Review of machine learning-based Mineral Resource estimation

by M.A. Mahboob[1,2], T. Celik[3], and B. Genc[1]

Affiliation:
[1]School of Mining Engineering, Faculty of Engineering and the Built Environment (FEBE). University of the Witwatersrand, Johannesburg, South Africa.
[2]Sibanye-Stillwater Digital Mining Laboratory (DigiMine), Wits Mining Institute (WMI), Faculty of Engineering and the Built Environment (FEBE), University of the Witwatersrand, Johannesburg, South Africa.
[3]School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg South Africa.

Correspondence to:
M.A. Mahboob

Email:
mahsan.mahboob@wits.ac.za
ahsan.igis@gmail.com

Dates:
Received: 11 Jun. 2020
Revised: 2 Aug. 2022
Accepted: 11 Aug. 2022
Published: November 2022

How to cite:
Mahboob, M.A., Celik, T., and Genc, B. 2022
Review of machine learning-based Mineral Resource estimation. Journal of the Southern African Institute of Mining and Metallurgy, vol. 122, no. 11, pp. 655–664

DOI ID:
http://dx.doi.org/10.17159/2411-9717/1250/2022

Synopsis

Mineral Resources estimation plays a crucial role in the profitability of the future of mining operations. The conventional geostatistical methods used for grade estimation require expertise, understanding and knowledge of the spatial statistics, resource modelling, geology, mining engineering as well as clean validated data to build accurate block models. However, the geostatistical models are sensitive to changes in data and would have to be rebuilt on newly acquired data with different characteristics, which has proved to be a time-consuming process. Machine learning methods have in recent years been proposed as an alternative to the geostatistical methods to alleviate the problems these might suffer from in Mineral Resource estimation. In this paper, a systematic literature review of machine learning methods used in Mineral Resource estimation is presented. This has been conducted on such studies published during the period 1990 to 2019. The types, performances, and capabilities, of several machine learning methods have been evaluated and compared against each other, and against the conventional geostatistical methods. The results, based on 31 research studies, show that the machine learning-based methods have outperformed the conventional grade estimation modelling methods. The review also shows there is active research on applying machine learning to grade estimation from exploration through to exploitation. Further improvements can be expected if advanced machine learning techniques are to be used.

Keywords

machine learning, artificial intelligence, Mineral Resources, grade estimation.

Introduction

Mineral Resources estimation (MRE) is one of the most important and critical stages in the mining value chain. The whole mining project depends on the reliable estimation of the grade of the mineralization. The spatial distribution of Mineral Resources (MR) depends on several known and unknown factors that cannot be incorporated in the traditional/conventional geostatistical models (Rossi and Deutsch, 2013; Hosseini, Asghari, and Emery, 2017). The basic assumption made by most of the mineral grade estimation models is that a spatial relationship exists between the grades at any two locations and that this relationship is a function of the distance between the two locations. Since the 1950s, many Mineral Resource Estimation (MRE) models have been proposed based on statistical methods, and later by incorporating the spatial dimensions in the estimations, which improved the results significantly. However, even the spatial estimation techniques are based on several assumptions and predict the MR values with some uncertainty levels. Many MRE models based on spatial statistics only incorporate the three-dimensional location of the measured value (X, Y, Z) along with the grade and thickness information. However, several other parameters like topographical variations, directions of geological structures, type of geology *etc.* are also very important for reliable MRE. These parameters are often neglected or not incorporated in the conventional statistical and spatial statistical methods. In the last 10 years, machine learning (ML)-based methods have become more popular in Resource estimation research. Several researchers (Samanta, Banopadhyay, and Ganduli, 2006; Chatterjee, 2010, Tahmasebi and Hezarkhani 2010, Zhang, 2017) have reported that ML-based methods have emerged as prediction models and as major alternatives to geostatistics for MRE.

Despite the huge number of research studies that used ML-based methods, inconsistent results have been reported regarding the accuracy of the methods, the comparison between ML and non-ML methods, and comparisons among several different ML-based methods. For example, in a comparison of ML and non-ML methods for MRE, Dutta, *et al.*, (2010) concluded that the ML methods produced more accurate results; however, Samanta, Ganguli, and Banopadhyay, (2005) showed that non-ML methods outperform the ML-based methods in producing reliable MREs. In a comparison among the ML-based methods,

Tahmasebi and Hezarkhani, (2010), showed that artificial neural networks perform better than regression models. Chatterjee and Bandopadhyay (2011) however, reported opposing results. Due to the discrepancies in the research studies which applied ML methods for MRE, practitioners in the fields of mining and geosciences may be hesitant to use ML models more practically and confidently. As opposed to other fields of study where the ML methods have been well tested and applied successfully, the applications of these methods in MRE face several challenges, like limited training data-sets, the uncertainty in existing data-sets and geological conditions, as well as the human factor. Although there is an increasing trend in academic research towards the applications of ML methods, most of the experts choose non-ML methods
for MRE.

To facilitate the applications of ML methods in the mining industry, it is very important to systematically summarize the empirical evidence on ML methods in current research and practice. The applications of ML methods are novel in the field of MRE, hence very little literature is available on the subject. There is no existing systematic literature review on the applications of ML methods in the field of mining related to MRE. In this paper, the literature review was performed on articles published from January 1 1990 to June 30 2019, related to ML methods applications in the field of MRE. The main purpose is to summarize the published work regarding the types of ML methods used in MRE, the comparisons between ML and non-ML methods, the performance evaluation of ML and non-ML methods, and the factors mainly considered in the application of ML methods.

Methodology

The systematic literature review methodology proposed by Kitchenham and Charters (2007) was used to conduct and report the review. The main steps include the definition of research questions, design of search strategy, selection criteria for studies, quality assessment, extraction of relevant data, and analysis, as given in Figure 1.

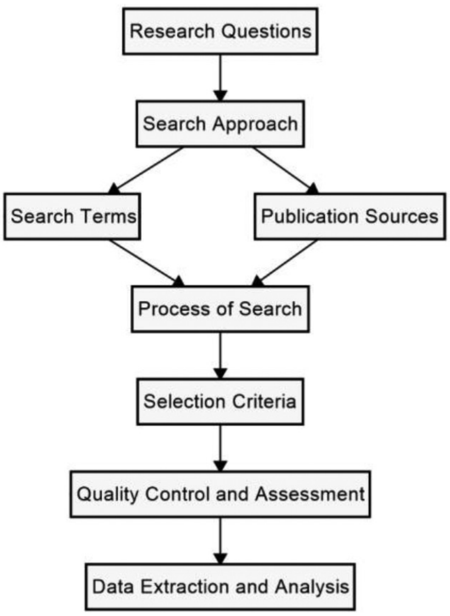

Figure 1—The methodology used for the systematic literature review (conceptualised from Wen, *et al*, 2012)

Research questions

Four research questions were defined based on the objective of summarizing the published work regarding the types of ML models used in MRE.

1. *Which ML methods/models have been used for MRE?*
   The aim was to identify the ML methods/models that have been used in MRE to provide MRE researchers and practitioners with a range of possible methodologies to consider.
2. *Does any publication on the comparison of ML against non-ML methods exist?*
   This question is concerned with the comparison of ML with non-ML methods in terms of accuracy, if performed in the studies.
3. *Do ML methods outperform non-ML methods?*
   The aim is to compare the accuracy of the ML methods against non-ML methods.
4. *Are there any ML methods that distinctly outperform other ML methods?*
   The comparison of different ML methods in order to identify these ML methods which perform better than others.

Search approach

The search strategy was based on the search terms, sources of publications, and process of search as explained below.

Search terms

The following steps were applied in order to search the terms for MRE, Wen, 2012:

(a) Selection of major keywords based on the research questions
(b) Identify possible different spellings and synonyms for major keywords
(c) Check the major keywords given in the relevant books and papers
(d) Usage of Boolean operator or to combine the different spellings and synonyms
(e) Usage of and operator to link the major keywords.

The following are the main keywords identified from the published work on machine learning models and techniques used for MRE.

> Mineral AND (*grade* OR *ore* OR *reserve* OR *resources*) AND (*estimation* OR *prediction*) AND (*machine learning* OR artificial *intelligence* OR *mining* OR *data mining*) AND (*geochemical* OR *exploration* OR *boreholes*) AND (*neural networks* OR *support vector machine* OR *support vector machine* OR *regression tree* OR *random forest* OR *Kriging* OR *nearest neighbour*) AND (*modelling* OR *spatial* OR *analysis*).

Although significant research has been done on the application of ML techniques in the oil and gas or petroleum industry, that was not considered in this literature review, being outside the scope of the paper.

Publication sources

Five reliable and widely searched electronic databases (IEEE Xplore, ScienceDirect, Springer Nature, Web of Science, and Google Scholar) were used to search the most relevant literature. All the other databases are largely covered by these five primary databases, and hence have also been used by many literature reviews studies in several fields. The main keywords developed previously were used to search for journal and conference papers in the five databases. Except for Google Scholar, the search was conducted on the title, abstract, and keywords. Google Scholar returned several millions of irrelevant records when terms were searched on full text, hence the search was limited to the titles of the publications. The searches were restricted to the time period from January 1 1990 to June 30 2019, as publications on the application of ML methods to MRE begin to appear in the early 1990s; for example, Wu and Zhou (1993) used neural networks for reserve estimation in 1993.

Search process

A comprehensive search of the relevant keywords among all the databases is very important. Therefore, the search process was designed and divided into two phases to identify relevant published papers.

Phase A: Search the five databases individually and list all the papers that resulted from the searches.

Phase B: The references in each resulting paper were scanned to identify additional relevant papers, which were added to the original. The inclusion and exclusion criteria set for the papers were as follows:

Inclusion criteria

1. ML methods used to pre-process the data
3. Applied more than one ML method and/or combined with non-ML methods
4. Comparative studies that compare different ML methods and/or with non-ML methods
5. Studies which contain both conference and journal publications; only journal papers were selected for inclusion
6. Studies which have multiple versions or duplications; only the most recent and complete study was included in the list.

Exclusion criteria

1. Qualitative studies without proper ML methodsdetails and used with other than borehole or geochemical data.
2. Review papers.

Quality control and assessment

The quality control process assists with the selection of the most relevant research papers. A series of research questions was formulated in order to assess the rigorousness, reliability, and relevance of the papers. The questions are given in Table I, based on the methodology developed by Wen *et al.* (2012). The three options were assigned to each question, *i.e.* yes, partially, and no. These options were scored as 1, 0.5, and 0 respectively as in by Wen *et al.* (2012). The quality was assessed by summing the scores for answers against each question. The studies with a minimum score of 3.0 (50% of a perfect score) were selected for data extraction and analysis to ensure the quality of this literature research.

By applying the selection criteria, 50 papers were identified. After scanning the references in these papers, eight additional relevant papers were found. Hence, a total of 58 relevant papers were initially considered. However, after applying the quality control criteria, only 31 papers were selected for data extraction. The quality assessment is discussed in detail in the following sections. The complete list of the 31 selected papers can be found in Appendix A.

Data extraction and analysis

The selected research studies were exploited in order to collect the data that can answer the specific research questions of this exercise. All the reliable and relevant research papers were divided into specific sections, as given in Table II.

These sections were analysed combinedly to provide more meaningful information and to enhance the understandings. The extracted data was both quantitative (number. of boreholes or data-sets) and qualitative (data type, publisher, publication type). Different visualization techniques such as bar charts, pie charts, *etc.* were also used to enhance the data extracted from the research studies.

The vote counting method was used in order to compare the accuracy and application of different ML models (Malhotra, 2015). The vote counting method counts the number of times a model *i.e.*, Model A outperformed Model B, or *vice versa*. With this method, a general idea of whether an estimation ML model outperforms another model in the estimation of mineral grades emerged.

Table I

Research Questions defined for the quality assessment of the published research studies

Code	Question
QA1	Are the aims of the research clearly defined?
QA2	Are the estimation methods well defined and deliberate?
QA3	Is the estimation accuracy measured and reported?
QA4	Is the proposed estimation method compared with other methods?
QA5	Are the findings of the study clearly stated and supported by the results?
QA6	Are the limitations of the study analysed thoroughly?

Table II

Information extracted from the research studies for the analysis

Sr. no	Extracted sections
1	Title of paper
2	Title of journal/conference
3	Publisher
4	Link
5	Year of publication
6	Type (journal/conference)
7	Data-set type (borehole/image)
10	Error assessment technique
11	Research question 1
12	Research question 2
13	Research question 3
14	Research question 4
15	Relevant to research
16	Publication domain

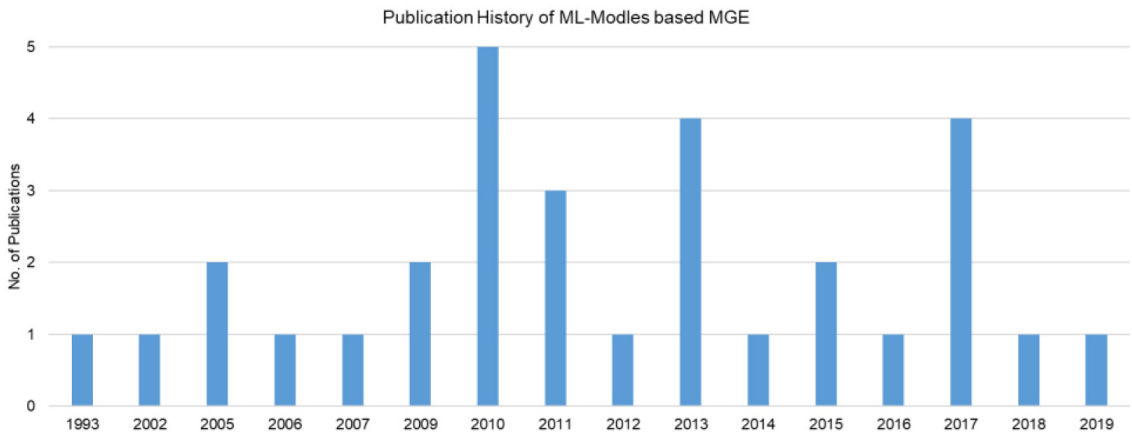

Figure 2—Annual ML research publications from 1990 to 2019

Results and discussion

A total of 31 research studies were identified that dealt with the application of ML methods in the field of MRE. These studies were published between 1990 and 2019 in journals and conference proceedings. Of these, 30 (approx. 97%) papers were published in journals, only one (approx. 3%) was published in conference proceedings, and none were found as book chapters. The places of publication are given in Appendix B. The publishers were mainly Elsevier, Springer, the IEEE, and Taylor & Francis. All of the research studies were experimental and none of them were survey research. In terms of quality control, only those studies were selected with a minimum quality score of 50%, hence all the studies were of a high-quality level.

The publication history is summarized in Figure 2 and shows that the oldest paper found was in the year 1993, followed by eight years wherein the selected publishers did not publish any paper until 2002. A significant increase in ML-based mineral grade estimation papers was found in 2010, with five papers published in that year. In both 2013 and 2017, four papers were published. However, only one paper was found to be published in 2018 and 2019. These statistics show that limited research has been conducted on the application, testing and validation ofML techniques in the field of mineral grade estimation.

Through this exercise, thirteen types of ML-based methods and techniques were identified that have been used for mineral resource exploration and estimation as listed below:

- ➤ Support vector machine (SVM)
- ➤ Support vector regression (SVR) – This is the order of discussion below
  - Artificial neural networks (ANN)
  - Adaptive neuro-fuzzy inference system (ANFIS)
  - Support victor regression (SVR)
- ➤ Local linear radial basis function (LLRBF) neural network
- ➤ Simultaneous perturbation artificial bee colony algorithm (SPABC)
- ➤ Back propagation (BP)
- ➤ Covariance matrix adaptation evolution strategy (CMAES)
- ➤ Particle swarm optimization (PSO)
- ➤ Naïve Bayes classifier (NBC)
- ➤ Radial basis function (RBF)
- ➤ Wavelet neural network (WNN)

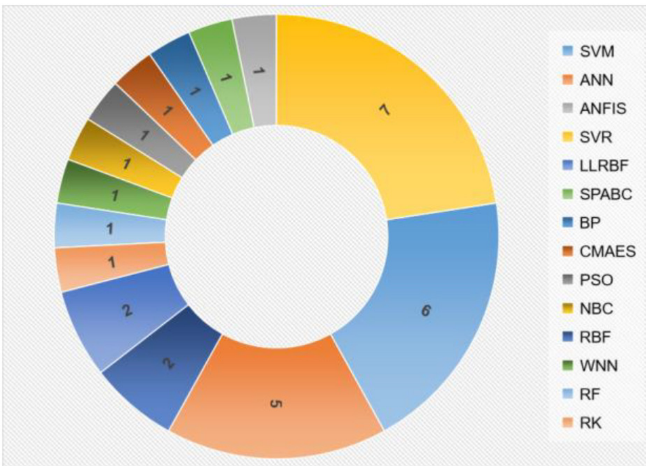

Figure 3— Frequency of use of ML models by the selected research studies

- ➤ Random forest (RF)
- ➤ Regression kriging (RK)

Among all these models and techniques, SVM, SVR, and ANN are the most commonly used and applied for MRE. Together they were found in 60% of the selected studies, as shown in Figure 3. Detailed information about which techniques were used in which study are given in Appendix C.

On the other hand, the non-ML methods mostly used in MRE are kriging, ordinary kriging, and inverse distance weight (IDW). Only 41% of studies (13 papers) compared the results of ML with non-ML methods in terms of mineral grade estimation. All of these studies concluded that ML models outperformed the non-ML methods except S-04, which concluded that ordinary kriging performed better than artificial neural networks. The ML methods which most frequently outperformed the other ML methods were SVM and SVR. A brief description of most the common ML techniques, including SVM, SVR, and ANN, are given in following sections.

The support vector machine is a supervised empirical machine-learning algorithm, based on statistical learning theory (Vapnik 1999). (SVM has recently been introduced in the field of mining and mineralization. The SVM is usually used for data classification and prediction; however, multi-class SVM can also be generated by combining multiple binary classifiers. It showed several unique advantages in a small data sample with nonlinear and multi-dimensional patterns recognition. The main objective of SVM is to locate a hyperplane that can separates data-points of one type from another. The best hyperplane is the one with the largest margin between two classes, and hence can separate the classes distinctively as shown in Figure 4.

The maximum distance between two parallel hyperplanes results in the minimum classification error. SVM has been extensively used in several field of engineering, science, and natural languages. In the mining industry, it has recently been used for mineral classification (Patel, Chatterjee, and Gorgi, 2017), mineral prospectivity (Abedi, Norouzi, and Bahroudi, 2012), and automatic lithological classifications (Yu *et al*. 2012).

Support vector regression differs greatly from other regression models. Whereas the other linear regression models try to minimize the difference between the estimated and the true value, SVR tries to fit the best line within a threshold value. In MRE, SVR tries to categorize all the estimation lines in two forms, those that pass through the threshold boundary and those that don't. The lines that do not pass the threshold boundary are not considered as the difference between the estimated grade value and the true grade value has exceeded the error threshold defined by ε (epsilon) as shown in Figure 5. On the other hand, the lines that pass are considered for a possible support vector to estimate the grade value at an unknown location.

Artificial neural networks is another strong machine-learning approach. This biologically inspired computational technique has wide applications in several science and engineering fields. A unique quality of ANNs is that they are able to create empirical relations between independent and dependent variables and extract the hidden variability and complex knowledge from training data-sets. The relations lies between independent and dependent variables can be built without assumptions about any mathematical depiction of the phenomena. ANN models have several benefits over regression-based ML methods, including their ability to deal with noisy data. An ANN model has thousands of interconnected artificial neurons made up of inputs and outputs. The input nodes receive the actual mineral grade values at known locations based on the internal weighting system and the neural network tries to learn the hidden patterns and produces the output, as shown in Figure 6.

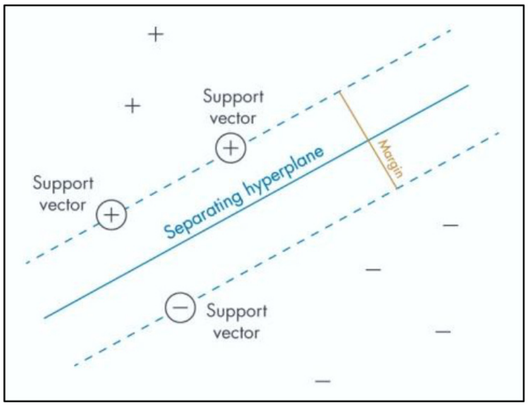

Figure 4—Hyperplane separating the support vectors during the classification by SVM (MATLAB, 2020)

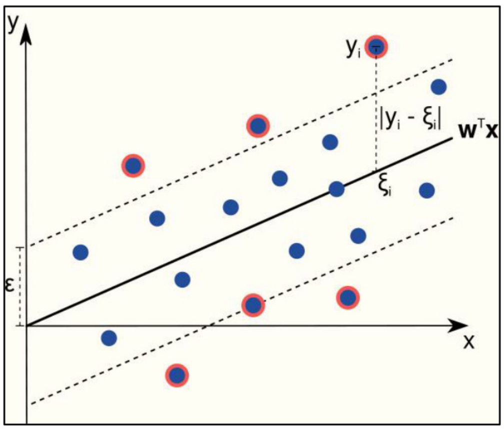

Figure 5—Illustration of an SVR regression function separated by the ξ band for data-sets (Rosenbaum *et al*., 2013)

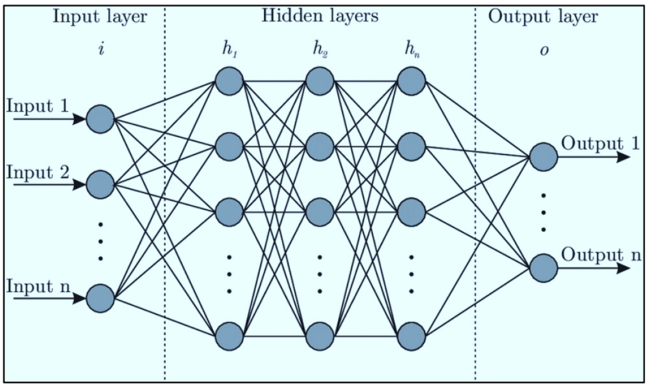

Figure 6—General architecture of an ANN network (Bre, Gimenez, and Fachinnoti, 2018)

An ANN model compares the actual output with what it was meant to produce, *i.e*., the desired output. The difference between both is corrected using back-propagation so the ANN works regressive, going from the final output node to the input nodes in order to fine-tune the weight of its connections until the variation between the actual grade value and estimated grade value produces the lowest possible error.

The types of data used for MRE estimation, regardless of technique, can be divided into three general categories:

- ➤ Exploratory boreholes
- ➤ Images
- ➤ Stream sediments.

Twenty-three studies used exploratory borehole data, six used images photographs, and only one study utilized stream sediment data for the mineral grade estimation. The study S-13 used 3,500 exploratory boreholes and concluded that ML-based SVR is he most accurate technique for mineral grade estimation compared to the non-ML based ordinary kriging. Similarly, research study S-08, in which the authors used stream sediments, also concluded that ANN, which is another ML method, outperformed kriging, a non-ML method, when used for mineral grade estimation. The studies in which images were used did not compare the results of ML models with non-ML models.

In terms of publication domains, the studies were divided into the following three main categories:

- ➤ Minerals/mining
- ➤ Computer science
- ➤ Geoscience

A total of seven studies were published in the field of minerals/mining, 12 in computer sciences, and 12 in the field of geosciences. This reveals that most of the publications fall into multidisciplinary applied domains. Another important aspect of this SLR is to identify the most applicable statistical/geostatistical technique used for the error assessment of the predicted mineral grades. More than 73% of the research studies used mean square error (MSE) and root mean square error (RMSE) as error assessment techniques, followed by the standard error (SE), and mean absolute error (MSE) with a total of 16%. The remaining 11% of the error assessment techniques include mainly generalization error, percentage of accuracy, co-efficient of determination, and out-of-bag error.

This review has found that most of the recent and advance ML methods and techniques such as deep convolution neural network, hierarchical convolutional deep maxout network, and hidden trajectory (generative) models, are seldom used within the field of mineral grade estimation. Therefore, researchers/practitioners are encouraged to apply these techniques and test their applicability for mineral grade estimation.

In addition, researchers/practitioners are also encouraged to explore the other, non-tested, ML methods for mineral grade estimation. In order to become acquainted with the unexplored ML methods and to apply them in a more efficient way, researchers/practitioners should keep a close watch on the related disciplines like machine learning, deep learning, data science, and artificial intelligence, as these disciplines may provide ideas for new ML methods and techniques (Wen *et al*., 2012). Even though this investigation has found that the ML models are more accurate and perform better than the non-ML methods, the results of error assessment between ML and non-ML methods, and between different ML methods, are still inconclusive. Hence, it is strongly recommended that the scientific community to develop a common framework for evaluating the performance of different ML techniques, as well as against non-ML methods. The results of the studies may vary because of different data-sets and/or different experimental designs.

Conclusion

This systematic literature review examined machine learning (ML)-based MRE models in terms of the type of ML methods or techniques, the error estimation of applied ML methods, comparison between different ML and non-ML methods, as well as ML methods with other ML methods. The extensive systematic literature review was based on research studies published in the period 1990-2019, with a total of 31 studies meeting the requirements of five research questions.

The key findings of the litrature review are that the significant ML methods applied for the MRE are, in order of application, SVM, ANN and SVR. Few studies actually compared the results of ML methods with non-ML methods for MRE. Those studies concluded that ML methods outperformed the non-ML methods in general. SVM and SVR are the most applied and tested ML methods, which yield much better results than other ML methods. Very few papers have been published in the fields of mineral and mining, whereas most of them were published in the computer and geosciences fields.

This review provides recommendations for researchers for future work as well as guidelines for practitioners. More research should be conducted detailing studies on the application of ML methods and drawing of conclusions in terms of their applicability, validation, and accuracy. Researchers should also develop a framework in terms of data usage and accuracy assessment against non-ML methods used for mineral grade estimation.

From this review, it is very clear that the application of ML methods in the industry for MRE are limited, and hence more studies should be conducted and analysis done in order to find the possible barriers to ML method applications. It is strongly recommended that ML models be used in parallel with non-ML (conventional statistical) models in the early stages of mineral grade estimations. After the error assessment and quality check, the validated ML methods can then be used for estimation of Mineral Resources. Moreover, it is advisable to consult researchers/experts from the fields of machine learning and/or data science in order to check the strengths and weaknesses of the potential ML methods and interpretation of results before and after application accordingly.

Acknowledgment

The work presented here is based on the PhD research study of the first author in the School of Mining Engineering at the University of the Witwatersrand, Johannesburg, South Africa. The authors would also like to acknowledge the valuable feedback provided by the reviewers, which has improve the presentation and quality of work.

References

Abbaszadeh, M., Hezarkhani, A., and Soltani-Mohammadi, S. 2013. An SVM-based machine learning method for the separation of alteration zones in Sungun porphyry copper deposit. *Geochemistry*, vol. 73, no. 4. pp. 545–554.

Abedi, M., Norouzi, G.-H., and Bahroudi, A. 2012. Support vector machine for multi-classification of mineral prospectivity areas. *Computers & Geosciences*, vol. 46. pp. 272–283.

Bre, F., Gimenez, J.M., and Fachinotti, V.D. 2018. Prediction of wind pressure coefficients on building surfaces using artificial neural networks. *Energy and Buildings*, vol. 158. pp. 1429–1441.

Chatterjee, S. and Bandopadhyay, S. 2011. Goodnews Bay platinum resource estimation using least squares support vector regression with selection of input space dimension and hyperparameters. *Natural Resources Research*, vol. 20, no. 2. pp. 117–129. not referenced in this version

Chatterjee, S., Bandopadhyay, S., and Machuca, D. 2010. Ore grade prediction using a genetic algorithm and clustering based ensemble neural network model. *Mathematical Geosciences*, vol. 42, no. 3. pp. 309–326.

Chatterjee, S., Bandopadhyay, S., and Rai, P. 2008. Genetic algorithm-based neural network learning parameter selection for ore grade evaluation of limestone deposit. *Mining Technology*, vol. 117, no. 4. pp. 178–190.

Chatterjee, S., Bhattacherjee, A., Samanta, B., and Pal, S.K. 2010. Image-based quality monitoring system of limestone ore grades. *Computers in Industry*, vol. 61, no. 5. pp. 391–408.

Das Goswami, A., Mishra, M., and Patra, D. 2017. Investigation of general regression neural network architecture for grade estimation of an Indian iron ore deposit. *Arabian Journal of Geosciences*, vol. 10, no. 4. pp. Article no. 80.

Dutta, S., Bandopadhyay, S., Ganguli, R., and Misra, D. 2010. Machine learning algorithms and their application to ore reserve estimation of sparse and imprecise data. *Journal of Intelligent Learning Systems and Applications*, vol. 2, no. 02. pp. doi:104236/j/lsa.2010.22012.

Horrocks, T., Wedge, D., Holden, E.-J., Kovesi, P., Clarke, N., and Vann, J. 2015. Classification of gold-bearing particles using visual cues and cost-sensitive machine learning. *Mathematical Geosciences*, vol. 47, no. 5. pp. 521–545.

Hosseini, S., Asghari, O., and Emery, X. 2017. Direct block-support simulation of grades in multi-element deposits: application to recoverable mineral resource estimation at Sungun porphyry copper-molybdenum deposit. *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 117, no. 6. pp. 577–585.

Jafrasteh, B. and Fathianpour, N. 2017. A hybrid simultaneous perturbation artificial bee colony and back-propagation algorithm for training a local linear radial basis neural network on ore grade estimation. *Neurocomputing*, vol. 235. pp. 217–227.

Jafrasteh, B., Fathianpour, N., and Suárez, A. 2018. Comparison of machine learning methods for copper ore grade estimation. *Computational Geosciences*, vol. 22, no. 5. pp. 1371–1388.

Jalloh, A.B., Kyuro, S., Jalloh, Y., and Barrie, A.K. 2016. Integrating artificial neural networks and geostatistics for optimum 3D geological block modeling in mineral reserve estimation: A case study. *International Journal of Mining Science and Technology,* vol. 26, no. 4. pp. 581–585.

Kitchenham, B. and Charters, S. 2007. Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report. EBSE-2007-01.

Li, X.-L., Li, L.-h., Zhang, B.-l., and Guo, Q.-j. 2013. Hybrid self-adaptive learning based particle swarm optimization and support vector regression model for grade estimation. *Neurocomputing,* vol. 118. pp. 179–190.

Li, X.-l., Xie, Y.-l., Guo, Q.-j., and Li, L.-h. 2010. Adaptive ore grade estimation method for the mineral deposit evaluation. *Mathematical and Computer Modelling,* vol. 52, no. 11–12. pp. 1947–1956.

Li, X., Ao, Y., Guo, S., and Zhum, L. 2019. Combining regression kriging with machine learning mapping for spatial variable estimation. *IEEE Geoscience and Remote Sensing Letters,* vol. 17, no. 1. pp. 27–31.

Malhotra, R. 2015. A systematic review of machine learning techniques for software fault prediction. *Applied Soft Computing*, vol. 27. pp. 504–518.

MATLAB. 2020. Support Vector Machine SVM. https://www.mathworks.com/discovery/support-vector-machine.html, [accessed 10 January 2021].

Patel, A.K., Chatterjee, S., and Gorai, A.K. 2017. Development of machine vision-based ore classification model using support vector machine (SVM) algorithm. *Arabian Journal of Geosciences*, vol. 10, no. 5. p. 107.

Patel, A.K., Chatterjee, S., and Gorai, A.K. 2017. Development of online machine vision system using support vector regression (SVR) algorithm for grade prediction of iron ores. *Fifteenth IAPR International Conference on Machine Vision Applications* (MVA), IEEE, New York. doi:10.23919/MVA.2017.1986823

Perez, C.A., Estévez, P.A., Vera, P.A., Castillo, L.E., Aravena, C.M., Schulz, D.A., and Medina, L.E. 2011. Ore grade estimation by feature selection and voting using boundary detection in digital image analysis. *International Journal of Mineral Processing*, vol. 101, no. 1-4. pp. 28–36.

Rosenbaum, L., Dörr, A., Bauer, M.R., Boeckler, F.M., and Zell, A. 2013. Inferring multi-target QSAR models with taxonomy-based multi-task learning. *Journal of Cheminformatics,* vol. 5, no. 1. p. 33.

Rossi, M E. and Deutsch, C.V. 2013. Mineral Resource Estimation. Springer Science & Business Media.

Samanta, B. 2010. Radial basis function network for ore grade estimation. *Natural Resources Research*, vol. 19, no. 2. pp. 91–102.

Samanta, B. and Bandopadhyay, S. 2009. Construction of a radial basis function network using an evolutionary algorithm for grade estimation in a placer gold deposit. *Computers & Geosciences,* vol. 35, no. 8. pp. 1592–1602.

Samanta, B., Bandopadhyay, S., and Ganguli, R. 2002. Data segmentation and genetic algorithms for sparse data division in Nome placer gold grade estimation using neural network and geostatistics. *Exploration and Mining Geology,* vol. 11, no. 1–4. pp. 69–76.

Samanta, B., Bandopadhyay, S., and Ganguli, R. 2006. Comparative evaluation of neural network learning algorithms for ore grade estimation. *Mathematical Geology,* vol. 38, no. 2. pp. 175–197.

Samanta, B., Banopadhyay, S., Ganguli, R., and Dutta, S. 2005. A comparative study of the performance of single neural network vs. Adaboost algorithm based combination of multiple neural networks for mineral resource estimation. *Journal of the Southern African Institute of Mining and Metallurgy,* vol. 105, no. 4. pp. 237–246.

Samanta, B., Ganguli, R., and Bandopadhyay, S. 2005. Comparing the predictive performance of neural networks with ordinary kriging in a bauxite deposit. *Mining Technology,* vol. 114, no. 3. pp. 129–139.

Savu-Krohn, C., Rantitsch, G., Auer, P., Melcher, F., and Graupner, T. 2011. Geochemical fingerprinting of coltan ores by machine learning on uneven datasets. *Natural Resources Research,* vol. 20, no. 3. pp. 177–191.

Sheng, L., Zhang, T., Niu, G., Wang, K., Tang, H., Duan, Y., and Li, H. 2015. Classification of iron ores by laser-induced breakdown spectroscopy (LIBS) combined with random forest (RF). *Journal of Analytical Atomic Spectrometry,* vol. 30, no. 2. pp. 453–458.

Tahmasebi, P. and Hezarkhani, A. 2010. Application of adaptive neuro-fuzzy inference system for grade estimation; case study, Sarcheshmeh porphyry copper deposit, Kerman, Iran. *Journal of Basic and Applied Sciences,* vol. 40, no. 3. pp. 408–420.

Tahmasebi, P. and Hezarkhani, A. 2012. A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. *Computers & Geosciences,* vol. 42. pp. 18–27.

Tessier, J., Duchesne, C., and Bartolacci, G. 2007. A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts. *Minerals Engineering,* vol. 20, no. 12. pp. 1129–1144.

Vapnik, V.N. 1999. An overview of statistical learning theory. *IEEE Transactions on Neural Networks,* vol. 10, no. 5. pp. 988–999.

Wen, J., Li, S., Lin, Z., Hu, Y., and Huang, C. 2012. Systematic literature review of machine learning based software development effort estimation models. *Information and Software Technology,* vol. 54, no. 1. pp. 41–59.

Wu, X. and Zhou, Y. 1993. Reserve estimation using neural network techniques. *Computers & Geosciences,* vol. 19, no. 4. pp. 567–575.

Yu, L., Porwal, A., Holden, E.-J., and Dentith, M.C. 2012. Towards automatic lithological classification from remote sensing data using support vector machines. *Computers & Geosciences,* vol. 45. pp. 229–239.

Zhang, X., Song, S., Li, J., and Wu, C. 2013. Robust LS-SVM regression for ore grade estimation in a seafloor hydrothermal sulphide deposit. *Acta Oceanologica Sinica,* vol. 32, no. 8. pp. 16–25.

Zhang, Y., Song, S., You, K., Zhang, X., and Wu, C. 2017. Relevance vector machines using weighted expected squared distance for ore grade estimation with incomplete data. *International Journal of Machine Learning and Cybernetics,* vol. 8, no. 5. pp. 1655–1666.

Zhang, Z., Yang, J., Wang, Y., Dou, D., and Xia, W. 2014. Ash content prediction of coarse coal by image analysis and GA-SVM. *Powder Technology,* vol. 268. pp. 429–435. ◆

APPENDIX A

List of research publications used in the systematic literature review after quality control and assessment		
Study ID	**Title**	**Reference**
S-01	Reserve estimation using neural network techniques.	Wu and Zhou, 1993
S-02	Data segmentation and genetic algorithms for sparse data division in nome placer gold grade estimation using neural network and geostatistics.	Samanta, Bandopadhyay, and Ganguli, 2002
S-03	A comparative study of the performance of single neural network vs. adaboost algorithm based combination of multiple neural networks for mineral resource estimation.	Samanta, Bandopadhyay, Ganguli, and Dutta, 2005
S-04	Comparing the predictive performance of neural networks with ordinary kriging in a bauxite deposit.	Samanta, Ganguli, and Bandopadhyay, 2005
S-05	Comparative evaluation of neural network learning algorithms for ore grade estimation.	Samanta, Bandopadhyay, and Ganguli, 2006
S-06	A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts.	Tessier, Duchesne, and Bartolocci, 2007
S-07	Genetic algorithm-based neural network learning parameter selection for ore grade evaluation of a limestone deposit.	Chatterjee, Bandopadhyay, and Rai, 2008

APPENDIX A (*continued*)

List of research publications used in the systematic literature review after quality control and assessment		
Study ID	**Title**	**Reference**
S-08	Ore grade prediction using a genetic algorithm and clustering based ensemble neural network model.	Chatterjee, Bandopadhyay, and Machuca, 2010
S-09	construction of a radial basis function network using an evolutionary algorithm for grade estimation in a placer gold deposit.	Samanta and Bandopadhyay, 2009
S-10	Application of adaptive neuro-fuzzy inference system for grade estimation; case study, sarcheshmeh porphyry copper deposit, Kerman, Iran.	Tahmasebi and Hezarkhani, 2010
S-11	Image-based quality monitoring system of limestone ore grades.	Chatterjee, Bhattacherjee, Samata, and Pal, 2010
S-12	Radial basis function network for ore grade estimation.	Samanta, 2010
S-13	Machine learning algorithms and their application to ore reserve estimation of sparse and imprecise data.	Dutta, Bandopadhyay, Ganguli, and Misra, 2010
S-14	Adaptive ore grade estimation method for the mineral deposit evaluation.	Li, Xie, Guo, and Li, 2010
S-15	Ore grade estimation by feature selection and voting using boundary detection in digital image analysis.	Perez, Estévez, Vera, Castillo, Aravena, Schultz, and Medina, 2011
S-16	geochemical fingerprinting of coltan ores by machine learning on uneven datasets.	Savu-Krohn, Rantitsch, Auer, Melcher, and Graupner, 2011
S-17	Goodnews Bay platinum resource estimation using least squares support vector regression with selection of input space dimension and hyperparameters.	Chatterjee and Bandopadhyay, 2011
S-18	A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation.	Tahmasebi and Hezarkhani, 2012
S-19	Hybrid self-adaptive learning based particle swarm optimization and support vector regression model for grade estimation.	Li, Li, Zhang, and Guo, 2013
S-20	An SVM-based machine learning method for the separation of alteration zones in sungun porphyry copper deposit.	Abbaszadeh, Hezarkhani , and Soltani-Mohammadi, 2013
S-21	Robust LS-SVM regression for ore grade estimation in a seafloor hydrothermal sulphide deposit.	Zhang, Song, You, Zhang, and Wu, 2013
S-22	Ash content prediction of coarse coal by image analysis and GA-SVM.	Zhang, Yang et al. 2014
S-23	Classification of iron ores by laser-induced breakdown spectroscopy (libs) combined with random forest (RF).	Sheng, Zhang, Niu, Wang, Tang, Duan, and Li, 2015
S-24	Classification of gold-bearing particles using visual cues and cost-sensitive machine learning.	Horrocks, Wedge, Holden, Kovesi, Clarke, and Vann, 2015
S-25	Integrating artificial neural networks and geostatistics for optimum 3d geological block modeling in mineral reserve estimation: A case study.	Jalloh, Kyuro, Jalloh, and Barrie, 2016
S-26	A hybrid simultaneous perturbation artificial bee colony and back-propagation algorithm for training a local linear radial basis neural network on ore grade estimation.	Jafrasteh and Fathianpour, 2017
S-27	Investigation of general regression neural network architecture for grade estimation of an indian iron ore deposit.	Das Goswami, Mishra, and Patra, 2017
S-28	Development of online machine vision system using support vector regression (svr) algorithm for grade prediction of iron ores.	Patel, Chatterjee, and Gorai, 2017
S-29	Relevance vector machines using weighted expected squared distance for ore grade estimation with incomplete data	Zhang, Song, You, Zhang, and Wu, 2017
S-30	Comparison of machine learning methods for copper ore grade estimation.	Jafrasteh, Fathianpour, and Suárez, 2018
S-31	Combining regression kriging with machine learning mapping for spatial variable estimation	Li, Ao, Guo, and Zhu, 2019

APPENDIX B

Publication domains of the selected research studies used for SLR		
Study ID	**Title of journal/cConference**	**Publisher**
S-01	Computers & Geosciences	Elsevier
S-02	Exploration and Mining Geology	Canadian Institute of Mining, Metallurgy and Petroleum
S-03	Journal of The Southern African Institute of Mining and Metallurgy	Southern African Institute of Mining and Metallurgy
S-04	Transactions of the Institutions of Mining and Metallurgy: Section A: Mining Technology	Taylor & Francis
S-05	Mathematical Geology	Springer
S-06	Minerals Engineering	Elsevier
S-07	Transactions of the Institutions of Mining and Metallurgy: Section A: Mining Technology	Taylor & Francis
S-08	Mathematical Geosciences	Springer

APPENDIX B (*continued*)

Publication domains of the selected research studies used for SLR		
Study ID	**Title of journal/cConference**	**Publisher**
S-09	Computers & Geosciences	Elsevier
S-10	Australian Journal of Basic and Applied Sciences	Australian Journal of Basic and Applied Sciences
S-11	Computers in Industry	Elsevier
S-12	Natural Resources Research	Springer
S-13	Journal of Intelligent Learning Systems and Applications	Scientific Research Publishing Inc
S-14	Mathematical and Computer Modelling	Elsevier
S-15	International Journal of Mineral Processing	Elsevier
S-16	Natural Resources Research	Springer
S-17	Natural Resources Research	Springer
S-18	Computers & Geosciences	Elsevier
S-19	Neurocomputing	Elsevier
S-20	Geochemistry	Elsevier
S-21	Acta Oceanologica Sinica	Springer
S-22	Powder Technology	Elsevier
S-23	Journal of Analytical Atomic Spectrometry	Royal Society of Chemistry
S-24	Mathematical Geosciences	Springer
S-25	International Journal of Mining Science and Technology	Elsevier
S-26	Neurocomputing	Elsevier
S-27	Arabian Journal of Geosciences	Springer
S-28	Fifteenth IAPR International Conference on Machine Vision Applications (MVA)	IEEE Xplore
S-29	International Journal of Machine Learning and Cybernetics	Springer
S-30	Computational Geosciences	Springer
S-31	IEEE Geoscience and Remote Sensing Letters	IEEE Xplore

APPENDIX C

The type of ML models applied in the research studies for MRE	
Study ID	**ML model applied for MRE**
S-01	Support vector machine (SVM)
S-02	Coactive neuro-fuzzy inference system based on artificial neural networks adaptive neuro-fuzzy inference system (ANFIS)
S-03	Adaptive neuro-fuzzy inference system (ANFIS) artificial neural networks (ANN) kriging
S-04	Support vector regression (SVR)
S-05	Support vector machine (SVM) Neural network kriging
S-06	Local linear radial basis function (LLRBF) neural network LLRBF network with skewed Gaussian activation function Simultaneous perturbation artificial bee colony (SPABC) algorithm Back propagation (BP) Standard artificial bee colony Covariance matrix adaptation evolution strategy (CMAES) Particle swarm optimization (PSO) Support vector machine (SVM)

APPENDIX C (*continued*)

The type of ML models applied in the research studies for MRE	
Study ID	**ML model applied for MRE**
S-07	Support vector machine (SVM Random forest (RF)
S-08	Multi-layer perceptron neural network
S-09	Support vector machine (SVM)
S-10	Support vector machine (SVM) A naive Bayes classifier A majority decision table
S-11	Least square support vector machine regression Inverse distance weight Ordinary kriging Back propagation neural network
S-12	Support vector machine (SVM)
S-13	Multi-layer feed forward neural network Simple kriging Ordinary kriging Kriging with linear drift function Kriging with a quadratic drift function Lognormal kriging
S-14	Support vector machines (SVM) Linear programming boosting
S-15	General regression neural network Multilayer perceptron neural network Ordinary kriging
S-16	Radial basis function (RBF)
S-17	Radial basis function network Ordinary kriging
S-18	Support vector machine (SVM) Regression
S-19	Single neural network Multiple neural network
S-20	Support vector machine (SVM)
S-21	Layered feedforward artificial neural network
S-22	Support vector regression (SVR)
S-23	Artificial neural networks (ANN) Geostatistics
S-24	Artificial neural networks (ANN) Ordinary kriging
S-25	Relevance vector machine Expected squared distance Weighted expected squared distance Inverse distance weighted
S-26	Neural networks Support vector regression (SVR) Ordinary kriging)
S-27	Neural networks
S-28	Wavelet neural network
S-29	Multilayer feedforward neural network
S-30	Neural networks Random forests (RF) Gaussian processes
S-31	Ordinary kriging Regression kriging Machine learning mapping Hybrid method