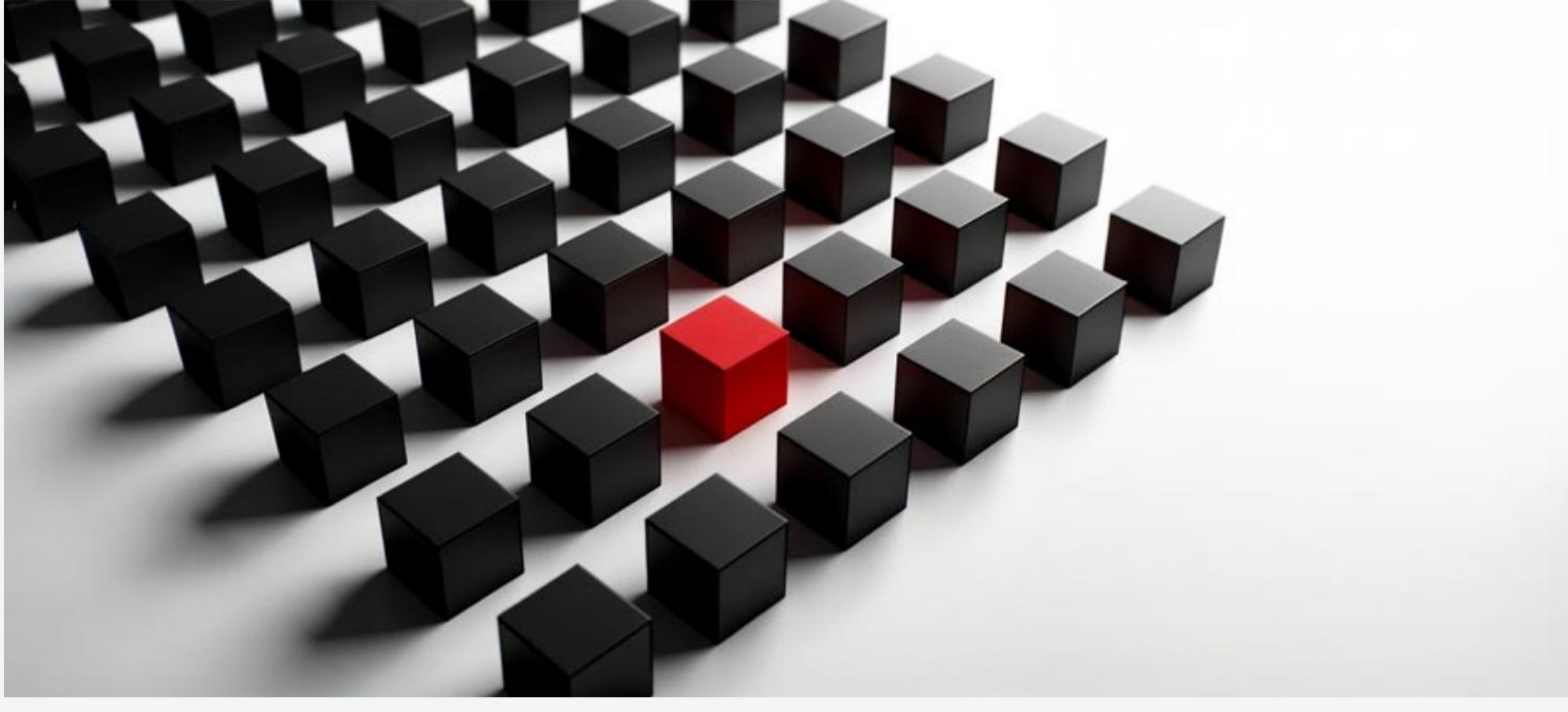


Erkennung von Anomalien in der industriellen Fertigung

ein Supper & Supper Use Case



Projektziel

Ziel dieses Projektes war es, anomales Verhalten sowohl in Maschinen als auch in mechanischen und industriellen Anlagen zu erkennen, ohne vorher zu wissen, wo das anomale Verhalten auftritt. Anomalieerkennung, im Englischen „Anomaly Detection“, birgt ein großes Potenzial für verschiedene Unternehmen mit industriellen Fertigungs- und Produktionsprozessen, um diese signifikant voranzubringen. In diesem Projekt haben wir modernste Werkzeuge des maschinellen Lernens eingesetzt, um eine präzise Anomalieerkennung in industriellen Prozessen zu gewährleisten, inklusive einer frühzeitigen Identifizierung dieser anomalen Abweichungen.

Was sind Anomalien in Fertigungsprozessen?

Anomalien in Fertigungs- und Produktionsprozessen beziehen sich auf Abweichungen im Betrieb eines (fertigungstechnischen/industriellen/technischen) Systems von seinem beabsichtigten oder normalen Ablauf. Solche Abweichungen können die Leistung beeinträchtigen und zu Instabilitäten, Sicherheitsproblemen und sogar Systemausfällen führen. Angesichts der komplexen Dynamik dieser Systeme kann es schwierig sein, die Ursachen solcher Anomalien zu identifizieren.

Genutzte Daten

Die Daten für die Anomalieerkennung waren multivariate Zeitreihen, die von Sensoren an Maschinenprüfständen erfasst wurden. Gesammelt wurden die anomalen Daten in Experimenten, bei denen das System über bestimmte Zeiträume hinweg absichtlich manipuliert wurde. Die verwendeten Algorithmen funktionieren äquivalent, wenn ausschließlich Daten aus dem normalen Ablauf des Systems zur Verfügung stehen. In diesem Fall können Anomalien in den Datensätzen synthetisch erzeugt werden, um die Anomalieerkennung zu testen. Der Trainingsdatensatz bestand aus 8.125 Datenpunkten, von denen 337 anomal waren.

Wenn bekannt ist, welche Datenpunkte Anomalien sind, können wir diese Information nutzen und Methoden des (semi-)überwachten Lernens einsetzen. In den meisten Anwendungen ist diese Information jedoch nicht verfügbar, sodass unüberwachtes Lernen angewandt werden muss. In diesem Projekt haben wir Methoden des unüberwachten Lernens verwendet. Bei der Überprüfung der Anomalieerkennung wurden die Informationen darüber, welche Datenpunkte anomal sind, ausschließlich zur Überprüfung der Leistungsfähigkeit und nicht zum Training des Modells verwendet.

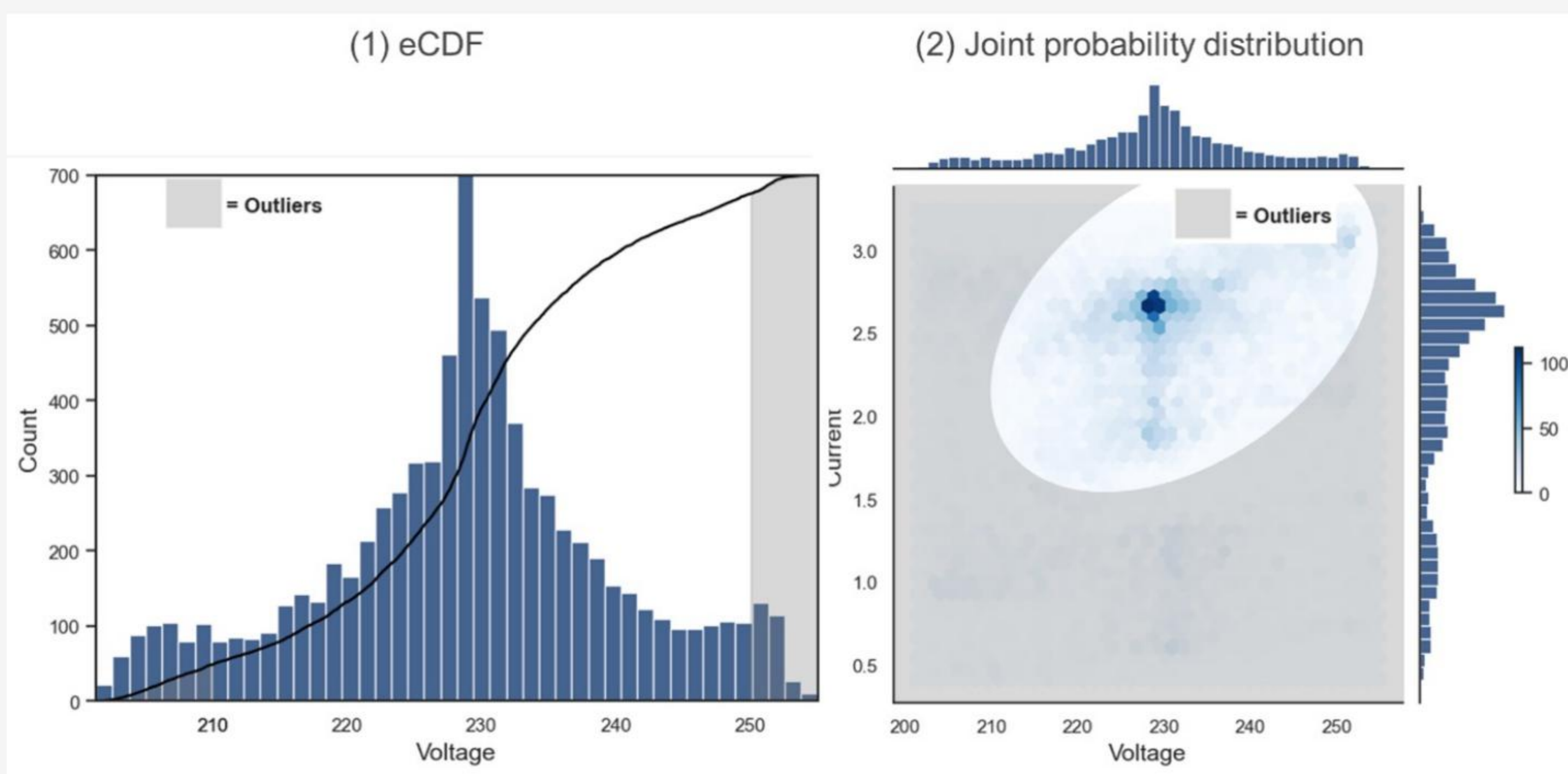
Herausforderungen

1. Eine große Herausforderung bestand darin, dass die anomalen Datenpunkte nur einen sehr kleinen Teil der Trainingsmenge (<5 %) ausgemacht haben. Entsprechend gab es nur wenige Informationen, aus denen man lernen konnte.
2. Darüber hinaus musste bei der Wahl des besten Algorithmus ein gewisser Kompromiss zwischen der Anzahl der falsch positiven (falsch erkannten Anomalien) und falsch negativen (übersehene Anomalien) Ergebnisse eingegangen werden. In diesem Projekt sind, wie bei vielen anderen industriellen Anwendungen auch, die Kosten einer übersehenen Anomalie höher als die eines Fehlalarms. Daher muss die optimale Methode die Anzahl der falsch negativen Meldungen minimieren und gleichzeitig eine gute Gesamtgenauigkeit bieten.
3. Schließlich sind viele etablierte Methoden des maschinellen Lernens sehr rechenintensiv und zeitaufwendig. Die optimale Methode sollte effizient und leichtgewichtig sein, um auf eingebetteten Geräten bzw. Edge-Geräten ausgeführt werden zu können.

Angewandte Methoden (Umsetzung)

Wir haben etablierte Methoden des maschinellen Lernens zur Anomalieerkennung (OneClassSVM, iForest) mit modernen Modellen verglichen, welche die oben genannten Herausforderungen adressieren (ECOD, COPOD).

- OneClass SVM: erkennt Anomalien durch das Lernen von Klassengrenzen, die Datenpunkte in anomale und nicht-anomal gruppieren
- Isolation Forest (iForest): erkennt Anomalien mithilfe von binären Entscheidungsbäumen
- ECOD: erkennt Anomalien mit Hilfe von empirischen kumulativen Verteilungsfunktionen (eCDFs)
- COPOD: erkennt Anomalien unter Verwendung empirischer Copulas, um multivariate Wahrscheinlichkeitsverteilungen zu erhalten



Statistical concepts behind (1) ECOD and (2) COPOD

Zur Bewertung der Algorithmen verwendeten wir:

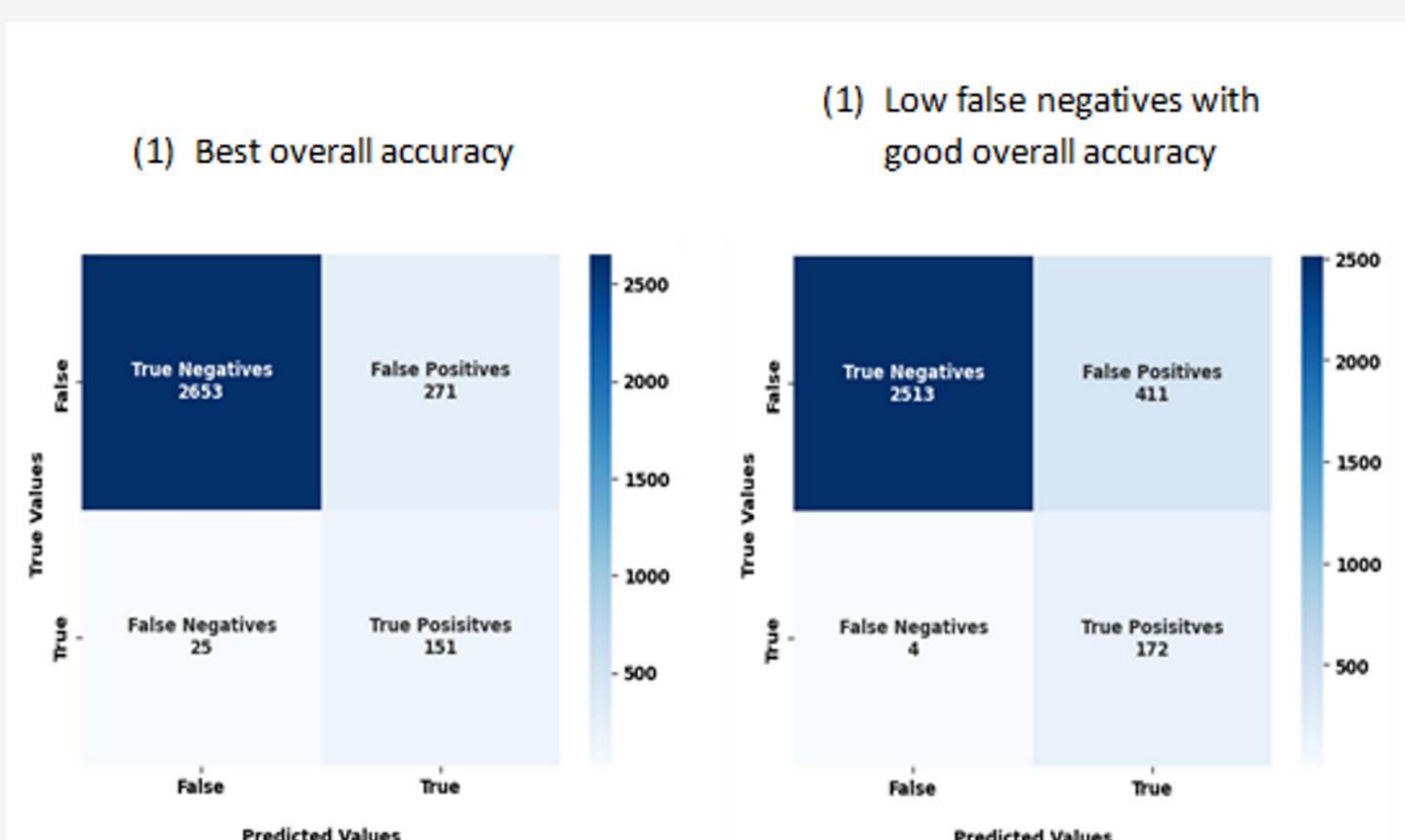
- die Missing Alarm (fehlender Alarm) Rate MAR = übersehene Anomalien / alle Anomalien
- Die False Alarm (Falschalarm) Rate FAR = falsch erkannte Anomalien / alle Nicht-Anomalien
- die F1 macro score: ein Maß für die Gesamtgenauigkeit des Modells

	FAR (%)	MAR (%)	F1 macro score
COPOD (F1 maximized)	9.3	14.2	0.73
COPOD (MAR minimized)	14.1	2.3	0.69
ECOD	12.7	8.0	0.69
iForest	12.1	16.5	0.68
OneClassSVM	13.7	38.6	0.62

Ergebnisse des Projektes

Die Modelle ECOD und COPOD haben in Bezug auf die Gesamtgenauigkeit, MAR und FAR am besten abgeschnitten. Sie sind leichtgewichtig sowie effizient und können daher auf eingebetteten Geräten ausgeführt werden.

- Das COPOD-Modell wurde auf (1) Gesamtgenauigkeit oder (2) geringe Anzahl an falsch negativen Ergebnissen (bei guter Gesamtgenauigkeit) optimiert.
- Das auf wenige falsch negative Ergebnisse optimierte Modell übersah nur 2,3 % der Anomalien (MAR) und klassifizierte 14,1 % aller nicht anomalen Datenpunkte fälschlicherweise als Anomalien (FAR).
- Das auf Gesamtgenauigkeit optimierte Modell übersah 14,2 % der Anomalien (MAR) und klassifizierte 9,3 % der Nicht-Anomalien falsch (FAR).



Kategorie

- Mechanical Engineering
- Predictive Maintenance

Download



Social Sharing

