ANOMALIAS CONTÁBEIS · REGRAS × ESTATÍSTICA

Regras pegam 2 em cada 3 anomalias.
O resto exige estatística.

Projeto open source em Python, plano de contas COSIF e dados 100% fictícios.

RF

Roberto Fogaça

Contabilidade · Tributário · Tesouraria

Arraste →

01 · O RESULTADO

Mais recall, menos precision

SÓ REGRAS

Recall 66,7%

Precision 88,9% · 800 de 1.200 anomalias

REGRAS + ESTATÍSTICA

Recall 100%

Precision 63,0% · 1.200 de 1.200 anomalias

20.400 lançamentos e 12 tipos de anomalia injetadas, em dados fictícios.

Recall e precision

Duas perguntas diferentes sobre os mesmos alertas:

Recall

Das anomalias que existem, quantas o método encontrou.

Precision

Dos alertas gerados, quantos eram anomalia de verdade.

Ganhar recall custa precision. O score organiza a fila de revisão.

Quatro tipos sem regra

Anomalias que nenhuma regra fixa enxerga:

Valor atípico para a conta
z-score robusto (mediana e MAD).

Usuário fora do seu perfil
Lança em conta que quase nunca movimenta.

Fracionamento abaixo da alçada
Perfil de primeiros dígitos (variante de Benford).

Pico mensal de volume
Quantidade de lançamentos fora do histórico da conta.

Cada detector marca um sinal; o score soma os pesos.

De onde vêm as falhas

Em dados fictícios, com a causa conhecida:

Recall de 66,7% só com regras

4 dos 12 tipos não tinham regra: ela só enxerga o que foi escrita para procurar.

Precision de 88,9% só com regras

100 falsos alarmes, todos de duplicidade: a regra marca o par inteiro.

Precision de 63,0% com estatística

705 falsos alarmes; 456 vêm do pico mensal, que marca a conta no mês inteiro.

Nível médio ou alto: 98,3% de precision (9 falsos em 516 alertas).

Duas regras do plano de contas

Dígito verificador

Conferido nas 4.401 contas do elenco.

Conta fora da vigência

Existe no plano, mas não valia na data do lançamento.

Aviso honesto: o recall de 100% vem de anomalias que eu mesmo injetei.

VAMOS TROCAR IDEIAS?

Qual anomalia em lançamentos mais aparece no seu fechamento?

Deixe seu comentário no post.
Conta como é na sua rotina.

Salve este carrossel. Código e resultados: link nos comentários.