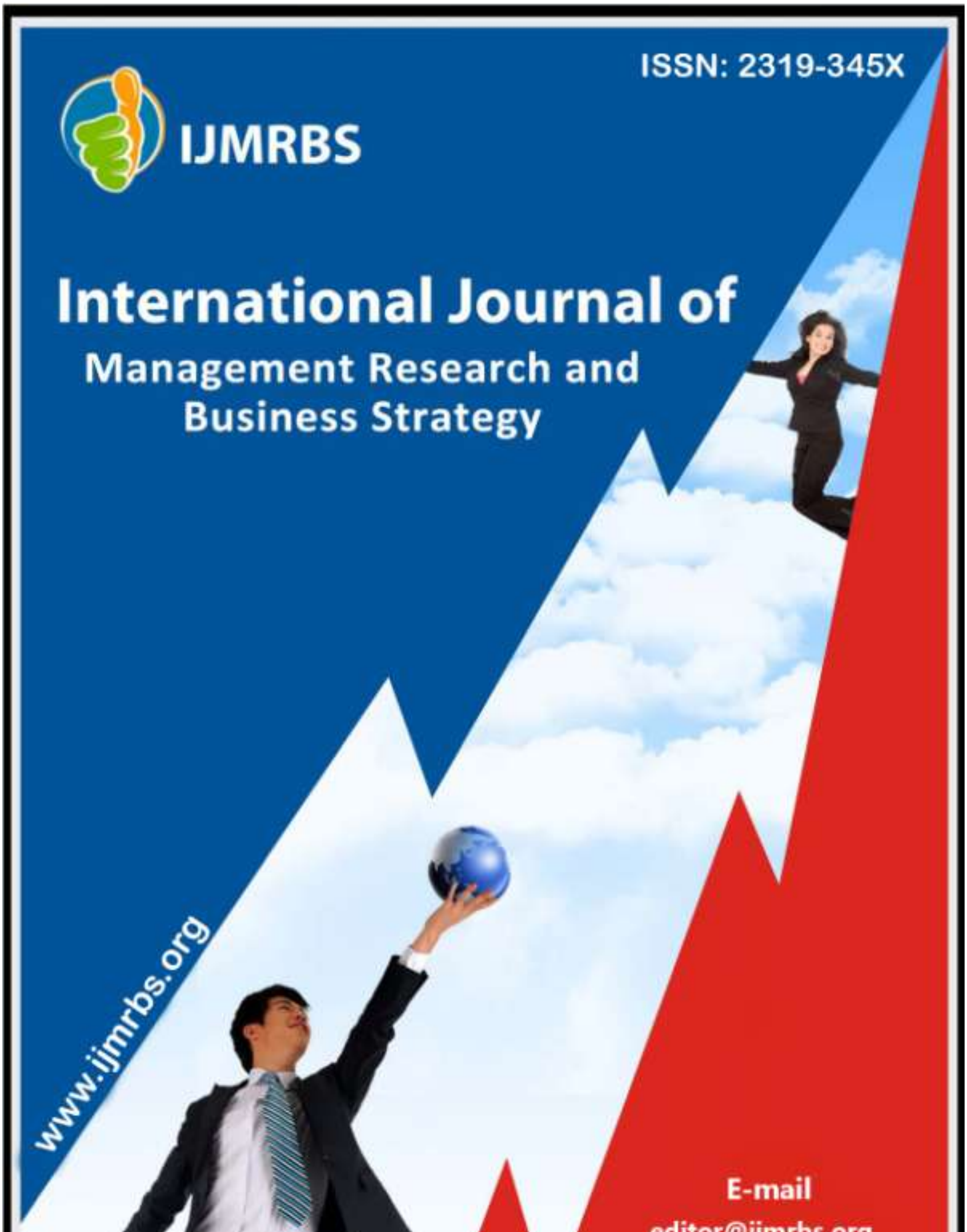
ISSN: 2319-345X
IJMRBS
International Journal of
Management Research and
Business Strategy
www.ijmrbs.org
E-mail
editor@ijmrbs.org
editor.ijmrbs@gmail.com

ISSN 2319-345X www.ijmrbs.com

Vol. 9, Issuse. 1, March 2021

ANALYSIS MODELS OF TECHNICAL AND ECONOMIC DATAOFMININGENTERPRISESBASEDONBIGDATA ANALYSIS

B.D.RAVEENDRABABU1,M.LAKSHMIMADHURI2

Abstract- Multidimensionality and nonlinearity are two of the characteristics of mining company data. An important economic indication for mining companies is the price data on mineral products, and an equally important technical indicator is the geological data itself. Using data mining and big data analysis techniques, researchers are looking at the best way to analyze technical and economic data. The pricing of mineral products is examined to determine its fluctuation pattern and affecting factors. Using an artificial neural network, the price of mineral goods is predicted. The findings reveal that the prediction model's practicality and accuracy are strong. In the course of mineral exploration, much geological data have been lost due to technical and equipment limitations, reducing the accuracy of orebody shapes and reserve estimations. Geostatistics and artificial neural networks are used to develop a model for predicting missing geological data. As a result of utilizing this model, it is possible to debate and assess the regularity of borehole geological data for single, multiple, and all boreholes. Predictions and interpolations for most of the missing geological data have been proven to be accurate and dependable, according to this study.

Keywords- data mining firms; technological and economic information; BP neural network; forecasting models

I. INTRODUCTION

It is difficult for mining companies to respond to changes in supply and demand because of the long production cycle. The company's operations strategy can't be produced rapidly enough to keep up with the globalization of the mining business. The creation of a mine operating plan is a critical component of mining company operations management. The mining company's information systems generate and collect vast amounts of data that are currently not examined professionally. Processes were unable to be adequately supported by thesedata.of the mining enterprise's production management and decision-making As a result, mining companies place a high value on developing models that can effectively analyze and anticipate mine technical and economic data. The mining company's information systems generate and collect vast amounts of data that are currently not examined professionally. The mining company's production management and decision-making processes were not adequately supported by these facts. For this reason, mining companies place a premium on developing robust models for analyzing and forecasting technical and economic data. The methods for analyzing and forecasting technological and economic data have been the subject of several investigations. The reference examines the methods for predicting and

1AssistantProfessor,DeptofCSE,CBIT,Proddutur,AP,India.
2AssistantProfessor,DeptofCSE,CBIT,Proddutur,AP,India.

interpolating missing economic data from the mining company, such as the mean method, the weighted average method, the linear regression method, the maximum expected method, and the multiple imputation method. method. Data from boreholes is analyzed by Reference to determine the worldwide trend and the aeolotropism in the data set. To increase interpolation precision, the data is transformed to normal distribution and the aeolotropism of the data is eliminated. Based on the characteristics of the technical and economic data of mining enterprises, a prediction model for mineral product prices and an interpolation model for geological missing data are developed using geostatistics and artificial neural networks, respectively, in order to improve mining companies' data analysis capabilities.

II. STATE-OF-THE-ARTANALYSIS

There are a slew of approaches in the current scientific literature addressing the challenge of large data stream mining while maintaining privacy. In the following, we'll go over some of the most noteworthy over the past several months. It offers a brand-new algorithm for erasing the traces of individuals. Gradual Trajectory Stream Anonymizer (ITSA), in particular, is incremental in nature and uses sliding windows to process target streams. This is a key feature of the proposed approach. Individuals joining and leaving the stream dynamically update these windows. As a result of End-to-end processing of enormous amounts of big data necessitates using effective data structures. The benefits of the proposed algorithm have been proven through extensive testing. It turns the attention to the context of electronic health data streams, whose privacy is carefully researched. The final proposal is called delay-free anonymization. The most important feature of this proposal is that input streams are immediately anonymized with fake values. The anonymization process is further improved by authors in order to increase the usefulness of the data. offer a new way of late validation that increases the overall benefit of privacy preservation. The privacy of trajectory streams is still taken into account. In this case, the answer is achieved through the implementation of appropriate data stream access control methods, as well as a new Privacy Protection Mechanism (PPM). Anonymity and diversity schemes over data streams, such as those mandated by PPM, are proven to meet solid privacy criteria by generalization.. When data streams are delayed in publication using the PPM methodology, false-negative results may be introduced that impact query processing. Authors thus acknowledge and present hardness results that have been enhanced by detailed experimental evaluation of a novel problem known as precision-bounded access control for privacy-preservation stream mining. Using a technique called Shadow Coding, it is able to protect the privacy of transmitted data and the recovery of collected data in dispersed data stream environments. They demonstrate that the hypothesis is correct. methodachieves privacy-preserving computationinadata-recoverable,efficient,and scalableway,beingscalabilityafirst-classrequirementforbigdataprocessing.Practical techniques that make Shadow Coding efficient and safeoverdatastreamsarealsoprovided.Authors complete their analytical contributions bymeans of anextensiveexperimentalstudyonalarge-scale reallifedataset.Itfocusesonthegeneralproblem ofestimatingthesortednessofdatastreamsina The Longest Increasing Subsequence (LIS) of target data streams can be used to preserve anonymity. Many modern applications, such as financial data stream monitoring, surveillance data stream processing and so on can benefit from this important aspect of the streams of information. Block decomposition and local approximation techniques are used as the foundation of the approach. Differential privacy strategies are used to address the challenge of ensuring user privacy while releasing geospatial data streams. The method is based on computing an appropriate synopsis with high utility to facilitate accurate query processing over geospatial data streams with well-defined dynamic scopes (thus the privacy need). TheRealtime Geospatial Publish (RGP) appears to be a viable method for distributing geospatial data in real time. There is no need to worry about categorical data streams, though. As a result, the authors offer a brand-new anonymization method for protecting sensitive data from being leaked or altered. An unique two-phase anonymization approach is included in the proposed technique. They demonstrate that this strategy is highly efficient in terms of speed and communication, and resistant to hypothetical tampering from adversaries. Extensive testing has shown that the proposed method is effective. Also investigates the interesting subject of safeguarding output-privacy (e.g., data stream privacy) of classification algorithms (such as privacy of the classifiers' outputs). Accordingly, the DAHOT algorithm, which combines Hoeffding tree's widely used data stream classification algorithm with appropriate variations of the k-anonymity and ldiversity principles, is proposed as a systematic approach by the authors.

CHALLENGESANDDIRECTIONS

Research in privacy-preserving big data stream mining opens the door to a variety of research difficulties and possibilities that should be examined in the near future. A few of these fascinating subjects will be discussed in the following paragraphs.

Emerging DomainsPrivacy-preserving big data stream mining algorithms are driven and determined by real applications and systems because of the specific research focus, i.e. privacy-preserving big data stream mining. There will be a major role for developing domains, such as social networks, TV provisioning and transportation systems, in the future.

Precision versus Confidentiality: Big data stream mining algorithms' accuracy and privacy are at odds. It's a vital scientific question to find the right balance between these two features. What is the best way to ensure privacy while yet maintaining accuracy? The latter is a question that should be addressed in the future.Concept-DriftIssues:Bigdatastreamsareaffected by concept-drift problems. This makes harder the privacy preserving requirement, because of, in general,preservingtheprivacyofdataisperformed in dependence on a predetermined set of attributes/concepts of the target data model.

Security Issues: Preserving the privacy of big data streams implies accessing big data streams, of course.Thisinvolvesinaproblematicside-effect:

how to ensure the security of the same big data streams while accessing them? Combining security and privacy (as well as privacy and security) is an annoying problem for big data stream mining research.

Cryptography: Model-based and algorithmic approaches have traditionally been used to solve the privacy-preserving massive data stream mining problem. Along with these efforts, cryptographic approaches are emerging as a promising avenue for future research.

Quality and Utility of Data:Ensuring the privacy of data streams while mining big data streams may deterioratethesamequalityand utilityofsuchdata. Asaconsequence, the latter arecritical issues for the future.

Stream Analytics: Models, techniques and algorithms proposed by active literature must converge in suitable unifying frameworks forfinally supporting privacy-preserving big data stream analytics, a critical research challenge at now. Here, several issues arise: from architectural requirements to parameter tuning, from framework trade-offs to performance, and so forth.

Performance: Last but not least, performance issues always arise when processing big data streams (to preserve their privacy, in this case). Asa consequence, devising models and optimizations that allow us to ensure performance of privacy-preservation mining methods over big data streams is a relevance challenge for the future.

III.**THE PREDICTION AND INTERPOLATIONMODELOFTHE GEOLOGICALMISSINGDATA**

A huge number of boreholes are drilled during the mineral development process to conduct mineral exploration. An analysis of drilled-core samples can reveal the varying depths of the core samples' ore grades. There are logs kept of core sample data that are used to assess mineral grades in the surrounding formation. Geostatistics can be used to determine the specific grade of ore in an area with few known sample values based on these geological facts. To make a 3D model of the orebody, the boundary of the orebody must be identified. Orebody delineation is the foundation for accurately characterizing the ore body's spatial distribution pattern and plays a vital role in reserve computation and mining design. A thorough database of geophysical and geological information obtained from surface and borehole data is required to properly restrict an orebody's structure and size. Many fundamental geological data have been lost as a result of the limitations of technical circumstances and equipment restrictions. After that, it became apparent.makes it impossible to provide the complete and accurate geological data for the modeling processof the orebody, which reduces the accuracy of the orebody shape and that of reserves estimation.

A. **TheEstabishmentoftheModel**

A neural network model for predicting geological missing data has been developed. The input vector for the back-propagation neural network model is established using the geological coordinate data. In order to create the output vector, the ore grade data is used. Three input neurons and one output neuron make up a single hidden layer of a three-layered neural network. Hidden layer neurons number 10 based on training samples and an empirical formula. The 'tansig' transfer function is utilized between the input and hidden layers, whereas the 'purelin' transfer function is used between the hidden and output layers. The 'trainlm' function is used for network training. For speedy learning and accurate function approximation, network structure is properly changed in the application process in light of big sample data..

B. TheTrainingandPredictionofthe Model

1) The regularity of geological data of all boreholes

2) For the input vector, the back-propagation neural network model uses the x, y, and z values of the samples centroid of all boreholes, while for the output vector, the corresponding ore grade data is employed. A low goodness-of-fit R value shows that the geological data from all of the boreholes does not show any substantial regularity, which is supported by the training and testing results, which demonstrate that the MSE is tiny and near to the goal value.

3) The regularity of geological data of singleborehole

A huge number of boreholes are drilled during the mineral development process to conduct mineral exploration. An analysis of drilled-core samples can reveal the varying depths of the core samples' ore grades. There are logs kept of core sample data that are used to assess mineral grades in the surrounding formation. Geostatistics can be used to determine the specific grade of ore in an area with few known sample values based on these geological facts. To make a 3D model of the orebody, the boundary of the orebody must be identified. Orebody delineation is the foundation for accurately characterizing the ore body's spatial distribution pattern and plays a vital role in reserve computation and mining design. A thorough database of geophysical and geological information obtained from surface and borehole data is required to properly restrict an orebody's structure and size. Many fundamental geological data have been lost as a result of the limitations of technical circumstances and equipment restrictions. After that, it became apparent.the x, y, z values of the samples centroid of boreholes A,B,C,D, as the input of the back-propagation neural network model, are used to establish the input vector, and corresponding ore grade data are used to establish the output vector. ThetrainingresultsshowthatRis0.8497andMSE is 2.26. The training effectiveness is good. The neural network models trained by the samples data are used to predict the missing data of 8 boreholes. The analysis of prediction results of boreholes A,B, C, D shows that R values are more than 0.9226, and MSE values arelessthan0.12. The Rvalues of prediction results of boreholes F, G, H are more than 0.9443, and MSE values are less than 0.19. Therefore, it is scientific that the data of boreholes with same characteristics are trained together.

4) The regularity of geological data of groupboreholes

There are five kinds of boreholes based on data features, the continuity of the ore body, and the distance from the main fault. Five sets of geological data are used to train the network models. Analysis of training outcomes shows that the R value is 0.9623, with 414 samples in the first group. With 1433 samples in the second group, a R value of 0.8791 has been determined by an examination of the training data set. Analysis of training results shows that the R value for the third group is 0.7449, with 10760 samples. There are almost forty thousand samples in the fourth and fifth groups, and an analysis of the training results shows that R values are quite near to zero there. To put it another way: The training efficacy of samples from the first three groups is significantly higher than that of samples from groups four and five.The analysis results show that the regularity of the datais notsignificant duetothe amountofsamples is too large. The results obtained by mathematical statistics show that the kurtosis values of the ore grade data of the fourth group and the fifth group are much more than 0. The regularity of the data is not significant, which make the training effectiveness of the data is worse. Therefore, if distributions of a large sample of the geological data are not normally distributed, the training effectiveness of the BP neural network will be not good. The results obtained by mathematical statistics show that the kurtosis values of natural logarithm values of ore grades of the fourth group samplesandthefifthgroupsamplesare0.697and 0.540 respectively. Due to the kurtosis values are close to 0, the ore grade data after taking natural logarithm are approximate normally distributed. The network models are trained by using the ore grade data after taking natural logarithm. The analysis of the training results of the fourth group samples and the fifth group samples shows that R valuesare0.5338and0.5702respectively.From

the comparison of the train results, it was known that the better fitting effect can be gotten using the ore grade data after taking natural logarithm totrain network.

IV. **CONCLUSIONS**

Mining companies' technical and economic data are studied to determine the best ways for predicting and interpolating their data.

In order to construct the mineral product pricing prediction model, we used a BP neural network. The outcomes of the predictions reveal that the prediction model is highly practicable and accurate in its predictions.

3) The geological missing data prediction and interpolation model is created utilizing geostatistics and BP neural network approaches. The model has been proven to be able to forecast and interpolate much of the geological data that is absent, and the findings are consistent.

REFERENCES

[1] Hu Nailian Ming, "The Enterprise Performance Management System for Mining Firms," Advanced Materials Research, Vols.433-440 (2012): 3276-3283, pp.3276-3283.

[2] Second International Conference on Information Science and Engineering (ICISE2010), Volume I, pp.742-747, 2010. [2] Ming Jian, Hu Nailian, "The operations plan management system of metallurgical mining firm based on Business Intelligence"

[3] 3 Using the Chaotie v-Support Vector Machine, Wu Qi and Yan Hongsen developed a product sales forecasting model. The Journal of Mechanical Engineering, 46(7):102-103, April 2010,

[4]Hu Nailian and ZuJianhua, Nonferrous Metals (Mining Section), No.6, pp.32, 1993.

[5] He Xichun and Hu Nailian report their findings in "Analysis of Mining Enterprise DSS Based on Data Warehouse," a research published in China Mining Magazine in April 2007.

Intelligent Decision System for Mining Plan, Dong Weijun, Metal Mine, No. 309, March 2002, pp. 10-16.Short-term wind speed forecast based on combined prediction, ActaEnergie Solaris Sinica, Vol. 32 No. 4, pp.543-547, April 2011.

Pan Guihao, Hu Nailian, Liu Huanzhong et al., "Empirical analysis of gold price using ARMA-GARCH model," Gold, vol. 31, no. 1, pg 5, 2010.

Hu Nailian and Liu Jiangtian, "Application of grey system theory in gold cost forecasting," China Mining Magazine, Vol. 8, No. 3, p. 65, 1999.