242 (2024) 213240

Contents lists available at ScienceDirect

Geoenergy Science and Engineering

journal homepage: www.sciencedirect.com/journal/geoenergy-science-and-engineering

Oil and gas flow anomaly detection on offshore naturally flowing wells using deep neural networks

Guzel Bayazitova [a,*], Maria Anastasiadou [a,b], Vitor Duarte dos Santos [a,b]

[a] NOVA Information Management School, Campus de Campolide, 1070-312, Lisboa, Portugal
[b] MagIC - Information Management Research Center, Lisboa, Portugal

A R T I C L E I N F O

Keywords:
Anomaly detection
Multivariate time series classification
Deep neural network
LSTM
Oil well monitoring
Genetic algorithm

A B S T R A C T

The oil and gas industry is changing. The drive towards cleaner and safer operations is becoming increasingly important. Researchers are looking for more efficient and accurate ways to detect faults that could lead to environmental and sustainability issues. This study aims to enhance the safety and sustainability of the oil and gas industry by improving existing artificial intelligence approaches to automate monitoring and detection of malfunctions. This article explores the application of deep neural networks for anomaly detection in monitoring oil and gas flow in natural flow offshore wells, proposing an innovative approach that takes advantage of the power of Genetic Algorithms and Gated Recurrent Units (GRU). The study aims to enhance the safety and sustainability of the oil and gas industry by leveraging artificial intelligence to automate the monitoring and detection malfunctions. Utilizing a comprehensive dataset from the 3W Petrobras project, which includes real-time data from 21 wells collected between 2012 and 2018, the research focuses on detecting various anomalies such as abrupt increases in basic sediment and water, spurious closures of downhole safety valves, severe slugging, flow instability, rapid productivity loss, quick restrictions in the production choke, scaling, and hydrate formation in production lines. The methodology integrates Long Short-Term Memory (LSTM) networks and GRU backbones with genetic algorithms to optimise model performance. Several hyperparameter optimisation tools were explored innovatively, focusing mainly on Genetic Algorithms, and it was possible to obtain an algorithm with 2 stacked GRU with better comparative performance compared to what is reported in the literature and producing an F1 equal to 0.97. The findings demonstrate the potential of AI to improve real-time anomaly detection, thereby reducing operational risks and contributing to the industry's transition towards greener practices. It also underscores the importance of open data and collaborative efforts in advancing AI applications in the oil and gas sector, aligning with the United Nations' Sustainable Development Goals to mitigate climate impact and promote responsible consumption and production.

1. Introduction

Artificial Intelligence (AI) has revolutionized the perspectives of industries and businesses worldwide, creating value by learning from data, accumulating knowledge from patterns and trends, simulating human logic and automating decision making. It has been used in many industries, such as finance, smart cities, healthcare, cybersecurity, education, criminal justice, etc. (Koroteev and Tekic, 2021). Domain knowledge and expertise are the main pillars of decision-making. However, AI brings huge benefits by automating many processes, improving accuracy, and further augmenting human intelligence.

The Oil and Gas industry has been somewhat relatively slow in applying AI to many parts of the Exploration and Production (E&P) life cycle, which would enable a reduction in costs and risks (Kuang et al., 2021) and thereby converting the industry into a greener version of itself. Nowadays, AI has entered all its branches, creating many "intelligent" versions of the sectors, such as intelligent drilling, intelligent development, intelligent exploration, intelligent production, etc. (Sircar et al., 2021). However, there is still massive potential for further AI development within the industry, such that the E&P industry can reach the point where many industries enjoy the full scale of revolution 4.0.

Due to the demanding nature of the Oil and Gas industry, consistent real time monitoring is essential, requiring constant monitoring by surveillance engineers. This is a formidable task that has been in

* Corresponding author.
E-mail addresses: m20210699@novaims.unl.pt (G. Bayazitova), manastasiadou@novaims.unl.pt (M. Anastasiadou), vsantos@novaims.unl.pt (V.D. dos Santos).

https://doi.org/10.1016/j.geoen.2024.213240
Received 16 October 2023; Received in revised form 2 August 2024; Accepted 13 August 2024
Available online 16 August 2024

existence for many years (Hasan et al., 2017). In the Upstream sector, one engineer might be in charge of not just one well but multiple wells within one field, monitoring many aspects of the production from these wells. Considering the limited time window available between the start of the issue and the actual failure, anomaly detection in real time can significantly alleviate the surveillance task.

The Upstream sector is the most capital-intensive and important sector compared to Midstream and Downstream segments (Koroteev and Tekic, 2021). However, it is the least automated in many aspects due to the activity's nature, often performed in harsh conditions, such as deep-water, arid deserts, arctic colds, extreme wind, etc. Drilling is a high risk and high capital expenditure operation that involves fire and explosion risks, operations with radioactive sources, the threat of gas leaks, and the movement of personnel by unconventional means of transport, such as helicopters and supply boats, which offer many possibilities of human error due to the complexity of the operations.

Digital oilfield is a new concept, symbolising a collection of automation and information technologies that revolutionise how the petroleum industry operates and allows for work processes to be conducted more efficiently (Pandey et al., 2020). It involves data management, automation, integrated production models, predictive maintenance, etc. – a combination of all the emerging technologies that assist in timely reaction and decision making.

A unique challenge for the Oil and Gas industry is the lack and availability of open data (Vargas et al., 2019a), which hinders further research and AI application advancement. The main reason is the confidentiality of high-cost information, whilst another is the difficulty in recognising and labelling all potential unlikely events from the available data (Soriano-Vargas et al., 2021).

Another issue is the absence of an author's active date network, which was used to research the subject matter. It slows any advancement because most papers were not produced by academic researchers but by company appointed professionals, applying AI to specific core activities and using proprietary data, which does not encourage further networking (D'Almeida et al., 2022). The lack of collaboration between oil companies, perceiving each other as competitors, does not help in AI advancement within the petroleum industry (Koroteev and Tekic, 2021). Most companies tend to follow a strategy of developing their own AI projects without experience and knowledge sharing.

This research aims to make a novel contribution to the application of artificial intelligence in the oil and gas industry by improving the accuracy of the existing algorithms to strongly mitigate its adverse climate effect while enhancing its safety and sustainability. We will focus on malfunctions within naturally flowing wells since the availability of open data dictates this research. However, most wells require Artificial lift methods. Hence, more failures are observed due to the failure of pumps or turbines. This research would deepen the knowledge of anomaly detection in the oil and gas industry to demonstrate the unlimited potential of data science applications in the hydrocarbon and fossil fuel domain and to encourage all industry key players to implement AI-based processes more extensively and willingly. To make a comparative analysis, the Petrobras 3W dataset, already studied in previous works, was taken as an example of labelled time series data. Ten initial RNN models with LSTM and GRU architectures were tested, and the best was optimised using Random Search, Hyperopt and three Genetic Algorithms. Compared to previous work on the same data set (Turan and Jaschke, 2021; Santos et al., 2021; Gatta et al., 2022), the results achieved a higher F1 score, up to 0.97%, in the optimised models, particularly those using Genetic Algorithms.

The United Nations "17 Sustainable Development Goals" calls for a global partnership in collaboration for a Better World with less poverty, increased economic growth, tackled climate change, preservation of nature, etc. (United Nations, 2022). The fossil fuels sector emits significant amounts of GHG, which affects local ecosystems and the environment via each oil spill, produced water and drilling waste discharge. The UN goals № 12 "Responsible consumption and production" and № 13 "Climate action" are the first two that need to be addressed by the petroleum business. With a better understanding of timely decision-making and the impact of predictive maintenance, this research could contribute to accomplishing the United Nations' goals and potentially make our world a greener and safer place.

This paper is organised as follows. Section 2 presents the background of the 3W Petrobras dataset project and its objectives. Section 3 analyses all the related work published up to December 2022, describes the application of PRISMA, details the collected data from the survey, and presents the results using the visualisation and bibliometric tool. Section 4 defines the adopted methodology of the research, followed by further data preprocessing steps. Section 5 discusses our findings, which are aligned with our research questions, while Section 6 presents our conclusions.

2. Background

The primary goal of the 3W Petrobras project is the development of a new automated AEM (Abnormal Event Management) process with machine learning algorithms, for which the 3W dataset was created by compiling accurate data from 21 wells during actual operations from 2012 to 2018 (Vargas et al., 2019b). The naturally flowing wells were selected as less complex and more suitable for research and innovation in predictive maintenance.

Naturally flowing wells are those in which the formation pressure is sufficient to produce oil commercially without requiring a pump. Most reservoirs at the initial stage of development have enough pressure for a natural flow and thus require less equipment and automation for control and successful oil and gas production. Fig. 1 presents the basic schema of an offshore platform connecting to a subsea christmas tree through a production line and subsequently to production tubing and the reservoir itself. Subsea christmas trees are a complex assembly installed on top of the wellhead to monitor and control the production while operated through an electro-hydraulic umbilical.

The 3W dataset combines measurements from topside and subsea sensors located in the production tubing (P-PDG), on the subsea christmas tree (P-TPT and T-TPT), the production line (P-MON-CKP and T-JUS-CKP), and the gas lift line (P-JUS-CKGL, T-JUS-CKGL, and QGL) (Santos et al., 2021) (Table 1).

The 3W Dataset is organised into folders according to the type of fault, with each event progressing from regular operation to transient condition, following through to a steady-state anomaly. The 8 types of recognised and labelled events are:

- Class 0 – Normal operation

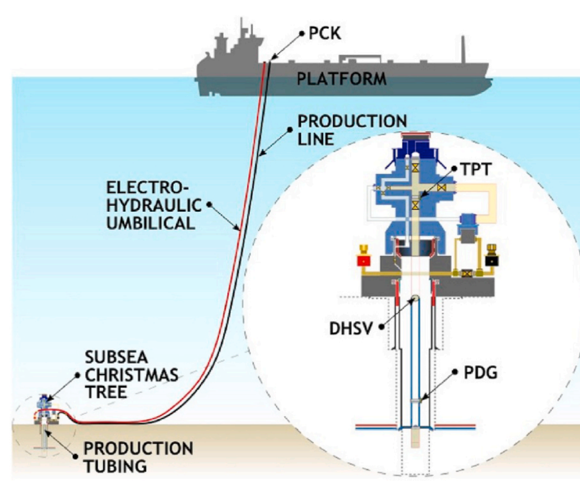


Fig. 1. Simplified schematic of a typical offshore naturally flowing well (Vargas et al., 2019b).

Table 1
3W dataset variables.

Number	Tag	Name	Unit
1	P-PDG	Pressure at the PDG	Pa
2	P-TPT	Pressure at the TPT	Pa
3	T-TPT	Temperature at the TPT	deg°C
4	P-MON-CKP	Pressure upstream of the PCK	Pa
5	T-JUS-CKP	Temperature downstream of the PCK	deg°C
6	P-JUS-CKGL	Pressure downstream of the GLCK	Pa
7	T-JUS-CKGL	Temperature downstream of the GLCK	deg°C
8	QGL	Gas lift flow rate	sm^3/s

- Class 1 – Abrupt increase of basic sediment and water (BSW) – suspended water, sediments and other impurities in the production measured as a percentage of the production stream (*The SLB Energy Glossary | Energy Glossary*, n.d.). The lifecycle of each well contains periods of increasing levels of BSW. However, an unexpected rise indicates a developing production issue, which needs to be remedied quickly.
- Class 2 – Spurious closure of the downhole safety valve (DHSV) – the valve isolates wellbore fluids in a catastrophic failure of surface equipment (*The SLB Energy Glossary | Energy Glossary*, n.d.). If the valve fails spuriously without any surface signs, it needs to be reopened; hence, an automatic event identification is essential.
- Class 3 – Severe slugging – an event in which large gas bubbles follow a sequence of liquid slugs. It is a cyclical phenomenon that can lead to wellhead and pipeline damage. Hence, it is considered a critical abnormality (Vargas et al., 2019b).
- Class 4 – Flow instability – pressure changes within acceptable thresholds, with differences due to slugging, representing the absence of cyclicity. This event can transform into slugging and then a severe variant, which requires imminent actions (Vargas et al., 2019b).
- Class 5 – Rapid productivity loss – flow loss due to changes in reservoir static pressure, alternating BSW percentage, production viscosity, changes in production line diameter, etc. (Vargas et al., 2019b).
- Class 6 – Quick restriction in the production choke (PCK) – a term Petrobras uses to indicate issues with a PCK valve installed at the beginning of the production line. Short restrictions might be observed when operated manually due to operational problems that need to be identified and reversed (Vargas et al., 2019b).
- Class 7 – Scaling in PCK – a mineral deposit, which can create a significant restriction or even a plug in the production tubing (*The SLB Energy Glossary | Energy Glossary*, n.d.). Thus, monitoring the production choke helps recognise the event and take appropriate actions, such as scale inhibitor injections (Vargas et al., 2019b).
- Class 8 – Hydrate in production line – compounds of complex ions formed by water and other substances at reduced temperatures and high pressure, which might lead to plugging of the pipelines (*The SLB Energy Glossary | Energy Glossary*, n.d.). It is one of the most significant issues in oil and gas production, and it can stop flow for an extended period; hence, it needs to be recognised immediately.

Two types of labelling are implemented on two levels: first by instance (a file within each folder, be it actual, simulated or hand-drawn) and second by observation (each row within each file also has a label according to the event).

The real ones were obtained from the actual wells; Schlumberger generated the simulated one through the OLGA system (*OLGA Dynamic Multiphase Flow Simulator*, n.d.); the hand-drawn ones were produced by the 3W database creators using expert knowledge so that the data mimics a typical sensor reading of the particular event type (Fig. 2).

As depicted in Figs. 3 and 4, each observation is labelled as a normal, faulty transient and faulty steady state according to the three periods. The faulty transient state is characterised by the development of undesirable events but still not reaching a failure condition and is labelled by three digits, with the last one corresponding to the event label (for example, 105 as faulty transient and 5 as steady state fault).

3. Related work

All the identified literature was classified according to industry division into 4 main groups: (1) Drilling and exploration, (2) Oil and Gas pipelines transportation system, (3) Production and reservoir management, and (4) 3W dataset. Most of the publications are related to the production sector and the least to the Oil and Gas pipeline and transportation equipment. The range of applied methods is broad, and many advanced techniques are implemented to enhance the result, such as using Genetic Algorithms (GA) for Machine Learning (ML) and Deep Learning (DL) model optimisation, creating stacked Autoencoders algorithms, applying Convolutional Neural Networks (CNN) for improved feature extraction and explaining black box models using Explainable Artificial Intelligence (XAI) techniques.

3.1. Drilling and exploration anomaly detection

There are 15 articles related to drilling operations, which focus on issues such as circulation loss, stuck pipe, washout, bit balling, drill pipe breaks, fluid show, potential kick and other downhole abnormalities. For anomaly detection, several unsupervised methods were applied to

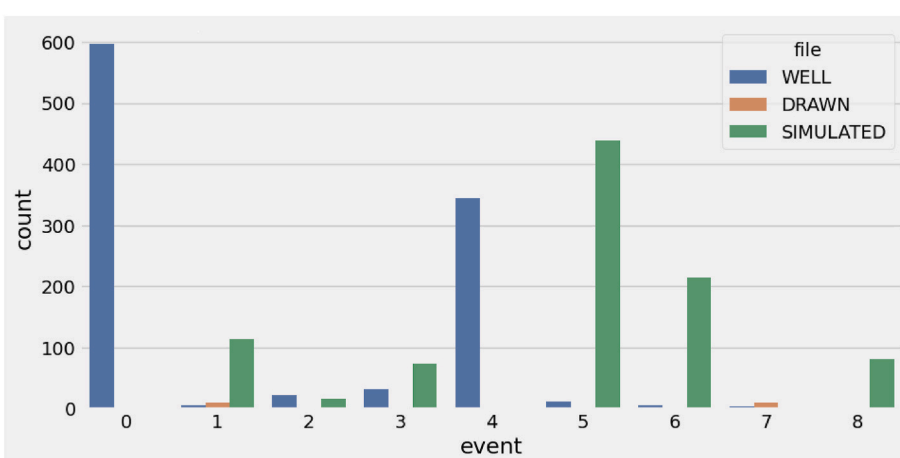


Fig. 2. – The number of instances in the 3W dataset.

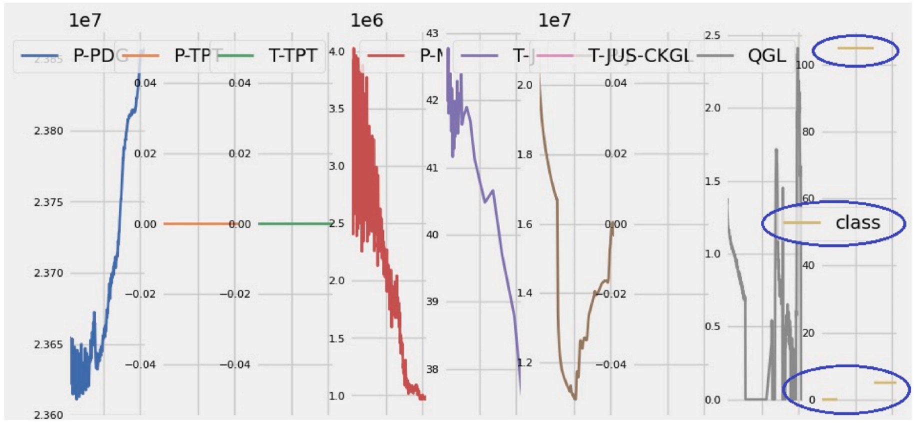


Fig. 3. Class 5 time series of the WELL-00015 instance.

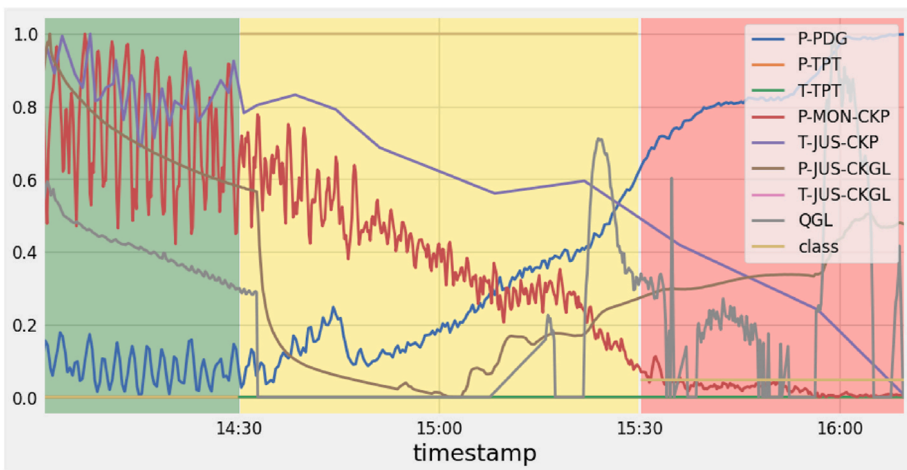


Fig. 4. Class 5 time series of the WELL-00015 instance with observations normal (green), faulty transient (yellow) and faulty steady state (red).

identify unusual data records in multivariate time series from downhole and rig floor sensors, such as Regression, K-Nearest Neighbor (KNN), K-Means, t-distributed Stochastic Neighbor Embedding (t-SNE), dendrograms clustering analysis, Recurrent Neural Networks (RNN), LSTM Autoencoder (LSTM-AE).

The following supervised Machine Learning methods were implemented to classify the abnormalities: Adaptive Neuro-Fuzzy Inference System (ANFIS), Random Forest (RF), Support vector machine (SVM), K-Nearest Neighbor (KNN), Gradient Boosting (GB), Shapley additive explanations (SHAP), Fully Connected network that has a multi-head attention mechanism (FCMH), eXtreme Gradient Boosting (XGBoost), Adaboost (ADA), Decision tree (DT), Multilayer Perceptron (MLP), Naïve Bayes Classifier (NBC) and Quadratic Discriminant Analysis (QDA).

In some publications DL methods were used either as an unsupervised learning tool to build Autoencoders (Mopuri et al., 2022) or for Classification: CNN, Artificial Neural Network (ANN), Functional Network (FN), Bag-of-features, the Feed Forward Back Propagation neural network (FFBPN), RNN with Long Short-Term Memory variant (LSTM-RNN) or Gated Recurrent Unit variant (GRU-RNN) type of architecture.

In a few cases, a Genetic Algorithm was applied to optimise the multilayer Back Propagation Neural Network, creating a GA-BP Neural Network ((Su et al., 2021; Li et al., 2022)).

3.2. Oil and gas pipelines transportation system anomaly detection

There are 8 articles with the research subject related to pipeline transportation systems. The significant problems highlighted are pipeline leakage due to corrosion and harsh environments, damaged insulation, equipment failure that provides pressure for oil and gas transportation, such as pumps and compressors, and the formation of gas hydrates due to low temperature and high pressure. Since most pipelines are unobservable to humans in real time, many remote surveillance algorithms using computer vision have been implemented.

Among the unsupervised learning methods applied for pattern recognition and clustering were the Gaussian mixture model (GMM) and K-Means. The supervised models employed are Random Forest (RF), SVM, KNN, Gradient Boosting (GB), Decision Tree (DT), Multiple Linear Regression, Neural Network and Multilayer perceptron (MLP).

The following Deep Learning methods were used: LSTM and Stacked Auto-Encoder (SAE) (Seo et al., 2021), CNN, Inception ResNet V2 and Visual Geometry Group with 16 layers (VGG16) (Vankov et al., 2020).

3.3. Production and reservoir management anomaly detection

24 articles are related to production and reservoir management, and it is the biggest group from the pool of the identified literature. Apart from naturally flowing wells, in which formation pressure is high enough to provide extraction without additional treatment, most wells require Artificial lift techniques: beam pumps (sucker rod system), electrical submersible pumping (ESP), gas lift systems, hydraulic pumps,

plungers and progressive cavity pumps (PCP) (*The Defining Series: Artificial Lift | SLB*, n.d.). Regardless of how robust and well maintained this equipment is, they are susceptible to many potential failures, and most of the research is focused on anomaly detection in this production branch of the industry.

The following supervised learning methods were applied: KNN, Logistic Regression (Logit), SVM, DT, RF, Rule Fit Classifier (RFC), Extreme Learning Machine (ELM), supervised shapelet-based classification algorithm Fast Shapelets, Naive Bayes (NB), Stochastic Gradient Descent (SGD), Quadratic discriminant analysis (QDA), Linear Discriminant Analysis (LDA), boosting techniques e.g., XGBoost), AdaBoost, and Categorical Boosting (CatBoost).

A few Genetic Algorithms were applied to optimise the SVM model due to the problem that the parameters are difficult to determine when classifying: Chicken Swarm Optimisation (CSO), differential mutation strategy, adaptive inertial strategy (DACSO), Particle Swarm Optimisation (PSO) and Bat Algorithm (BA) (J. Liu et al., 2019). Also, GA optimised Back Propagation neural network (GA-BP) was implemented for offshore submersible motor fault diagnosis (Y. Zhang and Yang, 2022).

Unsupervised machine learning algorithms that were used are Cluster based local outlier factor (CBLOF), Histogram-based Outlier Score (HBOS), Isolation Forest (IF), Median Absolute Deviation (MAD), Minimum Covariance Determinant (MCD), Principal Component Analysis (PCA), Gaussian Markov random fields (GMRF), graphical Gaussian model (GGM), sparse Principal Component Analysis (sPCA), sparse Autoencoder, Alternating Decision Tree (ADTree), SVM, Naïve Bayesian Network, Fuzzy C-means algorithm.

A semi-supervised method of Random peek was employed for the case of Artificial lift systems anomaly detection, where only a small number of samples is labelled, assuming that most of the unlabeled samples should be labelled normal ((Y. Liu et al., 2010) (Y. Liu et al., 2011)).

Out of the DL methods, the next were implemented: Back Propagation Neural Network (BPNN), CNN, Triplet network, i.e., an artificial neural network based on a Triplet loss metric, and other metric learning losses, such as Proxy-Anchor loss, Contrastive loss, Lifted Structured loss, CosFace loss (Mello et al., 2022), two stacked Autoencoders (Scoralick et al., 2021), Multilayer Feedforward Neural Network (MFNN), LSTM, Convolutional-LSTM (CONV-LSTM) (Sinha et al., 2020), CNN with backbones ResNet50, SE-ResNet50, ResNet50II, SE-ResNet50II, AlexNet (Tan et al., 2022), Deep-Broad Learning System (DBLS), Fast Fourier transform (FFT), Wavelet transformation (Wei and Gao, 2020).

3.4. 3W dataset anomaly detection and classification

To the best of our knowledge, 11 officially published articles about anomaly detection and Classification on the offshore naturally flowing wells, using 3W Dataset from Petrobras, were created by combining real, simulated and hand-drawn records written in English and having open access.

Many experiments were attempted to set up multiclass or binary classifications. An array of supervised and unsupervised learning methods was applied with outstanding results. Some researchers attempted multiclass Classification of undesirable events, while others selected one particular abnormality (ex., flow instability) and performed binary Classification against all the rest of the classes (Marins et al., 2021). Performed 3 experiments: (a) One-class classifier to identify normal vs abnormal events, thus combining all faults into one unique class; (b) Multiple binary classifiers with several classifiers discriminating each fault against normal events; and (c) Single multiclass classifier, identifying each fault against all events, as mentioned earlier.

The following supervised learning methods were applied: KNN, One Nearest Neighbor (1NN), Logistic regression (LR), Support Vector Classifier (SVC), LDA & QDA, DT, RF, ADA, GNB, Zero Rule (ZR), Extreme Learning Machine (ELM), MLP.

A few GA were used to optimise the algorithm: (Gatta et al., 2022) created a Convolutional 1D Autoencoder with a genetic approach for hyperparameters selection via Biased Random Key Genetic Algorithm (BRKGA), in which different combinations of hyperparameters are regarded as an individual of a population, and each hyperparameter is regarded as a gene of the individual.

Explainability of the classifiers was also researched, and three XAI techniques were applied to interpret black box models to understand the causes of abnormalities: global surrogate model using DT, Shapley Additive Explanation (SHAP), and Local Interpretable-Agnostic Explanation (LIME) (Aslam et al., 2022).

Unsupervised algorithms that were implemented are t-SNE, PCA, one-class SVM, Cluster-based Algorithm for Anomaly Detection in Time Series Using Mahalanobis Distance (C-AMDATS), Luminol Bitmap, SAX-REPEAT, KNN, Bootstrap, and Robust Random Cut Forest (RRCF).

Finally, the attempted deep learning methods were LSTM-AE and CNN 1D Autoencoder.

The systematic literature review following the PRISMA methodology allowed an insight into the current state of knowledge in the area of anomaly detection in the petroleum industry generally, as well as the 3W dataset. The recent publications, made in 2022, focused more on DL algorithms since they provide higher classification accuracy. Since it would be of great interest to further develop the latest processes that use DL algorithms, the authors of this research attempted other RNN configurations with LSTM and GRU architectures and GA hyperparameter optimisation.

The summary of all the publications with corresponding AI methods is represented in Table 2.

3.5. Visual analysis

The PRISMA bibliometric methodology selected the final 58 papers, from which 44 were journal articles and 14 were conference proceedings. The co-occurrence of keywords was performed using VOSviewer, a text mining software for creating maps based on network data. The analysis used the complete counting method with a minimum number of two keyword occurrences.

The top five keywords that were encountered most often are Machine learning (9 occurrences, 15 total link strength), Fault diagnosis (6 occurrences, 10 total link strength), Classification (3 occurrences, 9 total link strength), Oil well monitoring (3 occurrences, 8 total link strength) and Electrical submersible pump (2 occurrences, 7 total link strength).

As shown in Fig. 5, the keywords co-occurrence analysis revealed 5 clusters with 25 keywords, 52 links and 66 total line strength. The clusters are characterised by colours with the following major nodes:

- Machine learning – Red
- Fault diagnosis – Yellow
- Anomaly detection – Blue
- Unsupervised machine learning – Purple
- Convolutional neural network – Green.

The keyword co-occurrence network shows that clusters exhibit distinct separation with limited interconnections. Specifically, the blue cluster (major node Anomaly detection), the purple cluster (Unsupervised machine learning) and the green cluster (Convolutional neural network) link just to the yellow (Fault diagnosis) and red (Machine learning) clusters. None of them have any links to each other. The two biggest clusters, yellow (Fault diagnostics) and red (Machine learning), have multiple links in-between and with other clusters.

The keywords co-occurrence network by year overlay visualisation shows that Machine learning methods gained the most popularity from 2021, and there were many methodologies tried and implemented for anomaly detection, such as unsupervised machine learning methods, Random Forest, Support Vector Machine, the most recent being Autoencoder and Neural Network (Fig. 6).

Table 2
PRISMA method selected publications.

No	Publication	Research Question	Methods
Drilling and Exploration Anomaly Detection			
1	Application of adaptive neuro-fuzzy inference system and data mining approach to predict lost circulation using DOE technique (case study: Maroon oilfield) (Agin et al., 2020)	Prediction of lost circulation problem during drilling	Data mining (regression) and Adaptive Neuro-Fuzzy Inference System (ANFIS).
2	Deep Learning and Time-Series Analysis for the Early Detection of Lost Circulation Incidents during Drilling Operations (Aljubran et al., 2021)	Detection of lost circulation during drilling	Random Forest is used as a baseline, and deep learning methods include CNN, ANN, and LSTM.
3	Application of Machine Learning Methods in Modeling the Loss of Circulation Rate while Drilling Operation (Alsaihati et al., 2022)	Predicting the loss of circulation rate (LCR) while drilling	Support vector machine (SVM), Random Forest (RF), and K-Nearest Neighbor (KNN).
4	Use of Machine Learning and Data Analytics to Detect Downhole Abnormalities while Drilling Horizontal Wells, with Real Case Study (Alsaihati et al., 2021)	Continuous profile of the surface drilling torque (T&D) prediction to enable the detection of operational problems ahead of time.	Random forest (RF), Artificial Neural Network (ANN), and Functional Network (FN).
5	Forecasting the abnormal events at well drilling with machine learning (Gurina et al., 2022a)	Prediction of six types of drilling accident probabilities in real-time, using the data from the drilling telemetry representing the time-series.	Bag-of-features, K-Means, Gradient Boosting (GB), Convolution Neural Network (CNN)
6	Making the black-box brighter: Interpreting machine learning algorithm for forecasting drilling accidents (Gurina et al., 2022b)	Interpretability and development of the explanatory model of Bag-of-features approach, used for drilling accidents prediction.	Bag-of-features, Shapley additive explanations (SHAP), Fully connected network that has a multi-head attention mechanism (FCMH), T-SNE
7	Application of machine learning to accident detection at directional drilling (Gurina et al., 2020)	Development of data-driven Algorithm for anomaly alarming for directional drilling.	Gradient Boosting (GB), dendrograms clustering analysis
8	AI-Driven maintenance support for downhole tools and electronics operated in dynamic drilling environments (Kirschbaum et al., 2020)	Artificial Intelligence (AI)-driven Condition Maintenance (CBM), combining Bottom Hole Assembly (BHA) data with Big Data Analytics (BDA) for downhole electronics failure detection	Random Forest (RF), eXtreme Gradient Boosting (XGBoost)
9	Drilling performance monitoring and optimisation: a data-driven approach (Lashari et al., 2019)	Prediction of ROP, drilling performance monitoring and optimisation, identifying the bit malfunction or failure, like bit balling.	The feed forward back propagation neural network (FFBPN)

Table 2 (*continued*)

No	Publication	Research Question	Methods
10	A New Method for Intelligent Prediction of Drilling Overflow and Leakage Based on Multi-Parameter Fusion (Li et al., 2022)	Mud overflow and leakage prediction during drilling	Genetic algorithm to optimise the multilayer Back Propagation Neural Network (GA-BP Neural Network)
11	Well Control Space Out: A Deep-Learning Approach for the Optimisation of Drilling Safety Operations (Magana-Mora et al., 2021)	Surveillance method for drilling operations control using cameras and computer vision in real time. The model for tool joint detection is used to compute the location of the tool joint below the drill floor. In the case of an uncontrolled flow, the Well Control Space Out determines the appropriate measures to take.	Deep Learning methods: Regional Convolutional Neural Network (Faster-R-CNN), Single Shot Detector (SSD), You Only Look Once (YOLOv3), ResNet, DarkNet and Inception backbones.
12	Early sign detection for the stuck pipe scenarios using unsupervised deep learning (Mopuri et al., 2022)	Detecting early signs for the stuck events in drilling	Unsupervised learning: Recurrent Neural Networks (RNNs), LSTM Autoencoder (LSTM-AE)
13	Supervised data-driven approach to early kick detection during drilling operation (Muojeke et al., 2020)	Early kick detection during drilling for implementing the appropriate well control strategy to manage kick situations	Supervised models: Artificial Neural Network (ANN), Recurrent Neural Networks (RNN), Long Short-Term Memory variant of RNN (LSTM-RNN), Gated Recurrent Unit variant of RNN, (GRU-RNN)
14	Prediction of drilling leakage locations based on optimised neural networks and the standard random forest method (Su et al., 2021)	Creating a real time model for predicting leakage layer locations in drilled formations that cause potential circulation loss	Genetic Algorithm-Back Propagation (GA-BP) neural network, Random Forest (RF)
15	Effective prediction of lost circulation from multiple drilling variables: a class imbalance problem for machine and deep learning algorithms (Wood et al., 2022)	Prediction of lost circulation during drilling	8 Machine Learning methods: Adaboost (ADA), Decision tree (DT), K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP), Naïve Bayes Classifier (NBC), Quadratic Discriminant Analysis (QDA), Random Forest (RF) and Support Vector Classifier (SVR). 3 Deep Learning methods: Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM).
Production anomaly detection and Classification			
16	Explainable and Interpretable Anomaly Detection Models for Production Data (Alharbi et al., 2022)	Study of white-box and black-box classifiers for supervised anomaly detection on oil and gas production data.	K-Nearest Neighbor (KNN), Logistic Regression (Logit), Support Vector Machines (SVMs), Decision Tree (DT), Random Forest (RF), and Rule Fit Classifier (RFC). Further models analysis for explainability using LIME (local interpretable model-agnostic

(*continued on next page*)

Table 2 (*continued*)

No	Publication	Research Question	Methods
			explanations) and interpretability.
17	Self-Diagnosis of Multiphase Flow Meters through Machine Learning-Based Anomaly Detection (Barbariol et al., 2020)	Method AD4MPFM (Anomaly Detection for Multiphase Flow Meters), enabling the metrology system to detect outliers and provide a statistical level of confidence in the measures for oil production.	Unsupervised machine learning algorithms: Cluster based local outlier factor (CBLOF), Histogram-based Outlier Score (HBOS), Isolation Forest (IF), Median Absolute Deviation (MAD), Minimum Covariance Determinant (MCD), and Principal Component Analysis (PCA)
18	Sparse Gaussian Markov Random Field Mixtures for Anomaly Detection (Ide et al., 2017)	Anomaly detection of a compressor of offshore oil production from multivariate noisy sensor data.	Gaussian Markov random fields (GMRF), graphical Gaussian model (GGM), sparse Principal Component Analysis (sPCA), sparse Autoencoder
19	Fault Diagnosis of Rod Pumping Wells Based on Support Vector Machine Optimised by Improved Chicken Swarm Optimisation (J. Liu et al., 2019)	Diagnosis of the faults of pumping wells by classifying and identifying the indicator diagrams	Support vector machine (SVM), chicken swarm optimisation (CSO), differential mutation strategy, and adaptive inertial strategy (DACSO), particle swarm optimisation (PSO) and bat algorithm (BA)
20	Failure Prediction for Rod Pump Artificial Lift Systems (Y. Liu et al., 2010)	Prediction of Failure for Rod Pump Artificial Lift Systems	Unsupervised methods: Alternating Decision Tree (ADTree), Support Vector Machine (SVM), Naïve Bayesian Network. Semi-supervised: Random Peek.
21	Semi-supervised failure prediction for oil production wells (Y. Liu et al., 2011)	Development of Smart Engineering Apprentice (SEA) framework for Artificial Lift Systems failure prediction	Semi-supervised Classification using Random Peek, Support Vector Machines (SVM)
22	Adaptive fault diagnosis of sucker rod pump systems based on optimal perceptron and simulation data (X.-X. Lv et al., 2022)	The improved model of fault diagnosis for the sucker rod production system (SRPS)	Back Propagation Neural Network (BPNN), Extreme Learning Machine (ELM), and Support Vector Machine (SVM) with improved feature extraction
23	An evolutional SVM method based on incremental algorithm and simulated indicator diagrams for fault diagnosis in sucker rod pumping systems (X. Lv et al., 2021)	Fault diagnosis of the sucker rod pumping system (SRPS)	Evolutional SVM method based on incremental Algorithm and simulated IDs, ELM, PSO-ELM, BPNN and SVM as baselines
24	Anomaly Detection Based on Sensor Data in Petroleum Industry Applications (Martí et al., 2015)	Anomaly Detection in Offshore Oil Extraction Turbomachines	One-class support vector machine (SVM), Yet Another Segmentation Algorithm (YASA)
25	On the combination of support vector machines and segmentation algorithms for anomaly detection: A petroleum industry comparative study (Martí et al., 2017)	Anomaly detection of turbomachinery installed in offshore petroleum extraction platforms.	One-class Support Vector Machines (SVM), Kalman filters, Yet Another Segmentation Algorithm (YASA)
26	Metric Learning for Electrical Submersible Pump Fault Diagnosis (Mello et al., 2020)	Electrical Submersible Pump (ESP) fault diagnosis	Convolutional neural network (CNN) trained with a triplet loss learning for extracting relevant features, standard machine learning algorithms such as K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, Quadratic Discriminant Analysis and Naïve Bayes Classifier.
27	Ensemble metric learners for improving electrical submersible pump fault diagnosis (Mello et al., 2022)	Electrical Submersible Pump (ESP) fault diagnosis	Ensembles composed of deep neural networks (convolutional network (ConvNet) with 5 metrics: Triplet network, i.e., an artificial neural network based on a metric called Triplet loss, Proxy-Anchor loss, Contrastive loss, Lifted Structured loss, CosFace loss. Random Forest (RF), majority voting, Principal Component Analysis (PCA)
28	Unsupervised Methods to Classify Real Data from Offshore Wells (Orestes et al., 2021)	Anomalies detection during oil and gas production	Fuzzy C-means algorithm for Classification into clusters, Control Chart method, Random Forest (RF)
29	Predicting Compressor Valve Failures from Multi-Sensor Data (Patri et al., 2015)	Ranking sensor dimensions and finding signatures in compressor sensor data, which may aid in the prediction of valve failure	Decision Tree supervised shapelet-based classification algorithm Fast Shapelets.
30	Electric submersible pump broken shaft fault diagnosis based on principal component analysis (Peng et al., 2020)	Identify the cause and the time of ESP shaft fracture, predict the impending breakage time and determine the variable most responsible.	Principal Component Analysis (PCA)
31	Machine Learning Models to Predict Gas Hydrate Plugging Risks Using Flowloop and Field Data (Qin et al., 2019)	Evaluating gas hydrate risk based on measurable process parameters	Support vector classifier (SVC) with several kernels, such as linear, polynomial, radial basis functional (RBF), and artificial neural networks (ANN). Feature selection methods SelectKBest and ExtraTreesClassifier.
32	A novel machine learning model for autonomous analysis and diagnosis of well integrity failures in artificial-lift production systems (Salem et al., 2022)	Automated prediction of integrity failures in wells with Artificial Lift gas lift production method	Logistic regression, Naive Bayes (NB), Decision trees (DT), Random Forests (RF), KNN, SVM, Stochastic gradient descent (SGD), Quadratic discriminant analysis (QDA), boosting techniques e.g., Extreme Gradient Boosting (XGB), Adaptive Boosting (AdaBoost) and Categorical Boosting (CatBoost).
33	Fault detection with Stacked Autoencoders and pattern recognition techniques in gas lift operated oil wells (Scoralick et al., 2021)	Detection and Classification of failures in oil production wells operated with elevation by gas lift.	Two stacked autoencoders with 9 and 5 neurons, Decision Tree (DT), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), KNN

(*continued on next page*)

Table 2 (*continued*)

No	Publication	Research Question	Methods
34	Normal or abnormal? Machine learning for the leakage detection in carbon sequestration projects using pressure field data (Sinha et al., 2020)	Automation of the leakage detection process in carbon storage reservoirs using rates of (CO2) injection and pressure data measured by simple harmonic pulse testing (HPT).	Multilayer Feedforward Neural Networks (MFNN), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Convolutional-LSTM (CONV-LSTM)
35	A visual analytics approach to anomaly detection in hydrocarbon reservoir time series data (Soriano-Vargas et al., 2021b)	Anomaly detection in time series data of hydrocarbon reservoir using visual analytics approach based on interactive visualisations of time series connected with machine learning approaches.	Isolation Forest (IF)
36	Multiscale Normalization Method Combined with a Deep CNN Diagnosis Model of Dynamometer Card in SRP Well (Tan et al., 2022)	Development of a diagnosis model to identify the working condition of each sucker rod pumping (SRP) well	Four CNN backbones: ResNet50, SE-ResNet50, ResNet50II, SE-ResNet50II, SVM with radial basis function and particle swarm optimisation (PSO), AlexNet model for comparison
37	Fault Diagnosis of Sucker Rod Pump Based on Deep-Broad Learning Using Motor Data (Wei and Gao, 2020)	Fault diagnosis methods of sucker rod pump (SRP)	Convolutional Neural Network (CNN), Deep-Broad Learning System (DBLS), Fast Fourier transform (FFT), Wavelet transformation, Extreme Learning Machine (ELM), Support Vector Machine (SVM), Hidden Markov Model (HMM)
38	Fault Diagnosis of Submersible Motor on Offshore Platform Based on Multi-Signal Fusion (Y. Zhang and Yang, 2022)	Offshore submersible motor fault diagnosis	Back Propagation Neural Network (BP), Genetic Algorithm optimised Back Propagation neural network (GA-BP)
39	An intelligent diagnosis method of the working conditions in sucker-rod pump wells based on convolutional neural networks and transfer learning (R. Zhang et al., 2021)	Diagnosis of sucker-rod pump working conditions.	Transfer deep learning methods: AlexNet Network, GoogLeNet Network, shallow convolutional neural networks (CNN3 model and CNN2 model) and Fully Connected Neural Network model (FC model)
Oil Pipelines and Transportation System Anomaly Detection			
40	An Anomaly Detection Model for Oil and Gas Pipelines Using Machine Learning (Aljameel et al., 2022)	Oil pipeline leakage detection.	Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Gradient Boosting (GB), Decision Tree (DT).
41	A data-driven pipeline pressure procedure for remote monitoring of centrifugal pumps (Giro et al., 2021)	Automated strategy to remotely monitor the status of centrifugal pumps in pipeline transportation systems when the network of sensors is unavailable or not present.	Unsupervised clustering techniques: Gaussian mixture model (GMM)
42	Deep Learning Approach for Objects Detection in Underwater Pipeline Images (Gasparovic et al., 2022)	Underwater seafloor pipeline leakage detection, using images to verify their integrity and determine the need for maintenance	Convolutional Neural Network (CNN), Six different architectures: You Only Look Once (YOLO) architectures (YOLOv4, YOLOv4-Tiny, CSP-YOLOv4, YOLOv4@Resnet, YOLOv4@DenseNet), and one on the Faster Region-based CNN (RCNN) architecture.
43	DARTS-Drone and Artificial Intelligence Reconsolidated Technological Solution for Increasing the Oil and Gas Pipeline Resilience (Ravishankar et al., 2022)	Integrating drone technology and deep learning techniques to detect the targeted potential root problems that can cause critical pipeline failures and predict the progress of the detected problems by collecting and analysing image data periodically	Computer vision algorithm using deep learning neural network DeepLabV3+, data augmentation
44	Development of an AI-based diagnostic model for predicting hydrate in gas pipelines (Seo et al., 2021)	Diagnose hydrate for flow assurance purposes in gas pipelines.	Multilayer perceptron (MLP), Long Short-Term Memory LSTM, and Stacked Auto-Encoder (SAE)
45	Microwave Nondestructive Testing for Defect Detection in Composites Based on K-Means Clustering Algorithm (Shrifan et al., 2021)	Nondestructive testing (NDT) to detect the underneath defect in composites, used for insulation of steel pipelines in oil and gas industry, based on microwave reflection coefficients.	Unsupervised machine learning: K-Means clustering
46	Assessment of the condition of pipelines using convolutional neural networks (Vankov et al., 2020)	Analysis of amplitude-frequency measurements in pipelines to identify the presence of a defect and further clarify its variety.	Convolutional Neural Network (CNN), Inception ResNet V2, Visual Geometry Group with 16 layers (VGG16)
47	A minimalist approach for detecting sensor abnormality in oil and gas platforms (Wong et al., 2022)	Detecting abnormality of compressor's shaft's RPM sensor	Multiple Linear Regression, Neural Network
3W Dataset Anomaly Detection and Classification			
48	Proposal for two classifiers of offshore naturally flowing wells events using k-nearest neighbors, sliding windows and time multiscale (Vargas et al., 2017)	Identification of four anomalous events in oil wells for 3W Dataset: Spurious Closure of DHSV, Rapid Productivity Loss, Hydrates in Production Lines, Choke Valve Closure	KNN (k-Nearest Neighbors), t-distributed Stochastic Neighbor Embedding) (t-SNE)
49	Classification of undesirable events in oil well operation (Turan and Jaschke, 2021)	Multiclass classification of anomalous events in oil wells for 3W Dataset	Decision Tree, as baseline attempted Logistic Regression (LR), Support Vector Classifier (SVC), Linear and Quadratic Discriminant Analysis (LDA & QDA), Random Forest, AdaBoost (ADA), Principal Component Analysis (PCA)
50	Statistical analysis of offshore production sensors for failure detection	Identification of abnormal events in oil wells for 3W Dataset	Principal Component Analysis (PCA) and Logistic Regression (LR)

(*continued on next page*)

Table 2 (*continued*)

No	Publication	Research Question	Methods
	applications (Santos et al., 2021)		
51	Fault detection and Classification in oil wells and production/ service lines using random forest (Marins et al., 2021)	Development of CBM system for identification of anomalous events in oil wells for 3W Dataset	Random Forest, Principal component Analysis (PCA), Bayesian non-convex optimisation strategy
52	Improving the performance of one-class classifiers applied to anomaly detection in oil wells (Machado et al., 2022)	Identification of two types of faults in oil wells for 3W Dataset: Spurious closing of Downhole Safety Valves (DHSV) and Hydrate in Production Line.	Two unsupervised learning methods, a Long Short-Term Memory (LSTM) autoencoder and a one-class Support Vector Machine (OCSVM), are trained on faulty events as a target class.
53	Predictive maintenance for offshore oil wells by means of deep learning features extraction (Gatta et al., 2022)	Multiclass Classification of anomalous events in oil wells for 3W Dataset	Deep learning method for feature extraction: 1D AutoEncoder using Convolutional Neural Network. Machine learning classifiers: Random Forest, Nearest Neighbors, Gaussian Naive Bayes and Quadratic Discriminant Analysis, hyperparameters selection via Biased Random Key Genetic Algorithm (BRKGA).
54	Data-driven Detection and Identification of Undesirable Events in Subsea Oil Wells (Brønstad et al., 2021)	Development of a CBM system for identification of anomalous events in oil wells for 3W Dataset	Random Forest (RF), Principal Component Analysis (PCA)
55	Flow Instability Detection in Offshore Oil Wells with Multivariate Time Series Machine Learning Classifiers (Carvalho et al., 2021a,b)	3W Dataset Flow Instability prediction	Binary machine learning classifiers: One Nearest Neighbor (1NN), Gaussian Naïve Bayes (GNB), Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), Random Forest (RF). As a baseline, use the Zero Rule (ZR) classifier.
56	Hyperparameter Tuning and Feature Selection for Improving Flow Instability Detection in Offshore Oil Wells (Carvalho et al., 2021a,b).	Improvement of previous 3W Dataset Flow Instability prediction	Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Adaptive Boosting (ADA), Extreme Learning Machine (ELM) and Multilayer Perceptron (MLP), Zero-rule (ZR) classifier, Sequential Feature Selection SFS-F (forward), SFS-B (backward) and Genetic Algorithm for feature selection.
57	Detecting Interesting and Anomalous Patterns in Multivariate Time-Series Data in an Offshore Platform Using Unsupervised Learning (Figueirêdo et al., 2021)	A comparative evaluation performance of unsupervised learning algorithms for pattern recognition in 3W Dataset undesirable events, such as Spurious closure of DHSV and Quick restriction in PCK	Six unsupervised machine learning algorithms: Cluster-based Algorithm for Anomaly Detection in Time Series Using Mahalanobis Distance (C-AMDATS), Luminol Bitmap, SAX-REPEAT, KNN, Bootstrap, and Robust Random Cut Forest (RRCF).
58	Anomaly Detection Using Explainable Random Forest for the Prediction of Undesirable Events in Oil Wells (Aslam et al., 2022)	Identification of anomalous events in oil wells for 3W Dataset and model interpretation	Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and K-Nearest Neighbor (K-NN), SMOTE, Explainable Artificial Intelligence (XAI). Three XAI techniques: global surrogate model using DT, Shapley Additive Explanation (SHAP), and Local Interpretable-Agnostic Explanation (LIME).

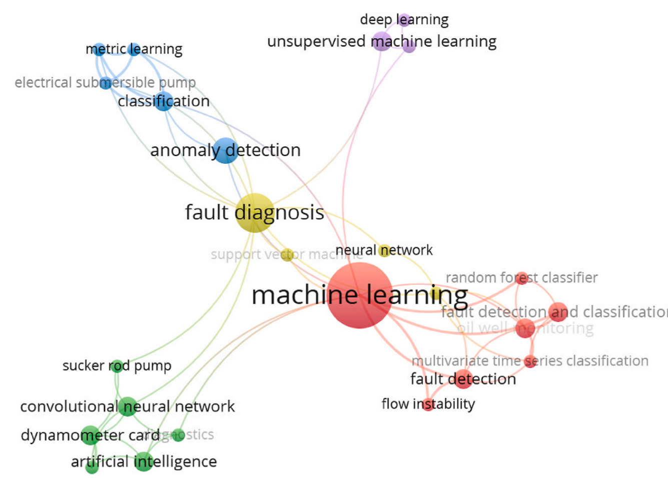


Fig. 5. Keywords co-occurrence network.

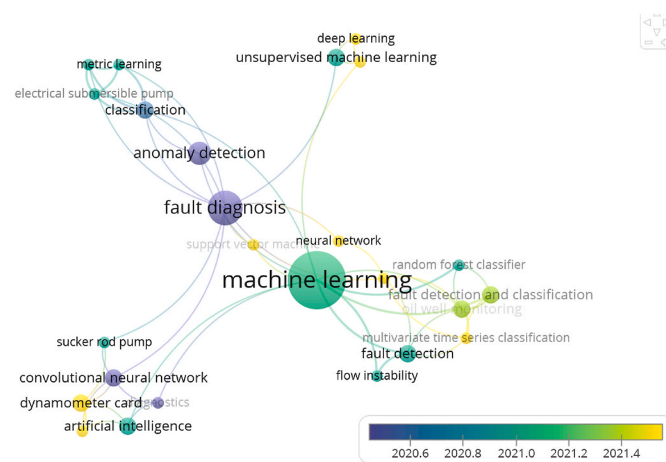


Fig. 6. Keywords co-occurrence by year.

The author's co-authorship was performed in VOSviewer using 25 maximum authors per document and a minimum number of documents of an author of 2. Out of 282 authors, only 20 meet the threshold.

In the author's co-authorship network, 6 clusters were identified with 20 items, 32 links and a total link strength of 65, as depicted in Fig. 7.

The red and yellow clusters are connected through Varejão, Flávio Miguel, who collaborated the most with other authors.

4. Methodology

AI and the Internet of Things application in the Oil and Gas industry have spurred increased interest in the research of recent ML, Data

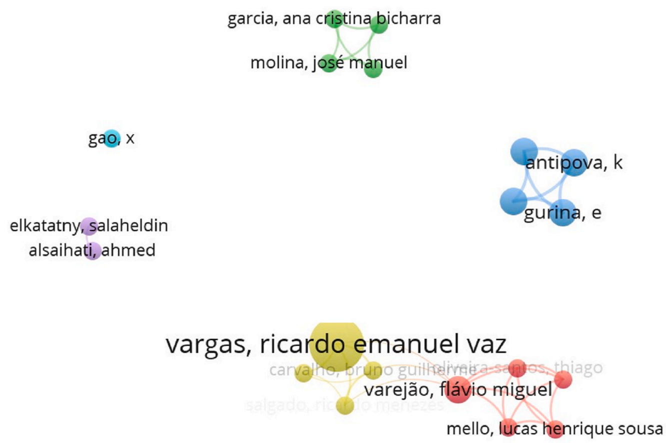


Fig. 7. Authors co-authorship network visualisation.

Mining (DM) and DL models to resolve its everyday demands. A growing number of publications in the last few years have created a necessity to understand better the most popular algorithms and the latest developments in the subject.

Anomaly detection is a statistical technique used to identify abnormal patterns in data that deviate from a priori expected behaviour (Martí et al., 2015). It is applied in many industries: manufacturing, aviation, transportation, banking, health, etc. The oil and gas industry has recently joined the trend and started to take the opportunity to identify anomalies in time and improve general performance, reducing potential downtime, minimising costs, and, on some occasions, saving lives. The framework of this analysis focuses on the Upstream and Midstream sectors of the petroleum industry, which involves Drilling and Exploration, Production and Transportation, where most of the failures occur.

The Systematic Literature Review (SLR) was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology, a well-known method for establishing a state of knowledge regarding specific topics. The quantitative and qualitative analysis was initialised by collecting 2814 articles published in the 2019–2022 frame for further analysis. Of the final selected publications, 44 were journal articles, and 14 were conference proceedings analysed.

4.1. Developed algorithms

RNN is a specialised neural network designed for processing sequential or time-series data. It features a feedback connection (Fig. 8) that feeds the output back into the network as input, along with new data at each time step. This feedback mechanism enables the network to remember past data when generating the following output (Das et al., 2023).

RNNs are called recurrent because they have a "memory" feature that allows the network to perform the same operation for each sequence element. One of the major disadvantages is a vanishing or exploding gradient, which can become very small or too large due to exponential backpropagation, and the training becomes insufficient. This issue is addressed by altering the gates within the network, leading to the creation of LSTM and GRU backbones.

LSTM is a type of RNN architecture that addresses the vanishing or exploding gradient by controlling the flow of information using additional gating units within each cell (Das et al., 2023). Three gates, namely the input gate, forget gate and output gate, create a "self-loop" that decides which information will be passed further and which can be overlooked (Figs. 9 and 10). The forget gate updates the internal state; the output gate determines if the value should be moved further or shut off (Das et al., 2023).

GRU is another variation of RNN, similar to LSTM, having an additional gating unit but with a reset gate and update gate that controls which information will be forgotten and updated simultaneously (Goodfellow et al., n.d.) (see Fig. 11). The reset gate controls the state of information being passed to the next stage, and the update gate determines how many long-term dependencies would be active to forward previous information (Das et al., 2023). It is faster and hence more efficient, as it doesn't possess internal memory, as LSTM and fewer parameters are used for backpropagation.

4.2. Research process

The research methodology is composed of 3 main phases, which are exploration, analytical and conductive phases. As the objective of the literature review process was to obtain insight into the state of the knowledge in the area of anomaly detection, identifying potentially applicable but not yet exploited algorithms, the result will be implemented for designing the methodology of the project.

The exploration phase starts with a literature review of the recent developments in AI in the oil and gas industry, particularly the 3W

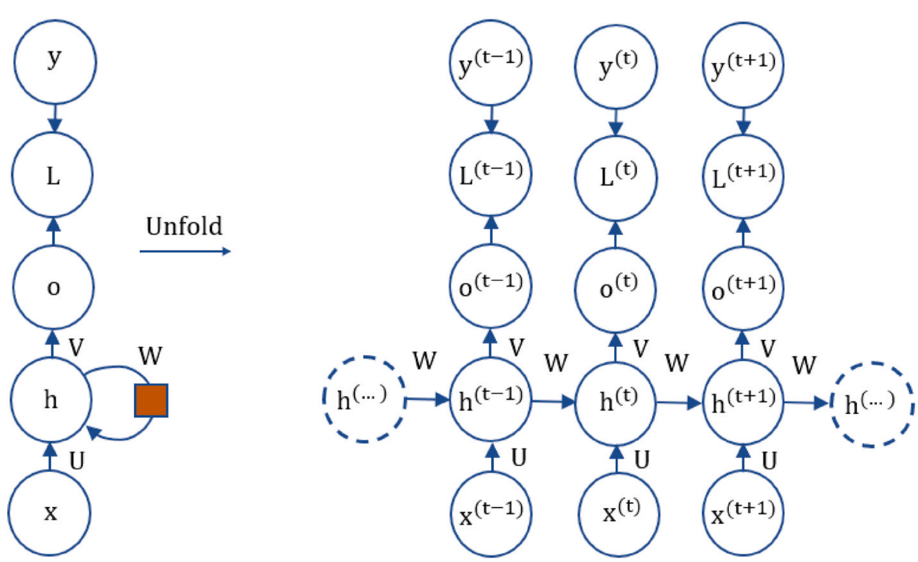


Fig. 8. – RNN schema (adapted from (Goodfellow et al., n.d.)).

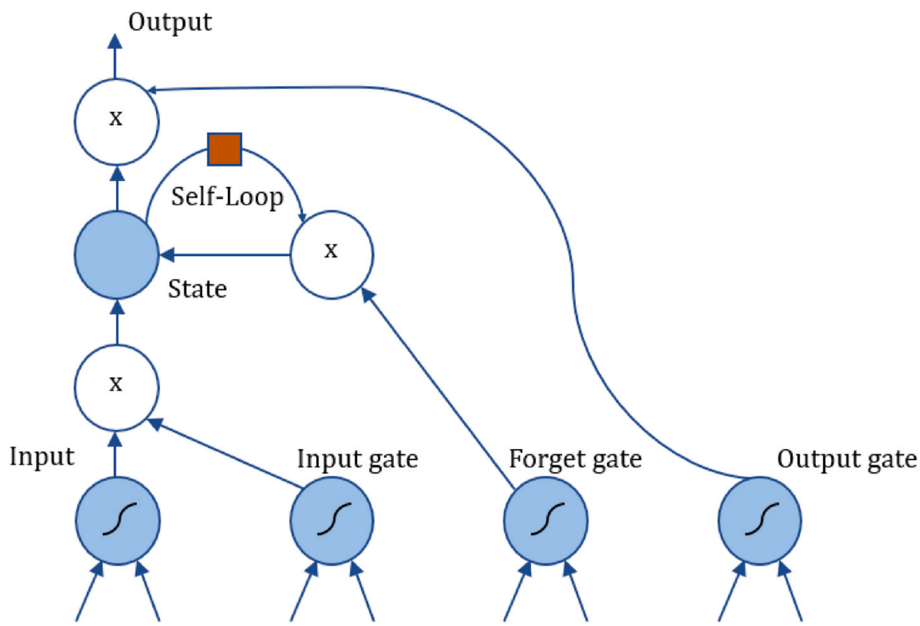


Fig. 9. LSTM schema (adapted from (Goodfellow et al., n.d.)).

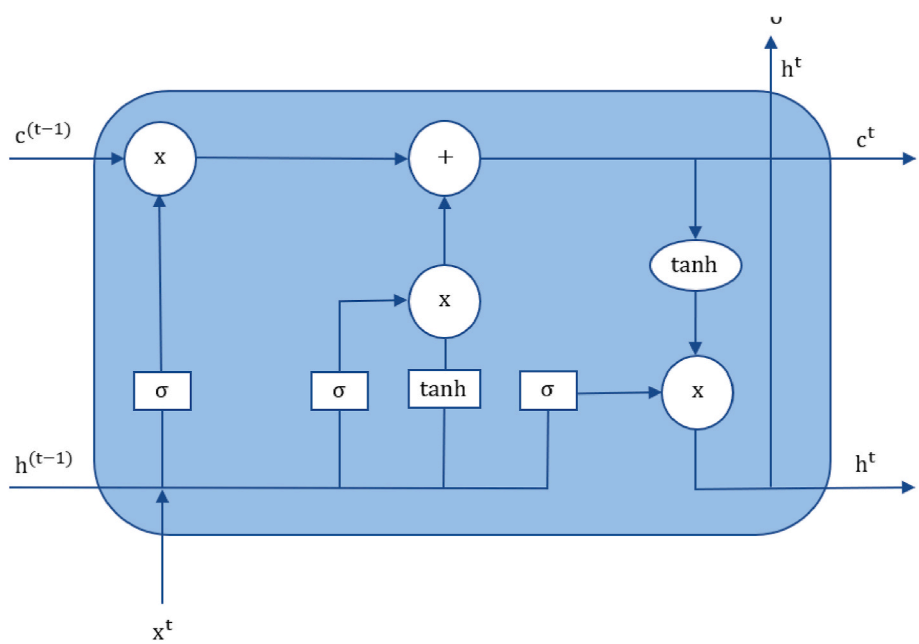


Fig. 10. LSTM schema (adapted from (Das et al., 2023)).

dataset (Fig. 12). It was recognised that DL methods were implemented only in 2 out of 11 official publications (taking into account those only written in English).

The suggested methodology of the research is based on further exploration of DL techniques for multiclass Classification, creating RNN configurations with LSTM and GRU architectures.

The analytical phase would include the following intermediate objectives:

- Preprocessing the 3W dataset by data cleaning, imputing or removing missing values, and standardising the data for better performance of the deep neural networks
- Data transformation by converting into the 3D matrix expected by LSTM and GRU backbones: [samples, timesteps, features]
- Based on RNN, developing the Algorithm with LSTM and GRU architectures to perform the abnormal events multiclass Classification
- Evaluating the results by comparing them with benchmarks from the previous research.

The suggested workflow is presented in Fig. 13, starting with the overview of the 3W Dataset. Its descriptive introduction and analysis are essential and detailed in the Data processing part of the research since it is a challenging dataset and requires a thorough grasp of the anomaly detection task in this project.

The research design runs in parallel, including an already performed literature review, an overview of the suggested algorithms, and an ongoing process of improving the model, with the consideration of potentially adding unsupervised algorithms for dimensionality

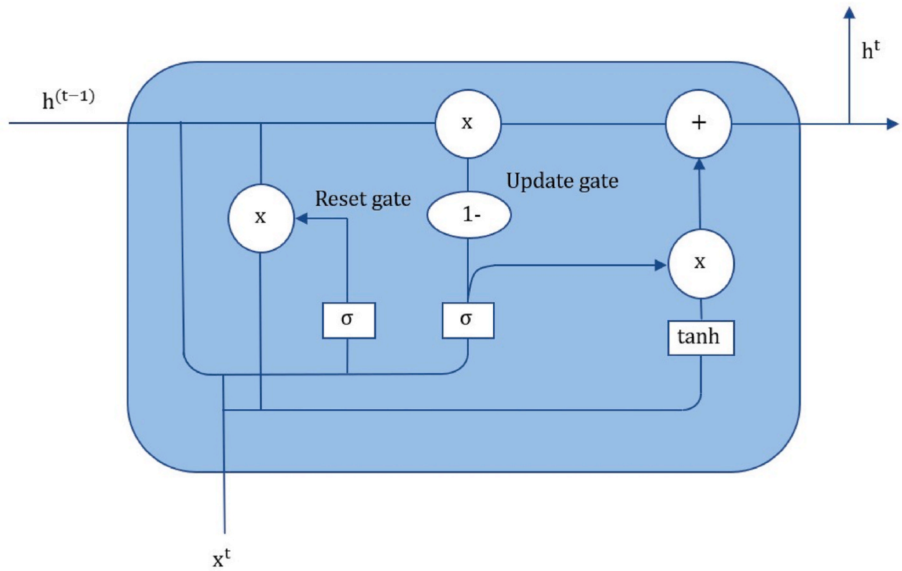


Fig. 11. GRU schema (adapted from (Das et al., 2023)).

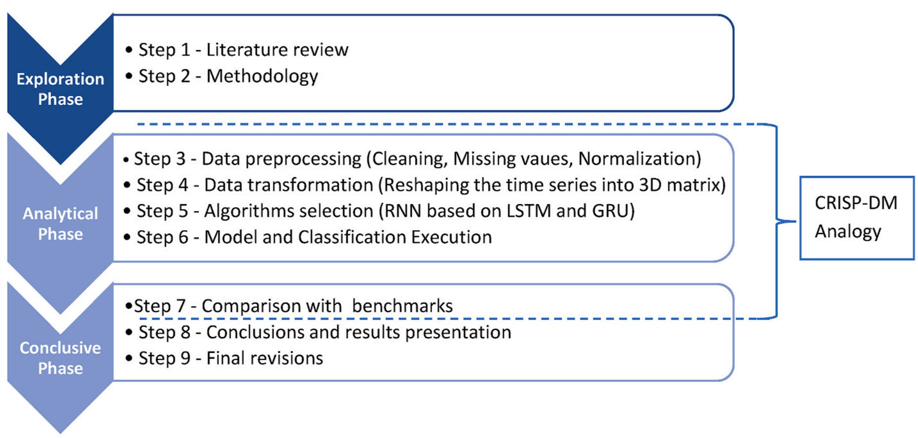


Fig. 12. Phases of the research.

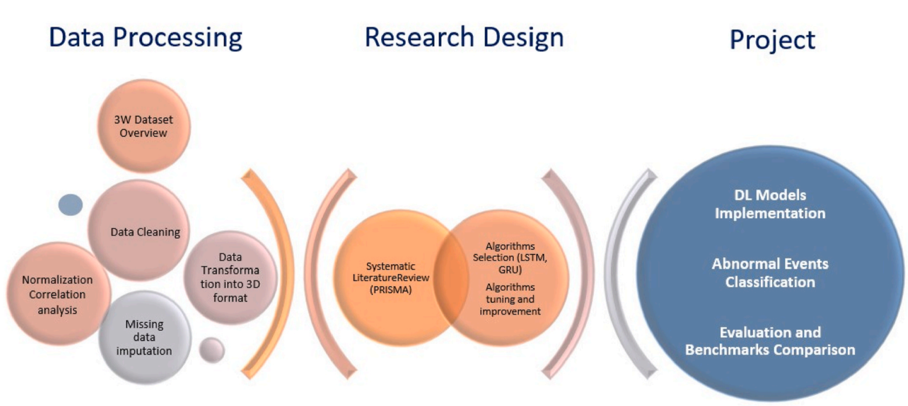


Fig. 13. Methodology overview.

reduction or genetic algorithms for hyperparameter tuning.

The core of the research is the project itself, where the pipeline of the algorithms will be set up, and Classification will be performed to detect undesirable events. Finally, a comparison with the results of the previous publication will be implemented, with the main focus on the papers which performed multiclass Classification.

4.3. Data preprocessing

Since the objective of this project is anomaly detection, only actual instances were considered for the analysis. All the simulated and drawn cases were ignored, leaving only files that start with "Well". The removal of synthetic data turns into a very imbalanced dataset, with classes 0 and 4 being the majority classes and the rest a minority (Fig. 14).

For this project, it is decided to treat faulty transients as faulty events since they naturally progress into failure, and this way, they will be recognised sooner. First, all the csv files were combined into one file, converting all faulty transient classes into corresponding steady state faulty. Considering them as faulty events, the Classification becomes multiclass with 9 classes identified. Also, while concatenating the files, they were down sampled to 1 min to decrease the calculation time. As demonstrated in Fig. 15, the final combined file shows the presence of multiple spikes and noise as well as frozen and missing data.

The "forward fill" method fills the missing values, in which the last valid observation is propagated forward. The final processed dataset is saved and divided into training and testing sets, stratifying by y to ensure that relative class frequencies are approximately preserved in each train and test split. The relationships between variables and classes of the final processed data are shown in Fig. 16.

Since RNN requires a 3D format [samples, timesteps, features], the data was converted into a 3D matrix with a window size equal to 30 as an initial experiment.

5. Results

The first attempt of an LSTM algorithm was run with the original processed data (without the train set being transformed by SMOTE) to evaluate the initial Classification. The network had just 2 stacked LSTM layers and one Dense layer, with the number of hidden units equal to 10, an activation function "tanh", a batch size equal to 30 and 10 epochs, resulting in a macro average F1 score of 0.75.

To understand the general response of different backbones and their parameters, the algorithms were run with the same batch size, timestep and number of epochs but with activation functions "relu", "softmax", "LeakyReLU" or "swish", with 10 or 20 number of hidden units, and a different number of LSTM or GRU layers, applied both before and after SMOTE oversampled train data (Table 3).

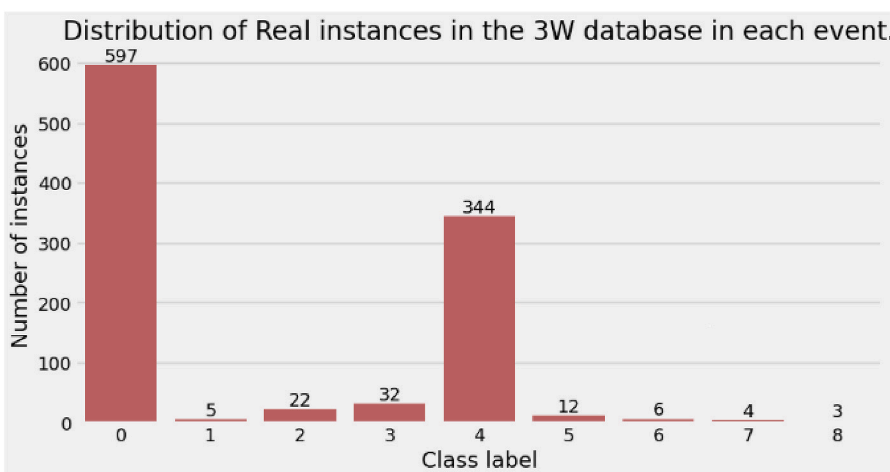
The best result was achieved by models with LSTM and GRU backbones of 2 layers with 20 hidden units each, giving an F1 score equal to 0.90 before SMOTE (Fig. 17) and 0.92 after SMOTE.

To summarise the observations from the initial algorithm settings, it is recognised that an increased number of layers doesn't improve the results. On the contrary, it makes them worse. A higher number of hidden units in each layer helps increase the classification metrics. However, it is unclear whether there is an optimal number that would significantly affect the results or whether 20 units is a plateau value, which has already produced the maximum possible F1 score.

5.1. Hyperparameters optimisation

The following hyperparameters were selected for the model optimisation:

- timestep (or window size)
- number of hidden units of each layer
- number of epochs



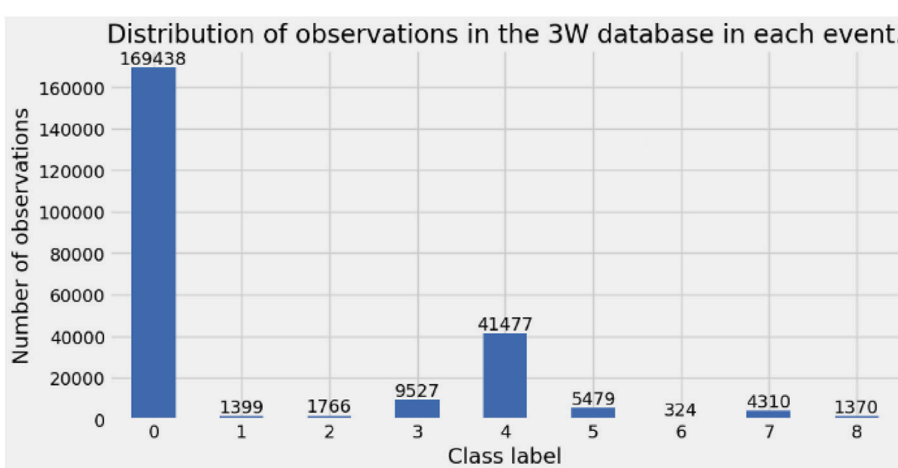


Fig. 14. Real instances and observations distribution according to fault events.

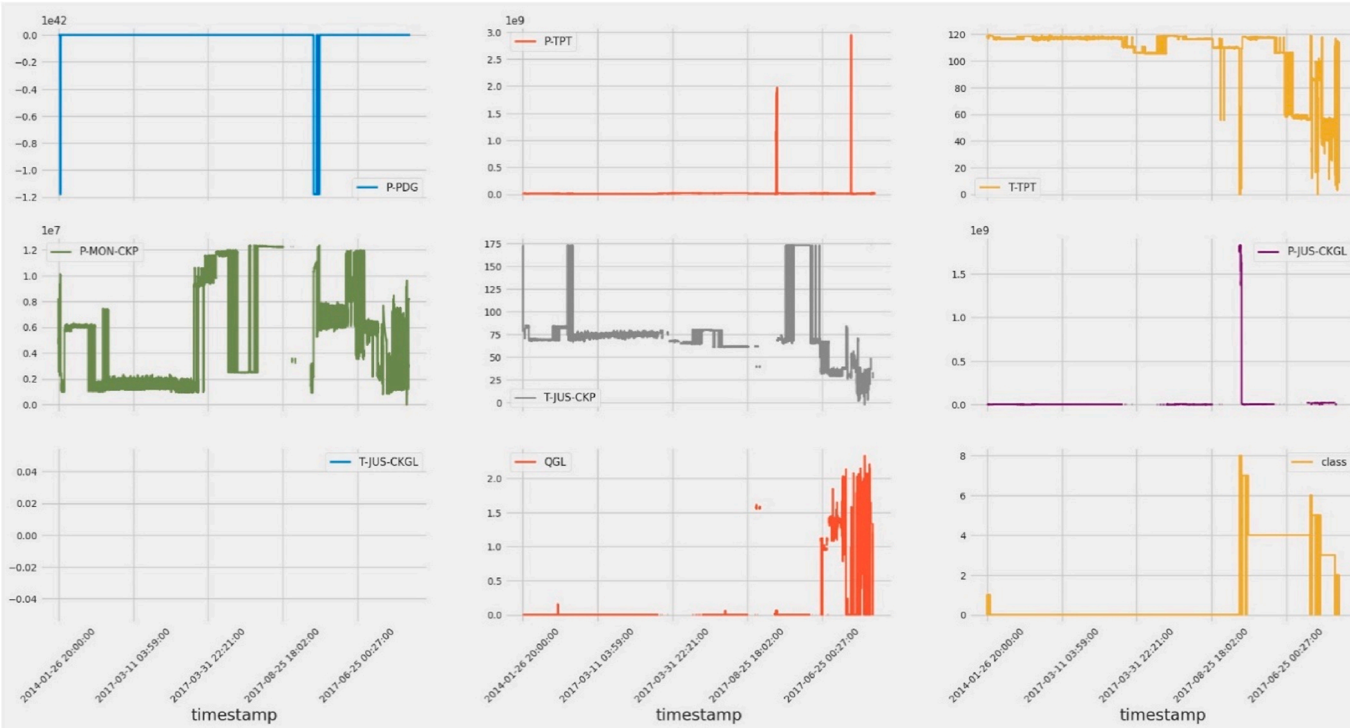


Fig. 15. Combined data visualisation.

- batch size.

Many other parameters, such as learning rate, activation function, optimiser type, and number of layers, could be optimised. Some have already been attempted to ensure the visibility and transparency of the model performance: both LSTM and GRU backbones were implemented with various activations, and several network layers were also built, which showed poorer results. To make computations less expensive, only quantitative parameters were estimated. RNN was run with just the GRU backbone since it is much faster and produces results similar to those of the LSTM. All attempts were made on the original without oversampling with SMOTE data for the abovementioned reasons. To investigate further effects of hyperparameter settings, their optimisation was performed using Random search, Hyperopt and GA to identify the best model for Classification. The Random Search and Hyperopt methods produced an F1 score equal to 0.94, an improvement from the original result.

5.2. GA 1

The DEAP framework was selected for this project since it provides a unique evolutionary algorithm that simplifies each step with its toolbox – a container of tools for all sorts of initialisers and genetic operators (*Overview — DEAP 1.3.3 Documentation*, n.d.). Each individual was encoded into a binary string of bit length 26, with the timestep equal to 8 bits, the number of hidden units 6, epoch 5 bits and a batch size of 7 bits. The gene initialisation values are chosen as the most appropriate for representing the decimal values of the hyperparameters. To identify the best model settings, the DEAP toolbox was run for 5 generations with a population size of 5 each. The hyperparameters range was not set to have limits, as was done with Random search and Hyperopt, since this method doesn't require an exhaustive search by iteration through the entire search space. The variables could have extreme values, and the best model achieved an F1 score of 0.96 with a window size equal to 1, a number of hidden units of 53, 29 epochs and a batch size of 25.

5.3. GA 2

Another approach is creating the GA by assigning all operators manually, which would give more transparency to the optimisation process and an opportunity to tailor the process for the task. The initial set up is similar to the previous experiment, with binary encoding, with each chromosome length equal to 26 bits. This time, the fitness function was selected as the F1 score of each model. The evolutionary algorithm included the selection of the fittest individual with maximum fitness in each population, one point crossover and mutation by flipping bits at the random change point.

The algorithm was run for 3 generations with 3 chromosomes in each population. It was highly desired to experiment with a higher number of individuals and iterations; however, due to technical limitations, it was not feasible. The best result was achieved with a timestep of 125, a number of hidden units of 61, epochs of 15 and a batch size of 99, resulting in a final F1 of 0.94.

5.4. GA 3

The third GA experiment was performed with value chromosome encoding, in which each phenotype is represented as a string of direct hyperparameter decimal values. The fitness function was set as a validation loss of each model. The evolutionary operators included tournament selection with size 3, one point crossover and random resetting mutation methods. The algorithm was attempted to run many times, with 5 generations and 3 generations and a corresponding population size; however, it was revealed that for successful evolution, the number of individuals should be sufficient for an increased chance of crossover and mutation. With a few chromosomes, the algorithm selected the best individual and converged early without attempting any other variations of hyperparameters. Increasing the number of generations and population size was not feasible, as it became extremely computationally expensive.

To overcome this issue, and considering GA 1 results with the DEAP toolbox, it was decided to decrease the timestep size to 1 and the number

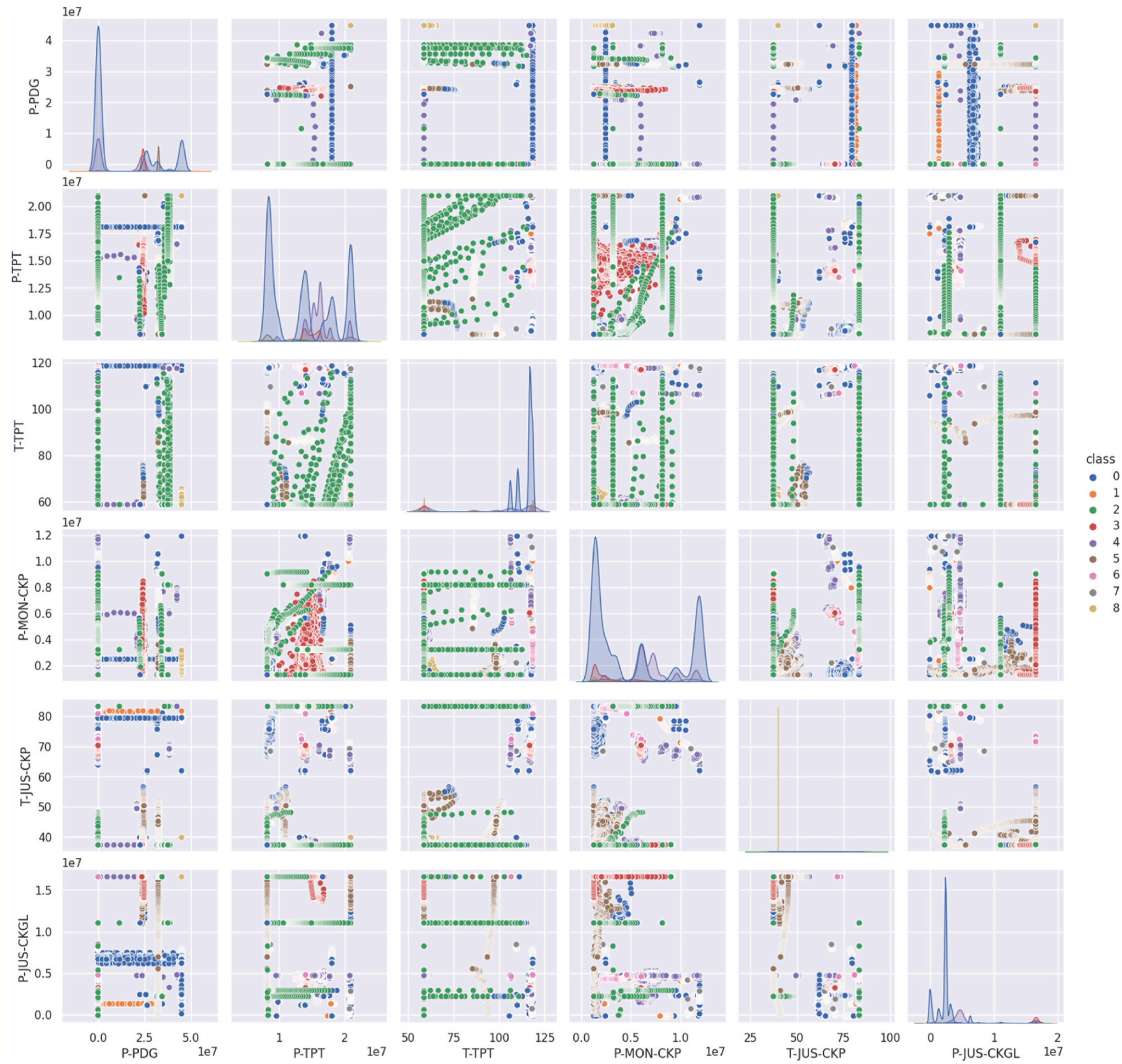


Fig. 16. Relationship between variables and classes.

of epochs to 1 for the initial set up, which would speed up the process and allow for increasing population size and the running of more generations. In this case, the objective was to optimise the number of hidden units and batch size, which could be implemented later for the final model evaluation with an arbitrary number of selected epochs.

The algorithm converged with the best model [1, 47, 1, 14] and the best model for all 5 generations. The variation in the best model fitness is due to the stochastic nature of RNN algorithms. However, the best fitness was still the smallest in each generation (Fig. 18).

To evaluate the final best model, it was run for an arbitrary 40 epochs, assigning the best model with a timestep of 1, the number of hidden units equal to 47 and a batch size of 14. As shown in Fig. 19, the F1 score achieved is 0.97, which is the best result for all the algorithms.

All hyperparameter tuning method outcomes are presented in Table 4, where the final F1 score per best model is calculated.

The best F1 score was achieved with GA. However, it is worth mentioning that a few limitations potentially affected the results. The first two Genetic Algorithms were performed with binary chromosome representation, so each parameter's upper and lower limits were not set but configured by selecting the number of bits for each gene. This allowed for an "accidental" application of the most extreme version of 3D GRU matrix selection, with a timestep equal to 1, i.e., selecting each observation for Classification rather than a window of several observations. Despite producing a good F1 score, it could create an "overfitting" issue if applied to new timeseries data.

For the current project, overfitting was carefully monitored in every application of LSTM and GRU by plotting training and validation loss, ensuring that the curves both decrease and converge after a selected number of epochs without further degradation in validation set performance. GRU, known for better handling the vanishing gradient problem, was chosen for the final evaluation. Additionally, regularisation techniques, such as early stopping, were applied during training and

Table 3
Deep neural networks model architectures F1 scores.

Model architecture	Before SMOTE F1	After SMOTE F1
Stacked 2 LSTM and 1 Dense, 10 hidden units, activation “tanh"	0.75	0.90
Stacked 2 LSTM and 1 Dense, 10 hidden units, activation “softmax"	0.19	0.71
Stacked 2 LSTM and 1 Dense, 10 hidden units, activation “relu"	0.85	0.88
Stacked 2 LSTM and 1 Dense, 20 hidden units, activation “relu"	0.90	0.92
Stacked 2 LSTM and 1 Dense, 10 hidden units, activation “LeakyReLU"	0.83	0.91
Stacked 2 LSTM and 1 Dense, 20 hidden units, activation “LeakyReLU"	0.85	0.85
Stacked 2 LSTM and 1 Dense, 10 hidden units, activation “swish"	0.87	0.88
Stacked 3 LSTM, 1 RepeatVector, 1 Dense, 10 hidden units, activation “relu"	0.82	0.90
Stacked 2 GRU and 1 Dense, 10 hidden units, activation “LeakyReLU"	0.86	0.87
Stacked 2 GRU and 1 Dense, 20 hidden units, activation “LeakyReLU"	0.90	0.92

implemented in each RNN model variation.

Another limitation was the difficulty of running all Genetic Algorithms for more than 5 generations. While DEAP allowed this to occur due to internal shortcuts through not running all individual evaluations, the second and third experiments struggled to run until the end. GA 2 could not be run for more than 3 generations, leading to poor individual crossover, mutation, and convergence with the best model on the 2nd generation. It could only be speculated that with higher computational resources, this algorithm could have been run for many more generations and individuals in each population, possibly resulting in a higher F1 score.

GA 3, in which each chromosome was represented as decimal values, was also attempted to run for more than 3 generations. However, since it also failed to run until the end, the decision was made to amalgamate the DEAP finding, where the best timestep was equal to 1, and rerun the GA 3 for just 1 epoch for each chromosome, but with an increased number of generations and individuals. This allowed for the simulation of the desired wide range of chromosome evaluation, assuming an increasing number of epochs on the most successful final model. As a result, the algorithm was allowed to create many variations of individuals by mutating and creating new children with crossover operators and produced the best F1 equal to 0.97.

This result might not be strictly comparable to the outcomes in other papers since each research applied different assumptions for evaluations, such as choice of training and testing sets treatment of faulty transient observations (in this project, they were combined with faulty steady state events and considered faults). Despite that, it can form the basis for developing further DL algorithms as a precedent with a confirmed good response to the task and a high attained outcome. Nevertheless, our model was better than the 2 other research that also did multiclass Classification based on the 3W Dataset, see Table 5. The optimised models, particularly those using GA, achieved a high F1 score of up to 0.97. This indicates a highly accurate anomaly detection system, which is the novelty of this work.

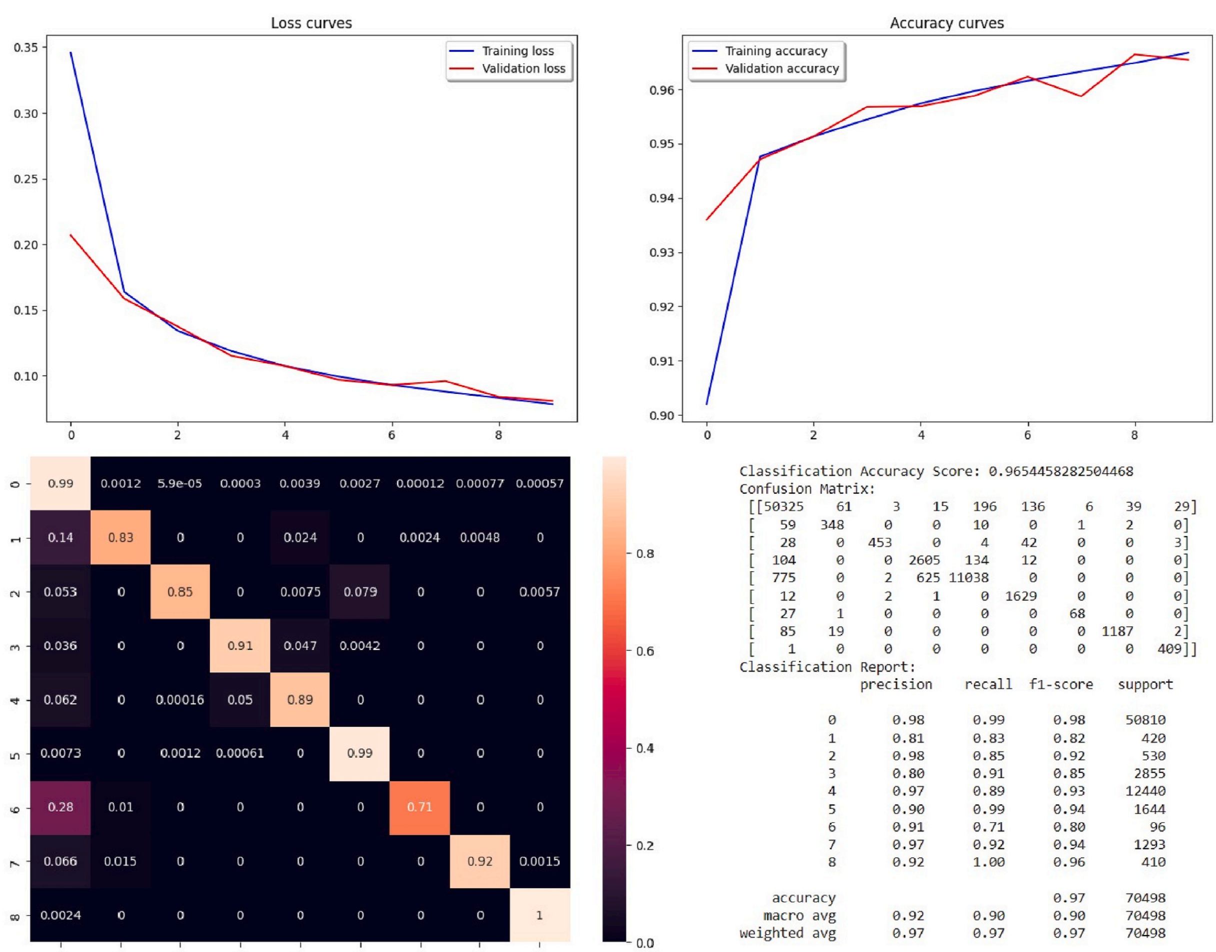


Fig. 17. LSTM with “relu” activation and 20 hidden units.

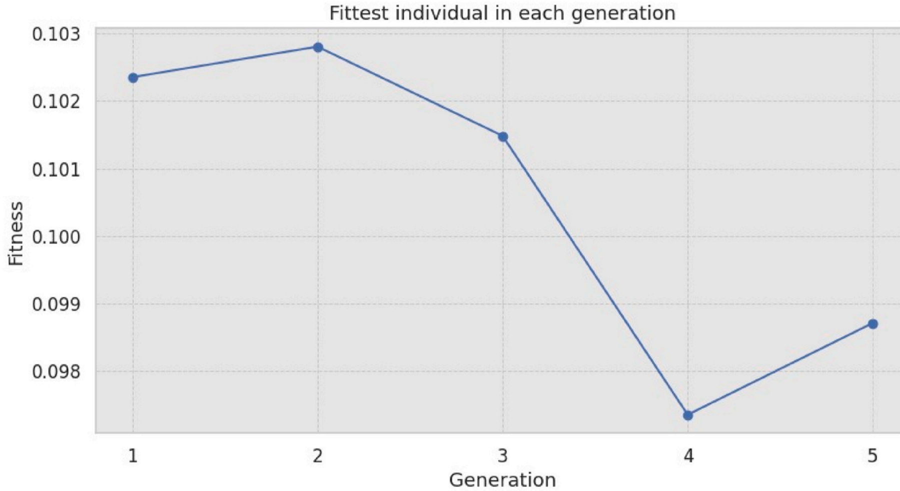


Fig. 18. GA 3 fitness evolution per generation.

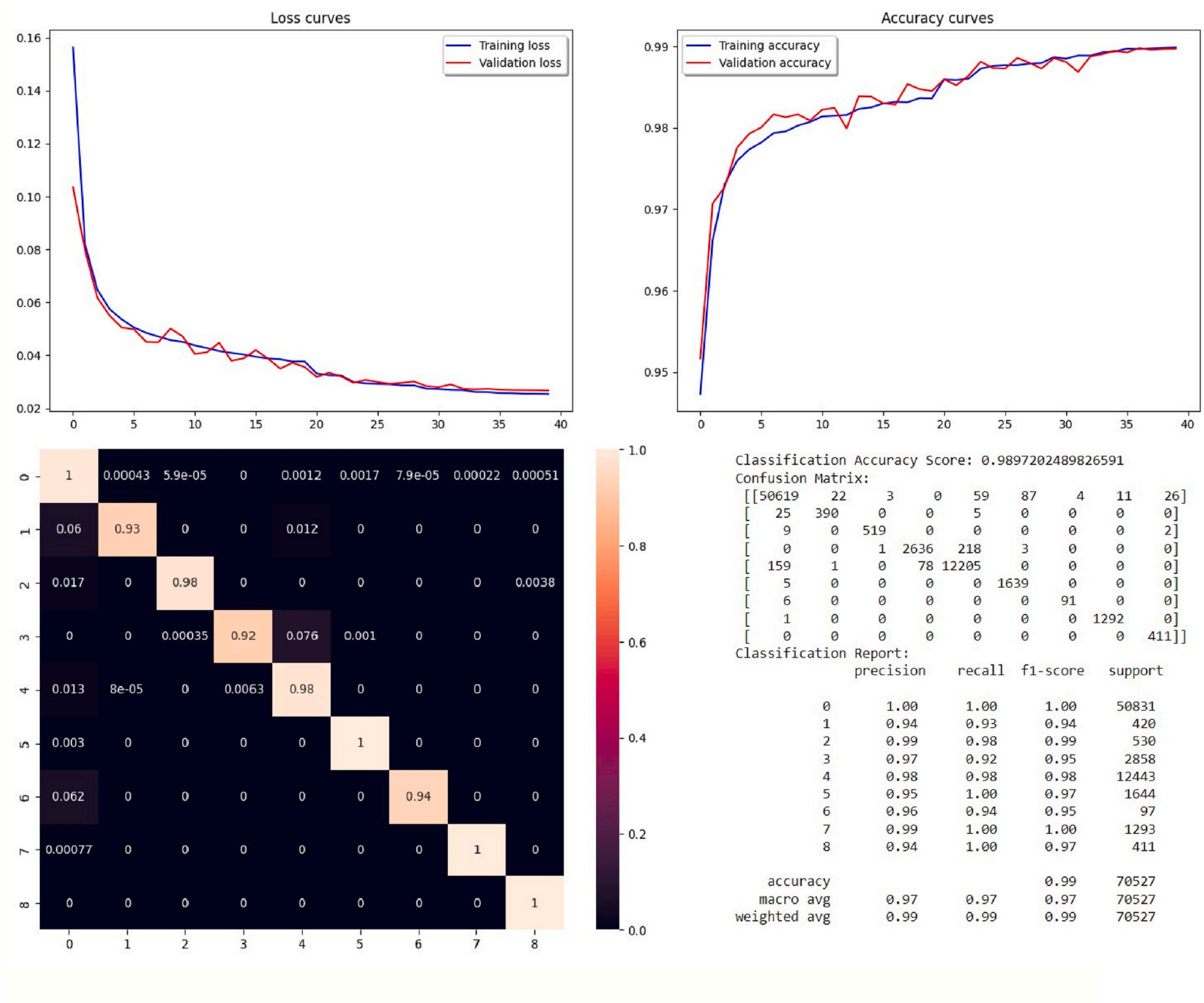


Fig. 19. GA 3 optimised model with F1 = 0.97.

6. Conclusions

In this project, detecting anomalies in oil and gas production was addressed using DL Neural Networks and GA for their optimisation. The 3W Dataset from Petrobras was taken as an example of labelled time series data, which can be applied to the pre-build algorithm and identification of anomalies with high accuracy.

For the project, two algorithms were selected as potentially most

Table 4
Hyperparameters optimisation best results.

Algorithm	Best model	F1 score
Random search	[23, 40, 16, 24]	0.94
Hyperopt	[10, 20, 20, 10]	0.94
GA 1	[1, 53, 29, 25]	0.96
GA 2	[125, 61, 15, 99]	0.94
GA 3	[1, 47, 40, 14]	0.97

Table 5
– Results comparison.

3W Dataset Anomaly Detection and Classification				
No	Publication	Research Question	Methods	Results
1	Classification of undesirable events in oil well operation (Turan and Jaschke, 2021)	Multiclass Classification of anomalous events in oil wells for 3W Dataset	Decision Tree, as baseline attempted Logistic Regression (LR), Support Vector Classifier (SVC), Linear and Quadratic Discriminant Analysis (LDA & QDA), Random Forest, AdaBoost (ADA), Principal Component Analysis (PCA)	Precision = 0.83 Recall = 0.88 F1 = 0.85
2	Predictive maintenance for offshore oil wells by means of deep learning features extraction (Gatta et al., 2022)	Multiclass Classification of anomalous events in oil wells for 3W Dataset	Deep learning method for feature extraction: 1D AutoEncoder using Convolutional Neural Network. Machine learning classifiers: Random Forest, Nearest Neighbors, Gaussian Naive Bayes and Quadratic Discriminant Analysis, hyperparameters selection via Biased Random Key Genetic Algorithm (BRKGA).	Precision = 0.95 Recall = 0.517 F1 = 0.898
3	**Our Model**	**Anomaly detection in the production of oil and gas using Deep Neural Networks and Genetic Algorithms for their optimisation**	**Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures, Optimized using Random Search, Hyperopt, and Genetic Algorithms**	**Precision = 0.96 Recall = 0.98 F1 = 0.97**

powerful for resolving such tasks, and they were currently unexplored, according to the literature review, namely RNN configurations with LSTM and GRU architectures.

The research was conducted through several steps, beginning with the basic setup of an LSTM backbone to evaluate the model's general response. The network comprised two stacked LSTM layers and one Dense layer, resulting in a macro average F1 score of 0.75.

To assess further the general response of various backbones and their parameters, the algorithms were executed with consistent batch size, timestep, and the number of epochs, but different activation functions ("relu", "softmax", "LeakyReLU", or "swish"), hidden units, and varying numbers of LSTM or GRU layers. These configurations were applied both before and after the SMOTE oversampling of the training data.

The initial algorithm settings revealed that increasing the number of layers did not improve the results. However, a higher number of hidden units per layer did enhance the classification metrics. Since it remained unclear whether there was an optimal number of hidden units that would significantly impact the results, further optimisation of the hyperparameter settings was performed using Random search, Hyperopt, and 3 versions of GA to identify the best model for Classification.

Timestep (or window size), number of hidden units of each layer, number of epochs and batch size were selected for the model optimisation. Both LSTM and GRU backbones were implemented with various activation functions, and several network layer configurations were tested. To reduce computational expense, only quantitative parameters were estimated. For the final evaluation, the RNN was run with the GRU backbone, as it is faster and yields results similar to those of the LSTM. All attempts were made using the original data without SMOTE oversampling for the reasons mentioned above.

Both random search and hyperopt showed the same result, achieving an F1 equal to 0.94. Next, three GA were performed to tune the hyperparameters that were encoded into vectors of decimals by an iterative process of improving an initial random population, using nature-inspired concepts like selection, crossover, and mutation. To apply genetic operators, each phenotype is converted to a genotype, entirely independent of fitness and represented in binary form using 0s and 1s.

GA 1 was performed using the DEAP framework, in which each individual was encoded into a binary string of 26 bits, with the timestep represented by 8 bits, the number of hidden units by 6 bits, the number of epochs by 5 bits, and the batch size by 7 bits, which were further decoded into integer values. To identify the optimal model settings, the DEAP toolbox was executed for 5 generations, each with a population size of 5, and the loss function was set up as validation loss. The best model achieved an F1 score of 0.96 with a window size of 1, 53 hidden units, 29 epochs, and a batch size 25.

GA 2 was designed using binary encoding; each chromosome was 26 bits in length. The fitness function was defined as the F1 score of each model. The evolutionary algorithm involved selecting the fittest individual with the highest fitness in each population, performing one-point crossover, and applying mutation by flipping bits at random points. Due to memory limitations, the algorithm was run for 3 generations and 3 individuals, resulting in a final F1 score of 0.94 with a timestep of 125, 61 hidden units, 15 epochs, and a batch size of 99.

GA 3 was an experiment with a value chromosome encoding; in this approach, each phenotype is represented as a string of direct hyperparameter decimal values. The fitness function was defined as the validation loss of each model. The evolutionary operators used included tournament selection with a size of 3, one-point crossover, and random resetting mutation methods. The best result, F1 equal to 0.97, was achieved with a timestep 1, 47 hidden units, 40 epochs and batch size 14. This accomplishment represents the highest F1 score attained among all other attempts, making it the most effective solution for the task at hand that can be implemented for other labelled time series anomalies detection and Classification, that can be implemented for other labelled time series anomalies detection and Classification.

The main limitation of the project, as with any DL and Big Data projects, is the technical capacity to perform the algorithms for the desired number of epochs and iterations. Since the main objective was creating a pipeline of models for detecting and classifying anomalies, Google Colab was sufficient to obtain the primary results and an F1 score for quality justification. However, more computational resources will be needed in the more robust calculation required by GA hyperparameter optimisation.

For future work, it would be highly interesting to employ XAI algorithms to interpret the DL algorithms as black box models due to the hidden nature of layers and neurons' actions to foster transparency. With a better understanding of which parameters reveal potential faults and the need to diagnose them sooner with higher precision within the scope of the suggested DL algorithm, anomaly detection can become a straightforward task for reservoir and production engineers in the Oil and Gas industry.

Funding

This work was supported by national funds through FCT (Fundação para a Ciência e a Tecnologia), under the project - UIDB/04152/2020 - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS (https://doi.org/10.54499/UIDB/04152/2020).

CRediT authorship contribution statement

Guzel Bayazitova: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis. **Maria Anastasiadou:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Vitor Duarte dos Santos:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Agin, F., Khosravanian, R., Karimifard, M., Jahanshahi, A., 2020. Application of adaptive neuro-fuzzy inference system and data mining approach to predict lost circulation using DOE technique (case study: Maroon oilfield). Petroleum 6 (4), 423–437. https://doi.org/10.1016/j.petlm.2018.07.005.

Alharbi, B., Liang, Z., Aljindan, J.M., Agnia, A.K., Zhang, X., 2022. Explainable and interpretable anomaly detection models for production data. SPE J. 27 (1), 349–363. https://doi.org/10.2118/208586-PA.

Aljameel, S.S., Alomari, D.M., Alismail, S., Khawaher, F., Alkhudhair, A.A., Aljubran, F., Alzannan, R.M., 2022. An anomaly detection model for oil and gas pipelines using machine learning. Computation 10 (8), 138. https://doi.org/10.3390/COMPUTATION10080138, 2022, Vol. 10, Page 138.

Aljubran, M., Ramasamy, J., Albassam, M., Magana-Mora, A., 2021. Deep learning and time-series analysis for the early detection of lost circulation incidents during drilling operations. IEEE Access 9, 76833–76846. https://doi.org/10.1109/ACCESS.2021.3082557.

Alsaihati, A., Abughaban, M., Elkatatny, S., Shehri, D. Al, 2022. Application of machine learning methods in modeling the loss of circulation rate while drilling operation. ACS Omega 7 (24), 20696–20709. https://doi.org/10.1021/ACSOMEGA.2C00970/ASSET/IMAGES/LARGE/AO2C00970_0009.JPEG.

Alsaihati, A., Elkatatny, S., Mahmoud, A.A., Abdulraheem, A., 2021. Use of machine learning and data analytics to detect downhole abnormalities while drilling horizontal wells, with real case study. Journal of Energy Resources Technology, Transactions of the ASME 143 (4). https://doi.org/10.1115/1.4048070/1086232.

Aslam, N., Khan, I.U., Alansari, A., Alrammah, M., Alghwairy, A., Alqahtani, R., Alqahtani, R., Almushikes, M., Hashim, M.A.L., 2022. Anomaly detection using explainable random forest for the prediction of undesirable events in oil wells. Applied Computational Intelligence and Soft Computing 2022. https://doi.org/10.1155/2022/1558381.

Barbariol, T., Feltresi, E., Susto, G.A., 2020. Self-diagnosis of Multiphase flow meters through machine learning-based anomaly detection. Energies 13 (12). https://doi.org/10.3390/en13123136.

Brønstad, C., Netto, S.L., Ramos, A.L.L., 2021. Data-driven detection and identification of undesirable events in subsea oil wells. In: SENSORDEVICES 2021 : the Twelfth International Conference on Sensor Device Technologies and Applications.

Carvalho, B.G., Vargas, R.E.V., Salgado, R.M., Munaro, C.J., Varejao, F.M., 2021a. Flow instability detection in offshore oil wells with multivariate time series machine learning classifiers. In: IEEE International Symposium on Industrial Electronics. https://doi.org/10.1109/ISIE45552.2021.9576310, 2021-June.

Carvalho, B.G., Vargas, R.E.V., Salgado, R.M., Munaro, C.J., Varejão, F.M., 2021b. Hyperparameter tuning and feature selection for improving flow instability detection in offshore oil wells. In: IEEE International Conference on Industrial Informatics (INDIN). https://doi.org/10.1109/INDIN45523.2021.9557415, 2021-July.

D'Almeida, A.L., Bergiante, N.C.R., de Souza Ferreira, G., Leta, F.R., de Campos Lima, C. B., Lima, G.B.A., 2022. Digital transformation: a review on artificial intelligence techniques in drilling and production applications. Int. J. Adv. Manuf. Technol. 119 (9–10), 5553–5582. https://doi.org/10.1007/S00170-021-08631-W.

Das, S., Tariq, A., Santos, T., Kantareddy, S.S., Banerjee, I., 2023. Recurrent neural networks (RNNs): architectures, training tricks, and introduction to influential research. Neuromethods 197, 117–138. https://doi.org/10.1007/978-1-0716-3195-9_4/FIGURES/9.

Figueirêdo, I.S., Carvalho, T.F., Silva, W.J.D., Guarieiro, L.L.N., Nascimento, E.G.S., 2021. Detecting interesting and anomalous patterns in multivariate time-series data in an offshore platform using unsupervised learning. Offshore Technology Conference. https://doi.org/10.4043/31297-MS.

Gasparovic, B., Lerga, J., Mausa, G., Ivasic-Kos, M., 2022. Deep learning approach for objects detection in underwater pipeline images. Appl. Artif. Intell. 36 (1) https://doi.org/10.1080/08839514.2022.2146853.

Gatta, F., Giampaolo, F., Chiaro, D., Piccialli, F., 2022. Predictive maintenance for offshore oil wells by means of deep learning features extraction. Expet Syst., e13128 https://doi.org/10.1111/EXSY.13128.

Giro, R.A., Bernasconi, G., Giunta, G., Cesari, S., 2021. A data-driven pipeline pressure procedure for remote monitoring of centrifugal pumps. J. Petrol. Sci. Eng. 205, 108845 https://doi.org/10.1016/j.petrol.2021.108845.

Goodfellow, I., Bengio, Y., & Courville, A. (n.d.). Deep Learning.

Gurina, E., Klyuchnikov, N., Antipova, K., Koroteev, D., 2022a. Forecasting the abnormal events at well drilling with machine learning. Appl. Intell. 52 (9), 9980–9995. https://doi.org/10.1007/s10489-021-03013-x.

Gurina, E., Klyuchnikov, N., Antipova, K., Koroteev, D., 2022b. Making the black-box brighter: interpreting machine learning algorithm for forecasting drilling accidents. J. Petrol. Sci. Eng. 218 https://doi.org/10.1016/j.petrol.2022.111041.

Gurina, E., Klyuchnikov, N., Zaytsev, A., Romanenkova, E., Antipova, K., Simon, I., Makarov, V., Koroteev, D., 2020. Application of machine learning to accidents detection at directional drilling. J. Petrol. Sci. Eng. 184 https://doi.org/10.1016/j.petrol.2019.106519.

Hasan, M.H., Malik, A.A., Jasamai, M., 2017. A review on anomaly detection methods for optimising oil well surveillance. IJCSNS International Journal of Computer Science and Network Security 17 (11).

Ide, T., Khandelwal, A., Kalagnanam, J., 2017. Sparse Gaussian Markov random field mixtures for anomaly detection. IEEE Xplore 955–960. https://doi.org/10.1109/ICDM.2016.0119.

Kirschbaum, L., Roman, D., Singh, G., Bruns, J., Robu, V., Flynn, D., 2020. AI-Driven maintenance support for downhole tools and electronics operated in dynamic drilling environments. IEEE Access 8, 78683–78701. https://doi.org/10.1109/ACCESS.2020.2990152.

Koroteev, D., Tekic, Z., 2021. Artificial intelligence in oil and gas upstream: trends, challenges, and scenarios for the future. Energy and AI 3, 100041. https://doi.org/10.1016/J.EGYAI.2020.100041.

Kuang, L., Liu, H., Ren, Y., Luo, K., Shi, M., Su, J., Li, X., 2021. Application and development trend of artificial intelligence in petroleum exploration and development. Petrol. Explor. Dev. 48 (1), 1–14. https://doi.org/10.1016/S1876-3804(21)60001-0.

Lashari, S.Z., Takbiri-Borujeni, A., Fathi, E., Sun, T., Rahmani, R., Khazaeli, M., 2019. Drilling performance monitoring and optimisation: a data-driven approach. J. Pet. Explor. Prod. Technol. 9 (4), 2747–2756. https://doi.org/10.1007/s13202-019-0657-2.

Li, M., Zhang, H.R., Zhao, Q., Liu, W., Song, X.Z., Ji, Y.Y., Wang, J.S., 2022. A new method for intelligent prediction of drilling overflow and leakage based on multi-parameter fusion. Energies 15 (16). https://doi.org/10.3390/en15165988.

Liu, J., Feng, J., Gao, X., 2019. Fault diagnosis of rod pumping wells based on support vector machine optimised by improved chicken Swarm optimisation. IEEE Access 7, 171598–171608. https://doi.org/10.1109/ACCESS.2019.2956221.

Liu, Y., Yao, K., Liu, S., Raghavendra, C.S., Lenz, T.L., Olabinjo, L., Seren, B., Seddighrad, C.S., Babu, C.G.D., 2010. Failure prediction for rod pump artificial lift systems. In: Society of Petroleum Engineers Western North American Regional Meeting 2010 - in Collaboration with the Joint Meetings of the Pacific Section AAPG and Cordilleran Section GSA, vol. 2, pp. 845–852. https://doi.org/10.2118/133545-MS.

Liu, Y., Yao, K.T., Liu, S., Raghavendra, C.S., Balogun, O., Olabinjo, L., 2011. Semi-supervised failure prediction for oil production wells. In: Proceedings - IEEE International Conference on Data Mining, ICDM, pp. 434–441. https://doi.org/10.1109/ICDMW.2011.151.

Lv, X., Wang, H., Zhang, X., Liu, Y., Jiang, D., Wei, B., 2021. An evolutional SVM method based on incremental algorithm and simulated indicator diagrams for fault diagnosis in sucker rod pumping systems. J. Petrol. Sci. Eng. 203 https://doi.org/10.1016/j.petrol.2021.108806.

Lv, X.-X., Wang, H.-X., Xin, Z., Liu, Y.-X., Zhao, P.-C., 2022. Adaptive fault diagnosis of sucker rod pump systems based on optimal perceptron and simulation data. Petrol. Sci. 19 (2), 743–760. https://doi.org/10.1016/j.petsci.2021.09.012.

Machado, A.P.F., Vargas, R.E.V., Ciarelli, P.M., Munaro, C.J., 2022. Improving performance of one-class classifiers applied to anomaly detection in oil wells. J. Petrol. Sci. Eng. 218, 110983 https://doi.org/10.1016/J.PETROL.2022.110983.

Magana-Mora, A., Affleck, M., Ibrahim, M., Makowski, G., Kapoor, H., Otalvora, W.C., Jamea, M.A., Umairin, I.S., Zhan, G., Gooneratne, C.P., 2021. Well control space out: a deep-learning approach for the optimisation of drilling safety operations. IEEE Access 9, 76479–76492. https://doi.org/10.1109/ACCESS.2021.3082661.

Marins, M.A., Barros, B.D., Santos, I.H., Barrionuevo, D.C., Vargas, R.E.V., de, M., Prego, T., de Lima, A.A., de Campos, M.L.R., da Silva, E.A.B., Netto, S.L., 2021. Fault detection and Classification in oil wells and production/service lines using random forest. J. Petrol. Sci. Eng. 197, 107879 https://doi.org/10.1016/J.PETROL.2020.107879.

Martí, L., Sanchez-Pi, N., Molina, J.M., Garcia, A.C.B., 2015. Anomaly detection based on sensor data in petroleum industry applications. Sensors 15 (2), 2774–2797. https://doi.org/10.3390/S150202774, 2015, Vol. 15, Pages 2774-2797.

Martí, L., Sanchez-Pi, N., Molina, J.M., Garcia, A.C.B., 2017. On the combination of support vector machines and segmentation algorithms for anomaly detection: a petroleum industry comparative study. J. Appl. Logic 24, 71–84. https://doi.org/10.1016/J.JAL.2016.11.015.

Mello, L.H.S., Oliveira-Santos, T., Varejão, F.M., Ribeiro, M.P., Rodrigues, A.L., 2022. Ensemble of metric learners for improving electrical submersible pump fault diagnosis. J. Petrol. Sci. Eng. 218, 110875 https://doi.org/10.1016/J.PETROL.2022.110875.

Mello, L.H.S., Ribeiro, M.P., Oliveira-Santos, T., Varejão, F.M., Rodrigues, A.L., 2020. Metric learning for electrical submersible pump Fault Diagnosis. In: Proceedings of the International Joint Conference on Neural Networks. https://doi.org/10.1109/IJCNN48605.2020.9207133.

Mopuri, K.R., Bilen, H., Tsuchihashi, N., Wada, R., Inoue, T., Kusanagi, K., Nishiyama, T., Tamamura, H., 2022. Early sign detection for the stuck pipe scenarios using unsupervised deep learning. J. Petrol. Sci. Eng. 208 https://doi.org/10.1016/j.petrol.2021.109489.

Muojeke, S., Venkatesan, R., Khan, F., 2020. Supervised data-driven approach to early kick detection during drilling operation. J. Petrol. Sci. Eng. 192 https://doi.org/10.1016/j.petrol.2020.107324.

OLGA Dynamic Multiphase Flow Simulator. (n.d.). Retrieved March 7, 2023, from https://www.software.slb.com/products/olga.

Orestes, A., Castro, D.S., De, M., Santos, J.R., Rodrigues Leta, F., Benevenuto, C., Lima, C., Brito, G., Lima, A., Jesus, D., Santos, R., Leta, M., Lima, F.R., Lima, C.B.C., 2021. Unsupervised methods to classify real data from offshore wells. Am. J. Oper. Res. 11 (5), 227–241. https://doi.org/10.4236/AJOR.2021.115014.

Overview — DEAP 1.3.3 documentation. (n.d.). Retrieved May 9, 2023, from https://deap.readthedocs.io/en/master/overview.html.

Pandey, Y.N., Rastogi, A., Kainkaryam, S., Bhattacharya, S., Saputelli, L., 2020. Machine learning in the oil and gas industry. In: Machine Learning in the Oil and Gas Industry. Apress. https://doi.org/10.1007/978-1-4842-6094-4.

Patri, O.P., Reyna, N., Panangadan, A., Prasanna, V., 2015. Predicting compressor valve failures from multi-sensor data. SPE Western Regional Meeting 2015: Old Horizons, New Horizons Through Enabling Technology 725–735. https://doi.org/10.2118/174044-MS.

Peng, L., Han, G., Landjobo Pagou, A., Shu, J., 2020. Electric submersible pump broken shaft fault diagnosis based on principal component analysis. J. Petrol. Sci. Eng. 191, 107154 https://doi.org/10.1016/J.PETROL.2020.107154.

Qin, H., Srivastava, V., Wang, H., Zerpa, L.E., Koh, C.A., 2019. Machine learning models to predict gas hydrate plugging risks using flowloop and field data. In: Proceedings of the Annual Offshore Technology Conference. https://doi.org/10.4043/29411-MS, 2019-May.

Ravishankar, P., Hwang, S., Zhang, J., Khalilullah, I.X., Eren-Tokgoz, B., 2022. DARTS-drone and artificial intelligence reconsolidated technological solution for increasing the oil and gas pipeline resilience. Int. J. Disaster Risk Sci. 13 (5), 810–821. https://doi.org/10.1007/s13753-022-00439-w.

Salem, A.M., Yakoot, M.S., Mahmoud, O., 2022. A novel machine learning model for autonomous analysis and diagnosis of well integrity failures in artificial-lift production systems. Advances in Geo-Energy Research 6 (2), 123–142. https://doi.org/10.46690/ager.2022.02.05.

Santos, M., de, J.R., Castro, A.O. de S., Leta, F.R., De Araujo, J.F.M., Ferreira, G. de S., Santos, R. de A., Lima, C.B. de C., Lima, G.B.A., 2021. Statistical analysis of offshore production sensors for failure detection applications/Análise estatística dos sensores de produção offshore para aplicações de detecção de falhas. Brazilian Journal of Development 7 (8), 85880–85898. https://doi.org/10.34117/BJDV7N8-681.

Scoralick, R., Scoralick Fontoura do Nascimento, R., Henrique Groenner, B., Vargas, R.E.V., Humberto Ferreira dos Santos, I., 2021. Fault detection with Stacked Autoencoders and pattern recognition techniques in gas lift operated oil wells. In: XLII Ibero-Latin-American Congress on Computational Methods in Engineering AndIII Pan-American Congress on Computational Mechanics. ABMEC-IACMRio de Janeiro, Brazil. November 9-12, 2021. https://www.researchgate.net/publication/363279803.

Seo, Y., Kim, B., Lee, J., Lee, Y., 2021. Development of ai-based diagnostic model for the prediction of hydrate in gas pipeline. Energies 14 (8). https://doi.org/10.3390/en14082313.

Shrifan, N.H.M.M., Jawad, G.N., Isa, N.A.M., Akbar, M.F., 2021. Microwave nondestructive testing for defect detection in composites based on K-means clustering algorithm. IEEE Access 9, 4820–4828. https://doi.org/10.1109/ACCESS.2020.3048147.

Sinha, S., de Lima, R.P., Lin, Y.Z., Sun, A.Y., Symons, N., Pawar, R., Guthrie, G., 2020. Normal or abnormal? Machine learning for the leakage detection in carbon sequestration projects using pressure field data. Int. J. Greenh. Gas Control 103. https://doi.org/10.1016/j.ijggc.2020.103189.

Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., Oza, H., 2021. Application of machine learning and artificial intelligence in oil and gas industry. Petroleum Research 6 (4), 379–391. https://doi.org/10.1016/J.PTLRS.2021.05.009.

Soriano-Vargas, A., Werneck, R., Moura, R., Mendes Júnior, P., Prates, R., Castro, M., Gonçalves, M., Hossain, M., Zampieri, M., Ferreira, A., Davólio, A., Hamann, B., Schiozer, D.J., Rocha, A., 2021a. A visual analytics approach to anomaly detection in hydrocarbon reservoir time series data. J. Petrol. Sci. Eng. 206, 108988 https://doi.org/10.1016/J.PETROL.2021.108988.

Soriano-Vargas, A., Werneck, R., Moura, R., Mendes Júnior, P., Prates, R., Castro, M., Gonçalves, M., Hossain, M., Zampieri, M., Ferreira, A., Davólio, A., Hamann, B., Schiozer, D.J., Rocha, A., 2021b. A visual analytics approach to anomaly detection in hydrocarbon reservoir time series data. J. Petrol. Sci. Eng. 206, 108988 https://doi.org/10.1016/J.PETROL.2021.108988.

Su, J., Zhao, Y., He, T., Luo, P., 2021. Prediction of drilling leakage locations based on optimised neural networks and the standard random forest method. Oil and Gas Science and Technology – Revue d'IFP Energies Nouvelles 76. https://doi.org/10.2516/ogst/2021003.

Tan, C., Chen, P., Feng, Z., Ai, X., Lu, M., Zhou, Q., Feng, G., 2022. Multiscale normalization method combined with a deep CNN diagnosis model of dynamometer card in SRP well. Front. Earth Sci. 10 https://doi.org/10.3389/feart.2022.852633.

The Defining Series: Artificial Lift | SLB. (n.d.). Retrieved January 30, 2023, from https://www.slb.com/resource-library/oilfield-review/defining-series/defining-artificial-lift.

The SLB Energy Glossary | Energy Glossary. (n.d.). Retrieved March 7, 2023, from https://glossary.slb.com/.

Turan, E.M., Jaschke, J., 2021. Classification of undesirable events in oil well operation. In: Proceedings of the 2021 23rd International Conference on Process Control, PC 2021, pp. 157–162. https://doi.org/10.1109/PC52310.2021.9447527.

United Nations, 2022. THE 17 GOALS | Sustainable Development, pp. 50–53. https://sdgs.un.org/goals.

Vankov, Y., Rumyantsev, A., Ziganshin, S., Politova, T., Minyazev, R., Zagretdinov, A., 2020. Assessment of the condition of pipelines using convolutional neural networks. Energies 13 (3). https://doi.org/10.3390/en13030618.

Vargas, R.E.V., Munaro, C.J., Ciarelli, P.M., De Araujo, J.C.D., 2017. Proposal for two classifiers of offshore naturally flowing wells events using k-nearest neighbors, sliding windows and time multiscale. In: 2017 6th International Symposium on Advanced Control of Industrial Processes, AdCONIP, pp. 209–214. https://doi.org/10.1109/ADCONIP.2017.7983782, 2017.

Vargas, R.E.V., Munaro, C.J., Ciarelli, P.M., Medeiros, A.G., Amaral, B. G. do, Barrionuevo, D.C., Araújo, J. C. D. de, Ribeiro, J.L., Magalhães, L.P., 2019a. A realistic and public dataset with rare undesirable real events in oil wells. J. Petrol. Sci. Eng. 181, 106223 https://doi.org/10.1016/J.PETROL.2019.106223.

Vargas, R.E.V., Munaro, C.J., Ciarelli, P.M., Medeiros, A.G., Amaral, B. G. do, Barrionuevo, D.C., Araújo, J. C. D. de, Ribeiro, J.L., Magalhães, L.P., 2019b. A realistic and public dataset with rare undesirable real events in oil wells. J. Petrol. Sci. Eng. 181, 106223 https://doi.org/10.1016/J.PETROL.2019.106223.

Wei, J., Gao, X., 2020. Fault diagnosis of sucker rod pump based on deep-broad learning using motor data. IEEE Access 8, 222562–222571. https://doi.org/10.1109/ACCESS.2020.3036078.

Wong, P., Wong, W.K., Juwono, F.H., Gopal, L., Yusoff, M.A., 2022. A minimalist approach for detecting sensor abnormality in oil and gas platforms. Petroleum Research 7 (2), 177–185. https://doi.org/10.1016/j.ptlrs.2021.09.007.

Wood, D.A., Mardanirad, S., Zakeri, H., 2022. Effective prediction of lost circulation from multiple drilling variables: a class imbalance problem for machine and deep learning algorithms. J. Pet. Explor. Prod. Technol. 12 (1), 83–98. https://doi.org/10.1007/s13202-021-01411-y.

Zhang, R., Wang, L., Chen, D., 2021. An intelligent diagnosis method of the working conditions in sucker-rod pump wells based on convolutional neural networks and transfer learning. Energy Eng. J. Assoc. Energy Eng.: Journal of the Association of Energy Engineering 118 (4), 1071–1082. https://doi.org/10.32604/EE.2021.014961.

Zhang, Y., Yang, K., 2022. Fault diagnosis of submersible motor on offshore platform based on multi-signal fusion. Energies 15 (3). https://doi.org/10.3390/en15030756.