92 % Genauigkeit. Was die Zahl nicht zeigt.

FIKTIVES BEISPIEL: HINWEISE AUF SPINDELSTÖRUNGEN

Das Rechenbeispiel betrachtet ein Modell für frühzeitige Störungshinweise. **1.000 Maschinenwochen** stehen für zehn Bearbeitungszentren mit je 100 Wochen: eine Beurteilung am Wochenanfang, ein Warnhorizont von sieben Tagen. Die Werte sind erfunden und rückblickend berechnet; das Modell gab damals keine Hinweise im Betrieb aus.

FEHLERMATRIX (CONFUSION-MATRIX)

92 %

korrekte Zuordnungen

(8 + 912) / 1.000

Die Prozentzahl zeigt nicht, welche Fehler für die Handlung zählen.

Dokumentierter Befund in den folgenden 7 Tagen	Modell hätte gewarnt	Modell hätte nicht gewarnt
Störung dokumentiert	**8** erkannt	**2** übersehen
Keine Störung dokumentiert	**78** Warnung	**912** keine Warnung

Nur acht der 86 Warnungen stehen einer dokumentierten Störung gegenüber. Eine Warnung ohne dokumentierte Störung ist nicht automatisch unnötig: Frühere Wartung könnte das Ereignis verhindert haben. Zehn Störungsfälle belegen keine stabile Leistung für andere Einsätze.

99 % **wären hier sogar ohne jede Warnung möglich.**
Immer „keine Störung“ ergibt 990 von 1.000 korrekte Zuordnungen, erkennt aber keine der zehn Störungen. Das ist eine Rechenreferenz, nicht der heutige Prozess.

DIESELBE WARNUNG - DREI UNTERSCHIEDLICHE HANDLUNGEN

Prüfhinweis

Fachleute prüfen weitere Signale. Der Versuch muss Vorwarnzeit, Hinweisqualität und Prüfkapazität untersuchen.

Geplante Inspektion

Eine Inspektion kann aus fachlicher Prüfung folgen. Als feste Reaktion auf jede Warnung ist sie noch nicht begründet.

Automatischer Stopp

Aus den Beispielzahlen nicht freigabefähig. Fehlerfolgen und Verhalten bei Unsicherheit sind noch offen.

VOR DEM NÄCHSTEN SCHRITT PRÜFEN

Passt der Test zum Einsatz?

Repräsentative Fälle und rechtzeitig verfügbare Daten; keine Informationen aus der Zukunft.

Ist die Reaktion machbar?

Vorwarnzeit, Hinweisvolumen, Zuständigkeit und Zusatzaufwand im Ablauf prüfen.

Gibt es echten Zusatznutzen?

Mit dem heutigen Prozess vergleichen. Modelltreffer belegen noch keine Wirkung des Eingriffs.

ERST DIE HANDLUNG, DANN DIE METRIK

Ein begrenzter Test als Prüfhinweis kann sinnvoll sein, sofern Hinweise rechtzeitig eintreffen und sorgfältig bearbeitet werden können. **92 % allein sind keine Automatisierungsfreigabe.**

Grundlagen: Google ML Crash Course (Klassifikation) · NIST AI RMF 1.0, Abschnitt 3.1