Variational Autoencoder for Anomaly Detection in Event Data in Online Process Mining

Philippe Krajsic[1][a] and Bogdan Franczyk[1,2][b]

[1]Business Information Systems Institute, Leipzig University, Grimmaische Straße 12, Leipzig, Germany

[2]Business Information Systems Institute, Wrocław University of Economics, ul. Komandorska 118-120, Poland

Keywords: Anomaly Detection, Business Process Management, Deep Generative Model, Process Mining, Variational Autoencoder.

Abstract: The analysis of event data recorded by information systems is becoming increasingly relevant. An increasing data-centric analysis of processes by using process mining techniques has a direct impact on the management of business processes. To achieve a positive impact on business process management, a high quality data basis is important. This paper presents an approach for the application of variational autoencoder for the filtering of anomalous event data in an online process mining environment, which help to improve the results of process mining techniques and thus positively influence business process management. For anomaly detection in an unsupervised environment, mass-volume and excess-mass scores are used as metrics. The results are compared on the basis of established algorithms such as one-class support vector machine, isolation forest and local outlier factor. These insights are used to highlight the benefits of this approach for process mining and business process management.

1 INTRODUCTION

Process Mining (van der Aalst, 2016) is a new analytical discipline for identifying, monitoring and improving real processes, where existing data is extracted from event logs provided by information systems. As real processes (e.g. logistics, loan application, payment) become more and more dynamic and complex, it is crucial to be able to analyze these processes in real time and to react adequately to deviations and inconsistencies. The real-time reaction to inconsistencies within the processes highlights new potentials, such as more efficient process design, which goes hand in hand with reducing and preventing losses. The associated analysis and cleansing of these event logs is becoming increasingly important. Existing analysis methods typically assume that the input event data is completely free of incorrect data and infrequent behavior, which does not usually correspond to reality (van der Aalst et al., 2004) (Leemans et al., 2013). Incorrect data in event streams can lead to incorrect results during further processing. For example, the accuracy of drift detection can be negatively affected by stochastic vibrations due to inaccurate event streams (Maaradji et al., 2017) (Ostovar et al., 2016). Approaches that have been conducted in this area using filtering techniques to eliminate erroneous events from event data show an improvement in the quality of process mining techniques which leads to an optimization of the analysis of the processes (Wang et al., 2015) (Conforti et al., 2017). The majority of these approaches on event data based anomaly detection addresses batch processing, i.e. the processing of historical data. In order to take full advantage of the possibilities offered by anomaly detection methods, it is necessary to transfer these methods to an online setting. This enables the use of anomaly detection methods and subsequent process mining techniques in operational support and allows processes to be influenced in real-time due to deviations in process flow.

To address this challenge this paper proposes a new approach to detect anomalies and infrequent behavior in an online process mining setting using deep generative models, especially variational

[a] https://orcid.org/0000-0002-1722-8752
[b] https://orcid.org/0000-0002-5740-2946

Krajsic, P. and Franczyk, B.
Variational Autoencoder for Anomaly Detection in Event Data in Online Process Mining.
DOI: 10.5220/0010375905670574
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 1*, pages 567-574
ISBN: 978-989-758-509-8

autoencoders. The contribution of this work to information system research can be summarized as follows: (i) demonstration of the use of a deep generative model, like the variational autoencoder used in this work, for unsupervised anomaly detection in process event data, (ii) its embedding into an online process mining environment for real-time operational use and (iii) highlighting the benefits of high-quality event data using variational autoencoder for process mining and business process management.

The remainder of the paper is structured as follows: Section 2 provides a background on process mining and event data and a look at related work in the area of anomaly detection approaches in event data in a process mining context. Section 3 presents the developed approach for anomaly detection in event data using a variational autoencoder. This section contains the methodology for the selection of the technology, the formal description of the variational autoencoder and the integration of the variational autoencoder in an online process mining environment. Section 4 presents the approach for anomaly detection. In Section 5 a preliminary technical experiment of the presented variational autoencoder is conducted. Section 6 evaluates the presented approach with regard to the benefits for process mining and related areas. The final section 7 concludes the paper and presents future work.

2 BACKGROUND AND RELATED WORK

In the management of business processes, new technologies have emerged in recent years that increase the quality of business processes. In the following, the extension of business process management with data-driven process mining and the event data that is central to it are briefly presented. Subsequently, existing approaches from the literature are presented, which describe the filtering of anomalies and incorrect behavior from event data and thus increase the process quality.

2.1 Business Process Management and Process Mining

Business Process Management (BPM) is a broad discipline and a systematic approach to planning, controlling, monitoring and improving business processes. The focus here is on the organization of a company's business processes with the aim of coordinating and improving them through appropriate planning and weighting in order to increase both effectiveness and efficiency in the manufacture of products and services (Gabler, 2020). The main focus of the current BPM literature is on the control flow perspective. With the ongoing development in data processing, this perspective does not lose its importance but has to be complemented by other data-centric components in order to lead to a real improvement of business processes (van der Aalst et al., 2016). An important step towards a symbiosis of model-driven and data-driven BPM is the use of the event data available in the information systems of the organizations. The idea behind the integration of data-centric analysis methods is to minimize the variability of the processes by early detection of anomalies and misbehavior in the event data.

From this necessity the process mining technology has emerged. The explorative, automated recognition of business processes is the focus of process mining. Furthermore, it bridges the gap between traditional model-based process analysis (e.g. business process improvement) and data-centric analysis techniques (e.g. machine learning) (van der Aalst, 2016). It also offers new ways of extracting knowledge from data generated and stored in the databases of (enterprise) information systems to generate event data and can be used in a variety of areas, such as automatically discover process models, checking compliances with reference models or determining the cause for different process variants.

Due to these characteristics and application possibilities process mining offers additional potentials that can be used in the context of BPM. For example, processes can be enhanced with additional information such as cycle times and resource utilization, which can be used to improve the process and contribute to the strategic objectives of the organization. Process Mining enables organizations to take a deeper look into their end-to-end processes. As a result, process mining methods are now used in all phases of the BPM life cycle to improve the actual processes with the help of the available event data (van der Aalst et al., 2016).

2.2 Event Data

Process Mining evaluates event data that was recorded during the execution of a process. An event is any data that is recorded during the execution of a process and is considered the smallest unit within a process. The granularity of an event depends on the application domain as well as the way it is recorded. For example, an event can describe which activity of a process was executed at what time. In the same way, it can also describe the different stages of the execution of an activity, e.g. events refer to the scheduling, starting, suspending, continuing or completing of an activity (van Zelst et al., 2018). An event is an assignment of values to a set of attributes.

In order to optimally use the potential of process mining technology for operational support, it is important to perform an analysis of the generated event data in real-time. Such real-time event streams are chronologically ordered sequences of unique events (van Zelst et al., 2018).

The event data extracted from the information systems (e.g., ERP systems) forms the basis for all further process mining activities. Therefore, the accuracy of this event data is essential for a high process quality.

2.3 Related Work

Regarding the approaches to filtering methods of event data in the context of process mining, the current state of research can be described as follows: With regard to detection and filtering of anomalies in event logs, there are some approaches described in the literature. In (Wang et al. 2015) a reference model is used to detect inappropriate behavior and repair the affected log. The approach proposed in (Conforti et al., 2017) is based on an automaton which is modeled on the frequent process behavior recorded in the logs. (Sani et al., 2017) proposes an approach that uses conditional probabilities between activity sequences to eliminate events that are unlikely in a particular sequence. In (Nolle et al., 2018) an approach using autoencoders for event classification is proposed. A stand-alone approach for filtering anomalies in event streams is offered in (van Zelst et al., 2018). It proposes an event processor that allows to effectively filter out unwanted events from an online event stream, based on probabilistic automaton.

With regard to the application of deep generative models in a business process environment the work of (Taymouri et al., 2020) should be mentioned that proposes an adversarial training framework based on an adaptation of generative adversarial networks to the realm of sequential temporal business event data.

Approaches that address the filtering of event data in an online setting are very limited (van Zelst et al., 2018). These mostly use statistical methods that make a priori assumptions about the underlying relationship of the variables used (Ahn et al., 2020). This leads to the difficulty that underlying probability distributions and variable relationships have to be adjusted for each use case. Furthermore, statistical models are often not suitable for processing high-dimensional data (Ahn et al., 2020). For these reasons, the advantages of deep generative models and its application in the field of anomaly detection in business event data are considered in this work.

3 VARIATIONAL AUTOENCODER FOR ANOMALY DETECTION

In order to optimize these weaknesses in the processing of event streams, a method is to be established that enables self-learning and unsupervised filtering of anomalies from event streams for an improved process flow. Therefore, the selection of the variational autoencoder used in this work is described in the following. Based on this, a formal description of the variational autoencoder and the integration of the variational autoencoder in an online process mining setting is given.

3.1 Deep Generative Models

In addition to the disadvantages of statistical techniques, the required specialized process knowledge and the extraction of suitable features from the examined data also poses a relevant challenge. The development of new models in the areas of deep learning and time series analysis make it possible to reduce precisely this need. With the help of special deep learning methods a specific selection of data features is possible. This includes methods such as convolutional neural networks, recurrent neural networks, generative adversarial networks and variational autoencoder. Especially in the context of anomaly detection in event streams, the superiority of deep learning techniques compared to statistical techniques is shown (Gamboa, 2017) (Ahn et al., 2019) (OMeara et al., 2018). So called deep generative models (DGM) are a special form of deep learning techniques. DGM use deep neural networks to parameterize the conditional distributions of the observed data. The advantage of these models is that they can be used for a variety of applications and, in contrast to discriminative models, can learn from both labelled and unlabelled data (Bishop, 2006). DGM include in particular the generative adversarial network (GAN) and variational autoencoder (VAE). The main difference of GANs compared to VAEs is that they are implicit, i.e. the likelihood of the samples produced cannot be evaluated directly (Goodfellow et al., 2016). Therefore, GANs are usually trained solely with adversarial procedures.

These implicit probabilities make it inherently difficult to treat their parameters in an appropriate manner. It is therefore unclear how to prioritize weightings and learn approximate posterior distributions. Out of these mentioned disadvantages of the GANs, VAEs are to be used for anomaly detection in process event streams addressed in this work.

3.2 Formal Description of VAE

The approach of an VAE was first introduced by (Kingma and Welling, 2014). In order to make the functioning of the VAE comprehensible, it will be briefly described here. For a more formal description of the VAE, please refer to (Kingma and Welling, 2014).

Figure 1 shows an exemplary architecture of a VAE. VAE are latent variable models (Doersch, 2016) (Kingma and Welling, 2019). Such models are based on the idea that the data generated by a model can be parameterized by a number of variables that create some specific characteristics of a given data point. These variables are called latent variables. One of the main ideas behind VAE is that instead of trying to explicitly construct a latent space (space of latent variables) and to sample from it in order to find samples that could actually produce correct outputs (as close as possible to our distribution), an encoder-decoder-like network that is divided into two parts is constructed (Kingma and Welling, 2019):

Probabilistic Encoder: The encoder inputs a datapoint x and outputs a hidden representation z, and has weights and biases θ. The encoder must learn an efficient compression of the data into this lower-dimensional space z. We denote the encoder $q_\theta(z|x)$.

The lower-dimensional space is stochastic, so the encoder outputs parameters to $q_\theta(z|x)$, which is a Gaussian probability density. The probabilistic encoder $q_\theta(z|x)$ produces the mean μ and standard deviation σ of a normal distribution. We can sample from this distribution to get noisy values of the representations z.

Probabilistic Decoder: Like the encoder, the decoder is another neural net. The decoder inputs the representation z and outputs the parameters to the probability distribution of the data, and has weights and biases ϕ. The decoder is denoted by $p_\phi(x|z)$. To find out how much information was lost during decoding we measure the reconstruction log-likelihood $p_\phi(x|z)$. This measures how effectively the decoder has learned to reconstruct an input x given its latent representation z.

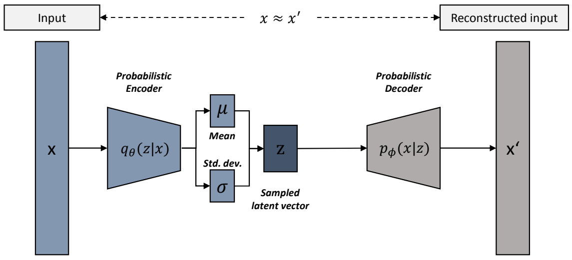

Figure 1: Example architecture of a variational autoencoder (Kingma and Welling, 2019).

The loss function of the VAE is calculated by the reconstruction loss or expected negative log-likelihood of the i-th datapoint in the first term and the Kullback-Leibler divergence (Kullback and Leibler, 1951) in the second term:

$$l_i(\theta,\phi) = -E_{z\sim q_\theta(z|x_i)}\left[\log p_\phi(x_i|z)\right] + KL(q_\theta(z|x_i)||p(z)) \tag{1}$$

The lost function is a method of evaluating how well the VAE models the given data and calculates how much information is lost during the reconstruction of the data. A bad reconstruction of the data automatically leads to high costs in the loss function and thus to a high reconstruction error. The aim is to keep this reconstruction error as low as possible so that $x \approx x'$

3.3 Using VAE in Online Process Mining

In this section the integration of the VAE in an online process mining setting is described. A suitable embedding in the process mining workflow leads to better process mining activities through filtered event streams and thus to an optimized BPM. The advantage of integrating the VAE into a lambda architecture (Marz and Warren, 2013) results from the processing of streaming data as well as historical data, which are relevant when using process mining techniques. A more detailed description of the use of a lambda architecture for anomaly detection in event streams is offered in (Krajsic and Franczyk, 2020). Figure 2 illustrates the embedding of the VAE (event filter) in an online process mining setting. A batch layer with the provision of historical data and the speed layer with the processing and filtering of streaming data from the real-time process are merged by the serving layer in which both offline and online process mining take place and lead to the discovery and analysis of process models.

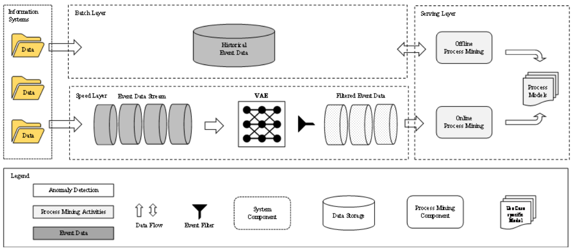

Figure 2: Integrated VAE in an online process mining setting (Krajsic and Franczyk, 2020).

The real-time event data filtered by the VAE and used during online process mining is then transferred to the historical data storage which can later be used as historical data for new process mining tasks.

In this representation of an online process mining environment, the VAE thus serves as a buffer between the streaming data from the speed layer and the historical data from the batch layer and ensures a clean data basis before merging the two layers, which leads to the discovery or analysis of process models in the serving layer through online and offline process mining activities.

4 APPROACH

Since the detection of anomalies is a classification problem, the desired output of a corresponding classification method is a class label. Through the use of neural networks it is possible to get such an output without training the neural network with class labels. This unsupervised learning corresponds more to the procedures in real world examples, since data that is to be generated and checked in real time is usually not labelled.

Before the VAE can be used for the unsupervised anomaly detection, the underlying model must be trained with normal, non-anomalous data. Once the VAE has been trained with the training data set and verified using test and validation data set, it can be used for anomaly detection. The anomaly detection process is described in more detail as follows.

For the purpose of anomaly detection, a latent space is initially created using the VAE. A clustering algorithm and different anomaly detection algorithms are applied to this latent space. For clustering the k-means algorithm (Lloyd, 1982) is used. For the purpose of anomaly detection Isolation Forrest (IF) (Liu et al., 2008), Local Outlier Factor (LOF) (Breunig et al., 2000) and One-Class Support Vector Machine (OCSVM) (Schölkopf et al., 2000) are used as anomaly detection algorithms.

4.1 Method of Anomaly Detection

In this unsupervised anomaly detection setting no ROC or Precision-Recall curves can be generated for evaluation purposes, because there is no possibility to check the results against a ground truth. For this purpose, other evaluation options, such as mass-volume and excess-mass (EM-MV) based criteria, are suitable (Goix, 2016). The EM-MV method can be formally described as follows (Goix, 2016):

The goal of this method is to estimate the density level curves of the probability distribution under observation, assuming that anomalies occur at the tail end of this distribution. We introduce f as a given constant $c, L_c(f) = \{(x_1, \ldots, x_n) | f(x_1, \ldots, x_n) = c\}$. In the presented case the function f is the probability density estimated by the VAE in latent space. To be able to determine the degree of anomaly, a score function s is introduced: $\mathbb{R}^d \to \mathbb{R}_+$ with data in $\mathbb{R}^d$. The MV and EM curves of s can be written as

$$MV_s(\alpha) = \inf_{u \geq 0} Leb(s \geq u) s.t. \;\; P(s(X) \geq u) \geq \alpha, \tag{2}$$

$$EM_s(t) = \sup_{u \geq 0} \{P(s(X) \geq u) - tLeb(x \geq u)\}, \tag{3}$$

where any $\alpha \in (0,1)$ and $t > 0$. With this knowledge the chosen method can be evaluated by calculating the distance between the level sets of f and s for EM and MV: $\|EM_s - EM_f\| L^1(I)$ and $\|MV_s - MV_f\| L^1(J)$ (Goix, 2016). I is defined as $I = [0.9, 0.999]$ and J as $J = [0, EM_s^{-1}(0.9)]$ where $EM_s^{-1}(0.9) = \inf\{t \geq 0, EM_s(t) \leq 0.9\}$ (Goix, 2016).

The measure is then determined based on the area under the EM_s and MV_s curves. For EM_s this area should be maximized and for MV_s minimized.

5 TECHNICAL EXPERIMENT

The goal of the VAE in this work is to detect anomalies in process event data. Thereby the VAE is trained with normal process data, in an unsupervised way, and finally applied to process data that contain additional abnormal data. The VAE then attempts to detect anomaly events by determining their EM and MV scores.

In order to illustrate the functionality of the VAE presented here, a technical experiment will be conducted and its results will be presented in this section. Based on the representation of latent space generated by the VAE, the k-means clustering algorithm is applied to form clusters from the data. Based on this, the EM-MV method is applied with different anomaly detection algorithms to detect the anomalies in the event data. The data used for this experiment is taken from an loan application dataset (4TU.ResearchData, 2013). The dataset is a collection of artificial event logs representing a simple loan application process, used in the form of a CSV file for the experiment.

5.1 Experimental Setup

Simulating a rapid deployment process, no hyper-parameter tuning was performed for this technical experiment. In an first preliminary implementation the dimension of the latent space of the VAE is set to 2. Furthermore, the VAE was trained on batches of size 64 for 100 epochs. Adam was used as an optimizer (Kingma and Ba, 2014). The learning rate for the model is set to 0.001 with the default settings of the adam optimizer. The selected data set was partitioned into training, test and validation data set, with the test set being 20% and the validation set being 10% of the full data.

For the experiment anomalous event data was added to the test and validation data set. For this purpose anomalous sequences were generated and integrated into the loan application data set for test and validation purposes. The generated noisy event logs are effected by the following mutations: *Event Skipping, Event Swapping or Event Duplicating*.

5.2 Experimental Results

In the following the results of the technical experiment are presented. Based on the latent space representation from the VAE we apply k-means clustering algorithm with a five clusters based on the number of different activities in the observed event log. We can apply the EM-MV method on the basis of the clusters. Three different anomaly detection algorithms, IF, LOF and OCSVM, are used. Table 1 gives a comparison of excess-mass (EM) and mass-volume (MV) scores for the different anomaly detection methods.

The goal of this approach is to maximize the EM score and minimize the MV score. As can be seen from Table 1, the OCSVM best meets this criteria.

Table 1: EM and MV scores based on VAE latent space.

Algorithm	$EM_S(10^{-3})$	MV_S
OCSVM	**6.738**	**1.152**
LOF	5.760	1.308
IF	5.349	1.377

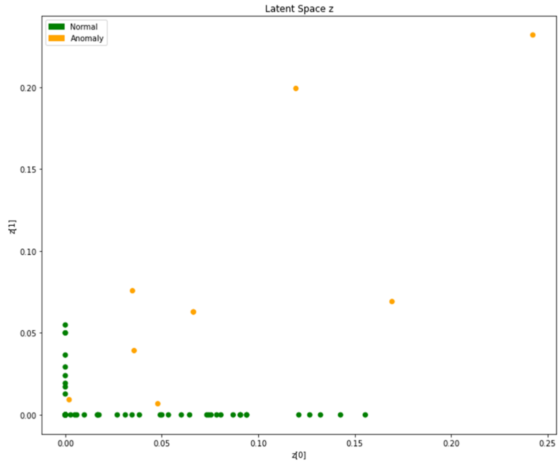

Figure 3: Representation of latent space.

Based on this results OCSVM is used as an algorithm for anomaly detection purposes on the generated latent space.

The OCSVM is an unsupervised learning algorithm that is trained only on non-anomalous data. To detect outliers the boundaries of these data points are learned to classify those data points as anomalies that lie outside these boundaries.

Figure 3 shows the latent space of the data space generated by the VAE. Data points within the latent space marked in green were recognized by the OCSVM algorithm as inconspicuous. The yellow marked data points, on the other hand, are structurally different due to their arrangement in latent space and lie outside of the boundaries defined by the OCSVM and are therefore classified as anomalous event data.

By filtering out the data points marked as anomalies, the data quality of the underlying event data used for process mining can be improved.

6 BENEFITS FOR PROCESS MINING AND BUSINESS PROCESS MANAGEMENT

The use of methods for filtering erroneous events and infrequent behavior from the event data collected in information systems leads to an improvement of the results in the analysis of processes.

The characteristics and advantages offered by the VAE have a considerable influence on the application area of process mining presented here and its comprehensive field of business process management. The advantages and potential benefits for process mining and business process management are outlined below.

Process Mining. The partially superficial approaches to the analysis of the underlying event data made a beneficial extraction of suitable process knowledge more difficult. In this context, process mining activities are not limited by the availability of data but by the quality of the underlying event data. This issue has received little attention in previous process mining research. Especially the application of process mining techniques to real world applications is made difficult by data quality challenges. For the extraction of meaningful process knowledge from the event data and high-quality process mining results, the improvement of the data basis is therefore indispensable. The presented approach contributes significantly to a filtering of faulty event data for an increased data quality. As an upstream step in the processing of real-time event data streams, it enables a higher quality of the data basis for downstream process mining activities by filtering out erroneous event data and incorrect behavior.

Business Process Management. Increasing the quality of the underlying event data for an efficient application of process mining techniques not only leads to an improvement of the results of the process mining activities but also has a direct impact on the higher level business process management and the implementation of and compliance with the strategic process and business goals. The inclusion of a high-quality data-centered view in the classic control flow perspective as a fixed pillar of BPM and workflow management research leads to a more realistic view through the inclusion of real world event data than the idealized as-is and target processes of an organization obtained in workshops and interviews (van der Aalst, 2011). In addition, the inclusion of filtered real world event data and the application of process mining techniques leads to the introduction of improved process analysis and process automation solutions that would have been difficult to achieve with model-based approaches. These developments in process mining, supported by high-quality event data, make it possible to replace the previous restrictions in BPM systems with partially hard-coded, custom-made processes (van der Aalst et al., 2016) with flexible, near-real-time processes and thus provide the organization with near-real-time operational support for BPM.

7 CONCLUSION AND FUTURE WORK

In this work, it was shown how variational autoencoder can be used for filtering erroneous event data and how this can affect process mining and the higher-level business process management.

For real-time process support, the VAE was integrated into an online process mining environment that filters incoming events for incorrect data and behavior. To demonstrate the functionality of the VAE, the VAE was applied to an artificial loan application process. The results of the conducted technical experiment have shown that the VAE enables the unsupervised detection of anomalies in process event data.

A preliminary limitation of the presented approach is that the model does not consider long-term dependencies in process data, which means that certain incorrect behavior cannot currently be detected. Furthermore, a transfer of the achieved results to the governance of business processes and derivation of recommendations for action is necessary. From the perspective of business process management, it would be necessary to conduct extensive case studies in order to be able to quantify the exact potential and added value of the approach presented.

The limitations mentioned above will be taken into account in future work. These include the integration of methods for the analysis of long-term dependencies, such as long short-term memory or transformers. This enables the recognition of more complex patterns and behavior. To quantify the business potential of the presented approach, case studies will be conducted with partners from science and practice to verify the added value of the approach under real world conditions.

ACKNOWLEDGEMENTS

This work was supported by the German Federal Ministry of Education and Research (BMBF, 01/S18026A-F) by funding the competence center for Big Data and AI “ScaDS.AI Dresden/Leipzig“.

REFERENCES

Ahn, H., Choi, H.-L., Kang, M., Moon, S., 2019. Learning-Based Anomaly Detection and Monitoring for Swarm Drone Flights. In Appl. Sci. 9, p. 5477.

Ahn, H., Jung, D., Choi, H.-L., 2020. Deep Generative Models-Based Anomaly Detection for Spacecraft Control Systems. In Sensors , 20, p. 1991.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer.

Breunig, M.M., Kriegel, H., Ng, R., Sander, J., 2000. LOF: Identifying Density-Based Local Outliers. In Proc. ACM Sigmod Int. Conf. on Management of Data, Dallas.

Conforti, R., La Rosa, M., ter Hofstede, A.H.M., 2017. Filtering Out Infrequent Behavior from Business Process Event Logs. In Transactions on Knowledge and Data Engineering 29 (2), pp. 300 – 314.

Doersch, C., 2016. Tutorial on variational autoencoders. In arXiv preprint, arXiv:1606.05908.

Gabler Banklexikon, https://www.gabler-banklexikon.de/definition/business-process-management-70709/version-377663 (Accessed: 12.10.2020).

Gamboa, J.C.B., 2017. Deep learning for time-series analysis. In arxiv preprint, arXiv:1701.01887.

Goix, N., 2016. How to Evaluate the Quality of Unsupervised Anomaly Detection Algorithms? In arXiv preprint, arXiv:1607.01152v1.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. In arXiv preprint, arXiv:abs/1412.6980.

Kingma, D.P., Welling, M., 2019. An introduction to variational autoencoders. In arXiv preprint, arXiv:1906.02691.

Kingma, D.P., Welling, M., 2014. Auto-encoding variational Bayes. In 2nd International Conference on Learning Representations, CoRR abs/1312.6114.

Krajsic, P., Franczyk, B. 2020. Lambda Architecture for Anomaly Detection in Online Process Mining using Autoencoders. In Hernes M., Wojtkiewicz K., Szczerbicki E. (eds) Advances in Computational Collective Intelligence. ICCCI 2020. Communications in Computer and Information Science, vol 1287, pp. 579-589, Springer.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. In Ann Math Stat 22, pp.79–86.

Leemans, S.J.J., Fahland, D., van der Aalst, W., 2013. Discovering Block-Structured Process Models from Event Logs – A Constructive Approach. In International Conference on Application and Theory of Petri Nets and Concurrency, pp. 311-329, Springer.

Liu, F.T., Ting, K.M., Zhou, Z.-H., 2008. Isolation forest. In Proc. of the 8th IEEE International Conference on Data Mining, Pisa, pp.413-422.

Lloyd, S., 1982. Least squares quantization in PCM. In IEEE transactions on information theory 28 (2), pp. 129-137.

Maaradji, A., Dumas, M., La Rosa, M., Ostovar, A., 2017. Detecting Sudden and Gradual Drifts in Business Processes from Execution Traces. In Transactions on Knowledge and Data Engineering 29 (10), pp. 2140 – 2154.

Marz, N., Warren, J., 2013. Big Data. Priciples and Best Practices of Scalable Realtime Data Systems, Manning.

Nolle, T., Luettgen, S., Seeliger, A., Mühlhäuser, M., 2018 Analyzing Business Process Anomalies using Autoencoders. In Mach Learn 107, pp. 1875–1893.

OMeara, C., Schlag, L., Wickler, M., 2018. Applications of Deep Learning Neural Networks to Satellite Telemetry Monitoring. In: Proceedings of the 2018 SpaceOps Conference, p. 2558.

Ostovar, A., Maaradji, A., La Rosa, M., ter Hofstede, A.H.M., van Dongen, B.F., 2016. Detecting Drift from Event Streams of Unpredictable Business Processes. In International Conference on Conceptual Modeling, pp. 330 – 346, Springer.

Sani, M.F., van Zelst, S.J., van der Aalst, W., 2017. Improving Process Discover Results by Filtering Outliers using Conditional Behavioral Probabilities. In Teniente E., Weidlich M. (eds.) Business Process Management Workshops. Lecture Notes in Business Information Processing, vol 308, pp. 216-229, Springer.

Schölkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., Platt, J., 2000. Support Vector Method for Novelty Detection. In Advances in Neural Information Processing Systems 12, pp. 582-588.

Taymouri, F., La Rosa, M., Erfani, S., Bozorgi, Z.D., Verenich, I., 2020. Predictive Business Process Monitoring via Generative Adversarial Nets: The Case of Next Event Prediction. In Fahland D., Ghidini C., Becker J., Dumas M. (eds) Business Process Management. BPM 2020. Lecture Notes in Computer Science, vol 12168. Springer, Cham.

van der Aalst, W.M.P., 2011. Process mining: discovery, conformance and enhancement of business processes. Springer.

van der Aalst, W.M.P., La Rosa, M., Santoro, F. M., 2016. Business Process Management. Don't Forget to Improve the Process! In Business & Information Systems Engineering 58 (1), pp. 1-6, Springer.

van der Aalst, W., 2006. Process Mining. Data Science in Action. Springer, Berlin, 2nd edition.

van der Aalst, W., Weijters, A.J.M.M., Maruster, L., 2004. Workflow Mining: Discovering Pro-cess Models From Event Logs. In Transactions on Knowledge and Data Engineering 16 (9), pp. 1128-1142.

van Zelst, S.J., Fani Sani, M., Ostovar A., Conforti, R., La Rosa, M., 2018. Filtering Spurious Events from Event Streams of Business Processes. In Advanced Information Systems Engineering. CAiSE 2018. Lecture Notes in Computer Science, vol, 10816, pp. 35-52, Tallin, Springer.

Wang, J., Song, S., Lin, X., Zhu, X., Pei, J., 2015. Cleaning Structured Event Logs: A Graph Re-pair Approach. In 31st International Conference on Data Engineering, pp. 30- 41, IEEE Press.

4TU.ResearchData, Loan application example, configuration 1., https://doi.org/10.4121/uuid:cdf3ba31-291d-468d-9712-3a58ac6da3fc (Accessed: 22.10.2020).