

# Projektauftrag: Verlustprävention an Selbstbedienungskassen im Einzelhandel

---

## 1. Projektbezeichnung

**Titel:** Betrug an Selbstbedienungskassen

---

## 2. Problemstellung und Ziele

**Ausgangslage:** Immer mehr Einzelhändler setzen auf Selbstbedienungskassen (SBK), um Personalressourcen zu sparen. Gleichzeitig ist die Minimierung von Verlusten durch fehlerhaftes Scannen eine Herausforderung. Verluste für den Anbieter entstehen durch das bewusste Auslassen von Scanvorgängen, und durch Bedien- oder technische Fehler.

**Ziel** des Projekts ist es, einerseits die Zusammenhänge von Verlusten an Selbstbedienungskassen herauszuarbeiten (z.B. bestimmte Produkte, bei denen es besonders oft zu Fehlern bzw. Betrug kommt), andererseits konkrete Algorithmen mithilfe des maschinellen Lernens zu entwickeln, die diese Zusammenhänge in praktische Handlungsempfehlungen übersetzen. Dies können z.B. Algorithmen sein, die bei bestimmten Einkäufen manuelle Kontrollen anregen. Die Qualität der Algorithmen ergibt sich aus vom Kunden vorgegebene Bewertungsparametern.

---

## 3. Domänenspezifika

Im Einzelhandel entstehen jährlich erhebliche Verluste durch Inventurdifferenzen. SBK-Kassen sind besonders anfällig, z. B. durch:

- das Scannen günstiger Artikel statt teurerer („Bananen-Trick“),
- das komplette Auslassen von Artikeln,
- oder Systemfehler.

Die xxxxxxxx GmbH setzt bislang manuelle Stichprobenkontrollen ein, die keiner detaillierten Systematik folgen. Somit ist weder klar, ob die Kontrollen überhaupt zum gewünschten Ziel führen (die Verluste zu reduzieren), noch ist eine konkrete Auswirkung auf die Verlusthöhe bezifferbar. Eine Kosten-Nutzen-Rechnung kann damit nicht durchgeführt werden. Darüber hinaus ist es unwahrscheinlich, dass diese manuellen Kontrollen die Ressourcen optimal allokalieren und Verluste minimieren, selbst wenn sie ansatzweise zu positiven Ergebnissen führen.

---

## 4. Beteiligte und Stakeholder

### Projektgruppe:

- Raphael Schaffarczyk (Experte für statistische Datenanalyse)
- David Zurschmitten (Experte für Programmierung und Softwareentwicklung)
- Matthias Bald (Experte für Dokumentation & Projektkoordination)

**Projektgeber:** xxxxxxxx GmbH

### Betreuung:

- Prof. Dr. Christian Beecks (Lehrgebiet Data Science)
- Frau Sabine Folz-Weinstein
- Herr Max Pernklau

---

## 5. Projektorganisation inkl. Zeitplan mit den Meilensteinen

Das Team arbeitet iterativ nach dem DASC-PM-Modell.

Die Kommunikation der Gruppe erfolgt über eine zu diesem Zweck eingerichtete WhatsApp-Gruppe sowie über die Softwareentwicklungsplattform GitHub.

Es finden wöchentliche interne Abstimmungen über Zoom sowie zweiwöchentliche Meetings mit der Betreuung statt.

### Geplante Meilensteine:

Meilenstein	Verantwortlich	Frist
Projektskizze	Matthias	17.04.2025
Explorative Datenanalyse	Raphael	24.04.2025
Analyseergebnis	Raphael	
Modelle/Verfahren/Systeme	David	
Abschlusspräsentation	Alle Teilnehmer	08./09.07.2025

---

## 6. Ressourcen

### Kenntnisse im Team:

- Statistik / maschinelles Lernen: Tiefgehende Kenntnisse in Statistik (Regressionen, Inferenzstatistik, Clustering, neuronale Netze)
- Mathematische Optimierung: Grundlegende Kenntnisse in Funktionsoptimierung
- Softwareentwicklung: Tiefgehende Kenntnisse der Softwareentwicklung und der Erstellung von strukturierten Codes (inkl. Versionsverwaltung in Github)
- Dokumentation / Koordination: Gute Kenntnisse im Projektmanagement und in der Kommunikation mit Stakeholdern

#### **Verwendete technische Infrastruktur:**

- Python (für Datenauswertungen, Statistik und maschinelles Lernen)
- Docker Container (aus dem Kurs Data Engineering für Data Science) zur Entwicklung von Infrastruktur unabhängigen Codes
- Apache (mit seinen verschiedenen Produkten) für NoSQL Auswertungen und Datenaufbereitungen / -vorverarbeitungen
- Github (für Dokumentenmanagement und Codeversionierung)
- Microsoft Office (für Dokumentation in Word und Erstellung von PowerPoint Präsentationen)

---

### **7. Risiken**

- Gelabelte Daten könnten nicht repräsentativ für den gesamten Datensatz sein
- Übertragbarkeit der Analyseergebnisse auf andere Filialen ggf. nicht gegeben
- Berücksichtigung externe Auflagen (gesetzliche Vorschriften, Gesellschaftsvertrag etc.) nicht abgedeckt
- Beeinflussung durch vom Unternehmen abgeschlossene (Diebstahl)- Versicherungen könne die Bewertungsfunktion ungeeignet machen für diesen konkreten Fall
- Technische Umsetzung der konkreten Algorithmen muss an den Selbstbedienungskassen geprüft werden

---

### **8. Vorerfahrungen**

**[Sollen hier wirklich noch einmal die Vorerfahrungen von uns rein? Das steht ja schon oben.]**

Alle Mitglieder verfügen über relevante Vorerfahrung in den Bereichen Datenanalyse, Python und Machine Learning.

- Raphael bringt vertiefte Kenntnisse in Mathematik und Statistik mit,
- David Erfahrung in Modellierung und Softwareentwicklung,
- Matthias verfügt über einen betriebswirtschaftlichen Hintergrund.