

X Escola Regional de
Alto Desempenho da
Região Sudeste

5 a 7 de Novembro

Explorando Modelos Preditivos de Tempo de Execução de Simulações de Reservatório em Clusters HPC

Alan L. Nunes¹, Lucas A. F. da Costa³, Felipe A. Portella²,
José Viterbo¹, Leonardo V. Neri³, Paulo J. B. Estrela²,
Luciana de C. Franci³, Renzo Q. Malini², Lúcia M. A. Drummond¹

¹Instituto de Computação – Universidade Federal Fluminense (UFF), Niterói – Brasil

²Petróleo Brasileiro S.A. (PETROBRAS), Rio de Janeiro – Brasil

³Centro de Estudos e Sistemas Avançados do Recife (CESAR), Recife – Brasil

Agenda

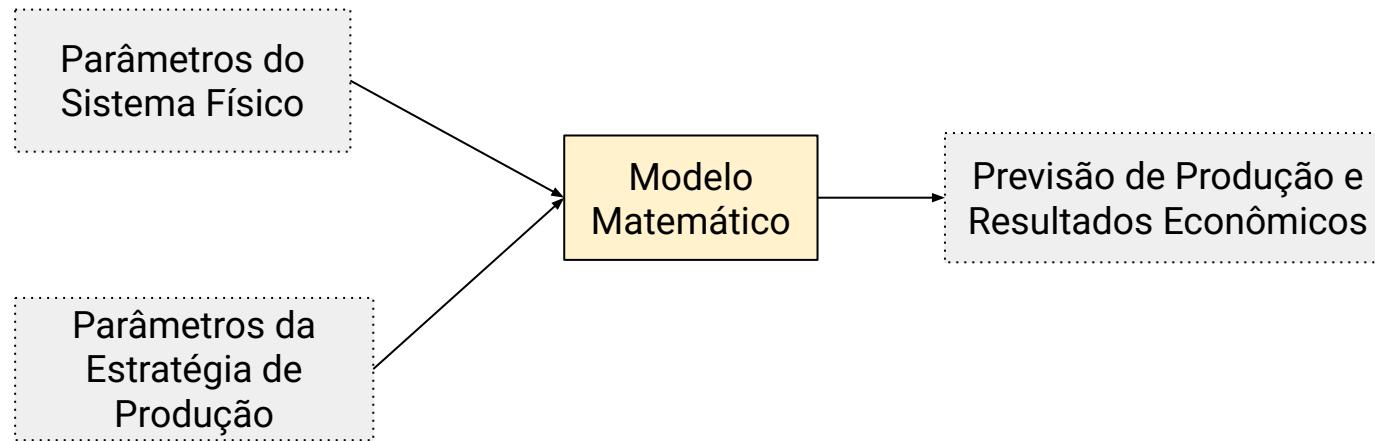
- ❖ Introdução
- ❖ Resultados Experimentais
- ❖ Conclusão e Direções Futuras

Introdução - Simulação de Reservatórios Petrolíferos

Simulação de Reservatórios é uma atividade padrão e essencial na indústria do petróleo:

- Para decidir quais as **melhores estratégias de exploração** (extração) de um campo, é necessário reproduzir (**simular**) qual seria o **comportamento** do reservatório para diferentes cenários, variáveis e incertezas.
 - **Poços**: Quantos? Onde perfurá-los? Verticais, Horizontais ou Multilaterais?
 - **Método de Recuperação**: Depleção Primária? Injeção de água ou gás? Métodos Térmicos? Como controlar os poços? Que vazões usar?
 - **Previsão de Comportamento**: Recuperação final? Fluxo de caixa? Reservas?
 - **Incertezas Geológicas**: Geometria externa; Distribuição espacial de propriedades; Falhas (localização, rejeito, transmissibilidades);
 - **Incertezas na Distribuição Inicial de Fluidos**: Posição dos contatos água-óleo e gás-óleo; Saturações residuais de água e óleo;
 - **Incertezas nas Propriedades dos Fluidos e Rocha**: Compressibilidade; Viscosidade; Permeabilidade relativa;
 - **Incertezas Econômicas**: Preço do petróleo;
 - ...

Introdução - Simulação de Reservatórios Petrolíferos



Parâmetros do Sistema Físico: Dados da geologia, dos fluidos, etc;

Parâmetros da Estratégia de Produção: Número e locação dos poços, vazões de injeção, etc;

Previsão de Produção e Resultados Econômicos: Produção acumulada de óleo e gás, Valor Presente Líquido (VPL), etc;

Modelo Matemático: Baseado em leis físicas (conservação de massa e energia, etc) e relações verificadas experimentalmente.

Introdução - Simulações de Reservatório em Clusters HPC

Por que rodar simulações de reservatório em clusters HPC (High-Performance Computing)?

➤ Alta complexidade dos modelos matemáticos

- Equações diferenciais parciais não lineares que descrevem o fluxo de fluidos (óleo, gás, água) em meios porosos;
- A discretização espacial e temporal dessas equações pode resultar em milhões ou bilhões de variáveis;

➤ Escalabilidade

- Divisão do domínio de simulação (malha do reservatório) entre centenas ou milhares de núcleos de CPU;
- A paralelização reduz drasticamente o tempo de execução, possibilitando análises que levariam dias ou semanas;

➤ Análise de incertezas com prazos viáveis

- Estudos de sensibilidade, history matching e otimização de produção exigem a execução de centenas ou milhares de simulações com diferentes parâmetros.

Introdução - Modelos Preditivos de Tempo de Execução de Simulações de Reservatório

Por que estimar o tempo de execução de jobs de simulação de reservatório em clusters HPC?

- **Otimização de recursos**
 - Agendamento mais inteligente de jobs (pode melhorar a taxa de utilização do cluster);
- **Gestão de custos**
 - Controle de orçamento e alocação eficiente de recursos;
- **Priorização de jobs**
 - Classificação e ordenação das execuções com base em critérios de urgência e disponibilidade de recursos;
- **Detecção de anomalias**
 - Estimativas inconsistentes com tempos reais podem indicar anomalias na execução.

Introdução - Objetivos Gerais

Objetivo: Avaliação de alguns modelos de aprendizado de máquina para a predição do tempo de execução de jobs de simulação de reservatórios petrolíferos em clusters HPC, utilizando logs reais da Petrobras.

Resultados Experimentais - Metodologia

- Dados reais do gerenciador de jobs SLURM do ambiente de HPC da Petrobras:
 - Período de junho/2022 a junho/2023;
 - O conjunto de dados final contém 1.305.452 registros de jobs após filtragem e pré-processamento;
 - 90% dos dados (1.174.186 jobs), de junho/2022 a maio/2023, foram usados para treinamento e validação cruzada;
 - 10% dos dados (131.266 jobs), de junho/2023, foram usados para teste, simulando um cenário real de predição futura sem vazamento de informação (Out-of-Sample).
- Dois modelos de classificação foram avaliados:
 - XGBoost, baseado em gradient boosting e de alto desempenho;
 - J48, implementação da árvore de decisão C4.5.
- Duas estratégias de codificação de atributos categóricos foram comparadas:
 - Target Encoder, que usa a média do valor alvo por categoria;
 - Ordinal Encoder, que atribui valores inteiros arbitrários às categorias.
- Dois esquemas de discretização do tempo de execução (atributo elapsed) foram analisados (a seguir).

Resultados Experimentais - Esquemas de Discretização do Atributo Alvo

Esquema de Discretização	Classe	Intervalo de Duração (em segundos)	Distribuição de Jobs	
			Treinamento	Teste
Esquema 1	0	$300 \leq \text{elapsed} < 1.800$	472.030	74.450
	1	$1.800 \leq \text{elapsed} < 3.600$	148.470	14.791
	2	$3.600 \leq \text{elapsed} < 7.200$	122.207	11.531
	3	$7.200 \leq \text{elapsed} < 14.400$	131.702	7.904
	4	$14.400 \leq \text{elapsed} < 21.600$	48.239	5.778
	5	$21.600 \leq \text{elapsed} < 28.800$	37.238	2.671
	6	$28.800 \leq \text{elapsed} < 36.000$	39.225	2.425
	7	$36.000 \leq \text{elapsed} < 43.200$	43.931	1.664
	8	$43.200 \leq \text{elapsed} < 50.400$	30.547	1.621
	9	$50.400 \leq \text{elapsed} < 57.600$	22.320	1.440
	10	$57.600 \leq \text{elapsed} < 64.800$	16.616	1.221
	11	$64.800 \leq \text{elapsed} < 72.000$	12.633	850
	12	$72.000 \leq \text{elapsed} < 79.200$	7.527	683
	13	$79.200 \leq \text{elapsed} < 86.400$	4.973	575
	14	$\text{elapsed} \geq 86.400$	36.528	3.662
Esquema 2	0	$300 \leq \text{elapsed} \leq 600$	153.811	17.474
	1	$601 \leq \text{elapsed} \leq 1.250$	230.894	50.662
	2	$1.251 \leq \text{elapsed} \leq 4.000$	254.793	23.285
	3	$4.001 \leq \text{elapsed} \leq 6.000$	71.276	7.664
	4	$\text{elapsed} \geq 6.001$	463.412	32.181

Resultados Experimentais - Esquemas de Discretização do Atributo Alvo

Esquema de Discretização	Classe	Intervalo de Duração (em segundos)	Distribuição de Jobs	
			Treinamento	Teste
Esquema 1	0	$300 \leq \text{elapsed} < 1.800$	472.030	74.450
	1	$1.800 \leq \text{elapsed} < 3.600$	148.470	14.791
	2	$3.600 \leq \text{elapsed} < 7.200$	122.207	11.531
	3	$7.200 \leq \text{elapsed} < 14.400$	131.702	7.904
	4	$14.400 \leq \text{elapsed} < 21.600$	48.239	5.778
	5	$21.600 \leq \text{elapsed} < 28.800$	37.238	2.671
	6	$28.800 \leq \text{elapsed} < 36.000$	39.225	2.425
	7	$36.000 \leq \text{elapsed} < 43.200$	43.931	1.664
	8	$43.200 \leq \text{elapsed} < 50.400$	30.547	1.621
	9	$50.400 \leq \text{elapsed} < 57.600$	22.320	1.440
	10	$57.600 \leq \text{elapsed} < 64.800$	16.616	1.221
	11	$64.800 \leq \text{elapsed} < 72.000$	12.633	850
	12	$72.000 \leq \text{elapsed} < 79.200$	7.527	683
	13	$79.200 \leq \text{elapsed} < 86.400$	4.973	575
	14	$\text{elapsed} \geq 86.400$	36.528	3.662
Esquema 2	0	$300 \leq \text{elapsed} \leq 600$	153.811	17.474
	1	$601 \leq \text{elapsed} \leq 1.250$	230.894	50.662
	2	$1.251 \leq \text{elapsed} \leq 4.000$	254.793	23.285
	3	$4.001 \leq \text{elapsed} \leq 6.000$	71.276	7.664
	4	$\text{elapsed} \geq 6.001$	463.412	32.181

Esquema 1

Vantagens:

- Maior granularidade;
- Os intervalos menores capturam melhor a variação de jobs mais rápidos;
- Pode capturar padrões mais sutis na distribuição de tempos;

Desvantagens:

- Mais classes podem dificultar a interpretação do modelo;
- Risco de overfitting (classes com poucas amostras);
- Desequilíbrio significativo (Classe 0: 472.030 jobs; Classe 13: 4.973 jobs → razão de 95:1).

Resultados Experimentais - Esquemas de Discretização do Atributo Alvo

Esquema de Discretização	Classe	Intervalo de Duração (em segundos)	Distribuição de Jobs	
			Treinamento	Teste
Esquema 1	0	$300 \leq \text{elapsed} < 1.800$	472.030	74.450
	1	$1.800 \leq \text{elapsed} < 3.600$	148.470	14.791
	2	$3.600 \leq \text{elapsed} < 7.200$	122.207	11.531
	3	$7.200 \leq \text{elapsed} < 14.400$	131.702	7.904
	4	$14.400 \leq \text{elapsed} < 21.600$	48.239	5.778
	5	$21.600 \leq \text{elapsed} < 28.800$	37.238	2.671
	6	$28.800 \leq \text{elapsed} < 36.000$	39.225	2.425
	7	$36.000 \leq \text{elapsed} < 43.200$	43.931	1.664
	8	$43.200 \leq \text{elapsed} < 50.400$	30.547	1.621
	9	$50.400 \leq \text{elapsed} < 57.600$	22.320	1.440
	10	$57.600 \leq \text{elapsed} < 64.800$	16.616	1.221
	11	$64.800 \leq \text{elapsed} < 72.000$	12.633	850
	12	$72.000 \leq \text{elapsed} < 79.200$	7.527	683
	13	$79.200 \leq \text{elapsed} < 86.400$	4.973	575
	14	$\text{elapsed} \geq 86.400$	36.528	3.662
Esquema 2	0	$300 \leq \text{elapsed} \leq 600$	153.811	17.474
	1	$601 \leq \text{elapsed} \leq 1.250$	230.894	50.662
	2	$1.251 \leq \text{elapsed} \leq 4.000$	254.793	23.285
	3	$4.001 \leq \text{elapsed} \leq 6.000$	71.276	7.664
	4	$\text{elapsed} \geq 6.001$	463.412	32.181

Esquema 2

Vantagens:

- Distribuição mais equilibrada entre classes;
- Menor risco de overfitting;
- Simplicidade de interpretação (apenas 5 classes);

Desvantagens:

- Intervalos muito amplos agrupando comportamentos distintos (ex: Classe 4);
- Pouca granularidade para jobs mais longos;
- Diferenças de comportamento entre jobs de 10 min e 1h são perdidas.

Resultados Experimentais - Análise do Conjunto de Teste Completo (Acurácia, MAE, MAPE)

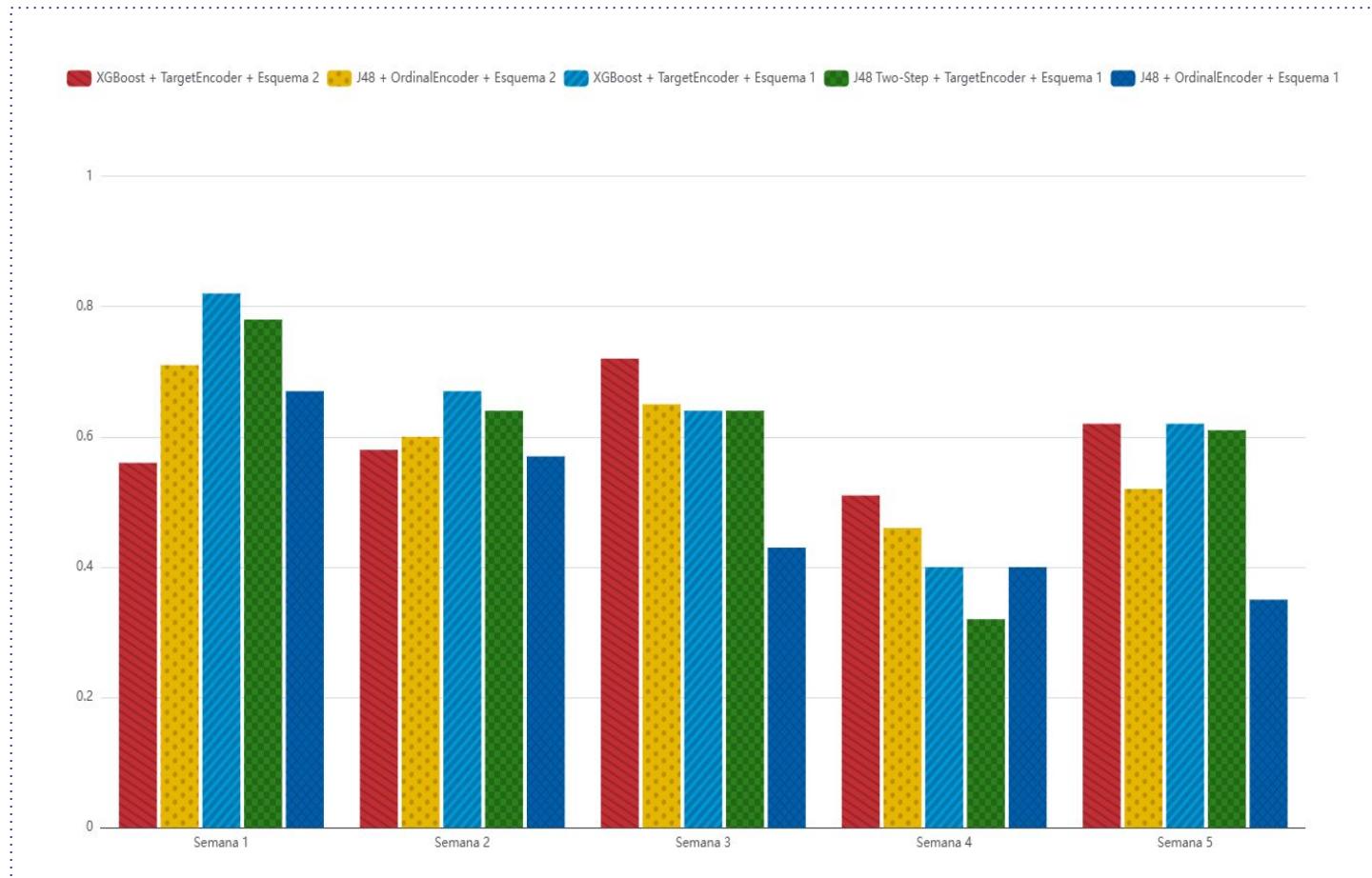
Modelo	Esquema de Discretização	Acurácia	MAE*	MAPE* (%)
XGBoost + TargetEncoder	Esquema 2	0,58	11577,62	82,14
J48 + OrdinalEncoder	Esquema 2	0,62	11562,17	77,91
XGBoost + TargetEncoder	Esquema 1	0,68	8642,40	131,58
J48 Two-Step + TargetEncoder	Esquema 1	0,64	10451,62	175,19
J48 + OrdinalEncoder	Esquema 1	0,53	10240,45	235,87

*MAE e MAPE foram calculados utilizando o valor do limite inferior do intervalo de tempo da classe predita.

Trade-off entre a acurácia e a magnitude do erro percentual:

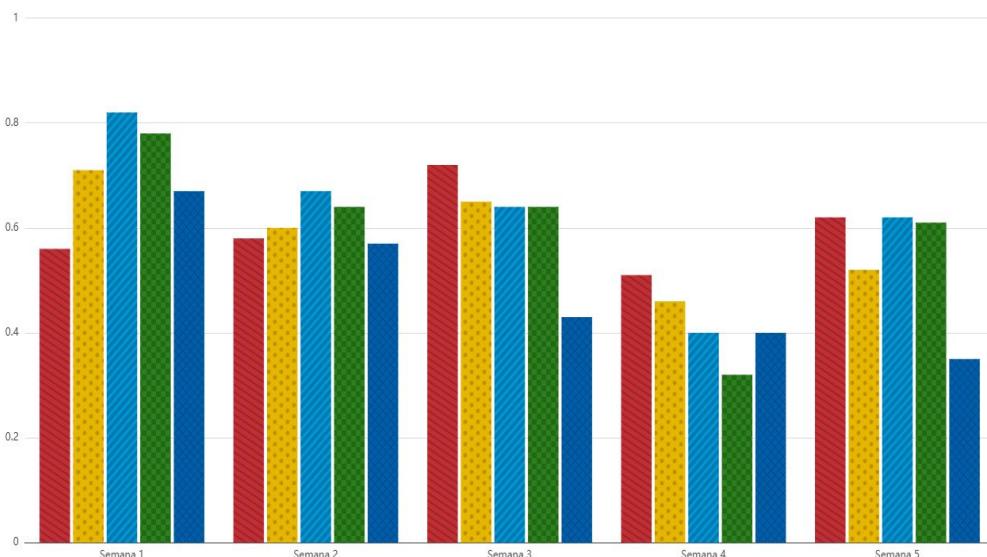
- XGBoost + TargetEncoder + Esquema 1: obteve a melhor acurácia (0.68) e MAE (≈ 8642), mas seu MAPE elevado (131.58%) indica baixa precisão relativa para jobs de longa duração;
- J48 + OrdinalEncoder + Esquema 2: obteve o menor erro percentual (MAPE=77,91%), em troca de uma menor precisão absoluta e acurácia;
- J48 + OrdinalEncoder + Esquema 1: obteve o pior desempenho de MAPE (235,87%), evidenciando a sensibilidade ao desbalanceamento do esquema de discretização granular.

Resultados Experimentais - Análise do Conjunto de Teste Por Semana (Acurácia)



Resultados Experimentais - Análise do Conjunto de Teste Por Semana (Acurácia)

XGBoost + TargetEncoder + Esquema 2 | J48 + OrdinalEncoder + Esquema 2 | XGBoost + TargetEncoder + Esquema 1 | J48 Two-Step + TargetEncoder + Esquema 1 | J48 + OrdinalEncoder + Esquema 1



XGBoost + TargetEncoder + Esquema 2:

Desempenho moderado (56%–72%), com pico na 3^a semana e leve recuperação na 5^a, sugerindo boa adaptação.

J48 + OrdinalEncoder + Esquema 2:

Desempenho sem picos extremos (71%–46%), com estabilidade relativa, mantendo comportamento previsível.

XGBoost + TargetEncoder + Esquema 1:

Alta variabilidade (82%–40%), embora com a melhor precisão geral, indicando vulnerabilidade a mudanças de distribuição.

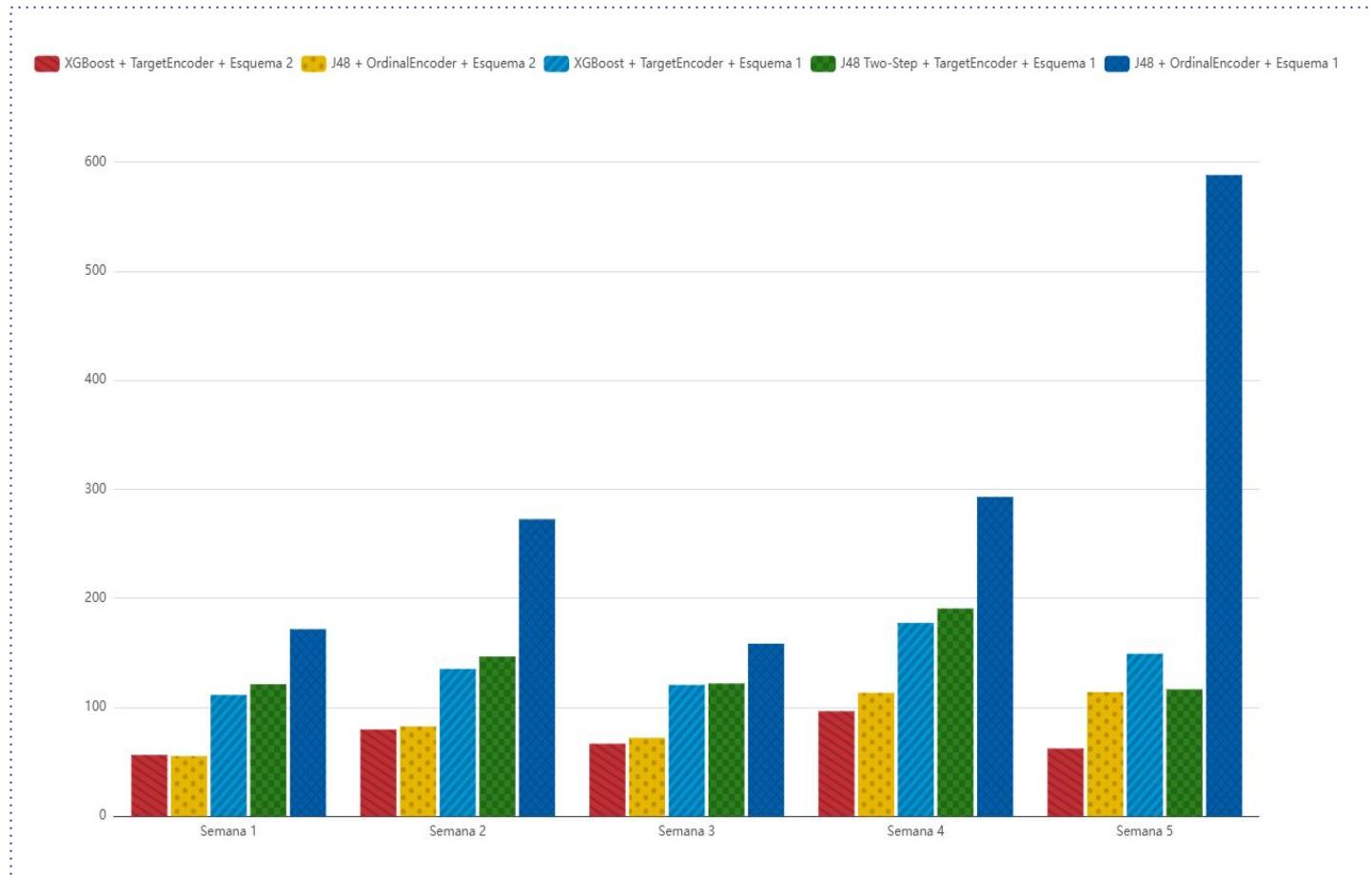
J48 Two-Step + TargetEncoder + Esquema 1:

Resultados consistentes até a 3^a semana (78%–64%), seguidos de queda acentuada na 4^a (32%), mostrando perda de generalização.

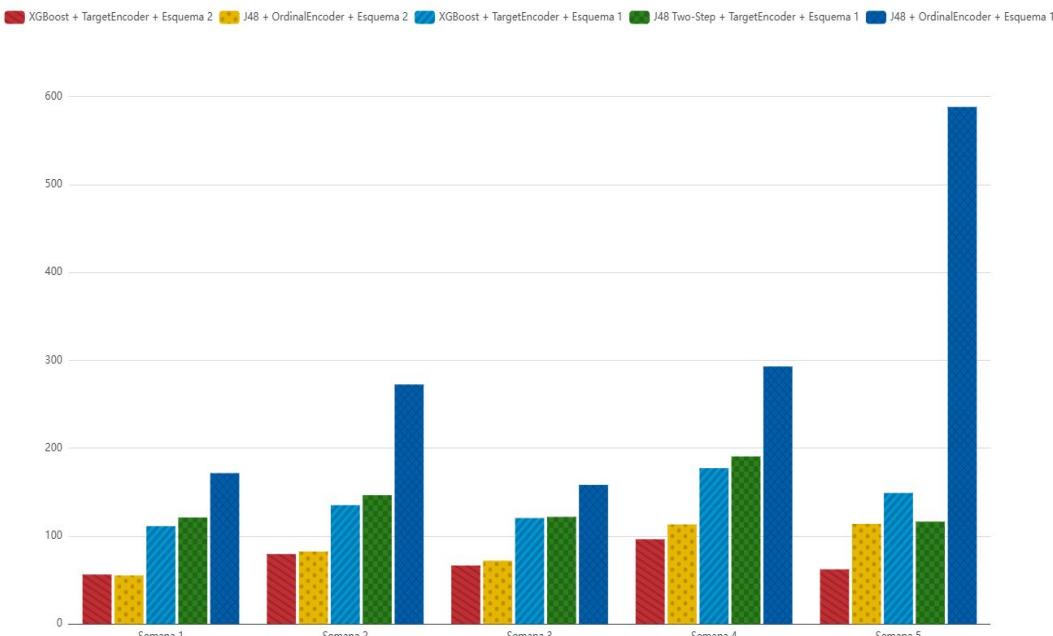
J48 + OrdinalEncoder + Esquema 1:

Baixa robustez e queda acentuada (35% na 5^a semana), confirmando sensibilidade ao desbalanceamento e degradação progressiva do desempenho preditivo.

Resultados Experimentais - Análise do Conjunto de Teste Por Semana (MAPE)



Resultados Experimentais - Análise do Conjunto de Teste Por Semana (MAPE)



XGBoost + TargetEncoder + Esquema 2:

Erros moderados (56%–96%), com bom equilíbrio geral, sugerindo maior consistência preditiva.

J48 + OrdinalEncoder + Esquema 2:

Erros crescentes nas semanas finais (até 113%), indicando leve instabilidade sob mudanças de cenário.

XGBoost / J48 Two-Step + TargetEncoder + Esquema 1:

MAPE consistentemente alto (111%–177% e 121%–190%), indicando sensibilidade ao esquema de discretização.

J48 + OrdinalEncoder + Esquema 1:

Pior desempenho geral, com MAPE explosivo na 5ª semana (588%), demonstrando forte instabilidade e baixa capacidade de generalização.

Conclusão

- A escolha do esquema de discretização é determinante para o desempenho de modelos (classificadores) de previsão de tempo de execução de jobs em ambientes HPC;
- Esquemas com maior granularidade podem atingir picos de acurácia, mas apresentam volatilidade e altos erros percentuais em semanas com distribuição atípica de jobs;
- A discretização com menor granularidade demonstrou maior robustez, mantendo o erro percentual médio em níveis mais baixos e estáveis ao longo do tempo;
- O modelo **XGBoost + TargetEncoder + Esquema 2** destacou-se como o mais equilibrado, sendo particularmente adequado a cenários de implantação onde a consistência das previsões supera a importância de acertos pontuais.

Direções Futuras

- Investigar métodos avançados de balanceamento por classe;
- Realizar uma análise detalhada dos erros por classe;
- Desenvolver abordagens adaptativas de discretização conforme a distribuição temporal dos jobs;
- Avaliar o desempenho de outros modelos preditores.

Obrigado!

alan_lira@id.uff.br

Resultados Experimentais - Metodologia

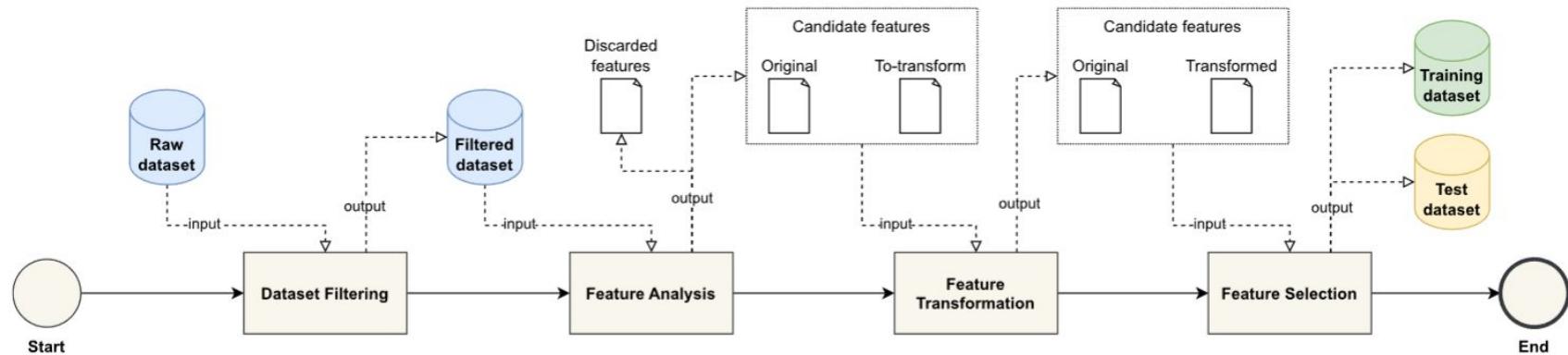


FIGURE 1 | Data preprocessing pipeline of the reservoir simulation jobs dataset.

Nunes, Alan L., et al. "Two-Step Estimation Strategy for Predicting Petroleum Reservoir Simulation Jobs Runtime on an HPC Cluster." *Concurrency and Computation: Practice and Experience* 37.4-5 (2025): e70026.