



Um Feature Store leve no Microsoft Fabric: scoring em tempo real sem plataformas pesadas

Desliza →

Por que um feature store leve?

Sem um ponto único de verdade, as equipes reescrevem transformações e gastam até 20% do tempo a corrigir discrepâncias.

Arquitetura minimalista

Lakehouse + Spark Notebooks + Warehouse
SQL + Pipelines. Protótipo viável em cerca de 4
semanas sem serviços externos.

Categorias de features

Estáticas, agregadas em batch e online. Ex.: purchases_30d_count recalculada a cada 30 minutos.

Materialização para serving

Tabelas otimizadas particionadas por user_id (p.ex. 256 partições) e cache ajustado para latências de consulta <100 ms.

Actualização incremental

Use Change Data Feed e MERGE; 20M eventos/dia podem ser agregados em 25–45 minutos num cluster médio (≈ 16 cores).

Fallbacks e caches

Fluxo típico: cache em memória (taxa de acerto 85–95%) → Warehouse (≤ 100 ms) → cálculo on-demand ou valor por defeito.

Operações e reprodutibilidade

Inclua manifestos de featurização, metadados de versão e checks (ex.: alertar se nulos $>2\%$ ou se job $>2\times$ tempo histórico).

CONTINUA NO BLOG

Por que um Feature Store leve é essencial para scoring em tempo real?

www.bconcepts.pt/insights