

Previsão de Séries Temporais na Presença de Mudança de Conceito

Gustavo H. Ferreira de Miranda Oliveira
gustavo.oliveira@penedo.ufal.br



Centro de Informática – UFPE
24/04/2025



UFAL
CAMPUS ARAPIRACA
U. E. PENEDO

Apresentação



Professor Gustavo Miranda:

- **Doutor** em Inteligência Computacional - **UFPE**
- **Mestre** em Inteligência Computacional - **UFPE**
- **Graduado** em Licenciatura em Computação - **UPE**
- **Experiência** nos seguintes campos:
 - Machine Learning
 - Ciência de Dados

Informações Adicionais

- <https://github.com/GustavoHFMO>



Tópicos

- Previsão de Séries Temporais
- Mudança de Conceito e Definições
- Detecção de Pontos de Mudança
- Revisão da Literatura
- Trabalhos do CIn na Área
- Temas de Pesquisa

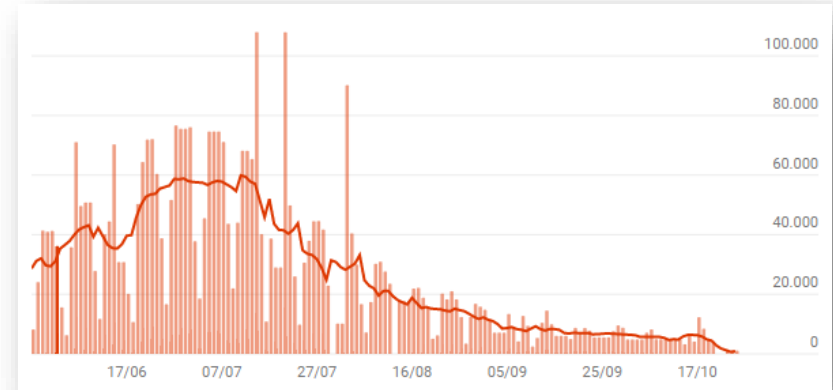
Séries Temporais



Séries Temporais

Uma série temporal é uma **coleção de observações** medidas **sequencialmente** ao longo do tempo:

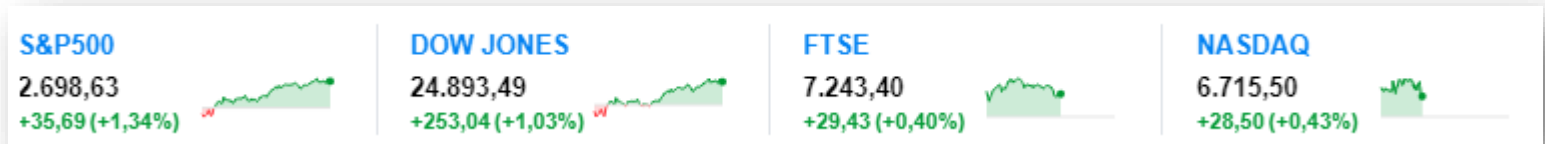
- Casos confirmados de COVID
- Mortes por COVID
- Movimento do preço de ações;
- Taxas de câmbio;
- Temperaturas de uma cidade;
- Entre outros.



Fonte: Dados OMS Brasil 2022

Previsão de Séries Temporais

- A maioria das abordagens de séries temporais **modelam os dados de forma estática**.
 - Processam a série temporal de **modo off-line**.
 - Processamento dos **dados históricos** em várias etapas.
- No entanto, na maioria das **aplicações de séries temporais do mundo real**, os dados chegam sequencialmente em um **fluxo de dados**.
 - Os dados podem fluir em alta velocidade e evoluir ao longo do tempo.



Representação da chegada de novos dados em séries do mundo real. **Fonte:** Yahoo Finanças.

Séries Temporais e Fluxos de Dados

- **Fluxos de Dados** (em inglês Data Streams) são caracterizados por **uma grande quantidade de instâncias** que chegam rapidamente e necessitam de uma **resposta em tempo real**. (Dongre et al, 2014).
- Com o **crescimento das tecnologias da internet** a quantidade de dados tornou-se **cada vez maior**. (Diztiler et al, 2015).



Desafios dos Fluxos de Dados

- Exemplos de aplicações:

- Análise de mercado;
- Filtragem de informações;
- Controle de processos industriais; e etc. (Minku et al, 2012)



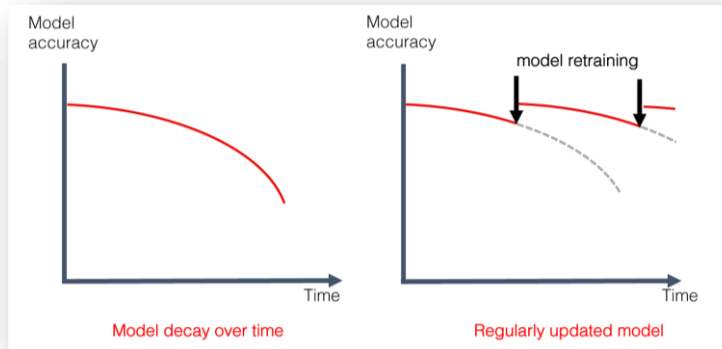
- Fluxos de Dados **representam novos desafios** para o campo de **aprendizagem de máquina** e mineração de dados, já que os métodos tradicionais foram projetados para conjuntos de dados estáticos. (Krawczyk et al, 2017)
- Um desses desafios é a **mudança de conceito** (em inglês Concept Drift). (Krawczyk et al, 2017)

Mudança de Conceito (Concept Drift)



Mudança de conceito

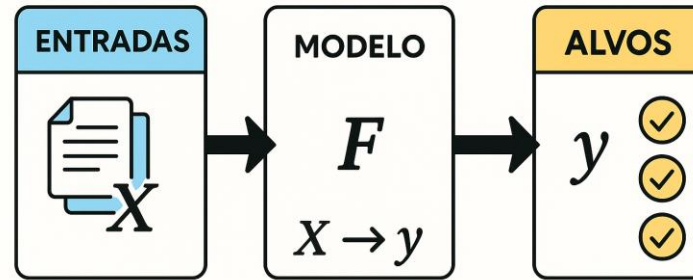
- Consiste em uma mudança na relação entre os dados de entrada e a variável alvo ao longo do tempo. (Cavalcante et al, 2015).
- Um **exemplo típico** de uma mudança de conceito real é uma **mudança nos interesses do usuário** ao seguir um fluxo de notícias on-line. (Gama et al, 2014).
- A evolução dos dados **impõe um grande desafio** aos algoritmos tradicionais de aprendizagem por lotes, já que o modelo aprendido a partir dos dados **pode tornar-se obsoleto** (Cavalcante et al, 2015).



Aprendizado de Modelos Supervisionados

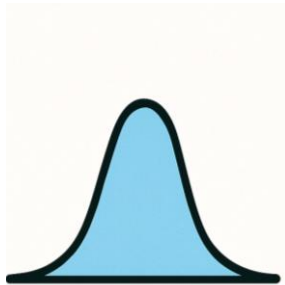
- Conjunto de entradas: X
- Conjunto de alvos: y
- Treinamento: $F: X \rightarrow y$

MODELO DE APRENDIZAGEM DE MÁQUINA SUPERVISIONADA



Processo Gerador dos Dados

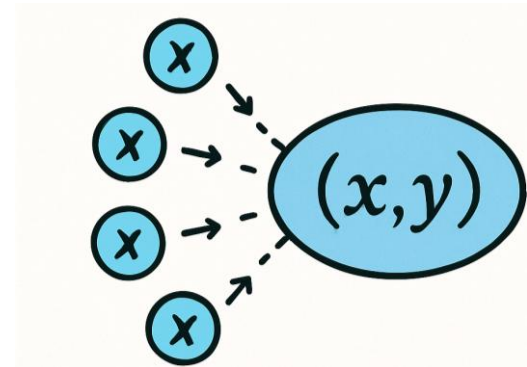
- Cada observação (x, y) vem de um fluxo dados gerado por uma distribuição de probabilidade $P(X,y)$.



Distribuição de probabilidade $P(X,y)$



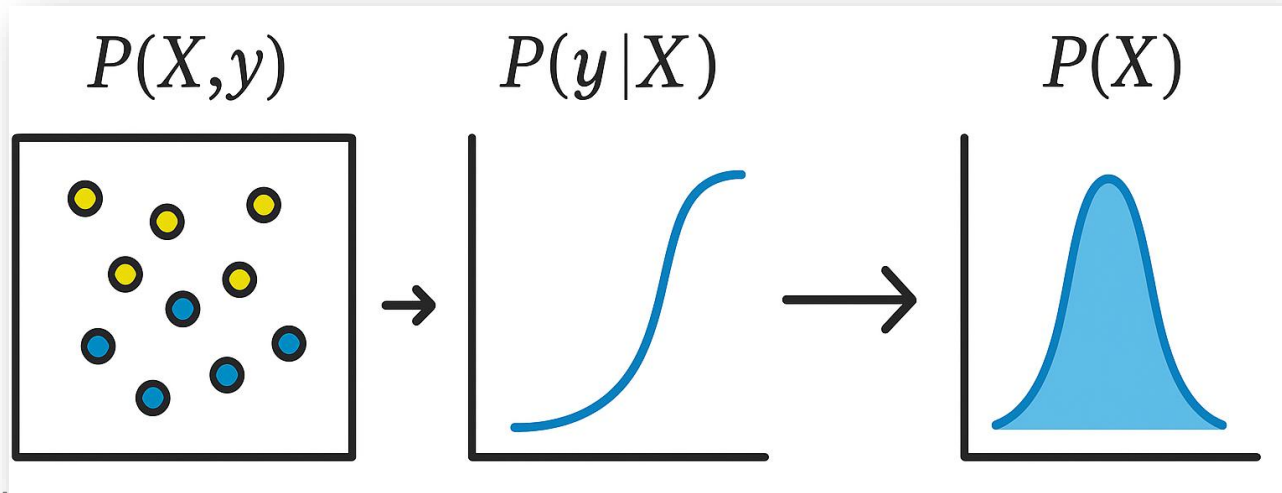
Observação (x, y)



fluxo de dados gerado por $P(X,y)$

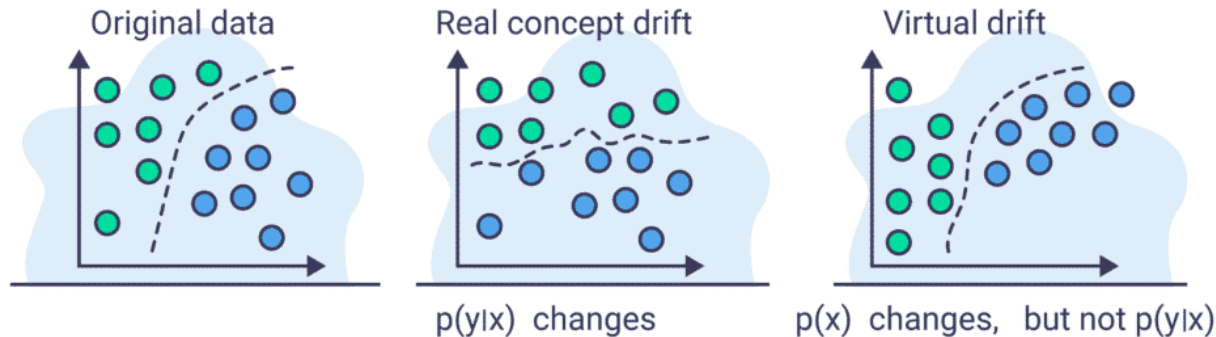
Distribuição de Probabilidade Conjunta

- Distribuição de probabilidade conjunta é $P(X,y) = P(y|X) P(X)$.
 - $P(y|X)$ é a função que relaciona o conjunto de entradas X com o conjunto de saídas y .
 - $P(X)$ é a distribuição dos componentes de entrada X .



Mudança de Conceito em Classificação

- **Mudança real** se refere a uma mudança em $P(y|x)$ que pode ser causada por uma **mudança na fronteira de decisão do problema**.
- **Mudança virtual** se refere a uma mudança em $P(x)$ em que a **distribuição das observações de chegada mudam**.

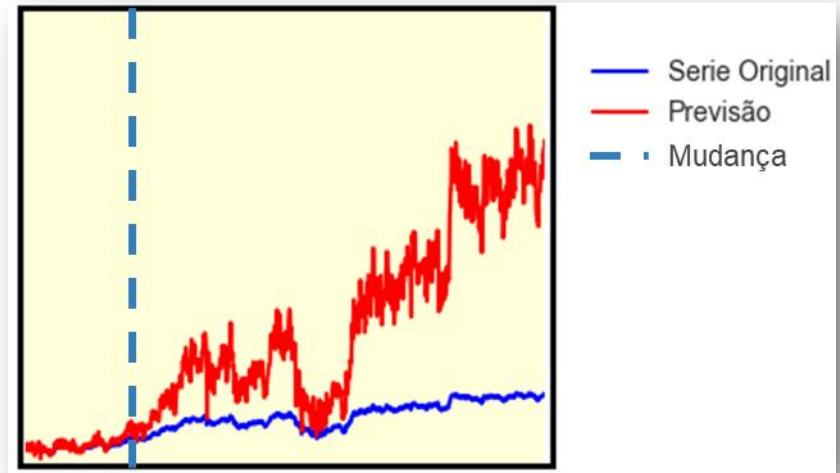


Mudança de Conceito Em Séries Temporais



Mudança de Conceito em Séries Temporais

- O processo de **geração dos dados** pode **mudar ao longo do tempo** (Cavalcante et al, 2015);
- **Observações históricas não são úteis** para definir comportamentos futuros (Oliveira et al, 2017).
 - **Prejudicam fortemente as previsões.**



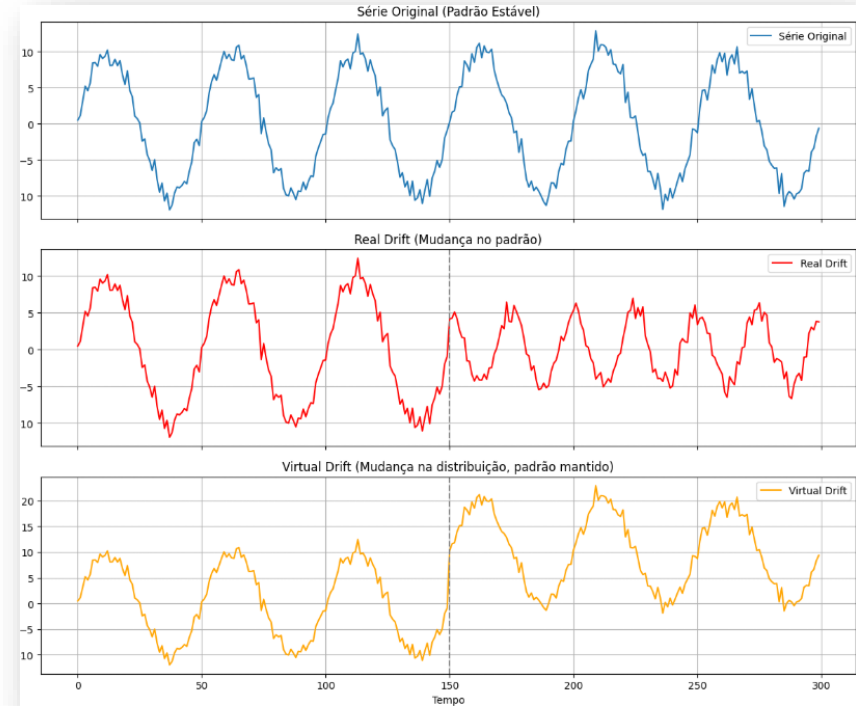
Tipos de Mudanças de Conceito em Séries Temporais

Real Drift (Mudança Real):

- Ocorre quando a **relação** entre as entradas e a saída muda ao longo do tempo, ou seja, a **função F** muda.

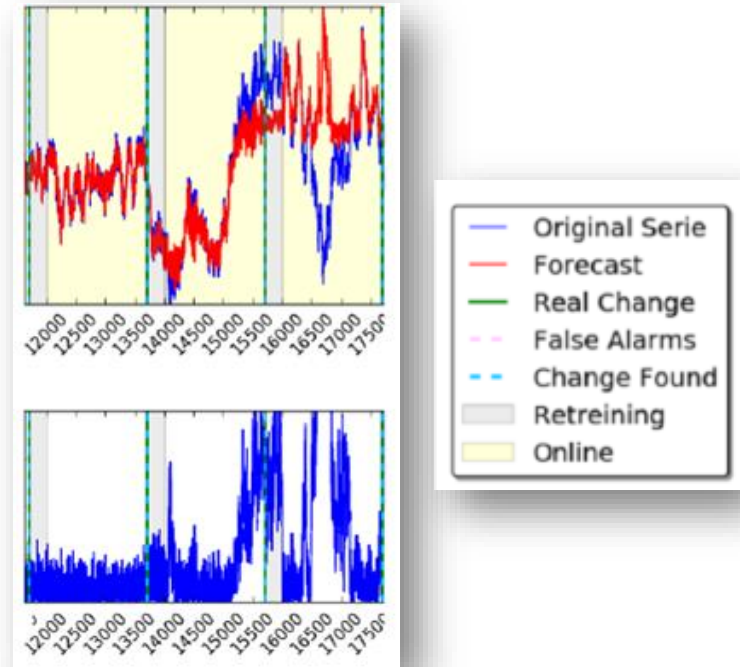
Virtual Drift (Mudança Virtual):

- Ocorre quando a **distribuição** das entradas muda, mas a **relação com a saída** continua a mesma.
- A **função F** é a mesma, mas o **espaço** onde ela atua mudou um pouco.



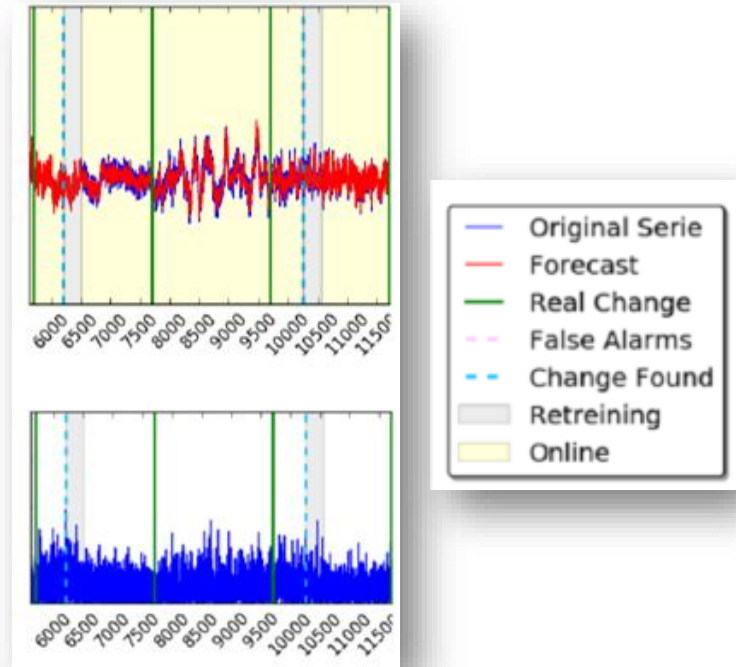
Exemplos de Mudanças Reais

- Uma mudança no comportamento do consumidor: antes a demanda dependia da temperatura, agora depende de promoções.
- Após a pandemia, o padrão de mobilidade urbana muda completamente — o mesmo histórico já não serve para prever o futuro.
- Mudança abrupta de política econômica que muda totalmente o padrão dos dados.



Exemplos de Mudanças Virtuais

- **A média ou variância da variável muda** (ex: dados de entrada estão mais altos, mas a regra de previsão ainda vale).
- **Uma mudança no ciclo de consumo** (ex: o padrão semanal se mantém, mas os valores estão todos com outro nível).
- **Uma estação do ano nova chega** — os dados mudam, mas os padrões (tipo tendência ou sazonalidade) ainda funcionam.



Tipos de Mudanças



Tipos de Mudanças

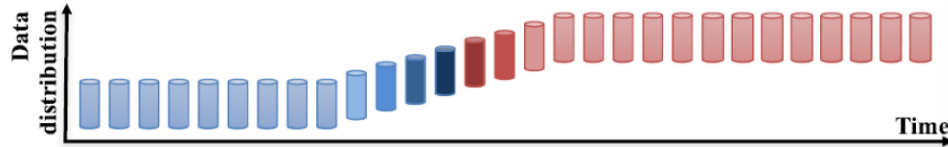


Figure 2.4: Example of incremental drift (LU et al., 2018).

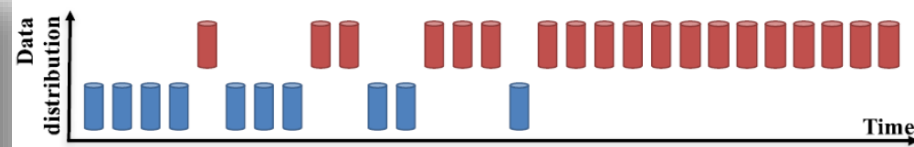


Figure 2.5: Example of gradual drift (LU et al., 2018).

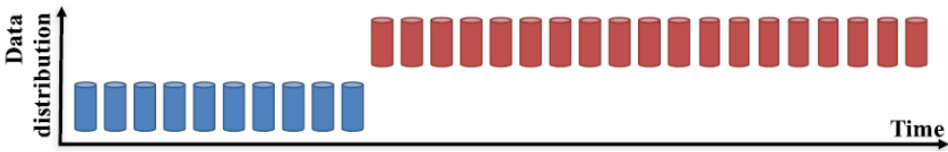


Figure 2.6: Example of abrupt drift (LU et al., 2018).

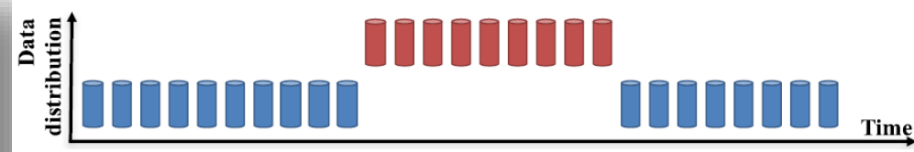


Figure 2.7: Example of recurrent drift (LU et al., 2018).

Severidade das Mudanças



Severidade das Mudanças

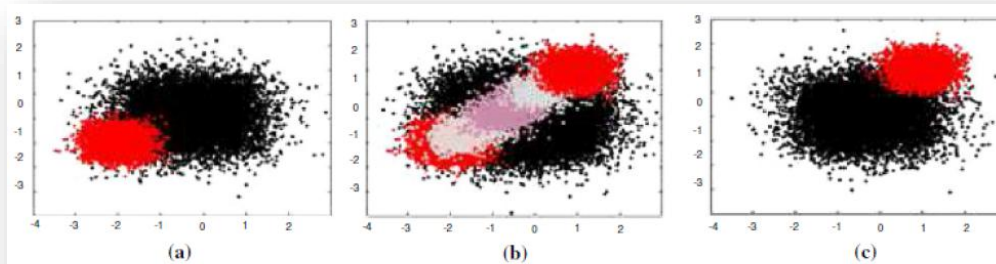


Figure 2.8: Example of severe drift (KHAMASSI et al., 2018). Fig a) shows two classes, one in black and one in red. Fig b) shows the red class evolving to another region of the feature space, where the gray part shows the path taken. Fig c) shows the complete evolution of the red class.

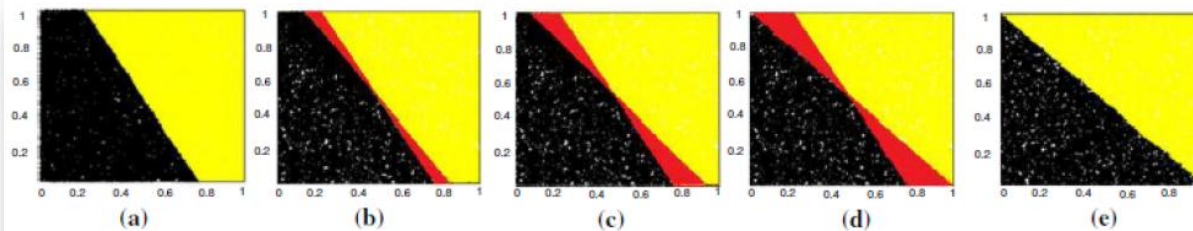


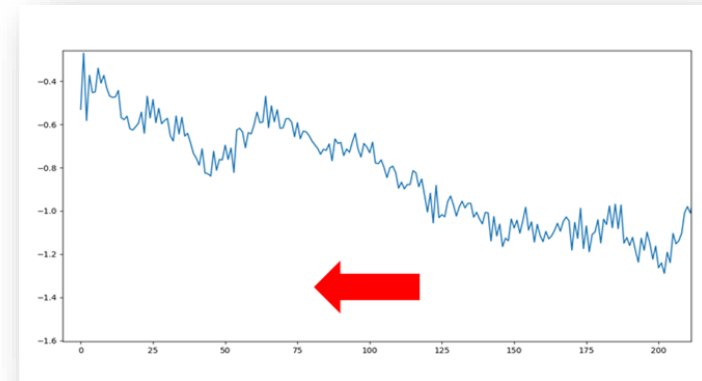
Figure 2.9: Example of non-severe drift (KHAMASSI et al., 2018). Fig a) presents two classes, one represented in black and the other in yellow. Fig b), c) and d), show in red the region of the feature space that had its class changed. Fig e) shows the complete drift.

Deteccção de Pontos de Mudança



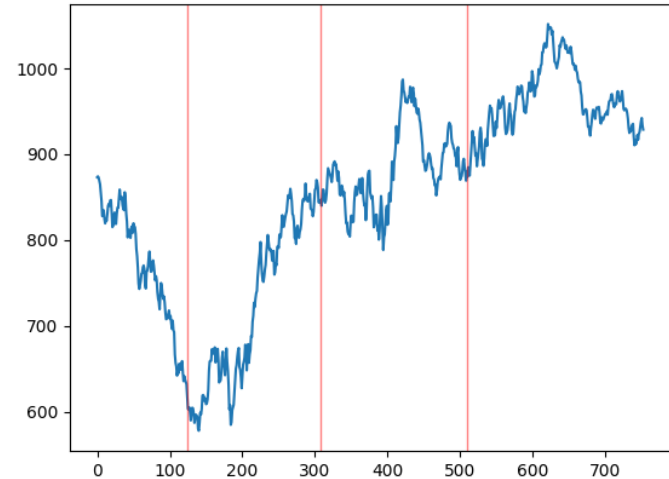
Detecção de Pontos de Mudanças em Séries Temporais

- São métodos que têm como objetivo principal a **detecção de pontos de mudança**, que são **instantes de tempo** quando ocorreram **mudanças de conceito**. (Cavalcante et al, 2016).
- A detecção de pontos de mudança é **útil na modelagem e previsão de séries temporais** e também em outras áreas de aplicação como **finanças, biometria e robótica**. (Adams et al, 2007).
- Essas abordagens são baseadas em uma **análise estatística retrospectiva** dos dados da série temporal ou dos resíduos para identificação de pontos de mudança. (Cavalcante et al, 2016).



Exemplo Real de Detecção de Pontos de Mudança – Política

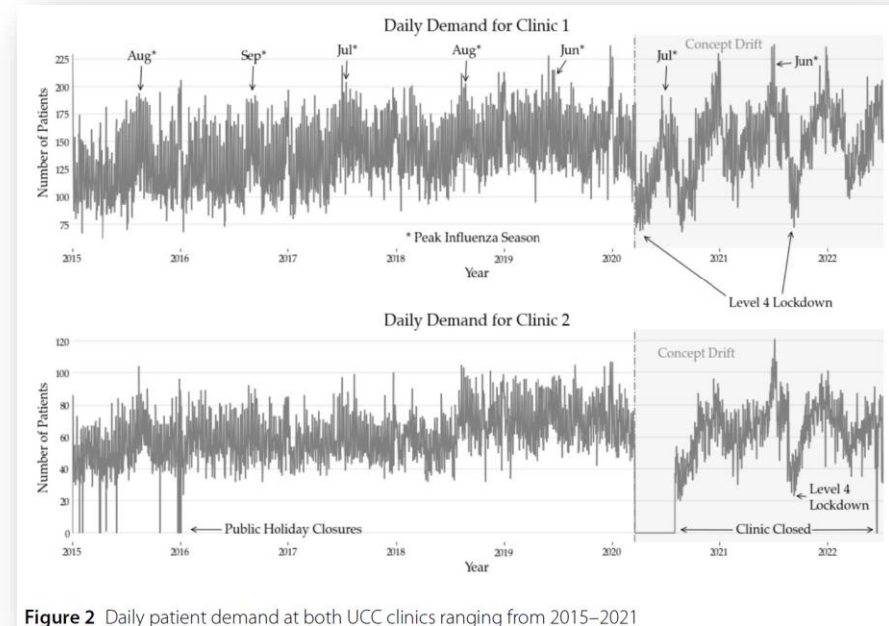
- 124: Condenação dos assessores do ex-presidente Nixon, G. Gordon Liddy e James W. McCord, Jr. Em 30 de Janeiro de 1973.
- 307: Inicio do embargo OPEC contra os Estados Unidos em 19 de outubro de 1973.
- 510: A renúncia do ex-presidente Nixon em 9 de agosto de 1974.



Dow Jones – Jan de 1973 à Jun de 1975.

Exemplo Real de Detecção de Pontos de Mudança - COVID

- Opontos exatos correspondentes às **determinações mais rigorosas referentes aos lockdowns da pandemia de COVID-19 são destacados**, bem como os fechamentos parciais periódicos da clínica menor durante partes desse período.



Susnjak et al. (2023)

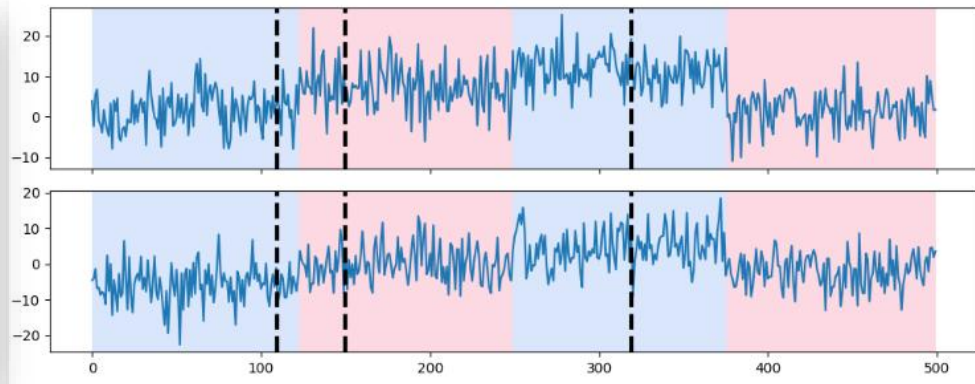
PELT – Change Point Detection

<https://centre-borelli.github.io/ruptures-docs/>

Welcome to ruptures

Maintained? **yes** build **failing** python **3.6 | 3.7 | 3.8 | 3.9** conda-forge **v1.1.9** docs **passing**
license **BSD-2-Clause** downloads **21M** code style **black** launch **binder** codecov **99%**

`ruptures` is a Python library for off-line change point detection. This package provides methods for the analysis and segmentation of non-stationary signals. Implemented algorithms include exact and approximate detection for various parametric and non-parametric models. `ruptures` focuses on ease of use by providing a well-documented and consistent interface. In addition, thanks to its modular structure, different algorithms and models can be connected and extended within this package.



Revisão da Literatura



Área de Pesquisa em Evidência

- A maior parte das pesquisas se concentra em tarefas de classificação. (Lima et al, 2022).

Received March 14, 2022, accepted April 17, 2022, date of publication April 22, 2022, date of current version May 3, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3169785

Learning Under Concept Drift for Regression—A Systematic Literature Review

MARÍLIA LIMA¹, MANOEL NETO¹, TELMO SILVA FILHO²,
AND ROBERTA A. DE A. FAGUNDES¹, (Member, IEEE)

¹Department of Computer Engineering, University of Pernambuco, Recife, Pernambuco 50720-001, Brazil

²Department of Statistics, Federal University of Paraíba, João Pessoa, Paraíba 58051-900, Brazil

Corresponding authors: Marília Lima (e-mail: mncal@ecomp.poli.br), Manoel Neto (e-mail: maan@ecomp.poli.br),
Telmo Silva Filho (e-mail: telmo@de.ufpb.br), and Roberta A. de A. Fagundes (e-mail: roberta.fagundes@upe.br)

This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, under Grant 001.

FIGURE 4. Number of articles selected in each step.

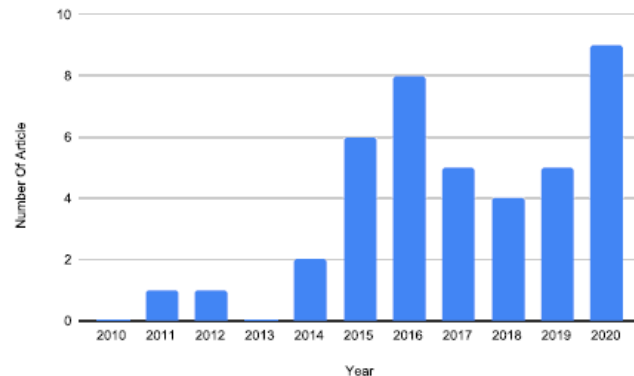


FIGURE 5. Number of articles per year (2010-2020).

Abordagens mais Investigadas

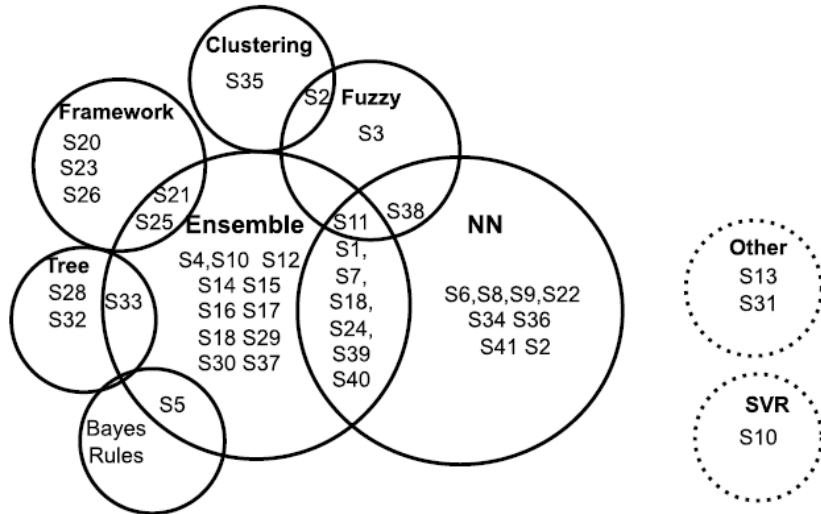


FIGURE 7. Distribution of studies by category.

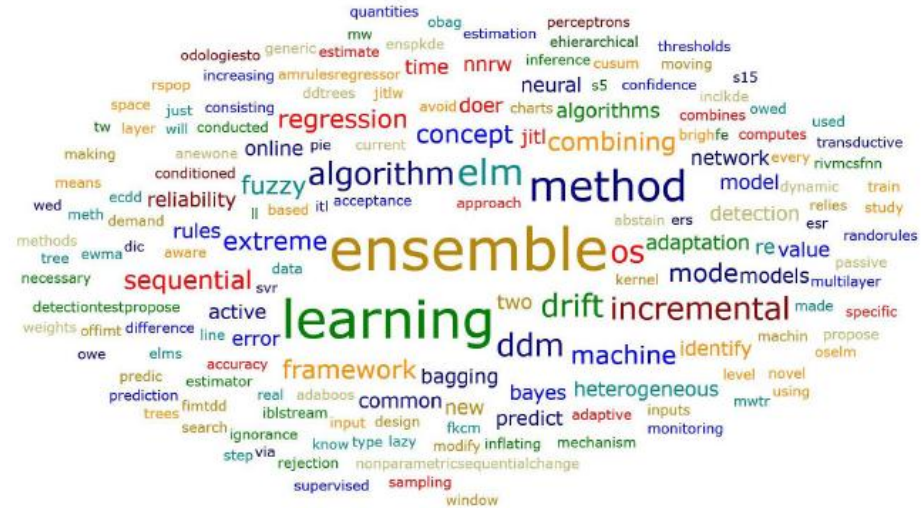
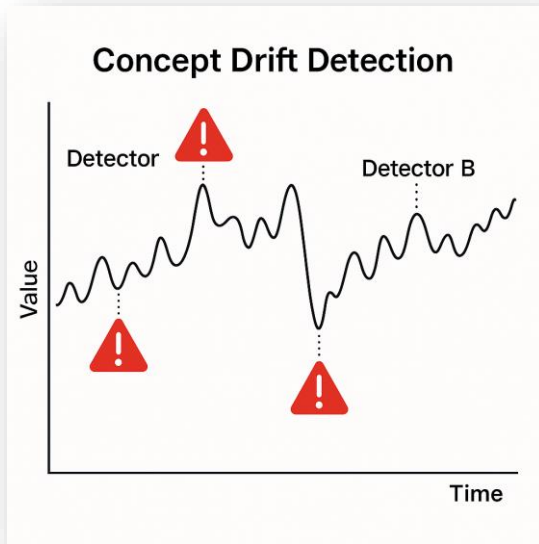


FIGURE 8. Word Cloud most used in study responses.

Aprendendo Sobre Mudança de Conceito



Aprendendo Sobre Mudança de Conceito

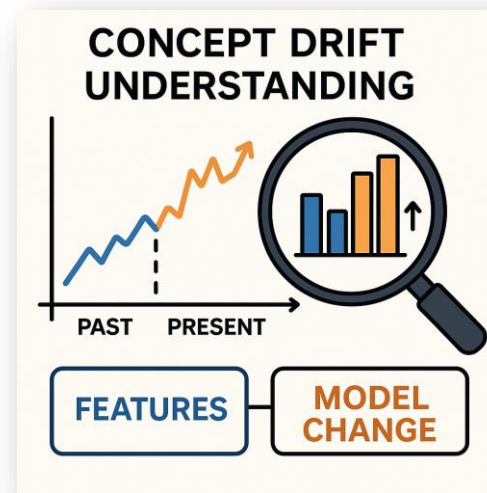
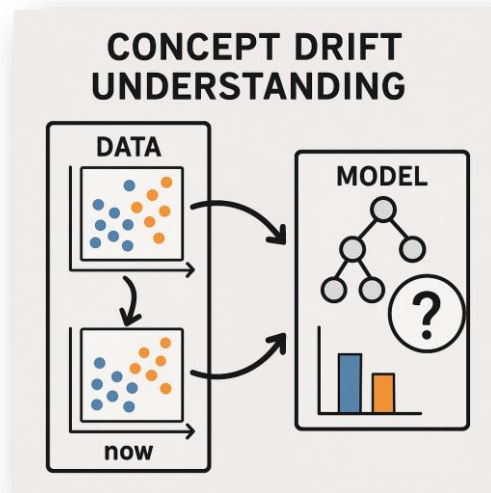


- **Detecção de Mudança de Conceito:**
 - Utiliza detectores para **identificar o momento em que a mudança ocorre** em tempo real.
 - Detectores podem ser baseados (i) no **erro dos modelos** ou (ii) **diretamente nos dados** do problema para

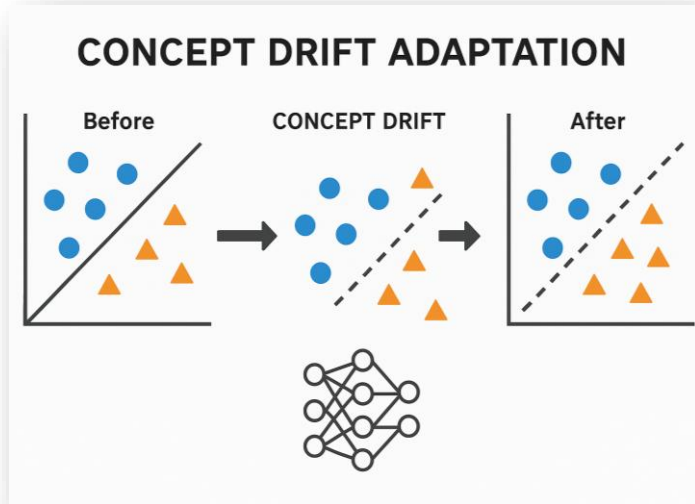
Aprendendo Sobre Mudança de Conceito

○ Compreensão da Mudança de Conceito

- Busca compreender qual foi a parcela dos dados que teve alteração para que a adaptação seja seletiva, removendo apenas as informações inúteis.



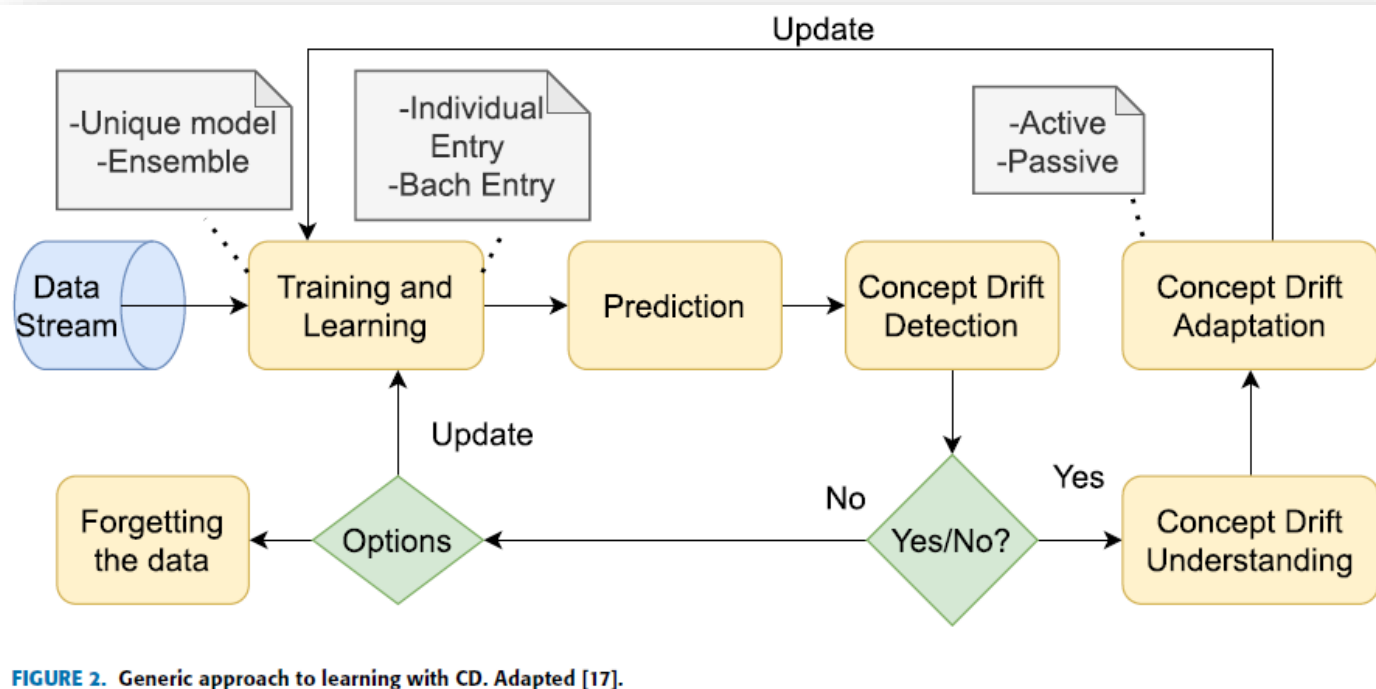
Aprendendo Sobre Mudança de Conceito



Adaptação a Mudança de Conceito

- Desenvolve **estratégias especializadas** por tipo de mudança que envolvem o **treinamento e atualização dos modelos** existentes.
- Envolve **abordagens** (i) **ativa**: atualização só após detecção de desvio e (ii) **passiva**: atualização periódica, com ou sem desvio real.

Framework Genérico de adaptação a Mudança de Conceito



Trabalhos do Cin na Área



Trabalhos Acadêmicos no CIn

1. Cavalcante, Rodolfo Carneiro. **"An adaptive learning system for time series forecasting in the presence of concept drift."** (2017). Tese de Doutorado. Universidade Federal de Pernambuco.
2. OLIVEIRA, Gustavo Henrique Ferreira de Miranda et al. **Previsão de Séries Temporais na Presença de Mudança de Conceito: uma Abordagem baseada em PSO.** 2018. Dissertação de Mestrado. Universidade Federal de Pernambuco.
3. SCHETTINI, Gabriela Mota de Lacerda Padilha et al. **Um Método Baseado em Correção de Erros Para Previsão de Séries Temporais em Ambiente Online e Na Presença De Concept Drift.** 2020. Dissertação de Mestrado. Universidade Federal de Pernambuco.
4. LEAL, Denisson Augusto Bastos et al. **Ensembles Dinâmicos para Detecção de Concept Drift em Séries Temporais.** 2022. Dissertação de Mestrado. Universidade Federal de Pernambuco.

Trabalhos Publicados

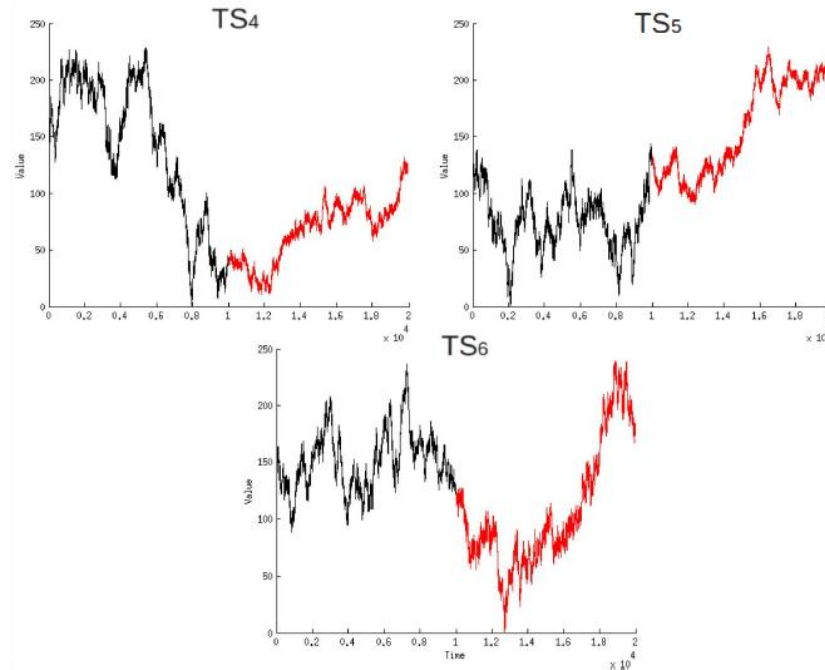
1. Cavalcante, Rodolfo C., and Adriano LI Oliveira. **"An approach to handle concept drift in financial time series based on Extreme Learning Machines and explicit Drift Detection."** *2015 international joint conference on neural networks (IJCNN)*. IEEE, 2015.
2. Cavalcante, Rodolfo C., Leandro L. Minku, and Adriano LI Oliveira. **"FEDD: Feature Extraction for Explicit Concept Drift Detection in time series."** *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016.
3. Oliveira, Gustavo H. F de M., Cavalcante, Rodolfo C., Cabral, George G., Leandro L. Minku, and Adriano LI Oliveira. **"Time Series Forecasting in the Presence of Concept Drift: A PSO-based Approach."** *Tools with Artificial Intelligence (ICTAI), 2017 IEE 29th International Conference on*. IEE, 2017.

1- Cavalcante et al (2015) - IJCNN

- **Objetivo:** Propõem uso de **mecanismos explícitos de detecção** de mudança de conceito a fim de **melhorar a previsão de séries temporais financeiras**.
- **Métodos Cegos ou Passivos:** OS-ELM.
○ **Métodos Ativos:** DDM e ECDD.
- **Hipóteses de pesquisa:**
 - Se usar **ELM em combinação com um mecanismo de detecção** de mudança para atualizar o modelo aprendido quando a mudança ocorrer, **ele pode obter precisão similar ao OS-ELM**, mas **com maior eficiência**.
 - Semelhante a primeira pergunta, **se atualizar OS-ELM apenas quando a mudança ocorrer**, apresentaria **melhor eficiência** enquanto mantém uma **boa precisão**.
- **Contribuições:** Uma abordagem que combina **OS-ELM com detecção de mudança de conceito**. O modelo **atualiza apenas na presença de mudança** de conceito. A abordagem proposta apresenta o **mesmo desempenho de um OS-ELM convencional** porém com **menos esforço** computacional.

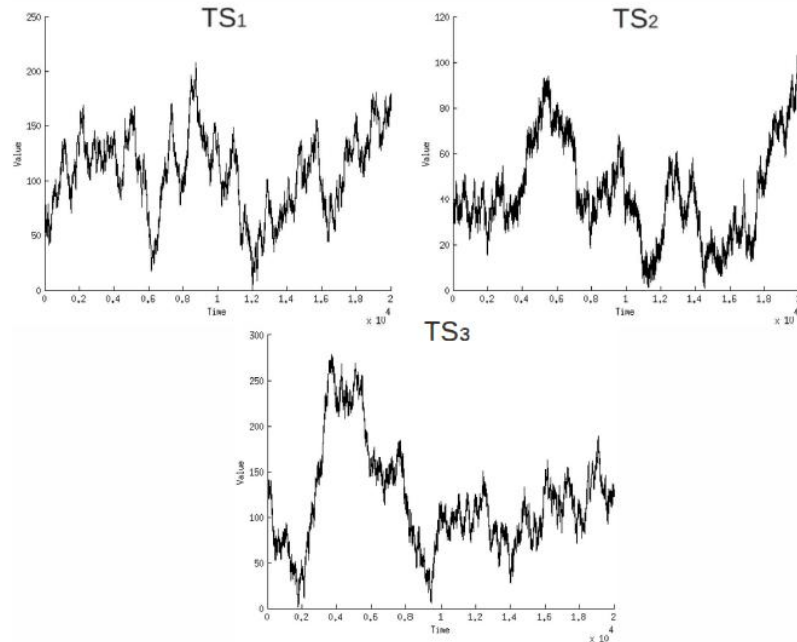
1- Cavalcante et al (2015) - IJCNN

○ Séries artificiais:



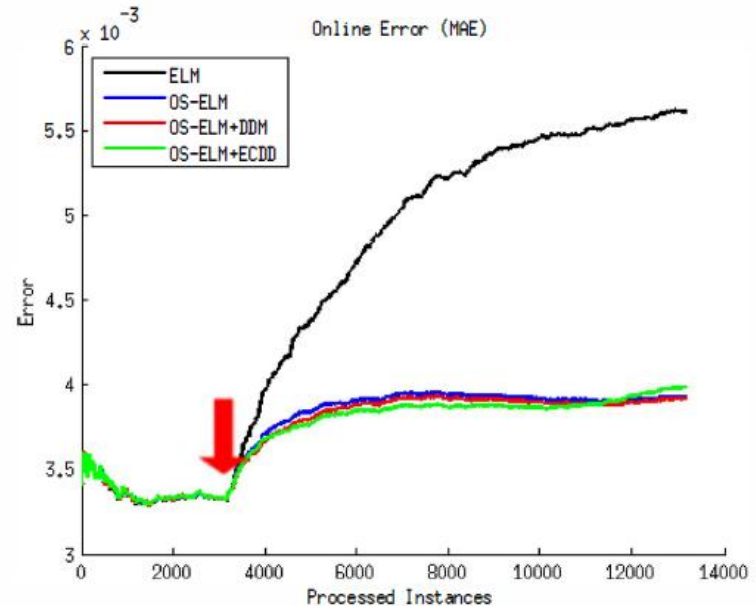
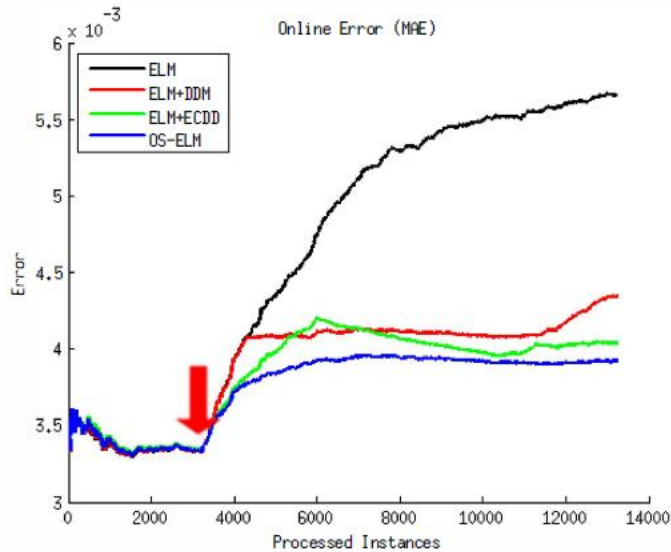
1- Cavalcante et al (2015) - IJCNN

○ Séries financeiras:



1- Cavalcante et al (2015) - IJCNN

Principais Resultados:

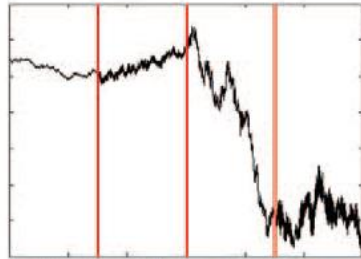


2- Cavalcante et al (2016) - IJCNN

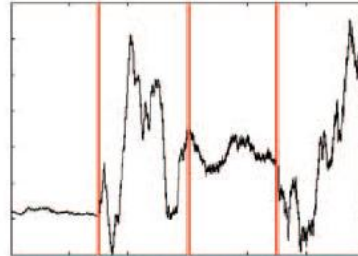
- **Objetivo:** Propõem um **método explícito de detecção** de mudança de conceito que identifica mudanças **monitorando características da própria série temporal**.
- Métodos de detecção **baseados no erro do modelo**: DDM, ECDD e PHt.
- Métodos de detecção **baseados nas características da série**: FEDD.
- **Hipóteses de pesquisa:**
 - Ao **monitorar características da série temporal** é possível construir um **método explícito de detecção** capaz de detectar mudanças de forma efetiva, **minimizando o atraso na detecção**, **alarmes falsos** e **taxas de detecção** em relação a **métodos baseados no erro** do modelo.
- **Contribuições:** A contribuição original deste trabalho é o **FEDD** (Feature Extraction for Explicit Concept Drift Detection), um **método de detecção explícito baseado em características de séries temporais**.

2- Cavalcante et al (2016) - IJCNN

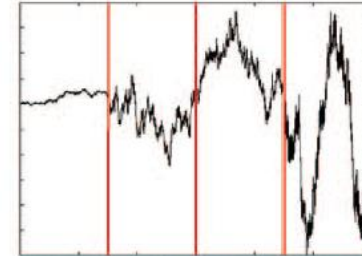
○ Séries Artificiais:



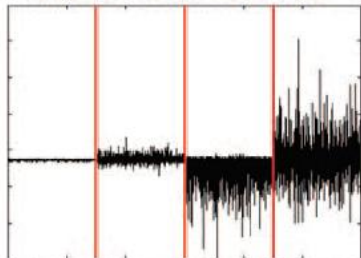
Linear Time Series1



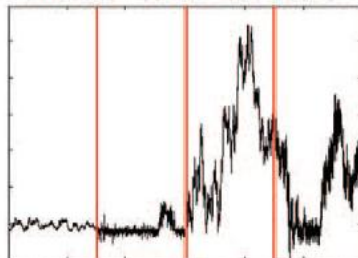
Linear Time Series2



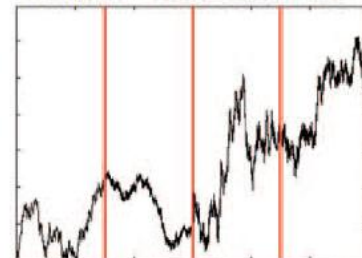
Linear Time Series3



Nonlinear Time Series1



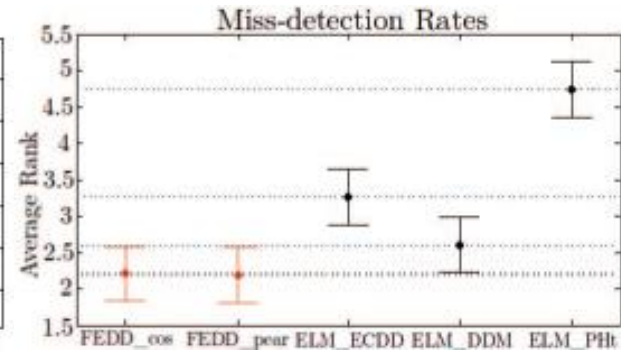
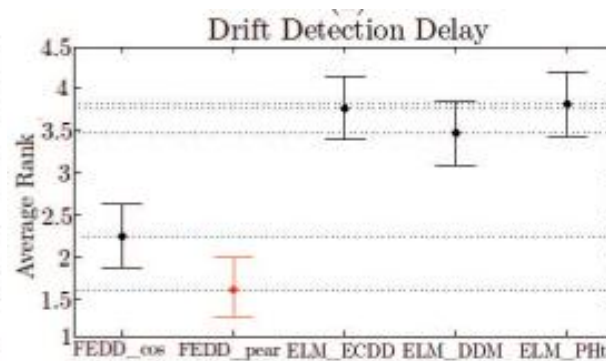
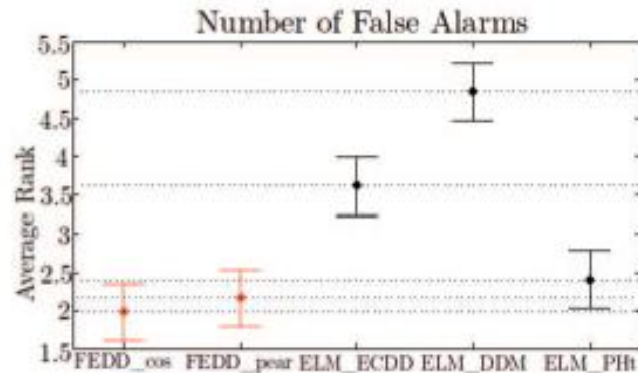
Nonlinear Time Series2



Nonlinear Time Series3

2- Cavalcante et al (2016) - IJCNN

○ Principais Resultados:

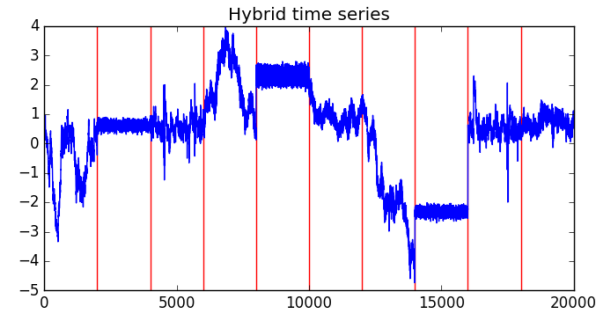
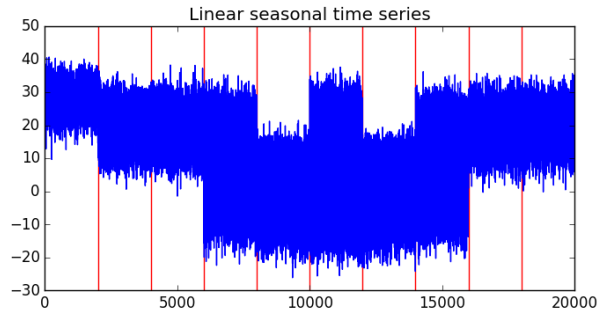
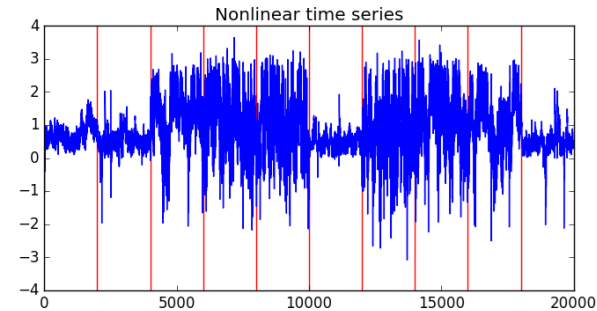
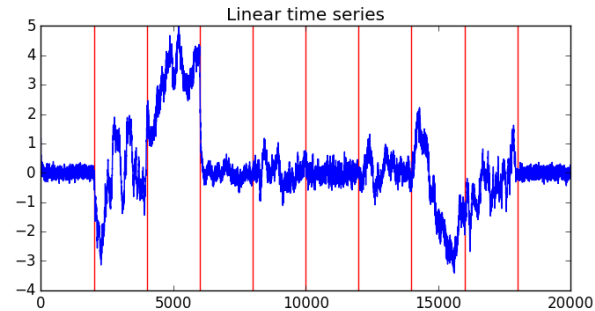


3- Oliveira et al (2017) - ICTAI

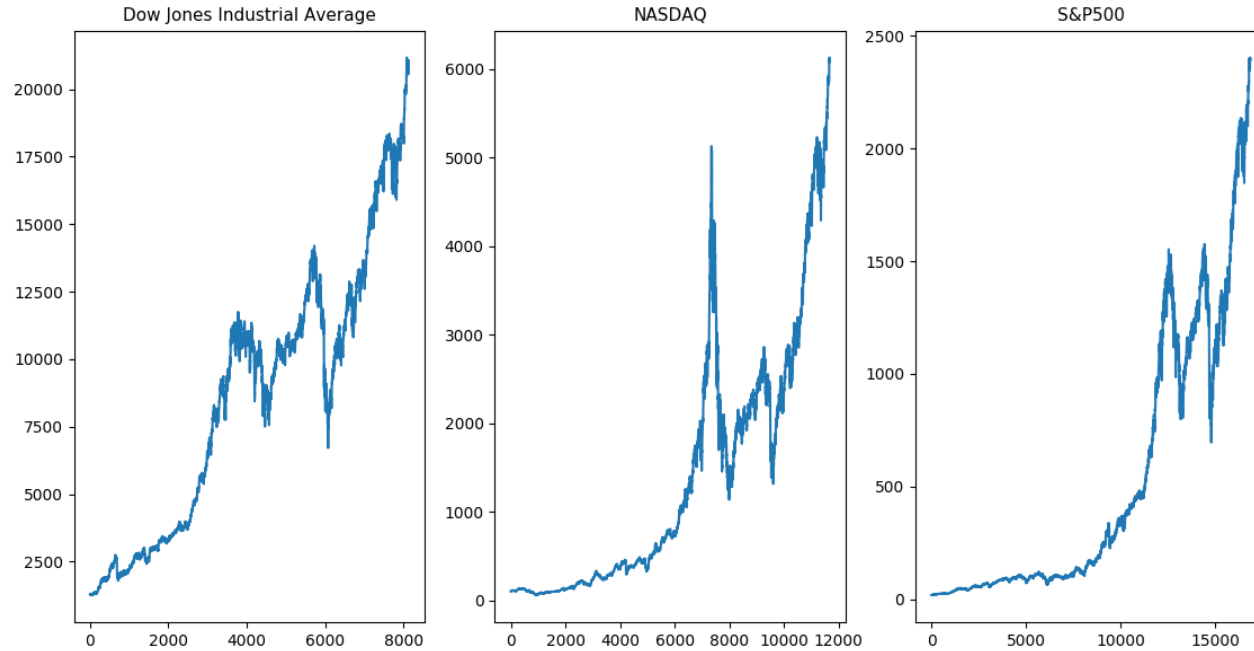
- **Objetivo:** Propõem um **sistema de aprendizagem adaptativo ativo** para previsão de séries temporais, com base no **monitoramento do erro de vários modelos de previsão** gerados por **inteligência de enxames**.
- Métodos de detecção **baseados em um modelo:** DDM, ECDD.
- Métodos de detecção **baseados em muitos modelos:** IDPSO-ELM-S, IDPSO-ELM-B.
- **Hipoteses de pesquisa:**
 - Quando ocorre uma **mudança no processo da geração de séries temporais** o **espaço de busca** enfrentado pelo PSO **reflete a mudança**. Assim, **em vez de monitorar o erro de um único método** de previsão, podemos **identificar a mudança** de conceito **ao monitorar uma área maior do espaço de busca** de otimização **através de vários modelos de previsão** gerados pelo PSO.
- **Contribuições:** Um **sistema de aprendizagem adaptativo** ativo que **monitora o erro de vários modelos** de previsão **baseado em inteligência de enxames**. O sistema proposto **atualiza o modelo preditivo** com inteligência de enxames **na presença de mudança de conceito**.

3- Oliveira et al (2017) - ICTAI

○ Séries artificiais:

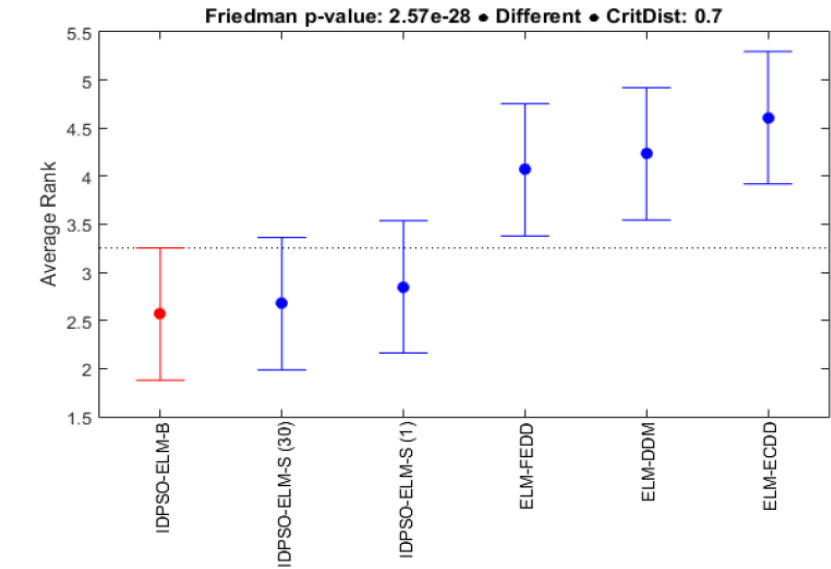


3- Oliveira et al (2017) - ICTAI

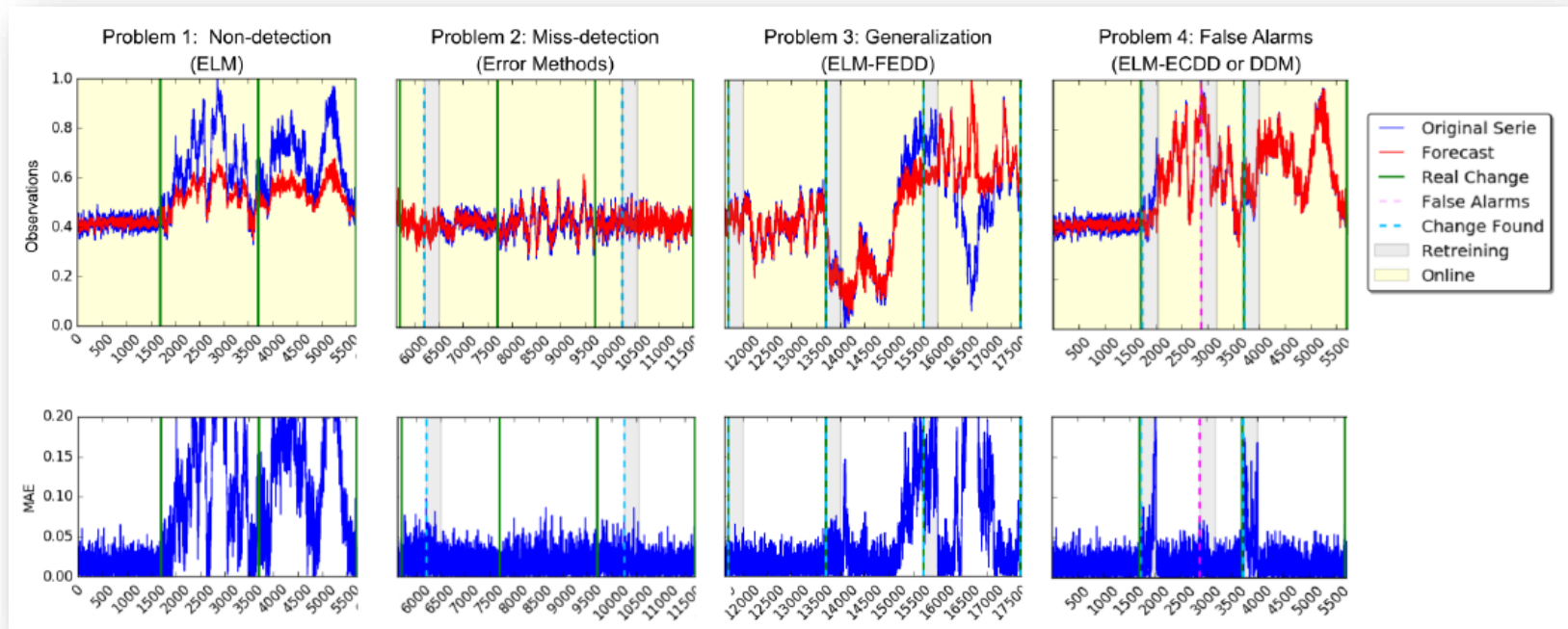


3- Oliveira et al (2017) - ICTAI

- Avaliação sobre a previsão de séries temporais artificiais com mudança de conceito.



3- Oliveira et al (2017) - ICTAI



Temas de Pesquisa



Como lidar com drifts virtuais e reais de forma simultânea em previsão de séries temporais?

Hipótese 1: Janela Adaptativa com Comparação de Séries.

O uso de uma janela dinâmica adaptativa, baseada na **similaridade entre trechos da série** temporal (ex: DTW ou métricas FEDD).

O objetivo é **preservar apenas os dados mais relevantes** após um conceito mudar, melhorando a adaptação do modelo a drifts reais.

Hipótese 2: Estratégia Dual de Erro para Adaptação a Drifts

A combinação **estratégias baseadas no erro do regressor** para lidar com drifts virtuais (atualização incremental via erro) e outra para drifts reais (troca seletiva de modelos via pool de históricos).

O objetivo é **aumentar a robustez e a acurácia de modelos** on-line frente a mudanças conceituais.

Referências

- Adams, Ryan Prescott, and David JC MacKay. "Bayesian online changepoint detection." arXiv preprint arXiv:0710.3742 (2007).
- Cavalcante, Rodolfo C., and Adriano LI Oliveira. "An approach to handle concept drift in financial time series based on Extreme Learning Machines and explicit Drift Detection." 2015 international joint conference on neural networks (IJCNN). IEEE, 2015.
- Cavalcante, Rodolfo C., et al. "Computational Intelligence and Financial Markets: A Survey and Future Directions." Expert Systems with Applications 55 (2016): 194-211.
- Cavalcante, Rodolfo C., Leandro L. Minku, and Adriano LI Oliveira. "FEDD: Feature Extraction for Explicit Concept Drift Detection in time series." Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE, 2016.
- Ditzler, Gregory, et al. "Learning in nonstationary environments: A survey." IEEE Computational Intelligence Magazine 10.4 (2015): 12-25.

Referências

- Dongre, Priyanka B., and Latesh G. Malik. "A review on real time data stream classification and adapting to various concept drift scenarios." Advance Computing Conference (IACC), 2014 IEEE International. IEEE, 2014.
- Gama, João, et al. "A survey on concept drift adaptation." ACM Computing Surveys (CSUR) 46.4 (2014): 44.
- Krawczyk, Bartosz, et al. "Ensemble learning for data stream analysis: a survey." Information Fusion 37 (2017): 132–156.
- LIMA, Marília et al. Learning under concept drift for regression—a systematic literature review. IEEE Access, v. 10, p. 45410–45429, 2022.
- Minku, Leandro L., and Xin Yao. "DDD: A new ensemble approach for dealing with concept drift." IEEE transactions on knowledge and data engineering 24.4 (2012): 619–633.
- SUSNJAK, Teo; MADDIGAN, Paula. Forecasting patient flows with pandemic induced concept drift using explainable machine learning. EPJ Data Science, v. 12, n. 1, p. 11, 2023.
- Zliobaite, Indre, et al. "Next challenges for adaptive learning systems." ACM SIGKDD Explorations Newsletter 14.1 (2012): 48–55.





Obrigado!

Overview Repositories 18 Projects Packages

Pinned

- IDPSO-ELM-S** (Public)
Algorithms proposed in the following paper: OLIVEIRA, Gustavo HFMO et al. Time series forecasting in the presence of concept drift: A pso-based approach. In: 2017 IEEE 29th International Conference...
Python 10
- TimeSeriesCln** (Public)
Repositório com os códigos para a monitoria da disciplina de séries temporais da pós-graduação em inteligência computacional do centro de informática (CIn) - UFPE.
Python 6 1
- GMM-VRD** (Public)
Algorithms proposed in the following paper: Oliveira, Gustavo HFM, Leandro L. Minku, and Adriano LI Oliveira. "GMM-VRD: A Gaussian Mixture Model for Dealing With Virtual and Real Concept Drifts." 2...
Python 4 2
- GrupoSerDisciplinas** (Public)
Repositórios de códigos para as disciplinas do curso de data science - EAD do grupo ser educacional.
Jupyter Notebook 3 2

Gustavo Henrique Ferreira de Miranda Oliveira
GustavoHFMO
Follow

Ph.D. in Artificial Intelligence.
Professor and Data Scientist.

Códigos dos artigos apresentados

Apresentação



UFAL
CAMPUS ARAPIRACA
U. E. PENEDO

<https://github.com/GustavoHFMO>

Previsão de Séries Temporais na Presença de Mudança de Conceito

Gustavo H. Ferreira de Miranda Oliveira
gustavo.oliveira@penedo.ufal.br



Centro de Informática – UFPE
24/04/2025