

**FACULDADE DOCTUM DE JOÃO MONLEVADE
INSTITUTO ENSINAR BRASIL - REDE DOCTUM DE ENSINO**

**ARTHUR FERNANDES MACIEL
FABRÍCIA CRISTINA DE SOUSA SILVA**

**APLICAÇÃO DA MINERAÇÃO DE DADOS COMO TÉCNICA DE ANÁLISE PARA
DIAGNÓSTICO DE FALHAS EM TRANSFORMADORES DE DISTRIBUIÇÃO**

**João Monlevade
2019**

ARTHUR FERNANDES MACIEL
FABRÍCIA CRISTINA DE SOUSA SILVA

**APLICAÇÃO DA MINERAÇÃO DE DADOS COMO TÉCNICA DE ANÁLISE PARA
DIAGNÓSTICO DE FALHAS EM TRANSFORMADORES DE DISTRIBUIÇÃO**

**Trabalho de Conclusão de Curso
apresentado à Coordenação do Curso
de Engenharia Elétrica da Faculdade
Doctum de João Monlevade, como
requisito parcial para a obtenção do
título de Bacharel em Engenharia
Elétrica.**

**Orientador: Prof. Dr^a Thaís de Fátima
Araújo Silva.**

**João Monlevade
2019**



FACULDADE DOCTUM DE JOÃO MONLEVADE

FOLHA DE APROVAÇÃO

O Trabalho de Conclusão de Curso intitulado APLICAÇÃO DA MINERAÇÃO DE DADOS COMO TÉCNICA DE ANÁLISE PARA DIAGNOSTICO DE FALHAS EM TRANSFORMADORES DE DISTRIBUIÇÃO, elaborado pelo aluno ARTHUR FERNANDES MACIEL, foi aprovado por todos os membros da Banca Examinadora e aceita pelo curso de Engenharia Elétrica da Faculdade Doctum de João Monlevade, como requisito parcial da obtenção do título de

BACHAREL EM ENGENHARIA ELÉTRICA.

João Monlevade, Treze de Dezembro de 2019

Thais de F. Araújo Silva
Prof. Dra. Thaís de Fátima Araújo Silva

Bráulio Frances Barcelos
Prof. Esp. Bráulio Frances Barcelos

Henrique de Miranda Sá Lopes
Esp. Henrique de Miranda Sá Lopes.



FACULDADE DOCTUM DE JOÃO MONLEVADE

FOLHA DE APROVAÇÃO

O Trabalho de Conclusão de Curso intitulado APLICAÇÃO DA MINERAÇÃO DE DADOS COMO TÉCNICA DE ANÁLISE PARA DIAGNOSTICO DE FALHAS EM TRANSFORMADORES DE DISTRIBUIÇÃO, elaborado pelo aluno FABRICIA CRISTINA DE SOUSA SILVA, foi aprovado por todos os membros da Banca Examinadora e aceita pelo curso de Engenharia Elétrica da Faculdade Doctum de João Monlevade, como requisito parcial da obtenção do título de

BACHAREL EM ENGENHARIA ELÉTRICA.

João Monlevade, Treze de Dezembro de 2019

Thais de F. Araújo Silva
Prof. Dra. Thais de Fátima Araújo Silva

Bráulio Frances Barcelos
Prof. Esp. Bráulio Frances Barcelos

Henrique de Miranda Sá Lopes
Esp. Henrique de Miranda Sá Lopes

AGRADECIMENTOS

Primeiramente agradeço a Deus por ter me proporcionado saúde e força para enfrentar esta jornada, foram inúmeros dias de viagens após um dia cansativo de serviço, inúmeras mudanças de cidade, e algumas mudanças de Faculdade devido ao trabalho, mais tudo está sendo vencido com muita determinação e disciplina.

À minha família que sempre me motivou a fazer o curso. Aos meus pais Geraldo Magela e Maria Aparecida, pela educação e todo o suporte que me deram, e dão até hoje. A minha irmã Patrícia que me incentivou e ajudou quando necessário.

A minha esposa Janaina e minhas filhas Alice e Laís, pelo amor, incentivo apoio e compreensão nos inúmeros momentos de cansaço e estresse.

A Dra. Thais de Fatima Araújo Silva pela Interface e apoio na criação desse projeto.

Aos professores e coordenação da faculdade Doctum, que me transmitiram inúmeros conhecimentos e desafios.

Agradeço a Fabrícia pelo apoio comprometimento e o aprendizado adquirido nesses dois períodos

Arthur Fernandes Maciel

AGRADECIMENTOS

Primeiramente agradeço a Deus, por que até aqui o Senhor me ajudou. Agradeço a Nossa Senhora minha Mãe por ter passado sempre à frente.

Agradeço aos meus pais Frederico Marinho e Jorgina Taciana, meus maiores exemplos. Sou grata pelo incentivo e todas as orações diárias que vocês me dedicaram. Obrigada por estarem sempre ao meu lado! Aos meus irmãos Eleno, Elton, Eric, Enio, Felipe e Fabiula e aos meus sobrinhos, tios, avós, primos que me apoiaram e sempre me passavam uma palavra de ânimo. Aos meus amigos (em especial Lucas, Thuany e Vanessa) que não me permitiram desistir.

Sou grata ao meu namorado Vagner, que me apoiou em todos os momentos, soube compreender quando eu não podia estar presente e me deu forças para vencer mais essa etapa da minha vida. A Alessandra Mara por sempre acreditar em mim e principalmente por me fazer acreditar que sou capaz. Sou grata pelos seus ensinamentos durante esses anos! Agradeço a todos os professores e coordenação da Faculdade Doctum que compartilharam seus conhecimentos e acompanharam a minha jornada nesses cinco anos. Sou grata especialmente aos professores Thais de Fatima Araújo Silva e Ana Regina Bretz que além de educadoras, foram um grande incentivo para essa conquista.

Agradeço o Arthur pela confiança, paciência e o aprendizado adquirido nesses dois longos períodos.

Fabírcia Cristina de Sousa Silva

RESUMO

Os dados originados através da análise dos gases dissolvidos (DGA), em transformadores, são consideráveis. Tarefas de análise e extração de informações desse banco de dados tornam-se extremamente custosas para um analista humano. No processo de Mineração de Dados a análise e extração do conhecimento é realizada procurando padrões consistentes e/ou relacionamentos sistemáticos entre as instâncias destes dados. Na implementação dessa técnica, métodos baseados na Inteligência Artificial são usados a fim de melhorar o processo de análise. Este projeto visa determinar padrões utilizando a ferramenta Diferencial Evolution com o intuito de minerar os dados, e identificar falhas em Transformadores de Distribuição. Sendo elas divididas em três parâmetros: falha térmica, falha elétrica e condição normal de operação. Conforme a literatura pesquisada a validação de um modelo pode ser descrita como o processo de aceitação ou rejeição do mesmo. Esta análise está condicionada a vários fatores, prevalecendo o confronto dos dados reais com os resultados obtidos a partir da simulação do modelo. Como um modelo matemático dificilmente descreve todas as características de um sistema real, ele pode ser considerado válido quando representar as características fundamentais do sistema que será modelado de forma satisfatória. A formulação desse problema considerou 224 amostras (valores de concentração dos gases em ppm), de análise de gases dissolvidos para transformadores de 13,8KV a 230Kv, com e sem comutador sob carga (OLTC). O desempenho do Algoritmo de Otimização Differential Evolution na resolução do Problema de Classificação de Padrões através da Mineração de Dados demonstrou ser válido, no entanto existem pontos que necessitam de melhoria, um exemplo seria a falha na identificação da Classe B. Vale ressaltar que o desempenho de um modelo matemático depende da característica dos dados selecionados e quanto maior for essa disponibilidade de dados melhor será a taxa de aprendizado do programa, com isso os resultados também tendem a melhorar consideravelmente.

Palavras-chave: Mineração de Dados. Descoberta de conhecimento em bases de dados. Evolução Diferencial. Banco de Dados.

ABSTRACT

The increasing production of data from transformer dissolved gas analysis (DGA) produces huge amounts of data. As a result, the information analysis and extraction tasks of these databases become extremely expensive for a human analyst. No Data Mining, Analysis, and Knowledge Extraction processes are performed, looking for consistent standards and / or systematic reporting across instances of this data. In implementing this technique, methods based on Artificial Intelligence are used to improve the analysis process. This project aims to determine the patterns of use of the Differential Evolution tool in order to mine the data, and identify failures in Distribution Transformers. They are divided into three parameters: thermal failure, electrical failure and normal operating condition. According to the researched literature, the validation of a model can be used as a process of acceptance or rejection of the same. This analysis is conditioned to several factors, prevailing or confronting real data with the results obtained from the simulation of the model. Since a mathematical model hardly describes all the characteristics of a real system, it can be considered valid when used as fundamental characteristics of the system that will be satisfactorily modeled. A problem problem in this paper is considered as constant information in a data group totaling 224 drinks (gas concentration values in ppm), these changes are transformed with contactless contact with load switch (OLTC) and distribution transformers of the Dissolved gas analysis for transformers of various voltage levels, from 13.8kV to 230kV. The problem was solved from the evolutionary algorithm. Performance of the optimization algorithm Differential evolution in solving the pattern classification problem using data mining proves to be valid, however, there are points that could improve, a serious example of class B identification failure, it is noteworthy that for a The mathematical model that performs well depends on the resources selected and the higher the data availability, the better the learning rate of the program, with the results likely to improve considerably.

Keywords: Data Mining. Knowledge discovery in databases. Differential evolution. Database.

LISTAS DE FIGURAS

Figura 1 - Esquema Básico do Sistema Elétrico	22
Figura 2 - Exemplificação de um transformador e seu funcionamento de acordo com a lei dos transformadores	23
Figura 3 – Transformador Trifásico de distribuição convencional	25
Figura 4 - Transformador Monofásico	25
Figura 5 - Transformador Trifásico convencional	26
Figura 6 - Óleo em transformador de distribuição	27
Figura 7 - Papel Kraft enrolado sob a bobina do transformador.....	27
Figura 8 - Núcleo de transformador trifásico identificando seus componentes	28
Figura 9 - Bobinas montadas sobre o núcleo de um transformador.....	28
Figura 10 - Carregamento Emergencial de Longa Duração NBR 5336	32
Figura 11 - Carregamento Emergencial de Longa Duração NBR 5356	32
Figura 12 - Carregamento Emergencial de Curta Duração NBR 5356.....	34
Figura 13 - Carregamento Emergencial de Curta Duração IEEE C57.91	34
Figura 14 - Termoscópio, criado por Galileu Galilei em 1595	36
Figura 15 - Termografia Aplicada em Transformadores.....	38
Figura 16- Desenho Esquemático do Sistema de Cromatografia a gás.....	40
Figura 17 - Análise cromatográfica do óleo isolante.....	41
Figura 18 - Percentual dos gases gerados durante arco elétrico no óleo	42
Figura 19 - Percentual de gases gerados durante descarga parcial no óleo	43
Figura 20 - Percentual de gases gerados durante sobre aquecimento	43
Figura 21 - Percentual de gases gerados na degradação da celulose.....	44
Figura 22 - Triângulo de Durval.....	48
Figura 23 - Fluxograma da composição IEEE e IEC para diagnóstico de falhas	49
Figura 24 - Monitor de gases on-line.....	52
Figura 25 - Kelman DGA 900	52
Figura 26 - Registro agrupados em três clusters.....	55
Figura 27 - Processo de Descoberta de Conhecimento através de Banco de Dados	56
Figura 28 - Processo CRISP	59
Figura 29 - Processo de comparação com algumas técnicas	61

Figura 30 - Conjuntos Aproximados	65
Figura 31 - Conjuntos Nebulosos	66
Figura 32 - Modelos matemáticos como prescritivos ou descritivos.....	68
Figura 33 - Processo de modelagem	70
Figura 34 - Pontos de agrupamentos	72
Figura 35 - Etapas do Processo de Descoberta de Conhecimento em Bases de Dados.....	76
Figura 36 - Valores Falso Negativo	86
Figura 37 - Valores Verdadeiro Positivo.....	86
Figura 38 - Valores Falso Positivo.....	87

LISTA DE TABELAS

Tabela 1 - Termografia Aplicada em Transformadores	39
Tabela 3- Relação de gases método Rogers	45
Tabela 4 - Identificação de falhas.....	45
Tabela 5 - Tabela de interpretação de análise de gases dissolvidos	46
Tabela 6 - Tabela de interpretação de análise de gases dissolvidos	47
Tabela 7 - Dados de concentração individual dos gases para verificação da validade de aplicação do método	47
Tabela 8 - Limite e taxas de geração de gases.....	48
Tabela 9 - Concentração de gás dissolvido.....	50
Tabela 10 - Ações baseadas em TDCG.....	50
Tabela 11 - Classificações finais e ações recomendadas	51
Tabela 12 - Banco de Dados do Problema	78
Tabela 13 - Conjunto de Agrupamentos (Clusters)	80
Tabela 14 - Parâmetros da matriz de confusão.....	84
Tabela 15 - Melhor Resultado Falso Negativo	85
Tabela 16 - Melhor Resultado Verdadeiro Positivo	86
Tabela 17 - Melhor Resultado falso Positivo	87
Tabela 18 - Parâmetros definidos para o modelo (%)	88
Tabela 19 - Diferença do valor médio função aptidão e valor de aptidão definido por execução	89
Tabela 20 - Diferença do valor médio Função aptidão e de aptidão definido por execução	90
Tabela 21 - Diferença do valor médio função aptidão e valor da função de aptidão definido por execução	90

ABREVIATURAS E SIGLAS

A.T – Alta Tensão
B – Fluxo magnético
B.T – Baixa Tensão
CFCA – Critério Flexível de Classificação de Acontecimento
CFC – Cromatografia em fase vapor
CPGL – Cromatografia de Partículas gás-líquido
CG – Cromatografia Gasosa
CH₄ – Metano
C₂H₂ – Acetileno
C₂H₄ – Etileno
C₂H₆ – Etano
CO – Monóxido de carbono
DGA – Analise de gás dissolvido
ED – Evolução diferencial
FM – Fase móvel
FE – Fase Estacionária
GE – General Electric
H₂ – Hidrogênio
I – Corrente elétrica
IEC – International Electrotechnical Commission
IED – Intelligent Electronic Devices
IEEE – Institute of Electrical and Electronics Engineers
KDD - Knowledge Discovery in Data Bases
kV - KiloVolt
MVA – Mega Volt ampère
NBR – Norma Brasileira Regulamentadora
NDIR – Non-Dispersive Infrared
ONAN – Óleo Natural, Ar natural
ONAF – Óleo Natural, Ar forçado
Ppm – Partículas por milhão
SEP – Sistema Elétrico de Potência
W – Watts

Sumário

1 INTRODUÇÃO.....	15
1.2 OBJETIVOS	18
1.2.1 <i>Objetivo Geral.....</i>	18
1.2.2 <i>Objetivos Específicos</i>	18
1.2.3 <i>Contribuições.....</i>	19
2 METODOLOGIA.....	19
3 MARCO TEORICO	20
3.1. 1 <i>Sistema elétrico de distribuição</i>	20
3.1.2 <i>História dos Transformadores</i>	21
3.1.3 <i>Transformadores Elevadores</i>	23
3.1.4 <i>Transformadores de Transmissão.....</i>	24
3.1.5 <i>Transformadores de Subtransmissão.....</i>	24
3.1.6 <i>Transformadores de Distribuição.....</i>	24
3.1.7 <i>Transformador de Distribuição Monofásico</i>	25
3.1.8 <i>Transformador de Distribuição Trifásico.....</i>	26
3.1.9 <i>Estrutura Básica</i>	26
3.2 Carregamento do Transformador	29
3.2.1 <i>Sobrecarga</i>	30
3.2.2 <i>Efeitos de um carregamento emergencial de longa duração.....</i>	31
3.2.3 <i>Efeitos de um carregamento emergencial de curta duração</i>	33
3.3 Manutenção em Transformadores	35
3.3.1 <i>Análise Termográfica.....</i>	36
3.3.2 <i>Análise Físico-química.....</i>	39
3.4 CROMATOGRAFIA GASOSA	40
3.4.1 <i>Análise de faltas em transformador</i>	42
3.5 Métodos de diagnósticos	44
3.5.1 <i>Método de Rogers</i>	44

3.5.2 Método NBR 7274.....	45
3.5.3 Método de Doernenburg.....	46
3.5.4 Método Triângulo de Duval.....	47
3.5.5 Guia IEE Norma IEC	49
3.6 REDES INTELIGENTES.....	52
3.7 MINERAÇÃO DE DADOS	53
3.7.1 Tarefas	53
3.7.2 Descoberta do Conhecimento	55
3.7.3 Processos da descoberta do conhecimento em banco de dados.....	55
3.7.4 Métodos (Técnicas)	60
3.7.5 Mineração de itens Frequentes	61
3.7.6 Árvores de Decisão	61
3.7.7 Classificação Bayesiana.....	62
3.7.8 Redes Neurais.....	62
3.7.9 SVM (Support Vector Machines)	63
3.7.10 Classificação por Regras de Associação (Classification by Association Rule).....	63
3.7.11 Aprendizado Tardio (Lazy Learners)	64
3.7.12 Algoritmo Genético (Genetic Algorithm)	64
3.7.13 Algoritmo Evolucionário (Differential Evolution).....	64
3.7.14 Conjuntos Aproximados (Rought Set)	65
3.7.15 Conjuntos Nebulosos (Fuzzy Set)	65
3.7.16 Métodos de Particionamento (Partitioning Methods)	66
4.0 Formulação dos problemas	66
4.1 O Processo de Modelagem	70
4.1.1 Agrupamento (Clustering).....	71
4.1.2 Medidas de distância	73
4.2 Diferencial Evolution	73

4.2.1 O Canônica DE Algoritmo.....	74
4.2.1.1 – Inicialização dos Vetores Decisão Variável	74
4.3.1.2 – Mutação com Vetores de Diferença.....	74
4.3.1.3 – Cruzamento	75
5 ESTUDO DE CASO.....	75
5.1 DADOS	77
5.2 SELEÇÃO DOS DADOS	77
5.3 PRÉ-PROCESSAMENTO	78
5.3.1 Limpeza dos dados	78
5.3.2 Enriquecimento e transformação dos dados	79
5.3.3 Normalização dos valores	79
5.3.4 Definição dos atributos	80
5.3.5 Formulação do Modelo Matemático	82
6 RESULTADOS E DISCUSSÕES	84
6.1.1 - Avaliação de Modelos	84
6.1.2 - Sensibilidade	85
6.1.3 - Resultados para a classe A de Verdadeiro positivos	88
6.1.4 Resultados para a classe B de Verdadeiro positivos.....	89
6.1.5 Resultados para a classe C de Verdadeiro positivos	90
6.1.6 Sugestões para melhoraria do modelo Matemático	91
6.1.7 Em relação ao objetivo geral	92
6.1.8 Em relação aos objetivos específicos.....	92
7 TRABALHOS FUTUROS	93
8 CONSIDERAÇÕES FINAIS.....	94
REFERENCIAS	95
ANEXO	104

1 INTRODUÇÃO

Com a descoberta da eletricidade a energia elétrica foi se tornando cada vez mais importante e imprescindível para as atividades humanas, fazendo-se presente desde um simples aparelho doméstico até grandes centros de pesquisas, hospitais, indústrias, entre outros.

Para chegar até o consumidor final, seja ele industrial ou residencial, a energia elétrica passa pelas etapas como geração, transmissão e distribuição. Na etapa de geração a energia elétrica é obtida por meio da conversão da energia mecânica, energia esta que pode ter origem em turbinas de hidroelétricas, queima de combustíveis fósseis, vapor superaquecido de usinas nucleares, conversão fotoelétrica ou ainda a força dos ventos. Como na maior parte dos casos as centrais de geração estão longe dos centros consumidores é necessário transmitir esta energia gerada.

A etapa de transmissão tem início nas subestações elevadoras localizadas próximo às centrais geradoras. Estas subestações têm como objetivo elevar as tensões geradas a valores da ordem de centenas de milhares de volts. Dessa forma, é possível reduzir o diâmetro dos condutores e simplificar as estruturas mecânicas do sistema.

A etapa de transmissão termina em subestações abaixadoras, onde os níveis de tensão são reduzidos. Esta fase dá início à etapa distribuição onde a energia é distribuída em níveis de tensão da ordem de milhares de volts, visando também uma economia neste estágio do sistema. Assim a energia é distribuída até os transformadores mais próximos as cargas, onde ainda acontece uma nova redução dos valores de tensão.

Um dos equipamentos primordiais na cadeia de distribuição de energia são os transformadores de distribuição. Falhas nos mesmos podem ocasionar colapso no sistema elétrico, fazendo com que inúmeras cidades fiquem sem energia elétrica.

Os transformadores, por serem equipamentos críticos, geralmente, possuem um processo de substituição de difícil realização e com altos custos financeiros. Sendo assim, a manutenção preditiva deve garantir o maior tempo possível de funcionamento e viabilidade, além, claro, da minimização de custos associados a interrupções e suas substituições desnecessárias. Dada a importância e custo do transformador no sistema elétrico, torna-se imprescindível a necessidade do seu

monitoramento, a fim de garantir sua maior vida útil, diagnosticando possíveis defeitos e/ou falhas e atuando antes que elas aconteçam.

Um dos métodos mais eficazes para diagnosticar possíveis desgastes ocasionados pelos efeitos do seu carregamento acima de sua capacidade nominal e até mesmo garantir o prolongamento da vida útil dos transformadores, é a Análise dos Gases Dissolvidos (DGA). Esta análise consiste na retirada de amostras do óleo isolante do transformador, na qual através dos dados coletados são feitos os testes cromatográficos, possibilitando a identificação da formação de alguns gases dissolvidos no óleo, tais como acetileno (C_2H_2), etano (C_2H_6), etileno, (C_2H_4), hidrogênio (H_2), metano (CH_4).

Sistemas de diagnósticos e monitoramento de equipamentos do sistema elétrico de distribuição sem a necessidade de interrupções no fornecimento de energia elétrica, têm sido propostos em níveis de pesquisas com implementações práticas finais por empresas do setor elétrico (fabricantes, transmissoras e distribuidoras).

Em consequência do cenário atual e a imersão da Indústria 4.0 juntamente com as redes inteligentes, a crescente demanda por eletricidade, os requisitos de melhoria da qualidade de energia por parte dos agentes reguladores e a inclusão da geração distribuída estão aumentando a complexidade dos sistemas de distribuição.

Como por exemplo, a aplicação do conceito de transformador inteligente que tem por objetivo melhorar o gerenciamento dos sistemas de distribuição, permitindo, por exemplo: atuação direta sobre os taps do transformador; monitoramento das variáveis e avaliação do carregamento do transformador; planejamento de expansões, manutenção e substituição de transformadores e cabeamento a partir das informações obtidas; entre outros.

Entretanto, a utilização de transformadores inteligentes requer um sistema de comunicação apto para suportar um conjunto de funções e a grande quantidade de dados que os sistemas de distribuição atuais não oferecem. Diante essas características estes sistemas podem implementar o uso de técnicas matemáticas sofisticadas, e recursos computacionais poderosos, como por exemplo, métodos aplicáveis da mineração de dados, um processo que consiste na realização da análise dos dados e na aplicação de modelos matemáticos, a fim de gerar um conjunto de padrões de dados estabelecidos.

Com base neste contexto o presente trabalho propõe avaliar a aplicação da técnica da Mineração de Dados KKD (Knowledge Discovery in Database), também conhecido como descoberta de conhecimento, para auxiliar métodos tradicionais de localização de faltas de técnicas de análises cromatográficas, bem como para as novas tecnologias de sistema inteligentes.

Na sessão 1 é apresentada uma introdução que destaca a necessidade de melhor aproveitamento da quantidade de dados gerada diariamente no Setor Elétrico. São apresentados de forma sintética os principais assuntos abordados neste trabalho e que estão relacionados ao processo de Descoberta de Conhecimento e ao Sistema Elétrico de Potência (SEP). A seguir é apresentado os objetivos deste trabalho que está subdividido em Objetivo Geral e Objetivos Específicos para melhor entendimento do leitor. Para finalizar estão listadas as contribuições provenientes do estudo proposto e também a Metodologia aplicada que serve como guia entre as diversas subseções descritas a seguir.

Na sessão 2 é apresentado o Marco Teórico que trata de forma mais detalhada os principais conceitos de transformadores, seu princípio de funcionamento, bem como suas nomenclaturas de acordo com sua função no sistema elétrico de potência apresentando algumas características dos mesmos.

Na sequência é descrito o processo de Descoberta de Conhecimento que é dividido em Redes Smart Grids e Mineração de Dados, incluindo os Dados e as Tarefas que nesse trabalho irá abordar as para alcançar os objetivos propostos.

Na sessão seguinte é descrito a formulação matemática para problemas de otimização e os conceitos do algoritmo evolucionários.

Na sequência é descrito o estudo de caso onde serão aplicados alguns conceitos e técnicas apresentadas em todo marco teórico. Por fim, para atender todos os objetivos propostos foi realizada uma série de simulações visando a verificação, validação e análise da formulação matemática do problema.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

O principal objetivo deste trabalho é a formulação matemática de um problema de mineração de dados que deve ser resolvido através de um algoritmo evolucionário, aplicado para a obtenção da classificação dos tipos de defeitos relacionados aos dados originados de processos de cromatografia a falhas dos transformadores de distribuição.

1.2.2 Objetivos Específicos

Visando atingir o objetivo geral deste trabalho, elencam-se os seguintes objetivos específicos:

- ✓ Realizar um estudo sobre o princípio de funcionamento dos transformadores de distribuição imerso à óleo, com ênfase nas variáveis que permitem relacionar os dados originados de processos de cromatografia à falhas;
- ✓ Realizar um estudo sobre o processo de Mineração de Dados e as principais tarefas e técnicas que possam ser aplicadas em base de dados do setor elétrico. Avaliando as que melhor se adaptam ao objetivo proposto;
- ✓ Analisar a base de dados de uma concessionária, visando sua organização e preparo dos dados para serem aplicadas as técnicas anteriormente analisadas e definidas;
- ✓ Implementar as tarefas e técnicas selecionados a partir de uma ferramenta computacional, e a partir dessa implementação definir o modelo matemático que será utilizado/proposto para verificar análise das variáveis que relacionam os dados originados de processos de cromatografia;
- ✓ Realizar uma série de simulações visando a verificação, validação e análise da formulação matemática do problema;

1.2.3 Contribuições

A principal contribuição deste trabalho é a disponibilização de um modelo matemático, validado com dados reais, que possibilita a análise do rendimento com ênfase nas variáveis que relacionam os dados originados de processos de cromatografia a falhas de um transformador alocado na cadeia de distribuição de energia elétrica de um sistema elétrico de potência.

A partir desta análise é possível realizar um melhor planejamento das rotinas de manutenção, e consequentemente otimizar a aplicação do investimento em ativos inseridos de forma a aumentar a disponibilidade física e confiabilidade da rede de distribuição de energia elétrica.

2 METODOLOGIA

Para a elaboração deste trabalho foi realizada uma pesquisa de caráter descritivo e explicativo, com a finalidade de analisar e qualificar dados de resultado de cromatografia gasosa em transformadores partindo de revisão bibliográfica e o desenvolvimento de um estudo de caso sobre o tema proposto: aplicação da mineração de dados como técnica de análise para diagnóstico de falhas em transformadores de distribuição.

Para o desenvolvimento do presente trabalho foi necessário um breve entendimento sobre os sistemas de distribuição de energia elétrica, bem como o transformador de distribuição, principal equipamento responsável por adequar os níveis de tensão de acordo com as necessidades de cada consumidor. Foi realizado um estudo sobre os métodos mais comuns utilizados para realizar detecção de falhas em transformadores e a importância do gerenciamento de riscos associados a falhas incipientes. Para formulação do objetivo proposto foi realizado um estudo criterioso de ferramentas matemáticas e computacionais utilizadas pelo processo de Mineração de Dados e as principais tarefas e técnicas que possam ser aplicadas em base de dados do setor elétrico.

Para aplicação dos conceitos e técnicas estudadas, foi realizado um estudo de caso seguindo as etapas da descoberta do conhecimento. Por fim, para atender todos os objetivos propostos foi realizada uma série de simulações visando a verificação, validação e análise da formulação matemática do problema.

3 MARCO TEORICO

Esta sessão demonstra algumas considerações que foram relevantes para a formulação do modelo matemático. Dentre elas, foi necessário um breve entendimento sobre os sistemas de distribuição de energia elétrica, bem como o transformador de distribuição, principal equipamento responsável por adequar os níveis de tensão de acordo com as necessidades de cada consumidor. Serão apresentados os métodos mais comuns utilizados para realizar detecção de falhas em transformadores e a importância do gerenciamento de riscos associados a falhas incipientes. São apresentadas também as ferramentas matemáticas e computacionais utilizadas pelo processo de Mineração de Dados e as principais tarefas e técnicas que possam ser aplicadas em base de dados do setor elétrico.

3.1. 1 Sistema elétrico de distribuição

O sistema de distribuição de energia elétrica pode ser facilmente identificado, uma vez que este muitas vezes acompanha a topografia das cidades, tem como ponto de partida as subestações e através dos circuitos elétricos, conhecidos também como circuitos alimentadores, se ramificam passando por avenidas e ruas, muitas vezes se conectando com os outros alimentadores e atualmente com unidades geradoras de médio e pequeno porte (ABRADEE, 2015).

Este setor de distribuição é visto como o segmento responsável por entregar a eletricidade ao consumidor final, além de ser a principal fonte de entrada de recursos, visto que a maior quantidade de consumidores se encontra neste setor (KVIATKOWSKI e GOZZI, 2005). Existem hoje no país aproximadamente 66 concessionárias distribuidoras de energia elétrica, e estas distribuíram um montante de 465,2 TWh equivalentes à demanda da população (EPE, 2016).

Tecnicamente, o sistema de distribuição é composto pela alta, média e baixa tensão. Silva (2018) caracteriza as linhas de transmissão sendo aquelas que possuem sua tensão superior ou igual a 230 kV, sendo estas as que fazem parte da rede básica do sistema brasileiro. Silva (2018) ainda completa que algumas transmissoras do Brasil possuem linhas com tensão inferior a 230 kV, chamadas assim de Demais Instalações da Transmissão – DIT, que é a grande parte das linhas de transmissão entre 69 kV e 138 kV.

Os sistemas de distribuição podem ser divididos em componentes, como por exemplo: subestações de distribuição, alimentadores de distribuição, transformadores de distribuição, sistemas de distribuição secundária, dentre outros, em que cada um desses tem suas especificações técnicas como tensão, capacidade de carga e estrutura física. No entanto, todos têm como objetivo o transporte com qualidade da energia elétrica para os consumidores.

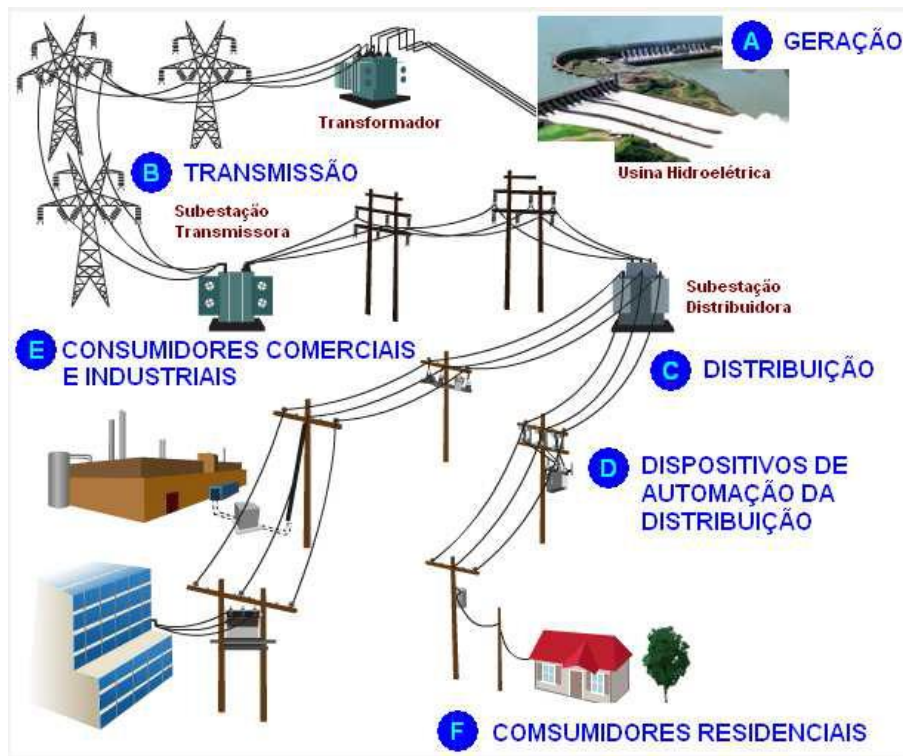
3.1.2 História dos Transformadores

Um dos principais equipamentos utilizados na área da Engenharia Elétrica é o transformador, devido ao fato de estar presente em todos os sistemas elétricos, eletrônicos e de telecomunicações, operando por meio de frequência industrial e radiofrequência (SEPÚLVEDA, 1981).

A história dos transformadores teve início em 1885 com a construção de um transformador nos Estados Unidos, com projeto a seco e utilização de ar como dielétrico. A ideia de que transformadores usando óleo mineral como dielétrico pudessem ser menores e mais eficientes foi patenteada em 1882, pelo professor Elihu Thompson, mas levou uma década até que fosse colocada em prática. Em 1892, a General Electric produziu a primeira aplicação reconhecida de óleo mineral em um transformador. Então, a indústria focou sua atenção em determinar quais as propriedades ideais para o óleo mineral para aplicações como dielétrico e, também, no desenvolvimento de processos para a produção de um Óleo Mineral Isolante OMI de melhor qualidade. As principais propriedades de desempenho foram identificadas e por volta de 1899 pelo menos uma refinaria começou a produzir óleo mineral especialmente desenvolvido para transformadores. (MC SHANE, 2002).

Como mostra a Figura 1, o transformador é um equipamento muito comum no Sistema Elétrico de Potência e tem como finalidade, adequar os níveis de tensão de acordo com as necessidades do consumidor, sendo ela na geração, transmissão e distribuição de energia elétrica.

Figura 1 - Esquema Básico do Sistema Elétrico



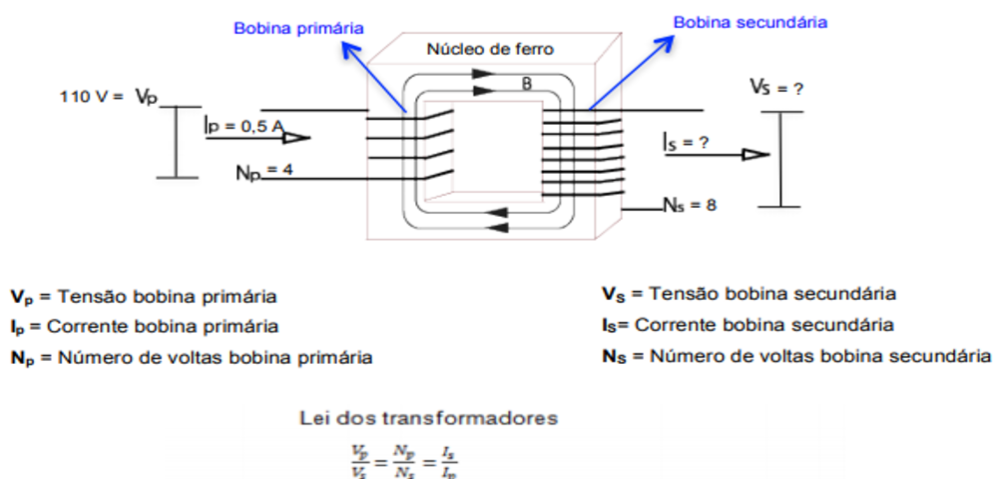
Fonte: Blatt (2016)

Segundo Milasch (1983), um transformador é formado por enrolamentos acoplados de cada lado de um núcleo de ferro, formando de um lado uma bobina primária e do outro uma secundária. Este núcleo é laminado e impregnado de grãos de aço silício orientados para que haja um aumento da corrente de magnetização, assim, uma corrente alternada na bobina primária produz um campo magnético que passará através do centro da bobina secundária e induzirá uma corrente alternada nesta. Os valores produzidos de corrente e tensão pelo transformador seguem a lei dos transformadores, demonstrando que a razão das tensões é proporcional à razão entre o número de voltas das espiras e inversamente proporcional à razão de correntes, como pode ser observado na Figura 2.

Os transformadores operam segundo a Lei de Faraday (1831): “Uma corrente elétrica é induzida em um circuito se este estiver sob a ação de um campo magnético variável”. Assim, pode-se dizer matematicamente que a força eletromotriz induzida aplicada no tempo é função do volume do núcleo de sua permissividade, originando seu vetor indução magnética e é função do número de espiras ou enrolamentos do primário e secundário do transformador (FERREIRA, 2018).

Todo este processo de transformação, como descreve Milasch (1983) contribui para o aumento da temperatura interna do transformador e que se não minimizado, pode levar a quebra da cadeia parafínica do OMI, e a consequente produção de gases que ficarão dissolvidos no óleo, podendo levá-lo a explosão.

Figura 2 - Exemplificação de um transformador e seu funcionamento de acordo com a lei dos transformadores



Fonte: Universidade Tecnológica Federal do Paraná (2009)

Para Milasch (1983), ao longo do funcionamento dos transformadores, muitos processos de desgaste e de envelhecimento ocorre no sistema de isolamento, os efeitos da fadiga térmica, química, elétrica e mecânica, tais como, pontos quentes, sobreaquecimentos, sobretensões e vibração são causadores de alterações do sistema isolante e precisam ser acompanhados para garantir a eficiência do equipamento, permitindo intervenções de manutenção preventiva e preditiva, a fim de evitar paradas de máquina e, conseqüentemente, aumento de custos.

3.1.3 Transformadores Elevadores

Geralmente são empregados para elevar as tensões produzidas pelos geradores de energia elétrica nas usinas, onde os terminais dos enrolamentos de baixa tensão são conectados ao gerador e os terminais dos enrolamentos de alta tensão são conectados à carga, que geralmente são as linhas de transmissão. São comumente encontrados em tensões de até 20kV nos enrolamentos de baixa tensão (B.T) e até 500kV nos enrolamentos de alta tensão (A.T).

3.1.4 Transformadores de Transmissão

Estão presentes nas subestações de transmissão, onde conectam linhas com diferentes níveis de tensão. Com uma forma construtiva mais complexa quando comparada aos transformadores elevadores, este tipo de equipamento tem maior quantidade de enrolamentos e sistema de regulação de tensão sob carga.

3.1.5 Transformadores de Subtransmissão

De forma análoga aos transformadores de transmissão, os de Subtransmissão conectam linhas com diferentes níveis de tensão, recebem as tensões das linhas de transmissão e alimentam as linhas de distribuição.

São providos geralmente de dois enrolamentos um de alta e outro de baixa tensão, e possuem sistema de regulação de tensão a vazio. Na maioria das vezes têm níveis de tensão de até 138kV nos enrolamentos primários e até 34,5kV no secundário.

3.1.6 Transformadores de Distribuição

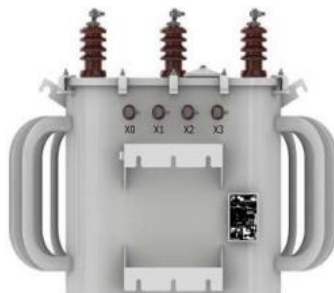
Caracterizados como de menor porte, os transformadores de distribuição são utilizados para converter os níveis de tensão recebidos das linhas de subtransmissão para valores de tensão entregues aos consumidores (residenciais, comerciais e industriais).

Segundo Caparó (2005), os transformadores de distribuição separam o sistema primário do sistema secundário. Os circuitos primários transmitem energia desde a subestação de distribuição para os transformadores de distribuição dos consumidores. As linhas de distribuição primária que se originam na subestação são chamadas de alimentadores primários ou circuitos primários. Os circuitos secundários transmitem energia desde o transformador de distribuição até a entrada do serviço dos consumidores, sendo as faixas de tensão de linha entre 110 e 380 volts.

Conforme demonstra a Figura 3, os transformadores de distribuição convencionais aplicados a classes de tensão de 15KV e 36,2KV são de características monofásicas e trifásicas. Os monofásicos são para sistemas de baixa

densidade de carga instalados em áreas urbanas e rurais, pois demandam baixa carga. Já os transformadores de distribuição trifásicos são instalados em áreas urbanas que demandam média e alta densidade de carga.

Figura 3 – Transformador Trifásico de distribuição convencional

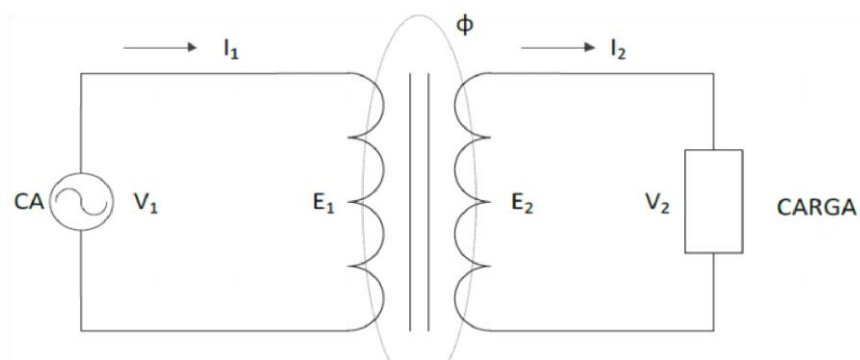


Fonte: ABB (2010)

3.1.7 Transformador de Distribuição Monofásico

A Figura 4 apresenta o esquema elétrico do transformador monofásico, onde possuem um conjunto de bobinas de alta e baixa tensão colocados sobre um núcleo, estes são alimentados por um sistema de corrente alternada monofásica. Este tipo de transformador é utilizado para consumidores de pequeno porte que necessitam uma tensão de 127 V a 380 V (CORRAL, 2008).

Figura 4 - Transformador Monofásico

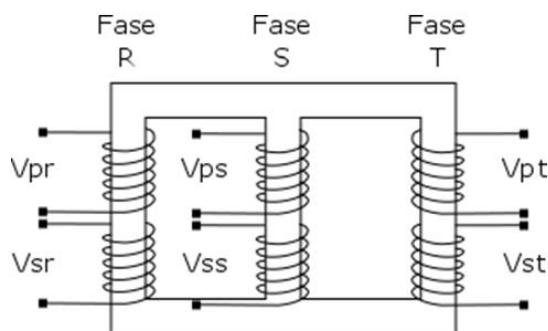


Fonte: Fitzgerald (2006)

3.1.8 Transformador de Distribuição Trifásico

A Figura 5 representa a construção de um transformador trifásico que tem como característica a junção de três transformadores monofásicos. Deste modo apenas as ligações das bobinas caracterizam um transformador trifásico, estes são constituídos por pelo menos três enrolamentos no primário e três enrolamentos no secundário (HARLOW, 2004).

Figura 5 - Transformador Trifásico convencional



Fonte: Entran (2018)

3.1.9 Estrutura Básica

A maioria dos transformadores em uso no mundo são isolados por quatro componentes básicos: óleo, resina, papel e ar (FERRAZ, 2015 apud HEATHCOTE, 2007).

✓ Óleo para Transformadores: Em conformidade com os óleos isolantes, na maioria de origem mineral tem o duplo objetivo: proporcionar isolamento e servir como meio de arrefecimento dissipando para o ambiente as perdas que são produzidos no transformador sob a forma de calor (Figura 6).

Figura 6 - Óleo em transformador de distribuição



Fonte: Fitzgerald (2006)

✓ Papel: O papel é um dos mais baratos e melhores materiais isolantes conhecidos. Como mostra a Figura 7, papel Kraft é, por definição, feito inteiramente de celulose de fibra longa não branqueada, fabricado pelo processo de sulfato cru, porque agentes clareadores residuais podem comprometer suas propriedades elétricas.

Figura 7 - Papel Kraft enrolado sob a bobina do transformador

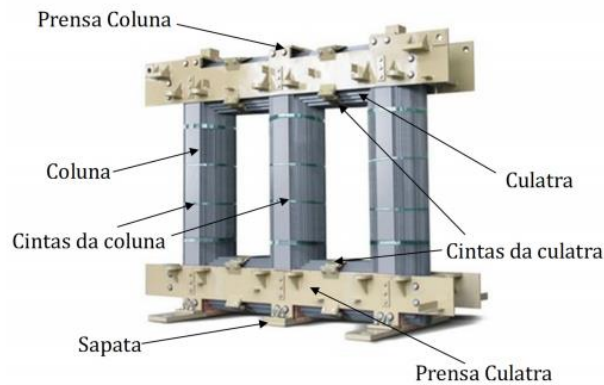


Fonte: Ferraz (2015)

✓ Núcleo: é constituído por lâminas de material ferromagnético que formam as colunas e culatras. É sobre as colunas que são montados os enrolamentos. As culatras completam o circuito magnético. As sapatas, prensa colunas, prensa culatras, cintas das colunas e as cintas das culatras são utilizadas para prensar o núcleo, e são fundamentais para reduzir o ruído gerado pelo choque entre as chapas que constituem o núcleo ao serem percorridas pelo fluxo magnético.

A Figura 8 mostra um núcleo de transformador trifásico montado, identificando seus componentes (VILLIBOR, 2019).

Figura 8 - Núcleo de transformador trifásico identificando seus componentes



Fonte: Vilibor (2019)

✓ Bobinas: é a primeira parte integrante de um dos transformadores de distribuição. é composta de fio de cobre revestido com esmalte (figura 9) e fica sobre um núcleo magnético. A entrada de baixa corrente, alta tensão de corrente alternada da usina flui através do fio, e no processo gera um fluxo magnético.

Figura 9 - Bobinas montadas sobre o núcleo de um transformador



Fonte: Villibor (2019)

Componentes como comutadores de tap, buchas, sistemas de refrigeração, dentre outros são considerados componentes auxiliares.

✓ Radiadores: tubos de resfriamento são usados para resfriar o óleo do transformador. O óleo do transformador é circulado através dos radiadores estes ficam dispostos na parte externa dos transformadores. Existem diferentes métodos de resfriamento de transformadores disponíveis são eles: óleo natural com resfriamento natural - ONAN (Óleo Natural, Ar Natural); óleo natural com ventilação forçada - ONAF (Óleo Natural, Ar Forçado); óleo com circulação forçada do líquido isolante e com ventilação forçada - OFAF (Óleo Forçado, Ar Forçado); óleo com circulação forçada do líquido isolante e com resfriamento a água - OFWF (Óleo Forçado, Água (*water*) Forçada; seco com resfriamento natural - AN (Ar Natural); seco com ventilação forçada - AF (Ar Forçado).

✓ Comutador: Dispositivo mecânico que permite variar o número de espiras dos enrolamentos de tensão. Sua finalidade é corrigir o desnível de tensão existente nas redes de distribuição, devido à queda de tensão ocorrida ao longo das mesmas.

Os componentes listados acima desempenham papel de elementos primários, secundários e dispositivos de proteção, fazendo parte da maioria dos transformadores. Os dispositivos primários e secundários que compõem os transformadores possibilitam que estes possam desempenhar sua função básica. Já os dispositivos de proteção têm como principal função alarmar ou desligar os transformadores em caso de situações anormais de funcionamento.

3.2 Carregamento do Transformador

Os efeitos de carregamento acima dos valores de placa do transformador podem proporcionar riscos ao equipamento. Para Ferreira (2013), o tempo de vida útil dos transformadores de distribuição está relacionado a muitos fatores, como por exemplo, o carregamento acima de sua capacidade nominal, cargas desbalanceadas, equipamentos mal projetados, sobre tensões internas e externas, sistemas de proteção inadequados e procedimentos de manutenção preventivas.

De acordo Bartley (2003), as falhas em transformadores de distribuição podem ser apresentadas por diversos fatores:

✓ Descargas atmosféricas (confundidas com surtos de linhas);

- ✓ Sobrecarga, situações nas quais os transformadores com carregamento acima da nominal têm sua vida útil reduzida;
- ✓ Falhas na Isolação, excluindo-se aquelas por descargas atmosféricas e sobretensões, concentrando-se naquelas por pirólise, oxidação, acidez e umidade;
- ✓ Sobretensões em linhas, oriundas de chaveamentos, curto-circuito e outras anormalidades das linhas de transmissão e distribuição;
- ✓ Umidade que penetrou nos transformadores devido a falhas nos anéis de vedação e outros orifícios causados por corrosão ou vandalismo;
- ✓ Erros de projeto, onde foram incluídas as perdas ou cargas nominais mal especificadas, isolação do núcleo inadequada, suportabilidade ao curto-circuito mal dimensionada e objetos estranhos no interior do tanque;
- ✓ Fogo externo, cujas temperaturas podem sobreaquecer o transformador iniciando uma falha interna. O fogo pode ser proveniente de queimadas ilegais em matas e plantações;
- ✓ Operações inadequadas e manutenções impróprias, resultando no acúmulo de sujeira, umidade, ninhos de pássaros e insetos, corrosão e perdas de conexões externas às carcaças dos transformadores;
- ✓ Contaminação do óleo, situações nas quais o óleo fica carbonizado ou perde suas propriedades físico-químicas;

3.2.1 Sobrecarga

Uma das origens de falhas mais comuns em transformadores de distribuição é a sobrecarga. De acordo com Fritz et al (2015), a sobrecarga em um transformador de distribuição é o resultado da circulação de um fluxo excessivo de corrente no enrolamento secundário do transformador, causando aumento adicional de temperatura. Esta elevação da temperatura pode causar a deterioração do material isolante, afetando a vida útil do equipamento. Em casos extremos, a sobrecarga pode provocar a avaria definitiva do transformador.

O carregamento dos transformadores convencionais aplicados a classes de tensão de 15KV e 36,2KV segundo Fritz et al (2015), podem variar de um carregamento inferior a 25% da sua potência nominal até um carregamento superior a 150% da sua potência nominal. Portanto, sobrecarga em termos do carregamento de um TD, (Transformador de Distribuição) pode ser definida como um

carregamento superior ao carregamento nominal imposto ao transformador (sobrecarga instantânea ou de longa duração).

As normas NBR 5356 (ABNT, 2017) e C57.91 (IEEE, 2011) abordam efeitos e perigos de duas condições de carregamento acima dos dados de placa que serão descritas em sequência.

3.2.2 Efeitos de um carregamento emergencial de longa duração

Segundo a NBR 5356 (ABNT, 2017), espera-se que o carregamento de longa duração ocorra raramente, pois não é uma condição de operação normal, podendo durar por várias horas, levando a um envelhecimento considerável do transformador.

O carregamento emergencial de longa duração tem como principais efeitos:

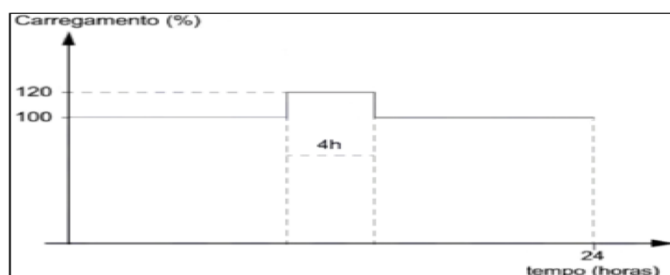
- ✓ Altas temperaturas são aceleradas pela deterioração das propriedades mecânicas de isolamento do condutor. A vida efetiva dos transformadores pode ser reduzida efetivamente se a deterioração durar por um longo período de tempo, principalmente se este é submetido a curto-circuito do sistema ou eventos de transporte;

- ✓ A taxa de envelhecimento de outras partes isolantes, especialmente as que sustentam a pressão axial do bloco de enrolamentos são aceleradas pelas altas temperaturas;

- ✓ O aumento da resistência de contato dos comutadores sob carga em correntes e em temperaturas elevadas leva à uma degradação térmica acelerada. As altas temperaturas ainda podem tornar os materiais de vedação do transformador frágeis e quebradiços;

A NBR 5356 (ABNT, 2017) considera que a expectativa de perda da vida útil do transformador não sofre alterações caso em seu ciclo diário o transformador atinja 120% de carregamento por no máximo 4 horas, tempo esse considerado carregamento de longa duração. A Figura 10, mostra o comportamento do carregamento de longa duração.

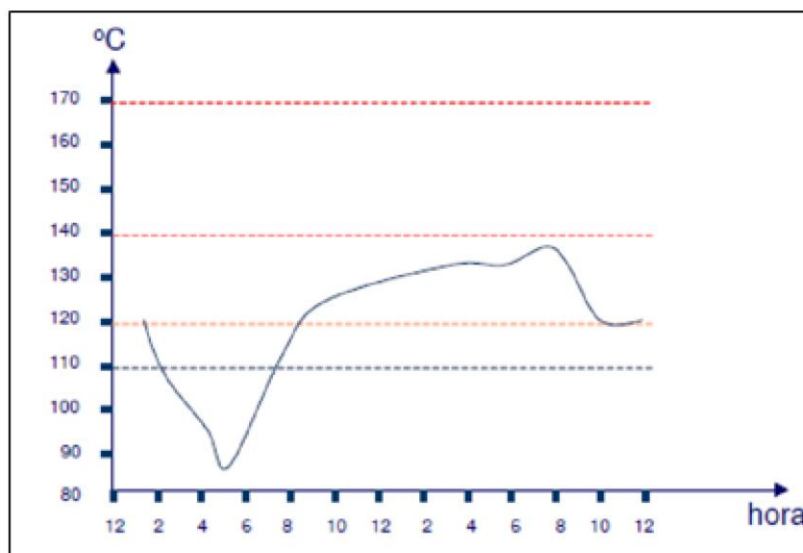
Figura 10 - Carregamento Emergencial de Longa Duração NBR 5336



Fonte: NBR 5356 (ABNT, 2017)

A C57.91 (IEEE, 2011) refere-se ao carregamento sendo limitado pela temperatura do ponto mais quente do transformador. A expectativa de perda da vida útil do transformador não sofre alterações caso em seu ciclo diário o transformador atinja 140°C por um curto período de tempo. A Figura 11, mostra o comportamento do carregamento de longa duração, onde a temperatura do ponto mais quente fica de 120 C° a 140C°.

Figura 11 - Carregamento Emergencial de Longa Duração NBR 5356



Fonte: Guia IEEE para carregamento de transformadores imersos em óleo mineral (2011)

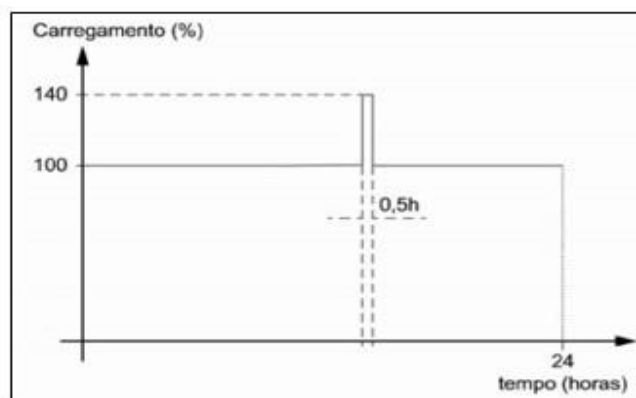
3.2.3 Efeitos de um carregamento emergencial de curta duração

Para a NBR 5356 (ABNT, 2017) o carregamento emergencial de curta duração resulta em aumento do risco de falha, levando o ponto quente do condutor a um nível que resultara na redução da suportabilidade dielétrica. Este tipo de condição pode ser melhor que perder a carga. Deve-se considerar o tempo do carregamento para evitar falhas no transformador, onde esse carregamento deve ocorrer raramente, sendo rapidamente reduzido ou que o transformador seja desconectado. O carregamento emergencial de curta duração tem como principais efeitos e perigos:

- ✓ Redução da suportabilidade dielétrica ocorre devida as altas solicitações elétricas, que produzem bolhas de gás nos enrolamentos e nas ligações, sendo a principal falha de curta duração. Essas bolhas ocorrem quando a temperatura do ponto mais quente excede 140°C para um transformador com percentual de umidade na isolação dos enrolamentos de aproximadamente 2%. Com o aumento da concentração de umidade esta temperatura critica diminui;
- ✓ A supersaturação do óleo ou o aquecimento nas superfícies de grandes partes metálicas desenvolvem bolhas de gás. Sendo desenvolvidas em regiões de baixa solicitação elétrica, para que houvesse redução significativa da suportabilidade dielétrica essas bolhas deveriam circular em regiões de alta solicitações elétricas;
- ✓ As partes metálicas que não estão em contato térmico direto com isolação celulósica e com o óleo transformador, podem atingir altos valores de temperatura, não excedendo 180°C;
- ✓ Altas temperaturas causam redução da suportabilidade a curtos-circuitos, devido as deteriorações temporárias das propriedades mecânicas da isolação sólida;
- ✓ Aumento de pressão nas buchas pode ocasionar vazamento de óleo. Se a temperatura da isolação exceder 140 °C pode ocorrer gases em buchas condensivas;
- ✓ A expansão do óleo pode causar um fluxo excessivo de óleo no conservador; Interrupção de altas correntes podem ser danosas ao comutador sob carga;

A NBR 5356 (ABNT, 2017) ainda aborda que a expectativa de perda da vida útil do transformador não sofre alterações caso em seu ciclo diário o transformador atinja 140% de carregamento por no máximo 30 minutos, tempo esse considerado carregamento de curta duração. A Figura 12 mostra o comportamento do carregamento de curta duração.

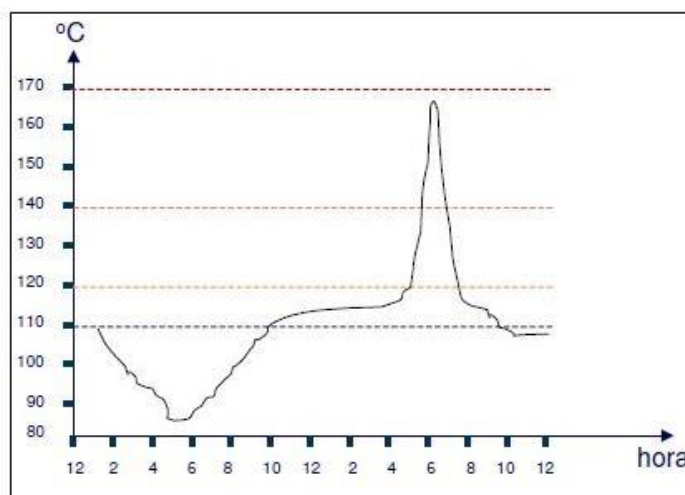
Figura 12 - Carregamento Emergencial de Curta Duração NBR 5356



Fonte: NBR 5356 (ABNT, 2017)

A C57.91 (IEEE, 2011) considera que a expectativa de perda da vida útil do transformador não sofre alterações caso em seu ciclo diário o transformador atinja 180°C por um curto período de tempo. A Figura 13 mostra o comportamento do carregamento de curta duração, onde a temperatura do ponto mais quente chega a 163°C em 1 hora.

Figura 13 - Carregamento Emergencial de Curta Duração IEEE C57.91



Fonte: Guia IEEE para carregamento de transformadores imersos em óleo mineral (2011)

O carregamento emergencial de curta duração pode causar danos maiores ao transformador do que o carregamento emergencial de longa duração, pois atinge níveis de temperatura maiores.

Freitas et al (2002) e o Laboratório de Alta Tensão da Universidade Federal de Itajubá (2014), afirmam que para se ter um bom conhecimento da carga conectada a um sistema elétrico e avaliar o perfil da curva de carregamento diária de um transformador de distribuição é necessário conhecer as características comportamental em função de seus consumidores específicos.

Ainda segundo o Laboratório de Alta Tensão da Universidade Federal de Itajubá (2014), para facilitar a descrição do comportamento variável das cargas, a curva de demanda de um consumidor é dividida em intervalos de tempo iguais.

O que se conclui que estas curvas não são exatamente iguais todos os dias do ano, e que, ao longo dos anos, muda o perfil dos consumidores.

3.3 Manutenção em Transformadores

Segundo Senna (2010), existem três técnicas de manutenção em transformadores: Preditivas, Preventivas e as Corretivas:

- ✓ A manutenção preventiva é realizada de forma programada e considera o envelhecimento do equipamento a partir do tempo de operação. Realiza ensaios e verificações com o equipamento energizado ou não;
- ✓ A manutenção preditiva baseada na condição determina o ponto ótimo para executar a manutenção preventiva, isto é, indica o ponto a partir do qual a probabilidade de falha no equipamento assume valores indesejáveis;
- ✓ Manutenção corretiva é realizada quando o equipamento deixa de cumprir sua função principal em virtude de uma falha;

Pelo Fato dos Transformadores de distribuição não possuírem proteções tão específicas e eficazes como as dos Transformadores de Potência, o acompanhamento periódico desses equipamentos é de suma importância. E a forma mais utilizada para esse fim seria a implementação da manutenção preditiva que possibilita realizar diagnóstico e análise de tendências a partir de resultados de

ensaios e da análise dos fenômenos que possam vir a ocorrer durante a operação de um equipamento.

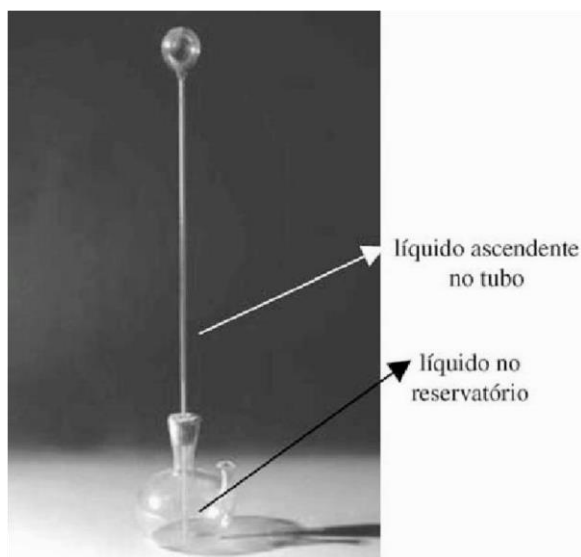
A seguir serão descritas algumas formas de execução da manutenção preditiva, no entanto o escopo deste trabalho é a análise cromatográfica dos gases dissolvidos no óleo.

3.3.1 Análise Termográfica

A manutenção preditiva utiliza de diversos métodos e ferramentas de análises, para detecção de falhas em estágios iniciais, dentre elas se destaca a utilização da Termografia.

A figura 14 mostra o primeiro instrumento com intuito de medição de temperatura, chamado de Termoscópio, criado 1595 por Galileu Galilei. Em 1645 foram fabricados os primeiros termômetros a base de álcool em tubos capilares selados hermeticamente. Já no ano de 1777 o pesquisador italiano Landrini observou que o aquecimento crescia do violeta para o vermelho.

Figura 14 - Termoscópio, criado por Galileu Galilei em 1595



Fonte: Researchgate (2019)

Foi William Herschel, no entanto, quem verificou, em 1800, que o máximo de aquecimento ocorria fora do espectro visível, além do vermelho. Em 1862, Melloni construiu a termopilha, composta de vários termopares ligados em série.

A sensibilidade alcançada pelo arranjo era quarenta vezes superior aos melhores termômetros da época, o suficiente para detectar uma pessoa a dez metros de distância (VERMINNEN *et al.* 1984).

Segundo Santos (2006) no Brasil, os primeiros termovisores chegaram nos anos 70. Eles empregavam detectores resfriados a nitrogênio líquido, sistemas de varredura óptico-mecânicos para formar as imagens térmicas e a temperatura do objeto inspecionado era através de cálculos.

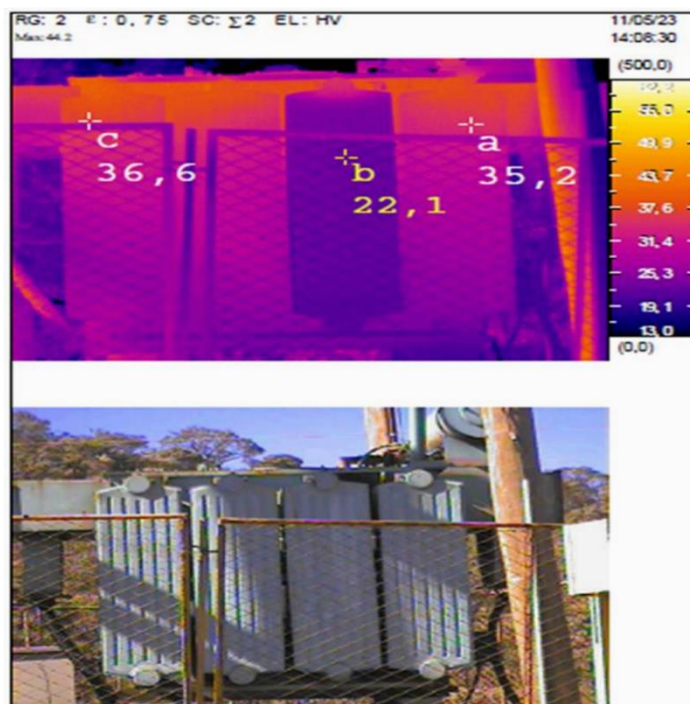
A termografia possibilita a medição de temperaturas e a formação de imagens térmicas de um determinado componente, equipamento ou processo, a partir da radiação infravermelha, naturalmente emitida pelos corpos. Atualmente seu uso na Engenharia abrange desde equipamentos mecânicos, equipamentos eletroeletrônicos, equipamentos e sistema elétricos. Nas instalações elétricas onde houver desequilíbrio de fases e ou sobrecargas, esses dois acabam gerando calor. Com a inspeção “sensitiva” (Usa os sentidos Humanos) esses problemas podem passar despercebidos.

Justamente nesse ponto a termografia pode ser utilizada, na identificação de problemas ela é considerada uma das formas de inspeção preditiva mais eficazes, em sistemas elétricos pois garante a segurança necessária até mesmo em grandes níveis de tensão, pois não é necessário o contato físico com os componentes. Algumas das vantagens da utilização desse método são:

- ✓ Identifica e localiza falhas pontuais;
- ✓ Possibilita a análise com o equipamento ou sistema elétrico em pleno funcionamento;
- ✓ Traz segurança a pessoa que faz a atividade, pois as inspeções são feitas a distância;
- ✓ Possibilita a determinação aproximada da temperatura dos objetos sem contato físico;
- ✓ Os dados são coletados e arquivados em tempo real.

A Figura 15 mostra a termovisão executada em transformador de 13,8kv onde foi constatado que um dos radiadores indicavam problemas.

Figura 15 - Termografia Aplicada em Transformadores



Fonte: Os Autores (2011)

Com o uso da Termografia, várias Técnicas e metodologias são adotadas por empresas. Uma delas considera o uso de um critério de aquecimento para definir um diagnóstico sobre o componente, o qual pode ser classificado como “Normal”, “Suspeita de Falha”, “Falha Potencial” ou “Falha Iminente” (VERATTI, 1999).

Essa metodologia, denominada critério Flexível de Classificação de Aquecimentos – CFCA, foi proposta em 1982 pelo Eng. Atílio Bruno Veratti, após a pesquisa em trabalhos semelhantes realizados por empresas norte-americanas, inglesas e francesas (ENGETERMO, 2009).

Para garantir a qualidade dessa técnica utiliza-se o termovisor como aparelho de medição que basicamente consiste em uma câmera infravermelha dotada de recursos para análise e determinação de regiões de diferentes temperaturas, que incorpora uma câmera visual de alta resolução com uma lâmpada de iluminação embutida, para ajudar a documentar claramente o trabalho e um apontador laser embutido para orientar em relação ao alvo.

Um importante parâmetro que deve ser observado durante as medições é a emissividade, pois os termovisores não medem temperatura diretamente mais sim a radiação. A emissividade é a relação entre a energia total irradiada por uma dada superfície em uma dada temperatura e a energia total irradiada por um corpo negro

a mesma temperatura. Depende da composição e da condição da superfície. Materiais polidos tem menor emissividade que materiais ásperos ou corroídos. Para facilitar na determinação dos componentes costuma-se dividi-los em grupos, segundo o material de que são feitos conforme demonstra a Tabela 1.

Tabela 1 - Termografia Aplicada em Transformadores

Material	Ligeiramente oxidado	Severamente oxidado
Alumínio e suas ligas	0,20 - 0,40	0,67 – 0,95
Cobre e suas ligas	0,39 – 0,50	0,78 – 0,95
Aço	0,52 – 0,60	0,82 – 0,94
Porcelana	0,10 – 0,20	0,80 – 0,95
Borracha	0,70 – 0,80	0,80 – 0,95

Fonte: Contemp (2017)

3.3.2 Análise Físico-química

Segundo Senna (2010), as características físico-químicas do óleo devem ser acompanhadas por meio de ensaios periódicos de forma a monitorar a degeneração do óleo. Esse processo possibilita identificar o grau de influência de contaminantes, tais como ácidos, partículas solidas em suspensão, água e borra, em suas propriedades dielétrica.

A coleta do óleo deve ser feita com muito cuidado evitando assim a contaminação do recipiente onde a amostra se encontra, pois, uma possível contaminação levaria a um resultado com alterações e o mesmo se tornaria insatisfatório.

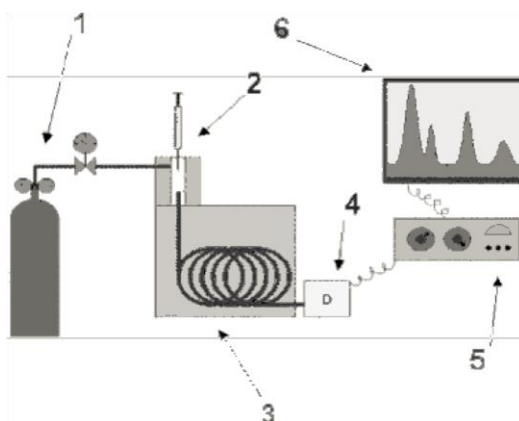
3.4 CROMATOGRAFIA GASOSA

Cromatografia gasosa é uma técnica onde gases ou substância voláteis podem ser separadas (LANÇAS, 1993). Essa separação se fundamenta na diferente distribuição das substâncias da amostra entre uma fase estacionária (sólida ou líquida) e uma fase móvel (gasosa) (COLLINS, 1993).

É também conhecida como (CFC), Cromatografia em fase vapor, ou (CPGL) cromatografia de partição gás-líquido. Essa segunda nomenclatura é a mais correta por isso é mais encontrada na literatura científica.

A Cromatografia Gasosa (CG) é uma técnica para separação e análise de misturas de substancias voláteis. A amostra (2) é vaporizada e introduzida em um fluxo de um gás adequado denominado de fase móvel (FM) ou gás de arraste (1) como ilustrado na figura 16.

Figura 16- Desenho Esquemático do Sistema de Cromatografia a gás



Fonte: Arantes (2005)

A figura acima demonstra o fluxo de gás vaporizado que caminha por um tubo (3) no mesmo está a fase estacionária FE (coluna cromatográfica), é lá que acontece a separação da mistura. Usualmente a FE é constituída por um filme de um líquido volátil (4) mais também é encontrado como um sólido adsorvente (Cromatografia Gás Sólido), que está sobre um sólido inerte (Cromatografia Gás-líquido com coluna recheada ou empacotada). Existem dois fatores que constituem a separação da amostra:

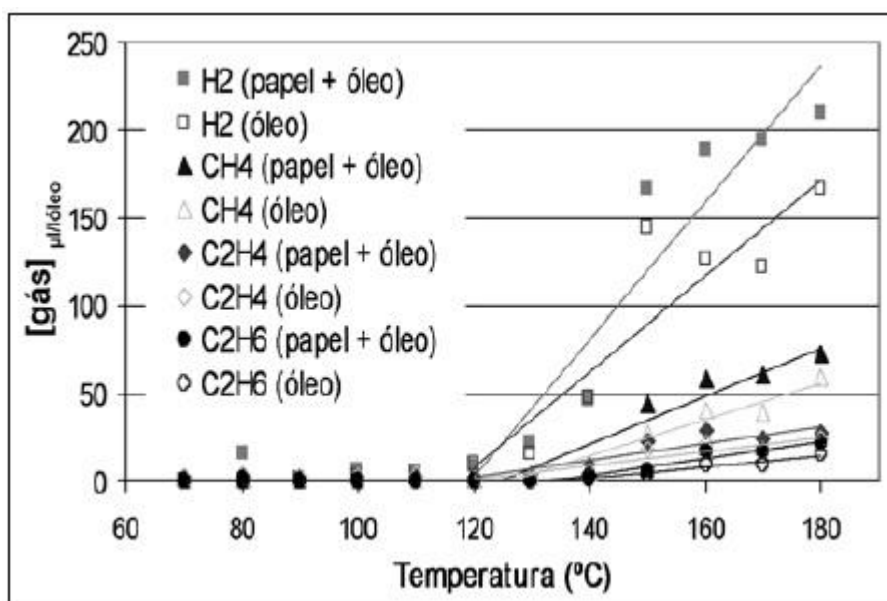
✓ Solubilidade na FE, quanto maior for a solubilidade mais devagar ela irá percorrer pela coluna.

✓ Volatilidade: quanto mais volátil for a substância, maior será a possibilidade de permanecer vaporizada e irá percorrer mais rápido pelo sistema.

No processo a substância sai da coluna separadamente, dissolvidas no gás de arraste e passam no dispositivo detector (5) que de acordo com a quantidade do fluido gera um sinal elétrico. Cromatograma (6) é o registro desse sinal em relação ao tempo.

Para analisar as condições do óleo são colhidas amostras e as mesmas são enviadas a um laboratório que irá analisar as condições desse óleo. Os diagnósticos a partir de uma amostra de óleo podem determinar causas e medidas a serem adotadas para evitar paradas indesejadas no equipamento como ilustra a figura 17.

Figura 17 - Análise cromatográfica do óleo isolante



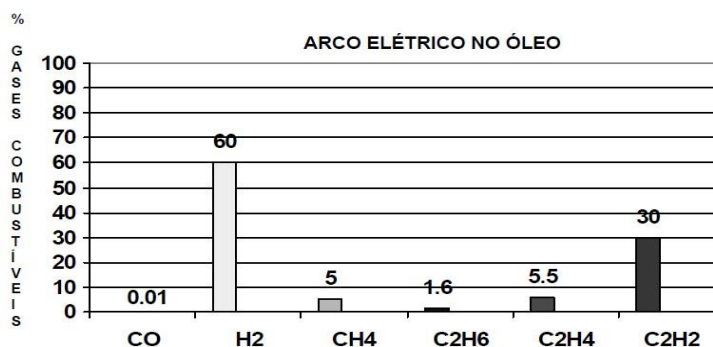
Fonte: Lisboa (2010)

3.4.1 Análise de faltas em transformador

É muito indicado e importante a análise periódica desses gases, com isso é possível encontrar falhas prematuramente, avaliar as condições de desgaste do equipamento e orientar sobre medidas a serem tomadas para evitar que esses danos possam causar maiores prejuízos ao proprietário do equipamento. Alguns dos problemas que são constatados através dessa análise são:

✓ **Arco elétrico:** Gera volume considerável de hidrogênio e acetileno também gera porém em volume menor o metano e etileno, os mesmos estão indicados na figura 18. Se envolver a celulose pode ocorrer a geração de monóxido e dióxido de carbono. Gás chave: acetileno;

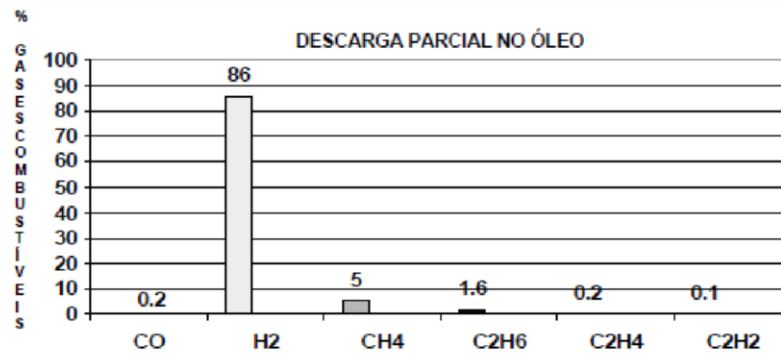
Figura 18 - Percentual dos gases gerados durante arco elétrico no óleo



Fonte: Arantes (2005)

✓ **Corona:** gera volume considerável de metano e hidrogênio e também menor quantidade de etano e etileno. Geram dióxido e monóxido e de carbono, devido a presença da celulose conforme ilustra a figura 19. Gás chave: hidrogênio.

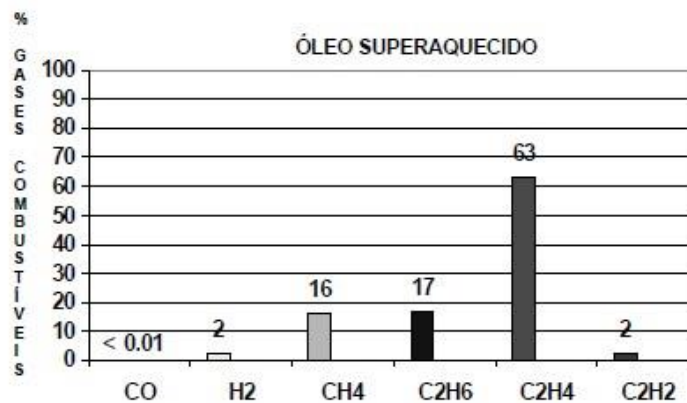
Figura 19 - Percentual de gases gerados durante descarga parcial no óleo



Fonte: Arantes (2005)

✓ **Aquecimento do óleo:** Gera Volume considerável de etileno e metano, e menor quantidade de hidrogênio se houver problema com isolamento do enrolamento. Pode haver geração de acetileno se o aquecimento for severo e houver contato elétrico conforme indicado na figura 20. Gás chave: etileno;

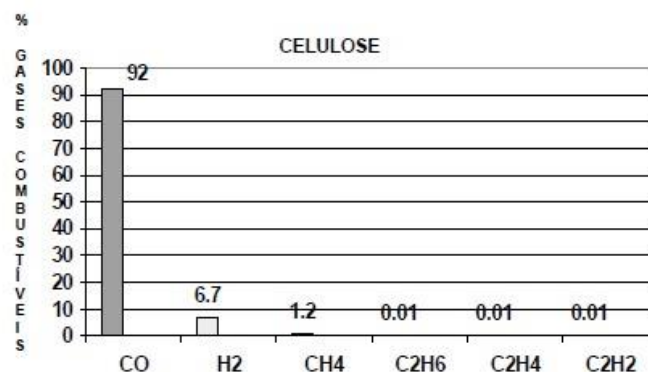
Figura 20 - Percentual de gases gerados durante sobre aquecimento



Fonte: Arantes (2005)

✓ **Degradação da celulose.** São gerados gases hidrocarbonetos, como metano e etileno, se houver a impregnação por óleo do papel isolante conforme indicado na figura 21. Gás chave: monóxido de carbono.

Figura 21 - Percentual de gases gerados na degradação da celulose



Fonte: Arantes (2005)

3.5 Métodos de diagnósticos

Para ajudar a detectar falhas no equipamento existem alguns métodos, de análises de gases presentes no óleo. Pode-se destacar os métodos de Rogers, NBR 7274, Doernenburg, Triângulo de Duval, que serão resumidamente descritos a seguir. Também uma pequena Abordagem do guia IEEE e da norma IEC realizada por Marques (2018) será apresentada.

3.5.1 Método de Rogers

Esse método é baseado no diagnóstico das relações dos gases CH_4/H_2 , $\text{C}_2\text{H}_6/\text{CH}_4$, $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$ e $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$ e limites de variação para as relações. O método *Rogers* possui várias alterações, em função de adaptações realizadas para ajustar a correlação do método com resultados obtidos das análises investigatórias nos equipamentos que sofrem falhas. (*Rogers*, 1975 *Apud* FILHO, 2012. p.47).

Com base no método de *Rogers* a IEEE C57.104, 2008 utiliza a Tabela 3 para relação de gases e a Tabela 4 para identificar as falhas, levando em consideração três relações de gases. A relação $\text{C}_2\text{H}_6/\text{CH}_4$ não é utilizada por indicar uma faixa limitada de temperatura de decomposição, não auxiliando na identificação de falhas.

Tabela 2- Relação de gases método Rogers

Código associado	R1(C ₂ H ₂ /C ₂ H ₄)	R2(CH ₄ /H ₂)	R5(C ₂ H ₄ /C ₂ H ₆)
0	0,1<R1<1,0	<0,1	<0,1
1	<0,1	0,1≤R2≤3,0	1,0≤R5≤3,0
2	>1,0	-	>3,0

Fonte: Guia IEEE para interpretação de gases em transformadores imersos em óleo mineral (2008)

Tabela 3 - Identificação de falhas

Código R2-R1-R5	Diagnóstico de falha sugerido
000	Nenhuma Falta Detectada
001	Baixa Temperatura de Sobrecarga
021	Falta Termica <700°C
022	Falta Termica >700°C
010	Descargas Parciais
102	Arco Elétrico

Fonte: Guia IEEE para carregamento de transformadores imersos em óleo mineral (2011)

3.5.2 Método NBR 7274

No método de diagnóstico de gases dissolvidos da NBR 7274 existem algumas limitações, mas em alguns casos como por exemplo problemas internos no equipamento com evolução mais lenta, ou seja, pontos quentes, sobreaquecimento nas conexões e demais cuja evolução não gera riscos maiores, o método da NBR 7274 se torna mais eficaz do que a cromatografia. Os defeitos com rápidas evoluções e com graves impactos para o equipamento, normalmente não são diagnosticados preventivamente pela cromatografia, pois geralmente estão ligados a fatores externos.

Quando é feita inspeções em equipamentos com problemas, geralmente são encontradas às seguinte classes de falhas:

✓ Descargas parciais do tipo de plasma frio (corona), resultando em possível cera depositada no isolamento celulósico ou do tipo de centelhamento, induzindo perfurações carbonizadas no papel, que podem não ser facilmente visualizadas;

- ✓ Centelhamento, ruptura dielétrica de alta densidade de ionização entre partes metálicas;
- ✓ Descargas de baixa energia em óleo e/ou papel, caracterizada pela larga carbonização perfurante através do papel, carbonização da superfície do papel ou partículas de carvão no óleo;
- ✓ Descargas de alta energia em óleo e/ou papel, com potência evidenciada pela extensiva destruição e carbonização do papel, fusão do metal, extensiva carbonização do óleo e em alguns casos, atuação da proteção do equipamento, confirmando a alta corrente;
- ✓ Sobreaquecimento, no óleo e/ou papel, abaixo de 300°C o papel torna-se escurecido (T1); acima de 300°C, o papel apresenta-se carbonizado (T2);
- ✓ Sobreaquecimento de temperatura acima de 700°C (T3), há evidência de carbonização do óleo, coloração do metal há 800°C ou fusão do metal acima de 1000°C.

Utilizando as três relações básicas de gases a ABNT 7274 conduz cada uma das sete falhas para uma caracterização padronizada, que é demonstrada na Tabela 5.

Tabela 4 - Tabela de interpretação de análise de gases dissolvidos

Caso	Falha característica	C ₂ H ₂ /C ₂ H ₄	CH ₄ /H ₂	C ₂ H ₄ /C ₂ H ₆
PD	Descargas parciais	NS	<0,1	<0,2
C1	Centelhamento	0 - 0,6	0,1 - 1	<1
D1	Descargas de baixa energia	<1	0,1 - 0,5	>1
D2	Descargas de alta energia	0,6 - 2,5	0,1 - 1	<2
T1	Sobreaquecimento, t<300°C	<0,01	>1	<1
T2	Sobreaquecimento, 300°C<t<700°C	<0,1	>1	1 - 4
T3	Sobreaquecimento, t>700°C	<0,2	>1	>4

Fonte: Guia de interpretação da análise dos gases livres e dissolvidos (ABNT, 1982)

3.5.3 Método de Doernenburg

O método de *Doernenburg* assemelha-se ao de *Rogers* utiliza relação da concentração dos gases dissolvidos. Utiliza duas relações principais CH₄/C₂H₂,

C_2H_2/C_2H_4 e duas auxiliares C_2H_6/C_2H_2 , C_2H_2/CH_4 . Estabelece condições para a validação do método em função da concentração dos gases componentes da relação, não estabelecendo critérios para caracterização do isolamento estando em condições normais ou não. Dados apresentados na Tabela 6 (ZIRBES *et al*, 2005).

Tabela 5 - Tabela de interpretação de análise de gases dissolvidos

Tipos de falha	Relações entre concentrações de gases			
	Principais		Auxiliares	
	CH_4/C_2H_2	C_2H_2/C_2H_4	C_2H_6/C_2H_2	C_2H_2/CH_4
Ponto quente	>1	<0,75	>0,4	<0,3
Descargas parciais	<0,1	Não significativo	>0,4	<0,3
Arco (PD de alta intensidade)	0,1 - 1	>0,75	<0,4	>0,3

Fonte: ZIRBES *et al* (2005)

Para utilizar o método das relações é necessário que um dos gases das relações principais tenha uma concentração superior ao dobro do valor na Tabela 5 e que para as demais relações ao menos um dos gases tenha uma concentração superior ao que consta na Tabela 7 (ZIRBES *et al*, 2005).

Tabela 6 - Dados de concentração individual dos gases para verificação da validade de aplicação do método

Tipo de gás	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Concentração ppm (v/v)	200	50	15	60	15

Fonte: ZIRBES *et al* (2005)

3.5.4 Método Triângulo de Duval

Desenvolvida em 1960 essa técnica analisa a concentração dos gases metano, etileno e acetileno dissolvidos no óleo, possibilitando assim a identificação de faltas no transformador. Esse método é aplicado nos transformadores quando os três gases citados aparecem juntos, mesmo em pequenas concentrações caracteriza-se um perigo eminente para o equipamento (ARANTES, 2005).

Para utilizar o método, primeiro deve-se determinar o problema utilizando os métodos de *Rogers e Doernenburg*. Após resultados é usada uma tabela para avaliar a possibilidade de aplicar o método de Duval. A Tabela 8 contempla os valores limites (L1) e as taxas de geração de gases (G1 e G2). As concentrações de gases devem ter pelo menos um gás no limite ou acima do nível (L1) e a taxa de geração estar ao menos no limite (G2).

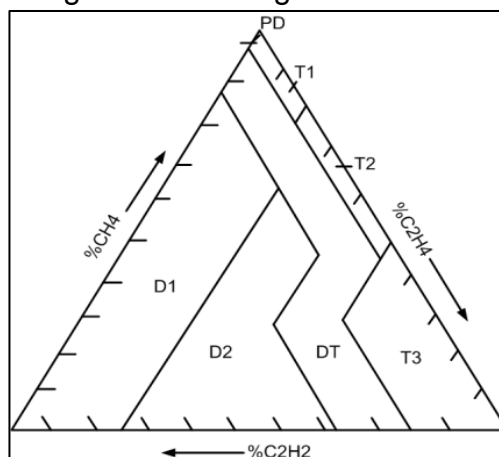
Tabela 7 - Limite e taxas de geração de gases.

Gás	Limite L1 (ppm/mês)	Limite G1 (ppm/mês)	Limite G2 (ppm/mês)
H ₂	100	10	50
CH ₄	75	8	38
C ₂ H ₂	3	3	3
C ₂ H ₄	75	8	38
C ₂ H ₆	75	8	38
CO	700	70	350
CO ₂	7000	700	3500

Fonte: Arantes (2005)

Ao se comprovar a existência da falha, o triângulo de Duval é criado pelas porcentagens de metano, etileno e acetileno. A falha é diagnosticada pelo Gráfico triangular, apresentado pela Figura 22. Permitindo identificar três falhas de origem elétrica, três de origem térmica e uma simultânea. Sendo identificadas a como descargas parciais (PD), faltas térmicas abaixo de 300 °C (T1), faltas térmicas entre 300 ° e 700 °C (T2), faltas térmicas maiores que 700 °C (T3), descarga de baixa energia (D1), descarga de alta energia (D2) e ocorrência simultânea de falta térmica e arco elétrico (DT).

Figura 22 - Triângulo de Duval



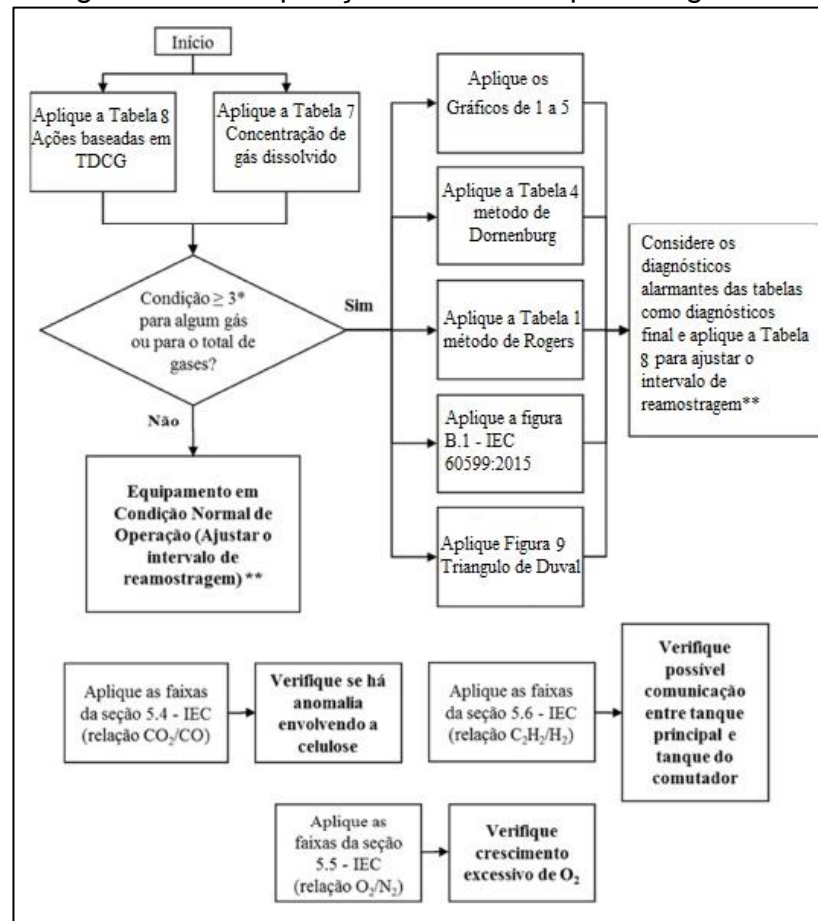
Fonte: Gomez (2013)

3.5.5 Guia IEE Norma IEC

Utiliza-se também a combinação do guia IEEE e da norma IEC para interpretação de análises de gases dissolvidos em óleo mineral isolante.

Esse método apresenta uma abordagem de interpretação utilizando o guia do IEEE C57.104-2008 e a norma IEC 60599: 2015. Segundo Marques (2018), nessa abordagem os resultados alcançados apresentam taxas de acerto mais elevadas do que a utilização de cada método separadamente, proporcionando uma melhor identificação das falhas, o fluxograma dessa metodologia é apresentado na Figura 23 e seus respectivos resultados apresentados nas tabelas 9, 10 e 11.

Figura 23 - Fluxograma da composição IEEE e IEC para diagnostico de falhas



Fonte: Marques (2018)

Tabela 8 - Concentração de gás dissolvido

Status	Limites de concentração de gás chave dissolvidos [$\mu\text{L/L}$ (ppm)]							
	Hidrogênio	Metano	Acetileno	Etileno	Etano	Monóxido de Carbono	Dióxido de Carbono	TDCG
Condição 1 Classificação A	100	120	1	50	65	365	2500	720
Condição 2 Classificação B	101-700	121-400	2-9	51-100	66-100	351-570	2500-4000	721-1920
Condição 3 Classificação C	701-1800	401-1000	10-35	101-200	101-150	571-1400	4001-10000	1921-4630
Condição 4 Classificação D	>1800	>1000	>35	>200	>150	>1400	>10000	>4630
Classificação E	Quando os valores excederem em muito aos correspondentes à classificação D							

Nota: A tabela 1 assume que nenhum teste anterior no transformador para análise de gás dissolvido foi feito ou que não existe histórico recente. Se houver uma análise anterior, ela deverá ser revisada para determinar se a situação é estável ou instável. Consulte 6.5.2 para a(s) ação(ões) apropriada(s) a ser(em) tomada(s).

Fonte: Marques (2018)

Tabela 9 - Ações baseadas em TDCG

	TDCG Níveis (µL/L)	TDCG Taxa (µL/L/dia)	Intervalos de amostragem e procedimentos operacionais para taxas de geração de gases	
			Intervalo de amostragem	Procedimentos operacionais
Classificação E	Quando os valores excederem em muito aos correspondentes à classificação D			
Condição 4 Classificação D	>4630	>30	Diário	Considere retira de serviço.
		10 a 30	Diário	Informar fabricante
		<10	Semanal	Tome muito cuidado. Analisar gases individualmente. Planeje a interrupção. Informar fabricante
Condição 3 Classificação C	1921 a 4630	>30	Semanal	Tome muito cuidado.
		10 a 30	Semanal	Analisar gases individualmente. Planeje a interrupção.
		<10	Mensal	Informar fabricante
Condição 2 Classificação B	721 a 1920	>30	Mensal	Tome cuidado.
		10 a 30	Mensal	Analisar gases individualmente.
		<10	Trimestral	Determinar a dependência de carga
Condição 1 Classificação A	720	>30	Mensal	Tome cuidado. Analisar gases individualmente. Determinar a dependência de carga
		10 a 30	Trimestral	
		<10	Anual	Continue a operação normalmente

Fonte: Marques (2018)

Tabela 10 - Classificações finais e ações recomendadas

Classificação	Significado	Ações Recomendadas
A	Excecente	Continuar a operar o equipamento normalmente
B	Bom	Continuar a operar o equipamento normalmente, estando atento à elevação de resultados nos próximos registros
C	Marginal	Investigar e realizar outros ensaios em curto prazo para confirmar resultados e tendências
D	Ruim	Programar retirada do equipamento de operação para inspeção interna, localizada e correção de defeito
E	Péssimo	Retirar o equipamento de operação em caráter de emergência para inspeção interna, localizada e correção de defeito

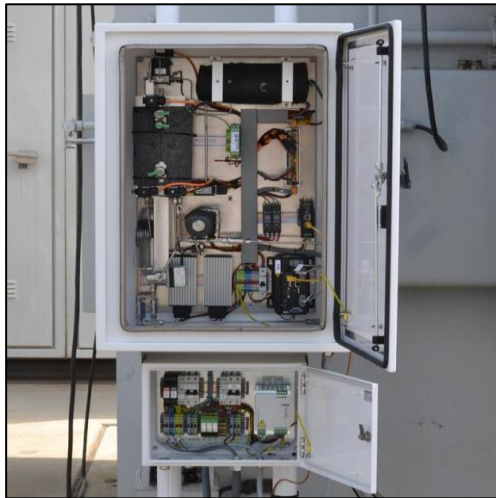
Fonte: Marques (2018)

Segundo Marques (2018), usando a proposta de combinar o guia IEEE e a norma IEC para análise de gases dissolvidos os resultados alcançados apresentam taxas de acerto de 92,3% para diagnósticos corretos, obtendo assim taxa de acerto mais elevada do que utilizar cada método separadamente.

O método do triangulo de Duval que teve o segundo melhor aproveitamento obteve taxa de acerto de 76,9% pode ser considerado como resultados aceitáveis, mais ainda inferiores a abordagem combinada. Ficando assim demonstrado ser apropriado o uso da abordagem combinada para análises de gases dissolvidos, por auxiliar de melhor forma na identificação da falha.

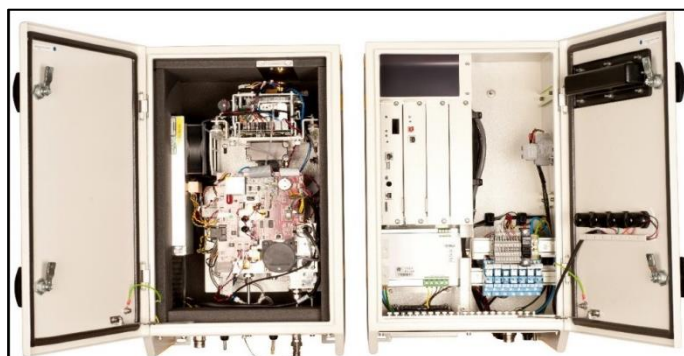
No entanto esses métodos mais avançados e precisos tem uma abrangência maior em transformadores de potência, além dos citados acima existem também o sistema de medição on-line de gás dissolvido, que utilizam aparelhos como o medidor OPT 100 via sala, além do sistema da GE, KELMAN DGA 900, mostrados na figura 24 e figura 25.

Figura 24 - Monitor de gases on-line



Fonte: Vaisala (2018)

Figura 25 - Kelman DGA 900



Fonte: GE GRID Solutions (2018)

3.6 REDES INTELIGENTES

A crescente demanda por eletricidade, os requisitos de melhoria da qualidade de energia por parte dos agentes reguladores e a inclusão da geração distribuída estão aumentando a complexidade dos sistemas de distribuição.

No contexto da busca em torno da melhoria contínua do fornecimento de energia elétrica, surge a ideia das Smart grids ou em uma tradução livre, Redes Inteligentes. As redes de energia inteligentes, no âmbito das redes de distribuição de energia elétrica, têm o objetivo de melhorar o fornecimento com o uso de equipamentos que no momento em que uma falha ocorra, a concessionária saiba

onde e como ocorreu a pane, deslocando equipes de forma mais eficiente para corrigir o problema e por consequência melhorando o fornecimento de energia e os índices coletivos de continuidade (ABOBOREIRA e CRUZ, 2016).

No contexto desse trabalho, os conceitos de redes e sistemas de inteligência computacional, podem ser utilizados, tanto para fazerem modelagem, quanto para fazerem classificação de padrões. Isso porque problemas de detecção de faltas podem ser tratados como um problema de classificação, e seus padrões podem ser obtidos por meio de modelos.

3.7 MINERAÇÃO DE DADOS

A mineração de dados é o processo de descoberta de informações acionáveis em grandes conjuntos de dados. A mineração de dados usa análise matemática para derivar padrões e tendências que existem nos dados. Normalmente, esses padrões não podem ser descobertos com a exploração de dados tradicional pelo fato de as relações serem muito complexas ou por haver muitos dados.

Segundo Possa et al (1998), é comprovado que o cérebro humano, consegue fazer até oito comparações ao mesmo tempo. A mineração de dados amplia essa comparação para “infinito” e tornar isso visível ao olho humano.

Segundo Castanheira (2009), pode-se concluir que a Mineração de dados se caracteriza pela existência de um algoritmo que, diante da tarefa proposta, será eficiente em extrair conhecimento implícito e útil de um banco de dados. Pode-se dizer que mineração de dados é a fase que transforma dados puros em informações úteis.

3.7.1 Tarefas

Segundo Larose (2005), a mineração de dados é comumente classificada pela sua capacidade em realizar determinadas tarefas. As mais comuns são:

- ✓ **Descrição (Description):** Descreve padrões e tendências descobertos pelos dados. A mesma disponibiliza uma possível interpretação para os resultados encontrados, ela pode ser utilizada em conjunto com as técnicas de análise exploratória, pois comprova a influência de certas variáveis no resultado encontrado;

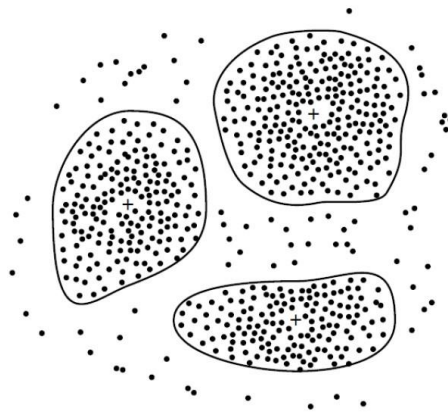
✓ **Classificação (Classification):** Identifica a que classe o registro pertence. Nessa etapa o modelo analisa os registros fornecidos com os demais a fim de identificar a qual classe o mesmo pertence, a fim de 'aprender' como classificar um novo registro. Abaixo alguns exemplos de como essa tarefa pode ser utilizada. Identificar quando uma transação no cartão de crédito pode ser uma fraude, diagnosticar onde uma determinada doença pode estar presente e identificar em uma escola, qual a turma mais indicada para aquele aluno;

- **Estimação (Estimation):** Essa se assemelha a classificação, mais é utilizada quando o registro é identificado por um valor numérico e não categórico. Essa tarefa pode ser usada por exemplo em: Estimar a quantia a ser gasta por uma família de 5 pessoas na volta às aulas e estipular a pressão ideal de um paciente baseando-se na idade, sexo e massa corporal;

- **Predição (Prediction):** Tem as mesmas características da classificação e estimação, porém ela tem a finalidade de descobrir o valor futuro de um determinado atributo. Exemplo: Predizer o campeão de um campeonato, baseando-se na estatística dos times e indicar o valor de uma ação três meses adiante;

✓ **Agrupamento (Clustering):** Visa encontrar e aproximar os registros similares. Ela é diferente da classificação, pois não precisa que os registros sejam categorizados. Ela apenas identifica os grupos de dados similares assim como mostra a figura 26 (HAN; KAMBER, 2006). Como por exemplo: Em auditoria, separando comportamentos suspeitos.

Figura 26 - Registro agrupados em três clusters



Fonte: Han Kamber (2006)

3.7.2 Descoberta do Conhecimento

Nas últimas décadas houve um salto na capacidade de processamento e armazenamento de dados, contribuindo para a informatização e automação dos processos nos mais diversos setores. Os processos de reconhecimento de padrões ficaram estagnados frente ao avanço dos sistemas informáticos, para isso, foram desenvolvidos processos fundamentados em KDD, auxiliando o reconhecimento de padrões e gestão dos sistemas. Processos esses, que fizeram a conexão entre os sistemas informáticos e o reconhecimento de padrões (CEMBRANEL, 2019).

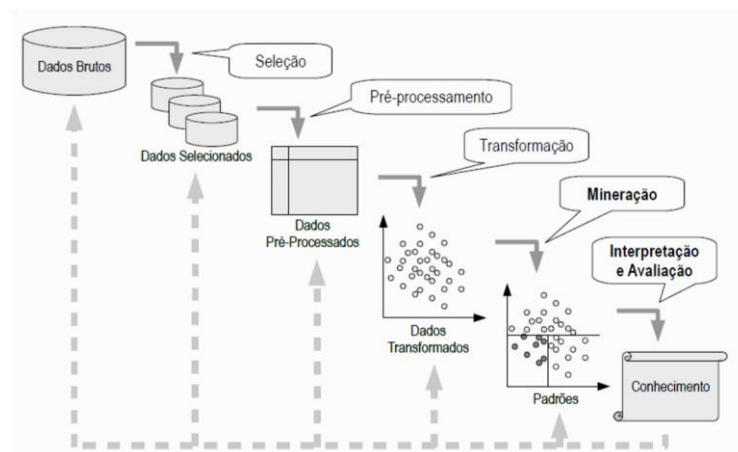
Segundo Cembranel (2019), DM nada mais é do que a formalização matemática de técnicas que auxiliam na identificação de padrões em bancos de dados, utilizada em processos que o trabalho humano de identificação de padrões é inviável em tempo hábil e errôneo (devido ao elevado número de dados analisados). Sua utilização no mundo real está bem consolidada apresentando diversas aplicações em áreas distintas, tais como: marketing, investimentos, detecção de fraudes, manufatura, telecomunicações, sistemas elétricos, entre outros.

3.7.3 Processos da descoberta do conhecimento em banco de dados

Devido às inconsistências presentes nos bancos de dados reais, os algoritmos de reconhecimento de padrão não podem ser aplicados diretamente aos dados. Por esse motivo foram desenvolvidos processos de KDD. As principais

inconsistências presentes nos bancos de dados são: ruídos, falta de valores e dados duplicados. Os erros nos dados acabam por distorcer os resultados e os padrões encontrados podem não retratar corretamente os perfis dos consumidores. Devido a isso, é feita a divisão do processo de KDD, diminuindo a chance de erros, e má interpretação dos dados ao longo do processo. A divisão do processo pode ser feita em cinco etapas como mostra a figura 27.

Figura 27 - Processo de Descoberta de Conhecimento através de Banco de Dados



Fonte: Santos (2008)

- ✓ **Seleção dos dados:** Nessa etapa os dados são selecionados de acordo com a análise a ser desenvolvida;
- ✓ **Pré-processamento:** Essa etapa é justificada pela existência de inconsistências nos bancos de dados, por exemplo ruídos, valores duplicados e valores inexistentes. Fundamental no processo de KDD, se o pré-processamento não for realizado corretamente pode acarretar na descoberta de padrões inexistentes ou inconsistentes, resultando em erros na interpretação do problema;
- ✓ **Estimativa de dados em falta:** Nesta etapa é feita a identificação e estimativa dos dados em falta, para corrigir os dados faltantes podem ser utilizadas diversas técnicas. São apresentadas algumas técnicas utilizadas na correção desse erro, tais como:

- ✓ Completar o valor em falta manualmente;
- ✓ Estimar uma constante global para substituir os valores faltantes;
- ✓ Usar a média dos pontos vizinhos para estimar o valor faltante;

- ✓ Calcular a média de todas as amostras e atribuir ao valor faltante;
- ✓ Usar técnicas de regressão para encontrar o valor mais provável e substituir o valor faltante;

✓ **Eliminação de ruídos:** Ruídos são variáveis aleatórias ou valores discrepantes que estão presentes nos bancos de dados reais. Os mesmos podem comprometer a mineração, pois esses dados aleatórios podem causar má interpretação dos dados reais. Nos sistemas elétricos os ruídos podem ocorrer por fatores como: perturbações na rede e defeitos nos equipamentos de medição;

✓ **Integração dos dados:** A integração de dados consiste em integrar dados com outros bancos de dados, pois apenas um sistema de descoberta independente pode ser suscetível a erros. A etapa de integração de dados inclui problemas devido à incompatibilidade dos bancos de dados. Devido aos problemas de integração, estima-se que 70% do tempo gasto no processo de KDD é devido à integração e preparação dos bancos de dados;

✓ **Redução dos dados:** Bancos de dados possuem grande volume de dados, tornando o processo de KDD complexo e demorado. Técnicas de redução de dados devem ser utilizadas para reduzir o volume de dados para um volume muito menor, facilitando a análise;

✓ **Transformação dos dados:** A transformação dos dados é efetuada para tornar o processo de mineração mais eficiente e facilitar a compreensão dos dados, a transformação pode ser feita por diversas técnicas, como:

✓ Suavização (smoothing): visa remover ruídos e utiliza métodos como binning, regressões e agrupamento;

✓ Construção de atributos: novos atributos são desenvolvidos e adicionados para auxiliar o processo de mineração;

✓ Agregação: utiliza a agregação dos dados, por exemplo, dados diários generalizados para dados semanais;

✓ Normalização: troca de escala para os dados, os mesmos são normalizados, por exemplo, para valores entre 0 e 1;

✓ Desratização: os valores numéricos são trocados por rótulos de intervalo, como exemplo 0-5, 6-10, etc.;

✓ Geração de hierarquia de conceito para dados nominais: os atributos são generalizados para que o processo possa ser mais eficiente.

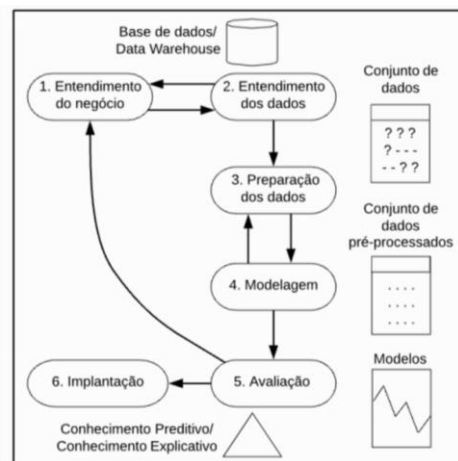
Devido à característica dos dados que serão estudados nesse trabalho (dados numéricos), a transformação de dados mais adequada na análise do trabalho em questão é a normalização. A normalização visa dar aos dados o mesmo peso, um dos benefícios da normalização é a transformação da escala dos dados.

✓ **Mineração dos dados:** A etapa de mineração de dados consiste na análise do banco de dados, e na construção de modelos explicativos. Nessa fase é feita a descoberta do conhecimento através da identificação dos padrões existentes nos dados. O modelo é avaliado, medindo a capacidade que o mesmo tem para explicar os dados conhecidos ou dados ainda desconhecidos.

Já o CRISP-DM é uma metodologia para Mineração de Dados que propõe aumentar a taxa de sucesso de processos de Data Mining. Segundo a IBM (2017), “CRISP-DM, que significa Processo Padrão de Vários Segmentos de Mercados para Mineração de Dados, é uma forma comprovada pelo mercado para orientar seus esforços de mineração de dados”.

O modelo proposto no CRISP-DM consiste de seis fases que são organizadas de maneira cíclica (OLSON, 2008). Segundo Moro (2011, p.2) “CRISP-DM define um projeto como um processo cíclico, em que várias interações podem ser usadas para permitir um resultado ajustado aos objetivos de negócio”. Além disso, ao se trabalhar como o modelo, é possível a transição entre fases sem seguir um fluxo, sendo assim um modelo unidirecional, conforme demonstra a figura 28.

Figura 28 - Processo CRISP



Fonte: Moro (2011)

As fases propostas na metodologia do CRISP-DM são: Entendimento do negócio, Entendimento dos Dados, Preparação dos Dados, Modelagem Avaliação e Implantação. Abaixo será apresentado de forma resumida as propostas da metodologia.

✓ **Entendimento dos Negócios:** O foco é entender qual o objetivo se quer atingir com a mineração dos dados. “A partir do entendimento do projeto e dos requisitos em uma perspectiva de negócio, converte-se esse conhecimento em uma definição de problema de data mining e em um plano preliminar para alcançar os objetivos” (CHAPMAN e outros, 2005, p. 10).

✓ **Entendimento dos Dados:** Deve-se descrever de forma clara o problema, identificar os dados relevantes para o problema em questão e certificar-se que as variáveis relevantes não são interdependentes. Conforme Camilo e Silva (2009) “As fontes fornecedoras dos dados podem vir de diversos locais e possuírem diversos formatos”. Com base nos dados adquiridos cabe ao analista separá-los e verificar se estes se encaixam no modelo proposto de trabalho no entendimento do negócio.

✓ **Preparação dos Dados:** seu propósito é limpar os dados selecionados de modo a obter-se melhor qualidade, tendo em vista que alguns dos dados selecionados podem seguir diferentes padrões por conta de serem coletados de diferentes fontes (OLSON, 2008). Nessa fase ocorre a seleção de atributos, tratamento de valores faltosos, erros em partes dos dados, padronização, formatação e a divisão dos dados em conjuntos de treinamento e conjuntos de teste

(Hiragi, 2009). Com isso os dados coletados nessa fase já estão prontos para serem utilizados na mineração de dados.

✓ **Modelagem:** ocorre a aplicação dos dados e das técnicas de mineração de dados são aplicadas. “Nesta fase, algoritmos de aprendizagem de máquina mais adequados a cada cenário são configurados para construir modelos aderentes e compatíveis com os dados preparados” (CARVALHO, 2016, p.11).

✓ **Avaliação:** “Considerada uma fase crítica do processo de mineração, a avaliação é a etapa que conta com a participação de especialistas nos dados, conhecedores do negócio e tomadores de decisão. ” (CAMILO; SILVA, 2009, p. 5). Como forma de auxílio ao processo da avaliação são utilizados gráficos para verificar os resultados obtidos nas etapas anteriores, dessa forma garante a confiabilidade dos modelos é indicada a realização de testes e validações dos modelos construídos. É nesta etapa que os conhecimentos dos dados minerados são assimilados e para isso há duas questões essenciais (ALVES, 2018).

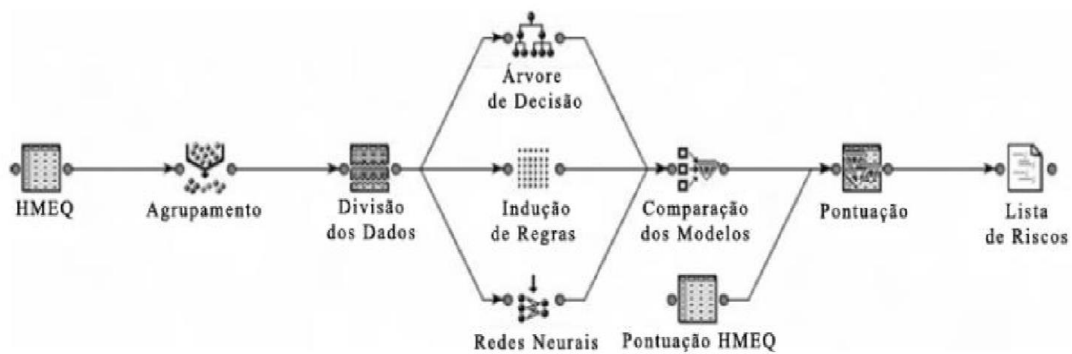
Diante disso, a implantação dos resultados do projeto pode atuar como fator crucial sobre os negócios dos envolvidos ao auxiliar na gerência do negócio. Desta forma Chapman e outros (2005) ressalta que dependendo dos requisitos, a fase de implantação pode ser tão simples como gerar um relatório ou tão complexa quanto implementar um processo repetitivo de mineração de dados por toda a empresa e que em muitos casos é o cliente, e não o analista de dados, que dá realiza os passos da implantação.

3.7.4 Métodos (Técnicas)

Normalmente os métodos de Mineração de dados são divididos em dois modelos, o aprendizado supervisionado (preditivo) e o não supervisionado (descritivo). Apesar do limite dessa divisão ser muito tênue (alguns métodos preditivos podem ser descritos e vice-versa) ele ainda é interessante para fins didáticos (FAYYAD,1996).

Durante o processo de mineração, diversas técnicas devem ser testadas e combinadas conforme mostra a figura 29, a fim de que comparações possam ser feitas e então a melhor técnica (ou combinação de técnicas) seja utilizada.

Figura 29 - Processo de comparação com algumas técnicas



Fonte: Mccue (2007)

3.7.5 Mineração de itens Frequentes

Essa técnica se distingue em duas etapas, na primeira o conjunto de itens frequentes é criado, logo após as regras de associação são criadas pela mineração desse conjunto. São utilizados conceitos de suportes e confiança para garantir resultados validos. Um dos mais tradicionais algoritmos de mineração utilizando a estratégia de itens frequentes é o Apriori (AGRAWAL,1993).

Diversas variações deste algoritmo, envolvendo o uso de técnicas da hast, redução de transações, particionamento e segmentação podem ser encontrados (AGRAWAL,1993). (CASANOVA,2005), usa o algoritmo da confiança inversa junto com a lógica Nebulosa para gerar regras mais precisas. Borgelt (2000), apresenta uma implementação do FP-growth e faz a comparação dele com outros três algoritmos, dentre eles o Apriori e o Eclat. Possas et al (2000), propõe uma variação do algoritmo Apriori a fim de que um número menor de regras seja gerado, os resultados apresentam até 15% de redução. Vasconcelos (2004), mostra o uso do Apriori para mineração de dados da Web.

3.7.6 Árvores de Decisão

Segundo Castanheira (2009), a árvore decisão são representações simples do conhecimento e um meio eficiente de construir classificadores que predizem classes baseadas nos valores de atributos de um conjunto de dados.

Uma árvore de decisão utiliza uma técnica chamada de Top-Down, que consiste em dividir para conquistar, a mesma divide um problema em diversos

subproblemas por diversas vezes até encontrar uma solução para o devido problema.

Um algoritmo que foi escolhido pela árvore compara e examina os resultados.

De acordo com Garcia (2003), os resultados obtidos após a construção de uma árvore de decisão são dados organizados de maneira compacta, com a árvore podendo ser utilizada para classificar novos casos.

Segundo Luciana Castanheira (2009), o conceito de entropia é uma medida de informação calculada pelas probabilidades de ocorrência de eventos individuais ou combinados. Pode-se dizer que a entropia é dada como medida da impureza em um conjunto arbitrário de amostras de treinamento. Pode ser considerada a medida da quantidade de desordem de um conjunto de amostras.

3.7.7 Classificação Bayesiana

É um método estatístico usado para revisar as estimativas de probabilidade a partir de novas informações. Revisar ou reavaliar as estimativas iniciais de probabilidade (risco) assim que novas informações tornam-se disponíveis são de grande importância para a tomada de decisão (aceitar ou rejeitar o projeto). A base fundamental para a Análise Bayesiana é um teorema desenvolvido pelo pastor e matemático inglês, Thomas Bayes, que é apropriadamente chamado de Teorema de Bayes (OLIVEIRA; RIBEIRO, 2018). O algoritmo de Naive Bayes parte do princípio que não exista relação de dependências entre os atributos. No entanto nem sempre é possível. Nesse caso uma variação conhecida como Bayesian Belief Networks, ou Bayesian Networks (NEIDERMAYER, 1998) deve ser utilizada.

3.7.8 Redes Neurais

Segundo Castanheira (2009), rede neural artificial (RNA) é uma técnica que constrói um modelo matemático, de um sistema neural biológico simplificado, com capacidade de aprendizado, generalização, associação e abstração. Assim como no cérebro humano, as redes neurais apresentam estrutura altamente paralelizada, composta por processadores simples (neurônios artificiais) conectados entre si.

As RNA vêm sendo desenvolvidas como generalização de modelos matemáticos do conhecimento humano, baseado nas seguintes suposições: O

processamento da informação ocorre através de muitos elementos simples, chamados neurônios; Os sinais são transmitidos entre os neurônios através de um elo de conexão; Cada elo de conexão tem um peso associado, o qual, em redes neurais típicas, é múltiplo do sinal transmitido; Cada neurônio aplica uma função de ativação (normalmente não linear) em sua rede de entrada (soma dos pesos multiplicados pelos sinais de entrada) para determinar seu sinal de saída (FAUSETT, 1994).

De acordo com Haykin (2001), uma propriedade importante das redes neurais é a sua habilidade para aprender a partir do ambiente na qual estão inseridas, ou ambiente de aprendizado, e melhorar seu desempenho através da aprendizagem.

As RNA's tentam aprender por experiência, ou seja, diretamente dos dados, através de um processo de repetidas apresentações dos dados a rede.

3.7.9 SVM (Support Vector Machines)

No ano de 1992 que o primeiro artigo referente a esse método foi apresentado por Vladimir Vapnik, Bernhard Boser e Isabelle Guyon (BOSER,1992), e uma técnica relativamente nova e tem alcançado ótimos resultados, o índice de assertividade é bem elevado, ela permite modelar situações não lineares complexas que geram modelos de interpretação simples. Essa técnica é utilizada na classificação como também na predição, o único problema atualmente é o tempo gasto na fase de aprendizado, mas existem estudos que visam solucionar esse problema.

3.7.10 Classificação por Regras de Associação (Classification by Association Rule)

Tem como foco buscar padrões de associações elevados entre os itens e as categorias (utilizando o conceito da frequência). Baseia-se em dois passos, no primeiro os dados de treinamento são observados para descobrir os itens mais frequentes, logo após esses itens são utilizados para criação das regras. Estudos demonstram que essa técnica está tendo mais sucesso do que algoritmos tradicionais como: CBA (Classification - Based Association) (LIU, B.HSU, W; MA, Y, 1998).

3.7.11 Aprendizado Tardio (Lazy Learners)

A maioria das técnicas de classificação usam um conjunto de dados de treinamento para classificar um registro, assim quando enxergam outro registro elas já conseguem classificar esse dado diretamente pois elas já aprenderam anteriormente. Mas também existe uma outra categoria de métodos que só realiza o aprendizado quando for solicitada, essa recebe o nome de aprendizado tardio. Seu tempo de treinamento é muito menor, no entanto ela precisa de muito recurso computacional para armazenar e recuperar os dados de aprendizagem. Mais esses métodos permitem o aprendizado incremental.

O algoritmo conhecido como KNN (K- Nearest Neighbor) armazena os dados de treinamento e quando um novo item é submetido para classificação, o algoritmo procura os k mais próximos desse novo registro. E o registro é classificado na classe mais comum entre os k mais próximos. Atualmente esse algoritmo tem sido muito utilizado na área de suporte a usuários no ramo da medicina, engenharia e direito. (CAMILO, 2009).

3.7.12 Algoritmo Genético (Genetic Algorithm)

Algoritmos genéticos tem como base a teoria da evolução, no início uma população é formada de forma aleatória. Seguindo a lei da evolução (mais forte), uma outra população é construída com base na atual, pois os indivíduos passam por processo de troca genética e mutação. Este processo continua até que populações com indivíduos mais fortes sejam gerados ou que se atinja algum critério de parada (CAMILO, 2009).

3.7.13 Algoritmo Evolucionário (Differential Evolution)

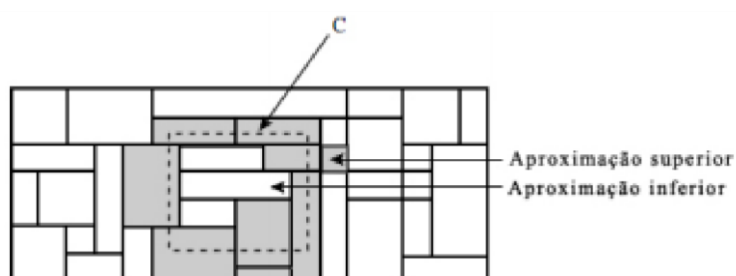
Diferentes tipos de problemas podem ser resolvidos pela computação evolucionária. Muitos problemas são de otimização (numérica ou combinatória), outros são de síntese de um objeto (programa de computador, circuito eletrônico) e, em outros contextos, procura-se um modelo que reproduza o comportamento de determinado fenômeno (machine learning).

Para vários desses problemas, é frequentemente possível encontrar um algoritmo que ofereça uma solução “ótima” ou uma boa aproximação desta solução. Enquanto alguns algoritmos requerem informações auxiliares, como derivadas, que muitas vezes não estão disponíveis ou são difíceis de se obter, a computação evolucionária dispensa informações auxiliares e oferece algoritmos gerais (i.e., algoritmos genéticos, programação genética, estratégias evolutivas e programação evolutiva) que são aplicados em problemas complexos, com grandes espaços de busca, de difícil modelagem, ou para os quais não há um algoritmo eficiente disponível.

3.7.14 Conjuntos Aproximados (Rought Set)

Consegue realizar a classificação mesmos se os dados estiverem imprecisos ou errados, a mesma é utilizada para valores discretos. Através da classe da equivalência eles usam a ideia de aproximação para criar as categorias, pois acreditam que os elementos da classe são indiscerníveis. (CAMILO, 2009). Uma representação de conjuntos aproximados pode ser vista na figura 30.

Figura 30 - Conjuntos Aproximados



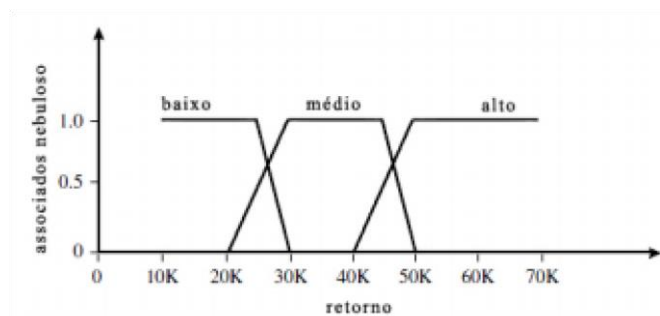
Fonte: Camilo (2009)

3.7.15 Conjuntos Nebulosos (Fuzzy Set)

O princípio dos conjuntos fuzzy é de que as variáveis sejam discretizadas em categorias e a lógica nebulosa aplicada para definição dos limites dessas categorias, ao invés de realizar o corte direto. Portanto não se consegue ter as categorias com

limites de corte bem definido, no entanto consegue-se um grau de flexibilidade entre as categorias conforme mostra a figura 31(CAMILO, 2009).

Figura 31 - Conjuntos Nebulosos



Fonte: Camilo (2009)

3.7.16 Métodos de Particionamento (Partitioning Methods)

Segundo Camilo (2009), K-Means é um algoritmo que utiliza o conceito de centroide, perante um conjunto de dados o algoritmo busca de forma aleatória k registros, cada um referente a um agrupamento, para os demais e calculado uma similaridade entre o registro analisado e o centro de cada agrupamento.

K- Medoids é a variação do K-Means. Nesse algoritmo, ao invés de calcular o centro do agrupamento e usá-lo como referência, trabalha-se com o conceito do objetivo mais central do agrupamento (CAMILO, 2009).

4.0 Formulação dos problemas

Segundo Moretti (2010), modelagem matemática é uma técnica de representação quantitativa de processos e problemas reais. Usualmente, classificam-se os modelos matemáticos como prescritivos ou descritivos como mostra a figura 32:

Figura 32 - Modelos matemáticos como prescritivos ou descritivos



Fonte: Moretti (2010)

Ainda segundo Moretti (2010), os modelos prescritivos baseiam-se na representação dos objetivos e restrições de um processo para o qual se deseja descobrir soluções otimizadas, ou seja, o modelo é escrito segundo uma técnica que permite encontrar a melhor solução ou política de ação para os condicionantes representados. Já os modelos descritivos são utilizados na representação de sistemas reais (ou propostos) e a experimentação de diferentes cenários e políticas de ação nos mesmos, onde permitem a flexibilidade na representação de modelos complexos e a facilidade de aplicação, possibilitando prever o comportamento de um sistema modelado em um horizonte de planejamento escolhido.

Na maioria das vezes, no processo de modelagem do problema, é necessário fazer simplificações, ou porque o problema não tem todos os dados conhecidos ou simplesmente para facilitar a resolução do modelo.

Para Holtz (2005) diante essas necessidades, um problema de modelo matemático de otimização está sempre associado a um problema de minimização ou maximização de uma ou mais funções, geralmente, existe um conjunto de restrições nas variáveis, como representado a formulação geral de um modelo matemático:

$$\text{Maximizar / Minimizar } f(x)$$

Sujeito a:

$$g(x) \geq a$$

$$h(x) \leq b$$

$$s(x) = c$$

$$l \leq x \leq u$$

Em que $f(x)$ é uma função, linear ou não linear, a ser otimizada (maximizada ou minimizada) sujeito às restrições (lineares ou não lineares) que indicam uma necessidade a ser satisfeita (“maior ou igual”), uma disponibilidade a ser respeitada (“menor ou igual”) ou um condicionante a ser satisfeito de maneira exata (“igual”).

Ainda segundo Holtz (2005), além dessas associações e estes conjuntos de regras, faz-se necessário o uso de algoritmos para a solução de um problema de otimização, que podem ser, basicamente, determinísticos ou probabilísticos, que serão detalhados nas próximas sessões.

Para melhor entendimento dos algoritmos de otimização, faz-se necessário o conhecimento de alguns conceitos e definições utilizados na literatura (BASTOS, 2004).

A seguir são listados alguns termos usualmente relacionados a um problema de otimização qualquer:

- ✓ **Variáveis de projeto:** São aquelas que se alteram durante o processo de otimização, podendo ser contínuas (reais), inteiras ou discretas;
- ✓ **Restrições:** São funções de igualdade ou desigualdade sobre as variáveis de projeto que descrevem situações de projeto consideradas não desejáveis;
- ✓ **Espaço de busca:** É o conjunto, espaço ou região que compreende as soluções possíveis ou viáveis sobre as variáveis do projeto do problema a ser otimizado, sendo delimitado pelas funções de restrição;
- ✓ **Função Objetivo:** É a função de uma ou mais variáveis de projeto que se quer otimizar, minimizando-a ou maximizando-a;
- ✓ **Ponto Ótimo:** É o ponto formado pelas variáveis de projeto que extremizam a função objetivo e satisfazem as restrições;
- ✓ **Valor Ótimo:** É o valor da função objetivo no ponto ótimo.

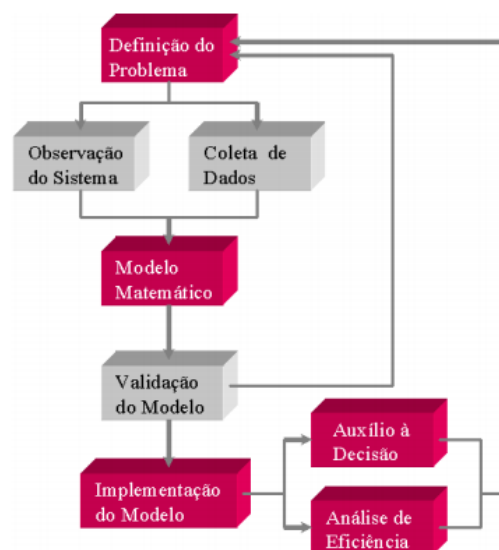
4.1 O Processo de Modelagem

O processo de modelagem envolve a observação do sistema real, a definição do problema a ser modelado (deve-se fazer a coleta de dados de maneira precisa), a definição do objetivo a ser otimizado e como as restrições interagem com este objetivo (MORETTI 2010).

De acordo com Holtz (2005) uma vez definida a função objetivo e suas restrições, o modelador escreve o modelo matemático que deverá ser validado através da aplicação em cenários reais. Após o modelo ser validado, faz-se a implementação final do mesmo, que poderá ser usado com uma ferramenta de tomada de decisões ou como uma ferramenta de análise de eficiência do sistema atual.

A modelagem é um processo contínuo e todo modelo deve ser elaborado com perspectivas de expansão. Neste caso, o modelador deve escolher ferramentas (linguagens algébricas e softwares) que possibilitem e facilitem alterações nas restrições e objetivos do modelo (MORETTI 2010). A figura 33 a seguir ilustra as várias etapas do processo de modelagem.

Figura 33 - Processo de modelagem



Fonte: Moretti (2010)

Segundo Cembranel (2019), as métricas de distâncias são: distância Euclidiana, distância de Minkowski, distância de Manhattan, distância Euclidiana quadrática, distância de Chebyshev, distância de Canberra, distância de Mahalanobis.

Distância Euclidiana: Métrica sugerida para formulação matemática, essa é provavelmente a mais conhecida e usada medida de distância. Ela simplesmente é a distância geométrica no espaço multidimensional. Ela é calculada como:

$$d_{Eucl(i,j)} = \left(\sum_{k=1}^n (x_{ik} - x_{jk})^2 \right)^{1/2} \quad (2)$$

Onde x_{ik} e x_{jk} representam os valores do k-ésimo atributo das séries temporais i e j, a distância Euclidiana é calculada pela raiz quadrada do somatório dos quadrados das diferenças entre os valores x_{ik} e x_{jk} (CEMBRANEL, 2015).

Ainda nesta sessão serão abordados de uma maneira mais detalhada, a tarefa de agrupamento citada brevemente na sessão 3.6.1, devido sua utilização como base para a formulação matemática e durante o processo de mineração de dados.

4.1.1 Agrupamento (Clustering)

Segundo Cembranel (2019), a prática de classificar objetos de acordo com as similaridades é base para muitas áreas da ciência e organizar dados em grupos é um dos métodos fundamentais para entender e aprender.

Análise de agrupamento ou Clustering é a formalização de algoritmos e métodos matemáticos para realizar agrupamento, ou classificar objetos seguindo uma metodologia, onde os grupos são formados de maneira a maximizar a similaridade entre os elementos de um grupo (similaridade intra-grupo) e minimizar a similaridade entre elementos de grupos diferentes (similaridade inter-grupos), (PERES E LIMA 2015).

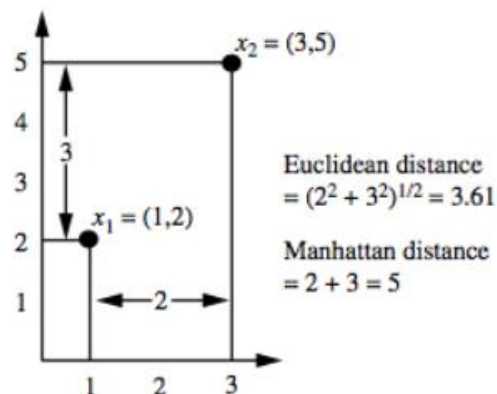
Os objetos também são denominados exemplos, tuplas e/ou registros. Cada objeto representa uma entrada de dados que pode ser constituída por um vetor de atributos que são campos numéricos ou categóricos (categórico é um tipo de campo que pode assumir um entre um conjunto de valores pré-definidos) (LINDEN 2009).

Conforme apresentado por Everitt (1974), um agrupamento pode ser definido de três maneiras:

- Um agrupamento pode ser definido como um conjunto de dados ou partições que são mais similares entre si do que com qualquer outro dado de outra partição;
- Pode ser definido como um conjunto de dados de forma que a distância entre quaisquer dois pontos no agrupamento é menor que a distância entre qualquer ponto no agrupamento e qualquer ponto que não esteja nele;
- Um agrupamento também pode ser descrito como uma região em um espaço multidimensional que contém uma densidade relativamente alta de pontos, separada por regiões que contém uma densidade relativamente baixa de pontos.

No presente trabalho o agrupamento sugerido para formulação matemática é de baseada nos parâmetros de medidas de similaridade entre vetores (dados em seu processo de execução). Essas medidas servem para guiar o processo de construção da superfície de decisão que determinará qual é a região de abrangência de cada grupo de dados (Figura34), (PERES E LIMA 2015).

Figura 34 - Pontos de agrupamentos



Fonte: Peres e Lima (2015)

4.1.2 Medidas de distância

Na tentativa de identificar agrupamentos que podem estar presentes nos dados é necessário ter o conhecimento de quão próximos os elementos estão uns com os outros, ou quão distantes estão.

Muitos métodos de Clustering tem como ponto de partida uma matriz $n \times n$, cujos elementos refletem, em certo sentido, uma medida quantitativa de proximidade. Dois indivíduos são “próximos” quando sua diferença ou distância é pequena, ou sua semelhança é grande (STAHL 2011).

As medidas de distância devem satisfazer três propriedades fundamentais, sendo elas: propriedade da não-negatividade, identidade de indiscernível e desigualdade triangular. Definindo um espaço n dimensional contendo os objetos i, j e k , em que d é a distância entre dois objetos, então as três propriedades segundo Pei (2011) são:

- ✓ **Não-negatividade:** $d(i, j) \geq 0$: a distância entre dois objetos no espaço deve ser um número não negativo.
- ✓ **Identidade de indiscerníveis:** $d(i, i) = 0$: a distância entre um objeto e ele mesmo deve ser igual a zero.
- ✓ **Desigualdade triangular:** $d(i, j) \leq d(i, k) + d(k, j)$: a distância de ir diretamente de um objeto i a um objeto j no espaço não é superior do que a distância entre esses dois objetos com um desvio sobre um outro objeto k , pois a menor distância entre dois pontos é uma reta.

4.2 Diferencial Evolution

O Differential Evolução (DE) é na atualidade uma das ferramentas mais poderosas e versáteis para otimização evolutiva, com o intuito de encontrar o ótimo global de funções não diferenciáveis não- linear, não convexo, multi-modal e definido no espaço de parâmetro contínuo (Storn e Price,1995) proporão a Evolução Diferencial (DE). Com isso diversas variantes de DE surgiram como uma das mais competitivas e versátil família de algoritmos de computação evolutiva e foram com sucesso aplicado para resolver vários problemas do mundo real de diversos

domínios da ciência e tecnologia (SUGANTHAN,2011 apud NERI E TIRRONEN,2010).

O DE básica destaca-se a ser um algoritmo muito simples, cuja implementação requer apenas algumas linhas de código em qualquer linguagem de programação padrão. Além disso, o DE canônica requer muito poucos parâmetros de controle (3 para ser mais preciso: o fator de escala, a taxa de crossover e o tamanho da população) - uma característica que o torna fácil de usar para os praticantes. No entanto, DE exibe desempenho notável, otimizando simultaneamente uma grande variedade de funções objetivo em termos de precisão final, velocidade computacional, e robustez.

4.2.1 O Canônica DE Algoritmo

Um Algoritmo DE padrão possui quatro passos básicos – inicialização, mutação, recombinação de cruzamento e seleção, no entanto somente os três últimos passos são repetidos. As iterações são repetidas até que o critério de parada seja satisfeito, ou se esgote a avaliação funcional máximo.

4.2.1.1 – Inicialização dos Vetores Decisão Variável

O DE começa com uma população escolhida aleatoriamente de N_p d - vetores dimensionais de decisão de valor real. Os vetores também são conhecidos como genoma / cromossomo e formam uma possível solução para o problema de otimização multidimensional. De acordo com as mudanças sofridas ao decorrer das diferentes interações, pode-se adotar a seguinte notação para representar o vetor da população atual.

$$Xi^{(t)} = (X_{i,1}^{(t)}, X_{i,2}^{(t)} X_{i,d}^{(t)}) \quad (4)$$

Em cada variável de decisão, pode haver um determinado intervalo dentro do qual o valor da variável de decisão tem que ser limitado, pois as variáveis de decisão estão relacionadas com componentes físicos ou medidas que tem limite natural.

4.3.1.2 – Mutação com Vetores de Diferença

Logo após sua inicialização o algoritmo cria um doador / mutante. São cinco as estratégias de mutação mais utilizadas, as mesmas estão listadas abaixo:

$$\text{"DE/rand/1"} : Vi^{(t)} = : X_{R1}^{(t)} + F (X_{R2}^{(t)} - X_{R3}^{i(t)}). \quad (4 \text{ a})$$

$$\text{"DE/best/1"} : Vi^{(t)} = : X_{best}^{(t)} + F (X_{R1}^{i(t)} - X_{R2}^{i(t)}). \quad (4 \text{ b})$$

$$\text{"DE/current-to-best/1"} : Vi^{(t)} = : X_i^{(t)} + F (X_{best}^{(t)} - X_i^{(t)}) + F (X_{R1}^{i(t)} - X_{R2}^{i(t)}). \quad (4 \text{ c})$$

$$\text{"DE/best/2"} : Vi^{(t)} = : X_{best}^{(t)} + F (X_{R1}^{i(t)} - X_{R2}^{i(t)}) + F (X_{R3}^{i(t)} - X_{R4}^{i(t)}). \quad (4 \text{ d})$$

$$\text{"DE/rand/2"} : Vi^{(t)} = : X_{R1}^{i(t)} + F (X_{R2}^{i(t)} - X_{R3}^{i(t)}) + F (X_{R4}^{i(t)} - X_{R5}^{i(t)}). \quad (4 \text{ e})$$

4.3.1.3 – Cruzamento

E por meio do cruzamento que o vetor doador mistura seus componentes com o vetor alvo para formar o vetor de ensaio ou seja os descendentes. Segundo Das e Suganthan (2011), a família de algoritmos DE, utiliza dois métodos de cruzamento sendo eles o exponencial (ou modulo dois-pontos) e o binomial (ou uniforme).

No modelo binomial, o cruzamento é realizado em todas as variáveis sempre que um número gerado aleatoriamente entre 0 e 1 é menor ou igual do que o valor definido. Nesse caso, o número de parâmetros herdado a partir do doador de uma quase distribuição binomial. O esquema pode ser expresso como:

$$Ui,j^{(t)} = \{Vi,j^{(t)} \text{ if } j = k \text{ or } rand_{i,j}[0,1] \leq Cr, \} \quad (4f)$$

4.3.1.4 – Seleção

A seleção determina se o alvo (progenitor) ou o processo (prole) sobrevive para a próxima iteração. Ou seja, $t=t+1$. A operação de seleção é descrito como:

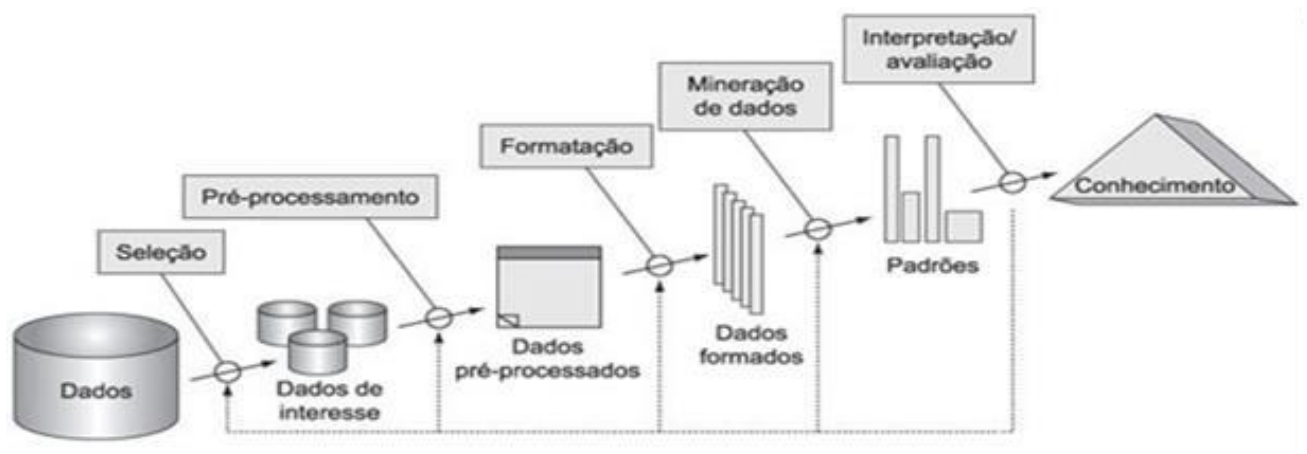
$$Xi^{(t+1)} = \begin{cases} u^{(t)} & \text{if } f(u^{(t)}) \leq f(Xi^{(t)}) \\ Xi^{(t)} & \end{cases} \quad (4g)$$

5 ESTUDO DE CASO

Nesta sessão apresenta-se o estudo de caso, serão aplicados alguns conceitos e técnicas apresentadas nas seções anteriores, com objetivo de identificar as variáveis que relacionam os dados originados de processos de cromatografia à falhas um de transformadores.

Foi considerado como referência trabalho já elaborado em que, focamos na alteração de valores e atributos, para posteriormente avaliarmos as reais chances de evolução do mesmo.

Figura 35 - Etapas do Processo de Descoberta de Conhecimento em Bases de Dados.



Fonte: Fayyad (1996)

5.1 DADOS

Para aplicação dos conceitos e técnicas estudadas, a base de dados utilizada no estudo de caso para a aplicação das técnicas de classificação de falhas utilizadas neste trabalho, foi publicada na dissertação de mestrado de (MORAIS 2004).

Nesta base de dados constam um conjunto de dados da análise de gases de transformadores com e sem contato com comutador sob carga (OLTC) e transformadores de distribuição de vários níveis de tensão, desde 13.8kV a 230kV, alocados na cadeia de distribuição de energia elétrica do Grupo Eletrobrás Cepel¹.

Nestes dados constam informações quanto à potência nominal do transformador, tempo de operação, tipo de cobertura (ar/conservador), tensões de trabalho e tipo de comutador (ZIRBES, 2003). A base é descrita no item abaixo:

5.2 SELEÇÃO DOS DADOS

Como citado na sessão anterior, os dados aqui tratados são de transformadores com e sem contato com comutador sob carga (OLTC) e transformadores de distribuição da análise de gases dissolvidos para transformadores de vários níveis de tensão, desde 13.8kV a 230kV. A intenção do uso deste conjunto neste formato é para estabelecer uma estimativa qualitativa da generalização possível entre grupos de características heterogêneas (MORAIS, 2004 apud ZIRBES, 2003).

Este grupo totaliza 224 amostras (valores de concentração dos gases em ppm) e para a apresentação dos mesmos como mostra a tabela 12, eles foram divididos em função da natureza das condições estabelecidas, para que contemplassem as saídas consideradas pelo modelo matemático em desenvolvimento (MORAIS 2004).

¹ O Centro de Pesquisas de Energia Elétrica (Eletrobras Cepel) foi criado em 1974 e sua sede é localizada no Rio de Janeiro (Ilha do Fundão). Faz parte do grupo Eletrobras e tem mais de 40 anos em pesquisa e desenvolvimento (P&D) relacionado a geração, transmissão e distribuição de energia elétrica e sendo considerado o maior centro de pesquisas em energia elétrica do Hemisfério Sul (Wikipédia, 2019).

Tabela 11 - Banco de Dados do Problema

BANCO DE DADOS DO PROBLEMA		
SAÍDAS CLUSTERS	DIAGNÓSTICOS	TOTAL DE AMOSTRAS
1	TRANSFORMADOR COM DIAGNÓSTICO DE NORMALIDADE	83
2	TRANSFORMADOR COM DIAGNÓSTICO DE FALHA ELÉTRICA	61
3	TRANSFORMADOR COM DIAGNÓSTICO DE FALHA TÉRMICA	80

Fonte: Moraes (2004).

5.3 PRÉ-PROCESSAMENTO

Descrita com detalhe na sessão 3.7.2, a propriedade na tarefa de pré-processamento tratada no cenário proposto utilizará a limpeza dos dados, afim de evitar que um volume grande de dados possa dificultar ou até mesmo tornar impraticável o processo de Descoberta de Conhecimento.

Vale salientar que está etapa descreve a metodologia proposta por MORAIS (2004) para etapas de pré-processamento dos dados.

5.3.1 Limpeza dos dados

Nesta etapa foi realizada a limpeza dos dados, visando principalmente evitar inconsistências e garantir a correta aplicação da Tarefa de Agrupamento citada item 3.7.1. A etapa de mineração de dados.

Para o tratamento dos dados de análise dos gases dissolvidos MORAIS (2004) destaca que para atender os objetivos específicos e tornar o modelo eficiente e eficaz as informações contidas via arquivo, dos dados históricos da análise dos gases dissolvidos no óleo, que necessariamente devem conter as informações de geração dos seguintes gases em ppm: hidrogênio (H₂), metano (CH₄), acetileno (C₂H₂), etileno (C₂H₄) e etano (C₂H₆), além do intervalo em dias entre as amostras.

No primeiro instante foram eliminados os valores inconsistentes para análise, devido a concentração de gases a eles associados ultrapassarem os valores normais de degradação da isolação estabelecidos em ensaios de laboratórios.

Posteriormente, os valores de concentração não detectáveis foram substituídos pelo valor de 0.4 ppm para os cálculos das relações (MORAIS, 2004).

5.3.2 Enriquecimento e transformação dos dados

Em detalhe na sessão 3.7.2 a etapa de transformação dos tipos de dados foi o principal fator de decisão para tornar o processo de mineração mais eficiente e facilitar a compreensão dos dados.

Devido à característica dos dados que serão estudados nesse trabalho (dados numéricos), a transformação de dados mais adequada na análise do trabalho em questão é a normalização que visa dar aos dados o mesmo peso e a construção de atributos adicionados para auxiliar o processo de mineração.

5.3.3 Normalização dos valores

Neste trabalho são consideradas as três classes definidas anteriormente. O algoritmo deve definir quais são os limites das taxas de gases para cada classe como mostra a tabela. Essas variáveis se referem à concentração dos gases, que são dados normalizados pela regra de normalização *min-max* que detalhada na sessão 3.7.2 tem como objetivo gerar valores correspondentes aos originais dentro de um intervalo determinado, colocando todos os valores existentes em uma mesma escala Avancini (2018). A normalização *min-max* é realizada através da equação 5:

$$z_i' = \frac{z_i - \min_A}{\max_A - \min_A} \quad (5)$$

Em que:

z_i' - Representa o dado normalizado;

z_i - Representa o dado a ser normalizado;

\min_A - Representa o menor valor encontrado no intervalo de dados a serem normalizados;

\max_A - Representa o maior valor encontrado no intervalo de dados a serem normalizados.

5.3.4 Definição dos atributos

Depois da consolidação de todos os dados na forma apropriada para a utilização do conjunto de técnicas de busca e otimização inspiradas na evolução natural das espécies de um algoritmo de evolução diferencial, a forma de representação dos atributos de entrada e de saída para um problema de aprendizado algoritmo de evolução diferencial é uma das decisões chave que influenciam na qualidade da solução que pode ser obtida.

Detalhado na sessão 3.7.2 os atributos para a formulação de um modelo matemático podem ser discretos, ou contínuos.

Desde modo, nesta etapa a tarefa de agrupamento (Tabela 13) ficou definida a partir da observação dos seguintes gases no óleo dos transformadores de distribuição, como mostra a tabela abaixo e em seguida os atributos do problema em questão:

Tabela 12 - Conjunto de Agrupamentos (Clusters)

CONJUNTO DE AGRUPAMENTOS (CLUSTERS)	
Hidrogênio	Grandes quantidades associadas com descarga parcial
Hidrogênio, Etano, Metano e etileno	Decomposição térmica do óleo, contato do óleo isolante com partes quentes
CLUSTERS 3	Gás associado com arco elétrico no óleo

V_{HNmin}	Concentração mínima de Hidrogênio para que se classifique o transformador como normal
V_{HNmax}	Concentração máxima de Hidrogênio para que se classifique o transformador como normal
V_{ENmin}	Concentração mínima de etano para que se classifique o transformador como normal;
V_{ENmax}	Concentração máxima de etano para que se classifique o transformador como normal;
V_{MNmin}	Concentração mínima de metano para que se classifique o transformador como normal;
V_{MNmax}	Concentração máxima de metano para que se classifique o transformador como normal;

V_{EtNmin}	Concentração mínima de etileno para que se classifique o transformador como normal;
V_{EtNmax}	Concentração máxima de etileno para que se classifique o transformador como normal;
V_{ACNmin}	Concentração mínima de acetileno para que se classifique o transformador como normal;
V_{ACNmax}	Concentração máxima de acetileno para que se classifique o transformador como normal;
V_{HEmin}	Concentração mínima de hidrogênio para que se classifique o transformador com falha elétrica;
V_{HEmax}	Concentração máxima de hidrogênio para que se classifique o transformador com falha elétrica;
V_{ACEmin}	Concentração mínima de acetileno para que se classifique o transformador com falha elétrica;
V_{ACEmax}	Concentração máxima de acetileno para que se classifique o transformador com falha elétrica;
V_{HTmin}	Concentração mínima de hidrogênio para que se classifique o transformador com falha térmica;
V_{HTmax}	Concentração máxima de hidrogênio para que se classifique o transformador com falha térmica;
V_{EtTmin}	Concentração mínima de etileno para que se classifique o transformador com falha térmica;
V_{EtTmax}	Concentração máxima de etileno para que se classifique o transformador com falha térmica;
V_{MTmin}	Concentração mínima de metano para que se classifique o transformador com falha térmica;
V_{MTmax}	Concentração máxima de metano para que se classifique o transformador com falha térmica;
V_{ETmin}	Concentração mínima de etano para que se classifique o transformador com falha térmica;
V_{ETmax}	Concentração máxima de etano para que se classifique o transformador com falha térmica;

Depois de todos os dados preparados, é nesta fase que as técnicas (algoritmos) de mineração serão aplicadas. A escolha da técnica depende do objetivo desejado, como apresentados nas sessões anteriores.

Como citado na sessão 4.2 um problema de otimização está sempre associado a um problema de minimização ou maximização de uma ou mais funções, que geralmente, existe um conjunto de restrições nas variáveis.

5.3.5 Formulação do Modelo Matemático

Para Hasse (2000) a formulação de um problema matemático de mineração de dados para um algoritmo evolucionário deve se iniciar a partir da equação (6):

$$r_1 \wedge r_2 \wedge \dots \wedge r_n \Rightarrow C \quad (6)$$

Onde,

r_i : restrição para o valor atribuído a i , sendo $1 \leq i \leq n$.

C : classe que será dada como resultados se somente se todas as restrições da regra forem atendidas.

Ainda segundo Hasse (2000), a função de aptidão deve ser capaz de classificar as regras como classificadores parciais. Dessa forma, a precisão da regra é mais importante do que sua capacidade de cobrir todos os exemplos positivos encontrados nos dados de treinamento. Isto quer dizer que a precisão da regra (percentual de acerto) deve ser considerada juntamente com algum tipo de compensação pela abrangência da regra, a fim de evitar o uso de regras que cubram poucos exemplos, isto é, representem apenas ruídos nos dados.

Com base na consideração feita por Hasse (2000) para a formulação matemática o presente trabalho baseia-se nos parâmetros de medidas de similaridade entre vetores (dados) em seu processo de execução das métricas da distância euclidiana.

$$d_{Eul(i,j)} = \sqrt{\sum_{i=1}^{nc} \sum_{j=1}^{ng}} (Vpmax_{ij} - max_{ij})^2 + (Vpmin_{ij} - min_{ij})^2 \quad (7)$$

$d_{Eul(i,j)}$ Função a ser minimizada

nc Número de classes

ng Quantidade de gases usados para a classificação

np Número de classes

$Vpmax$ Limites máximos dos atributos de uma classe para cada indivíduo

$Vpmin$ Limites mínimos dos atributos de uma classe para cada indivíduo

Sujeita às seguintes condições:

$$\text{Classe A se: } V_{HNmin} \leq V_{HN} \leq V_{HNmax} \wedge V_{ENmin} \leq V_{EN} \leq V_{ENmax} \wedge V_{MNmin} \leq V_{MN} \leq V_{MNmax} \wedge V_{EtNmin} \leq V_{EtN} \leq V_{EtNmax} \wedge V_{ACNmin} \leq V_{ACN} \leq V_{ACNmax} \quad (7a)$$

$$\text{Classe B se: } V_{HEmin} \leq V_{HE} \leq V_{HEmax} \text{ OU } / \wedge V_{ACEmin} \leq V_{ACE} \leq V_{ACEmax} \quad (7b)$$

$$\text{Classe C se: } V_{HTmin} \leq V_{HT} \leq V_{HTmax} \wedge V_{ETmin} \leq V_{ET} \leq V_{ETmax} \wedge V_{MTmin} \leq V_{MT} \leq V_{MTmax} \wedge V_{EtTmin} \leq V_{EtT} \leq V_{EtTmax} \quad (7c)$$

Essa função deve ser minimizada, para que gere regras mais precisas, mais abrangentes e significativas. A significância da regra é medida comparando-se o número dos exemplos classificados corretamente com aqueles classificados corretamente pelas regras geradas para classificar a mesma classe.

Para fechar o ciclo da mineração de dados é feita a etapa de classificação onde foi utilizada para treinamento da base de dados as técnicas de algoritmos evolucionários.

6 RESULTADOS E DISCUSSÕES

Nesta sessão são apresentados os resultados das simulações obtidos utilizando o conjunto de transformadores com e sem contato com comutador sob carga (OLTC) e transformadores de distribuição da análise de gases dissolvidos, totalizando 224 amostras (valores de concentração dos gases em ppm) apresentadas no estudo de casos para construção do modelo matemático.

Todas as simulações foram realizadas no Matlab utilizando Algoritmo de Otimização Differential Evolution “best 2” para calcular e obter o rendimento do modelo matemático.

O desempenho do modelo matemático foi avaliado através dos parâmetros utilizados nas Métricas derivadas da matriz de confusão.

$$\text{“DE/best/2”} : Vi^{(t)} = : X_{best}^{(t)} + F (X_{R1}^{i(t)} - X_{R2}^{i(t)}) + F (X_{R3}^{i(t)} - X_{R4}^{i(t)}). \quad (8)$$

6.1.1 - Avaliação de Modelos

O método de avaliação mais utilizado é a matriz de confusão de acertos e erros associados à previsão dos modelos. Esta matriz é utilizada para quantificar a qualidade do modelo ajustado e seu formato é apresentado na tabela abaixo:

Tabela 13 - Parâmetros da matriz de confusão

PARAMÊTROS		
PREVISÃO DO MODELO	AMOSTRAS PRESENTE	
AUSENTE		
PRESENTE	VP	FP
AUSENTE	VN	FN

Fonte: Os Autores (2019)

Os valores VP (Verdadeiro Positivo) e VN (Verdadeiro Negativo) são predições corretas. FP (Falso Positivo) e FN (Falso Negativo) são considerados erros de predição. Os erros do tipo FP também são conhecidos como erros de comissão ou superestimativa, enquanto os erros do tipo FN são conhecidos como erros de omissão.

6.1.2 - Sensibilidade

Com o intuito de averiguar a capacidade de determinado método corresponder melhor a restrições, escolheu-se investigar o comportamento pela métrica da sensibilidade, bem como a taxa de falso negativo.

A sensibilidade descreve a probabilidade de um legítimo verdadeiro positivo ser classificado como verdadeiro positivo e é a obtida pela seguinte equação:

$$\textbf{Sensibilidade} = \frac{(VP)}{(VP+FN)} \quad (9)$$

Os erros do tipo FN (omissão) são considerados erros graves. Ou seja, um local onde se sabe que a espécie é encontrada e está sendo predito como ausente.

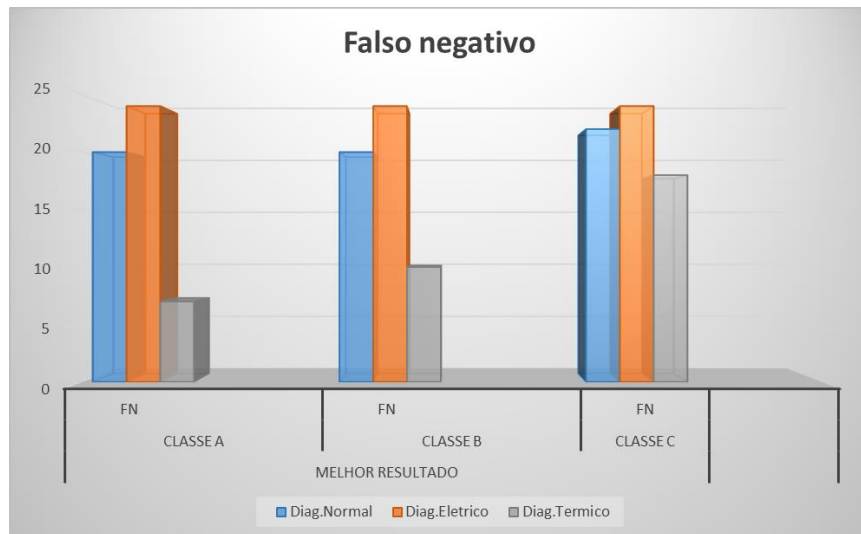
$$\textbf{Taxa de falso negativo (omissão)} = \frac{(FN)}{(VP+FN)} \quad (10)$$

Tabela 14 - Melhor Resultado Falso Negativo

Melhor resultado Falso negativo		
Classe A	Classe B	Classe C
Fn	Fn	Fn
20	20	22
24	24	24
7	10	18

Fonte: Os autores (2019)

Figura 36 - Valores Falso Negativo



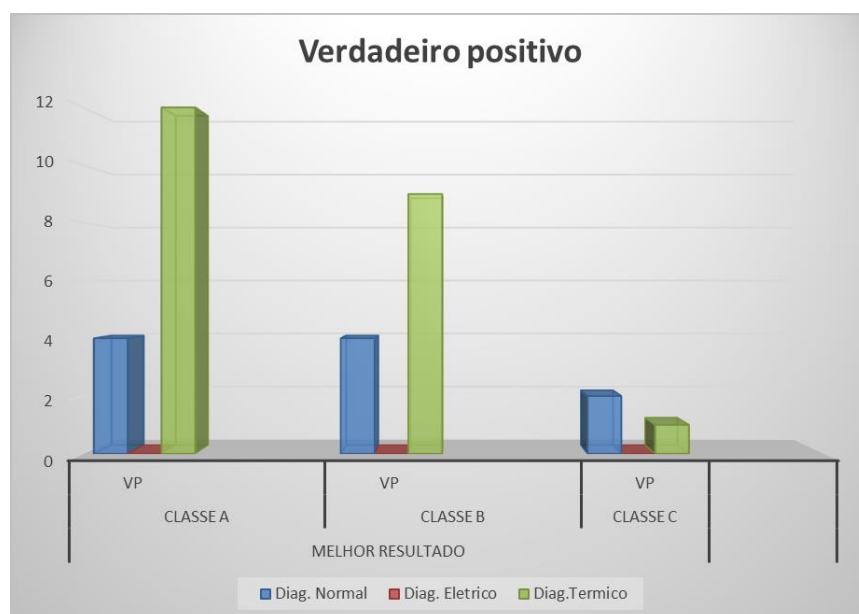
Fonte: Os autores (2019)

Tabela 15 - Melhor Resultado Verdadeiro Positivo

Melhor resultado Verdadeiro positivo		
Classe A	Classe B	Classe C
Vp	Vp	Vp
4	4	2
0	0	0
12	9	1

Fonte: Os autores (2019)

Figura 37 - Valores Verdadeiro Positivo



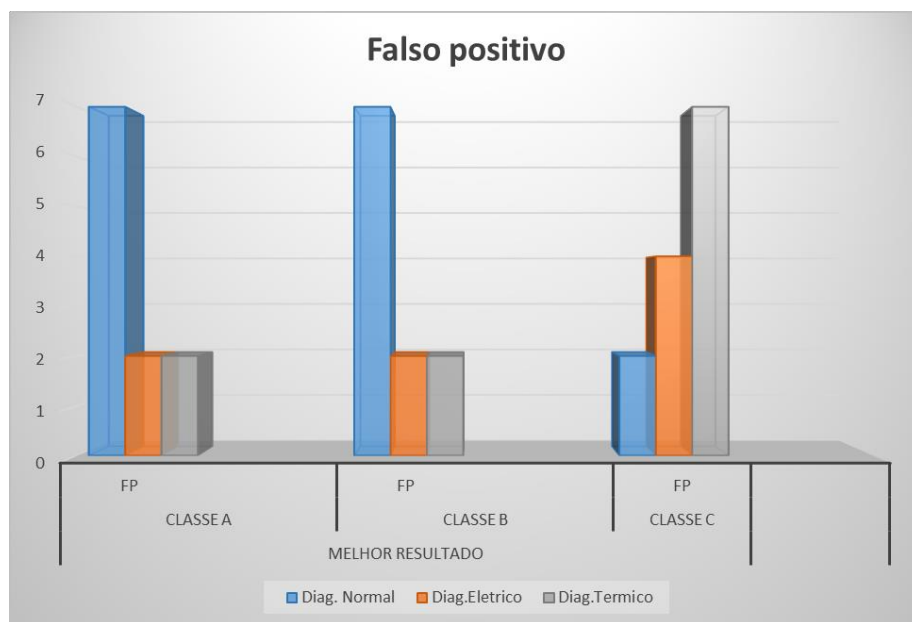
Fonte: Os autores (2019)

Tabela 16 - Melhor Resultado falso Positivo

Melhor resultado Falso Positivo		
Classe A	Classe B	Classe C
Fp	Fp	Fp
7	7	2
2	2	4
2	2	7

Fonte: Os autores (2019)

Figura 38 - Valores Falso Positivo



Fonte: Os autores (2019)

Tabela 17 - Parâmetros definidos para o modelo (%)

Parâmetros definidos para o modelo (%)	
Tamanho da População	20
Número de gerações	400
Taxa de renovação	0,80
Taxa de cruzamento	0,9
Taxa de mutação	0,2
Significância	0,4
Quantidade de gases	5
Número de classes	3
Limite inferior	5,1,3,30
Limite superior	5,1,3,30
Número de execução	30

Fonte: Os autores (2019)

6.1.3 - Resultados para a classe A de Verdadeiro positivos

$$\text{Sensibilidade} = \frac{(VP)}{(VP+FN)} = \frac{(16)}{(16+51)} = 0,2388 \times 100 = 23,88\% \quad (11)$$

$$\text{Taxa de falso negativo (omissão)} = \frac{(FN)}{(VP+FN)} = \frac{(51)}{(16+51)} = 0,7611 \times 100 = 76,1194\% \quad (12)$$

- Número médio de verdadeiros positivos = 3.7667
- Número médio de falsos positivos = 4.7000
- Número médio de falsos negativos = 20.2333
- Valor médio da função de aptidão = 3.0561
- Valor da Função de Aptidão – Melhor resultado definido por 3 execução
- 2.6164
- 2.9225
- 2.8473

:

Tabela 18 - Diferença do valor médio função aptidão e valor de aptidão definido por execução

Melhor índice (1/3) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.6164 - 3.0561 = -0.4397$
Melhor índice (2/2) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.9225 - 3.0561 = -0.1336$
Melhor índice (3/3) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.8473 - 3.0561 = -0.2088$

Fonte: Os autores (2019)

6.1.4 Resultados para a classe B de Verdadeiro positivos

$$\text{Sensibilidade} = \frac{(VP)}{(VP+FN)} = \frac{(13)}{(13+54)} = 0,1940 \times 100 = 19,4029\% \quad (13)$$

$$\text{Taxa de falso negativo (omissão)} = \frac{(FN)}{(VP+FN)} = \frac{(54)}{(13+54)} = 0,8059 \times 100 = 80,59\% \quad (14)$$

- Número médio de verdadeiros positivos = 0
- Número médio de falsos positivos = 6.1333
- Número médio de falsos negativos = 24
- Valor médio da função de aptidão = 2.9143
- Valor da Função de Aptidão - Melhor resultado definido por 3 execução
- 3.2990
- 2.6838
- 2.8378

Tabela 19 - Diferença do valor médio Função aptidão e de aptidão definido por execução

Melhor índice (1/3) = $VA_{\text{resultado}} - VA_{\text{médio}} = 3.2990 - 2.9143 = -0.3847$
Melhor índice (2/2) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.6838 - 2.9143 = -0.2305$
Melhor índice (3/3) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.8378 - 2.9143 = -0.0765$

Fonte: Os autores (2019)

6.1.5 Resultados para a classe C de Verdadeiro positivos

$$\text{Sensibilidade} = \frac{(VP)}{(VP+FN)} = \frac{(3)}{(3+64)} = 0,0447 \times 100 = 4,4776\% \quad (15)$$

$$\text{Taxa de falso negativo (omissão)} = \frac{(FN)}{(VP+FN)} = \frac{(64)}{(3+64)} = 0,9552 \times 95,522\% \quad (16)$$

- Número médio de verdadeiros positivos = 7.7667
- Número médio de falsos positivos = 3.1000
- Número médio de falsos negativos = 11.233
- Valor médio da função de aptidão = 2.8502
- Valor da Função de Aptidão - Melhor resultado definido por 3 execução
3.3121
2.9776
2.8001

Tabela 20 - Diferença do valor médio função aptidão e valor da função de aptidão definido por execução

Melhor índice (1/3) = $VA_{\text{resultado}} - VA_{\text{médio}} = 3.3121 - 2.8502 = -0.4619$
Melhor índice (2/2) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.9776 - 2.8502 = 0.1274$
Melhor índice (3/3) = $VA_{\text{resultado}} - VA_{\text{médio}} = 2.8001 - 2.8502 = -0.0501$

Fonte: Os autores (2019)

Os resultados apresentados demonstram o desempenho do modelo matemático, para o treinamento foi utilizada uma amostra de 30% dos dados e 70% para a validação do modelo.

De acordo com os valores calculados para cada classe pode perceber que a classe A referente ao transformador sem falha elétrica, teve maior número de validação, onde os valores de Verdadeiros positivos foram 23,88%, o valor de valor da Taxa de falso negativo (omissão) 76,11%.

Já a classe B ficando com 19,4029% e com o valor da *Taxa de falso negativo (omissão)* 80,59% ou seja esses valores para classe B se justificam devido os valores limite mínimos serão os mesmo para classe A.

Pode-se considerar que a classe C o algoritmo teve desempenho satisfatório, pois os seus valores de limites mínimos foram diferentes, o que mostra que ao validar, mesmo com valores inferiores as demais classes, a validação foi feita.

Já os valores para função objetivo conclui se que a classe A na execução 2 obteve melhor valor em relação à média da função. Já as classes B e C pode-se considerar a terceira execução.

Em relação aos valores limites, foi observado que a grande maioria ficou dentro dos parâmetros estabelecidos, vale ressaltar que para a classe B não apresentaram um bom desempenho devido os seus parâmetros serem os mesmo da classe A, o que é validar rever os limites estabelecidos.

6.1.6 Sugestões para melhoraria do modelo Matemático

Alterar os valores de taxas de população, pois a mesma interfere diretamente no teste, nessa execução foi feito com valor de 20%, para melhores resultados de validação pode-se aumentar a taxa em aproximadamente 10% e observar se houve melhoria.

Alterar os valores da taxa de cruzamento, aumentando até o limite em relação a quantidade de dados pois existe a probabilidade de que quanto maior for esta taxa, mais rapidamente novas estruturas serão introduzidas na população.

Testar outros valores da taxa de mutação, uma baixa taxa de mutação previne a convergência prematura para um ótimo local, possibilitando ao algoritmo explorar melhor todo o espaço de busca. Uma taxa de mutação muito alta faz com que o processo de busca se torne essencialmente aleatório.

6.1.7 Em relação ao objetivo geral

Em relação ao objetivo geral pode-se concluir que a proposta do mesmo era a é a formulação matemática de um problema de mineração de dados que deve ser resolvido através de um algoritmo evolucionário.

Conforme a literatura pesquisada a validação de um modelo pode ser descrita como o processo de aceitação ou rejeição do mesmo. Esta análise está condicionada a vários fatores, prevalecendo o confronto dos dados reais com os resultados obtidos a partir da simulação do modelo (DESTEFANI, 2018). Como um modelo matemático dificilmente descreve todas as características de um sistema real, ele pode ser considerado válido quando representar as características fundamentais do sistema que será modelado de forma satisfatória. A formulação do problema neste trabalho considerou as informações constantes em Castanheiras (2008) e que o problema deverá resolvido a partir de algoritmo evolucionário.

6.1.8 Em relação aos objetivos específicos

As referências encontradas e apresentadas com ênfase nas variáveis que permitem relacionar os dados originados de processos de cromatografia à falhas, demonstraram que a aplicação de técnicas e métodos para solução de problemas como o apresentado neste trabalho podem garantir o diagnostico eficaz.

Em relação aos valores apresentados por Castanheira (2008), com técnicas diferentes das utilizadas neste trabalho, pode se concluir que ambos métodos se ajustados corretamente, podem auxiliar na atividade de tomada de decisões na área de engenharia elétrica.

Vale ressaltar que segundo Bala *et al.*(1995), a seleção de um conjunto de recursos apropriados para a classificação é uma das tarefas mais difíceis de um sistema de classificação de padrões.

7 TRABALHOS FUTUROS

Com o conhecimento adquirido durante a execução deste trabalho foi possível perceber que alguns pontos referentes ao tema podem ser aprimorados, ficando como sugestão para trabalhos futuros.

Conduzir estudos para instalar medidores de gases *on-line*, para auxiliar no diagnóstico preditivo de forma mais eficaz. A instalação desse sistema também otimizará o algoritmo evolucionário, que conseguirá otimizar seu aprendizado e assertividade, pois o banco de dados seria abastecido constantemente.

8 CONSIDERAÇÕES FINAIS

Com o conhecimento adquirido durante a execução deste trabalho foi possível perceber que alguns pontos referentes ao tema podem ser aprimorados, ficando como sugestão para trabalhos futuros.

Conduzir estudos para instalar medidores de gases *on-line*, para auxiliar no diagnóstico preditivo de forma mais eficaz. A instalação desse sistema também otimizará o algoritmo evolucionário, que conseguirá otimizar seu aprendizado e assertividade, pois o banco de dados seria abastecido constantemente.

Devido a importância dos transformadores para os sistemas elétricos, indústria e outros segmentos da sociedade, um acompanhamento este ativo se faz muito importante, uma vez que dependendo do tamanho e importância deste transformador uma falha nele pode trazer grandes prejuízos.

Algumas variáveis são importantes para que se possa fazer um acompanhamento dos transformadores. Estas variáveis estão diretamente relacionadas com o funcionamento dos transformadores. São elas temperatura, carregamento, teor de gases dissolvidos, teor de furfuraldeídos, entre outras.

A execução deste trabalho possibilitou conhecer um pouco mais a respeito dos transformadores, mineração de dados, algoritmos evolucionários, inteligência artificial dentre outras tecnologias que estão surgindo com a inserção da Indústria 4.0. Foi possível aprofundar em assuntos relacionados ao aproveitamento dos dados originados da análise cromatográfica dos gases, que foram transformados em informações importantes que podem identificar e conseqüentemente antecipar a tomada de decisão para minimização de impactos.

REFERÊNCIAS

ABB, **Transformadores de Potência e Reatores de Derivação**, catálogo, 1999. P.7.

ABB, **Sistema de Monitoração Geração de Transformadores**. Disponível em: <http://www.eyeproject.cl/files/42dc39_ABB-BR-34-ABBTMonitor-SistMonitoracao.pdf>. Acesso em: 29 jul. 2019.

ABB, Manual Técnico. **Transformadores de Distribuição**. Santa Catarina, 2010.

ABRADEE. **Visão Geral do Setor**, 2015. Disponível em: <http://www.abradee.com.br/setoreletrico/redes-de-energia-eletrica>, acesso em 01 dez. 2019.

ABREU, J.P.G.; COGO, J.R.; OLIVEIRA, J. C. **Transformadores: teoria e ensaios**. São Paulo: Edgard Blücher, 1984.

AGRAWAL, R; IMIELINSKI, T; SWAMI, A. **Mining association rules between sets of items in large databases**. Proc. Of the ACM SIGMOD, p. 207-216,1993.

AGRAWAL, R; SRIKANT, R. **Fast algorithms for mining association rules**. 20th International Conference on Very Large Data Bases, p. 487-499,1994.

ALVES, M. E. G.; MOURA, G. A.; VASCONCELOS, V.. **Especificação de Sistemas de Monitoramento On-line para Trans Potência Baseada em Arquitetura Descentralizada**. CLADE, 2008.

ALVES, Renato Marinho. **Plataforma de mineração e apresentação dos dados do ENADE da área de Computação dos anos de 2008 a 2014**. Centro Universitário Luterano de Palmas, 2018.

ARANTES, Jose Geraldo. **Diagnóstico de Falhas em Transformadores de Potência pela Análise de gases Dissolvidos em óleo isolante através de redes neurais**. 2005. Dissertação (Mestrado) - Universidade Federal de Itajubá, Itajubá, 2005.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5356: Transformador de potência: especificação**. Rio de Janeiro, 1993.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14725-4.**
Confiabilidade e manutenibilidade. Rio de Janeiro, 1993.

AVELAR, C. F. P. D.; ROCHA, A. H.; CRUZ, F. J. S. **Mineração de Dados: uma revisão da literatura em Administração.** *Vianna Sapiens*, Juiz de Fora, v.8. n.2, julho/dezembro 2017.

BARTLEY, P. **Analysis of transformer failures.** [S.1:S. n.]. p. 1-13.

BECHARA, Ricardo. **Análise de falhas em transformadores de potência.**
Dissertação de Mestrado – Escola Politécnica da Universidade de São Paulo.

BLATT, JAIR JOSÉ FRITZEN. **PROPOSTA DE INSTALAÇÃO DE REGULADOR DE TENSÃO EM UMA REDE DE DISTRIBUIÇÃO PRIMÁRIA A PARTIR DA ANÁLISE DO PERFIL DE TENSÃO.** Universidade regional do noroeste do Estado do Rio Grande do Sul. Santa Rosa, Rio Grande do Sul, 2016.

BORGELT, C. **An implementation of the FP-growth algorithm,** 2005.

BOSE, Bernhard E.; GUYON, Isabelle M.; VAPNIK, Vladimir N. **A Training Algorithm for Optimal Margin Classifiers.** In: ANNUAL WORKSHOP ON COMPUTATIONAL LEARNING, 5, 1992, Pittsburgh. ACM Press. Pittsburgh: Haussler D, jul 1992. p.144-152

CAMILO, Cássio Oliveira; SILVA, João Carlos. **Mineração de Dados: Conceitos, tarefas, métodos e ferramentas.** 2009, [s/l]. Disponível em:
<[http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios Tecnicos/RTINF_001-09.pdf](http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios_Tecnicos/RTINF_001-09.pdf)> . Acesso em: 12 jun.2019

CAPARÓ, José Luís Choque. **Modelagem de Transformadores de Distribuição para Aplicação em Algoritmos de Fluxo de Potência Trifásico.** Faculdade de Engenharia de Ilha Solteira. Ilha Solteira, São Paulo, 2005.

CARVALHO, Ricardo Silva. **Modelos preditivos para avaliação de risco de corrupção de servidores públicos federais.**2016. 119 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Universidade de Brasília, Brasília, 2016.

CASANOVA, A. A; LABIDI, S. A. **Identificação de competências em currículos usando ontologias: uma abordagem teórica**, 2006.

CASTANHEIRA, Luciana Gomes. Aplicação da Mineração de dados a análise das condições de operação de transformadores. **Aplicação da Mineração de dados a análise das condições de operação de transformadores**, Revista Eletrônica Produção e Engenharia, 12 jan. 2009.

CEMBRANEL, Samuel Sandmann. **Utilização de Técnicas de mineração de dados na caracterização perfis típicos de carga**. Mestrado em Engenharia Eletrotécnica. Instituto Trabalho de Conclusão de Curso (TCC) - Instituto Federal de Santa de Santa Catarina, de Santa Catarina, Florianópolis, 2019.

CHAPMAN, Pet,e et al. CRISP-DM 1.0: Step-by-step data mining guide. [s/l]: Spss, 2000. 76 p. Disponível em: <<https://www.the-modeling-agency.com/crisp-dm.pdf>>. Acesso em: 12 jun.2019

COLLINS, C. H. e outros **Introdução a Métodos Cromatográficos**. 5ª edição. Unicamp, Campinas, 1993.

Congresso Brasileiro de Química. 53, 2013, Rio de Janeiro. **Determinação de gases dissolvidos em óleo vegetal utilizado como fluido isolante em transformador de distribuição**. ISBN: 978-85-85905-06-4

CORRAL, Maria Filomena Brito. **Gestão da Manutenção de Transformadores de Distribuição visando à Qualidade no Fornecimento de Energia Elétrica**. Universidade Federal do Pará. Belém, Pará, 2008.

DESTEFANI, Luciano. Desenvolvimento de um Modelo Matemático utilizando a Descoberta de Conhecimento Aplicado a Transformadores Trifásicos. **Desenvolvimento de um Modelo Matemático utilizando a Descoberta de Conhecimento Aplicado a Transformadores Trifásicos**, Ijuí- RS, Julho 2018.

ENGETERMO. **Soluções em Termografia**, 2009.

EPE. **Anuário Estatístico de Energia Elétrica 2016**. Empresa de Pesquisa Energética, p. 232, 2016.

FAUSETT, L.; 1994. **"Fundamentals of Neural Networks"**. New Jersey: Prentice Hall Inc.FAUSETT, L.; 1994.

FAYYAD, U; SHAPIRO, G.P; SMYTH, P. UTHURUSAMY, R. **From Data Mining to Knowledge, Discovery in Databases**. *American Association for Artificial Intelligence*, 1996.

Felipe Luz Aboboreira e Antônia Ferreira dos Santos Cruz² “A IMPORTÂNCIA DO SMART GRID NA REDE ELÉTRICA DE DISTRIBUIÇÃO DO BRASIL. Disponível em: < <https://revistas.unifacs.br/index.php/sepa/article/viewFile/4375/3023>> Acesso em: 12 jun.2019.

FERREIRA, Paulo Roberto Nunes. **Estudo de viabilidade técnica na utilização de óleo vegetal como agente isolante e refrigerante em transformadores de distribuição**. Dissertação de Mestrado em Agroenergia. Universidade Federal do Tocantins. Palmas, 2018.

FILHO, Geraldo Lupi,. **Comparação entre os critérios de diagnósticos por análise cromatográfica de gases dissolvidos em óleo isolante de transformador de potencia**. Dissertação de Mestrado em Ciências. Escola de Engenharia de São Carlos da Universidade de São Paulo, 2012.

FITZGERALD, A. E.; KINGSLEY, Charles; UMANS, Stephen D. **Máquinas elétricas: com introdução à eletrônica de potência**. 6. ed. Porto Alegre: Bookman, 2006.

FONSECA, Tales Lima. **Integração de ferramentas computacionais para a solução de problemas de otimização em engenharia**. Universidade Federal de Juiz de Fora. Juiz de Fora, 2016.

GARCIA, Simone Carboni. **O uso de árvores de decisão na Descoberta de Conhecimento na Área da Saúde**. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul. Porto Alegre, Rio Grande do Sul, 2003.

GE GRID SOLUTIONS, 2018 “GE GRID SOLUTIONS. Disponível em: < https://www.gegridsolutions.com/press/News_Events.htmf> Acesso em: 12 jun.2019.

GÓMEZ, N.A.G. **Diagnóstico de falhas em transformadores isolados com óleo vegetalisolante a partir da análise de Gases dissolvidos**.2013. 124p. Dissertação (Engenhariae Ciência dos Materiais). Universidade Federal do Paraná.Curitiba.2018.

HAYKIN, S. **Redes Neurais- Princípios e Práticas**. BOOKMAN, São Paulo, 2ª ed. 2001. 900 p.

HAN, J; KAMBER, M. **Data Mining: Concepts and Techniques**. Elsevier, 2006.

HARLOW, JAMES H. **Electric power transformer engineering**. CRC Press, Florida, 2004.

HASSE, M., **Mineração de Dados Usando Algoritmos Genéticos**, Dissertação de Mestrado Em Informática, Universidade Federal do Paraná, Agosto de 2000.

HIRAGI, Gilberto de Oliveira. **Mineração de dados em base de germoplasma**. 2008. 107 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Universidade de Brasília, Brasília, 2008.

IBM. **Visão geral da ajuda do CRISP-DM**. *IBM Knowledge Center*. 2017. Disponível em: <https://www.ibm.com/support/knowledgecenter/pt-br/SS3RA7_17.1.0/modeler_crispdm_ddita/clementine/crisp_help/crisp_overview.html>. Acesso em: 12 jun.2019.

IEEE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE C57-104**: Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers. New York: Ieee, 2008. 28 p.

IEEE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. . **IEEE C57.91**: Guide for Loading MineralOil-Immersed Transformers and Step-Voltage Regulators. New York: Ieee, 2011. 15p.

JÚNIOR, José Arnóbio. **Modelo para Predição de Indicadores de Continuidade em um Sistema de Distribuição de Energia Elétrica, uma Aplicação à Gestão de Manutenção com a Perspectiva do uso da Termografia**. 2016. Dissertação (Pós Graduação em Engenharia Industrial) - Universidade Federal da Bahia, Salvador, 2016.

KOBAYASHI *et al.* **Latest Diagnostic Methods of Gasin-oil Analysis for Oil-filled transformer in Japan**. Proceedings of 13th International Conference on Dielectric Liquids, Nara, Japan, 1999.

KVIATKOWSKI, Márcio Augusto; GOZZI, Sérgio. **Impactos da automação de subestações de energia em indicadores técnicos de qualidade da ANEEL – uma abordagem crítica**. XII SIMPEP – Bauru, SP, 2005.

LANÇAS, F.M. **Cromatografia em Fase Gasosa**. Suprema, São Carlos, 1993.

LAROSE, D. T. **Discovering Knowledge in Data: An Introduction to Data Mining**. John Wiley and Sons, Inc, 2005.

LIU, B; HSU, W; MA, Y. **Integrating classification and association rule mining**. AAAI Workshop of Knowledge Discovery in Databases, 1998.

MARQUES, A.P. **Diagnóstico otimizado de transformadores de potência mediante a integração de técnicas preditivas**. 2018.195p.Dissertação. (Doutorado Engenharia Elétrica e de Computação). Universidade Federal de Goiás. Goiânia. 2018.

MCCUE, C. **Data Mining and Predictive Analysis - Intelligence Gathering and Crime Analysis**. Elsevier, 2007.

MELO, A.S. **Identificação de Pontos Quentes em Transformadores de Potência por Meio de Técnicas Não Invasivas**. 2017. 70p.(Mestrado em Ciências). Escola Politécnica da Universidade de São Paulo, São Paulo. 2017.

MILASCH M., **Manutenção de Transformadores em Líquido Isolante**, 5 edição, Florianópolis-SC. Edgard Blucher, 1998. 354p.

MORAIS, Diego Roberto. **Ferramenta inteligente para detecção de falhas incipientes em transformadores baseada na análise de gases dissolvidos no óleo isolante**. Dissertação de Mestrado em Engenharia Elétrica. Universidade Federal de Santa Catarina. Florianópolis, 2004.

Moretti (2010) "Modelagem matemática Disponível em: <
http://www.fsma.edu.br/si/edicao4/FSMA_SI_2009_2_Tutorial.pdf> Acesso em: 12 jun.2019.

MORO, Sérgio; LAUREANO, Raul; CORTEZ, Paulo. **Using data mining for bank direct marketing: an application of the CRISP-DM methodology**. In: EUROPEAN SIMULATION AND MODELLING CONFERENCE, 25., 2011, Guimarães. Proceedings of European Simulation and Modelling Conference - ESM'2011. Guimarães: Esm, 2011. p. 117 - 121. Disponível em:<
<https://repositorium.sdum.uminho.pt/handle/1822/14838>> . Acesso em: 12 jun. 2019.

MCSHANE, C. P. - **Vegetable-oil-based dielectric coolants**. IEEE Applications Magazine, 2002.

NACARATO, Ricardo Luiz. **Investigação de falha em transformadores de potência pela análise dos gases dissolvidos no óleo isolante utilizando simulação em redes neurais artificiais**. 2018. 83p. Dissertação (Mestrado) -- Universidade Santa Cecília, Programa de Pós-Graduação em Engenharia Mecânica, Santos, SP, 2018.

NBR 5362, (2017). **“Transformador de Potência”**. ABNT, 2017.

NBR 7274, (1982). **“Guia de interpretação da análise dos gases livres e dissolvidos”**. ABNT, 1982.

NEPOMUCENO, L.X. **Técnicas de Manutenção Preditiva, V;2** São Paulo, Edgard Blucher, 1999.

NUNES FERREIRA, P.R, **Estudo de Viabilidade Técnica na utilização de óleo vegetal como agente isolante e refrigerante em substituição ao óleo mineral em transformadores de distribuição**. 2018. 75 f. Dissertação (Mestrado em Agroenergia) - Programa de Pós-Graduação em Agroenergia, Universidade Federal do Tocantins, Palmas, 2018.

OLIVEIRA, Felipe Aballo de Figueiredo Relvas de; RIBEIRO, Henrique Lemos **Análise De Risco De Projetos De Exploração e Produção De Óleo e Gás: Aplicação De Árvore De Decisão Com Teorema De Bayes**. Projeto de Graduação em Engenharia de Petróleo. Universidade Federal do Rio de Janeiro, 2018.

OLSON, David L; DELEN, Dursun. **Advanced Data Mining Techniques**. Berlin: Springer, 2008. 182p. Disponível em: <<https://pdfs.semanticscholar.org/c1c7/4829d6430d468a1fe1f75eae217325253baf.pdf>> Acesso em: 12 jun. 2019.

POSSA, B.A V.; CARVALHO, M.L.B. de; REZENDE, R.S. F.; MEIRA JR., W Data mining: técnicas para exploração de dados. Belo Horizonte: UFMG. 1998

QUEIROS, L. M. O. **Estimação e Análise das Perdas Técnicas na Distribuição de Energia Elétrica**. Universidade Estadual de Campinas, Campinas-SP, Brasil, 2010.

Ricardo Linden “Técnicas de Agrupamento 2009. Disponível em: <http://www.fsma.edu.br/si/edicao4/FSMA_SI_2009_2_Tutorial.pdf> Acesso em: 12 jun. 2019.

SWAGATAM, Das; SUGANTHAN, Ponnuthurai Nagaratnam, 2011. Differential evolution: A survey of the state-of-the-art. IEEE Transactions on Evolutionary Computation, vol.15(1), pp:4-31.

SANTOS, Gustavo Maciel. **TERMOGRAFIA E INTELIGÊNCIA ARTIFICIAL NA DETECÇÃO DE FALHAS EM TRANSFORMADORES. 2017. TERMOGRAFIA E INTELIGÊNCIA ARTIFICIAL NA DETECÇÃO DE FALHAS EM TRANSFORMADORES (Pós Graduação em Engenharia Elétrica)** - Universidade Federal de Pernambuco, Recife, 2017.

SANTOS, Laerte. **Termografia em Subestações de Alta Tensão Desabrigadas.** Dissertação apresentada ao programa de Pós-Graduação em Engenharia da Energia da Universidade Federal de Itajubá. 2006.

SANTOS, R. (2008). **“Conceitos de Mineração de Dados Multimídia”**. WebMedia, 2008. Disponível em:
<<http://www.lac.inpe.br/~rafael.santos/Docs/WebMedia/2008/mmdm.pdf>> Acesso em: 12 jun.2019.

Sarajane M. Peres e Clodoaldo A. M. Lima “Técnicas de Agrupamento (Clustering)”. 2015. Disponível em: < <http://each.uspnet.usp.br/sarajane/wp-content/uploads/2015/09/clustering1.pdf>. > Acesso em: 12 jun.2019.

SENNA, Sidney Lima. **Computação evolucionária Aplicada ao Diagnóstico de Falhas Incipientes em Transformadores de Potência utilizando dados de Cromatografia.** 2010. 101f. Dissecção (Mestrado em Engenharia Elétrica) – Universidade Federal de Minas Gerais, 2010.

SEPÚLVEDA, H.L. **O Transformador e sua Influência nas Características de Funcionamento de um Sistema Elétrico de Potência: Uma experiência profissional.** Universidade Federal de Minas Gerais, 1981.

VAISALA. *The Vaisala Optimus™ DGA Monitor Technology Overview.*2017. Disponível em: <<https://www.vaisala.com/sites/default/files/documents/CEN-G-Optimus-DGA-Technology-in-nut-shell-B211588EN.pdf>> acessado dia 29 de outubro de 2018.

VASCONCELOS, L. M. R; CARVALHO, C. L. **Aplicação de regras de associação para mineração de dados na web.** Technical report, Universidade Federal de Goiás, 2004.

VERATTI, A. B. **Sistema Básico de Inspeção Termográfica**. p. 6, 1999.

VERMINNEN, P., VERATTI, A. B.; Junior, O. R. "**Termografia: Princípios e Aplicações**". Brasília: AGA S.A. Sistemas Infravermelhos. 1984.

ZIRBES, R. **Metodologias para Avaliação e Diagnóstico do Estado de Isolamentos de Papel Impregnado com Óleo Mineral**. 2003. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina.

ANEXO

CONCENTRAÇÃO DE GASES ENCONTRADOS (CEPEL)

HIDROGÊNIO (H2)	METANO(CH4)	ACETILENO(C2H2)	ETILENO(C2H4)	ETANOLC2H6)
22	6	4	6	5
36	5	0.4	20	6
180	11	0.4	27	18
18	3	0.4	15	0.4
100	62	0.4	7	12
12	6	0.4	25	1
12	13	1	58	18
30	6	0.4	0.4	5
120	42	0.4	46	80
64	48	49	160	26
1600	2200	74	2700	430
48	20	120	130	21
7	2	0.4	2	1
32	6	0.4	7	5
33	25	63	16	7
23	6	0.4	2	3
9	5	0.4	17	4
140	110	74	250	36
14	200	0.4	41	520
10	24	3	24	12
47	35	11	78	49
34	5	0.4	5	3
17	4	0.4	11	2
7	8	0.4	3	13
59	16	0.4	12	14
150	54	0.4	9	61
82	20	74	33	16
40	62	1	11	51
20	3	0.4	5	2
19	9	0.4	7	20
13	10	0.4	55	3
290	16	0.4	12	24
14	3	0.4	12	3
11	79	0.4	14	130
49	14	1	24	12
8	3	0.4	4	2
610	1800	270	6300	960
1	0.4	0.4	0.4	0.4
5	0.4	1	1	1
62	28	260	180	30

4	1	4	2	0.4
3	1	1	1	0.4
4	2	1	1	1
23	8	1	8	9
11	23	0.4	36	9
19	140	0.4	20	330
75	92	26	250	240
35	8	0.4	3	3
22	19	0.4	40	75
30	8	1	200	96
15	2	6	48	3
62	11	0.4	59	8
20	29	130	77	5
22	4	0.4	8	12
20	40	0.4	62	5
8	4	0.4	4	9
23	30	0.4	86	14
61	310	330	3000	570
88	6	0.4	15	23
350	350	5	860	290
28	6	2	5	4
9	4	0.4	15	3
21	8	0.4	17	25
69	43	0.4	20	32
50	50	5	66	36
430	2300	10	3000	1900
25	3	0.4	2	1
16	3	0.4	4	10
110	8	0.4	4	10
1	0.4	0.4	0.4	4
52	3	0.4	4	3
430	2300	10	3000	1900
7	0.4	0.4	0.4	0,4
14	38	0.4	22	55
51	97	0.4	3	110
48	29	0.4	20	17
30	76	0.4	48	220
600	1800	130	3800	520
24	3	0.4	8	4
16	1	0.4	1	1
48	12	0.4	18	14
56	6	10	7	5
25	18	1	39	33
590	8400	710	62000	92000
310	1200	5	1200	320
260	100	26	2200	210

150	340	14	1400	220
1700	1600	0.4	2100	440
980	1800	22	3200	460
1	1800	5	6100	3400
29	440	0.4	180	2900
130	220	0.4	87	1700
87	430	0.4	190	570
42	320	0.4	840	820
4800	39000	20	70000	19000
300	1400	9	2100	720
64	840	37	5100	820
430	450	0.4	2000	1000
380	280	0.4	330	280
460	570	38	720	100
150	380	49	2300	320
46	960	0.4	1000	720
420	800	0.4	370	100
99	480	0.4	220	2600
290	110	0.4	14	320
47	74	0.4	100	150
200	1200	75	2700	800
190	820	0.4	1500	780
51	53	0.4	29	200
45	900	3	8700	7900
28	320	13	1200	250
170	1200	24	5200	940
280	160	39	1000	610
6	2	0.4	7	5
30	520	0.4	1000	310
1600	55000	9300	74000	42000
44	95	0.4	13	33
420	1400	7	1500	640
34	41	0.4	5	69
19	4	1	20	6
610	1900	33	14000	2400
67	9	22	7	9
57	46	0.4	14	58
17	3	1	4	9
110	2	3	15	2
66	38	0.4	39	120
6000	14000	47	19000	18000
410	90	0.4	36	53
410	11	30	330	300
2500	3200	220	5300	620
12	5	1	4	3
64	160	1	68	1200

4	19	0.4	20	1
7	22	0.4	32	35
450	250	510	300	130
82	13	0.4	8	23
20	11	1	14	14
110	46	0.4	9	52
1400	460	0.4	310	940
1400	3500	3600	7900	2500
13	290	13	4100	2400
220	750	7	1200	300
250	430	9	700	120
36	250	0.4	410	130
27	130	1	280	110
2200	120	0.4	5	63
23	5	1	8	6
54	3300	0.4	3000	2200
11	8	6	39	11
580	880	160	1100	190
300	250	32	560	160
470	2800	0.4	4300	1400
440	3600	96	6500	1500
61	240	1	320	130
12	110	1	660	190
1000	1200	1200	2000	140
45	240	18	1100	260
57	29	0.4	190	220
85	160	2	38	270
32	51	0.4	52	52
72	27	140	220	26
260	650	3	950	220
3800	1100	1700	1300	300
550	67	240	89	30
78	580	1200	1500	130
5800	560	4700	830	31
430	40	480	68	13
3800	1700	4200	1300	200
6600	3300	6200	4200	620
4500	2100	1300	3000	250
7200	7200	1100	7700	3800
5000	2000	5500	3100	1900
1500	110	910	91	140
20000	1700	11000	6600	940
7	4	60	94	49
2100	210	1200	230	160
770	86	370	170	29
540	26	280	21	11

1500	640	1700	810	50
1900	1700	5300	2900	130
2400	770	1900	850	68
530	260	250	210	14
16	3	5	2	4
36	67	380	290	110
250	20	200	31	5
2600	660	810	750	47
49	4	0.4	1	3
200	240	100	380	74
970	240	1100	390	18
1200	2500	120	3700	710
500	160	240	140	48
630	670	1700	1100	81
110	860	1800	1500	200
200	110	230	150	12
94	5	20	25	4
4000	1200	4400	1900	130
230	120	740	220	14
6600	2300	3400	3500	210
11000	1300	4600	1400	140
9	2	52	5	1
51	6	11	6	1
16000	9400	22000	17000	1400
7	2	0.4	1	1
97	12	21	25	2
150	38	110	54	5
35	110	540	380	150
32	32	29	47	5
79	100	230	83	410
20000	6400	10000	7100	2600
1300	8400	5500	2200	5900
1100	140	350	2	14
200	6	8	8	5
39	4	15	4	0.4
6000	1900	2700	2400	280
81	21	29	46	11
14	7	22	15	1
19	0.4	0.4	4	0.4
380	250	1000	590	34
67	6	0.4	21	15
15000	10000	23000	21000	2200
4700	240	300	120	0.4
170	98	140	45	220
210	65	190	100	22
