



Data Analytics
Stiftung Universität Hildesheim
Marienburger Platz 22
31141 Hildesheim
Prof. Dr. Dr. Lars Schmidt-Thieme

Thesis

Unsupervised Real-Time Time-Series Anomaly Detection

Abdul Rehman Liaqat
271336, Liaqat@uni-hildesheim.de

Abstract

Anomaly detection is a crucial task for machine learning due to wide-spread usage and type. In particular, it is worth noting that most data arising in industrial setups are of a streaming nature, thus restricting the range of standard anomaly detection tools. This thesis will identify the potential approaches to learn the identification of abnormal behavior from large-scale streaming data. An empirical comparison of state-of-the-art methods will to be extended by a novel technical contribution. In this thesis, the focus is particularly on streaming time-series Anomaly Detection which changes in nature with time and novel contribution will especially try to target this dynamic nature of time-series.

Contents

1	Introduction	4
1.1	Motivation	5
1.2	Objective	6
2	Related Work	7
3	Unsupervised Anomaly detection with recency	8
4	Experiments	9
4.1	Data	10
4.1.1	Numenta Anomaly Benchmark (NAb)	10
5	Execution and Results	11
6	Discussion	12
7	Experiment Infrastructure	13
7.1	Experiment Management using MLflow	14
7.2	Parallel execution using Docker	15
8	Best practices	16
9	Reference Usage	17
10	References	18

1 Introduction

1.1 Motivation

1.2 Objective

2 Related Work

3 Unsupervised Anomaly detection with recency

4 Experiments

4.1 Data

4.1.1 Numenta Anomaly Benchmark (NAb)

5 Execution and Results

6 Discussion

7 Experiment Infrastructure

7.1 Experiment Management using MLflow

7.2 Parallel execution using Docker

8 Best practices

Following steps were taken to maximize the efficiency and speed of research:

1. Use version control to track the code and share between different devices.
- 2.
- 3.
- 4.

9 Reference Usage

10 References

,