

KI im Leasing

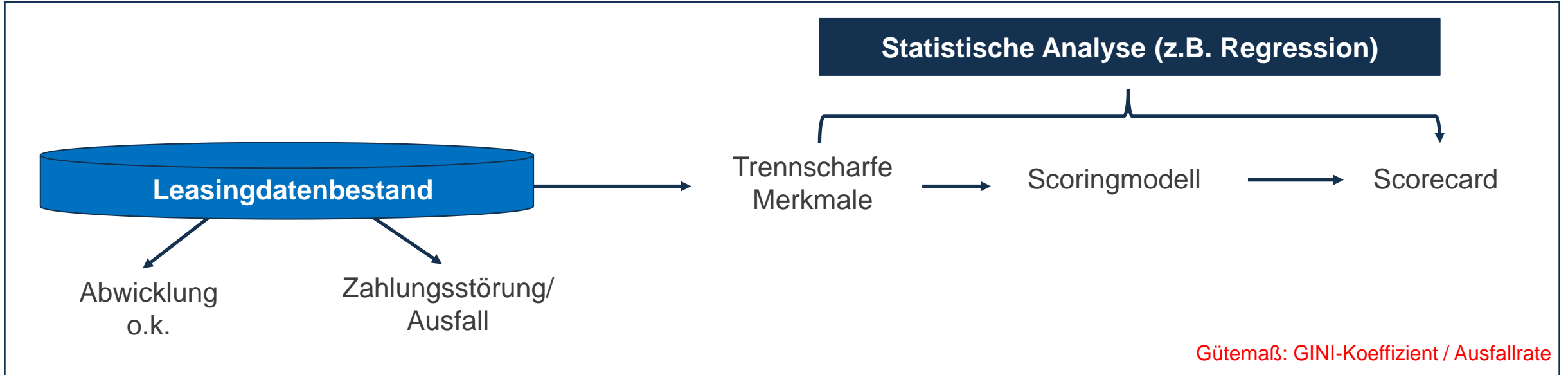
Wie verändert sich das Kreditantragsrating?

Prof. Dr. Matthias Schumann

Enabling success.

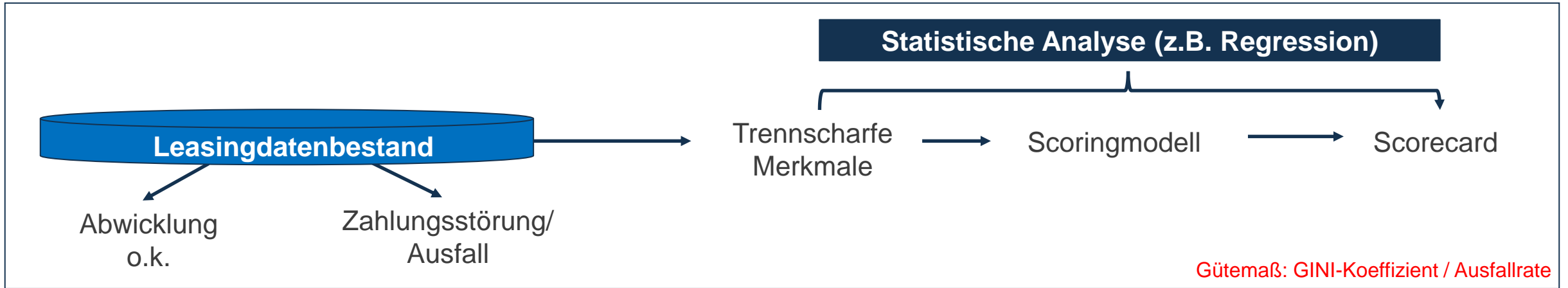
Entwicklung des Kredit Scorings

Klassisches Scoringverfahren

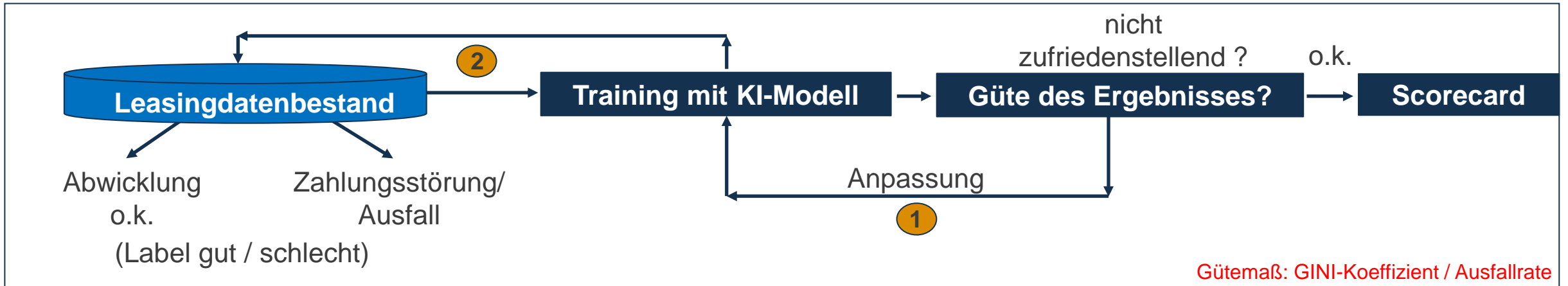


Entwicklung des Kredit Scorings

Klassisches Scoringverfahren



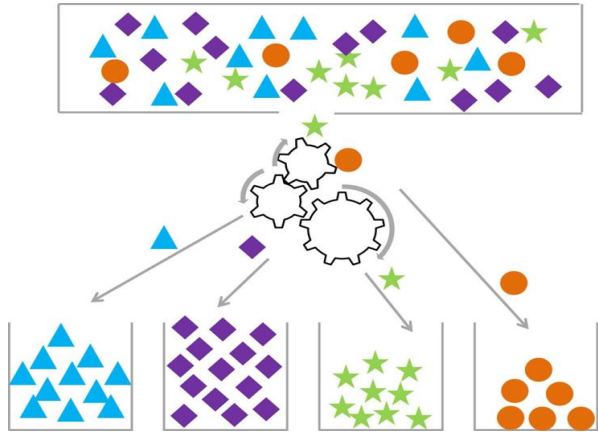
KI-basierte Scoringverfahren



KI-basierte Modellansätze für das Kredit scoring

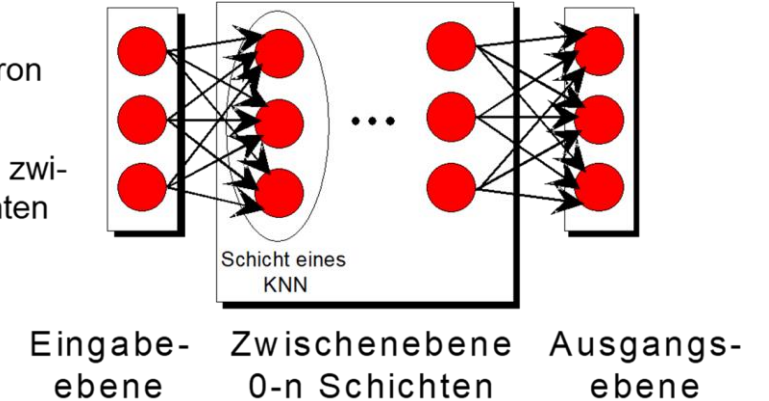
KI-basierte Scoringverfahren

Entscheidungsbaum-orientierte Verfahren



Neuronale Netzbasierte Verfahren (incl. Deep Learning)

● = Element, Neuron eines KNN
→ = Verbindungen zwischen Elementen



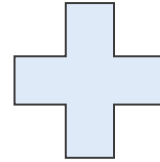
- Es werden große Trainingsdatenmengen benötigt
- Möglichst ausgeglichene positive und negative Datensätze im Trainingsdatenbestand (kein Bias)
 - reichen die „ausgefallenen Datensätze“? (Anreichern mit Ähnlichkeitsverfahren (Oversampling))
- Umgang mit fehlenden Werten in Datensätzen?

Daten des Privatkundenscorings

Haushaltsrechnung / frei verfügbares Einkommen

Scoringfaktoren

- Privatkundenauskunft
- Berufsgruppe
- Familienstand
- Dauer Tätigkeit letzter Arbeitgeber
- Dauer letzte Wohnadresse
- Alter
- Nationalität



Erweiterung der Datentypen

- Social Media-Daten
 - Aktivitäten
 - Themen
 - Soziales Umfeld
- LinkedIn & XING-Aktivitäten
- Scoring der Girokontotransaktionen

Scorekarten jeweils für unterschiedliche Objektklassen

Daten des Unternehmensscorings

Cash Flow-Rechnung / Kapitaldienstfähigkeit inkl. Prognose

Rating des Jahresabschlusses

- Liquiditätskennzahlen
- Eigen- und Fremdkapitalkennzahlen
- Erfolgskennzahlen
(zumeist zwischen 8 und 15 Kennzahlen)

Firmenauskunft

- Scorewert oder
- Einzelne Komponenten der
Auskunft individuell bewertet

Weitere Datentypen für das Scoring von Unternehmen

Nutzung von Berichtstexten/Anhängen von Jahresabschlüssen und Analystenreports

Ergebnisse von großzahligen Versuchen zeigen nur geringes Verbesserungspotenzial



- Linguistische Korpusanalyse (wir gehen davon aus...)
- Passive Sprache, sprachliches hedging, Sentiments, Komma-Rate
- Textlänge
- Zeitraum zwischen Publikationen

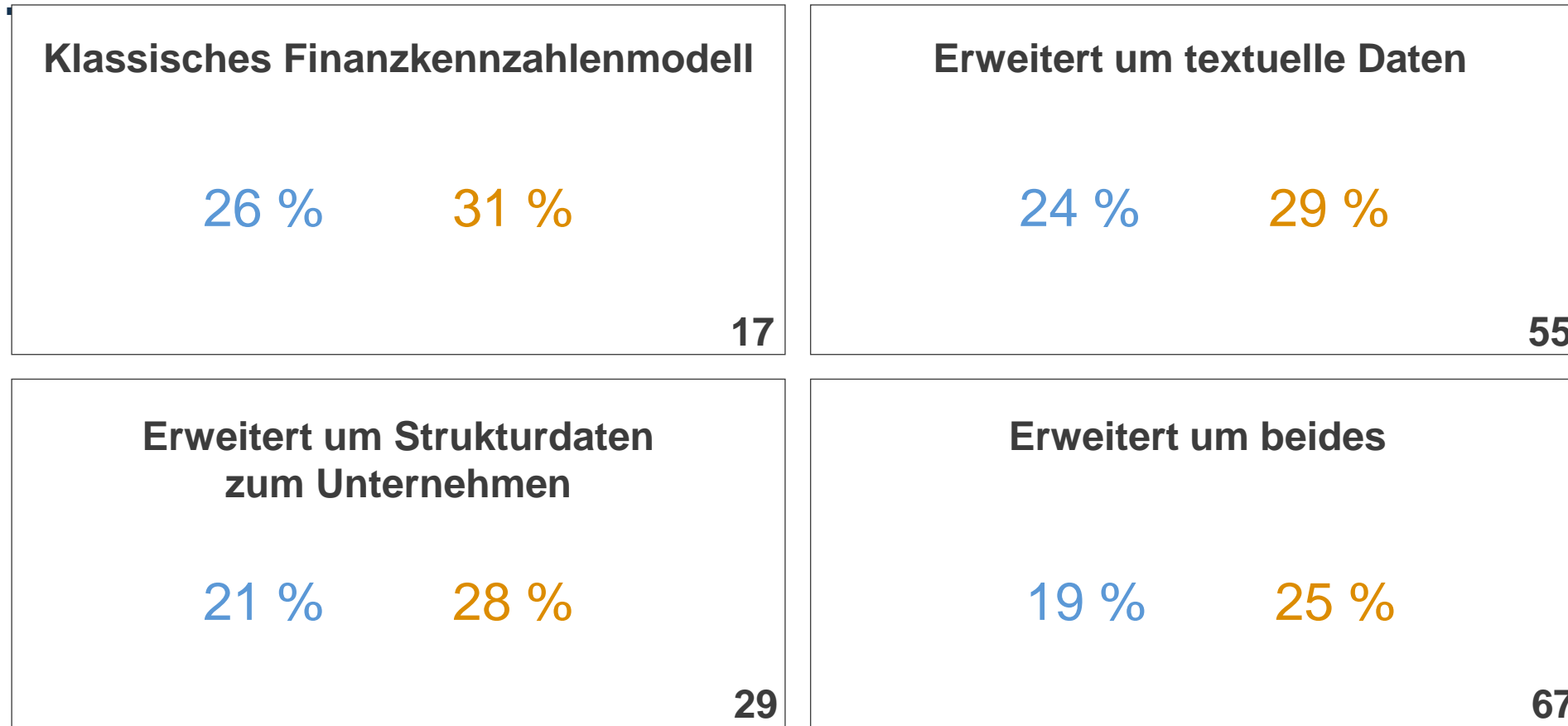
Strukturdaten zum Unternehmen

Welche Daten sind verfügbar?



- Integration von Daten aus Arbeitgeberbewertungsportalen (z. B. Kununu)
- Integration von Veränderungen auf LinkedIn
- Geschäftsführer-/Führungskreiswechsel (Häufigkeit..)
- Branchen- / Wettbewerbsintensität und -entwicklung
- Länderrisikoentwicklung

Potenziale unterschiedlicher Modelle – verschiedene



Fehler 1. Art (False-Positive) ■

Fehler 2. Art (False-Negative) ■

Anzahl der Variablen (N) ■

Ergebnisse beim Test KI-basierter Verfahren

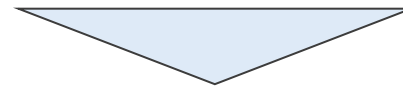
1

Training von KI-Modellen mit klassischen Datenbeständen für Scoring-Anwendungen zeigt, dass KI-Modelle bei der Trennschärfe kaum Verbesserungen bringen

2

Umfangreiches KI-basiertes Scoring von Jahresabschlüssen mit großem Kennzahlenpool

- Marginal verbessertes Scoring (Fehler)
- Ergebnisse verbessern sich bei dem Vorliegen von Zeitreihen



Häufig beste Ergebnisse mit XGBoost-Modellen (Ensemble-Learning)

These 1

Signifikante Verbesserungen der Trennschärfe sind von KI-Modellen nur zu erwarten, wenn bisher nicht verwendete Datentypen in die Modelle integriert werden

Rahmenbedingungen für KI-basierte Ansätze

BAFIN: Modelle können nachvollzogen werden

EU: Bei ablehnenden Entscheidungen im Privatkundenbereich muss die Ursache dafür erklärt werden

Entscheidungsbaumbasierte Verfahren

- Modelle einfach nachvollziehbar, Ergebnisse leicht erklärbar

Komplexe Neuronale Netze

- Erklärungsfähigkeit eingeschränkt
- Xplainable KI als Ansatz? → Methoden, die versuchen Modelle und/oder Ergebnisse auch von Neuronalen Netzen zu erklären

Selbst-lernende KI-Systeme?

1

Gini ↓ oder Ausfallrate ↑

- neues oder ergänzendes Training
- jüngere Datensätze werden starker gewichtet (kein Bias in Trainingsdaten!)

2

Kreditentscheider überstimmt den Vorschlag des Systems

- Liegen die zur Entscheidung verwendeten Daten digital vor?
- War die Entscheidung richtig?

➤ Daten können erst mit zeitlichem Abstand zum Nachlernen verwendet werden

➤ Evtl. ist zu entscheiden, ob die Datenbasis erweitert wird

These 2

Für selbstlernende Scoring-Systeme müssen komplexe Herausforderungen bewältigt werden!

Entscheidungsunterstützung durch KI

These 3

Im ersten Schritt werden Kreditentscheider mit KI-basierten Lösungen unterstützt, die Daten aus unterschiedlichen Internetquellen interpretieren (z. B. Kununu, LinkedIn, Branchenforen usw.)



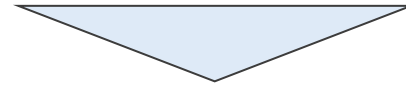
Es werden negative Entwicklungen aufgrund dieser Daten aufgezeigt (Alert-System)

Diese Daten werden (noch) nicht zur Entwicklung des Scoring-Modells herangezogen

Neue Gefahren durch KI?

Generative KI-Methoden erlauben das textuelle und Bild-gestützte Erstellen künstlicher Profile (Personen und Unternehmen)

Bots können genutzt werden, um unrichtige Informationen auf Internet-Portalen zu platzieren



These 4

Es sind verstärkt Systeme zu entwickeln, die es erlauben zwischen realen und fake-Daten zu unterscheiden

Zusammenfassung



Die KI wird für die Prozesse des Kredit-Scorings graduelle Veränderungen bringen.

Die dafür notwendigen Aktivitäten dürfen vom Aufwand nicht unterschätzt werden.

Disruptive Ergebnisse sind auch mittelfristig nicht zu erwarten.

Prof. Dr. Matthias Schumann

m.schumann@prof-schumann.de

www.prof-schumann.com

+49 551 38 3150

