Anomaly Detection - Review of Methods, Tools and Algorithms

Roberts Volkovičs
Modelling of sociotechnical systems.
Vidzeme University of Applied Sciences
Valmiera, Latvia
roberts.volkovics@gmail.com

This paper contains review of algorithms, methods and tools nowadays used for anomaly detection.

Anomaly detection is used in many domains of science and industry, some authors classify anomaly detection as data mining and data science tool, others state it is decision support tool under artificial intelligence domain and indeed the use cases of anomaly detection are very different.

The article describes the main algorithms used for anomaly detection from perspective of theory of computer science and practical use cases of anomaly detection in different domains of industry. Several paragraphs are dedicated to the frameworks used by one of the most popular and powerful anomaly detection tools available in the market - Microsoft Anomaly detector.

Keywords:* *Algorithms, Anomaly detection, Novelty detection, Outlier detection.

I. INTRODUCTION

Anomaly detection refers to the problem of finding patterns in data that do not conform to the expected behaviour. These non-conforming patterns are often referred to as anomalies, outliers, discordant observations, exceptions, aberrations, surprises, peculiarities or contaminants in different application domains [1].

Anomaly detection is used in many areas:

- Financial fraud detection [2].
- Anti-money laundering [3].
- Disease detection.
- Network security threat detection.
- Detection of anomalies in the logs of software.
- Monitoring of nature, weather or climate change.
- Monitoring of industrial equipment faults.
- Defect detection.
- Detecting faulty products from vision sensors [4].
- Traffic prediction [5].

In this article methods and tools used during anomaly detection will be described for different use case scenarios.

Purposes and goals of usage of anomaly detection could be:

- **Product performance:** An anomaly detection paired with machine learning can correlate the existing data to be cross-checked while maintaining generalization and finding odd standing products with complete knowledge of what makes them an anomaly [6].
- **Technical performance:** For example, any faults in your own deployed system may leave your server to be vulnerable to active DDoS attacks. Such errors can also be proactively avoided and treated at the root using machine learning integrated into the DevOps pipeline [6].
- **Training performance:** During the pre-training phase, anomaly detection can come in handy, pointing out irregularities in the data set, which may cause the model to over-fit and, in turn, act poorly [6].

Print ISSN 1691-5402
Online ISSN 2256-070X
https://doi.org/10.17770/etr2023vol2.7283

Anomalies could be recognized by:

- **Statistical analysis:** By calculating measures such as mean, median, and standard deviation, you can identify data points significantly different from the rest of the sample [5].
- **Machine learning algorithms:** Various machine learning algorithms can be used to identify anomalies, such as clustering, classification, and density-based methods [5].
- **Data visualization:** By creating charts or plots of your data, you can visually identify anomalies by looking for points or trends that stand out [5].
- **Rule based systems:** You can set up rules or thresholds to flag data points that fall outside a specific range or violate certain conditions [5].
- **Human inspection:** In some cases, manually reviewing the data may be necessary to identify anomalies that may not be easily detected through automated methods [5].

Anomalies could be further classified on the way the data of anomalies differ from normal data:

- **Outliers:** Short/small anomalous patterns that appear in a non-systematic way in data collection [7].
- **Change in Events:** Systematic or sudden change from the previous normal behaviour [7].
- **Drifts:** Slow, unidirectional, long-term change in the data [7].

Anomaly classification from data perspective "with real world context":

- **Point Anomaly:** A tuple within the data set that can be an anomaly if it is far from the trend set by the other data points [6].
- **Contextual Anomaly:** Contextual anomalies can be considered an anomaly only if taken in a particular context and may even be valid if taken from another context [6].
- **Collective Anomaly:** Such anomalies occur when data points in a whole collection of points act strange towards the other values, making the subset a complete rarity [6].

Data of anomalies are usually interesting for further studies, but sometimes anomaly detection methods and tools are used to remove the "noisy" data from the data set before the further analysis. This might be in the case when sensor data return anomalies due to broken or malfunctioning sensors.

Natural sciences with anomaly detection tasks work in different ways. In the scope of this work we focus on anomaly processing with methods and tools available in mathematics, algorithms and artificial intelligence rather than physics which tend to improve measurement results from sensors by combining sensor data to reduce the number of anomalies in final calculated sensor measurements.

For example, satellites can have many sensors to monitor climate change on the earth, but some particular measurement could be done only by combining data from many sensors, reducing the noise, removing anomalies, removing data which are not correct due to bad visibility (clouds, fog, etc.).

More on remote sensing and satellite data processing one can find in the research aimed to develop a split-window (SW) algorithm to estimate land surface temperature (LST) from two-channel thermal infrared (TIR) and one-channel middle infrared (MIR) images of SLSTR observation [8].

II. Materials and methods

Anomaly Detection in Machine Learning (ML) is the complete procedure of dealing with anomalies and irregularities in a data set. These can either be outliers or data points significantly different from the usual trend that the other data points follow. These irregularities tend to become an issue while training causing unwanted skewing in the model predictions [6].

Benefits and importance of machine learning in anomaly detection:

- **It makes scaling anomaly detection easier:** by automating the identification of patterns and anomalies without requiring explicit programming [9].
- **Highly adaptable:** Machine Learning algorithms are highly adaptable to changing data set patterns, making them highly efficient and robust with time [9].
- **Easily handles large and complex data sets:** making anomaly detection efficient despite the data set complexity [9].
- **Ensures early anomaly identification and detection:** by identifying anomalies as they happen, saving time and resources [9].
- **Higher levels of accuracy:** Machine Learning based anomaly detection systems help to achieve higher levels of accuracy in anomaly detection compared to traditional methods [9].

Anomaly Detection in machine learning, either pre-model or post-model development and deployment, is an essential task to ensure the smooth running of the ML operations pipeline. With small, skewed values in the data pre-training or frauds and misuse of your services, anomaly detection goes a long way to cut cost, time and boost performance [6].

In Layman's terms, Anomaly Detection is ultimately the task of training a machine to gauge the ability to define what is expected. Still, when paired with machine learning, it also ensures that the model does not lose its ability to generalize [6].

From machine learning techniques which we can use for anomaly detection we can mention:

- **Normal-only anomaly detection:** Several approaches to designing anomaly detection algorithms require little or no anomalous data. These "normal-only" methods train an algorithm on normal data only, and identify data outside those norms as anomalous [10].
- **Unsupervised learning:** In addition to the unsupervised learning techniques, i.e., K-means, Gaussian mixture techniques, K-medians, etc., Anomaly Detection for unsupervised learning also deals with unlabelled data - anomaly detection for such knowledge also works by figuring out the pattern the unlabelled points are following. A vast selection of unsupervised learning algorithms works on the concept of clustering techniques [6].
- **Isolation forest:** The premise of the Isolation Forest algorithm is that anomalous data points are easier to separate from the rest of the sample.

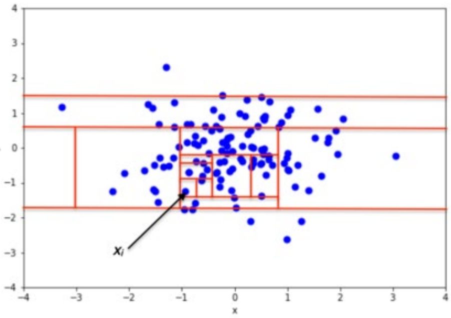


Fig. 1. Isolation of non-anomalous data point [11]

In order to isolate a data point, the algorithm recursively generates partitions on the sample by randomly selecting an attribute and then randomly selecting a split value between the minimum and maximum values allowed for that attribute [11].

In the example, "Fig. 1.", we see isolation of non-anomalous data point [11]. In the example, "Fig. 2.", we see the isolation of anomalous data point [11].

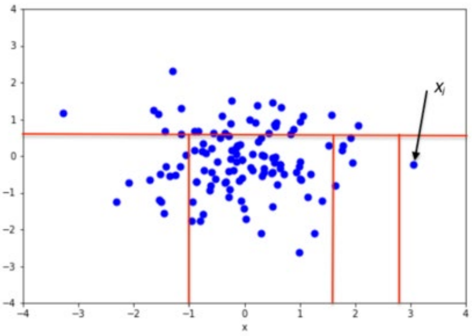


Fig. 2. Isolation of anomalous data point [11]

- **Mahalanobis Distance:** In statistics, we sometimes measure "nearness" or "farness" in terms of the scale of the data. Often "scale" means "standard deviation". For uni-variate data, we say that an observation that is one standard deviation from the mean is closer to the mean than an observation that is three standard deviations away. For many distributions, such as the normal distribution, this choice of scale also makes a statement about probability. Specifically, it is more likely to observe an observation that is about one standard deviation from the mean than it is to observe one that is several standard deviations away [12]. Mahalanobis Distance is a multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. This distance is zero for P at the mean of D and grows as P moves away from the mean along each principal component axis. If each of these axes is re-scaled to have unit variance, then the Mahalanobis distance corresponds to standard Euclidean distance in the transformed space. The Mahalanobis distance is thus unit-less, scale-invariant, and considers the correlations of the data set [13].
- **Autoencoders:** This approach uses artificial neural networks to compress the data into lower dimensions in order to encode it. The data is then decoded by ANNs to recreate the initial input. The rules are already recognized in the compressed data, so when we lower the dimensionality, we don't lose the necessary information [14].

Machines cannot learn a function that translates input features to outputs using unsupervised machine learning because they lack examples of input-output pairings. Instead, they discover structure within the input features and use that to learn. Unsupervised methods are more widely used in the field of anomaly identification than supervised ones because, as was already said, labelled anomalous data is comparatively uncommon. However, the type of anomalies one expects to find is frequently very particular. As a result, many of the abnormalities discovered in an entirely unsupervised approach may simply be noise and may not be relevant to the task at hand [14].

- **Supervised learning:** Since supervised learning relies on labelled data, so do the techniques used to detect anomalies in such models. However, detecting anomalies in such labelled data can be much easier than doing so in unsupervised learning data sets; these techniques hold great potential to be automated and made more efficient [6]. Machines learn a function that maps input features to outputs based on sample input-output pairings while they are learning under supervision. Adopting application-specific knowledge into the process of anomaly detection is the aim of supervised anomaly detection algorithms [14].
- **Support Vector Machines:** Another supervised machine learning approach that is frequently used

for classification is the support vector machine (SVM). SVMs categorize data points using hyper-planes in multidimensional space. The threshold for outliers that must be manually selected is the hyper-parameter [14].

- **K Nearest Neighbours:** A popular supervised machine learning approach for classification is kNN. KNN is a helpful tool when used to solve anomaly detection difficulties since it makes it simple to see the data points on the scatter-plot and makes anomaly identification much more understandable. The fact that kNN performs well on both small and large data sets is an additional advantage. In order to tackle the categorization problem, kNN doesn't actually learn any 'normal' and 'abnormal' values. Therefore, kNN functions as an unsupervised machine learning method for anomaly detection. A range of normal and abnormal values is explicitly defined by a machine learning expert, and the algorithm automatically divides this representation into classes [14].
- **Semi-supervised learning:** It is a blend of the supervised and unsupervised learning, typically, it occurs when there are marked input data but no identified outliers. The model will learn the trends of the standard data from the labelled training data and find anomalies in the unlabelled data that exceed this threshold [6]. Semi-supervised machine learning strategies use a variety of techniques that can benefit from both huge volumes of unlabelled data and sparsely labelled data, acting as a type of middle ground. Due to the abundance of normal instances from which to learn and the dearth of examples of the more unusual or abnormal classes of interest, many real-world anomaly detections use cases are well suited to semi-supervised machine learning. One can train a reliable model on an unlabelled data set and assess its performance using a small quantity of labelled data on the presumption that the majority of the data points in an unlabelled data set are normal [14].
- **Local outlier factor (LOF):** The most popular method for anomaly identification is likely the local outlier factor. The idea of local density serves as the foundation for this method. It contrasts an object's local density with the densities of the nearby data points. A data point is deemed an outlier if its density is lower than that of its neighbours [14]. The local outlier factor is based on a concept of a local density, where locality is given by k nearest neighbours, whose distance is used to estimate the density [15].
- **Robust Covariance:** For gaussian independent features, simple statistical techniques can be employed to detect anomalies in the dataset. For a gaussian/normal distribution, the data points lying away from 3rd deviation can be considered as anomalies. For a dataset having all the feature gaussian in nature, then the statistical approach can be generalized by defining an elliptical hypersphere that covers most of the regular data points, and the data points that lie away from the hypersphere can be considered as anomalies [7].
- **DBSCAN:** This approach uses unsupervised machine learning and is based on the density principle. By examining the local density of the data points, DBSCAN may find clusters in sizable spatial data sets and generally produces positive findings when used for anomaly identification [14].

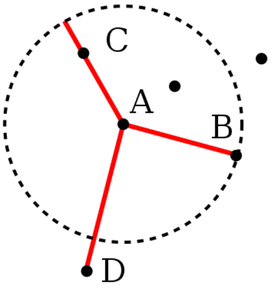
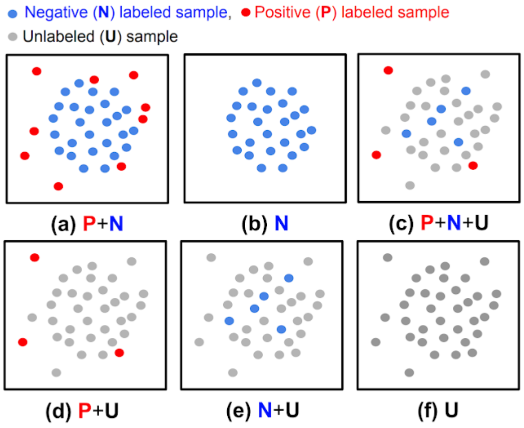


Fig. 3. Reachability distance [15]

- **Bayesian networks:** Machine learning engineers may find anomalies even in high dimensional data thanks to Bayesian networks. When the anomalies we're seeking are subtler and challenging to spot and visualizing them on the plot might not yield the expected results, we employ this strategy [14].

Depending on the availability of the type of data — negative (normal) vs. positive (anomalous) and the availability of their labels — the task of AD involves different challenges [16].


● Negative (N) labeled sample, ● Positive (P) labeled sample
● Unlabeled (U) sample

(a) P+N (b) N (c) P+N+U

(d) P+U (e) N+U (f) U

Fig. 4. Labelled data anomaly detection [16]

While most previous works were shown to be effective for cases with fully-labelled data (either (a) or (b) in the above figure), such settings are less common in practice because labels are particularly tedious to obtain. In most scenarios users have a limited labelling budget, and sometimes there are not even any labelled samples during training. Furthermore, even when labelled data are available, there could be biases in the way samples are labelled, causing distribution differences. Such real-world data challenges limit the achievable accuracy of prior methods in detecting anomalies [16].

As well to many algorithms one can find for anomaly detection, there exist methods to estimate the performance of each algorithm and usability of algorithms for some particular task.

NAB is a standard open-source framework for evaluating real-time anomaly detection algorithms. In simple words, it is a repository that you can easily find on Kaggle. NAB consists of two main components: a dataset containing labelled real-world time series data and a scoring system designed for streaming data. The dataset contains 58 labelled files (approximately 365,000 data points) from various sources such as IT, industrial machine sensors, and social media [5].

A. Anomaly detection in Microsoft Azure cloud "Cognitive services"

In next sections the relationship of modern framework of anomaly detection tool which is part of Cognitive Services (Microsoft AI solution in Azure cloud) and machine learning will be described.

Anomaly detection finds application in many domains including cyber-security, medicine, machine vision, statistics, neuroscience, law enforcement and financial fraud to name only a few. Anomalies were initially searched for clear rejection or omission from the data to aid statistical analysis, for example to compute the mean or standard deviation. They were also removed to better predictions from models such as linear regression, and more recently their removal aids the performance of machine learning algorithms. However, in many applications anomalies themselves are of interest and are the observations most desirous in the entire data set, which need to be identified and separated from noise or irrelevant outliers [17].

Azure Cognitive Services are AI solutions available in Microsoft Azure Cloud. Cognitive Services allow application developers to use those via API calls and hide the complexity of AI algorithms. Developers are able to use AI solutions in their applications in a standard way designed by Microsoft as software vendor and focus on business logic automation tasks instead.

Azure Cognitive Services provide solutions in following AI domains [18]:

- Speech:
- Speech to text;
- Text to speech;
- Speech translation;
- Speaker recognition;
- Language:
- Entity recognition;
- Sentiment analysis;
- Question answering;
- Conversational language understanding;
- Vision:
- Computer vision;
- Custom vision;
- Face API;
- Decision:
- **Anomaly detector;**
- Content moderator;
- Personalizer;
- Open AI service.

As we see “Anomaly detector” goes under AI services which support decision making.

In mathematics, a time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus, it is a sequence of discrete-time data. Examples of time series are heights of ocean tides, counts of sunspots, and the daily closing value of the Dow Jones Industrial Average [19].

Time series analysis comprises methods for analysing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values [19].

Anomaly detector service helps to detect anomalies in:

- uni-variate time series data;
- multivariate time series data.

As for all the other Cognitive Services the complexity and all the algorithms used to detect anomalies are hidden from developers. Developers can use API calls towards the solution and get the results from it.

In “Fig. 5” one can see an example of uni-variate time series graphical representation [18].

According to Microsoft documentation machine learning algorithms are used only when working with multivariate time series. For uni-variate time series Anomaly detector is processing using many different algorithms and mathematical methods which could be a part of processing multivariate time series as well.

Tony Xing describes the algorithms used in uni-variate anomaly detection process: “State of art anomaly detection system often uses a one size fit all approach. Which mean they apply some specific algorithm on all types of time series. Our learning is that each algorithm can handle some specific type of time series better. Our innovation is that we provide a generic framework to plug in different algorithm ensembles to handle a wide spectrum of different time series [20].

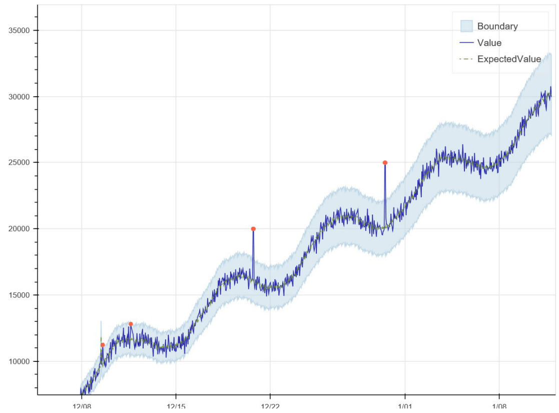


Fig. 5. Example of univariate time series [18]

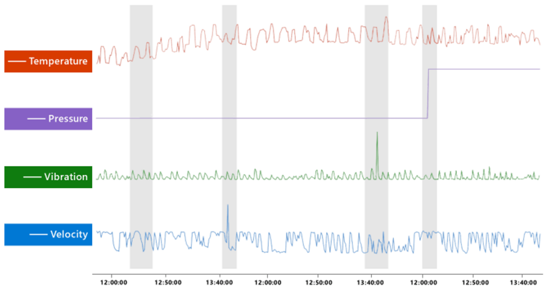


Fig. 6. Example of multi-variate time series [18]

In our different ensembles, we have following algorithms used [20]:

- Fourier Transformation;
- Extreme Studentized Deviate (ESD);
- STL Decomposition;
- Dynamic Threshold;
- Z-score detector;
- SR-CNN.

In “Fig. 7. and Fig. 8.” one can see algorithm selection process for uni-variate time-series [20]:

Detecting all kinds of anomalies through one single endpoint “Fig. 9.” [20]:

Besides spikes and dips, Anomaly detector also detects many other kinds of anomalies, such as trend change and off-cycle softness, all in one single API endpoint.

Microsoft Anomaly detector is available in Microsoft Azure cloud. It has API for such programming languages as Python, C sharp, etc. If organization plans to use Microsoft Azure cloud and build solutions using Anomaly detector, first it has to check availability of Anomaly detector in particular region of the Microsoft Azure cloud [21].

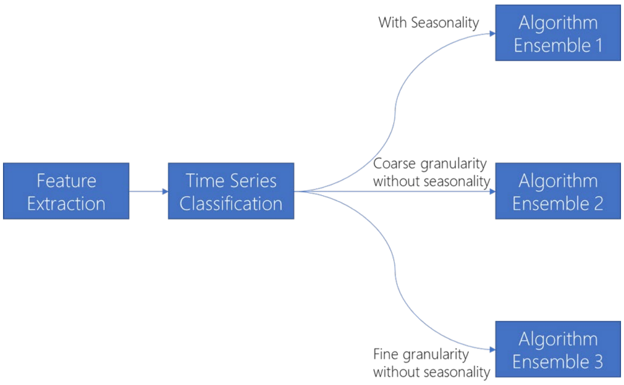

Feature Extraction
Time Series Classification
With Seasonality
Coarse granularity without seasonality
Fine granularity without seasonality
Algorithm Ensemble 1
Algorithm Ensemble 2
Algorithm Ensemble 3

Fig. 7. Algorithm selection process for univariate timeseries [20]

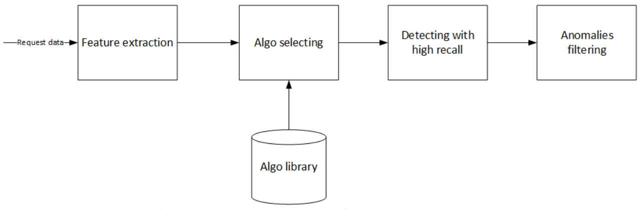


Fig. 8. Algorithm selection process for univariate timeseries [20]

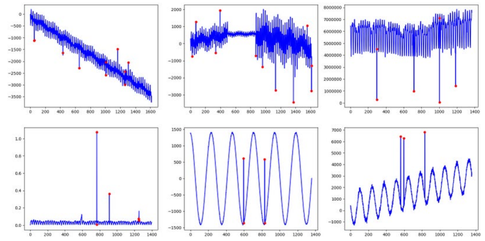


Fig. 9. Detecting all kinds of anomalies in univariate timeseries [18]

Recently Microsoft has released the graphical representation “Fig. 10.” of the framework which they do use in the process of multivariate anomaly detection. It was done by one of Anomaly detector project leaders via blog [22]. This is not the part of Anomaly detector documentation it is rather as description of the framework for customers and explanation on how the results are achieved.

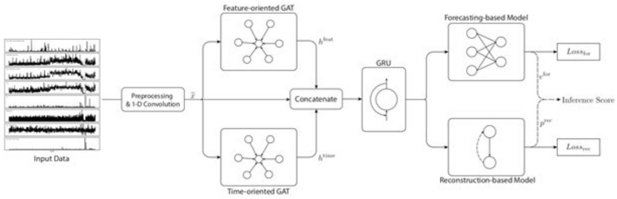


Fig. 10. Detecting anomalies in multi-variate timeseries [22]

Tony Xing describes the multivariate anomaly detection process: "In this newly introduced feature, we productized a novel framework - MTAD-GAT (Multivariate Time-series Anomaly Detection via Graph Attention Network), to tackle the limitations of previous solutions. Our method considers each univariate time-series as an individual feature and tries to model the correlations between different features explicitly, while the temporal dependencies within each time-series are modelled at the same time. The key ingredients in our model are two graph attention layers, namely the feature-oriented graph attention layer and the time-oriented graph attention layer. The feature-oriented graph attention layer captures the causal relationships between multiple features, and the time-oriented graph attention layer underlines the dependencies along the temporal dimension. In addition, we jointly train a forecasting-based model and a reconstruction-based model for better representations of time-series data. The two models can be optimized simultaneously by a joint objective function [22].

Graph Attention Networks is a subclass of Graph Neural Networks first presented to wider audience as a conference paper at ICLR 2018 (Sixth International Conference on Learning Representations, Vancouver Convention Center, Vancouver CANADA) [23].

From the above we can conclude that Anomaly detector in Microsoft cognitive services is built upon ideas from the latest scientific discoveries. When we see the API documentation we can notice that many parts of API are still developed further and documentation has the warning that API may change in the future.

The idea is that for GANs during matrix computations behind the scenes the weight of attention is calculated for a node on how much it should consider other nodes in the neighbourhood during the decision-making process and that is used as learning mechanism. Weights could change during the calculation process as new information comes from the neighbourhood.

III. Results and discussion

Anomaly detection is data processing challenge in different areas of natural sciences where results of research depend on data collection from sensors, databases, etc. Anomalies can be present in any data and could correspond to the data facts that scientists are interested into or could be the noise in data scientists would like to exclude from their processing and further analysis. The anomaly detection field in science is wide so in next section we will look into some anomaly detection real life usage scenarios.

A. Anomaly detection use case scenarios

In the real world, popular anomaly detection applications in deep learning include detecting spam or fraudulent bank transactions. Systems are already in place in most major banks where the authorities are alerted when unusually high spending or credit activity occurs on someone's account. The term "unusually high" can be defined on a user-to-user basis or collectively based on account type [24].

In industries, anomaly detection applications attached with machinery can help flag irregular or dangerous temperature levels or movement in parts or filter faulty materials (like filtering strange-looking food ingredients before they are processed and packed). Given that data can back the decision and sufficiently reliable data is available, anomaly detection can be potentially life-saving [24].

In a different use case, anomaly detection machine learning algorithms can also be used for classification tasks when the class imbalance in the training data is high. For instance, one can gather images of various species of flowers and plants for a multi-class classification task. However, substantially insufficient data is likely available for one particular species, thus resulting in an imbalance in the dataset. In such a case, the model can treat that class as an anomaly and classify the species differently. This is particularly relevant for medical diagnosis where there are only a few samples (images or test reports) where the disease is present, with the majority being benign. Anomaly detection can again be a life-saver in these cases [24].

HSBC, one of the largest banks in the world (more than 38 million customers), uses anomaly detection to deal with anti-money laundering (AML) issues. HSBC representative says: "It's a game changer because it's something we can scale when you operate in more than 60 jurisdictions. It will enable us to have consistency in how we do anti-money laundering" [3]. The bank said it would have an estimated 100 petabytes of data on Google Cloud by the end of 2018.

We can read about usage of anomaly detection for production line failure prediction in following paper [25]. As One of the main challenges here is mentioned that usually one production line could be used for production of many items based on the predefined input parameters. As production line could be changed by input parameters so the results of the sensor measurements do change when the input parameters are changed and this requires special learning cycle for anomaly detection for each product production line can produce.

Another example of anomaly detection is mentioned in article [26], where scientists review satellite data and anomalies in satellite data to correctly estimate biodiversity and other ecological factors of National forests in United States.

In article [27] Google historical stock data analysis is described with the help of Python library Scikit-learn [28].

With these examples we have covered just a small amount of use cases of anomaly detection and there is a space for further studies and further use case scenarios using different methods, tools and algorithms.

Author of this article plans to continue his further research in anomaly detection by finding solutions for different use case scenarios with the help of existing anomaly detection tools available in the market, for example, Microsoft Anomaly detector.

REFERENCES

[1] V. Chandola, A. Banerjee, V. Kumar, "Anomaly detection: A survey," ACM Comput. Surv., 2009, pp. 41, 1–72.

[2] P. Kattamuri, "How to build a serverless real-time credit card fraud detection solution," March, 2021. [Online] Available: https://cloud.google.com/blog/products/data-analytics/how-to-build-a-fraud-detection-solution [Accessed: Feb. 18, 2023].

[3] HSBC Holdings plc, "HSBC to launch AML system with google cloud," February, 2023. [Online] Available: https://www.pymnts.com/google/2023/google-pay-ditches-the-cvv-with-virtual-card-numbers-for-amex-holders/ [Accessed: Feb. 18, 2023].

[4] Google LLC, "Visual inspection ai," [Online] Available: https://cloud.google.com/solutions/visual-inspection-ai [Accessed: Feb. 18, 2023].

[5] M. Narang, "Anomaly detection in machine learning," March, 2023. [Online] Available: https://www.shiksha.com/online-courses/articles/anomaly-detection/ [Accessed: Feb. 23, 2023].

[6] A. Kargwal, "Anomaly detection in machine learning," August, 2022. [Online] Available: https://nimblebox.ai/blog/anomaly-detection-machine-learning [Accessed: Feb. 18, 2023].

[7] S. Kumar, "5 anomaly detection algorithms every data scientist should know," December, 2021. [Online] Available: https://towardsdatascience.com/5-anomaly-detection-algorithms-every-data-scientist-should-know-b36c3605ea16 [Accessed: Feb. 15, 2023].

[8] Y. Zheng, J. Guo, D. Ghent, K. Tansey, X. Hu, J. Nie, & S. Chen, "Land surface temperature retrieval from sentinel-3 a sea and land surface temperature radiometer, using a split-window algorithm. Remote Sensing, " 2019, pp. 11, 650. https://doi.org/10.3390/rs11060650

[9] T. Sushir, "Anomaly detection: Guide to prevent network intrusions," January, 2023. [Online] Available: https://geekflare.com/anomaly-detection/ [Accessed: Feb. 26, 2023].

[10] The MathWorks, Inc, "Identify unexpected events and departures from normal behavior," [Online] Available: https://la.mathworks.com/discovery/anomaly-detection.html [Accessed: Feb. 23, 2023].

[11] Wikipedia, "Isolation forest," [Online] Available: https://en.wikipedia.org/wiki/Isolation_forest [Accessed: Feb. 22, 2023].

[12] Wikipedia, "Anomaly detection," [Online] Available: https://en.wikipedia.org/wiki/Anomaly_detection [Accessed: Feb. 15, 2023].

[13] Wikipedia, "Mahalanobis distance," [Online] Available: https://en.wikipedia.org/wiki/Mahalanobis_distance [Accessed: Feb. 22, 2023].

[14] Dataconomy Media GmbH, "Anomaly detection in machine learning", [Online] Available:

[15] https://dataconomy.com/2022/10/machine-learning-anomaly-detection/ [Accessed: Feb. 20, 2023].

[16] Wikipedia, "Local outlier factor," [Online] Available: https://en.wikipedia.org/wiki/Local_outlier_factor [Accessed: Feb. 22, 2023].

[17] J. Yoon, S. O. Arik. "Unsupervised and semi-supervised anomaly detection with data-centric ML," February, 2023. https://ai.googleblog.com/2023/02/unsupervised-and-semi-supervised.html [Accessed: Feb. 18, 2023].

[18] Wikipedia, "Anomaly detection," [Online] Available: https://en.wikipedia.org/wiki/Anomaly_detection [Accessed: Feb. 15, 2023].

[19] Microsoft Corporation, "Azure cognitive services – overview," [Online] Available: https://azure.microsoft.com/en-us/products/cognitive-services/#overview [Accessed: Feb. 22, 2023].

[20] Wikipedia, "Time series," [Online] Available: https://en.wikipedia.org/wiki/Time_series [Accessed: Feb. 22, 2023].

[21] T. Xing, "Introducing azure anomaly detector api," April, 2019. [Online] Available: https://techcommunity.microsoft.com/t5/ai-customer-engineering-team/introducing-azure-anomaly-detector-api/ba-p/490162 [Accessed: Feb. 22, 2023].

[22] Microsoft Corporation, "Azure cognitive services – overview," [Online] Available: https://azure.microsoft.com/en-us/explore/global-infrastructure/products-by-region/?products=cognitive-services&regions=all [Accessed: Mar. 5, 2023].

[23] T. Xing. "Introducing multivariate anomaly detection," April, 2021. [Online] Available: https://techcommunity.microsoft.com/t5/ai-cognitive-services-blog/introducing-multivariate-anomaly-detection/ba-p/2260679 [Accessed: Feb. 22, 2023].

[24] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, "Graph attention networks," February, 2018. [Online] Available: https://arxiv.org/abs/1710.10903 [Accessed: Mar. 27, 2023].

[25] Iconiq Inc., "Anomaly detection using machine learning in python," February, 2023. [Online] Available: https://www.projectpro.io/article/anomaly-detection-using-machine-learning-in-python-with-example/555 [Accessed: Feb. 23, 2023].

[26] A. Graß, C. Beecks, J. A. C. Soto, "Unsupervised anomaly detection in production lines," In J. Beyerer, C. Kühnert, O. Niggemann (Eds.), "Machine learning for cyber physical systems", Springer Berlin Heidelberg, 2019, pp. 18–25.

[27] J. A. Knott, G. C. Liknes, C. L. Giebink, S. Oh, G. M. Domke, R. E. McRoberts, V. F. Quirino, B. F. Walters, "Effects of outliers on remote sensing-assisted forest biomass estimation: A case study from the United States national forest inventory," March, 2023. [Online] Available: https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.14084 [Accessed: Feb. 23, 2023].

[28] A. Naib, "Anomaly Detection on Google Stock Data 2014 – 2022," February, 2023. [Online] Available: https://www.analyticsvidhya.com/blog/2023/02/anomaly-detection-on-google-stock-data-2014-2022/ [Accessed: Mar. 5, 2023].

[29] A. Naib, "Complete guide on How to learn Scikit-Learn for Data Science," August, 2021. [Online] Available: https://www.analyticsvidhya.com/blog/2021/08/complete-guide-on-how-to-learn-scikit-learn-for-data-science/ [Accessed: Mar. 5, 2023].