019.

SHIRZAD, A.; TABESH, M.; FARMANI, R. A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. **KSCE Journal of Civil Engineering**, v. 18, n. 4, p. 941–948, 2014.

SILVA, A. H. M.; DA HORA, H. R. M.; FERNANDES, R. M. **ANÁLISE DE CONFIABILIDADE DE UM SISTEMA DE CONTROLE DE POÇOS SUBMARINOS DE PETRÓLEO**. Campos dos Goytacazes: Instituto Federal de Educação, Ciência e Tecnologia Fluminense - Programa de Pós-graduação em Sistemas Aplicados à Engenharia e Gestão, 2017.

SNEED, J. **Predicting ESP lifespan with machine learning**. SPE/AAPG/SEG Unconv. Resour. Technol. Conf. **Anais...Unconventional Resources Technology Conference (URTEC)**,

2017Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85051671524&doi=10.15530%2furtec-2017-2669988&partnerID=40&md5=106107473da0ad3ecc9960269c2573e6>>

SUGUMARAN, V.; MURALIDHARAN, V.; RAMACHANDRAN, K. I. Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing. **Mechanical Systems and Signal Processing**, v. 21, n. 2, p. 930–942, fev. 2007.

SWANSON, L. Linking maintenance strategies to performance. **International Journal of Production Economics**, v. 70, n. 3, p. 237–244, 18 abr. 2001.

TAITEL, Y.; DUKLER, A. E. A model for slug frequency during gas-liquid flow in horizontal and near horizontal pipes. **International Journal of Multiphase Flow**, v. 3, n. 6, p. 585–596, 1 dez. 1977.

VASSILIADIS, P. A Survey of Extract–Transform–Load Technology: **International Journal of Data Warehousing and Mining**, v. 5, n. 3, p. 1–27, jul. 2009.

VELLOSO, H. M.; DA HORA, H. R. M. Utilização De Técnicas De Machine Learning Na Previsão De Falhas Em Equipamentos: Uma Revisão Sistemática. 2020.

VELLOSO, H. M.; DA HORA, H. R. M.; CARVALHO, R. A. DE. **Inteligência Artificial Para Detecção De Falhas Em Equipamentos: Um Estudo Bibliométrico No Setor De Óleo E Gás**. . In: XV SEGET. Penedo/RJ: 31 nov. 2018

WANG, G. et al. **AdaBoost and Least Square Based Failure Prediction of Railway Turnouts**. Proc. - Int. Symp. Comput. Intell. Des., ISCID. **Anais...Institute of Electrical and Electronics Engineers Inc.**, 2017Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85013642211&doi=10.1109%2fISCID.2016.1107&partnerID=40&md5=fc6a59069dd96d38df3e42a8dd91ac5a>>

WANG, L. **Support Vector Machines: Theory and Applications**. [s.l.] Springer Science & Business Media, 2005.

YAO, Y. et al. **Fault Prognosis Method of Industrial Process Based on PSO-SVR**. [s.l.] Springer Verlag, 2019. v. 1043

ZABARAS, G. J. **Prediction of Slug Frequency for Gas-Liquid Flows**. SPE Annual Technical Conference and Exhibition. **Anais...** In: SPE ANNUAL TECHNICAL CONFERENCE AND EXHIBITION. Houston, Texas: Society of Petroleum Engineers, 1999Disponível em: <<http://www.onepetro.org/doi/10.2118/56462-MS>>. Acesso em: 11 set. 2019

ZHOU, D. P.; HU, Q.; TOMLIN, C. J. **Quantitative comparison of data-driven and physics-based models for commercial building HVAC systems**. 2017 American Control Conference (ACC). **Anais...** In: 2017 AMERICAN CONTROL CONFERENCE (ACC). maio 2017

5 ARTIGO C - PREVISÃO DE SLUGS EM PIPELINES DE PRODUÇÃO DE ÓLEO E GÁS - UM MODELO UTILIZANDO SVM

5.1 Resumo

A indústria do petróleo e gás é uma das maiores, mais complexas e importantes indústrias globais, sendo responsável por 54% de toda a energia primária fornecida globalmente em 2017 e tendo altos custos para extração e produção. A utilização de bombas, chamadas ESP, é indicada para aumentar a produção de campos de petróleo, principalmente em campos de produção de fluidos multifásicos. A fim de garantir uma maior vida útil para as ESPs, faz-se importante prever a formação de bolsões de gás chamados *slugs*. Diversos modelos já foram criados a fim de realizar essa predição, porém a comunidade científica deu maior ênfase para os modelos *physics-based* de alta dimensionalidade e custo computacional. Assim, este trabalho propõe a criação de um modelo *data-driven* utilizando a ferramenta de *Support Vector Machines* (SVM). Foram coletados 172.800 *data points* de um campo de petróleo localizado na costa do Brasil. Os dados representam a produção de quatro poços e de uma ESP. Após a coleta, os mesmos passaram por etapas de tratamento para então serem submetidos à técnica de SVM utilizando três *kernels* diferentes: linear, rbf e polinomial. Os resultados encontrados mostram que o modelo possui uma performance melhor do que os modelos encontrados na literatura, sendo o modelo com *kernel* rbf o de melhor performance, com uma acurácia geral de 99,76%, seguido do *kernel* polinomial, acurácia de 99,63%, e do *kernel* linear, acurácia de 98,30%. Este trabalho se mostrou de grande valia para a indústria, sendo capaz de prever *slugs* e ainda fazer simulações de situações futuras.

Palavras-chave: petróleo e gás, ESP, predição de falhas, SVM

5.2 Abstract

The oil and gas industry is one of the biggest, more complex and important global industries, being responsible for 54% of the entire primary energy supplied globally in 2017 and recognized by its high extraction and production cost. The using of pumps, called ESPs, is suitable to increase the oil filed production, especially in production field with multiphase fluids. To ensure a greater life to the ESPs, it is important to predict gas slugs. Many models have been created to do this prediction, but the scientific community gave more emphasis to physics-based models with high dimensionality and computational cost. Therefore, this paper aims to create a data-driven model using the Support Vector Machine (SVM) tool. 172,800 data points were collected from an offshore oil field located in Brasil coast. The data represents the production of four wells and an ESP. After the data collection, the data were submitted to data treatment steps and then to the SVM technique using three different kernels: linear, rbf and polynomial. The findings show that the model outperforms the models found in literature, where

the rbf kernel demonstrated to have the best performance, with an accuracy score of 99.76%, followed by polynomial kernel, accuracy of 99.63%, and linear kernel, accuracy of 98.30%. This paper reached its goal and showed itself to be of great value to industry, being able to predict slugs and simulate future situations.

Keywords: oil and gas, ESP, failure prediction, SVM.

5.3 Introdução

A indústria do petróleo e gás é uma das maiores, mais complexas e importantes indústrias globais (INKPEN; MOFFETT, 2011), sendo responsável por 54% de toda a energia primária fornecida globalmente em 2017 (IEA, 2019). Além dessa relevância, os valores para a extração do petróleo e do gás são protagonistas nessa indústria. A perfuração de um único poço na camada pré-sal custa US\$ 100 milhões (GOUVEIA, 2010), além dos custos de instalação dos equipamentos para a produção.

Dado o alto valor econômico da produção de petróleo e gás, o Brasil busca cada vez mais reduzir os custos e tempo para a perfuração de poços (GOUVEIA, 2010) assim como aumentar a produção de petróleo, principalmente de poços submarinos (GEREVINI *et al.*, 2016). Porém, ao longo da vida produtiva de um poço, sua pressão interna, que o faz produzir de forma natural, começa a cair, causando uma diminuição do fluxo que chega à superfície (BASTOS, 2017). A fim de manter a produção, são aplicados métodos de elevação artificial que podem ser separados em três classificações: (i) utilização de bomba de fundo, (ii) injeção contínua de gás comprimido no fundo do poço para aumentar a pressão no reservatório e (iii) injeção de gás a alta pressão na coluna de produção para diminuir a densidade do fluido (NASCIMENTO, 2005).

O método de elevação através de bombas de fundo, também conhecidas como bombas elétricas submarinas ou ESP (sigla para *electric submersible pump*), se apresenta como uma solução viável por ocupar um espaço reduzido, ser altamente eficiente e resistente em condições severas (BROWN, 1977), além de ser um método ainda mais apropriado para casos de fluidos multifásicos (BAI; BAI, 2018) e tem a ESP como componente de maior criticidade, uma vez que ela fica instalada no leito marinho. Uma configuração comumente encontrada nos campos de produção que possuem ESPs é formada por diversos poços que possuem uma válvula reguladora de vazão (*choke*) e enviam sua produção para uma bomba, como mostrado na Figura 5-1.

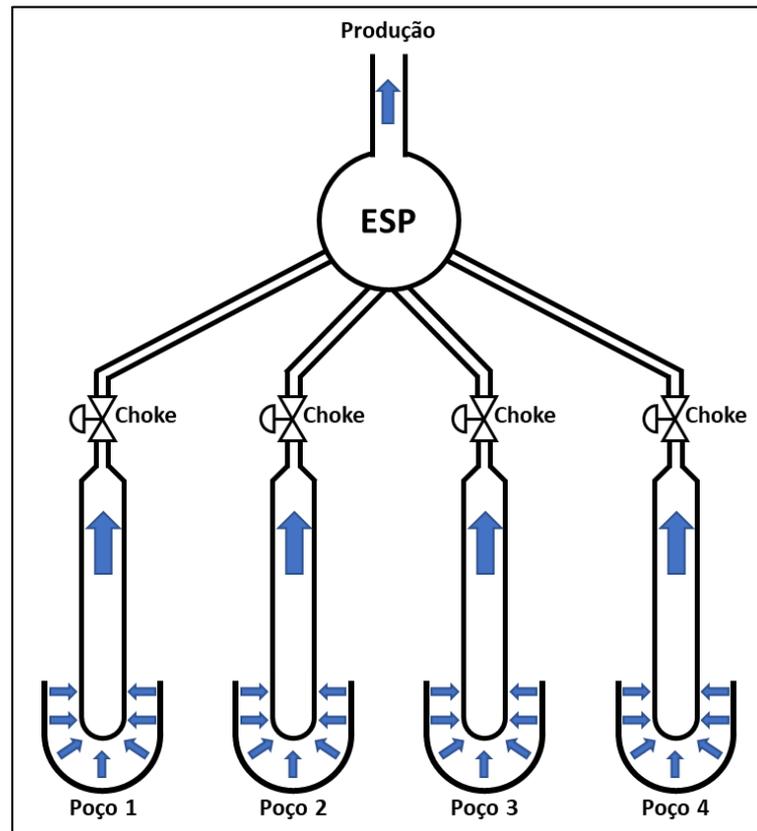


Figura 5-1 - Esquema de arranjo submarino.

Fonte: Dos autores (2020).

Em sistemas com presença de fluidos multifásicos, apesar das ESPs serem o método de elevação artificial mais indicado, a fração volumétrica de gás (GVF) se apresenta como uma variável de grande importância, pois essas bombas são projetadas tendo um valor máximo de GVF como premissa de projeto. Valores acima do limite operacional da bomba podem acarretar em mudanças bruscas no torque (BAI; BAI, 2018) e sobreaquecimento, diminuindo assim a vida útil do equipamento. Porém, mesmo que a bomba seja compatível com os valores de GVF de uma determinada produção de petróleo, o gás produzido pode se acumular nas linhas de fluxo horizontais que interligam os poços e a bomba, formando bolsões de gás conhecidos na indústria como *slugs* (BELGACEM *et al.*, 2013). Ao chegar na bomba, o *slug* eleva o valor de GVF do fluido que está sendo bombeado, acarretando em um possível dano no equipamento. Para evitar danos ao equipamento, um sistema de proteção pode ser configurado para que, ao atingir determinada temperatura de operação, a bomba desligue de forma automática.

A fim de prever quando um *slug* irá acontecer, Zabarás (1999) apresenta métodos de predição, entre correlações e modelos numéricos, já consolidados e utiliza o modelo de frequência mecânica de *slugs* de Taitel e Dukler (1977) para 399 pontos de dados, sendo 194 pontos coletados da literatura e 205 de uma instalação da Shell's Westhollow Technology Center. A correlação encontrada apresentou um erro médio de 0% e um erro médio absoluto de

60%, representando uma melhora significativa na precisão de predição de *slugs*. Entretanto, o autor descreve o resultado de erro médio como sendo um resultado enviesado.

Em Belgacem *et al.* (2013), um fluxo multifásico contendo líquido e gás dentro de um tubo de 14m de comprimento e 40mm de diâmetro interno foi gerado a fim de estudar o processo de formação de *slugs*. Os dados de formação do *slug*, variação de pressão e frequência e comprimento do *slug* foram captados através de sensores acoplados ao tubo. Nos resultados os dados coletados foram confrontados com a velocidade superficial do fluido e do gás, tendo a frequência atingido um valor de até quatro *slugs* por segundo e o comprimento variado de 15d a 36d. Por fim, foi observado que a frequência dos *slugs* é invariante quanto a pressão mas é linearmente proporcional à velocidade superficial do líquido e que o comprimento do *slug* diminui conforme a velocidade superficial do gás aumenta.

Al-Safran (2009) investigou fenomenologicamente os parâmetros que controlam a frequência de *slugs* formados em fluxos de duas fases em tubos horizontais. Assim como em Zabaraz (1999), foram apresentados os modelos de predição de *slugs* já presentes na literatura e realizado um experimento com um circuito de tubos de 420m de comprimento e 50,8mm diâmetro na Tulsa University Fluid Flow Projects (TUFPF). Com os dados do experimento, foi possível identificar características relevantes sobre os mecanismos de formação dos *slugs* no início do *pipeline* e o desenvolvimento durante o fluxo do fluido dentro do *pipeline*. Após a coleta e análise dos dados, foram identificadas as correlações entre as variáveis do processo, o que permitiu a criação de um novo modelo de correlação. Para tal, foram utilizados 230 *data points* coletados de estudos já publicados, sendo 90% para teste e 10% para treino. Inicialmente, o modelo obteve um erro médio percentual (APE) de 35%, um erro médio absoluto percentual (AAPE) de 47% e um desvio padrão (SD) de 68%, mas, após um ajuste utilizando intervalos de confiança estatísticos dos coeficientes, foram alcançados os resultados de 19% APE, 43% AAPE e 62% SD. Apesar do SD não ter uma redução significativa e permanecer com um valor alto, o APE é o melhor valor e o AAPE atingiu um valor nivelado, se comparados aos outros modelos encontrados na literatura.

Todos os trabalhos selecionados apresentam modelos analíticos e numéricos já consolidados na academia, que são *physics-based*, ou seja, baseados nas propriedades físicas do *pipeline* e a forma como ele foi instalado. Apesar desses modelos normalmente serem altamente acurados, eles possuem alta dimensionalidade e são computacionalmente custosos (ZHOU; HU; TOMLIN, 2017). Na contramão dos modelos *physics-based*, os modelos *data-driven*, baseados apenas nos dados capturados no processo, visam baixa dimensionalidade e diminuir a

complexidade computacional, podendo substituir os modelos *physics-based* com uma perda de acurácia insignificante (ZHOU; HU; TOMLIN, 2017).

Uma das formas de criação de um modelo *data-driven* é a utilização de ferramentas de inteligência artificial (IA), por serem capazes de lidar com um grande volume de dados e terem se mostrado promissoras na detecção e reconhecimento de padrões, onde o sistema deve observar o ambiente, distinguir sobre quais variáveis possuem padrão e tomar decisões razoáveis (BASU; BHATTACHARYYA; KIM, 2010; LIU *et al.*, 2018b).

Dada a literatura apresentada neste trabalho, nota-se que, apesar dos modelos de previsão de *slugs* serem estudados e desenvolvidos por décadas, não foi encontrado um modelo de predição utilizando ferramentas de inteligência artificial. Assim, este trabalho tem como objetivo desenvolver um algoritmo de previsão de *slugs* para um sistema de bombeio submarino utilizando ferramentas de inteligência artificial com base nos dados de produção dos poços.

O presente trabalho justifica-se devido a quantidade cada vez maior de poços que irão necessitar de métodos de elevação artificial como as ESPs, uma vez que sua vazão de produção diminui ao longo de sua vida útil.

5.4 Metodologia

5.4.1 COLETA DE DADOS

Neste trabalho, foram utilizados dados de produção de uma ESP instalada em um campo submarino de produção de petróleo na costa do Brasil. No período coletado, a bomba estava interligada a quatro poços produtores que serão identificados a partir deste ponto como poços A, B, C, e D. Foram coletados dados de produção de um período de oito dias, entre os datas 30 de maio de 2019 e 6 de junho de 2019, com uma taxa de 15 amostras/min. Este período foi escolhido por ter sido o espaço de tempo com maior frequência de geração de *slugs* (cinco *slugs*).

As variáveis coletadas dos poços foram: (i) valor de GVF da produção do poço em porcentagem e (ii) porcentagem de abertura da válvula *choke*. Já da ESP, foi coletada a variável de (iii) frequência do controlador de velocidade da bomba em hertz (Hz). O Quadro 5-1 apresenta um resumo das variáveis coletadas. No total, 172.800 *data points* foram coletados.

Quadro 5-1 - Variáveis coletadas

Identificação	Variável	Fonte	Unidade	Tipo de Variável
(i)	GVF da produção	Poços	%	Numérica
(ii)	Abertura do <i>choke</i>	Poços	%	Numérica
(iii)	Frequência do controlador de velocidade da bomba	ESP	%	Numérica

Fonte: Dos autores (2020).

5.4.2 TRATAMENTO DE DADOS

O tratamento ou transformação dos dados é a etapa logo após a coleta, onde são utilizadas regras para limpeza, normalização, ordenação e agrupamento a fim de passar o banco de dados para a estrutura desejada (BANSAL; KAGEMANN, 2015). Segundo Vassiliadis (2009), este passo é importante para remover o "ruído" presente nos dados operacionais que vão desde erros de digitação em atributos textuais a valores inconsistentes.

O primeiro tratamento aplicado foi a adição de um novo atributo (iv) para contabilizar a velocidade da bomba em RPM. O valor deste atributo é do tipo numérico real dependente do atributo (iii), conforme demonstrado na Equação 5-1.

Equação 5-1 - Cálculo dos valores presentes na coluna (v)

$$velocidade\ da\ bomba\ (RPM) = Freq.do\ controlador\ de\ velocidade\ (Hz) * 60$$

Fonte: Dos autores (2020).

A fim de determinar se houve um desligamento da bomba devido à ocorrência de *slugs*, um novo atributo (v) foi adicionada. Este atributo é baseado no valor do atributo (iv), pois considera-se que a bomba está operacional apenas quando sua velocidade de rotação é superior a 1600 RPM. Assim os valores categóricos serão: *RUNNING* para quando a bomba estiver funcionando, e *NOT RUNNING* para quando a bomba estiver parada, conforme Quadro 5-2.

Quadro 5-2 - Resumo dos valores da coluna (v)

Valor	Condição
<i>RUNNING</i>	> 1600 RPM
<i>NOT RUNNING</i>	< 1600 RPM

Fonte: Dos autores (2020).

Visando diminuir o esforço computacional empregado para processar este banco de dados, os atributos (iii) e (iv) foram excluídos, já que foram utilizadas apenas como suporte para a criação do atributo (v). Assim, o novo formato do banco de dados possui 9 atributos, conforme demonstrado no Quadro 5-3.

Quadro 5-3 - Resumo das colunas do banco de dados

Id.	Variável	Unidade	Tipo de Variável
(a)	Estado da bomba	N/A	Catégorica
(b)	GVF do poço A	%	Numérica
(c)	Abertura do <i>choke</i> do poço A	%	Numérica
(d)	GVF do poço B	%	Numérica
(e)	Abertura do <i>choke</i> do poço B	%	Numérica
(f)	GVF do poço C	%	Numérica
(g)	Abertura do <i>choke</i> do poço C	%	Numérica
(h)	GVF do poço D	%	Numérica
(i)	Abertura do <i>choke</i> do poço D	%	Numérica

Fonte: Dos autores (2020).

Uma vez que todos os dados necessários já estejam presentes no banco de dados, o passo seguinte se dá através de um ajuste na organização dos dados. Como o objetivo do trabalho se baseia na previsão de chegada de *slugs* na bomba utilizando as informações de produção dos poços, faz-se necessário que cada uma das linhas do banco de dados represente as variáveis da bomba em um determinado instante e as variáveis do poço que levaram àquela situação. Considerando-se o diâmetro $d = 8''$ das linhas que interligam os poços à bomba e que cada poço fica a uma distância específica D da bomba e possui um vazão média de Q , conforme Quadro 5-4, é possível calcular qual é o volume V a ser deslocado por cada poço para chegar à bomba (equações $A_{seção} = \frac{\pi d^2}{4}$ (5-2)

$$V = A_{seção} \times D \quad (5-3)$$

$$T = \frac{V}{Q} \quad (5-4)$$

e **Error! Reference source not found.**), quanto tempo T o poço leva para produzir este volume (**Error! Reference source not found.**) e, conseqüentemente, qual é a diferença de amostras $\Delta_{amostras}$ entre os poços e a bomba (Equação 5-5). Os valores encontrados estão presentes no Quadro 5-5.

Quadro 5-4 - Distâncias e vazão dos poços

Poço	D (m)	Q (m ³ /h)
Poço A	70	27,62
Poço B	70	18,00
Poço C	5594	22,19
Poço D	5594	24,50

Fonte: Dos autores (2020).

$$A_{seção} = \frac{\pi d^2}{4} \quad (5-2)$$

$$V = A_{seção} \times D \quad (5-3)$$

$$T = \frac{V}{Q} \quad (5-4)$$

Equação 5-5 - Diferença de amostras entre o poço e a bomba

$$\Delta_{amostras} = \frac{T \times 3600}{4}$$

Fonte: Dos autores (2020).

Quadro 5-5 - Dados dos poços

Poço	D (m)	Q (m³/h)	V (m³)	T (h)	$\Delta_{amostras}$
Poço A	70	27,62	2,24	0,08	73
Poço B	70	18,00	2,24	0,12	112
Poço C	5594	22,19	179,01	8,06	7259
Poço D	5594	24,50	179,01	7,31	6577

Fonte: Dos autores (2020).

Com todas as informações já apresentadas, nota-se que o poço C é o poço que possui maior $\Delta_{amostras}$. Assim, os dados dos outros poços e da bomba devem ser adiantados uma quantidade de amostras definida pela Equação 5-6, para que cada linha do banco de dados possa representar uma situação de causa (poços) e efeito (bomba). O Quadro 5-6 apresenta quantas amostras devem ser cronologicamente corrigidas para cada poço e bomba.

Equação 5-6 - Quantidade de amostras a serem cronologicamente corrigidas

$$qtd\ amostras = \max(\Delta_{amostras}) - \Delta_{amostras}$$

Fonte: Dos autores (2020).

Quadro 5-6 - Quantidade de amostras corrigidas para cada poço e bomba

Poço	Qtd. de amostras a serem adiantadas
Poço A	7186
Poço B	7147
Poço C	0
Poço D	682
Bomba	7259

Fonte: Dos autores (2020).

Após as transformações, as variáveis referentes ao poço C foram as únicas que seguiram sem alteração e as variáveis da bomba as mais impactadas, uma vez que tiveram um maior número de amostras adiantadas, resultando em uma maior quantidade de linhas sem valores.

Para manter um banco de dados sem dados em branco, as últimas 7.259 linhas serão desconsideradas, resultando em um banco de dados com 165.541 *data points*.

5.4.3 APLICAÇÃO DA FERRAMENTA DE INTELIGÊNCIA ARTIFICIAL

Segundo Velloso e Hora (2020), as técnicas de IA podem ser separadas entre as técnicas de classificação e regressão, sendo a técnica de *Support Vector Machine* (SVM) a mais utilizada na academia para as duas categorias. Considerando a relevância desta técnica e que ela é conhecida por ser um aproximador universal de qualquer função multivariável para qualquer grau de acurácia (WANG, 2005), ela será aplicada neste trabalho.

Segundo Amari e Wu (1999), ao contrário de outros métodos que visam minimizar os erros durante o treinamento, o SVM visa separar os conjuntos de dados através de hiperplanos e maximizar a distância entre o hiperplano e os dados. Amari e Wu (1999) também afirmam que a performance dessa técnica é dependente da escolha de seu *kernel*, uma vez que, escolhendo diferentes tipos de *kernel*, o SVM pode realizar classificações do tipo, *Radial Basis Function* (RBF) e polinomial, por exemplo. Com a finalidade de encontrar o melhor modelo utilizando o SVM, serão gerados três modelos diferentes utilizando os *kernels* linear, polinomial e rbf. Todos os outros parâmetros serão mantidos para todos os *kernels*: $C=1$, $degree = 3$, $gamma = 'scale'$ e $coef0 = 0$.

Como entrada para o algoritmo, serão utilizadas os atributos de (b) a (i), conforme Quadro 5-3, e o atributo (a) como desfecho desejado. Para treino e validação do modelo desenvolvido, os dados serão divididos na proporção 70%/30%, respectivamente. Este método de validação foi escolhido devido ao baixo custo computacional envolvido se comparado com outros métodos como o *cross-validation* para bases de dados muito grandes.

5.5 Resultados e Discussões

Os modelos gerados para todos os *kernels* apresentam bom desempenho, tendo o *kernel* rbf atingido um resultado ligeiramente melhor do que os outros dois. O modelo utilizando rbf atingiu uma acurácia geral de 99,76%, seguido do *kernel* polinomial (99,63%) e do *kernel* linear (98,30%). Além dos altos desempenhos gerais, os modelos obtiveram bons resultados de precisão, sensibilidade e *f1 score*. Os resultados de classificação para cada modelo podem ser encontrados no Quadro 5-7 e as matrizes de confusão estão apresentadas nos quadros Quadro 5-8, Quadro 5-9 e Quadro 5-10.

Quadro 5-7 - Resultado das classificações

<i>Kernel</i>	Classificação	Precisão	Sensibilidade	<i>F1 score</i>
linear	NOT RUNNING	94,57%	92,01%	93,27%
	RUNNING	98,83%	99,22%	99,02%
rbf	NOT RUNNING	99,23%	98,93%	99,08%
	RUNNING	99,84%	99,89%	99,86%
polinomial	NOT RUNNING	98,94%	98,18%	98,56%
	RUNNING	99,73%	99,85%	99,79%

Fonte: Dos autores (2020).

Quadro 5-8 - Matriz de confusão para o *kernel* linear

	NOT RUNNING	RUNNING
NOT RUNNING	5865	509
RUNNING	337	42952

Fonte: Dos autores (2020).

Quadro 5-9 - Matriz de confusão para o *kernel* rbf

	NOT RUNNING	RUNNING
NOT RUNNING	6306	68
RUNNING	49	43240

Fonte: Dos autores (2020).

Quadro 5-10 - Matriz de confusão para o *kernel* polinomial

	NOT RUNNING	RUNNING
NOT RUNNING	6258	116
RUNNING	67	43222

Fonte: Dos autores (2020).

Apesar de desempenhos semelhantes, os *kernels* linear e rbf se destacam devido ao alto custo computacional demandado para a geração do modelo polinomial. Além de ter uma acurácia menor do que o rbf, o modelo polinomial performou melhor do que o modelo linear em apenas 1,33% dos casos, não justificando o alto tempo de processamento. Esta desvantagem se torna ainda mais significativa quando aplicada em um sistema que pode sofrer alterações nos seus parâmetros ao longo da operação, demandando assim a atualização do modelo de predição.

Em Zabarás (1999), Belgacem *et al.* (2013) e Al-Safran (2009), todos os modelos de predição possuem como objetivo encontrar o valor da frequência de ocorrência de *slugs*, objetivo que vai no sentido contrário a este trabalho que gera um modelo de predição baseado em classificação.

Belgacem *et al.* (2013) apresentou seu resultado analisando apenas a correlação das variáveis do processo, impossibilitando assim uma comparação direta quanto ao modelo

proposto neste trabalho. Já em Zabaras (1999) e Al-Safran (2009), são apresentados os valores de 0% e 19% de erro médio percentual, respectivamente. Uma vez que o resultado do primeiro artigo é considerado enviesado, todos os modelos propostos neste trabalho performam melhor do que os modelos *physics-based* encontrado na literatura.

Além de um melhor desempenho, o modelo proposto, utilizando os *kernels* linear e rbf, apresentou baixo custo computacional, mesmo utilizando uma quantidade de dados muito superior à utilizada por Zabaras (1999) e Al-Safran (2009), 115.879 *data points* contra 399 e 207, respectivamente.

5.6 Conclusão

Este trabalho se mostra de grande valia para a indústria, uma vez que a predição de *slugs* dá a possibilidade de que manobras sejam realizadas a fim de manter a ESP funcionando sem anomalias e a produção ininterrupta de óleo. A janela de tempo de predição vai depender da distância dos poços até a bomba, da vazão dos poços e qual deles irá colaborar para a geração do *slug*. Nos dados utilizados neste trabalho, a previsão pode ter uma janela de tempo de até 8h.

Além de conseguir prever o estado futuro da formação de *slugs*, o modelo apresentado neste trabalho auxilia na simulação de cenários para possíveis mudanças dos parâmetros do poço, como aumento ou diminuição da vazão através da abertura e fechamento dos *chokes*, podendo estar presente em sistemas online que fazem a leitura dos parâmetros de campo em tempo real.

Como trabalho futuro, outras ferramentas de classificação podem ser utilizadas e seus desempenho comparados. Uma outra opção é recolher uma maior quantidade de variáveis e realizar um estudo de quais são as variáveis que mais tem relevância na formação de *slugs*.

5.7 Referências

- ACCORSI, R. et al. Data Mining and Machine Learning for Condition-based Maintenance. **Procedia Manufacturing**, v. 11, p. 1153–1161, 2017.
- ALI, A. The Impact of Innovativeness and Development Time on New Product Performance for Small Firms. **Marketing Letters**, v. 11, n. 2, p. 151–163, 1 maio 2000.
- AL-SAFRAN, E. Investigation and prediction of slug frequency in gas/liquid horizontal pipe flow. **Journal of Petroleum Science and Engineering**, v. 69, n. 1–2, p. 143–155, nov. 2009.
- AMARI, S.; WU, S. Improving support vector machine classifiers by modifying kernel functions. **Neural Networks**, v. 12, n. 6, p. 783–789, 1 jul. 1999.

ANP. **Tabela de Poços 2019**, 6 fev. 2019a. Disponível em: <<http://www.anp.gov.br/arquivos/dados-tecnicos/acd/2019-tabela-de-pocos.xlsx>>. Acesso em: 15 jun. 2019

ANP. **Boletim da Produção de Petróleo e Gás Natural**. [s.l.] ANP, 31 maio 2019b.

API. **Impacts Of The Natural Gas And Oil Industry On The Us Economy In 2015**. EUA: API, 2017.

ATS. **Four Ways Predictive Maintenance Can Cut Costs and Improve Productivity**, 2018. Disponível em: <<https://www.advancedtech.com/wp-content/uploads/2019/06/Four-Ways-Predictive-Maintenance-Can-Cut-Costs-and-Improve-Productivity.pdf>>. Acesso em: 15 jun. 2019

BAI, Y.; BAI, Q. **Subsea Engineering Handbook**. [s.l.] Gulf Professional Publishing, 2018.

BANSAL, S. K.; KAGEMANN, S. Integrating Big Data: A Semantic Extract-Transform-Load Framework. **Computer**, v. 48, n. 3, p. 42–50, mar. 2015.

BASTOS, I. C. Eficiência Energética em Sistemas de Elevação Artificial – Bombeio Mecânico. **HOLOS**, v. 1, n. 0, p. 255–263, 18 jul. 2017.

BASU, J. K.; BHATTACHARYYA, D.; KIM, T. Use of artificial neural network in pattern recognition. **International journal of software engineering and its applications**, v. 4, n. 2, 2010.

BELGACEM, I. et al. Experimental investigation of slug pattern in a horizontal two-phase flow. In: **MULTIPHASE FLOW 2013**. A Coruña, Spain: 3 jul. 2013 Disponível em: <<http://library.witpress.com/viewpaper.asp?pcode=MPF13-035-1>>. Acesso em: 10 dez. 2019

BROWN, K. E. Technology of artificial lift methods. Volume 1. **Inflow performance, multiphase flow in pipes, the flowing well**. 1 jan. 1977.

DI BITETTI, M. S.; FERRERAS, J. A. Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. **Ambio**, v. 46, n. 1, p. 121–127, 1 fev. 2017.

EL-ABBASY, M. S. et al. Artificial neural network models for predicting condition of offshore oil and gas pipelines. **Automation in Construction**, v. 45, p. 50–65, set. 2014.

EPE. **Balanco Energético Nacional 2018: Ano base 2017**. Rio de Janeiro: EPE, 2018.

GALVÃO, T. F.; PANSANI, T. DE S. A. Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA. **Epidemiologia e Serviços de Saúde**, v. 24, n. 2, p. 335–342, jun. 2015.

GALVÃO, T. F.; PEREIRA, M. G. Revisões sistemáticas da literatura: passos para sua elaboração. **Epidemiologia e Serviços de Saúde**, v. 23, n. 1, p. 183–184, mar. 2014.

GANTZ, J.; REINSEL, D. **Extracting Value from Chaos**. p. 12, 2011.

GANTZ, J.; REINSEL, D. **THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East**. p. 16, 2012.

GE AUTOMATION. **The Rise of Industrial Big Data**. Disponível em: <<http://www.geautomation.com/download/rise-industrial-big-data>>. Acesso em: 8 nov. 2019.

GEREVINI, G. et al. **Increase of Oil Production in Offshore Production Systems through NMPC Control**. 8 out. 2016.

GOUVEIA, F. Tecnologia nacional para extrair petróleo e gás do pré-sal. **Conhecimento & Inovação**, v. 6, n. 1, p. 30–35, 2010.

GOVERNO DO BRASIL. **Setor de petróleo e gás chega a 13% do PIB brasileiro**. Disponível em: <<http://www.brasil.gov.br/economia-e-emprego/2014/06/setor-de-petroleo-e-gas-chega-a-13-do-pib-brasileiro>>. Acesso em: 14 jun. 2019.

GUO, D. et al. **Data driven approach to failure prediction for electrical submersible pump systems**. SPE West. Reg. Meet.: Old Horizons, New Horizons Enabling Technol. **Anais...Society of Petroleum Engineers**, 2015 Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84960337869&partnerID=40&md5=0d287ddaaf9a34f98e0f02727fdbbfe6>>

HAN, Q. et al. **On fault prediction based on industrial big data**. 2017 36th Chinese Control Conference (CCC). **Anais... In: 2017 36TH CHINESE CONTROL CONFERENCE (CCC)**. jul. 2017

IEA. **World Energy Balances 2019**. Disponível em: <<https://webstore.iea.org/world-energy-balances-2019>>. Acesso em: 20 jan. 2020.

IEEE. **About IEEE Xplore**. Disponível em: <<https://ieeexplore.ieee.org/Xplorehelp/#/overview-of-ieee-xplore/about-ieee-xplore>>. Acesso em: 12 nov. 2019.

INKPEN, A. C.; MOFFETT, M. H. **The Global Oil & Gas Industry: Management, Strategy & Finance**. [s.l.] PennWell Books, 2011.

INTERNATIONAL ENERGY AGENCY. **WORLD ENERGY BALANCES 2018**. Place of publication not identified: ORGANIZATION FOR ECONOMIC, 2018.

JALALI, A. et al. **Predicting time-to-failure of plasma etching equipment using machine learning**. IEEE Int. Conf. Progn. Health Manag., ICPHM. **Anais...Institute of Electrical and Electronics Engineers Inc.**, 2019 Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072772020&doi=10.1109%2fICPHM.2019.8819404&partnerID=40&md5=badc7477077c62e549bb902519aaf8b0>>

LEAHY, K. et al. Diagnosing wind turbine faults using machine learning techniques applied to operational data. **2016 IEEE International Conference on Prognostics and Health Management (ICPHM)**. jun. 2016

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33–47, 2018a.

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33–47, 1 ago. 2018b.

MARCORIN, W. R.; LIMA, C. R. C. Análise dos Custos de Manutenção e de Não-manutenção de Equipamentos Produtivos. v. 11, p. 8, 2003.

MOHAMMADZAHERI, M. et al. An intelligent approach to optimize multiphase subsea oil fields lifted by electrical submersible pumps. **Journal of Computational Science**, v. 15, p. 50–59, jul. 2016.

MONDAL, K. C.; BARUA, H. B. **Fault Analysis and Trend Prediction in Telecommunication Using Pattern Detection: Architecture, Case Study and Experimentation**. [s.l.] Springer Verlag, 2019. v. 1030

MUNION, M. C. **Predictive vs. preventive: The debate—and the future**. Disponível em: <<https://www.plantengineering.com/articles/predictive-vs-preventive-the-debate-and-the-future/>>. Acesso em: 27 dez. 2018.

NASCIMENTO, J. M. A. DO. Simulador computacional para poços de petróleo com método de elevação artificial por bombeio mecânico. 11 mar. 2005.

OKOLI, P.; VEGA, J. C.; SHOR, R. **Estimating downhole vibration via machine learning techniques using only surface drilling parameters**. SPE West. Reg. Meet. Proc. **Anais...Society of Petroleum Engineers (SPE)**, 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066314507&partnerID=40&md5=a61e71ee673b18e1b2481072ee69fbd&id=1>>

OUADAH, A. **Pipeline Defects Risk Assessment Using Machine Learning and Analytical Hierarchy Process**. Proc. Int. Conf. Appl. Smart Syst., ICASS. **Anais...Institute of Electrical and Electronics Engineers Inc.**, 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063503902&doi=10.1109%2FICASS.2018.8651970&partnerID=40&md5=e7cf9cf9b2d5a24347cc1d7526929c0d&id=1>>

QIANG, S.; CHENG, S.; JING, K. **Hybrid SVM/HMM Method for Tool Wear Intelligence Measure in Cutting Process**. 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering. **Anais... In: 2010 3RD INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING**. nov. 2010

RICCOMINI, C.; SANT'ANNA, L. G.; TASSINARI, C. C. G. Pré-sal: geologia e exploração. **Revista USP**, v. 0, n. 95, p. 33, 30 nov. 2012.

SAMPAIO, R.; MANCINI, M. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Revista Brasileira de Fisioterapia**, v. 11, n. 1, p. 83–89, fev. 2007.

SCHEFFER, C.; GIRDHAR, P. **Practical Machinery Vibration Analysis and Predictive Maintenance**. [s.l.] Elsevier, 2004.

SCHUTTE, G. R. PANORAMA DO PRÉ-SAL: DESAFIOS E OPORTUNIDADES. p. 79, 2012.

SCOPUS. **The largest database of peer-reviewed literature - Scopus | Elsevier Solutions**. Disponível em: <<https://www.elsevier.com/solutions/scopus>>. Acesso em: 12 nov. 2019.

SHIRZAD, A.; TABESH, M.; FARMANI, R. A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. **KSCE Journal of Civil Engineering**, v. 18, n. 4, p. 941–948, 2014.

SILVA, A. H. M.; DA HORA, H. R. M.; FERNANDES, R. M. **ANÁLISE DE CONFIABILIDADE DE UM SISTEMA DE CONTROLE DE POÇOS SUBMARINOS DE PETRÓLEO**. Campos dos Goytacazes: Instituto Federal de Educação, Ciência e Tecnologia Fluminense - Programa de Pós-graduação em Sistemas Aplicados à Engenharia e Gestão, 2017.

SNEED, J. **Predicting ESP lifespan with machine learning**. SPE/AAPG/SEG Unconv. Resour. Technol. Conf. **Anais...Unconventional Resources Technology Conference (URTEC)**, 2017Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85051671524&doi=10.15530%2furtec-2017-2669988&partnerID=40&md5=106107473da0ad3ecc9960269c2573e6>>

SUGUMARAN, V.; MURALIDHARAN, V.; RAMACHANDRAN, K. I. Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing. **Mechanical Systems and Signal Processing**, v. 21, n. 2, p. 930–942, fev. 2007.

SWANSON, L. Linking maintenance strategies to performance. **International Journal of Production Economics**, v. 70, n. 3, p. 237–244, 18 abr. 2001.

TAITEL, Y.; DUKLER, A. E. A model for slug frequency during gas-liquid flow in horizontal and near horizontal pipes. **International Journal of Multiphase Flow**, v. 3, n. 6, p. 585–596, 1 dez. 1977.

VASSILIADIS, P. A Survey of Extract–Transform–Load Technology: **International Journal of Data Warehousing and Mining**, v. 5, n. 3, p. 1–27, jul. 2009.

VELLOSO, H. M.; DA HORA, H. R. M. Utilização De Técnicas De Machine Learning Na Previsão De Falhas Em Equipamentos: Uma Revisão Sistemática. 2020.

VELLOSO, H. M.; DA HORA, H. R. M.; CARVALHO, R. A. DE. **Inteligência Artificial Para Detecção De Falhas Em Equipamentos: Um Estudo Bibliométrico No Setor De Óleo E Gás**. . In: XV SEGET. Penedo/RJ: 31 nov. 2018

WANG, G. et al. **AdaBoost and Least Square Based Failure Prediction of Railway Turnouts**. Proc. - Int. Symp. Comput. Intell. Des., ISCID. **Anais...Institute of Electrical and Electronics Engineers Inc.**, 2017Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85013642211&doi=10.1109%2fISCID.2016.1107&partnerID=40&md5=fc6a59069dd96d38df3e42a8dd91ac5a>>

WANG, L. **Support Vector Machines: Theory and Applications**. [s.l.] Springer Science & Business Media, 2005.

YAO, Y. et al. **Fault Prognosis Method of Industrial Process Based on PSO-SVR**. [s.l.] Springer Verlag, 2019. v. 1043

ZABARAS, G. J. Prediction of Slug Frequency for Gas-Liquid Flows. **SPE Annual Technical Conference and Exhibition**. Houston, Texas: Society of Petroleum Engineers, 1999. Disponível em: <<http://www.onepetro.org/doi/10.2118/56462-MS>>. Acesso em: 11 set. 2019.

ZHOU, D. P.; HU, Q.; TOMLIN, C. J. Quantitative comparison of data-driven and physics-based models for commercial building HVAC systems. **2017 American Control Conference (ACC)**. Maio 2017.

GLOSSÁRIO

A

ANM - Equipamento composto por válvulas e sensores que direcionam o fluxo da produção de um poço de petróleo.

B

BAP - Equipamento responsável por servir como base para outros equipamentos submarinos como ANM, MCV e TH.

F

FLOWLINES - Dutos que interligam os poços às unidades de produção com a finalidade de transportar fluidos.

M

MCV - Equipamento responsável por fazer a interligação dos poços com as *flowlines*.

T

TH - Equipamento responsável por sustentar a coluna de produção.

Referências

- ACCORSI, R. et al. Data Mining and Machine Learning for Condition-based Maintenance. **Procedia Manufacturing**, v. 11, p. 1153–1161, 2017.
- ALI, A. The Impact of Innovativeness and Development Time on New Product Performance for Small Firms. **Marketing Letters**, v. 11, n. 2, p. 151–163, 1 maio 2000.
- AL-SAFRAN, E. Investigation and prediction of slug frequency in gas/liquid horizontal pipe flow. **Journal of Petroleum Science and Engineering**, v. 69, n. 1–2, p. 143–155, nov. 2009.
- AMARI, S.; WU, S. Improving support vector machine classifiers by modifying kernel functions. **Neural Networks**, v. 12, n. 6, p. 783–789, 1 jul. 1999.
- ANP. **Tabela de Poços 2019**, 6 fev. 2019a. Disponível em: <<http://www.anp.gov.br/arquivos/dados-tecnicos/acd/2019-tabela-de-pocos.xlsx>>. Acesso em: 15 jun. 2019
- ANP. **Boletim da Produção de Petróleo e Gás Natural**. [s.l.] ANP, 31 maio 2019b.
- API. **Impacts Of The Natural Gas And Oil Industry On The Us Economy In 2015**. EUA: API, 2017.
- ATS. **Four Ways Predictive Maintenance Can Cut Costs and Improve Productivity**, 2018. Disponível em: <<https://www.advancedtech.com/wp-content/uploads/2019/06/Four-Ways-Predictive-Maintenance-Can-Cut-Costs-and-Improve-Productivity.pdf>>. Acesso em: 15 jun. 2019
- BAI, Y.; BAI, Q. **Subsea Engineering Handbook**. [s.l.] Gulf Professional Publishing, 2018.
- BANSAL, S. K.; KAGEMANN, S. Integrating Big Data: A Semantic Extract-Transform-Load Framework. **Computer**, v. 48, n. 3, p. 42–50, mar. 2015.
- BASTOS, I. C. Eficiência Energética em Sistemas de Elevação Artificial – Bombeio Mecânico. **HOLOS**, v. 1, n. 0, p. 255–263, 18 jul. 2017.
- BASU, J. K.; BHATTACHARYYA, D.; KIM, T. Use of artificial neural network in pattern recognition. **International journal of software engineering and its applications**, v. 4, n. 2, 2010.
- BELGACEM, I. et al. **Experimental investigation of slug pattern in a horizontal two-phase flow**. . In: MULTIPHASE FLOW 2013. A Coruña, Spain: 3 jul. 2013 Disponível em: <<http://library.witpress.com/viewpaper.asp?pcode=MPF13-035-1>>. Acesso em: 10 dez. 2019
- BROWN, K. E. Technology of artificial lift methods. Volume 1. Inflow performance, multiphase flow in pipes, the flowing well. 1 jan. 1977.
- DI BITETTI, M. S.; FERRERAS, J. A. Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. **Ambio**, v. 46, n. 1, p. 121–127, 1 fev. 2017.

EL-ABBASY, M. S. et al. Artificial neural network models for predicting condition of offshore oil and gas pipelines. **Automation in Construction**, v. 45, p. 50–65, set. 2014.

EPE. **Balço Energético Nacional 2018: Ano base 2017**. Rio de Janeiro: EPE, 2018.

GALVÃO, T. F.; PANSANI, T. DE S. A. Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA. **Epidemiologia e Serviços de Saúde**, v. 24, n. 2, p. 335–342, jun. 2015.

GALVÃO, T. F.; PEREIRA, M. G. Revisões sistemáticas da literatura: passos para sua elaboração. **Epidemiologia e Serviços de Saúde**, v. 23, n. 1, p. 183–184, mar. 2014.

GANTZ, J.; REINSEL, D. Extracting Value from Chaos. p. 12, 2011.

GANTZ, J.; REINSEL, D. THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. p. 16, 2012.

GE AUTOMATION. **The Rise of Industrial Big Data**. Disponível em: <<http://www.geautomation.com/download/rise-industrial-big-data>>. Acesso em: 8 nov. 2019.

GEREVINI, G. et al. **Increase of Oil Production in Offshore Production Systems through NMPC Control**. 8 out. 2016

GOUVEIA, F. Tecnologia nacional para extrair petróleo e gás do pré-sal. **Conhecimento & Inovação**, v. 6, n. 1, p. 30–35, 2010.

GOVERNO DO BRASIL. **Setor de petróleo e gás chega a 13% do PIB brasileiro**. Disponível em: <<http://www.brasil.gov.br/economia-e-emprego/2014/06/setor-de-petroleo-e-gas-chega-a-13-do-pib-brasileiro>>. Acesso em: 14 jun. 2019.

GUO, D. et al. **Data driven approach to failure prediction for electrical submersible pump systems**. SPE West. Reg. Meet.: Old Horizons, New Horizons Enabling Technol. **Anais...Society of Petroleum Engineers**, 2015 Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84960337869&partnerID=40&md5=0d287ddaaf9a34f98e0f02727fdbbfe6>>

HAN, Q. et al. **On fault prediction based on industrial big data**. 2017 36th Chinese Control Conference (CCC). **Anais... In: 2017 36TH CHINESE CONTROL CONFERENCE (CCC)**. jul. 2017

IEA. **World Energy Balances 2019**. Disponível em: <<https://webstore.iea.org/world-energy-balances-2019>>. Acesso em: 20 jan. 2020.

IEEE. **About IEEE Xplore**. Disponível em: <<https://ieeexplore.ieee.org/Xplorehelp/#/overview-of-ieee-xplore/about-ieee-xplore>>. Acesso em: 12 nov. 2019.

INKPEN, A. C.; MOFFETT, M. H. **The Global Oil & Gas Industry: Management, Strategy & Finance**. [s.l.] PennWell Books, 2011.

INTERNATIONAL ENERGY AGENCY. **WORLD ENERGY BALANCES 2018**. Place of publication not identified: ORGANIZATION FOR ECONOMIC, 2018.

JALALI, A. et al. **Predicting time-to-failure of plasma etching equipment using machine learning**. IEEE Int. Conf. Progn. Health Manag., ICPHM. *Anais...*Institute of Electrical and Electronics Engineers Inc., 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072772020&doi=10.1109%2fICPHM.2019.8819404&partnerID=40&md5=badc7477077c62e549bb902519aaf8b0>>

LEAHY, K. et al. **Diagnosing wind turbine faults using machine learning techniques applied to operational data**. 2016 IEEE International Conference on Prognostics and Health Management (ICPHM). *Anais...* In: 2016 IEEE INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT (ICPHM). jun. 2016

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33–47, 2018a.

LIU, R. et al. Artificial intelligence for fault diagnosis of rotating machinery: A review. **Mechanical Systems and Signal Processing**, v. 108, p. 33–47, 1 ago. 2018b.

MARCORIN, W. R.; LIMA, C. R. C. Análise dos Custos de Manutenção e de Não-manutenção de Equipamentos Produtivos. v. 11, p. 8, 2003.

MOHAMMADZAHERI, M. et al. An intelligent approach to optimize multiphase subsea oil fields lifted by electrical submersible pumps. **Journal of Computational Science**, v. 15, p. 50–59, jul. 2016.

MONDAL, K. C.; BARUA, H. B. **Fault Analysis and Trend Prediction in Telecommunication Using Pattern Detection: Architecture, Case Study and Experimentation**. [s.l.] Springer Verlag, 2019. v. 1030

MUNION, M. C. **Predictive vs. preventive: The debate—and the future**. Disponível em: <<https://www.plantengineering.com/articles/predictive-vs-preventive-the-debate-and-the-future/>>. Acesso em: 27 dez. 2018.

NASCIMENTO, J. M. A. DO. Simulador computacional para poços de petróleo com método de elevação artificial por bombeio mecânico. 11 mar. 2005.

OKOLI, P.; VEGA, J. C.; SHOR, R. **Estimating downhole vibration via machine learning techniques using only surface drilling parameters**. SPE West. Reg. Meet. Proc. *Anais...*Society of Petroleum Engineers (SPE), 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066314507&partnerID=40&md5=a61e71ee673b18e1b2481072ee69fbd>>

OUADAH, A. **Pipeline Defects Risk Assessment Using Machine Learning and Analytical Hierarchy Process**. Proc. Int. Conf. Appl. Smart Syst., ICASS. *Anais...*Institute of Electrical and Electronics Engineers Inc., 2019Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063503902&doi=10.1109%2fICASS.2018.8651970&partnerID=40&md5=e7cf9cf9b2d5a24347cc1d7526929c0d>>

QIANG, S.; CHENG, S.; JING, K. **Hybrid SVM/HMM Method for Tool Wear Intelligence Measure in Cutting Process**. 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering. *Anais...* In: 2010 3RD

INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING. nov. 2010

RICCOMINI, C.; SANT'ANNA, L. G.; TASSINARI, C. C. G. Pré-sal: geologia e exploração. **Revista USP**, v. 0, n. 95, p. 33, 30 nov. 2012.

SAMPAIO, R.; MANCINI, M. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Revista Brasileira de Fisioterapia**, v. 11, n. 1, p. 83–89, fev. 2007.

SCHEFFER, C.; GIRDHAR, P. **Practical Machinery Vibration Analysis and Predictive Maintenance**. [s.l.] Elsevier, 2004.

SCHUTTE, G. R. PANORAMA DO PRÉ-SAL: DESAFIOS E OPORTUNIDADES. p. 79, 2012.

SCOPUS. **The largest database of peer-reviewed literature - Scopus | Elsevier Solutions**. Disponível em: <<https://www.elsevier.com/solutions/scopus>>. Acesso em: 12 nov. 2019.

SHIRZAD, A.; TABESH, M.; FARMANI, R. A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. **KSCE Journal of Civil Engineering**, v. 18, n. 4, p. 941–948, 2014.

SILVA, A. H. M.; DA HORA, H. R. M.; FERNANDES, R. M. **ANÁLISE DE CONFIABILIDADE DE UM SISTEMA DE CONTROLE DE POÇOS SUBMARINOS DE PETRÓLEO**. Campos dos Goytacazes: Instituto Federal de Educação, Ciência e Tecnologia Fluminense - Programa de Pós-graduação em Sistemas Aplicados à Engenharia e Gestão, 2017.

SNEED, J. **Predicting ESP lifespan with machine learning**. SPE/AAPG/SEG Unconv. Resour. Technol. Conf. **Anais...Unconventional Resources Technology Conference (URTEC)**, 2017 Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85051671524&doi=10.15530%2Furtec-2017-2669988&partnerID=40&md5=106107473da0ad3ecc9960269c2573e6>>

SUGUMARAN, V.; MURALIDHARAN, V.; RAMACHANDRAN, K. I. Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing. **Mechanical Systems and Signal Processing**, v. 21, n. 2, p. 930–942, fev. 2007.

SWANSON, L. Linking maintenance strategies to performance. **International Journal of Production Economics**, v. 70, n. 3, p. 237–244, 18 abr. 2001.

TAITEL, Y.; DUKLER, A. E. A model for slug frequency during gas-liquid flow in horizontal and near horizontal pipes. **International Journal of Multiphase Flow**, v. 3, n. 6, p. 585–596, 1 dez. 1977.

VASSILIADIS, P. A Survey of Extract–Transform–Load Technology: **International Journal of Data Warehousing and Mining**, v. 5, n. 3, p. 1–27, jul. 2009.

VELLOSO, H. M.; DA HORA, H. R. M. Utilização De Técnicas De Machine Learning Na Previsão De Falhas Em Equipamentos: Uma Revisão Sistemática. 2020.

VELLOSO, H. M.; DA HORA, H. R. M.; CARVALHO, R. A. DE. **Inteligência Artificial Para Detecção De Falhas Em Equipamentos: Um Estudo Bibliométrico No Setor De Óleo E Gás**. . In: XV SEGET. Penedo/RJ: 31 nov. 2018

WANG, G. et al. **AdaBoost and Least Square Based Failure Prediction of Railway Turnouts**. Proc. - Int. Symp. Comput. Intell. Des., ISCID. **Anais**...Institute of Electrical and Electronics Engineers Inc., 2017Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85013642211&doi=10.1109%2fISCID.2016.1107&partnerID=40&md5=fc6a59069dd96d38df3e42a8dd91ac5a>>

WANG, L. **Support Vector Machines: Theory and Applications**. [s.l.] Springer Science & Business Media, 2005.

YAO, Y. et al. **Fault Prognosis Method of Industrial Process Based on PSO-SVR**. [s.l.] Springer Verlag, 2019. v. 1043

ZABARAS, G. J. **Prediction of Slug Frequency for Gas-Liquid Flows**. SPE Annual Technical Conference and Exhibition. **Anais**... In: SPE ANNUAL TECHNICAL CONFERENCE AND EXHIBITION. Houston, Texas: Society of Petroleum Engineers, 1999Disponível em: <<http://www.onepetro.org/doi/10.2118/56462-MS>>. Acesso em: 11 set. 2019

ZHOU, D. P.; HU, Q.; TOMLIN, C. J. **Quantitative comparison of data-driven and physics-based models for commercial building HVAC systems**. 2017 American Control Conference (ACC). **Anais**... In: 2017 AMERICAN CONTROL CONFERENCE (ACC). maio 2017