VOLUME-8, ISSUE-7, JULY-2019 • PRINT ISSN No. 2277 - 8160

Original Research Paper | Information Technology

ANALYSIS AND CALCULATION OF SALT MINE CAVERNS USING MACHINE LEARNING

Jasmin Katanić International Burch University, Sarajevo, Bosnia and Herzegovina, Faculty of Engineering and Natural Sciences, Department of Information Technologies

Nejdet Doğru* Ph.D., International Burch University, Sarajevo, Bosnia and Herzegovina, Faculty of Engineering and Natural Sciences, Department of Information Technologies *Corresponding Author

ABSTRACT Mining is a commodity that all countries find essential for maintaining and improving their standards of living. Mining is globally ranked as a third most profitable industry and it plays a vital role in economic development of many countries. As mining is such an important aspect of economy and is a foundation of many countries it has constantly been developed and evolved, because of these developments now there is abundance of data that can be processed. This study evaluates performance of machine learning algorithms C4.5, naive Bayes, k-NN and ensemble learning methods bagging, adaboost and voting. It has to be pointed out that research was done using only datasets composed out of data provided from a salt mine Tuzla and from boreholes located on its exploitation field. This study's aim is to predict diameters within salt caverns and to compare predicted results to actual data. In the study it was possible to achieve approximate 84% accuracy with Boosting method.

KEYWORDS : Salt mine caverns, Machine learning methods, Accuracy, Sensitivity, Specificity, Ensemble learning

INTRODUCTION

Regardless of which mineral is being mined there are always two main priorities that must be addressed. First priority is to ensure safety of people and equipment located on the extraction site. Once the safety is covered second most important thing is to improve the production process and to make it as productive as possible and to make it as cost effective as possible. Vacuum salt mining process is no different; it possesses less risk than traditional underground mining. However there is a possibility of underground caverns collapsing which could cause landslides on the surface. This can occur if condition of underground caverns is not checked and checking is done by preforming echolocation measurements. Problem with preforming these measurements is that they take time and slowdown production process a bit, and sometimes a problem can occur in between scheduled measurements.

Aim of this study is to predict diameters of underground caverns based on previous echolocation measurements, seven machine learning algorithms are used in this study. C4.5, naive Bayes, k-NN and ensemble learning methods bagging, adaboost and voting.

LITERATURE REVIEW

Use of machine learning in mining industry is not very common occurrence as one would think. However machine learning is popular in geophysics and it is gaining more popularity with each passing day. Inspiration for this study came from research conducted by Jong-Se Lim back in 2005, while study is not in the same field it proposes a similar approach and that is to use machine learning in combination with well logs [1]. From that point a citation work done by Tarek Helmy, Anifowose Fatai and Kanaan Faisal published in 2010, focused on computation of oil and gas reservoirs gave an insight into application of machine learning in mining [2]. Finally a push into a right direction was a study published by Wenlong Wang, Jianwei Ma and Fangshu Yang in 2018, which focused on salt detection with use of deep learning and calculation of salt deposits which was relatable to this study [3].

This study builds upon a previous study conducted by Salt Mine Tuzla in association with CHEMKOP institute in Krakow, Poland. Study was conducted in 1984 and was named Exploration of the underground excavations of the Tušanj salt mine in Yugoslavia. Materials from the symposium Mining of chemical raw materials, underground reservoirs, environmental protection [4]. Later in 1989 a new study was conducted named Main mining project for exploitation of salt deposit on location Tetima [5]. Study in 1989 was conducted by CHEMKOP institute and in later years institutes echolocation technology was used for measuring and controlling the production process. These early studies and projects are a key factor in this study as their data and referrals, which are still used in the field, are applied in this study.

METHODOLOGY

Dataset used in this study was constructed from 3 boreholes echolocation measurements and well logs located on the exploitation field Tetima at Salt mine Tuzla. Seven machine learning algorithms C4.5, naive Bayes, k-NN and ensemble learning methods bagging, adaboost and voting were used on the original dataset. Same algorithms and ensemble learning methods were used after outliers and extreme values were removed. Models were built using 10 fold cross validation as a testing option.

RESULTS

Table 1 shows accuracy comparison between different algorithms used on an original dataset and after outliers and extreme values have been removed. Highest accuracy has been achieved by a Boosting method when used on a original dataset, the lowest accuracy rate has been recorded when C 4.5 (decision tree) algorithm was used on an original dataset. It has to be noted that all the algorithms except for the Boosting method had a constant accuracy rate increase.

TABLE 1 ACCURACY COMPARISON BETWEEN DIFFERENT ALGORITHAMS

Classification Algorithm	Original Dataset	Removed Outliers and Extreme Values
Naive Bayes	0.667	0.673
C 4.5	0.614	0.615
k-NN	0.772	0.827
Boosting	0.842	0.807
Voting	0.654	0.712
Bagging	0.711	0.798

Table 2 shows specificity rate comparison when using different algorithms.

TABLE 2 SPECIFICITY COMPARISON BETWEEN DIFFERENT ALGORITHAMS

Classification Algorithm	Original Dataset	Removed Outliers and Extreme Values
Naive Bayes	0.425	0.435
C 4.5	0.154	0.202
k-NN	0.271	0.362
Boosting	0.307	0.300
Voting	0.447	0.350
Bagging	0.360	0.398

Table 3 shows sensitivity rate comparison when using different algorithms.

TABLE 3 SENSITIVITY COMPARISON BETWEEN DIFFERENT ALGORITHAMS

Classification Algorithm	Original Dataset	Removed Outliers and Extreme Values
Naive Bayes	0.660	0.667
C 4.5	0.971	0.971
k-NN	0.957	0.935
Boosting	0.944	0.929
Voting	0.958	0.931
Bagging	0.955	0.922

Predicted results were compared to actual echolocation measurements and this was done in such a way that predictions were made before the measurements and after measurements were conducted and interpreted by a unbiased engineer results of the measurement and predicted results were compared.

CONCLUSIONS

Comparing all the rates, the final results showed that Boosting method used on a original dataset had the best ACC rate of 0.842, best PPV rate of 0.892 was achieved by Naive Bayes when outliers and extreme values are removed. Highest NPV rate of 0.720 was achieved by Boosting method when used on a original dataset and highest ROC area score was achieved by Boosting method when it was used on a original dataset. However best comparison of predicted results was done when predicted data was compared to actual echolocation measurements. Overall best result was achieved by a Boosting method compared to actual echolocation measurements, 80 diameters were predicted correctly out of 94. All of the algorithms made correct diameter prediction in the central portion of the cavern while there were deviation in the roof of the chamber and in the bottom of the chamber. This can be linked to the fact that the central portion of the cavern has the most measurement data while the roof and the bottom of the chamber have less data.

It must be noted that in this research data of only three boreholes was used with an average chamber size of a 70 meters. It also must be noted that all the boreholes in this research had a proper chamber development without any major irregularities.

This research can be furthered and overall accuracy can be increased with the inclusion of additional attributes, which at the time of this research were not available. These attributes include previously mentioned irregular diameters developments, geological profile, extraction rate or more specifically a speed at which rack salt is melted and turned into brine. All of these attributes should contribute to increase in prediction accuracy and as such it would give a more reliable data which could be used to draw conclusions, which would impact the production process directly.

REFERENCES:

[1] Jong-Se Lim, (2005), Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea, Journal of Petroleum Science and Engineering, https:// www. sciencedirect.com/ science/article/abs/pii/S0920410505001403

[2] Tarek Helmy, Anifowose Fatai and Kanaan Faisal, (2010), Hybrid computational models for the characterization of oil and gas reservoirs, Expert Systems with Applications, https:// www.sciencedirect .com/science/article/pii/S0957417410000230

[3] Wenlong Wang, Jianwei Ma and Fangshu Yang, (2018), Salt detection with deep learning, https:/ /www. researchgate. net/publication/ 329778853_salt_detection_with_deeplearning

[4] Muhamed A, Bra ka S, Alfons R. (1984), Exploration of the underground excavations of the Tušanj salt mine in Yugoslavia. Materials from the symposium Mining of chemical raw materials, underground reservoirs, environmental protection, http://chemkop.pl/12/publikacje/

[5] CHEMKOP (1989), Main mining project for exploitation of salt deposit on location Tetima, http://chemkop.pl/12/publikacje/