Gaussian Processes for Anomaly Description in Production Environments

Christian Beecks
University of Münster and Fraunhofer Institute for Applied Information Technology FIT, Germany
christian.beecks@uni-muenster.de

Kjeld Willy Schmidt
University of Münster, Germany
kjeld.schmidt@uni-muenster.de

Fabian Berns
University of Münster, Germany
fabian.berns@uni-muenster.de

Alexander Grass
Fraunhofer Institute for Applied Information Technology FIT, Germany
alexander.grass@fit.fraunhofer.de

ABSTRACT

Concomitant with the rapid spread of cyber-physical systems and the advancement of technologies from the Internet of Things, many modern production environments are characterized by vast amounts of sensor data which are generated throughout different stages of production processes. In this paper, we propose a novel method for discovering the inherent structures of anomalies arising in IoT sensor data. Our idea consists in modeling and describing anomalies by means of kernel expressions, which are combinations of well-known kernels. The results of our empirical analysis show that our proposal is suitable for modeling differently structured anomalies. Moreover, the results indicate that Gaussian processes provide a powerful tool for future algorithmic investigations of IoT sensor data.

1 INTRODUCTION

Concomitant with the rapid spread of cyber-physical systems and the advancement of technologies from the Internet of Things (IoT), many modern production environments are characterized by vast amounts of sensor data which are generated throughout different stages of production processes. These sensor data streams are often considered as valuable information sources with a high economic potential and are characterized by high volume, velocity and variety. Their data-driven value is indisputable for optimizing and fine-tuning industrial production processes.

Monitoring sensor data from complex production processes in order to detect outliers or low-performing production behavior caused by undesired drifts and trends, which we summarize as *anomalies*, is a challenging task. Not only due to the massive amount of sensor data but also due to different types of anomalies, which are potentially unknown in advance, manual or automatic inspection systems are frequently supported by anomaly detection algorithms. While the last years have witnessed the development of different anomaly detection algorithms, cf. the work of Renaudie et al. [21] for a recent performance evaluation in an industrial context, only less effort has been spent to the investigation of the inherent structure of an anomaly.

In this paper, we thus propose a novel method to discover the inherent structure of an anomaly. Our idea consists in modeling and describing anomalies by means of kernel expressions, which are combinations of well-known kernels. By fitting kernel expressions to the corresponding sensor data, we are able to decompose the inherent structure of an anomaly and to describe its individual behavior such as linearity and periodicity by natural language. For this purpose, we make use of Gaussian processes [20] and the Compositional Kernel Search model [11]. We carry out our analysis on the recently proposed IoT dataset [5], a real-world industry 4.0 dataset, which has been collected within the EU project MONSOON[1]. To sum up, we make the following contributions:

- We propose a machine-learning-based method in order to model anomalies and to describe their inherent components.
- We enrich the MONSOON IoT dataset with a novel ground truth derived from domain experts in order to further stimulate research of anomaly detection algorithms on this real-world dataset.

The paper is structured as follows. In Section 2, we outline related work. In Section 3, we briefly introduce Gaussian processes and their application to adapt kernel expressions to sensor data. The preliminary results of our proposed method are reported and discussed in Section 4, before we conclude our paper with an outlook on future research directions in Section 5.

2 RELATED WORK

Strongly related to our approach are anomaly detection algorithms. There is a plethora of these algorithms including Z-Score [10], Mahalanobis Distance-Based, Empirical Covariance Estimation [18] [9], Mahalanobis Distance-Based, Robust Covariance Estimation [22] [9], Subspace-based PCA Anomaly Detector [9], One-Class SVM [23] [18] [9] [12], Isolation Forest (I-Forest) [16] [18], Gaussian Mixture Model [18] [9] [19], Deep Auto-Encoder [8], Local Outlier Factor [7] [18] [9] [1], Least Squares Anomaly Detector [24], GADPL [14] and k-nearest Neighbour [13] [1] [12].

While these algorithms are all possible options for anomaly detection, as shown in different surveys such as [13], [19] and [9], they are not directly suited for describing the inherent structure of anomalies, which is the major focus of this paper. We choose the means of Gaussian processes for anomaly description due to their capability to not only gather statistical indicators, but deliver the very characteristics of specific anomalous behavior from the data [20].

For describing these characteristics, Lloyd et al. [17] have proposed the Automatic Bayesian Covariance Discovery System that adapts the Compositional Kernel Search Algorithm [11] by

First International Workshop on Data Science for Industry 4.0.

[1] www.spire2030.eu/monsoon

Published in the Workshop Proceedings of the EDBT/ICDT 2019 Joint Conference (March 26, 2019, Lisbon, Portugal) on CEUR-WS.org.

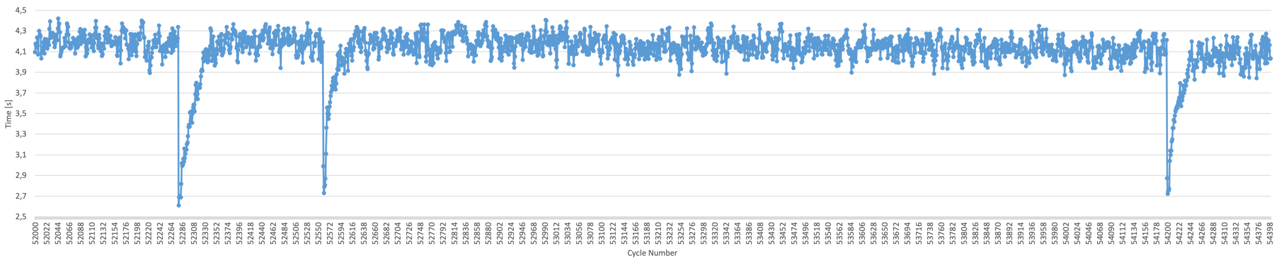

Figure 1: An example of the MONSOON IoT dataset with three anomalies.

adding intuitive natural language descriptions of the function classes described by their models. In [15], these models are expanded to discover kernel structures which are able to explain multiple time series at once.

In this work, we make use of these algorithms in order to describe the inherent structures of anomalies, as shown in the following section.

3 GAUSSIAN PROCESSES

In this section, we describe the analysis of anomalies in sensor data via Gaussian processes. To this end, we assume the sensor data to be univariate[2] and an anomaly A to be a finite subsequence of timestamp-value pairs $A = \{(t_i, v_i)\}_{i=i}^{n}$ with timestamps $t_i \in \mathbb{T}$ and values $v_i \in \mathbb{R}$.

As we do not know in advance the number of values and the distances between individual timestamps, we can also thought of an anomaly A as a mathematical function $A : \mathbb{T} \to \mathbb{R}$, which assigns every timestamp $t \in \mathbb{T}$ a real-valued value $v(t) \in \mathbb{R}$. By considering the individual values $v(t)$ to be random variables following a Gaussian distribution, we can formalize the Gaussian process as

$$v(t) \sim GP(m(t), k(t, t')),$$

where $m(t) = \mathbb{E}[v(t)]$ is the mean function and $k(t, t') = \mathbb{E}[(v(t) - m(t)) \cdot (v(t') - m(t'))]$ is the covariance function $k : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$. In other words, a Gaussian process is a stochastic process over random variables, where every subset of random variables from the Gaussian process follows a normal distribution. The distribution of the Gaussian process is the joint distribution of all of these random variables and it is thus a probability distribution over (the space of) functions in $\mathbb{R}^{\mathbb{T}}$.

While the covariance function k defined above is a general way to model the behavior of data, we aim to describe each anomaly A by its own covariance function k_A. That is, we aim to learn a covariance function k_A, which is then also denoted as kernel expression in the domain of machine learning, by fitting combinations of well-known kernels, such as

- the constant kernel $k_{\mathrm{C}}(t, t') = \lambda \in \mathbb{R}$,
- the linear kernel $k_{\mathrm{LIN}}(t, t') = (t - l) \cdot (t' - l)$,
- the squared exponential kernel $k_{\mathrm{SE}}(t, t') = \exp -\frac{|t-t'|^2}{2l^2}$,
- or the periodic kernel $k_{\mathrm{PER}}(t, t') = \exp \frac{2 \sin^2 \frac{t-t'}{2}}{l^2}$.

In order to individually fit a kernel expression to each anomaly based on the aforementioned kernels, we use the compositional kernel model, as utilized for instance in [17]. This allows us to decompose an anomaly into individual components, which can be ranked by their contribution towards explaining the data. As an example, an anomaly A with a highly weighted linear kernel k_{LIN} indicates a hidden linearity component while a highly weighted periodic kernel k_{PER} indicates an inherent periodicity in the anomaly.

The resulting kernel expressions are reported and discussed in the next section.

[2] It is noteworthy that this approach also applies to multivariate data.

Anomaly	BIC	Kernel Expression
0	-799	C*PER + C*PER + C*PER
1	-706	C*SE*PER + C*SE + C
2	-604	C*PER + C*PER + C*PER + C
3	-921	C*SE*PER + C*PER + C
4	-742	C*PER + C*PER + C*SE + C
5	-543	C*SE*LIN + C*SE + C*WN + C
6	-630	C*PER + C*SE + C*WN + C
7	-1020	C*PER + C*PER + C*PER + C*SE + C
8	-762	C*SE*PER + C*PER + C
9	-1025	C*PER + C*PER + C*SE + C
10	-424	C*PER + C*SE + C*SE
11	-849	C*PER + C*PER + C*SE + C
12	-311	C*SE*PER + C*PER + C
13	-860	C*LIN + C*PER + C*PER + C*PER + C
14	-339	C*PER + C*SE + C*SE
15	-590	C*SE*PER + C*PER + C*SE
16	-503	C*PER + C*SE + C
17	-602	C*SE*PER + C*SE + C*WN + C
18	-545	C*PER + C*SE + C*SE + C
19	-804	C*PER + C*SE + C*WN + C
20	-281	C*PER + C*SE + C*SE
21	-426	C*PER + C*PER + C*SE
22	-425	C*SE*PER + C*PER + C*SE
23	-975	C*SE*PER + C*PER + C
24	-1181	C*PER*LIN + C*PER + C*SE
25	-880	C*PER*PER + C*PER + C*PER + C
26	-455	C*PER + C*PER + C*SE
27	-542	C*PER + C*SE + C*SE

Table 1: Discovered kernel structures and the Bayesian Information Criterion (BIC) for the encountered 28 anomalies.

4 PRELIMINARY RESULTS

In this section, we report and discuss the results of our preliminary performance evaluation. For this purpose, we use the recently introduced MONSOON IoT dataset [5] which comprises 357,383 data records in total. This dataset is based on a real production line of coffee capsules and the attribute under observation is the plastification time, that is the time which is needed to melt (plastify) the plastic melt for the actual injection molding cycle. More information about this process can be found in [3].

An overview of this attribute value, i.e. the pastification time, as a function of the cycle number is shown in Figure 1. As can be seen in the figure, while the normal plastification time is at approximately 4.2 seconds, it drops down to less then 3 seconds in case of an anomaly. Supported by domain experts, we figured out 28 anomalies in total in this dataset, of which three are shown in the above figure.

In the first series of experiments, we computed the best fitting kernel expressions by means of the ABCD algorithm. The results are shown in Table 1 for each anomaly. Together with the kernel expression of the corresponding anomaly, we also show the Bayesian Information Criterion (BIC) value which models the trade-off between model accuracy and size. As can be seen in the table, all anomalies are well described by their corresponding kernel expression (lower BIC values indicate better fit and vice versa). Surprisingly many kernel expressions do not show a linear component k_{LIN}, although some anomalies clearly show this linear tendency. We figure out that this is due to overfitting of the kernel expression in the ABCD algorithm. We aim to address this issue in future research.

In the second series of experiments, we evaluated how suitable a kernel expression of a certain anomaly fits to other anomalies. The results in form of the corresponding BIC values are summarized in Table 2. As can be seen in this table, kernel expressions of a certain anomaly do in general not fit to other anomalies. One reason for this behavior is the high degree of idiosyncrasy of the anomalies. Another reason might be the overfitting issue mentioned above.

To sum up, we have investigated the potential of describing anomalies in IoT sensor data by means of kernel expressions. Our preliminary results indicate that our proposal is well suited for this purpose. As one major challenge, we figure out that the problem of overfitting needs to be addressed in future research.

5 CONCLUSIONS AND FUTURE WORK

In this paper, we have addressed the problem of discovering the inherent structures of anomalies arising in IoT sensor data. To this end, we have proposed to model and describe anomalies by means of kernel expressions, which are combinations of well-known kernels. The results of our empirical analysis show that our proposal is suitable for modeling differently structured anomalies. Moreover, the results indicate that Gaussian processes provide a powerful tool for future algorithmic investigations of IoT sensor data.

In future work, we aim to address the problem of overfitting by modifying the grammar used within the ABCD algorithm for computing the kernel expressions. In addition, we aim to further develop our proposal in order to not only describe anomalies but also detect anomalies (which is not the focus of the current paper). For this purpose, we aim to measure similarity in IoT sensor data by incorporating Gaussian processes into adaptive distance-based similarity models, such as the Signature Matching Distance [6], and query processing algorithms [2, 4].

ACKNOWLEDGMENTS

The project underlying this paper has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 723650 (MONSOON). This paper reflects only the authors' views and the commission is not

Kernel	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
0	-799	7434	41283	15859	1493	15065	60870	104642	12783	201580	19	4323	-290	14637	-138	768	-336	1010	-453	-74	299	429	212	26246	85201	22461	-242	-410
1	-690	-706	-371	-874	-613	-427	-384	-973	-705	-1003	-323	-780	-291	-775	-334	-540	-456	-453	-536	-783	-169	-271	-310	-922	-1008	-849	-452	-524
2	-644	-706	-604	-307	-650	-522	12783	1261	-741	1050	-369	-745	-266	507	-339	-586	-471	-552	-518	-732	-269	-417	-423	25	328	-542	-463	-515
3	-746	-689	-49	-921	-575	-252	-131	1428	-634	656	-220	-819	-304	244	-315	-451	-448	-370	-534	-825	-41	-81	-150	-826	-524	-781	-397	-507
4	-757	64	8037	510	-742	-188	8272	1161	2069	2158	-188	-178	-317	960	-359	180	-517	-347	-554	-825	-47	232	72	1452	733	1311	-437	-360
5	-664	-712	-603	-861	-664	-543	-596	-1012	-726	-988	-340	-762	-278	-837	-343	-574	-481	-564	-364	-748	-262	-413	-425	-113	-1008	-862	-461	-522
6	-630	-690	-598	-812	-645	-519	-630	-942	-711	-955	-353	-727	-265	-803	-336	-576	-467	-555	-505	-711	-269	-415	-425	-866	-948	-818	-454	-498
7	-710	-628	-203	-879	-571	-277	40	-1020	-610	-605	-280	-786	-289	-690	-308	-470	-443	-386	-532	-797	-105	-124	-165	-901	-979	-812	-434	-513
8	-705	-695	-251	-37	-654	-435	26932	4043	-762	3202	-346	-799	-295	1808	-339	-521	-473	-514	-552	-790	-228	-308	-316	1284	1396	-311	-456	-543
9	-707	-676	-341	-876	-603	-353	-321	-949	-680	-1025	-295	-789	-295	-739	-315	-524	-445	-417	-531	-796	-150	-188	-245	-942	-1030	-833	-417	-516
10	1100	17963	39634	11743	9764	28200	42825	21496	17842	23668	-424	8490	-256	21496	-321	14360	744	16475	1186	6690	-280	27159	22641	15147	16334	16285	1852	5048
11	-768	-264	35152	4123	-613	-275	38575	16915	1813	4547	-205	-849	-308	5682	-282	-366	-453	-369	-525	-843	-44	-26	-97	4963	3229	1179	-393	-511
12	3497	12806	43065	8404	10581	20436	43440	13649	15087	12678	12444	7084	-311	19649	2792	22502	6864	8959	5856	3580	12974	23140	20507	10828	10761	14404	7576	8797
13	-725	-690	-329	-897	-661	-390	-274	-531	-707	-44	-338	-807	-298	-860	-349	-553	-495	-471	-560	-825	-195	-315	-302	-868	-766	-824	-477	-545
14	782	10272	34307	5692	10540	17399	39816	11292	15353	11941	10135	4633	-286	16934	-339	12064	3486	13876	4713	3175	3996	19687	17287	8695	8431	10076	5533	7290
15	-682	-142	18892	2896	-648	-460	25790	2757	3848	4590	-349	109	-284	1571	-334	-590	-464	-486	-536	-740	-212	-320	-341	3514	2919	4823	-462	-531
16	841	16636	45060	7627	6465	22063	46105	23901	15094	9764	-353	8375	-304	18458	-344	18424	-503	15531	1180	5836	-224	18906	21613	9877	12135	16161	216	2171
17	-728	-706	43	-883	-687	-192	-253	-1034	-724	-1057	-152	-795	-293	-826	-353	-157	-488	-602	-524	-813	-97	146	-294	-940	-1073	-879	-403	-509
18	239	8646	30046	4008	2545	13235	31476	11202	6420	8144	-273	3878	-289	11521	-305	10184	-431	12023	-545	1630	-85	10539	11217	6903	6388	10154	-430	1030
19	-666	595	4854	-26	-303	-70	8137	595	1317	1111	-53	-205	-257	863	-237	-67	-313	-188	-467	-804	127	198	57	999	610	1062	-305	-411
20	1073	22702	53438	15472	14644	32643	45911	22105	15895	22178	8444	6579	-257	24654	-319	13959	2244	15564	4203	9121	-281	34709	20454	14474	15040	16554	5304	9165
21	167	6959	22082	2677	844	6647	25113	8595	7569	6198	-348	2186	-232	5298	-300	7729	-410	1939	-458	632	-259	-426	6228	4740	4675	5462	-408	-468
22	-550	618	4611	422	212	1187	5830	473	419	897	-330	-205	-261	807	-334	392	-459	-37	-504	-545	-251	-413	-425	819	647	406	-446	-485
23	-740	-690	-188	-922	-622	-337	-189	-699	-687	-615	-254	-828	-305	-757	-318	-487	-446	-388	-551	-842	-69	-124	-180	-975	-798	-853	-438	-523
24	749	618	108	-307	833	386	938	610	652	1209	34	972	295	141	496	412	494	802	497	866	35	80	461	781	-1181	461	508	-438
25	-721	-648	-171	-860	-589	-274	-99	1349	-616	1530	-269	-773	-290	304	-302	-433	-438	-387	-532	-809	-77	-127	-190	-326	-942	-880	-425	-513
26	108	9489	32787	5935	4124	14088	43508	10486	13005	11777	-335	6449	-280	9810	-322	13738	-450	12148	439	1934	-185	15172	15764	7871	5359	10290	-455	792
27	188	10101	32689	4425	2295	8295	35339	11450	12860	9609	-325	5189	-294	9651	-318	9626	-455	6945	-535	1349	-151	-231	7129	7497	8193	10342	-437	-542

Table 2: Evaluation of the BIC for every kernel expression against every anomaly.

responsible for any use that may be made of the information it contains.

REFERENCES

[1] Bryan Auslander, Kalyan Moy Gupta, and David W. Aha. 2011. A comparative evaluation of anomaly detection algorithms for maritime video surveillance. In *Proc. SPIE 8019, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X (SPIE Proceedings)*, Edward M. Carapezza (Ed.). SPIE, 801907. https://doi.org/10.1117/12.883535

[2] Christian Beecks and Max Berrendorf. 2018. Optimal k-Nearest-Neighbor Query Processing via Multiple Lower Bound Approximations. In *IEEE International Conference on Big Data, Big Data 2018, Seattle, WA, USA, December 10-13, 2018*. IEEE, 614–623. https://doi.org/10.1109/BigData.2018.8622493

[3] Christian Beecks, Shreekantha Devasya, and Ruben Schlutter. 2019. Machine Learning for Enhanced Waste Quantity Reduction: Insights from the MONSOON Industry 4.0 Project. In *Machine Learning for Cyber Physical Systems*, Jürgen Beyerer, Christian Kühnert, and Oliver Niggemann (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–6.

[4] Christian Beecks and Alexander Graß. 2016. Multi-step threshold algorithm for efficient feature-based query processing in large-scale multimedia databases. In *2016 IEEE International Conference on Big Data, BigData 2016, Washington DC, USA, December 5-8, 2016*. IEEE, 596–605. https://doi.org/10.1109/BigData.2016.7840652

[5] Christian Beecks, Alexander Grass, and Shreekantha Devasya. 2018. Metric Indexing for Efficient Data Access in the Internet of Things. In *IEEE International Conference on Big Data, Big Data 2018, Seattle, WA, USA, December 10-13, 2018*. IEEE, 5132–5136. https://doi.org/10.1109/BigData.2018.8622387

[6] Christian Beecks, Steffen Kirchhoff, and Thomas Seidl. 2013. Signature matching distance for content-based image retrieval. In *International Conference on Multimedia Retrieval, ICMR'13, Dallas, TX, USA, April 16-19, 2013*. ACM, 41–48. https://doi.org/10.1145/2461466.2461474

[7] Markus Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. 2000. LOF: Identifying Density-Based Local Outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*. ACM, 93–104.

[8] Arno Candel, Erin LeDell, Viraj Parmar, and Anisha Arora. 2018. Deep Learning with H2O. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/DeepLearningBooklet.pdf. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/DeepLearningBooklet.pdf (Accessed on 01/08/2019).

[9] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly Detection: A Survey. *ACM Comput. Surv.* 41, 3, Article 15 (July 2009), 58 pages. https://doi.org/10.1145/1541880.1541882

[10] R. Domingues, F. Buonora, R. Senesi, and O. Thonnard. 2016. An Application of Unsupervised Fraud Detection to Passenger Name Records. In *2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshop (DSN-W)*. 54–59. https://doi.org/10.1109/DSN-W.2016.21

[11] David Duvenaud, James Robert Lloyd, Roger Grosse, Joshua B. Tenenbaum, and Zoubin Ghahramani. 2013. Structure Discovery in Nonparametric Regression through Compositional Kernel Search. arXiv:arXiv:1302.4922

[12] Eleazar Eskin, Andrew Arnold, Michael Prerau, Leonid Portnoy, and Sal Stolfo. 2002. A Geometric Framework for Unsupervised Anomaly Detection. In *Applications of Data Mining in Computer Security*, Daniel Barbará and Sushil Jajodia (Eds.). Advances in Information Security, 1568-2633, Vol. 6. Springer US and Imprint and Springer, Boston, MA, 77–101. https://doi.org/10.1007/978-1-4615-0953-0{_}4

[13] Markus Goldstein and Seiichi Uchida. 2016. A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data. (2016).

[14] Alexander Graß, Christian Beecks, and Jose Angel Carvajal Soto. 2019. Unsupervised Anomaly Detection in Production Lines. In *Machine Learning for Cyber Physical Systems*, Jürgen Beyerer, Christian Kühnert, and Oliver Niggemann (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 18–25.

[15] Yunseong Hwang, Anh Tong, and Jaesik Choi. 2016. Automatic Construction of Nonparametric Relational Regression Models for Multiple Time Series. In *ICML 2016: Proceedings of the 33rd International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Maria Florina Balcan and Kilian Q. Weinberger (Eds.), Vol. 48. PLMR, 3030–3039.

[16] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. 2008. Isolation Forest. In *Eighth IEEE International Conference on Data Mining, 2008*, Fosca Giannotti (Ed.). IEEE, Piscataway, NJ, 413–422. https://doi.org/10.1109/ICDM.2008.17

[17] James Robert Lloyd, David Duvenaud, Roger Grosse, Joshua B. Tenenbaum, and Zoubin Ghahramani. 2014. Automatic Construction and Natural-Language Description of Nonparametric Regression Models. arXiv:arXiv:1402.4304

[18] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12 (Nov. 2011), 2825–2830. http://dl.acm.org/citation.cfm?id=1953048.2078195

[19] Clifton Phua, Vincent C. S. Lee, Kate Smith-Miles, and Ross W. Gayler. 2010. A Comprehensive Survey of Data Mining-based Fraud Detection Research. *CoRR* abs/1009.6119 (2010). arXiv:1009.6119 http://arxiv.org/abs/1009.6119

[20] Carl Edward Rasmussen and Christopher K. I. Williams. 2006. *Gaussian Processes for Machine Learning (Adaptive Computation And Machine Learning)*. The MIT Press.

[21] David Renaudie, Maria A. Zuluaga, and Rodrigo Acuna-Agost. 2018. Benchmarking Anomaly Detection Algorithms in an Industrial Context: Dealing with Scarce Labels and Multiple Positive Types. In *IEEE International Conference on Big Data*. 1227–1236.

[22] Peter J Rousseeuw. 1984. Least median of squares regression. *Journal of the American statistical association* 79, 388 (1984), 871–880.

[23] Bernhard Schölkopf, John C. Platt, John C. Shawe-Taylor, Alex J. Smola, and Robert C. Williamson. 2001. Estimating the Support of a High-Dimensional Distribution. *Neural Comput.* 13, 7 (July 2001), 1443–1471. https://doi.org/10.1162/089976601750264965

[24] M. Tavallaee, N. Stakhanova, and A. A. Ghorbani. 2010. Toward Credible Evaluation of Anomaly-Based Intrusion-Detection Methods. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40, 5 (September 2010), 516–524. https://doi.org/10.1109/TSMCC.2010.2048428