National Research University Higher School of Economics

as a manuscript

Ryzhikov Artem

Deep generative models for anomaly detection

PhD Dissertation Summary

for the purpose of obtaining academic degree
Doctor of Philosophy in Computer Science

Academic Supervisor:
PhD
Derkach Denis Aleksandrovich

Moscow – 2024

DISSERTATION TOPIC

Machine learning (ML) has a wide range of applications and remains an essential field of scientific and computer research. In most of these applications, ML provides consistent and promising results. In particular, supervised learning (SL) algorithms that incorporate classification tasks are one of the most well-studied ML fields.

However, in some realistic scenarios, these algorithms remain suboptimal. The high imbalance of datasets and the lack of data samples from some classes or domains may be one of the reasons for such suboptimal behavior. In the first case, such rare class samples are called *anomalies*, and the corresponding problem is called *anomaly detection*. In the second case, the problem is called *domain adaptation*.

The anomaly detection appears in many real scenarios and fields, such as particle identification [1], change point detection [2], chemical process control [3], credit card fraud detection [4], complex system failure predictions [5], video scene analysis [6], novelty detection in time series data [7], data quality certification [8], detection of climate changes [9], finding rare specific cases of diseases in medicine [10], production quality control [11], aircraft monitoring [12], vibration monitoring of mechanical systems [13], seismic signal processing [14], human motion and health state analysis [15], detection of cyberattacks [16], audio signal segmentation [17], and many others [18].

This area continues to be underresearched, with many contemporary anomaly detection algorithms failing to fully harness the capabilities of deep learning (DL) techniques. Specifically, the use of generative models in anomaly detection has not been extensively explored yet.

In this work, the goal is to explore the ability to utilize generative models for anomaly detection and domain adaptation problems to reveal the potential of DL and SL in these tasks. This research investigates various methodologies including comprehensive DL-based anomaly detection and domain adaptation techniques, distinct deep generative models for creating surrogate anomalies to enrich datasets, and their applications and enhancements in actual anomaly detection scenarios, such as improving the anomaly detection inference speeds. The methodologies introduced cater to image, tabular, and time-series data, showcasing substantial improvements compared to current leading methods.

THEORETICAL AND PRACTICAL SIGNIFICANCE

The thesis presents significant theoretical and practical advancements in the fields of anomaly detection and change point detection (CPD) using generative deep learning (DL) algorithms.

The theoretical significance lies in the introduction of novel anomaly detection algorithms. The $(1 + \varepsilon)$-class classification method introduces a new family of algorithms that efficiently address problems that fall between one-class and two-class settings. This method can incorporate any number of known anomalous examples without requiring a representative sample of anomalous data, marking a significant advancement over traditional methods. Another significant contribution is the NFAD method, which uses normalizing flows. This method significantly outperforms previous algorithms, including the $(1 + \varepsilon)$-class method, across various scenarios. It is model-agnostic and can be integrated with any supervised learning algorithms. Additionally, the application of latent neural stochastic differential equations (SDEs) to time series anomaly detection is a pioneering effort in the field. This method significantly outperforms existing CPD algorithms, providing a robust and scalable solution for detecting change points in multivariate time series data. The thesis also provides detailed descriptions of the introduced algorithms, including theorems and proofs that ensure their functionality. This rigorous theoretical foundation guarantees the robustness and reliability of the proposed methods.

On the practical side, the thesis includes comprehensive evaluations on various datasets, demonstrating that the proposed algorithms consistently outperform existing baselines. This empirical evidence supports the practical applicability of the methods. The thesis includes software implementations of the proposed algorithms along with scripts to reproduce the experiments. This transparency and reproducibility are crucial for practical applications and further research. The implementations are published and accessible, facilitating their use and adaptation in other projects. The research also explores domain adaptation with gradient reversal for high-energy physics applications, and variational dropout sparsification techniques for speeding up neural networks. These studies demonstrate the applicability of the proposed methods in real-world scenarios, particularly in the field of particle identification at LHCb.

The thesis outlines several promising directions for future research, including enhancing model robustness and generalization capabilities, developing techniques to improve the interpretability and explainability of deep generative models for anomaly detection, exploring methods for dynamic anomaly detection that can adapt to evolving data distributions and anomaly patterns over time, validating the efficacy of deep generative models in real-world applications across various domains, and investigating human-in-the-loop approaches to leverage human expertise in complementing and validating model predictions.

In conclusion, the thesis makes substantial contributions to the fields of anomaly detection and change point detection by introducing novel generative deep learning-based methods that are theoretically sound and practically effective. The research offers a solid foundation for further exploration and innovation, paving the way for advanced applications in various domains. The proposed algorithms and their implementations are expected to have a significant impact on both academic research and practical applications.

KEY RESULTS

The first work "$(1 + \varepsilon)$-class Classification: an Anomaly Detection Method for Highly Imbalanced or Incomplete Data Sets" [19] of this part represents a novel DL generative approach to anomaly detection, which utilizes the MCMC (Markov Chain Monte Carlo, [20]) sampling technique for surrogate anomalies generation. In this work, a new family of anomaly detection algorithms is presented. It can be efficiently applied to problems intermediate between one-class and two-class settings. The solutions produced by these methods combine the best features of one-class and two-class approaches. In contrast to conventional one-class approaches, proposed methods can effectively take into account any number of known anomalous examples, and, unlike conventional two-class classification, do not require a representative sample of anomalous data. The experiments show better or comparable performance to conventional two-class and one-class algorithms. The approach is especially beneficial for anomaly detection problems, in which anomalous data is non-representative or might evolve over time. The algorithm significantly outperforms existing approaches and introduces a conceptually novel approach to the anomaly detection problem, making the introduced $1 + \varepsilon$ method work well on any fraction of anomalies in the training dataset.

The second work is "NFAD: Fixing anomaly detection using normalizing flows" [21] evolves the first work with normalizing flows [22]. In this research, a novel model-agnostic training scheme for anomaly detection is introduced. Theoretical and practical evaluations demonstrate its effectiveness in addressing challenges that are typically difficult for both one-class and two-class methods. This method merges the advantages of both one-class and two-class algorithms. Unlike one-class methods, this new approach allows the classifier to effectively use any available anomalous examples without needing a large dataset of anomalies, as required by traditional two-class approaches. The algorithm introduced outperforms current anomaly detection techniques in most practical scenarios. The introduced approach is quick, robust, and adaptable during both training and inference phases. Its comprehensive augmentation strategy broadens the possibilities for ongoing research in anomaly detection problem and enables the application of any classifiers to any types of data. Furthermore, image dataset results can be enhanced through the adoption of emerging normalizing flow techniques. The related algorithm is called *NFAD* and drastically outperforms all previously existing anomaly detection algorithms, including $1 + \varepsilon$. Additionally, the introduced method is model-agnostic and can be used with any supervised learning algorithms, including conventional ones.

The third work "Latent Neural Stochastic Differential Equations for Change Point Detection" [23] introduces the first deep learning anomaly (change point) detection algorithm for time series data based on neural stochastic differential equations.The work is aimed to designing an efficient DL generalization of the conventional likelihood ratio CPD approaches based on stochastic differential equations. To this end, a first study of Latent SDE in a change-point detection setting is presented. As a result of this work, a novel CPD algorithm on the edge of modern deep learning approaches and conventional CPD methods is introduced. This is the first deep learning modification of the stochastic differential equations approach to change point detection. Both theoretical and experimental evidence demonstrate that the proposed method effectively identifies all principal change point types in multivariate time series data, including trend, mean, and volatility shifts. In most of the scenarios and metrics, the model shows high robustness and a performance which is strongly higher than other state-of-the-art CPD algorithms used in this work. With all the aforementioned, the proposed algorithm represents a great interest from a theoretical and performance perspective for change point detection problem, which occurs in many real time series analysis scenarios. The introduced algorithm also drastically outperforms the existing change point detection algorithms and has a great research potential as the first application of neural SDE and one of the first efficient deep learning applications to the change point detection problem.

The next paper is called "Domain adaptation with gradient reversal for MC/real data calibration" [24] and incorporates a study of domain adaptation [25] technique for the PID anomaly detection task. The study of this work shows that the domain adaptation technique can be effectively applied to real anomaly detection problem, namely, to Particle Identification (PID) in High Energy Physics. The obtained results have great potential, proving that the domain adaptation approach can be applied effectively to the problem, preserving the neural network classifier from the overfitting on the training domain.

The fifth paper "Variational Dropout Sparsification for Particle Identification speed-up" [26] represents a study of neural network sparsification and speed-up techniques in application to the PID problem. The results show that Variational Dropout Sparsification technique provides the best results for the given problem. In the PID problem, the studied technique gives an impressive 16 times speedup without any loss of quality, which is of great interest for research and application in many other use cases.

Finally, the paper titled "Robust Neural Particle Identification Models" [27], introduces a technique utilizing the Common Specific Decomposition [28]. In this work, the recent common-specific low-rank decomposition (CSD) algorithm is studied in application to the real anomaly detection task (PID). The algorithm is capable of selecting common features even for decays that are not present in the original domains. The algorithm obtained shows higher stability with respect to the previously presented algorithms, thus demonstrating a substantial increase in the quality of the solution for the particular case and demonstrating a substantial interest in many machine learning applications. This method accounts for distinct domains within the training data by separating the shared and unique decay elements of the input feature set. It effectively reduces the decline in the performance of anomaly detection algorithms in practical PID scenarios.

PUBLICATIONS AND APPROBATION OF RESEARCH

The author of the thesis is a main author of the most and coauthor of the rest proposed anomaly detection algorithms [19], [21], [23] and related application studies [24], [26], [27].

These algorithms are based on deep generative models and provide scientific novelty in the given area. The thesis author performed the methodological design of the aforementioned anomaly detection algorithms and related experiments, the technical implementation of them, and the analysis of the results obtained. As a result, three novel anomaly detection algorithms were published in respected Q1-Q2 journals. In two of them (Q1 and Q2), the dissertator is a main author. In the third paper (Q1), the dissertator is a second author and has the same impact as the first author (Maxim Borisyak, HSE) of the paper.

In addition, three studies on anomaly detection applications were published in Q4 journals. In two of them, the dissertator is a main author. In the third paper, the dissertator is a second coauthor.

First-tier publications[1]

[19] M. Borisyak, A. Ryzhikov, A. Ustyuzhanin, D. Derkach, F. Ratnikov and O. Mineeva, '(1 + epsilon)-class classification: An anomaly detection method for highly imbalanced

[1] First-tier publications include papers indexed in the Web of Science (Q1 or Q2) or Scopus (Q1 or Q2) databases, as well as peer-reviewed collections of conferences that appear in CORE rankings (ranks A and A*).

or incomplete data sets,' 2020. [Online]. Available: https://jmlr.csail.mit.edu/papers/volume21/19-514/19-514.pdf.

[21] A. Ryzhikov, M. Borisyak, A. Ustyuzhanin and D. Derkach, *Nfad: Fixing anomaly detection using normalizing flows*, 2021. DOI: 10.7717/peerj-cs.757. [Online]. Available: https://peerj.com/articles/cs-757/.

[23] A. Ryzhikov, M. Hushchyn and D. Derkach, *Latent stochastic differential equations for change point detection*, 2023. DOI: 10.1109/ACCESS.2023.3318318. [Online]. Available: https://ieeexplore.ieee.org/document/10261192.

Second-tier publications[2]

[24] A. Ryzhikov and A. Ustyuzhanin, 'Domain adaptation with gradient reversal for mc/real data calibration,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1085, 2018, p. 042 018.

[26] A. Ryzhikov, D. Derkach, M. Hushchyn, L. Collaboration *et al.*, 'Variational dropout sparsification for particle identification speed-up,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1525, 2020, p. 012 099.

[27] A. Ryzhikov, A. Temirkhanov, D. Derkach *et al.*, 'Robust neural particle identification models,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2438, 2023, p. 012 119.

CONTENTS

The first and main part of the work comprises three novel DL anomaly detection algorithms [19], [21], [23]. The elaborated methods allow fitting tabular, image, and time series anomaly detection methods in either supervised, semi-supervised, or unsupervised manners when an arbitrary number of anomalies is given in the training dataset. The algorithms are based on deep generative models and allow utilizing any neural network architecture behind. In

[2] Second-tier publications are papers published in journals included on HSE's list of high quality journals or indexed in the Web of Science (Q3 or Q4) or Scopus (Q3 or Q4) databases, as well as peer-reviewed collections of conferences appearing in CORE rankings (rank B).

comparison with conventional anomaly detection algorithms, the introduced methods are designed to work the same well with a variate anomaly fraction in the training dataset. Also, they have no limitations on the neural network architecture used in the backbone.

The second part of the work comprises three application studies to the real anomaly detection experiment (PID [1]). This part comprises two domain adaptation studies [24], [27] applied to the experiment, and one study of existing neural network speedup approaches for the PID anomaly detection problem.

CONCLUSION

This thesis comprises a complete set of scientifically novel anomaly detection studies, along with application studies of existing methods in anomaly detection and related fields.

The new set of DL anomaly detection algorithms introduced in this thesis deals efficiently with hard-to-address problems both by one-class or two-class methods. These anomaly detection solutions are the first which combine the best features of one-class and two-class approaches with a power of deep generative models. It forms a new, strong, and promising direction in anomaly detection research.

The results obtained with these algorithms significantly outperform the existing anomaly detection approaches in most anomaly detection scenarios. It is the first well-proved and effective application of deep generative models to real anomaly detection challenges, with a wide further application potential.

The thesis presents several innovative contributions to the field of anomaly detection, summarized as follows:

- a brand-new generative approach to anomaly detection problem is designed and published in top Q1 journal (the thesis author is a second co-author). It is a first approach to anomaly detection, which works well both in one-class, two-class, and intermediate scenarios. The algorithm is called "$1 + \varepsilon$" and is based on MCMC sampling of surrogate anomalies around the classification neural network boundaries. The algorithm showed an outstanding performance on all the aforementioned scenarios.

- another deep-generative approach based on normalizing flow called *NFAD* is published in the Q2 journal (the thesis author is the main coauthor). It is a first approach which utilizes normalizing flows in anomaly detection setting, and beats all the previously existing algorithms including "$q + \varepsilon$" on the most part of experiments.
- a first time series anomaly (change point) detection approach based on neural (latent) stochastic differential equations is published in the top Q1 journal (the thesis author is a main coauthor). The method drastically outperforms all the existing change point detection algorithms on almost all the benchmark corpuses and metrics studied.
- two domain adaptation training and one sparsification techniques are first studied in application to real anomaly detection problem. Each of the three studies is published in Q4 proceedings.

Bibliography

[1] C. Lippmann, ‘Particle identification,’ *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 666, pp. 148–172, 2012, Advanced Instrumentation, ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2011.03.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900211005419.

[2] S. Aminikhanghahi and D. J. Cook, ‘A survey of methods for time series change point detection,’ *Knowledge and information systems*, vol. 51, no. 2, pp. 339–367, 2017.

[3] B. Huang, ‘Detection of abrupt changes of total least squares models and application in fault detection,’ *IEEE Transactions on Control Systems Technology*, vol. 9, no. 2, pp. 357–367, 2001. DOI: 10.1109/87.911387.

[4] E. Aleskerov, B. Freisleben and B. Rao, ‘Cardwatch: A neural network based database mining system for credit card fraud detection,’ pp. 220–226, Apr. 1997.

[5] J. Xu and H. Li, ‘The Failure Prediction of Cluster Systems Based on System Logs,’ *Wang M. (eds) Knowledge Science, Engineering and Management. KSEM 2013. Lecture Notes in Computer Science, vol 8041.*, 2013.

[6] Z. Gao, G. Lu, C. lv and P. Yan, ‘Key-frame selection for automatic summarization of surveillance videos: A method of multiple change-point detection,’ *Machine Vision and Applications*, vol. 29, Oct. 2018. DOI: 10.1007/s00138-018-0954-7.

[7] M. Schmidt and M. Simic, ‘Normalizing flows for novelty detection in industrial time series data,’ 2019. eprint: arXiv:1906.06904.

[8] M. Borisyak, F. Ratnikov, D. Derkach and A. Ustyuzhanin, ‘Towards automation of data quality system for cern cms experiment,’ *Journal of Physics: Conference Series*, vol. 898, no. 9, p. 092 041, 2017.

[9] J. Reeves, J. Chen, X. L. Wang, R. Lund and Q. Q. Lu, 'A review and comparison of changepoint detection techniques for climate data,' *Journal of Applied Meteorology and Climatology*, vol. 46, no. 6, pp. 900–915, 2007. DOI: 10.1175/JAM2493.1. eprint: https://doi.org/10.1175/JAM2493.1. [Online]. Available: https://doi.org/10.1175/JAM2493.1.

[10] C. Spence, L. Parra and P. Sajda, 'Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model,' pp. 3–10, 2001. DOI: 10.1109/MMBIA.2001.991693.

[11] A. F. Bissell, 'Cusum techniques for quality control,' *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 18, no. 1, pp. 1–30, 1969, ISSN: 00359254, 14679876. [Online]. Available: http://www.jstor.org/stable/2346436.

[12] D. Henry, S. Simani and R. J. Patton, 'Fault detection and diagnosis for aeronautic and aerospace missions,' in *Fault Tolerant Flight Control: A Benchmark Challenge*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 91–128, ISBN: 978-3-642-11690-2. DOI: 10.1007/978-3-642-11690-2_3. [Online]. Available: https://doi.org/10.1007/978-3-642-11690-2_3.

[13] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone and A. Beghi, 'Machine learning for predictive maintenance: A multiple classifier approach,' *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 812–820, 2015.

[14] M. Basseville and I. V. Nikiforov, 'Detection of abrupt changes: Theory and application,' *Technometrics*, vol. 36, p. 550, 1993.

[15] A. Briassouli, T. Vagia and I. Kompatsiaris, 'Human motion analysis via statistical motion processing and sequential change detection,' *EURASIP Journal on Image and Video Processing*, vol. 2009, Jan. 2009. DOI: 10.1155/2009/652050.

[16] A. G. Tartakovsky, B. Rozovskii, R. B. Blazek and H. Kim, 'Detection of intrusions in information systems by sequential change-point methods,' *Statistical Methodology*, vol. 3, pp. 252–293, 2006.

[17] Z. Shuyang, T. Heittola and T. Virtanen, 'Active learning for sound event detection,' *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2895–2905, 2020.

[18] A. Tartakovsky, I. Nikiforov and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. Aug. 2014, ISBN: 9781439838204. DOI: 10.1201/b17279.

[19] M. Borisyak, A. Ryzhikov, A. Ustyuzhanin, D. Derkach, F. Ratnikov and O. Mineeva, '(1 + epsilon)-class classification: An anomaly detection method for highly imbalanced or incomplete data sets,' 2020. [Online]. Available: https://jmlr.csail.mit.edu/papers/volume21/19-514/19-514.pdf.

[20] F. Septier and G. W. Peters, 'Langevin and hamiltonian based sequential mcmc for efficient bayesian filtering in high-dimensional spaces,' 2015. DOI: 10.1109/JSTSP.2015.2497211. eprint: arXiv:1504.05715.

[21] A. Ryzhikov, M. Borisyak, A. Ustyuzhanin and D. Derkach, *Nfad: Fixing anomaly detection using normalizing flows*, 2021. DOI: 10.7717/peerj-cs.757. [Online]. Available: https://peerj.com/articles/cs-757/.

[22] D. J. Rezende and S. Mohamed, 'Variational inference with normalizing flows,' 2015. eprint: arXiv:1505.05770.

[23] A. Ryzhikov, M. Hushchyn and D. Derkach, *Latent stochastic differential equations for change point detection*, 2023. DOI: 10.1109/ACCESS.2023.3318318. [Online]. Available: https://ieeexplore.ieee.org/document/10261192.

[24] A. Ryzhikov and A. Ustyuzhanin, 'Domain adaptation with gradient reversal for mc/real data calibration,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1085, 2018, p. 042 018.

[25] Y. Ganin and V. Lempitsky, 'Unsupervised domain adaptation by backpropagation,' in *International conference on machine learning*, PMLR, 2015, pp. 1180–1189.

[26] A. Ryzhikov, D. Derkach, M. Hushchyn, L. Collaboration *et al.*, 'Variational dropout sparsification for particle identification speed-up,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1525, 2020, p. 012 099.

[27] A. Ryzhikov, A. Temirkhanov, D. Derkach *et al.*, 'Robust neural particle identification models,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2438, 2023, p. 012 119.

[28] V. Piratla, P. Netrapalli and S. Sarawagi, *Efficient domain generalization via common-specific low-rank decomposition*, 2020. arXiv: 2003.12815 [cs.LG].