

# Detecção de Anomalias em Máquinas de Empacotamento

Alexandre G. Peixoto

Inteligência Artificial e Generativa  
Universidade Católica de Santa Catarina – Joinville, SC – Brazil

alexandre.peixoto@catolicasc.edu.br

**Abstract.** *This article describes the development of an anomaly detection system using Support Vector Machines (SVM) in Python. The implementation utilizes OneClassSVM on sensing data from a packaging machine, which is analyzed and standardized for training purposes. The prediction is integrated into the dataset and presented graphically.*

**Resumo.** *Este artigo descreve o desenvolvimento de um sistema de detecção de anomalias utilizando Support Vector Machines (SVM) em python. A implementação utiliza OneClassSVM sobre os dados de sensoramento de uma máquina de empacotamento, que são analisados e padronizados para realização do treinamento. A predição é incorporada ao dataset e apresentada graficamente.*

## 1. Introdução

A detecção precoce de anomalias em equipamentos industriais desempenha um papel fundamental na garantia da eficiência operacional, segurança e sustentabilidade nas indústrias modernas. Anomalias são variações no comportamento normal de um equipamento que podem indicar um problema potencial e podem ser causadas por diversos fatores, como desgaste, componentes defeituosos, condições ambientais adversas ou operação inadequada. Detectar esses problemas para atuar na correção significa reduzir significativamente os prejuízos.

## 2. Implementação

O código do sistema de detecção foi desenvolvido em Python, linguagem comumente utilizada para análise de dados. O modelo aplicado se dá através da biblioteca “sklearn.svm.OneClassSVM”, um dos métodos de classificação mais comuns[2], capaz de lidar eficientemente com a detecção de uma única classe. O objetivo deste modelo é distinguir um conjunto de dados de outros objetos[1]. Sendo assim, torna-se especialmente útil em cenários em que a maioria dos dados é considerada normal e apenas uma pequena porção é potencialmente anômala.

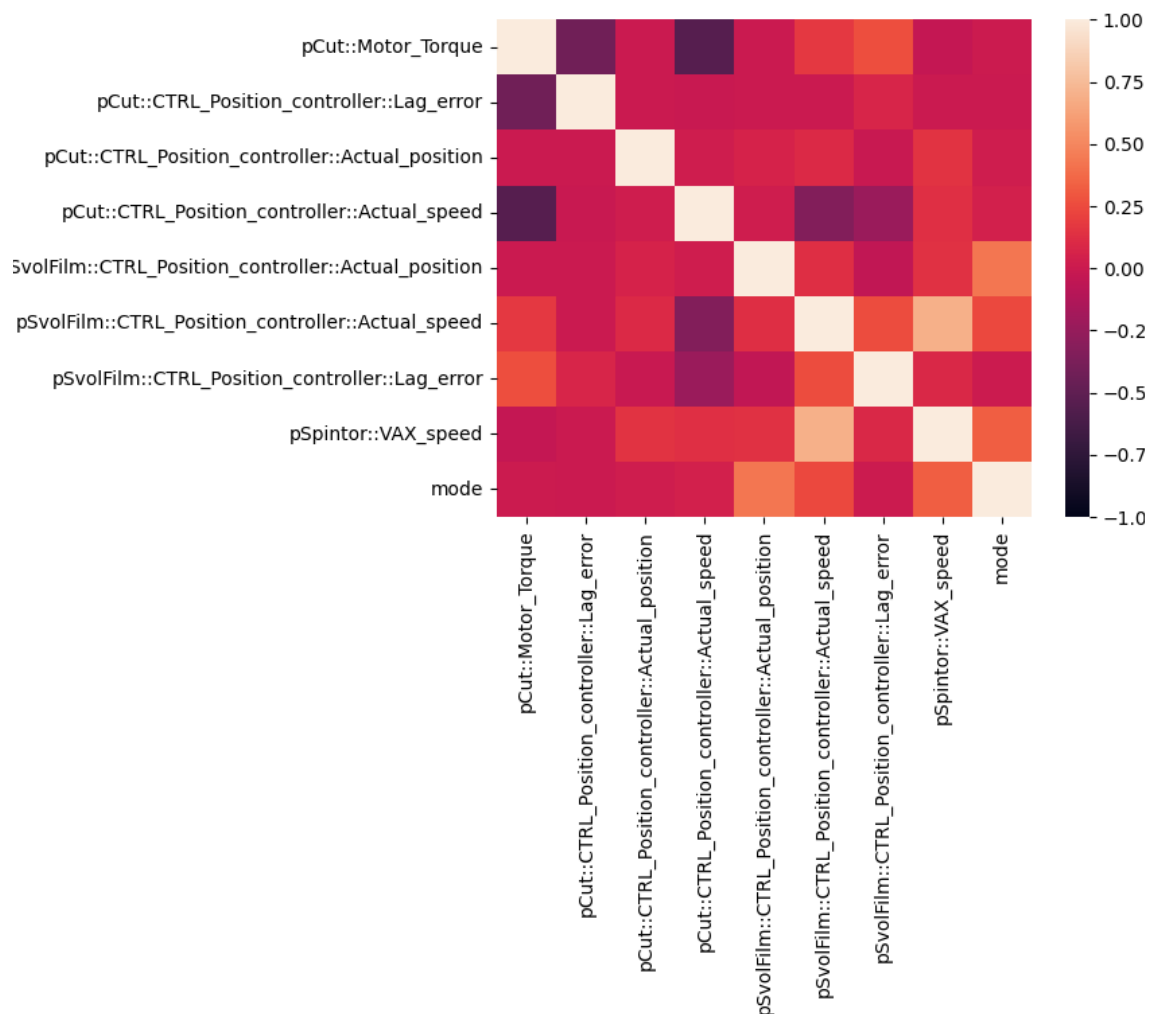
### 2.1. Leitura e Formatação dos Dados

Uma vez importadas as bibliotecas necessárias para análise dos dados, tais como matplotlib, pandas e scikit-learn, é realizada a leitura do dataset através de arquivo CSV. o dataset carregado é explorado, obtendo estatísticas descritivas e buscando dados faltantes. Depois, o método “convertNonNumeric” lida com colunas cujos valores não são do tipo Int ou Float. No dataset aplicado neste artigo, apenas a coluna “mode” atende a estes requisitos e a mesma não pode ser descartada pois o modo de funcionamento da máquina de corte influencia diretamente os dados, por isso é necessário substituir os dados da coluna com um identificador numérico para cada

modo. Após esta etapa outras colunas que não serão utilizadas são descartadas durante a limpeza do dataset.

## 2.2. Análise de Correlações

O próximo passo envolve a análise de correlações entre o restante das variáveis do conjunto de dados. Para isso, é utilizado um mapa de calor correlacional, gerado através da biblioteca seaborn. Isso é fundamental para entender a interdependência das variáveis e selecionar aquelas mais relevantes para o modelo de detecção de anomalias.



**Figura 1. Mapa de calor com interdependência das variáveis**

## 2.3. Padronização

O conjunto de dados é normalizado e padronizado utilizando Min-Max Scaling e Z-Score. Estas são etapas cruciais no pré-processamento dos dados, pois garantem que todas as variáveis do conjunto de dados estejam em uma escala uniforme, permitindo que o modelo faça previsões mais precisas. O Min-Max Scaling é uma técnica que redimensiona os valores de cada variável para um intervalo específico, como 1 e -1. Já o

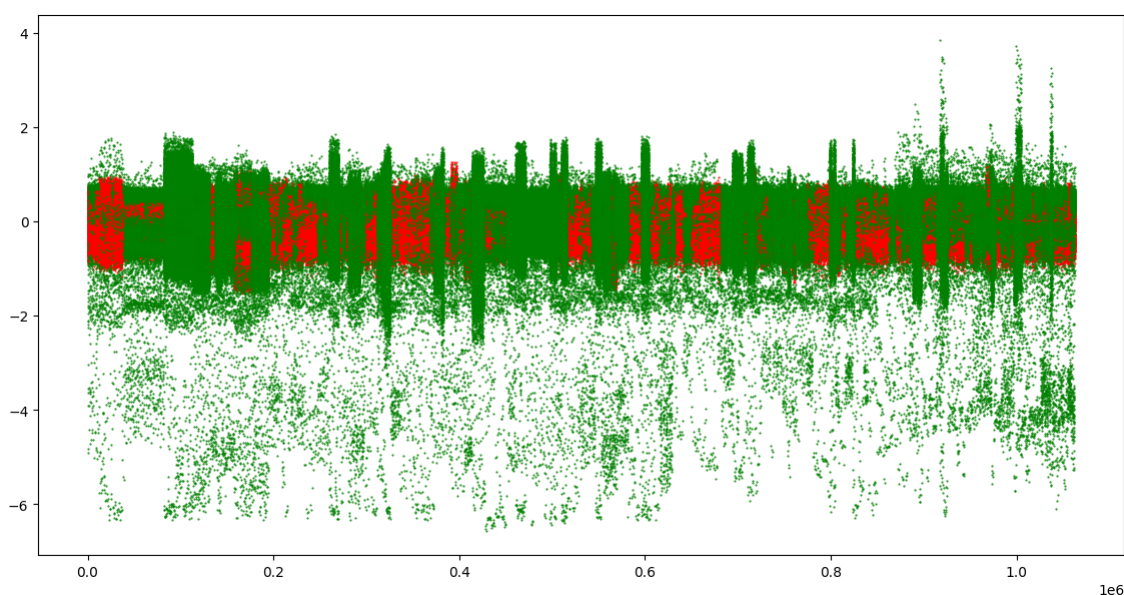
Z-Score Normalization padroniza os dados em uma escala comum com média zero e desvio padrão unitário.

## 2.4. Implementação do SVM

Com o dataset devidamente preparado, é realizada a implementação do Support Vector Machine, especificamente através do OneClassSVM. Uma porção do dataset é separada para compor o conjunto de treinamento, neste caso, a amostra foi de 200.000 registros. Para a realização do treinamento, o OneClassSVM foi configurado com “nu” igual a 0.25 e “gamma” igual a 0.05. O parâmetro "nu" é um valor que representa a fração de margem de erro permitida pelo modelo. Em outras palavras, nu controla a quantidade de pontos que o modelo considera como anomalias. O parâmetro "gamma" é um coeficiente que influencia a forma como o modelo trata a influência de cada ponto de dados. A escolha adequada de gamma é crucial para evitar overfitting ou underfitting do modelo.

## 2.5. Detecção de Anomalias

A detecção foi realizada prevendo as anomalias no conjunto de dados usando o modelo treinado. Uma cópia do dataset foi criada, recebendo uma nova coluna “anomaly” que armazena sua classificação, sendo o valor “1” para anomalia e “-1” para conforme. Ao plotar o gráfico da figura a seguir, as anomalias são representadas na cor vermelha e os dados conformes na cor verde.



**Figura 2. Gráfico de identificação de anomalias**

## 3. Aplicação

Na indústria de manufatura, a detecção de anomalias pode ser crucial para prever falhas em máquinas e equipamentos. A aplicação desta técnica impacta diretamente no KPI OEE (Índice de Eficiência Global de Equipamentos) da empresa. Identificando comportamentos anômalos em sensores e processos de produção, as empresas podem realizar manutenção preditiva, reduzindo o tempo de inatividade não planejado e

aumentando a disponibilidade dos equipamentos, o que consequentemente melhora o componente de disponibilidade no cálculo do OEE. Além disso, identificar anomalias em tempo real contribui para a melhoria da qualidade do produto, outro indicador do índice. Não menos importante, torna-se possível aumentar a eficiência operacional através da otimização de processos, melhorando o indicador de desempenho. Por fim, o impacto da redução de custos de manutenção acontece através da troca da abordagem reativa pela preditiva.

#### **4. Conclusão**

O desenvolvimento deste sistema de detecção de anomalias utilizando SVM oferece uma abordagem eficaz para identificar padrões não usuais em conjuntos de dados complexos. A compreensão das correlações entre variáveis e a escolha cuidadosa de parâmetros são essenciais para o sucesso da detecção de anomalias. Este código serve como ponto de partida para aplicações mais específicas e pode ser aprimorado com técnicas avançadas de machine learning.

#### **References**

- [1] Rana, Divya. "One Class SVM Vs SVM Classification." (2015).
- [2] Omar, Salima, Asri Ngadi, and Hamid H. Jebur. "Machine learning techniques for anomaly detection: an overview." *International Journal of Computer Applications* 79.2 (2013).