VUTTIPITTAYAMONGKOL, P., TUNG, A. and ELYAN, E. 2021. Towards machine learning-driven practices for oil and gas decommissioning: introduction of a new offshore pipeline dataset. In *Proceedings of the 9th International conference on computer and communications management (ICCCM 2021)*, 16-18 July 2021, [virtual event]. New York: ACM [online], pages 111-116. Available from: https://doi.org/10.1145/3479162.3479179

Towards machine learning-driven practices for oil and gas decommissioning: introduction of a new offshore pipeline dataset.

VUTTIPITTAYAMONGKOL, P., TUNG, A. and ELYAN, E.

2021

Towards Machine Learning-Driven Practices for Oil and Gas Decommissioning – Introduction of a New Offshore Pipeline Dataset

Pattaramon Vuttipittayamongkol
pattaramon.vut@mfu.ac.th
School of Information Technology, Mae Fah Luang University
Chiang Rai, Thailand

Aaron Tung
aaron.tung@postgrad.curtin.edu.au
Aberdeen University Centre for Energy Law, King's College
Aberdeen, UK

Eyad Elyan
e.elyan@rgu.ac.uk
School of Computing, Robert Gordon University
Aberdeen, UK

ABSTRACT

Thousands of offshore oil and gas structures worldwide are approaching the end of their operating lifespan. Decommissioning processes are expensive and normally take years to finish as various options need to be analysed based on numerous stakeholders' preferences. Despite recent and significant progress in machine learning and data-driven applications in the oil and gas industry, very little work has been done in the area of using machine learning to inform the decommissioning processes and operations. This can be attributed to the lack of relevant public datasets with sufficient samples. In this paper, we present a new oil and gas decommissioning dataset comprised of 708 real samples extracted from over a hundred company proposals and reports. A supervised learning algorithm was applied to the dataset to predict the decommissioning option. Experiments and results suggest that a machine learning approach can greatly help shorten the traditional analysis process while providing decent accuracy. The classification results of this work serve as a baseline to motivate further experiments and enable the research community to broaden and advance the knowledge in this prominent and timely topic.

CCS CONCEPTS

• **Computing methodologies → Machine learning**; • **Applied computing → Decision analysis**.

KEYWORDS

Oil and gas, Decommissioning, Machine learning, Decision support solution, Stakeholder management, Multi-criteria decision analysis tool, Offshore structures

ACM Reference Format:
Pattaramon Vuttipittayamongkol, Aaron Tung, and Eyad Elyan. 2021. Towards Machine Learning-Driven Practices for Oil and Gas Decommissioning – Introduction of a New Offshore Pipeline Dataset. In *The 2021 9th International Conference on Computer and Communications Management*

ICCCM '21, July 16–18, 2021, Singapore, Singapore

1 INTRODUCTION

Decommissioning is currently a challenging topic in the oil and gas industry. The number of offshore installations worldwide is rapidly growing to over 12,000, many of which are soon to be removed [7]. Tens of billions US dollars of expenses is expected as all related costs could be approximately half of the oil and gas industry's total debt [6]. Decommissioning projects are not only expensive but also complicated and extremely time-consuming. Making a decision of dismantling, completely removing or leaving in place the facility involves several stakeholders, who pull their preferred decommissioning option in multiple different directions [9]. The process of engaging stakeholders is labour intensive and takes years to complete [19]. An example of a Shell's Brent field decommissioning programme lasted ten years before sufficient information was gathered from stakeholders to generate the consultation draft [1].

Current regulatory frameworks in many jurisdictions require those who are undertaking oil and gas decommissioning activities to extensively engage stakeholders [33]. Fig. 1 shows this large quantum of oil and gas decommissioning stakeholders discussed in the literature [14, 34]. Accordingly, in a decommissioning project, there will be a large number of stakeholders' preferences to take into consideration. To handle this, multi-criteria decision analysis (MCDA) tools are used to produce the decommissioning option. Well-known MCDA tools include Comparative Assessment (CA), Net Environmental Benefit Analysis (NEBA) and Best Practicable Environment Outcome (BPEO) [20, 27]. However, these existing tools are formula-based [23], where the user needs to empirically input the threshold and assign a score for each parameter. The output is prone to manipulation by the user as there is no clear prescriptive guidance for scoring and assessing. This may result in inaccurate decommissioning decisions. Such a problem can lead to devastating economic, environmental and societal consequences as evident by a past event of Brent Spar [11] and a recent one of Echo Yodel [26]. These affect the performance of oil and gas organisations and also the well-being of stakeholders.

Despite extensive research in the field of oil and gas decommissioning, very little progress on the use of machine learning as a decision support tool has been made [24]. Supervised machine

Figure 1: Oil and gas decommissioning stakeholders

learning has proved its usefulness in discovering patterns and providing accurate predictive outputs of complex problems [13, 36, 38]. To the best of our knowledge, only recent work of Martins et al. [24] has evidenced the employment of machine learning algorithms in the problem of oil and gas decommissioning. However, their experiment was based on a very small number of real-world samples, which could result in biased or inaccurate conclusions. This suggests the lack of sufficient training data in this field, which is often a requirement to produce well-performing models [10]. Such an issue may explain the limited number of research work in this problem domain.

In this paper, we present a new real-world dataset of oil and gas decommissioning activities. The dataset consists of 708 unique samples, each of which represents the decommissioning information of an offshore structure. Compared to the previous work [24], which considered 14 real-world samples, our dataset will contribute to the research community further. Moreover, we propose the use of machine learning techniques with our introduced dataset to develop a predictive model for oil and gas decommissioning. The learned features will not include manually assigned assessment scores used in MCDA tools, rather, structure information is considered.

The rest of this paper is structured as follows. Section 2 elaborates the discussion of relevant literature including existing MCDA tools and the use of machine learning in oil and gas decommissioning. In Section 3, the dataset is presented along with the description of the classification method and experimental setup. Results and discussion are provided in Section 4. Finally, Chapter 5 concludes the work and suggests potential future directions.

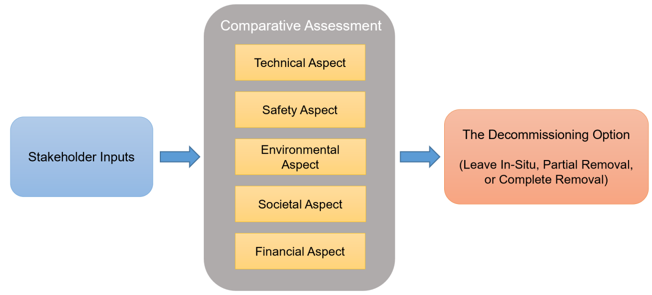

Figure 2: The comparative assessment process

2 DECOMMISSIONING STAKEHOLDERS ANALYSIS

While the process of decommissioning an oil and gas appears to be simple from a theoretical point of view, amongst other considerations, there is wide range of different stakeholders that have significant interests in decommissioning decisions [32, 34]. As shown in Fig. 1, several groups of stakeholders are involved [14, 34]. It is challenging to manage these different stakeholder groups as they have different interests and concerns. Their preferred decommissioning options are often pulled in multiple directions [32, 34]. Commercial fishermen, for example, prefer oil and gas facilities to be removed in order to ensure safety in navigation [30]. On the other hand, some environmental organisations prefer them to be left in place to prevent damaging existing marine ecosystems [25].

To handle such complex requirements, MCDA tools are used. Quantitative and qualitative data from stakeholders are input into an MCDA tool to determine the decommissioning option [23]. Different choices of tools are preferred in different regions. For example, in UK, the Offshore Petroleum Regulator for Environment and Decommissioning (OPRED) recommends oil and gas operators to use CA [4]. Meanwhile, BPEO is a mandatory requirement for regulatory approval in Thailand [8]. Since UK is among the world's most mature oil and gas landscapes [21], it can be said that CA is one of the most commonly-used MCDA tools for determining the decommissioning option of oil and gas infrastructures [15, 33].

Fig. 2 illustrates the five aspects considered in CA. These are technical, safety, environmental, societal and financial aspects [28]. The assessment of these criteria, however, heavily relies on the appraisers' judgement [24, 28, 33]. The users empirically input scores of different attributes without clear prescriptive guidance [18]. Thus, the resulting decommissioning option can be subjective. Another main issue of using MCDA tools is an extensive amount of resources and time spent to collect sufficient data from stakeholders. Employing machine learning techniques in this problem domain has shown to help scale down these issues [24].

While the application of machine learning has been explored extensively in various aspects of oil and gas exploration and production [16, 17], its application in decommissioning has only been recently seen though the work of Martin et al. [24]. The authors proposed to use machine learning techniques to reduce the problem's dimension and learn the patterns to predict the decommissioning option. The finding of this work suggests that using machine learning is a viable approach in generating the decommissioning decision. However, several limitations need to be addressed. Firstly, their experiments were based on 14 real-world samples. The use of small sample size is known to cause errors and overfitting in classification tasks [35]. Although they applied bootstrapping on the 14 samples to obtain 1,313 synthetic samples, a large generalization error in classification is likely to occur [22]. Secondly, the attributes used in model training also included the assessment scores of the five aspects in CA. This practice does not help reducing the resources and time in collecting stakeholders' preferences. In addition, the problem of assigning assessment scores has not been tackled.

3 METHODOLOGY

This section presents our novel dataset of oil and gas decommissioning activities. First, a brief description of how the dataset was collected is given. This is followed by a detailed description and characteristics of the dataset. Then, a brief discussion related to the classification method used to benchmark this dataset is provided.

3.1 Data Collection and Extraction

We collected data on offshore decommissioning activities from 111 oil and gas decommissioning programme reports. These reports were published on the government official website of the UK's OPRED [4]. Only the approved decommissioning programmes were considered as they had been authorised by the experts. Each report was examined and analysed by content analysis, a well-proven qualitative analysis research technique [31], to extract relevant information regarding (i) the technical specifications of the offshore pipeline and (ii) the final decommissioning decision. The approvals of these programmes were granted between 2000 and 2020. These include the reports of over 30 well-known oil and gas companies such as Shell, Repsol Sinopec, Chrysaor, TAQA, Canadian Natural Resources Company, Spirit Energy, ConocoPhillips, etc. The reason for conducting this study in the UK landscape is that its oil and gas decommissioning is much more mature compared to other landscapes around the world [21]. As such, there is more data available, allowing creation of more accurate machine learning models.

Typically, different criteria for determining the decommissioning options are used for different types of offshore structures[2]. In this work, we are interested in pipelines, which is the most common type of structures [33] and hence has the most samples available [4]. There are a total of 708 pipelines structures addressed in these 111 reports. Their decommissioning decision analyses were based on CA. Interestingly, each of the five aspects described in Section 2 of all 708 pipeline structures has the same analysed value. That is, *Full Removal* for the technical, environmental and societal aspects; *Partial Removal* for the safety aspect; and *Leave In-situ* for the financial aspect. This is because the analysis of the five attributes is highly dependent on the type of structures [33]. In the view of the project assessors, the final proposed decommissioning options of pipelines structures with identical values in the five attributes can be different. Similarly, in the machine learning point of view, the five attributes, therefore, do not contribute to the final decision, and hence we have dropped them.

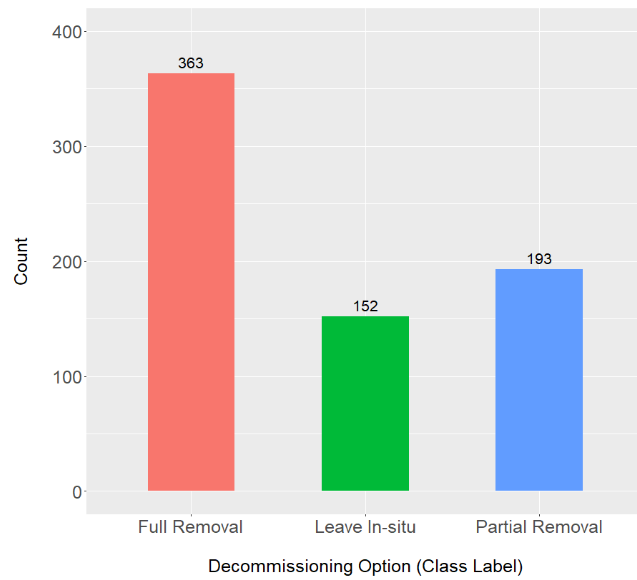

Figure 3: Class distribution in the dataset

Based on the review of decommissioning guidelines offered by various industrial representative bodies around the world [2, 3, 5], the decommissioning option was found to be influenced by structure's size, weight, material (e.g. steel, concrete or plastic), residues (e.g. hydrocarbons, chemicals), and burial status (e.g. surface laid and buried). Therefore, we selected the following features for the classification task: diameter, length, material, residue, burial status.

3.2 Dataset

The dataset is comprised of 708 unique samples with 10 attributes and the class label. Each sample represents the technical specifications of an offshore pipeline structure (attributes) and its decommissioning option (class label). The distribution of the classes, namely *Full Removal*, *Leave In-situ* and *Partial Removal* is displayed in Fig. 3. Table 1 shows some samples of the dataset, where the description of each attribute is given in Table 2.

As can be seen in Table 2, we have transformed some categorical variables, e.g. structure's materials, type of residual fluid, into binary variables. This is because some machine learning algorithms can learn only numerical data and not categorical data [29]. To avoid biases in recognising different values of a categorical variable as different numbers, an encoding technique was applied. The one-hot encoding technique, one of the most widely used encoding schemes [29], was employed for this purpose. The technique transforms a categorical variable with n distinct values into n binary variables, each of which indicates the presence or absence of the new binary variable. There are 12 missing values in *Diameter* and 2 in *Length*. We have made this dataset publicly accessible for benchmarking purposes (See GitHub). Please kindly cite this paper when making use of the dataset.

Table 1: A selection of samples from the dataset

Structure	Diameter	Length	Metal	Plastic	Concrete	Chemicals	Hydrocarbon	Bury	DecommissioningOption
PL23837	0.5	0.1	0	1	0	1	0	0	PartialRemoval
PL1057	12	19.8	1	0	1	0	1	1	LeaveInSitu
PL1059.1-2	8	0.1	1	0	0	0	1	0	LeaveInSitu
PL1099	4	15.1	1	1	0	0	0	1	PartialRemoval
PL111	8	5.28	1	0	1	1	0	1	LeaveInSitu
PL112A	6	1.55	1	1	0	0	1	1	PartialRemoval
PL115	16	19.1	1	0	1	0	1	0	PartialRemoval
PL126A	12.75	0.08	1	1	0	0	1	0	FullRemoval

Table 2: Attribute descriptions

Feature	Description
Structure	Name of the structure
Diameter	Diameter of pipeline (inch)
Length	Length of pipeline (km)
Metal	The structure was made of metal: 0 = no, 1 = yes
Plastic	The structure was made of plastic: 0 = no, 1 = yes
Concrete	The structure was made of concrete: 0 = no, 1 = yes
Chemicals	The residual fluid (if any) was chemicals: 0 = no, 1 = yes
Hydrocarbon	The residual fluid (if any) was hydrocarbons: 0 = no, 1 = yes
Bury	Burial status of the structure: 0 = surface, 1 = buried
DecommissioningOption	Outcome of the CA process: full removal, partial removal, leave in-situ

4 CLASSIFICATION METHOD AND SETUP

To demonstrate the use of machine learning on the novel dataset to build a predictive model, Random Forest (RF) was chosen. It is important that a proper and well-performed baseline for future studies on this oil and gas decommissioning dataset is provided. This justifies the choice of RF, which is among the top performing traditional learning algorithms that often provide promising classification results [12] and is accessible to any researchers, for the baseline model. To serve such a purpose, the default parameter settings of RF in *caret* package in *R* was used. This includes the setting of the number of trees (mtree) to 500. The number of features determined at each split, mtry, was automatically tuned as we applied 10-fold cross-validation during model training. This allows plenty of rooms for improvement on the results for further research.

The dataset was partitioned into training and testing sets at the ratio of 70 to 30. The training set was then split into 90% and 10% for training and validation purposes. The testing set was only used for model evaluation and result report in Section 5. Numerical variables, *Diameter* and *Length* in particular, were scaled using standard scores. Missing values in the two variables were handled using mean imputation.

5 EXPERIMENTAL RESULTS AND DISCUSSION

Our classification of the decommissioning option achieved the overall accuracy of 77.88%. Table 3 shows the classification results. Table 4 provides more detailed performance of the predictive model across different classes in terms of sensitivity, specificity, precision, F1-score, and balanced accuracy.

It can be seen in Table 4 that the dominating class in the dataset, which is *Full removal*, had the highest class accuracy of 85.98%. The prediction accuracy in *Leave in-situ* and *Partial removal* classes were 71.11% and 67.86%, respectively. Considering the distribution ratio of the three classes, which are 363:152:193 for *Full removal:Leave in-situ:Partial removal*, the difference in the class accuracies could be attributed to the unequal class distribution. This is because traditional learning algorithms are generally designed to maximise the overall accuracy [37]. Accordingly, the accuracy of the dominating class will likely to be higher than those of minority classes. This issue is recognised in the literature as the problem of class imbalance [37]. Similarly, the performance of our model in other metrics as shown in Table 4 could be explained by the imbalance situation. The results in precision and F1-score were in the same direction as sensitivity. Although the results in specificity were in the opposite direction of the class distribution, this is not surprising as the metric considers the prediction results on instances of the other classes.

Despite no optimisation or parameter tuning, we have shown promising classification results using machine leaning practices. Moreover, it was prove that favourable results can yet be obtained even with the exclusion of some features used in the traditional analysis process. Specifically, we found that assessment scores from the CA process, which are one of the key factors to be considered in the traditional approach, do not at all contribute to the classification

Table 3: Classification results

		Actual Class		
		Full Removal	Leave In-situ	Partial Removal
Predicted	Full Removal	92	8	8
	Leave In-situ	2	32	10
	Partial Removal	13	5	38

Table 4: Performance across the decommissioning options

Decommissioning Option	Sensitivity	Specificity	Precision	F1-score	Balanced Accuracy
Full Removal	85.98	84.16	85.19	85.58	85.07
Leave In-situ	71.11	92.64	72.73	71.91	81.87
Partial Removal	67.86	88.16	67.86	67.86	78.01

since the given scores in the five aspects are dependent on the type of structures.

Furthermore, these results are comparable to the results reported by Martin et al. [24]. However, our results were based on 708 real data of oil and gas decommissioning activities as compared to 14 real sample in their work. This is highly likely to make our decision support model more generalised to real-world problems. Another main different of our approach to their approach is that we do not consider any assessment scores from the CA process, which usually takes years to complete. Thus, the time and cost in producing the decommissioning option can be significantly reduced.

6 CONCLUSIONS

In this paper, we presented a new oil and gas decommissioning option dataset to the machine learning research community. The dataset was collected and extracted from over a hundred reports of UK oil and gas companies. It contains information of 708 pipeline offshore infrastructures including the approved decommissioning decisions. Moreover, we demonstrated the use of this dataset to build a classification model for suggesting the decommissioning option. The model achieved 77.88% overall accuracy with the highest per-class accuracy of 85.98%. These results were comparable to the previous work reported. However, our model was trained based on a significantly higher number of real samples. This allows better generalisation of the model in the real-world oil and gas decommissioning problem. Another advantage of our method is that we have excluded the assessment scores from the CA process in developing the decision support tool. By doing so, human biases in assigning assessment scores are eliminated. Also, the cost and time required in the comparative assessment or any other MCDA processes will be greatly reduced. Despite the growing interests in oil and gas decommissioning worldwide, there is a lack of research and data in this topic in the field of machine learning. No real-world dataset with sufficient samples for machine learning tasks has been available before. We see the importance of making our dataset publicly available and demonstrating a machine learning-driven approach for oil and gas decommissioning. This serves as a baseline to the research community and motivate further experiments to broaden the knowledge and research in this field. Potential future direction will include improving the performance of the predictive model by addressing the class imbalance problem. Emerging methods such as evolutionary algorithms and GANs may be utilised as resampling techniques to tackle the imbalanced distribution of data. Another interesting future direction is to expand the work by also considering other types of oil and gas infrastructures. This will enable the development of machine learning-based decision support tools to be more generalised across the problem domain.

REFERENCES

[1] 2017. Brent Decommissioning Stakeholder Engagement Report: A Supporting Document to the Brent Field Decommissioning Programmes. *Shell UK, London* (2017).
[2] 2019 (Retrieved February 1, 2021). Guidelines for Comparative Assessment in Decommissioning Programmes - Issue 1. https://oilandgasuk.co.uk/product/guidelines-for-comparative-assessment-in-decomm-programmes/
[3] (Retrieved February 1, 2021). Decommissioning of Offshore Concrete Gravity Based Structures (CGBS) in the OSPAR Maritime Area/ Other Global Regions. https://www.iogp.org/bookstore/product/decommissioning-of-offshore-concrete-gravity-based-structures-cgbs-in-the-ospar-maritime-area/
[4] (Retrieved February 1, 2021). Oil and gas: decommissioning of offshore installations and pipelines. https://www.gov.uk/guidance/oil-and-gas-decommissioning-of-offshore-installations-and-pipelines
[5] (Retrieved February 1, 2021). Options for decommissioning subsea bundles. https://www.iogp.org/bookstore/product/options-for-decommissioning-subsea-bundles/
[6] Hafez Abdo, Musa Mangena, Graham Needham, and David Hunt. 2018. Disclosure of provisions for decommissioning costs in annual reports of oil and gas companies: A content analysis and stakeholder views. In *Accounting forum*, Vol. 42. Elsevier, 341–358.
[7] Francois Ars, Renato Rios, et al. 2017. Decommissioning: A call for a new approach. In *Offshore Technology Conference*. Offshore Technology Conference.
[8] D Beckstead. 2018. Lessons from Thailand on the Importance of Devising and Implementing Detailed Decommissioning Regimes. *Oil, Gas & Energy Law Journal (OGEL)* 16, 2 (2018).
[9] John Chandler, David White, Erika J Techera, Susan Gourvenec, and Scott Draper. 2017. Engineering and legal considerations for decommissioning of offshore oil and gas infrastructure in Australia. *Ocean Engineering* 131 (2017), 338–347.
[10] Guiguang Ding, Yuchen Guo, Kai Chen, Chaoqun Chu, Jungong Han, and Qionghai Dai. 2019. DECODE: Deep confidence network for robust image classification. *IEEE Transactions on Image Processing* 28, 8 (2019), 3752–3765.
[11] John Elkington and Alex Trisoglio. 1996. Developing realistic scenarios for the environment: Lessons from Brent Spar. *Long Range Planning* 29, 6 (1996), 762–769.
[12] Eyad Elyan and Mohamed Medhat Gaber. 2017. A genetic algorithm approach to optimising random forests applied to class engineered data. *Information sciences* 384 (2017), 220–234.
[13] Eyad Elyan, Carlos Francisco Moreno-García, and Pamela Johnston. 2020. Symbols in Engineering Drawings (SiED): An Imbalanced Dataset Benchmarked by Convolutional Neural Networks. In *International Conference on Engineering Applications of Neural Networks*. Springer, 215–224.
[14] Sabrina Genter et al. 2019. Stakeholder Engagement in the Decommissioning Process. In *SPE Symposium: Decommissioning and Abandonment*. Society of Petroleum Engineers.

[15] Paul L Gordon, Edward H Poot, and Grace S O'Connor. 2019. Rigorous decommissioning decision making with strong stakeholder engagement using comparative assessment. *The APPEA Journal* 59, 2 (2019), 591–595.
[16] Yasin Hajizadeh. 2019. Machine learning in oil and gas; a SWOT analysis approach. *Journal of Petroleum Science and Engineering* 176 (2019), 661–663.
[17] Khadijah M Hanga and Yevgeniya Kovalchuk. 2019. Machine learning and multi-agent systems in oil and gas industry applications: A survey. *Computer Science Review* 34 (2019), 100191.
[18] Max Henrion, Brock Bernstein, and Surya Swamy. 2015. A multi-attribute decision analysis for decommissioning offshore oil and gas platforms. *Integrated environmental assessment and management* 11, 4 (2015), 594–609.
[19] Kayleigh Hughes, Sylvia Jagerroos, et al. 2019. Must Haves for a Credible Comparative Assessment. In *SPE Symposium: Decommissioning and Abandonment*. Society of Petroleum Engineers.
[20] Sylvia Jagerroos, Hughes Kayleigh, et al. 2019. Emerging Decommissioning Trends in South East Asia:-Local Interpretation and Implementation of Recently Updated Legislative Framework and Guidelines.. In *SPE Symposium: Decommissioning and Abandonment*. Society of Petroleum Engineers.
[21] Jia Li, Yun Peng, Ming Zhu, Kai Wang, and Jiexin Yi. 2019. Decommission in Petroleum Industry: Current Status, Future Trends and Policy Advices. In *IOP Conference Series: Earth and Environmental Science*, Vol. 237. IOP Publishing, 042013.
[22] Xiaoxu Li, Dongliang Chang, Zhanyu Ma, Zheng-Hua Tan, Jing-Hao Xue, Jie Cao, and Jun Guo. 2020. Deep InterBoost Networks for Small-sample Image Classification. *Neurocomputing* (2020).
[23] ID Martins, FF Moraes, G Távora, HLF Soares, CE Infante, EF Arruda, L Bahiense, J Caprace, and MI Lourenço. 2020. A review of the multicriteria decision analysis applied to oil and gas decommissioning problems. *Ocean & Coastal Management* 184 (2020), 105000.
[24] Isabelle D Martins, Laura Bahiense, Carlos ED Infante, and Edilson F Arruda. 2020. Dimensionality reduction for multi-criteria problems: An application to the decommissioning of oil and gas installations. *Expert Systems with Applications* 148 (2020), 113236.
[25] DL McLean, T Bond, JC Partridge, S Rouse, and M Love. 2020. The impact of oil and gas infrastructure in marine ecosystems: a global vision for informing decommissioning decisions. *The APPEA Journal* 60, 2 (2020), 476–479.
[26] Peter Milne. 2020 (Retrieved February 1, 2021). Woodside abandons abandonment for Echo Yodel. https://www.boilingcold.com.au/woodside-shouts-abandon-abandonment-for-echo-yodel/
[27] Patrick E O'Connor, Brian R Corr, Steve Palmer, Robert C Byrd, et al. 2004. Comparative assessment of decommissioning applications of typical north sea and gulf of mexico approaches to several categories of offshore platforms in the middle east. In *The Fourteenth International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers.
[28] David Palandro, Azivy Aziz, et al. 2018. Overview of Decommissioning Option Assessment: A Case for Comparative Assessment. In *SPE Symposium: Decommissioning and Abandonment*. Society of Petroleum Engineers.
[29] Kedar Potdar, Taher S Pardawala, and Chinmay D Pai. 2017. A comparative study of categorical variable encoding techniques for neural network classifiers. *International journal of computer applications* 175, 4 (2017), 7–9.
[30] Sally Rouse, Andronikos Kafas, Rui Catarino, and Hayes Peter. 2018. Commercial fisheries interactions with oil and gas pipelines in the North Sea: considerations for decommissioning. *ICES Journal of Marine Science* 75, 1 (2018), 279–286.
[31] Mark Saunders, Philip Lewis, and Adrian Thornhill. 2019 (Retrieved February 1, 2021). Research methods for business students.
[32] JL Shaw, P Seares, and SJ Newman. 2018. Decommissioning Offshore Infrastructure: a Review of Stakeholder Views and Science Priorities. *WAMSI: Perth, WA, Australia* (2018).
[33] Brigitte Sommer, Ashley M Fowler, Peter I Macreadie, David A Palandro, Azivy C Aziz, and David J Booth. 2019. Decommissioning of offshore oil and gas structures–Environmental opportunities and challenges. *Science of the total environment* 658 (2019), 973–981.
[34] Aaron Tung et al. 2020. A Comparison of Stakeholder Engagement Strategies for Offshore Decommissioning Projects in the United Kingdom and Australian Landscape. In *Offshore Technology Conference*. Offshore Technology Conference.
[35] Gaël Varoquaux. 2018. Cross-validation failure: small sample sizes lead to large error bars. *Neuroimage* 180 (2018), 68–77.
[36] Pattaramon Vuttipittayamongkol and Eyad Elyan. 2020. Improved Overlap-based Undersampling for Imbalanced Dataset Classification with Application to Epilepsy and Parkinson's Disease. *International journal of neural systems* 30, 08 (2020), 2050043.
[37] Pattaramon Vuttipittayamongkol and Eyad Elyan. 2020. Neighbourhood-based undersampling approach for handling imbalanced and overlapped data. *Information Sciences* 509 (2020), 47–70.
[38] Pattaramon Vuttipittayamongkol and Eyad Elyan. 2020. Overlap-Based Undersampling Method for Classification of Imbalanced Medical Datasets. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 358–369.