## Glossar The Elephant in the Room: Towards A Reliable Time-Series Anomaly Detection Benchmark

|  |  |  |
| --- | --- | --- |
| **Begriff** | **Bedeutung** | **Erklärung** |
| Time Series | Zeitreihe | Geordnete Folge von Datenpunkten in aufeinanderfolgenden Zeitintervallen |
| Anomaly Detection | Anomalieerkennung | Datenpunkte die nicht das erwartete Muster oder Verhalten zeigen |
| Benchmarking | Leistungsbewertung | Systematischer Vergleich von Methoden oder Systemen |
| TSB-AD | Time-Series Benchmark for Anomaly Detection | Benchmark, der drauf abzielt, bestehende Probleme in der Anomalieerkennung von Zeitreihen zu adressieren |
| Dataset Integrity | Datensatzintegrität | Qualität und Zuverlässigkeit eines Datensatzes |
| Metric Reliability | Metrikzuverlässigkeit | Maß, für konsistente und genaue Ergebnisse zur Bewertung von Anomalien |
| VUS-PR | Volume Under the Surface – Precision-Recall | Metrik zur Bewertung der Leistung von Anomalieerkennungsmodellen |
| IoT | Internet of Things | Netzwert aus physischen Gräten, Fahrzeugen,... die mit Sensoren, Software und andern Technologien ausgestattet sind, um Daten zu sammeln und auszutauschen |
| SOTA methods | State-of-the-Art Methods | Fortschrittlichsten und leistungsfähigsten Methoden oder Algorithmen die in einem bestimmten Forschungsfeld verfügbar sind. Z. B. Methoden, die auf deep-Learning oder statistischen Modellen basieren |
| UCR dataset |  | Sammlung von Zeitreihendatensätzen von der University of California (Riverside). Standard-Benchmark für verschiedene Aufgaben der Zeitreihenanalyse |
| AUC-ROC | Area Under the Curve – Receiver Operating Characteristic | Metrik zur Bewertung der Leistung eines binären Klassifikators. Sensitivität (TPR) gegen Spezifität (1-Falsch Positive Rate) aufgetragen. Hohe AUC-ROC = bessere Trennung zwischen Klassen |
| TPR | True Positiv Rate | Anteil der tatsächlich positiven Ergebnisse geteilt durch die als korrekt positiv erkannt wurden |
| AUC-PR | Area Under the Curve -Precision-Recall | Metrik zur Bewertung eines Modells bei stark unausgeglichenen Datensätzen. Misst Fläche unter er Precision-Recall-Kurve. Stellt Zusammenhang zwischen Präzision und Sensitivität dar |
| BUS-PR | Balanced Utility Surface – Presision-Recall | Metrik, die die Ausgewogenheit zwischen Precision und Recall in einem mehrdimensionalen Kontext berücksichtigt. Robuste Bewertung von Anomalieerkennungsmodelle. Minimiert Verzerrungen in unbalancierten Datensätzen |
| Hyperparameter-Tuning |  | Anpassen der Parameter eines Modells, die nicht während des Trainings gelernt werden |
| Foundation Model |  | Vortrainierte Modelle, die auf spezifische Aufgaben angepasst werden können |
| Data Contamination | Datenkontamination | Unbeabsichtigte Einfließen von Informationen aus dem Testdatensatz in den Trainingsdatensatz, führt zu überoptimisitschen Leistungsschätzungen |
| Unsupervised Learning | Unüberwachtes Lernen | Maschinelles Lernverfahren, Modelle finden Muster in Daten, ohne vorherige Labels oder Kategorien |
| IForest | Isolation Forest | Algorithmus zu Anomalieerkennung, identifiziert Anomalien indem er Datenpunkte isoliert |
| Chronos |  | Framework oder Tool zur Zeitreihenanalyse und -vorhersage zur Anomalieerkennungsaufgabe |
| MOMENT |  | Modell zur Anomalieerkennung, welches in der Implementierung des TSB-AD verwendet wird |
| TSAD | Time Series Anomaly Detection | Identifizierung von ungewöhnlichen Mustern |
| Ground Truth |  | Tatsächliche Labels oder Fakten, die verwendet werden, um Leistung eines Modells zu bewerten. Oft manuell durch Experten annotiert |
| Precision |  | Anteil der korrekt vorhergesagten positiven Ergebnisse an allen vorhergesagten positiven Ergebnissen |
| Recall |  | Tatsächlich positive Fälle, die vom Modell korrekt erkannt wurden |
| FPR | False Positive Rate | Anteil der negativ gelabelten Instanzen, die fälschlicherweise als positiv klassifiziert wurde |
| Overfitting |  | Trainingsdaten wurden zu genau gelernt, dadurch entsteht schlechte Generalisierung |
| Multivariate Time Series |  | Datensätze mit mehreren Variablen oder Dimensionen, die häufig miteinander korreliert sind |
| Window Size |  | Lände des Zeitfensters die für einen Analyse einer Zeitreihe verwendet wird |
| Sysnthetic Data |  | Künstlich erzeugte Datensätze, um Modelle zu trainieren oder zu bewerten, wenn reale Daten unzureichend oder schwierig zu handhaben sind |
| Class Imbalance |  | Problem bei dem die Anzahl der normalen Ereignisse viel größer ist als die Anzahl der Anomalien |
| Sliding Window |  | Technik in der ein Algorithmus in kleine, sich überlappende Zeitfenstern analysiert wird, um Anomalien zu erkennen |
| Evaluation Bias |  | Verzerrung in der Bewertung eines Modells, durch Datenkontamination oder unzureichende Metriken |
| Robustness |  | Modell kann gut auf unbekannten oder verrauschten Daten performen ohne signifikanten Leistungsabfall |
| Interpretability |  | Wie Entscheidungen oder Vorhersagen eines Modells verstanden und erklärt werden können |
| Dynamic Thresholding |  | Verwendung von sich anpassenden Schwellwerten anstelle von festen werten, um Anomalie in Zeitreihen zu erkennen |
| Latency |  | Verzögerung zwischen Auftreten einer Anomalie und der Erkennung durch ein Modell. Wichtiger Faktor für Echtzeitanwendungen |