



CONGRESSO BRASILEIRO
DE ENGENHARIA QUÍMICA EM
INICIAÇÃO CIENTÍFICA

21-24 Julho de 2019
Uberlândia/MG



APLICAÇÃO DE SUPPORT VECTOR MACHINE NA DETECÇÃO DE FALHAS EM SENSORES DE UMA COLUNA DEBUTANIZADORA

L. C. BARRETO¹, B. C. V. SANTOS¹, A. P. M. R. SOARES¹, G. D. SILVA¹ e
F. O. CARVALHO¹

¹ Universidade Federal de Alagoas, Centro de Tecnologia
E-mail para contato: lucascosta_108@hotmail.com

RESUMO – Com a automatização de processos tecnológicos dentro do ambiente industrial, atrelada a presença de sistemas de controle e monitoramento das plantas, faz-se necessário a garantia da integridade das informações enviadas a rede de controle e monitoramento, para que ela possa operar sem falhas que comprometam a produção. Visando atender a essa necessidade, esse trabalho propôs a utilização da técnica de inteligência artificial Máquina de Vetores Suporte (do inglês *Support Vector Machine*, (SVM)), para criação de um sistema de detecção de falhas em sensores industriais. O estudo de caso foi realizado em sensores instalados em uma coluna debutanizadora que opera em uma planta de dessulfurização. O modelo se mostrou satisfatório na detecção das falhas nos sensores. Os resultados mostraram que a *Support Vector Machine* obteve desempenho satisfatório na detecção de falhas para este tipo de problema.

1. INTRODUÇÃO

Segundo Chiang *et al.* (2001), o aumento da procura por sistemas de detecção de falhas é justificado pelo fato de os processos industriais estarem se tornando cada vez mais instrumentados, o que resulta em uma grande quantidade de dados disponíveis para utilização no monitoramento. Estes dados ficam armazenados em computadores robustos, sob a forma de históricos do processo, e são extremamente necessários para análise da operação. Isto levou ao desenvolvimento de ferramentas de detecção e diagnóstico de falhas baseadas na análise destes históricos, eliminando a necessidade de se utilizar modelagem fenomenológica do processo. Dentre estas ferramentas, pode-se citar a *support vector machine* (SVM), que foi utilizada neste trabalho.

Devido ao seu forte respaldo matemático, tendo construção mais geral e aplicabilidade mais ampla, e por ter apresentado alto desempenho em aplicações práticas, inclusive superior a muitos algoritmos de aprendizagem conhecidos, as SVMs têm atraído o interesse da comunidade de aprendizagem de máquina (SOUZA, 2011).

Sendo assim, esse trabalho buscou investigar a aplicação das SVM, técnica baseada em inteligência artificial oriunda do conceito estatístico, para o problema de detecção de falhas em sensores. Para tal, foram utilizados registros de sensores de uma coluna debutanizadora, obtidos em Fortuna *et al.* (2007).

1.1. Máquinas de Vetores Suporte

As máquinas de vetor de suporte (ou SVMs, do inglês *support vector machines*), propostas inicialmente por Vapnik (1999), consistem em uma categoria de aprendizado de máquina que tem fundamentação teórica na teoria de aprendizagem estatística desenvolvida por Vapnik (1999) e podem ser utilizadas na classificação de conjunto de dados linearmente e não linearmente separáveis, especificamente, no caso desse trabalho na detecção de falhas, como reportado por Silva (2012) na detecção de falhas em motores elétricos e por Granzotto e Oliveira-Lopes (2015) na detecção de falhas em um sistema para a produção de ciclopentenol em um CSTR.

O fundamento básico do funcionamento de uma *support vector machine* é a construção de um hiperplano (objeto matemático), o qual pode possuir uma dimensão R^n , possibilitando a separação dos dados em classes distintas. Essa separação tem abordagem baseada na Teoria de Aprendizagem Estatística, utilizando a otimização matemática, para promover a busca de um hiperplano ótimo, adotando os mínimos quadrados por meio da minimização estrutural de risco. A otimização é realizada por meio de programação quadrática ou convexa, baseadas no método dos multiplicadores de Lagrange (SARAIVA, 2017).

A equação geral de um hiperplano é dada pela equação 1, onde w é o vetor peso e b , o bias, que são ajustados de acordo com os dados fornecidos na etapa de treinamento(x), procedimento que se repete até que o hiperplano seja capaz de executar a separação dos dados em diferentes classes (BEN-HUR, WESTON, 2010).

$$f(x) = w \cdot x + b \quad (1)$$

Para a resolução de problemas de natureza não linear, é utilizado um artifício matemático de transformação de espaço vetorial chamado de uma função *kernel*. Desta forma, essas funções são utilizadas na SVM para realizar o mapeamento e a mudança do espaço característico do problema para um de maior dimensão, em que os dados possam ser linearmente separáveis (SARAIVA, 2017). Os tipos de função *kernel* mais utilizadas são: Linear, Polinomial e Gaussiana (RBF), dadas pelas equações 2, 3 e 4, respectivamente.

$$K = x_j^T x_j \quad (2)$$

$$K = (x_i^T x_j - k)^p \quad (3)$$

$$K = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (4)$$

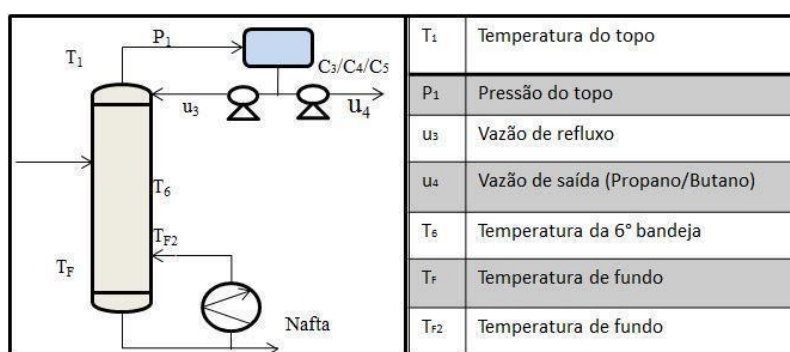
Para o *kernel* Polinomial (Equação 3), p e k devem ser especificadas e para o Gaussiano (Equação 4), a amplitude σ^2 deve ser especificada.

2. METODOLOGIA

2.1. Descrição do Conjunto de Dados

A coluna debutanizadora em estudo opera em uma planta de dessulfurização e separação de nafta, a Figura 1 mostra o esquema da coluna e as variáveis envolvidas nessa etapa do processo.

Figura 1 – Esquema de uma coluna debutanizadora e variáveis inerentes ao processo.



Devido a frequente ocorrência de falhas em sensores de temperatura e pressão, estes foram considerados para o estudo de caso. No total, o conjunto de dados continha 2394 registros referentes às medições contínuas realizadas no equipamento dentro de um determinado intervalo de tempo. Todo esse conjunto foi obtido em Fortuna *et al.* (2007), vale salientar que os dados disponibilizados na literatura citada já estavam normalizados e tratados estatisticamente e prontos para utilização para construção do modelo de detecção de falhas.

2.2. Simulação das Falhas e Construção dos Modelos

Foram inseridas, aleatoriamente, falhas nas medições dos sensores de temperatura de topo e pressão de topo dentro de um intervalo de $\pm 30\%$ em relação ao valor real. Após essa etapa, o conjunto total dos dados (com e sem falha) passou a ter 7182 registros, os quais poderiam ser classificados dentro dos seguintes grupos:

- Grupo 1: Falha no sensor de pressão de topo;
- Grupo 2: Falha no sensor de temperatura de topo;
- Grupo 3: Falha simultânea dos sensores;
- Grupo 4: Funcionamento normal do sistema.

2.3. Treinamento, Validação e Topologia dos Modelos

O modelo desenvolvido com a utilização das Máquinas de Vetores suporte, foi treinado apenas para identificar a ocorrência de falhas nos sensores, sem necessariamente discernir qual dos sensores apresentava a falha, o que seria o diagnóstico. Isso ocorreu em razão das limitações computacionais encontradas no software utilizado, para essa ferramenta em específico. Dessa forma, os grupos 1, 2 e 3 se tornaram uma só classe e o grupo 4 uma outra classe. Sendo assim, o problema passou a ter apenas duas classes (dados com ou sem falha) sendo a classe 0 relativa aos dados sem a inserção de falhas e a classe 1 (mostradas na matriz de confusão Figura 2) relativa aos dados que representavam a operação da coluna com falhas operacionais.

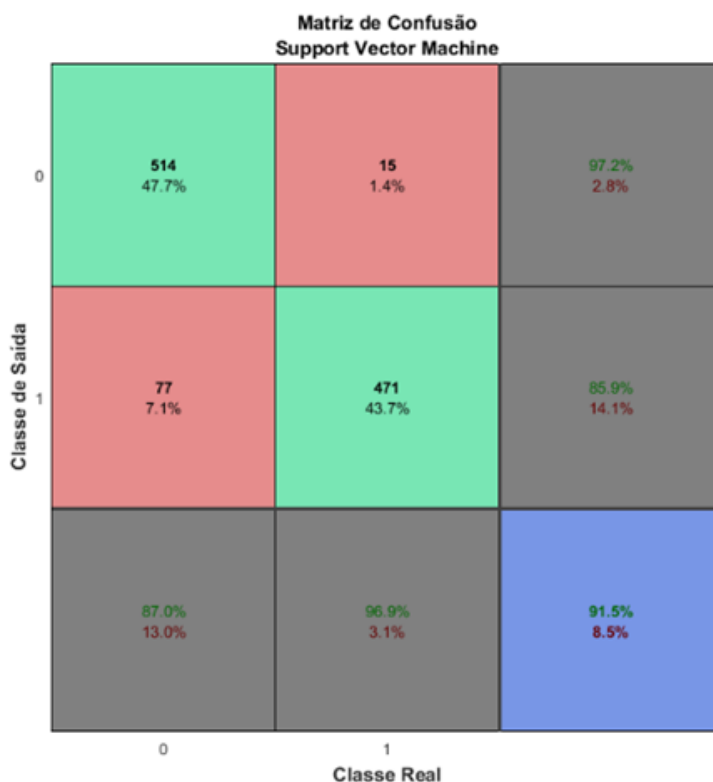
Vale salientar que a escolha da função *kernel* do modelo SVM foi selecionada através de métodos empíricos. Desta forma, a função *kernel* Gaussiana (RBF) foi a que apresentou melhor desempenho para o conjunto de dados analisados neste trabalho.

Todo o sistema de detecção de falhas, desde a simulação das falhas e treinamento dos modelos, foi desenvolvido utilizando a linguagem de programação Python®. Vale salientar que na etapa de validação do modelo, foram utilizados 15% dos dados do conjunto original, registros esses que não foram utilizados na etapa de treino. Desta forma, foi possível avaliar a capacidade de generalização do modelo desenvolvido.

3. RESULTADOS

A avaliação do modelo desenvolvidos foi realizada a partir da observação dos resultados de desempenho mostrados na matriz de confusão geradas na etapa de validação (Figura 2).

Figura 2 – Matriz de Confusão para o sistema de detecção de falhas com o modelo *Máquinas de Vetores Suporte* (SVM).



A partir da Figura 2 (Modelo *Máquinas de Vetores Suporte*) é possível ver o percentual de acertos na classificação do modelo na fase de validação. Desta forma, podemos perceber que dos 529 casos de teste para a classe 1, que representa o perfeito funcionamento do sistema, 514 (97,2%) foram classificados corretamente. A classe 2 por sua vez, foi identificada corretamente 471 vezes (85,9%), dos 548 casos testados. Dessa forma, o modelo obteve uma taxa de acerto global de 91,5%, desempenho que denota a eficiência e aplicabilidade do modelo de classificação SVM para o estudo de caso avaliado neste trabalho.

4. CONCLUSÃO

Neste trabalho foi demonstrada a construção de um sistema de detecção de falhas em sensores industriais. A análise dos resultados obtidos nos permite dizer que o modelo de detecção de falhas desenvolvido neste trabalho a partir da técnica de classificação com a *Máquinas de Vetores Suporte* teve uma alta taxa de acerto, além disso, poucos foram os registros de alarmes falsos e erros de classificação por parte dessa técnica. Dessa forma, podemos concluir que a SVM se mostrou uma técnica bastante promissora no que diz respeito a sua utilização na rede de comunicação industrial. Sendo assim, os resultados obtidos neste trabalho demonstram que o modelo desenvolvido tem potencial de utilização para o estudo de caso analisado. Entretanto, para trabalhos posteriores pensa-se no desenvolvimento de um modelo que execute também o diagnóstico das falhas, de modo que seja possível não somente identifica-las como também saber o equipamento ou sensor que deu origem a essa falha.

5. REFERÊNCIAS

- CHIANG, L. H.; RUSSELL, E. L.; BRAATZ, R. D. *Fault Detection and Diagnosis in Industrial Systems*. Londres: Springer, 2001.
- BEM-HUR, A.; WESTSON J. A User's guide to Support Vector Machine data maning techniques for the Life Sciences. *Methods in Molecular biology*, v. 609, 2010.
- FORTUNA, L.; GRAZIANI, S.; RIZZO, A.; XIBILIA, M. G. *Soft Sensors for Monitoring and Control of Industrial Processes*. Londres: Springer, 2007.
- GRANZOTTO, M. H.; OLIVEIRA-LOPES, L. C.; "Desenvolvimento de Sistema de Detecção de Falhas Baseado em Aprendizado Estatístico de Máquinas de Vetores de Suporte", p. 11819-11828. In: *Anais do XX Congresso Brasileiro de Engenharia Química - COBEQ 2014 [= Blucher Chemical Engineering Proceedings, v.1, n.2]*. São Paulo: Blucher, 2015.
- SARAIVA, S. V. *Estudo de um controlador preditivo baseado em sistemas inteligentes*. Trabalho de conclusão de curso. Universidade Federal de Alagoas, Maceió, 2017.
- SILVA, V. A. D. *Detecção de falhas em motores elétricos através das máquinas de vetores de suporte*. Dissertação (Mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica, Campinas, 2012. Disponível em: <<http://www.repositorio.unicamp.br/handle/REPOSIP/265564>>. Acesso em: 13 mar. 2019.
- SOUZA, F. C. M. *Técnicas de busca heurística para otimização de parâmetros de Máquinas de Vetores Suportes*. Dissertação submetida ao programa de Pós-Graduação em Ciência da Computação. Universidade Federal de Pernambuco, Recife, 2011.
- VAPNIK, V. N. *The Nature of Statistical Learning Theory*. Springer Verlag, New York, 2nd edition, 1999.