

Método de concept drift contextual para aprendizado online com redes neurais na Fog Computing



Brenno de Mello Alencar
Orientador: Ricardo Rios
Coorientador: Cássio Prazeres



Método de concept drift contextual para aprendizado online com redes neurais na Fog Computing

Nome	Brenno de Mello Alencar
Curso	Doutorado
Orientador	Ricardo Rios
Ingresso	20/11/2020
Qualificação	04/2023
Defesa	11/2023
Bolsista?	Sim

Contexto

- Internet das Coisas possibilita a conexão, orquestração e coordenação de um grande número de dispositivos à Internet.
- Exemplos de dispositivos e domínios de aplicações IoT:
 - Monitor cardíaco, automóvel, saúde, agricultura, casas inteligentes
- Os dispositivos possuem a capacidade de comunicar e trocar informações entre si para atingir objetivos comuns

Big Data Stream

- Os fluxos de dados IoT apresentam as seguintes características:
 - Fluxos ilimitados continuamente ao longo do tempo;
 - Alta velocidade;
 - Mudam o seu comportamento com o passar do tempo;
 - Não são considerados independentes e identicamente distribuídos.
- Características diferenciadas para os Vs do Big Data:
 - Volume, velocidade, variedade, Veracidade
- Definição de Concept Drift $\exists X : p_0(X, y) \neq p_1(X, y)$

Motivação/Problema

- O desafio de processar e analisar dados no *Big Data Stream* da IoT apresenta as seguintes características:
 - Processar e inspecionar um exemplo de cada vez apenas uma vez e na ordem de chegada;
 - Utilizar uma quantidade limitada de memória;
 - A complexidade temporal e espacial deve ser linear ou sublinear;
 - Os algoritmos devem responder a qualquer momento.

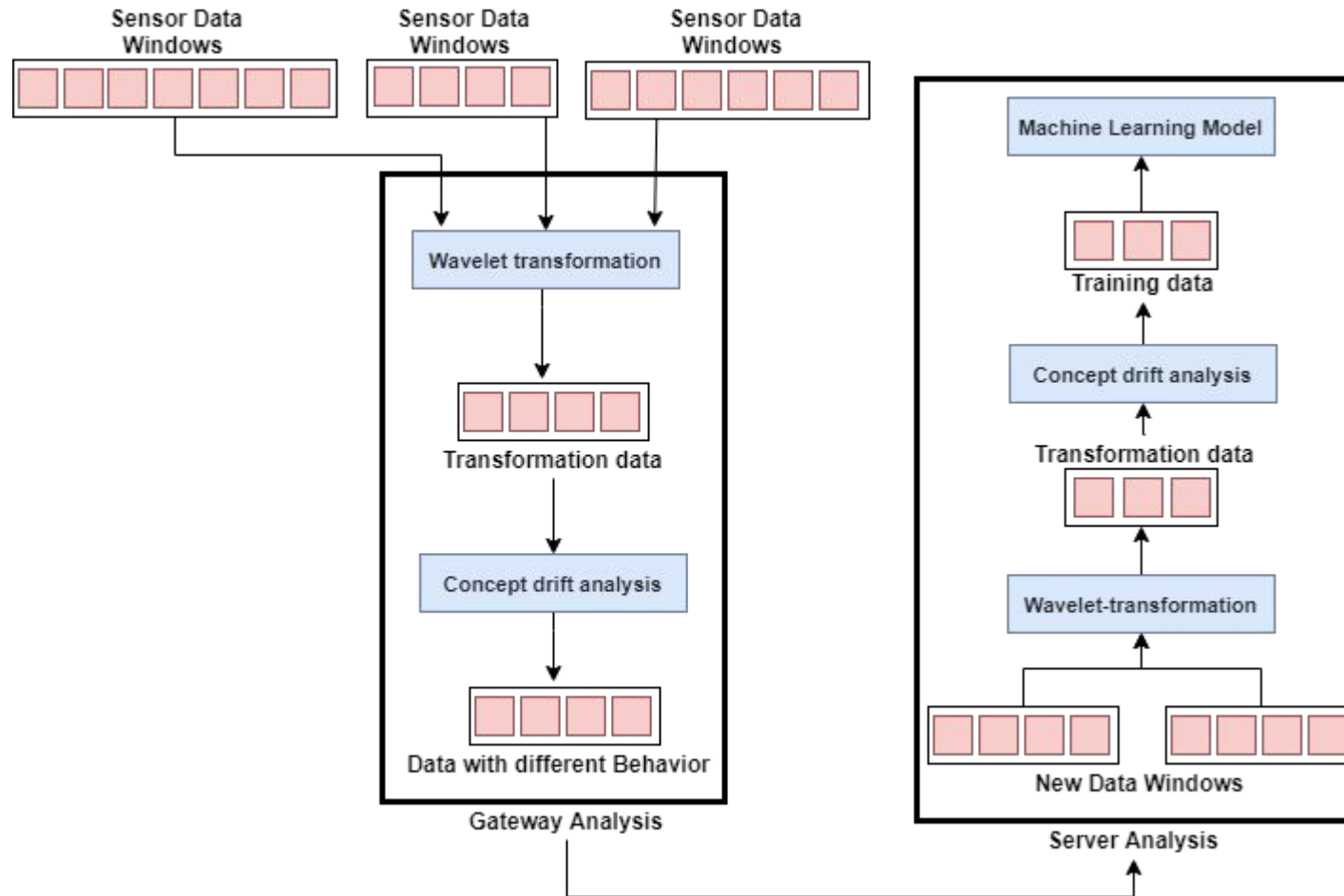
Problema de Pesquisa

- Como processar e analisar fluxos de dados em tempo real na *Fog Computing* para Internet das Coisas para:
 - Processar dados com baixo atraso a partir de modelos de redes neurais.
 - Reduzir a quantidade de dados trafegados na rede.
 - Sem necessidade de conectividade constante com Internet.

Hipótese

- O desenvolvimento de um método de detecção de *concept drift* baseado em contexto para fluxos de dados aplicado na *Fog Computing* pode permitir que dados sejam processados, filtrados e analisados em tempo real (*online*) a partir de modelos de redes neurais. Além disso, pode:
 - Reduzir quantidade de dados processados.
 - Reduzir a utilização da rede.
 - Obter respostas mais rápidas para alterações no fluxo de dados.
 - Manutenção da acurácia de predição na aplicação de algoritmo de redes neurais na *Fog Computing*.

Proposta



Planejamento de Experimentos Ambiente Simulado

	Sensores	Gateways	Algoritmo Concept Drift	Configuração
Execução 1	10	2	CUSUM	1
Execução 2	20	4	CUSUM	2
Execução 3	10	2	Page Hinkley	1
Execução 4	20	4	Page Hinkley	2
Execução 5	10	2	EWMA	1
Execução 6	20	4	EWMA	2

Análise dos Resultados Ambiente Simulado

- Resultados para execuções 1, 3 e 5

	Gateway 1			Gateway 2		
	RMSE	MAE	MSE	RMSE	MAE	MSE
Execução 1	0.47	0.32	0.35	0.36	0.26	0.39
Execução 3	0.46	0.29	0.31	0.28	0.19	0.14
Execução 5	0.58	0.36	0.59	0.33	0.21	0.26

- Resultados para execuções 2, 4 e 6

	Gateway 1			Gateway 2		
	RMSE	MAE	MSE	RMSE	MAE	MSE
Execução 2	0.82	0.53	1.09	0.42	0.26	0.29
Execução 4	1.02	0.85	1.69	0.47	0.31	0.37
Execução 6	0.63	0.38	0.54	0.35	0.22	0.21

Análise dos Resultados Ambiente Simulado

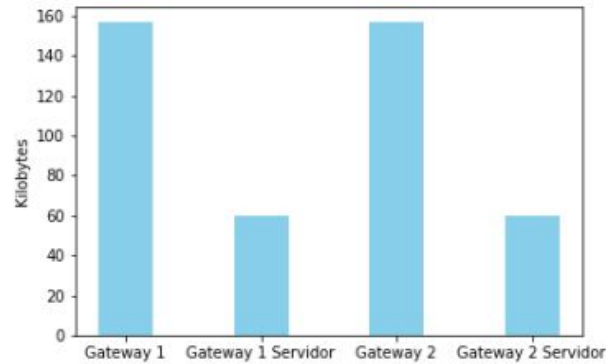
- Resultados para execuções 2, 4 e 6

	Gateway 3			Gateway 4		
	RMSE	MAE	MSE	RMSE	MAE	MSE
Execução 2	1.01	0.62	1.47	0.77	0.56	1.21
Execução 4	1.17	0.84	2.01	0.66	0.40	0.70
Execução 6	0.97	0.56	1.24	0.51	0.31	0.46

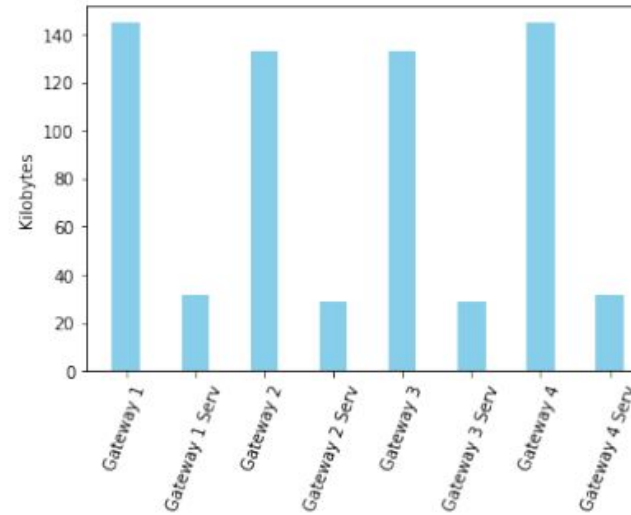
Análise dos Resultados Ambiente Simulado

- Quantidade de dados processados experimentos 1, 2, 3 e 4.

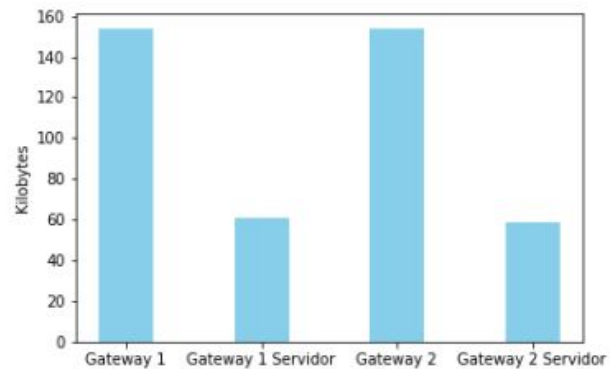
(a)



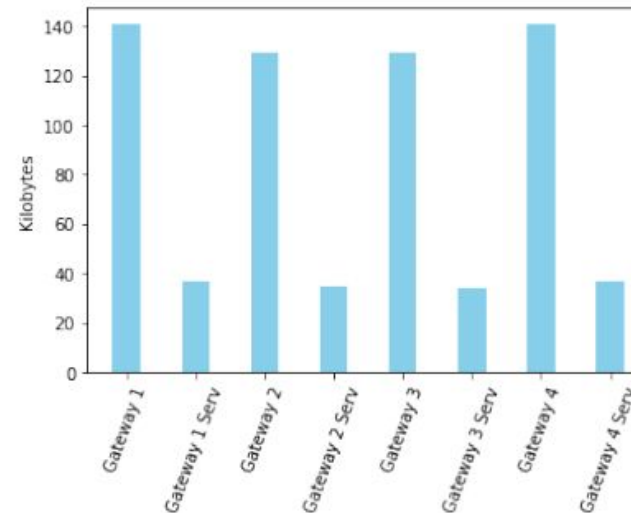
(b)



(c)



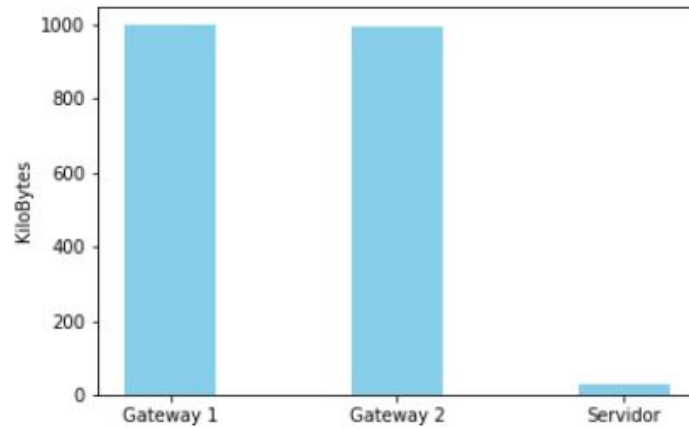
(d)



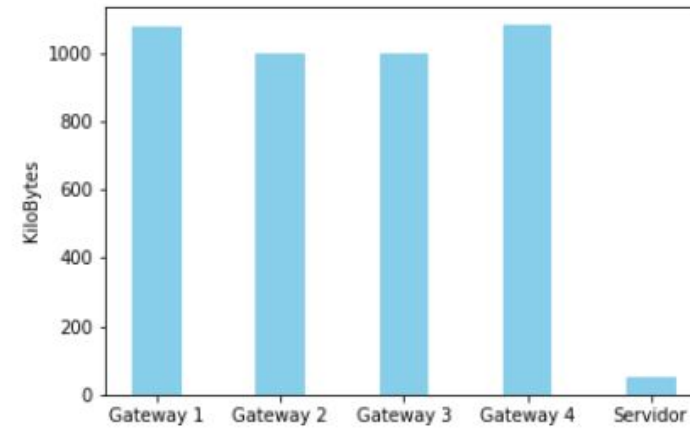
Análise dos Resultados Ambiente Simulado

- Quantidade de dados trafegados na rede experimentos 1, 2, 3 e 4.

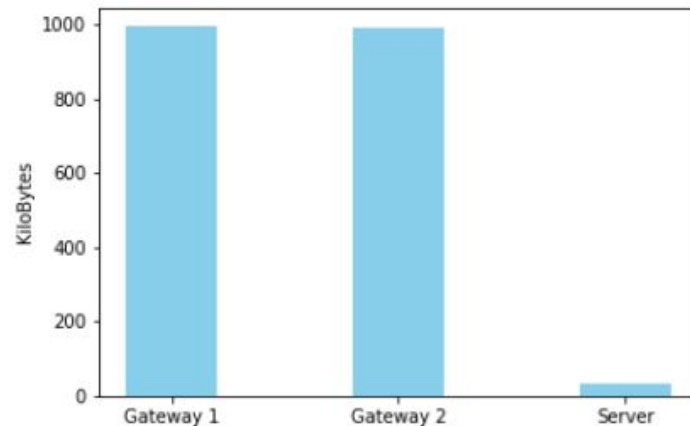
(a)



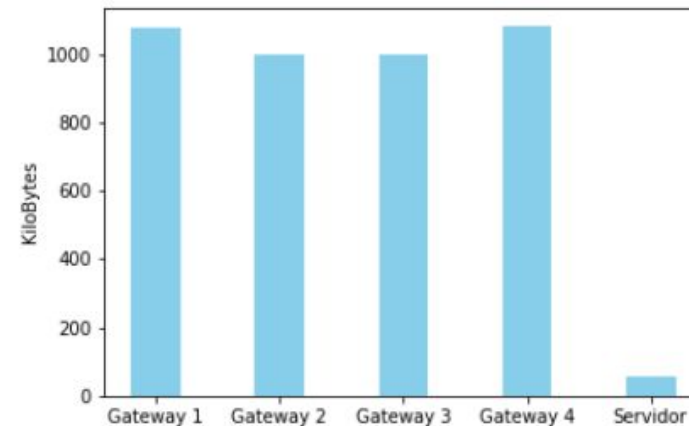
(b)



(c)



(d)



Conclusão

- Limitações:
 - Métricas de desempenho não foram avaliadas nos experimentos
 - Os experimentos emulados foram executados na mesma máquina
 - A abordagem não foi implantada e testada a partir de dispositivos mais limitados (Microcontroladores)
- Trabalhos Futuros:
 - Implementação da abordagem para ser eficiente aplicada em dispositivos com menor poder de processamento.
 - Novos experimentos em cenários reais e emulados para verificar o consumo de recursos computacionais pela abordagem, como, por exemplo, CPU e memória.
 - Realizar testes em larga escala para algoritmos que utilizem modelos de previsão sofisticados com análises a partir dos sensores.

Referências Bibliográficas

AFICAS AL-FUQAHA, A. et al. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials* , IEEE, v. 17, n. 4, p. 2347–2376, 2015.

ANAND, N. et al. Practical edge analytics: Architectural approach and use cases. In: IEEE. *Edge Computing (EDGE), 2017 IEEE International Conference on* . [S.l.], 2017. p. 236–239.

ATZORI, L.; IERA, A.; MORABITO, G. The internet of things: A survey. *Computer networks* , Elsevier, v. 54, n. 15, p. 2787–2805, 2010.

BIFET, A. Mining big data in real time. *Informatica (Slovenia)* , v. 37, n. 1, p. 15–20, 2013.

BIFET, A.; KIRKBY, R. *Data stream mining: a practical approach* . [S.l.], 2009.