

# Ferramenta de Aprendizado de Máquina para Previsão de Falha de Transformadores de Rede Elétrica

*Machine Learning Tool for Fault Prediction in Electric Grid Transformers*

Leonardo Fagundes Luz Serrano <sup>1</sup>  [orcid.org/0000-0001-9659-9895](https://orcid.org/0000-0001-9659-9895)

Victor Mendonça de Azevêdo <sup>2</sup>  [orcid.org/0000-0003-2943-4622](https://orcid.org/0000-0003-2943-4622)

Anthony José da Cunha Carneiro Lins <sup>3</sup>  [orcid.org/0000-0002-7153-841X](https://orcid.org/0000-0002-7153-841X)

<sup>1</sup> Residência em Inteligência Artificial, Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil

<sup>2</sup> Fundação para Inovações Tecnológicas, Recife, Brasil

<sup>3</sup> Departamento de Comunicação Social, Universidade Católica de Pernambuco, Recife, Brasil

E-mail do autor principal: Leonardo F. L. Serrano [lfls2@poli.br](mailto:lfls2@poli.br)

## Resumo

---

Este trabalho tem como objetivo o desenvolvimento de uma ferramenta de previsão de falhas de transformadores de rede elétrica para suporte da equipe técnica de manutenção preventiva. Com esse propósito, é feito o tratamento de uma base de registros de falha, conversão dos dados para o formato de uma série histórica mensal de falhas por transformador, cálculo de novos atributos a partir das informações da base, seguido pela estimativa do risco de falha no mês seguinte de cada transformador da rede e pelo ranqueamento dos transformadores com base nesse risco. A estimativa de risco é feita baseando na metodologia Aprendendo a Ranquear (*Learning To Rank - LTR*), ponto-a-ponto, na qual os itens de uma lista são ranqueados com base numa métrica calculada a partir de algoritmos de classificação ou regressão. É demonstrado um ganho significativo no desempenho do preditor devido ao processo de engenharia de atributos.

**Palavras-Chave:** Previsão de Falhas de Transformadores; Aprendendo a Ranquear; Engenharia de Atributos

## Abstract

---

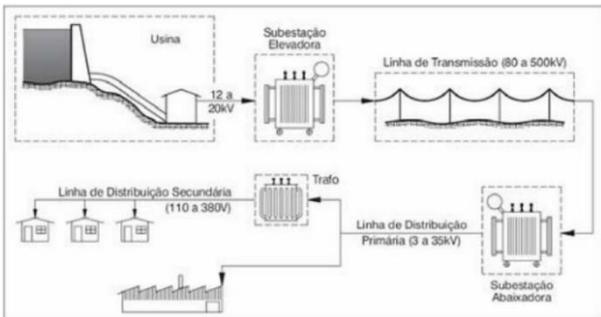
*This work aims for the development of a tool for forecasting faults within the transformers connected to the electric grid in order to support the maintenance team. With this purpose, the fault database is treated and converted to a list of monthly time series, one for each transformer, new features are calculated based on the information on the database, followed by the fault risk estimation for the next month in the series for each transformer and the ranking of the transformers. The risk estimation is done using the Pointwise Learning to Rank (LTR) methodology, in which the items on a list are ranked based on metric calculated using classifiers or regressors. A significant gain in performance is demonstrated due to the feature engineering process.*

**Key-words:** Transformer Fault Prediction; Learning to Rank; Feature Engineering

## 1 INTRODUÇÃO

O transformador é um dispositivo muito útil em circuitos elétricos, pois permite que haja fluxo de potência de um circuito primário para um circuito secundário adequando os níveis de tensão e corrente para os valores apropriados às funções de cada circuito [1]. No caso de sistemas de distribuição de energia elétrica, transformadores minimizam as perdas na distribuição ao permitir a transmissão em alta tensão e baixa corrente, reduzindo o aquecimento dos condutores por efeito Joule [2]. Dessa forma, torna-se viável que o ponto de geração e os consumidores estejam bastante afastados, em alguns casos por milhares de quilômetros.

Na Figura 1, é mostrado o diagrama de uma rede elétrica. A tensão é elevada para transmissão pelas subestações elevadoras. Nas subestações abaixadoras, a tensão é reduzida, mas mantida em níveis relativamente altos para fluxo nas cidades até, finalmente, chegar aos transformadores locais e ser reduzida aos valores padrões para os consumidores.



**Figura 1** - Esquemático simplificado de um sistema de fornecimento de energia elétrica.

De maneira simplificada, o transformador é composto internamente por um núcleo metálico envolto por espiras condutoras. O núcleo é submerso em óleo cuja função é servir de isolante elétrico e facilitador de resfriamento, dissipando o calor por convecção [3]. Usualmente transformadores são acompanhados por isoladores cerâmicos e dispositivos elétricos auxiliares.

Todos esses componentes estão sujeitos à depreciação e defeitos que podem levar a irregularidades no fornecimento de energia,

como tensões mais altas ou mais baixas que a nominal, ou até interrupção do fornecimento de energia [13]. Essas irregularidades acarretam prejuízos financeiros para a concessionária de energia, não só pela receita perdida com a interrupção do serviço como também por não atender às metas de qualidade exigidas pela Agência Nacional de Energia Elétrica (ANEEL) [4]. Além disso, transformadores são equipamentos caros e de grande porte, de modo que a substituição não programada em caso de quebra é custosa e requer bastante tempo para sua conclusão. Dessa forma, observa-se a necessidade de técnicas de diagnóstico e manutenção preventiva.

O estado do óleo de resfriamento e a temperatura do equipamento são bons indicativos da necessidade de manutenção, de modo que os principais métodos de diagnóstico de um transformador são a análise físico-química do óleo de resfriamento, a análise de gás dissolvido no óleo, imagens térmicas para detecção de pontos de aquecimento e a análise da resposta em frequência [5]. A manutenção da rede pode ser feita de modo periódico, com a realização de inspeções e desligamentos agendados em cada um dos transformadores da rede elétrica. No entanto, uma abordagem mais eficiente é a estimativa do risco de falha de cada transformador e a construção de uma lista ranqueada de prioridade de manutenção preventiva. Esta última pode ser beneficiada pela capacidade de detecção de padrões provida por técnicas de aprendizado de máquina.

De posse de informações sobre os transformadores, tais como histórico de falhas, idade do transformador, carga, temperatura, fabricante, e quaisquer outros dados disponíveis, é possível gerar um algoritmo capaz de estimar o risco de falha [6]. A precisão dos resultados depende da qualidade dos dados disponíveis e dos métodos utilizados.

Neste trabalho, almeja-se aumento na eficácia das equipes de manutenção preventiva por meio do desenvolvimento de uma ferramenta auxiliar capaz de estimar e ranquear o risco de falha dos transformadores. É realizado o pré-processamento de uma base de dados real contendo registros de falha, seguido pela aplicação de uma metodologia de ranqueamento em aprendizado de máquina chamada *Learning To Rank* (Aprendendo A Ranquear - LTR) [7, 8].

O restante do trabalho é organizado da seguinte forma: na sessão II é feita uma revisão do estado da arte; na sessão III é explanada a abordagem utilizada neste trabalho; na sessão IV são apresentados os resultados obtidos e na sessão V são dadas as conclusões deste trabalho.

## 2 PREVISÃO DE FALHAS DE TRANSFORMADORES

A utilização de algoritmos de aprendizado de máquina na predição de falhas de transformadores é uma área de pesquisa que vem sendo explorada há décadas. A literatura relacionada a esse problema é vasta e as soluções apresentadas são variadas e dependem dos dados disponíveis ao pesquisador e do objetivo do problema. Há grande diversidade nas abordagens citadas na literatura.

Em 2006, pesquisadores da Companhia Edison de Nova Iorque, desenvolveram um sistema que objetiva a previsão de falhas nos cabos alimentadores, que levam energia até os transformadores locais [6]. A base de dados conta com 150 atributos para cada alimentador da rede. O sistema é baseado em *boosting*, no qual uma série de modelos simples é combinado em um único modelo, mais sofisticado. Cada um dos modelos base gera uma previsão usando apenas um dos atributos disponíveis, o que permite a verificação de importância de cada atributo. Dessa forma, pode ser estimada uma causa de falha a ser repassada para a equipe de manutenção. A transparência nas razões da decisão do modelo serve também para aumentar a confiança dos usuários na análise do sistema. Como resultado, 75% dos alimentadores que falharam num período de 3 meses estavam listados dentre os 25% de maior risco. Os pesquisadores observaram que a temperatura ambiente é um dos fatores de maior influência na quantidade de falhas, de modo que o verão é o período de maior risco. A métrica de desempenho utilizada no estudo é a Média Normalizada do Ranque de Falhas (MNRF), definida pela Equação (1):

$$\text{MNRF} = 1 - \left[ \frac{\sum_i R_i}{N_F N_T} \right] \quad (1)$$

na qual  $N_T$  é o número total de elementos na lista,  $N_F$  é o número total de falhas no período alvo e  $R_i$  é o ranque estimado pelo preditor para os elementos que de fato falharam. Essa métrica apresenta valores na faixa de zero a um, na qual valor próximo a um indica que as falhas estão próximas ao topo do ranqueamento. Caso a métrica MNRF tenha valor próximo a 0,5, isso indica que as falhas estão distribuídas de maneira uniforme ao longo do ranqueamento.

Em 2005, Unsal *et al.* utilizaram a distribuição de Poisson para previsão do número de falhas numa rede elétrica [9]. Em 2009, Liu e Wang utilizaram Máquinas de Vetores de Suporte (Support Vector Machine - SVM) para previsão de falhas de transformadores [10]. Em 2018, Lin *et al.* aplicaram uma categoria de Rede Neural Artificial Recorrentes chamada LSTM-DBN (*Long Short-Term Memory - Deep Belief Network*) para previsão de falhas por meio da modelagem da tendência da concentração de gás no óleo [11].

Em 2019, Peimankar *et al.* introduziram uma versão binária da Otimização por Enxame de Partículas Multiobjectivo (*Multi-Objective Particle Swarm Optimization - MOPSO*) como forma de seleção dos atributos a serem usados no treinamento de um ensemble de classificadores para diagnóstico de transformadores [12]. O algoritmo usa dados obtidos a partir da análise de gás dissolvido no óleo de resfriamento. O sistema foi analisado com e sem a seleção de atributos proposta, apresentando ganho de desempenho com a utilização da seleção de atributos.

## 3 METODOLOGIA

A base de dados é formada pelos registros de ocorrência de falha nos transformadores monitorados de Janeiro de 2012 a Dezembro de 2014. Cada registro contempla informações sobre as falhas ocorridas ao longo de um determinado mês em transformador por uma causa específica. São citados a subestação, alimentador e bloco ao qual o transformador está conectado, bem como sua potência nominal.

Sobre as falhas, são informados a Duração de Interrupção Individual por Unidade Consumidora (DIC) e a Frequência de Interrupção Individual por Unidade Consumidora (FIC). O DIC indica quanto tempo, em média, o consumidor ficou

**DOI: 10.25286/rep.v5i2.1351**

sem energia e o FIC indica quantas vezes o fornecimento de energia foi interrompido [4].

A base apresenta algumas limitações significativas quando comparadas as bases descritas na literatura de previsão de falhas de transformadores, como por exemplo a ausência de informações sobre o ponto de operação, tensões e correntes de entrada e saída, carga, e a temperatura. Os registros não especificam se a falha provocou comportamento anormal do equipamento ou completa interrupção de funcionamento. Também não há informação sobre transformadores que não apresentaram falhas e de quantos transformadores foram monitorados no total.

Os 35 mil registros são divididos entre 141 causas de falha seguindo uma distribuição de pareto, de modo que as 10 causas mais frequências contemplam quase metade dos registros. A análise dessas causas indica que a maioria das falhas é causada por eventos não previsíveis, como objetos, árvores e animais que caem na rede elétrica, postes abalroados ou vandalismo. Interrupções para manutenção preventiva também estão inclusas nas causas de falha. Considerando que o propósito do projeto é o auxílio às equipes de manutenção preventiva, são retirados da base os registros de falhas provocadas por fatores externos à rede elétrica pois estes podem ser considerados como ruído para o preditor, restando, assim, em torno de 10 mil registros.

Na etapa de aprendizagem de máquina, optou-se pela metodologia *Learning To Rank* (Aprendendo A Ranquear - LTR) ponto-a-ponto, na qual os itens de uma lista são ranqueados com base numa métrica calculada a partir de algoritmos de classificação ou regressão. Esses algoritmos são supervisionados e, portanto, para seu treinamento, requerem um conjunto de amostras compostas por um vetor de atributos e uma saída esperada.

Para geração dessas amostras, é construída uma série histórica de falhas para cada transformador, seguido pelo janelamento dessas séries, de modo que cada amostra tenha como saída esperada o número de falhas em um determinado mês em determinado transformador e seus atributos incluem o número de falhas ocorridas nos 12 meses anteriores. Também são acrescentadas as informações de potência e subestação.

Além disso, é realizado um processo de engenharia de atributos, no qual informações são extraídas dos dados disponíveis e reformatadas de modo a aumentar o desempenho do preditor. São acrescentadas colunas relativas a: número de meses desde a última falha; a percentagem de meses que o transformador apresentou falha desde o início do monitoramento; o tempo médio entre falhas do transformador; uma estimativa de risco em função da subestação a que está conectado o transformador, dada pela razão do número de falhas em todos os transformadores conectados à subestação dividido pelo número de transformadores e pelo número de meses monitorados; e, finalmente, uma estimativa a-priori do risco de falha, descrita pela Equação (2):

$$\text{Risco a priori} = (1 - F)^n \cdot F \quad (2)$$

na qual  $F$  é o inverso do tempo médio entre falhas do transformador e  $n$  é o número de meses desde a última falha.

Finalmente, as amostras são usadas para o treinamento do modelo-base do LTR. Provisoriamente, foi utilizado o classificador *Random Forest* para verificar a contribuição dos novos atributos no desempenho do preditor por seu treinamento ser relativamente rápido, acelerando os ajustes no desenvolvimento da abordagem.

O modelo-base é retreinado com todos os dados históricos disponíveis até, e somente até, o mês alvo da previsão. Os atributos adicionais também são calculados de acordo. Dessa forma, todas as informações úteis são aproveitadas no modelo, com o cuidado de não contaminar os resultados pelo uso de informações futuras ao alvo.

## 4 RESULTADOS

A análise do desempenho do preditor foi feita com base em três métricas: a acurácia considerando as previsões para todos os transformadores; a métrica MNRF, indicadora da concentração de falhas corretamente previstas no topo do ranqueamento; e a acurácia considerando apenas os 100 transformadores apontados pelo preditor como os de

maior risco de falha no mês alvo. Esta última métrica é relevante pois é razoável supor que a equipe de manutenção pode atuar apenas em um número limitado de transformadores em um mês, optando por aqueles que estão no topo do ranqueamento do preditor. É realizada a média de 30 repetições de treinamento e teste no cálculo das métricas.

Na Tabela I, são apresentados os resultados obtidos nas previsões para os 12 meses do ano de 2014 quando a base de dados não tem atributos adicionais, apenas os disponíveis originalmente. Observa-se que, apesar da acurácia total acima de 80%, a acurácia de previsão nos 100 trafos apontados como de maior risco foi apenas de aproximadamente 40%. Isso indica que os atributos usados pelo preditor na tomada de decisão não estão fortemente relacionados com as ocorrências de falha, caso contrário a segunda métrica seria maior, não menor, que a primeira. Além disso, a métrica MNRF aponta que falhas estão uniformemente distribuídas ao longo do ranqueamento.

**Tabela I** - Métricas de desempenho do preditor treinado com dados sem engenharia de atributos

|               | <b>Acurácia</b> | <b>Acurácia Top100</b> | <b>MNR F</b> |
|---------------|-----------------|------------------------|--------------|
| Jan           | 87,7%           | 87,7%                  | 0.49         |
| Fev           | 72,8%           | 41,3%                  | 0.38         |
| Mar           | 86,5%           | 77,4%                  | 0.52         |
| Abr           | 89,6%           | 12,2%                  | 0.57         |
| Mai           | 95,0%           | 87,7%                  | 0.59         |
| Jun           | 86,4%           | 41,6%                  | 0.50         |
| Jul           | 55,5%           | 11,5%                  | 0.48         |
| Ago           | 92,6%           | 01,6%                  | 0.62         |
| Set           | 86,6%           | 01,7%                  | 0.52         |
| Out           | 88,1%           | 18,4%                  | 0.43         |
| Nov           | 96,0%           | 76,1%                  | 0.59         |
| Dez           | 75,7%           | 12,4%                  | 0.62         |
| <b>Médi a</b> | <b>84,4%</b>    | <b>39,1%</b>           | <b>0.53</b>  |

Na Tabela II, são apresentados os resultados utilizando a base adaptada, com os atributos adicionais. A acurácia nos 100 transformadores de maior risco é em média de aproximadamente 80%, um aumento de 40% em relação ao resultado obtido sem os novos atributos. A métrica MNRF continua próxima aos 0,5, indicando que as falhas estão uniformemente distribuídas ao longo do ranqueamento. Uma provável razão é que os novos

atributos estão altamente correlacionados com algumas das falhas ocorridas, mas não com outras.

**Tabela II** - Métricas de desempenho do preditor treinado com dados com engenharia de atributos.

|              | <b>Acuráci a</b> | <b>Acurácia Top100</b> | <b>MNR F</b> |
|--------------|------------------|------------------------|--------------|
| Jan          | 87,8%            | 82,8%                  | 0.54         |
| Fev          | 75,0%            | 84,9%                  | 0.42         |
| Mar          | 86,7%            | 78,8%                  | 0.57         |
| Abr          | 94,2%            | 81,1%                  | 0.68         |
| Mai          | 95,1%            | 89,1%                  | 0.56         |
| Jun          | 89,2%            | 78,2%                  | 0.55         |
| Jul          | 84,3%            | 23,3%                  | 0.61         |
| Ago          | 98,2%            | 86,6%                  | 0.69         |
| Set          | 95,6%            | 89,9%                  | 0.55         |
| Out          | 92,1%            | 93,3%                  | 0.37         |
| Nov          | 96,7%            | 85,1%                  | 0.61         |
| Dez          | 80,2%            | 80,5%                  | 0.55         |
| <b>Média</b> | <b>89,6%</b>     | <b>79,5%</b>           | <b>0.56</b>  |

## 5 CONCLUSÕES

Neste trabalho, foi desenvolvida uma metodologia para estimação do risco de falha dos transformadores de uma rede elétrica a partir de uma base de dados real provida por uma concessionária de energia de uma capital brasileira, com o objetivo de auxiliar a equipe técnica de manutenção preventiva. A metodologia indica uma ordem de prioridade para diagnóstico dos transformadores da rede.

Os registros da base contêm um grande número de causas de falha relacionadas a eventos externos à rede elétrica, tais como queda de árvores, postes abalroados, vandalismo, animais e objetos que entraram em contato com a rede. Esses registros podem ser considerados como ruído para o treinamento de um preditor, considerando que os registros da base de dados contam apenas com informações relativas ao sistema elétrico. Sendo assim, esses registros são removidos.

Com relação aos atributos disponíveis, a concessionária ainda não dispõe de informações sobre a carga e temperatura dos transformadores ou de análises técnicas do óleo em seus registros. Os trabalhos relacionados à previsão de falhas em rede elétrica encontrados na literatura, em geral, utilizam essas informações como alicerce de seus estudos,

apontando a qualidade do óleo, carga e a temperatura como fatores de grande influência na ocorrência de falhas.

Para compensar a ausência desses dados e extrair previsões mais acuradas, é feito um pré-processamento específico adaptado à base de dados disponível com adição e reformatação de atributos, de modo que conhecimentos técnicos e de modelos matemáticos relacionados ao problema possam ser considerados pelo algoritmo de aprendizado de máquina.

Comparando os resultados obtidos com e sem a engenharia de atributos, nota-se uma evolução da acurácia no topo do ranque, indicando que o processo contribuiu para a previsão mais acurada de alguns tipos de falhas e apontando transformadores que de fato necessitam de manutenção preventiva no topo da lista.

Porém, observa-se pela métrica MNRF que os transformadores defeituosos são distribuídos de modo uniforme ao longo do ranqueamento, indicando que, apesar da acurácia próxima à 90% na classificação dos transformadores quanto à ocorrência ou não de falha, a estimativa de risco não atinge os níveis de precisão necessários para um ranqueamento de todos os transformadores da base. Ou seja, boa parte das causas de falha tem pouca conexão com os atributos disponíveis mesmo após a engenharia de atributos.

Como trabalhos futuros, tem-se execução de um procedimento mais minucioso para determinação do modelo-base do LTR e seus hiper parâmetros, bem como a implementação de outras metodologias e comparação de desempenho. Dentre as metodologias alternativas, tem-se análise de sobrevivência, detecção de anomalias e outras abordagens de ranqueamento, como LTR em pares (*pairwise LTR*) e LTR em listas (*listwise LTR*).

## AGRADECIMENTOS

Os autores desde trabalho gostariam de agradecer à Fundação para Inovações Tecnológicas (FITec), à Secretaria de Ciência, Tecnologia e Inovação do Estado de Pernambuco (SECTI), ao Centro de Manufatura Avançada (CMA-Parqtel), à Universidade de Pernambuco (UPE) e à Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE) por prover

a estrutura, didática e financiamento que tornaram esse trabalho possível.

## Referências

- [1] FITZGERALD, A. E.; KINGSLEY, C.; STEPHEN D. **Electric machinery**. Boston, Mass: McGraw-Hill, 2009.
- [2] NILSSON, J. W.; SUSAN A. **Electric circuits**. Upper Sadle River: Prentice-Hall, 2000.
- [3] HARLOW, J. H.; GRIGSBY, L. L. **Electric power transformer engineering**. Boca Raton: CRC Press, Taylor & Francis Group, 2012.
- [4] AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST**. 2015.
- [5] SOUZA; D. C. P. **Falhas e defeitos ocorridos em transformadores de potência do sistema elétrico da Celg, nos últimos 28 anos: um estudo de caso**. 2008. **Dissertação**. (Mestrado em Engenharia Elétrica e Computação) -- Escola de Engenharia Elétrica e de Computação, Universidade de Goiás, Goiânia.
- [6] GROSS, P. et al. **Predicting electricity distribution feeder failures using machine learning susceptibility analysis**. Eighteenth Conference on Innovative Applications of Artificial Intelligence IAAI-06. 2006.
- [7] LIU, T.-Y. **Learning to rank for information retrieval**. Boston: Now the essence of knowledge, 2009.
- [8] LIU, T.-Y. **Learning to Rank for Information Retrieval**. 2009. Disponível em: <<http://www.conference.org/www2009/pdf/T7A-LEARNING%20TO%20RANK%20TUTORIAL.pdf>>. Acesso em: 12 nov 2019.
- [9] UNSAL, A. et al. **Predicting the failures of transformers in a power system using the poisson distribution: A case study**. ELECO 2005 International Conference on Electrical and Electronics Engineering. 2005. Disponível em: <[https://www.researchgate.net/publication/303824402\\_PREDICTING\\_THE\\_FAILURES\\_OF\\_TRANSFORMERS\\_IN\\_A\\_POWER\\_SYSTEM\\_USING\\_THE\\_POISSON\\_DISTRIBUTION\\_A\\_CASE\\_STUDY](https://www.researchgate.net/publication/303824402_PREDICTING_THE_FAILURES_OF_TRANSFORMERS_IN_A_POWER_SYSTEM_USING_THE_POISSON_DISTRIBUTION_A_CASE_STUDY)>
- [10] LIU, T.; WANG, Z. **Design of power transformer fault diagnosis model based on support vector machine**. 2009 International

Symposium on Intelligent Ubiquitous Computing and Education. IEEE, 2009. Disponível em: <<https://doi.org/10.1109/iuce.2009.59>>.

**[11]** LIN, J. et al. **Prediction method for power transformer running state based on LSTM\_DBN network.** Energies, MDPI AG, v. 11, n. 7, p. 1880, jul. 2018. Disponível em <<https://doi.org/10.3390/en11071880>>.

**[12]** PEIMANKAR, A. et al. **Evolutionary multi-objective fault diagnosis of power transformers.** Swarm and Evolutionary Computation, v. 36, p. 62-75, 2017.

**[13]** KIMMET, M. et al. **Predicting Distribution Transformer Failures. Transmission & Distribution World.** Disponível em: <<https://www.tdworld.com/asset-management-service/predicting-distribution-transformer-failures>>. Acesso em: 11 nov. 2019.