



Universiteit
Leiden

The Netherlands

Unsupervised representation learning for time series anomaly detection and beyond

Ferreira Correia; L.

Citation

Unsupervised representation learning for time series anomaly detection and beyond. (2026, September 8). *Unsupervised representation learning for time series anomaly detection and beyond*. Retrieved from <https://hdl.handle.net/1887/4309435>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4309435>

Note: To cite this publication please use the final published version (if applicable).

Unsupervised Representation Learning for Time Series Anomaly Detection and Beyond

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 8 september 2026
klokke 14.30 uur

door

Lucas Ferreira Correia

geboren te Braga, Portugal

in 1998

Promotor:

Prof. dr. T.H.W. Bäck

Co-promotor:

Dr. A.V. Kononova

Promotiecomissie:

Prof. dr. M.M. Bonsangue

Prof. dr. A. Plaat

Dr. Y. Fan

Dr. N. van Stein

Prof. dr. K.M. Malan (University of South Africa)

Prof. dr. B. Sendhoff (Technical University Darmstadt)

Copyright © 2026 *Lucas Ferreira Correia*

Cover: *Naila Lakehsar, The JetBrains Mono Project Authors*

Contents

1	Introduction	1
1.1	Background	1
1.2	Research Questions	1
1.3	Structure of this Thesis and Publications	3
2	Background and Theory	7
2.1	Time Series Theory	7
2.2	Anomaly Detection	8
2.2.1	Anomaly Types	9
2.2.2	Concept of Supervision	12
2.2.3	Continuous- and Discrete-sequence Problems	12
2.2.4	Online Anomaly Detection	13
2.3	Long Short-term Memory Cells	13
2.4	Variational Autoencoders	16
2.5	Multi-head Attention	17
2.6	Active Learning	18
2.7	Dynamic Time Warping	19
3	Related Work	21
3.1	Benchmarking Time Series Anomaly Detection Methods	21
3.1.1	Datasets	22
3.1.2	Evaluation Metrics	26
3.1.3	End-to-end Benchmarking Frameworks	37
3.2	Navigating Through the Time Series Anomaly Detection Landscape	38
4	Datasets	43
4.1	Simulation-based Publicly Available Dataset	43

4.1.1	Simulation Model	44
4.1.2	Data Set Generation	48
4.1.3	Usage of the Data Set	57
4.1.4	Preliminary Analysis	58
4.2	Real-world Application: Endurance Testing	62
4.2.1	Automotive Powertrain Testing Setup	62
4.2.2	Data Set	63
4.2.3	Preliminary Analysis	66
5	Online Multivariate Time Series Anomaly Detection	69
5.1	Introduction	69
5.2	Proposed Methodology	70
5.2.1	Temporal Variational Autoencoder	70
5.2.2	Bypass Phenomenon	73
5.2.3	Mapping Fixed-length Windows to Variable-length Sequences .	74
5.2.4	Generalisation Capabilities	77
5.2.5	Experimental Setup	77
5.3	Results	81
5.3.1	Online Anomaly Detection	81
5.3.2	Ablation Study for Bypass Phenomenon	85
5.3.3	Impact of Reverse-window Method	86
5.3.4	Generalisation Study	87
5.4	Conclusions	89
6	Interactive Threshold Search Using Active Learning	91
6.1	Introduction	91
6.2	Related Work	92
6.3	Proposed Methodology	94
6.3.1	Dissimilarity-based Querying Strategy	94
6.3.2	Experimental Setup	96
6.4	Results	98
6.4.1	Baselines without Active Learning	98
6.4.2	Impact of Query Budgets	99
6.4.3	Impact of Mislabelling Probability	101
6.5	Conclusions	102

7	Extending Representation Learning to Predictive Maintenance	105
7.1	Introduction	105
7.2	Experimental Setup	106
7.3	Results	108
7.4	Conclusions	109
8	Conclusions and Outlook	111
	Repositories	125
	List of Abbreviations	127
	Summary	129
	Samenvatting	133
	Acknowledgments	137
	About the Author	139

