

# Comparação de Modelos de Deep Learning para Detecção de Anomalias em Poços de Petróleo

CPE 727 - Deep Learning

**Ana Clara Loureiro Cruz** ([anaclaralcruz@poli.ufrj.br](mailto:anaclaralcruz@poli.ufrj.br))

December 11, 2025

# Table of Contents

## 1 Introdução

- ▶ Introdução
- ▶ Revisão bibliográfica
- ▶ Metodologia
- ▶ Resultados
- ▶ Conclusão
- ▶ Referências

# Descrição do problema

## 1 Introdução

- Grandes volumes de dados de processo na indústria de óleo e gás (pressão, vazão, temperatura, etc.)
- Necessidade de monitoramento dos dados
- Dificuldade de identificar anomalias no processo

### Questão central

Modelos de *deep learning* podem melhorar a detecção precoce de anomalias em poços de petróleo?

# Motivação

## 1 Introdução

- Falhas em poços e linhas de produção podem resultar em:
  - perdas significativas de produção;
  - aumento de custos de manutenção corretiva;
  - riscos à segurança operacional e ao meio ambiente.
- Modelos de aprendizado de máquina clássicos podem não capturar adequadamente padrões temporais complexos e relações não lineares entre sensores.

# Objetivo

## 1 Introdução

- Selecionar e implementar 3 arquiteturas de *deep learning* adequadas à modelagem de séries temporais de poços.
- Comparar os resultados e propor melhorias para os sistemas de detecção de anomalias em séries temporais.

# Table of Contents

## 2 Revisão bibliográfica

- ▶ Introdução
- ▶ Revisão bibliográfica
- ▶ Metodologia
- ▶ Resultados
- ▶ Conclusão
- ▶ Referências

# Métodos propostos

## 2 Revisão bibliográfica

### RF com atributos estatísticos [1]

- Método: extração de estatísticas em janelas deslizantes das séries de poços, normalização e classificador *Random Forest* (RF) para detecção e classificação de falhas na base 3W [2].
- Resultado: *balanced accuracy* de aproximadamente 94,2% no classificador multiclasses.

### SWT + arquitetura modular [3]

- Método: proposta de um sistema modular e extração de atributos baseados em decomposição *stationary wavelet transform* (SWT), combinados a atributos estatísticos, para classificar eventos da base 3W.
- Resultado: *balanced accuracy* de 98,6% no problema multiclasses.

## ThreeWToolkit [4]

### 2 Revisão bibliográfica

- Pacote em Python 3 para experimentos com o 3W Dataset.
- Focado em detecção e classificação de eventos indesejáveis em poços offshore.
- Ferramentas:
  - Dataloaders para 3W Dataset;
  - Pre-processors for 3W Dataset;
  - Modelo MLP configurável;

# Table of Contents

## 3 Metodologia

- ▶ Introdução
- ▶ Revisão bibliográfica
- ▶ Metodologia
- ▶ Resultados
- ▶ Conclusão
- ▶ Referências

# Descrição do Dataset

## 3 Metodologia

- Classificação de séries temporais multivariadas
- Séries temporais anotadas divididas em 10 classes

No escopo deste trabalho: subset de 3 classes Janelamento de passo 1 e tamanho 100

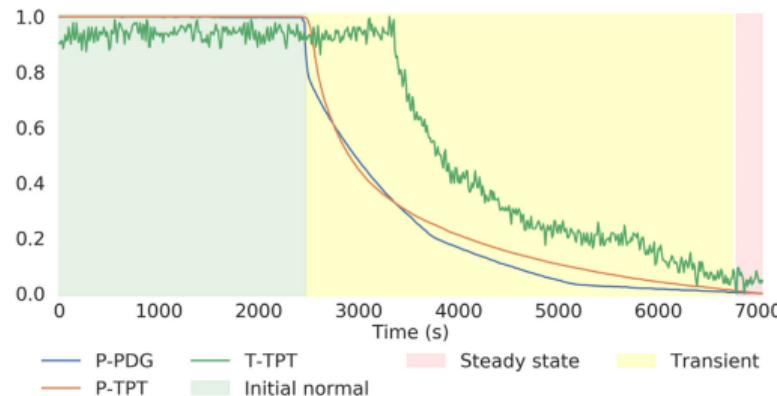


Figure: Exemplo de série temporal anotada (Imagen retirada de [1]).

# Estratégia de comparação

## 3 Metodologia

### Modelos escolhidos

- Multi-layer Perceptron (MLP) implementado em [4] (baseline)
- MLP com regularização (dropout)
- Convolutional Neural Network (CNN) [5]
- Temporal Convolutional Network (TCN) [6]

### Métricas de avaliação

- Acurácia
- Precisão
- Recall
- F1-Score
- Tempo de execução

# Padronização das pipelines

## 3 Metodologia

### Hiperparâmetros de treinamento [4]

- Otimizador: Adam
- Função de perda: Cross-entropy
- Taxa de aprendizado: 0,0005
- Batch size: 64
- Cross-validation: ativado
- 70% treino 15% teste 15% validação

# Padronização das pipelines

## 3 Metodologia

### Pré-processamento [4]

- Imputação de dados faltantes:
  - Estratégia: média (*mean*)
  - Aplicado a todas as colunas
- Normalização:
  - Tipo: norma L<sub>2</sub>
- Janela temporal:
  - Tamanho da janela: 100 amostras

# Hiperparâmetros dos Modelos

## 3 Metodologia

### MLP (com e sem regularização)

Dados de entrada: Extração de dados estatísticos da janela temporal

- Tipo: MLP totalmente conectado
- Camadas escondidas: (128, 64, 32)
- Tamanho da saída: 3 classes (0, 1, 2)
- Função de ativação: ReLU
- Regularização: L2

# Hiperparâmetros dos Modelos

## 3 Metodologia

### CNN 1D e TCN

- Número de canais de entrada: 1
- Filtros por camada: (32, 64, 128)
- Tamanho do kernel: 3
- Função de ativação: ReLU
- Taxa de dropout: 0.1

# Table of Contents

## 4 Resultados

- ▶ Introdução
- ▶ Revisão bibliográfica
- ▶ Metodologia
- ▶ Resultados
- ▶ Conclusão
- ▶ Referências

# Comparação dos modelos

## 4 Resultados

Modelo	Acurácia balanceada	Precisão	Recall	F1	Tempo (s)
MLP puro	0.9395	0.9743	0.9744	0.9742	1479
MLP com regularização	0.9311	0.9743	0.9744	0.9742	?
CNN	0.8824	0.9539	0.9541	0.9538	1654
TCN	0.9147	0.9884	0.9472	0.9474	605

# Exemplos de evolução da função-custo

## 4 Resultados

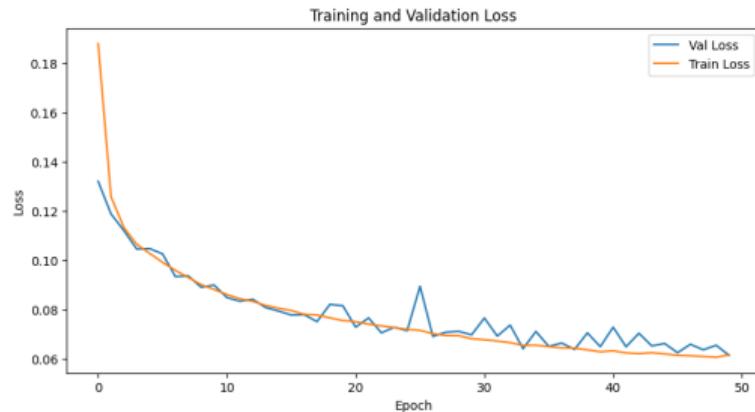


Figure: Exemplo de evolução da função-custo no MLP puro

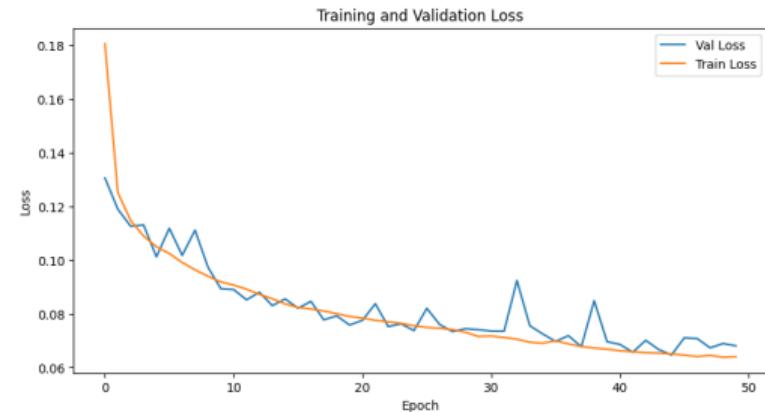


Figure: Exemplo de evolução da função-custo no MLP com dropout

# Exemplos de evolução da função-custo

## 4 Resultados

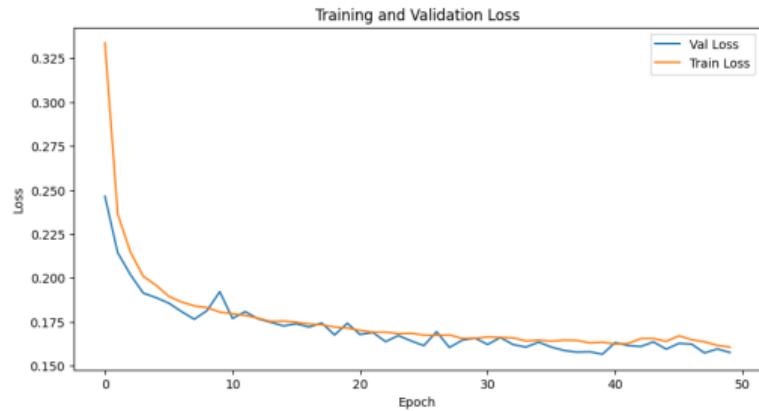


Figure: Exemplo de evolução da função-custo na CNN

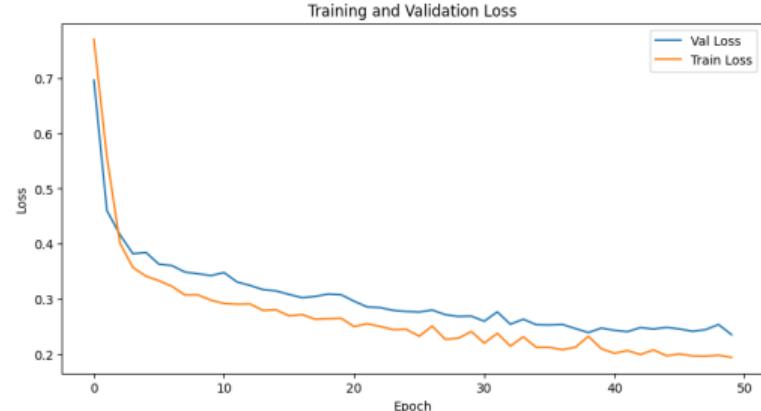


Figure: Exemplo de evolução da função-custo na TCN

# Table of Contents

## 5 Conclusão

- ▶ Introdução
- ▶ Revisão bibliográfica
- ▶ Metodologia
- ▶ Resultados
- ▶ Conclusão
- ▶ Referências

# Discussão dos resultados

## 5 Conclusão

- Os MLP baseado em características estatísticas não foi superado pelos modelos convolucionais
- A CNN talvez se beneficiaria de mais poder computacional
- A TCN mostrou bons resultados no trade-off entre desempenho e custo
- Aumentar o subset para obter melhores resultados

# Table of Contents

## 6 Referências

- ▶ Introdução
- ▶ Revisão bibliográfica
- ▶ Metodologia
- ▶ Resultados
- ▶ Conclusão
- ▶ Referências

# Bibliography

6 Referências

- [1] M. A. Marins, B. D. Barros, I. H. Santos, D. C. Barrionuevo, R. E. V. Vargas, T. d. M. Prego, A. A. de Lima, M. L. R. de Campos, E. A. B. da Silva, and S. L. Netto, "Fault detection and classification in oil wells and production/service lines using random forest," *Journal of Petroleum Science and Engineering*, vol. 197, p. 107879, 2021.
- [2] R. E. V. Vargas, C. J. Munaro, P. M. Ciarelli, M. A. Marins, B. D. Barros, I. H. Santos, M. L. R. de Campos, E. A. B. da Silva, and S. L. Netto, "A realistic and public dataset with rare undesirable real events in oil wells," *Journal of Petroleum Science and Engineering*, vol. 181, p. 106223, 2019.
- [3] T. L. B. Dias, M. A. Marins, C. L. Pagliari, R. M. E. Barbosa, M. L. R. de Campos, E. A. B. da Silva, and S. L. Netto, "Development of oilwell fault classifiers using a wavelet-based multivariable approach in a modular architecture," *SPE Journal*, 2024.

# Bibliography

6 Referências

- [4] U. Petrobras, "Threew toolkit." <https://github.com/petrobras/3W>. Acessado em: 10 dez. 2025.
- [5] K. O'shea and R. Nash, "An introduction to convolutional neural networks," *arXiv preprint arXiv:1511.08458*, 2015.
- [6] C. Lea, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks: A unified approach to action segmentation," in *Computer Vision – ECCV 2016 Workshops* (G. Hua and H. Jégou, eds.), (Cham), pp. 47–54, Springer International Publishing, 2016.

# Comparação de Modelos de Deep Learning para Detecção de Anomalias em Poços de Petróleo

*Obrigada!  
Perguntas?*