IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A SYSTEMATIC REVIEW ON EXPLORATIONS IN ARTITIFICAL INTELLIGENCE AND MACHINE LEARNING FOR VARIOUS FIELD OF APPLICATIONS

[1]Dr. L. Chandra Sekhar Reddy, [2]Gundu Rithindatta

Asst. Prof. in Computer Science Engineering (Data Science) Department, CMR College of Engineering & Technology, KANDLAKOYA(v), Medchal, Hyderabad, India

Student, Computer Science Engineering Department, CMR College of Engineering & Technology, KANDLAKOYA(v), Medchal, Hyderabad, India

ABSTRACT

Machine learning (ML) and other forms of artificial intelligence (AI) are finding more and more uses in production environments. The study provides an in-depth analysis of the current state of machine learning (ML) applications in manufacturing settings. Technology advancements in smart mining have made real-time data production, gathering, and exchange possible. For this reason, the mining sector is actively engaged in ML-based research that makes use of this data. In addition, this research provided a comprehensive review of the uses of data sciences and ML in several areas of petroleum engineering and geosciences, including petroleum exploration, reservoir characterization, oil well drilling, production, and well stimulation, with special attention paid to the rapidly expanding field of unconventional reservoirs. A look into the future of data science and ML in the oil and gas business is provided, along with an examination of the features of ML that are needed to improve prediction. Different ML methods with application to the oil and gas sector are compared in detail in this paper. This article also discussed how the advent of AI and ML has sparked a paradigm change in healthcare, opening up new avenues of analysis and prediction within the realm of medical data. The application of AI and ML to combat COVID-19 has recently been demonstrated in a number of studies. The purpose of this article is to provide reviewers with a review of recent research that has used AI and ML in various applications.

Key words: Artificial Intelligence, Machine Learning, Oil Industry applications, Health care related applications, Mining based applications.

1. INTRODUCTION

Several studies have looked back at the development of AI, including machine learning and data mining; for example, HARDING ET AL. provides an overview of AI applications from 1987 to 2005[1]. The growing trend of digitalization is only fueling the unstoppable rise of artificial intelligence and machine learning. Companies are seeing great promise in data-driven strategies, as seen by the large cash prizes being offered in recent online competitions (e.g., Kaggle.com). According to GOODFELLOW, AI is most effectively put to use when it's applied to situations that can't be reduced to code but can be solved by intuition alone [2].

This article shows how the study and implementation of machine learning (ML), a branch of AI, has recently shifted from the academy to the commercial world. Due to the rapid digitalization of industry, more data is being generated, and as a result, there are more datasets available for an ML application to employ in order to learn from the past. Convolutional neural networks (CNNs) and other deep neural network architectures (NNs) are examples of deep learning (DL), a subfield of machine learning and artificial intelligence, that has been used to solve problems like identifying pictures and finding objects (e.g. [3]). Starting in 2015, there will be a shift from research-only solutions to solutions that can be used in the real world. This article gives a thorough look at several fields and gives repeated examples of how they can be used.

Subfields of ML include supervised learning, unsupervised learning, and reinforcement learning. The article's conclusions are summarised in Fig. 1. Most machine learning techniques now rely on some form of supervision, as seen by the prevalence of supervised approaches. What's more, Fig. Fig. 1 demonstrates how NN have been utilised to address industrial domain problems in recent years across a range of architectural styles.

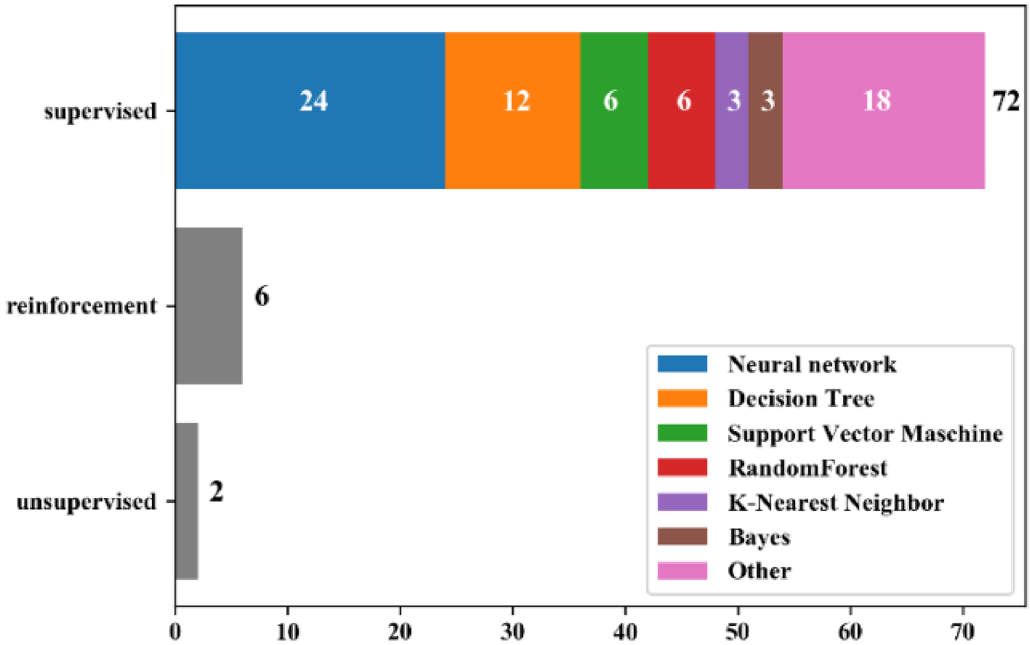

Fig. 1: Application of structured review techniques

Since humans are rational beings, they take in knowledge from a variety of sources, including formal schooling and self-reflection. Yet, in contrast to people, computers can learn through a series of preprogrammed instructions called an algorithm. This method is known as "machine learning" (ML). In order to assist computers better understand and take advantage of their surroundings, researchers in the field of artificial intelligence developed a technique called machine learning (ML), which mimics the learning processes of humans using a series of algorithms. One definition of ML is, "A computer programme is said to learn from experience E relating to given class of tasks T and performance measure P, if its performance at tasks T, as measured by P, improves with experience E." Although ML's foundational ideas date back to the 1950s, it wasn't until the 1990s that the discipline was recognised as a separate academic discipline. The domains of computer science [4], healthcare [5], the environment, medicine, energy, and services, and so on all make use of ML algorithms.

2. REVIEW OF LITERATURE

2.1 ML Methods for manufacturing processes – identifying artificial intelligence (ai) methods for field application

The SLR follows the framework of posing a research question, identifying the traits to be emphasised, narrowing the focus to the appropriate literature, and then drawing conclusions and synthesising the findings. These recommendations were adapted from KITCHENHAM ET AL. [6]. To begin, here is the formal research question that guided the structured review that was carried out: Finding the manufacturing industry's use of machine learning applications during the past five years. Additional criteria for primary attention in a typical factory or especially a learning factory environment include manufacturing applications such as manufacturing process planning, quality control, predictive maintenance, logistics, robotics, assistance and learning systems, ML-training concepts in learning factories, and process control and optimization. An overview of each work is provided, along with a mention of the algorithm that was implemented or used to address the issue raised in the study.

The next step in the structured review is to do a keyword search using the previously indicated subtopics, as well as terms from the "manufacturing" and "factory" domains. The literature on AI techniques in practise, particularly those used in ML applications, was thoroughly reviewed. Tabulated below is a summary and analysis of the remaining articles' use of ML applications.

Table 1. Overview of applications and algorithms.

Subtopic	**Application**	**Algorithm**	**Literature**
Manufacturing process planning	Scheduling	Q-learning, RF, decision tree	**[7]**
	Cost & energy prediction	NN, SVR, GBT, RF, others	**[8]**
	System modeling	log. Regr., RF, decision tree, bayes	**[9]**
Quality control	Quality cost reduction	decision tree, NN, SVM, others	**[10]**
Predictive maintenance	remaining useful life &	Decision tree, NN, PCA, KNN, others	**[11]**
Logistics	Scheduling	NN, Q-learning, deep q-learning, RF	**[12]**
Robotics	Human robot collaboration	hidden markov model, KNN, clustering, NN	**[13]**
Assistance and learning systems	assembly assistance	NN	[14]
AI-training concepts in learning factories	ObjectRecognition	NN	[15]
Process control & optimization	Production line	GBT	[16]

3. A MACHINE LEARNING ILLUSTRATION FOR MEDICAL

More over 1.6 million people have lost their lives as of December 22, 2020 due to the SARS-COV-2 pandemic, which has also put a significant burden on the global economy and healthcare systems. A global death toll of around 6,000 per day, with no cure in sight and the chance of novel virus emergence, might kill around 2.2 million people annually. Global mortality and prevalence curves have not decreased [17] despite continued attempts at prevention and social isolation. More focus on clinical therapy in the earliest stages could help reduce fatality rates. A patient in critical condition requires immediate transfer to the intensive care unit (ICU) and insertion of a ventilator. About 54% of critically ill patients in China did not have fast access to the Intensive Care Unit [18], and 30% of those who did not survive did not get prompt mechanical ventilation. In cases where there are numerous patients, medical staff is overworked, and not enough resources to treat each patient, rapid identification of those at high mortality risk becomes critical.

When COVID-19 patients are admitted to the hospital, it is generally too late for doctors to give an accurate diagnosis. To add insult to injury, the course of COVID-19 can take unexpected twists, in which a previously stable patient's status abruptly deteriorates to a critical state [19]; this could surprise even the most seasoned doctors off guard. As AI models are able to spot intricate patterns in massive datasets, they may prove to be invaluable helpers in clinical prediction, a task at which the human brain excels not. From large-scale epidemiological modelling to fine-grained individual diagnosis and prognosis prediction, AI techniques have been used in the fight against COVID-19 [20]. Although a number of prognostic models for COVID-19 have been presented [21], the predictive ability of non-invasive and invasive characteristics has not been systematically examined.

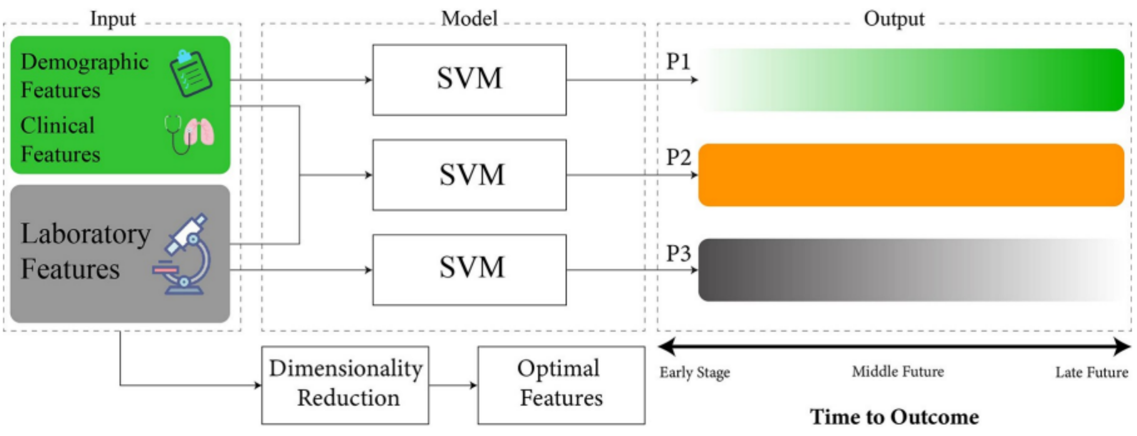

Fig 2. Illustration of the modeling framework.

This study aimed to develop a mortality prediction model using routine clinical data from the first day of admission, investigate the feasibility of predicting COVID-19 mortality outcome using non-invasive patient features, and provide a direct comparison of the mortality prediction powers of non-invasive and invasive features. The invasiveness of laboratory testing and other distinguishing features were used to classify patients into distinct groups. A total of three machine learning models were developed, two using each feature group individually and one using them all together, to evaluate and contrast the projected efficacy of the aforementioned features (Fig 2). Numerous COVID-19 patients have been reported to have experienced an exacerbation episode during the hours of 24 and 48 following hospital admission [22]. As a result, we built our model using information collected on the first day of patients' stays in order to give a practical resource.

Using the feature weights from the SVM method, features of low relevance are progressively removed in the Support Vector Machine-recursive feature elimination (SVM-RFE) feature assessment approach [23]. Forty iterations of specialised MATLAB code were utilised to develop SVM-RFE for this study. Using the fitcsvm function and the KFold as input, we trained a linear support vector machine without optimising the training data over a period of 10 iterations. To evaluate performance, we used the kfoldPredict method to create predictions based on the trained algorithm. Predictions are made for each in-fold case using the algorithm trained on all out-fold instances, and a combined label of predictions for up to 10 folds is provided by the method. Retraining the algorithm with fewer features and a new partitioning method followed by the elimination of the feature with the lowest absolute weight from the input dataset. All but one feature from the original dataset was removed using the methods described above. The accuracy was averaged across iterations to determine a mean and standard deviation.

3.1 Sparsity analysis

Consider using a support vector machine. Twenty-five constant, logarithmically-spaced lambda values were used in each iteration. Each lambda was trained with a non-optimized linear SVM model using the fitclinear function with all data instances as input and 10 folds (KFold input of the fitclinear was set to 10). Lattice was selected as the Regularization input. Using the trained algorithm's kfoldLoss technique, which provides the loss as a mean across all folds, we were able to acquire a classification loss. The number of features for each lambda that had not had their SVM weights set to zero was counted as the total number of features. To assess the importance of feature subsets in a computationally efficient manner, researchers have turned to sparsity analysis using linear SVM (i.e., Sparse Linear SVM) [23]. Together, the least absolute shrinkage and selection operator (Lasso)—a sparse regularisation framework—and liner SVM were used to assess the predictive information richness of invasive and non-invasive variables with respect to the result. To implement 100 iterations of Sparse Linear in MATLAB, a new programme was created.

4. EXPLORATION, EXPLOITATION OF MACHINE LEARNING APPLICATIONS IN MINING:

The mining sector is beginning to examine adopting "smart mining" practises. The term "smart mining technology" [24] refers to the utilisation of cutting-edge ICTs such the Internet of Things (IoT), big data, mobile, AI, augmented reality, and virtual reality in the extraction of mineral resources. A massive amount of information is being generated, gathered, and disseminated in real time as a result of the evolution of smart mining technologies. Several factors have contributed to the recent surge in popularity of data science in mining. These include the emergence of novel data types (such as drilling data, sensor data, and measurement data), the maturation of AI methods, the enhancement of computers' processing power, and the advent of machine learning (ML).

Application of ML in the mining industry is the focus of a number of ongoing initiatives. In order to extract more gold from Canada's Red Lake mine, Goldcorp's geologists employed IBM's Watson AI Supercomputer to analyse exploration data and locate previously unknown amounts of the precious metal [25]. Heavy equipment manufacturers like Komatsu and NVIDIA have collaborated on projects to improve the mining industry's ability to track its employees and their machinery. New technologies allow for the detection of inefficient machinery and the maintenance of a secure working environment [26]. Pilot projects are being conducted by the Canadian artificial intelligence research institute IVADO and the underground mining safety and operations management services provider Newtrax to collect big data from sensors installed in mining equipment, analyse the data using ML, and predict when the equipment has failed and needs maintenance [27].

The use of ML in mineral processing, the application of soft computing technology in exploration, the latest developments in digitalization, and the automation of the mining sector are only some of the subjects discussed in [28]. The study covers mineral exploration, mining, and mine reclamation; nevertheless, a more thorough introduction of ML's applications in the mining industry would be helpful.

Structured data (such as tables of numbers) and unstructured data (such as images and papers) make up the bulk of what is commonly referred to as "big data" [29]. In other words, it's a data-intensive technology. By their very nature, ML algorithms get better as more data is fed into them. In this way, data plays an important role in ML studies. Data in abundance is useful and transferable to other settings like hospitals and universities. We looked at research with big data-related discussions or comments.

5. MACHINE LEARNING APPLICATIONS TO THE OIL AND GAS INDUSTRY

The most common problem addressed by ML petroleum researchers is how to generalise the results of ML models, which are typically exclusive to the data set evaluated. Overfitting, coincidence, excessive training, lack of interpretability of outcomes, and bias are only a few of the ML applications' widespread limits and problems that prevent the globalisation of the produced models. Furthermore, these models call for a substantial amount of data that is often lacking.

When there is no predetermined point at which training should end, overtraining can occur. It's possible that changing the model structure, including the weights, will continue the trend of decreasing error. The true danger then is that the model becomes overfit to a particular dataset, making further generalisation unfeasible. Early stopping is a training strategy that employs a control set to keep an eye on how things are going during the training process and stop when it looks like enough is enough. The training process will be terminated prematurely if the rate of error increases. To conserve time and energy, reinforcement learning with instream supervision is utilised. Generative adversarial networks, which take a look at the development of two rival networks to comprehend the model concept, are conceptually similar to this method [30].

Generalizability issues in existing AI models are a major hurdle to AI's mainstream adoption in the oil and gas industry. Many models struggle when applied to conditions outside of those used to create them. When training on new datasets, additional resources should be used even if the dataset is comparable to one previously trained on [31]. Further complicating matters is the fact that reusing ML models is quite difficult. In most cases, the performance of models trained on one geological field suffers when applied to another. When the given dataset's input parameters are within the range of the input parameters on which the model is to be built, it is strongly advised to do so [32].

Last but not least, bias has a real impact that must be taken into account, even if it is difficult to identify and counteract. A large number of scientists are working to address the problems of AI bias by deciphering the motivation behind the model and the outcomes it produces. To reduce the impact of biases, model-independent perturbations can be used by replacing the inputs with random values drawn from a normal distribution (Samek et al., 2018) [33]. All of the problems that can arise with AI and ML models are outlined in Table 2.

Table 2: A summary of the limitations of AI and ML models

Limitation	Reason	Solution	References
Overftting	The model's complexity can be increased to better fit a given dataset as long as the error rate continues to fall as a result of structural updates.	An approach to training known as "early stopping" A good example of in-stream reinforcement learning supervision is the generative adversarial networks.	Hossain (2018) [30]
Data availability	Data collection can have its limitations.	One-time training and improvement with additional data is a type of machine learning known as "single-shot learning."	Weyrauch and Herstatt (2016) [34]
Interpretability	Results are affected not by individual model connections but by the whole of all model linkages.	A locally interpretable model, and the agnostic justifications behind it Using GAMs, or generalised additive models	Shabbir et al. (2018) [35]
Generalization	Failure of the model to predict results in conditions other than those used to create it.	New datasets require training, which will require the allocation of more resources.	Ramamoorthy and Yampolskiy (2018) [31]
Bias	Black-box models are susceptible to biases due to their inherent nature.	Using model-independent perturbations	Samek et al. (2018) [33]

CONCLUSION

The relevant SLR examined the state of the art during the past few years (2015-2022), with a particular emphasis on concrete examples and the specific nomenclature of applied machine learning algorithms in the manufacturing context of each piece of literature. Numerous publications make use of these technologies to create novel applications that improve upon existing industrial methods. With regards to the machine learning approach, supervised approaches are now considered state-of-the-art, while reinforcement learning approaches have received greater research attention over the past few years. Simply two references were found to unsupervised approaches, which may be because many papers only explained the implemented regression or classification job in detail, leaving out the actual data preparation and investigation. The authors believe that unsupervised approaches are used extensively throughout the data analysis process, even if they are not specifically specified. Researchers in the mining, oil, and gas industries, as well as those working in the medical field, have all done systematic assessments of the most up-to-date ML research to determine where the field is headed. The review serves as a summary of previous studies and a guide for additional studies.

REFERENCES

1. [1] Bajic, B., Cosic, I., Lazarevic, M., Sremcev, N., Rikalovic, A., 2018. Machine Learning Techniques for Smart Manufacturing: Applications and Challenges in Industry 4.0. 9th International Scientific and Expert Conference TEAM 2018.
2. [4] Leo Kumar, S.P., 2017. State of The Art-Intense Review on Artificial Intelligence Systems Application in Process Planning and Manufacturing. Engineering Applications of Artificial Intelligence 65, 294–329.
3. [11] Ma, L., Xie, W., Zhang, Y., 2019. Blister Defect Detection Based on Convolutional Neural Network for Polymer Lithium-Ion Battery. Applied Sciences 9 (6), 1085.
4. Handelman, G.S.; Kok, H.K.; Chandra, R.V.; Razavi, A.H.; Huang, S.; Brooks, M.; Lee, M.J.; Asadi, H. Peering Into the Black Box of Artificial Intelligence: Evaluation Metrics of Machine Learning Methods. Am. J. Roentgenol. 2019, 212, 38–43. [CrossRef] [PubMed]
5. Spasi´c, I.; Nenadic, G. Clinical Text Data in Machine Learning: Systematic Review. JMIR Med. Inform. 2020, 8, e17984. [CrossRef] [PubMed]
6. [13]Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S., 2009. Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology 51 (1), 7–15.
7. [14]Kim, D., Lee, T., Kim, S., Lee, B., Youn, H.Y., 2018. Adaptive Packet Scheduling in IoT Environment Based on Q-learning. Procedia Computer Science 141, 247–254
8. [19]Walther, J., Spanier, D., Panten, N., Abele, E., 2019. Very short-term load forecasting on factory level – A machine learning approach. Procedia CIRP 80, 705–710.
9. [18]Brik, B., Bettayeb, B., Sahnoun, M.'h., Duval, F., 2019. Towards Predicting System Disruption in Industry 4.0: Machine Learning-Based Approach. Procedia Computer Science 151, 667–674.
10. [12]Lavrik, E., Panasenko, I., Schmidt, H.R., 2019. Advanced Methods for the Optical Quality Assurance of Silicon Sensors. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 922, 336–344
11. [26]Pinto, R., Cerquitelli, T., 2019. Robot fault detection and remaining life estimation for predictive maintenance. Procedia Computer Science 151, 709–716.
12. [30]Kuhnle, A., Schäfer, L., Stricker, N., Lanza, G., 2019. Design, Implementation and Evaluation of Reinforcement Learning for an Adaptive Order Dispatching in Job Shop Manufacturing Systems. Procedia CIRP 81, 234–239
13. [35]Duckworth, P., Hogg, D.C., Cohn, A.G., 2019. Unsupervised human activity analysis for intelligent mobile robots. Artificial Intelligence 270, 67–92.
14. [48]Haslgrubler, M., Gollan, B., Ferscha, A., 2018. A Cognitive Assistance Framework for Supporting Human Workers in Industrial Tasks. IT Prof. 20 (5), 48–56
15. [50]Oberc, H., Fahle, S., Prinz, C., Kuhlenkötter, B., 2020. A Practical Training Approach in Learning Factories to Make Artificial Intelligence Tangible. CMS (Accepted for publication).
16. [8] BOSCH. Bosch Production Line Performance: Reduce manufacturing failures. https://www.kaggle.com/c/bosch-production-line-performance. [9] Daimler, 2017. Mercedes-Benz Greener Manufacturing
17. 1. WHO. Who coronavirus disease (covid-19) dashboard. URL https://covid19.who.int. Available at https://covid19.who.int. Accessed on 12.22.2020.
18. 2. Quah P., Li A. & Phua J. Mortality rates of patients with covid-19 in the intensive care unit: a systematic review of the emerging literature. Critical Care 24, 1–4

(2020). https://doi.org/10.1186/s13054-019- 2683-3 PMID: 31898531

19. Cascella M., Rajnik M., Cuomo A., Dulebohn S. C. & Di Napoli R. Features, evaluation and treatment coronavirus (covid-19). In Statpearls [internet] (StatPearls Publishing, 2020).
20. Alimadadi Ahmad, et al. "Artificial intelligence and machine learning to fight COVID-19." Physiol. Genomics, vol. 52, no. 4, pp. 200–202 (2020). https://doi.org/10.1152/physiolgenomics. 00029.2020 PMID: 32216577
21. Wynants L. et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. bmj 369 (2020). https://doi.org/10.1136/bmj.m1328 PMID: 32265220
22. Wu Z. & McGoogan J. M. Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the Chinese center for disease control and prevention. Jama 323, 1239–1242 (2020). https://doi.org/10.1001/jama.2020.2648 PMID: 32091533
23. Bi J., Bennett K., Embrechts M., Breneman C. & Song M. Dimensionality reduction via sparse support vector machines. Journal of Machine Learning Research 3, 1229–1243 (2003).
24. Choi, Y.; Lee, H.W. Trends in Mineral Resources Development Technology Using Artificial Intelligence. ITFIND 2020, 1935, 13–24.
25. MINING. Goldcorp partners with IBM to hunt for exploration targets at Red Lake. 2017. Available online: www.mining.com/ goldcorppartners-ibm-hunt-exploration-targets-red-lake/ (accessed on 15 February 2019)
26. Forbes. NVIDIA and Komatsu Partner on AI-Based Intelligent Equipment for Improved Safety and Efficiency. 2017. Available online: www.forbes.com/sites/tiriasresearch/2017/12/12/nvidia-and-komatsu-partner-on-ai-based-intelligent-equipment/ #63ad3365665b (accessed on 15 February 2019).
27. Mining Magazine. NEWTRAX. 2018. Available online: www.miningmagazine.com/partners/partner-content/1332132/thefuture-of-mining-isunderground (accessed on 15 February 2019).
28. Ali, D.; Frimpong, S. Artificial intelligence, machine learning and process automation: Existing knowledge frontier and way forward for mining sector. Artif. Intell. Rev. 2020, 53, 6025–6042.
29. 130. IBM. Analytics: The Real-World Use of Big Data. Available online: http://www.informationweek.com/pdf_whitepapers/ approved/1372892704_analytics_the_real_world_use_of_big_data.pdf (accessed on 10 June 2019).
30. Hossain M (2018) Frugal innovation: a review and research agenda. J Clean Prod 182:926–936
31. Ramamoorthy A, Yampolskiy R (2018) Beyond map?: the race for artifcial general Intelligence. ITU J 1(1):77–84
32. Mohaghegh SD (2017) Shale analytics: data-driven analytics in unconventional resources. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-48753-3
33. Samek W, Wiegand T, Muller KR (2018) Explainable artifcial intelligence: Understanding, visualizing and interpreting deep learning models. ITU J 1(1):39
34. Weyrauch T, Herstatt C (2016) What is frugal innovation? Three defning criteria. J Frugal Innov 2(1):1–17
35. Shabbir J, Anwer T (2018) Artifcial intelligence and its role in near future. J Latex Class Files 1(8):1–11
36. L. Chandra Sekhar Reddy, “A Review on Mobile App Ranking Review and Rating Fraud Detection in Big Data”, Springer Nature Singapore Pte Ltd. 2019.
37. Pradyumna Kumar Tripathy, “Federated learning algorithm based on matrix mapping for data privacy over edge

computing", International Journal of Pervasive Computing and Communications © Emerald Publishing Limited 1742-7371 DOI 10.1108/IJPCC-03-2022-0113

38. L. Chandra Sekhar Reddy, "YouTube: big data analytics using Hadoop and map reduce", International Journal of Engineering & Technology, 7 (3.29) (2018) 12-15 International Journal of Engineering & Technology Website: www.sciencepubco.com/index.php/IJET
39. L. Chandra Sekhar Reddy, "LOG DATA PROCESSING WITH MAPREDUCE", Volume 118 No. 14 2018, 229-233 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version) url: http://www.ijpam.eu Special Issue
40. Saroj Kumar, "Deep Learning based model for classification of COVID-19 images for healthcare research progress", Available on 18 may 2022.