FACULTY OF ENGINEERING

DEPARTMENT OF MINING ENGINEERING

FINAL YEAR PROJECT REPORT

PREDICTION OF POTENTIAL GOLD LOCATION USING MACHINE LEARNING TECHNIQUES

CASE STUDY: BUSIA DISTRICT

By

AINEMBABAZI TRUST RACHEAL

REGISTRATION NO: BU/UG/2019/0087

Contact: +256785102730

Email: ainetrustracheal4@gmail.com

A final year project report submitted to the Department of Mining Engineering as a partial fulfillment of the requirements for the award of a Bachelor of Science degree in Mining Engineering.

SEPTEMBER 2023

ABSTRACT

This study explores the use of machine learning techniques in mineral exploration, specifically in predicting potential locations of gold mineralization in Busia District, Uganda. The study uses geological data, including lithology, geophysical data, geochemical data as well and geographical data to develop and test a predictive model. The results show that the developed model can accurately predict areas with potential gold mineralization, with an accuracy of 87.5%. The study's significance lies in its potential to reduce the time and cost of identifying new gold mineralized areas, contributing to sustainable consumption and production patterns, and promoting sustained, inclusive, and sustainable economic growth. The study recommends validating the accuracy of predictive models, exploring the use of non-geological data sources, monitoring and mitigating the environmental impacts of gold mining, and sharing the findings of this study with relevant stakeholders.

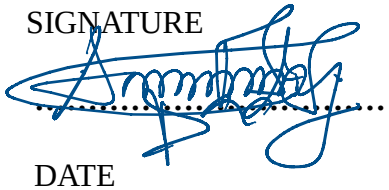
DECLARATION

I, Ainembabazi Trust Racheal, solemnly declare that this work is entirely my own and original. It has not been previously submitted for any academic award at any institution.

NAME

AINEMBABAZI TRUST RACHEAL

SIGNATURE



……

DATE

18/02/2024

……..

APPROVAL

This final year project report has been submitted to the Faculty of Engineering for examination with the consent and approval of my supervisors.

Main Supervisor

Name: Mr. KIDEGA RICHARD

Busitema University

Signature: .....

Date: 18/02/2024

Co-supervisor

Name: Mr. NASASIRA BAKAMAA MICHAEL

Department: Mining Engineering

Signature: .....

Date: 17/02/2024

ACKNOWLEDGEMENT

I would like to begin by expressing my profound gratitude to the Almighty God for blessing me with the gift of life, wisdom, knowledge, understanding, and good health, which have enabled me to embark on this academic journey.

I extend my heartfelt thanks to my parents for their unwavering support, both emotionally and financially. Their encouragement and belief in my abilities have been my driving force.

I am deeply indebted to my dedicated supervisors, Mr. Kidega Richard and Mr. Nasasira Bakamaa Michael, for their invaluable guidance, mentorship, and constructive feedback throughout the course of this project. Their expertise and willingness to invest their time in my academic growth have been instrumental in shaping the outcome of this work.

A special word of appreciation goes to my best friend, Lule David, for his selfless assistance and unwavering encouragement. His support during the challenging phases of this project has been immeasurable.

I also want to express my sincere gratitude to all my classmates for the mutual support I received from them during the final year project report compilation. Our shared experiences and collaborative efforts made this journey more manageable, and I am thankful for the camaraderie we shared.

I would also like to express my gratitude to all those who contributed in various ways, no matter how small, to the successful completion of this project. Your support has not gone unnoticed, and I am sincerely thankful for it.

Lastly, to all my colleagues, friends, and family members who stood by me during this journey, I extend my warmest thanks. Your belief in me and your words of encouragement were sources of motivation and inspiration.

This project would not have been possible without the collective support and encouragement of all those mentioned above. Thank you from the bottom of my heart.

DEDICATION

I dedicate this report to my beloved parents, Mr. and Mrs. Ruteisire, for their unwavering love, encouragement, and steadfast support throughout my academic journey. Their belief in me has been my greatest motivation, and this achievement is a testament to their dedication as parents. I thank them being my guiding light, may the Almighty God reward them abundantly.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ASGM - Artisanal and Small-Scale Gold Mining

SVM - Support Vector Machine

RF - Random Forest

ML-Machine Learning

KNN - K-Nearest Neighbors

GIS - Geographic Information System

DEM - Digital Elevation Model

ANN - Artificial Neural Network

CNN - Convolutional Neural Network

F1 - F1 Score

AUC - Area Under the Curve

GPS - Global Positioning System

ROC - Receiver Operating Characteristic

Rho…m -Resistivity

ppm -Parts per million

B..nt -Magnetism

M..Vs.Vo -Chargeability

EDA -Exploratory data analysis

IP -Induced polarization

DT -Decision tree

SOM -Self organized maps

CHAPTER ONE

1.1 Introduction

The discovery of gold deposits in Busia district has been of great interest to both local and international investors. However, the search for gold deposits is an expensive and time-consuming task that requires extensive exploration since most outcropping deposits have been discovered. Therefore, this research aimed to predict potential locations of gold in Busia district using machine learning techniques. By leveraging historical geological data and satellite imagery, a predictive model was developed to identify areas with the highest likelihood of containing gold deposits. The outcome of this research can provide valuable insights to investors and mining companies, enabling them to make informed decisions on where to invest their resources.

1.2 Background

Globally, the search for gold has been an important part of human history for thousands of years(Macdonald, 2007). The discovery of gold has fueled exploration, settlement, and economic growth throughout the world(Macdonald, 2007). However, finding new gold deposits is becoming increasingly difficult since there are limited outcrops for ease of discovery and exploitation(Yousefi & Nykänen, 2017) .Gold exploration has a long and rich history dating back to ancient times when gold was first discovered and used for decorative and ceremonial purposes. Egyptian artisans were known to use gold lavishly in decorating tombs and temples over 5,000 years ago (Harold Kirkemo, William L. Newman, 2023). Today, gold exploration remains a significant part of human civilization as gold continues to be a highly valued commodity. However, exploration success has decreased in recent decades, leading to an 80% fall in gold production since the 1970s, and a major decrease in revenue and employment in the industry (Brown & Vearncombe, 2014). Gold mineralization can be found in a variety of geological settings, from oceanic arcs to orogenic belts and back-arc basins (Torvela et al., 2022). Over the years, there have been numerous advances in mineral exploration technology, from theoretical advances in economic geology to breakthroughs in the methods or technologies for discovering and defining ore deposits (Okada, 2021).

REFERENCES

Aravind, C. (2023). *Exploring Clustering Algorithms : Explanation and Use Cases What are clustering algorithms ? Types of clustering algorithms and how to select one for your use case*. 1–37.

Babic, B., Cohen, I. G., Evgeniou, T., & Gerke, S. (2021). When machine learning goes off the rails a guide to managing the risks. *Harvard Business Review*, *2021*(January-February), 1–10.

Badr, I. (2009). *Gold Mineralization*. https://doi.org/10.13140/RG.2.2.32854.86082

Bahiru, E. A., & Woldai, T. (2016). Integrated geological mapping approach and gold mineralization in Buhweju area, Uganda. *Ore Geology Reviews*, *72*, 777–793.

BELL, C. C. (2013). *Bell, Christopher C. (2013) Structural controls on gold mineralization at the Homestake Mine and their implications for the geology of the Black Hills . PhD thesis, James Cook University.*

Brown, L., & Vearncombe, J. (2014). Critical analysis of successful gold exploration methods. *Applied Earth Science*, *123*(1), 18–24.

Caté, A., Perozzi, L., Gloaguen, E., & Blouin, M. (2017). Machine learning as a tool for geologists. *Leading Edge*, *36*(3), 215–219. https://doi.org/10.1190/tle36030215.1

Coursera. (2022). *7 Machine Learning Algorithms to Know: A Beginner's Guide*. 1–9. https://www.coursera.org/articles/machine-learning-algorithms

Das, S., Nayak, J., Kamesh Rao, B., Vakula, K., & Ranjan Routray, A. (2022). Gold Price Forecasting Using Machine Learning Techniques: Review of a Decade. *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2021*, 679–695.

Dickson, B. (2021). The future of deep learning, according to its pioneers. *TechTalks*, 1–8. https://bdtechtalks.com/2021/07/01/deep-learning-future-bengio-hinton-lecun/amp/

Eculidean. (2019). Pros and cons of common Machine Learning algorithms. *Medium*, 5–6. https://medium.com/@eculidean/pros-and-cons-of-common-machine-learning-algorithms-

45e05423264f

Elite Data Science. (2019). *Modern Machine Learning Algorithms: Strengths and Weaknesses*. 1–12. https://elitedatascience.com/machine-learning-algorithms

Ezugwu, A. E., Shukla, A. K., Agbaje, M. B., Oyelade, O. N., José-García, A., & Agushaka, J. O. (2021). Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature. *Neural Computing and Applications*, *33*(11), 6247–6306. https://doi.org/10.1007/s00521-020-05395-4

Foote, K. D. (2021). *Data Topics*. 2021.

Franklin, R. (2022). *How Data Quality Impacts Machine Learning*. https://www.precisely.com/blog/data-quality/data-quality-impact-machine-learning

Géron, A., & Russell, R. (2019). Machine learning step-by-step guide to implement machine learning algorithms with Python. *O'Reilly Media, Inc*, 106.

Ghamari, P. (2023). *Gold Prospecting and Exploration : Techniques , Tools , and Environmental Sustainability*. 4–5.

Ghosh, P. (Guha). (2018). *The Impact of Data Quality in the Machine Learning Era*. 1–3.

Ghule, R., Gadhave, A., Dubey, M., Kharade, J., & Mumbai, N. (2022). Gold Price Prediction using Machine Learning. *International Journal of Environmental Engineering*, *6*. https://doi.org/10.55041/IJSREM15027

Gold, H., Lifecycle, T., & Mine, G. (2017). *Gold Mine Exploration : 1 - 10 years Gold Mine Development : 1 - 5 years Gold Mining Operation : 10 - 30 years Gold Mine Closure and Decommissioning : 1 - 5 years Gold Mine Reclamation and Post-Closure*. 30–33.

Greenwald, H. S., & Oertel, C. K. (2017). Future directions in machine learning. *Frontiers in Robotics and AI*, *3*, 79.

Greenwood, M. (2019). Is Machine Learning the Future of Bioinformatics? *Life Sciences*, 1–4.

Guanabara, E., Ltda, K., Guanabara, E., & Ltda, K. (2019). *machine learning and data mining*.

Hammond, J. (2020). A Beginner ' s Guide to Data & Analytics. *Harvard Business School*

Online, *0*(1), 1–9.

Harold Kirkemo, William L. Newman, and R. P. A. (2023). *A brief history of gold uses, prospecting, mining and production*. 1–9.

Hassan, M., & Kotelnikov, A. (2020). *The geological and structural controls of gold mineralization in Qala en Nahal-Um Sagata region, South Gedarif, Sudan*. https://doi.org/10.21440/2307-2091-2020-3-19-26

Hillier, W. (2021). What Is a Decision Tree and How Is It Used? *Https://Careerfoundry.Com/En/Blog/Data-Analytics/What-Is-a-Decision-Tree/*, 1–12.

IBM. (2012). *1. What is a decision tree?* 3–6.

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology*, *26*(9), 1011–1013. https://doi.org/10.1038/nbt0908-1011

KPMG. (2021). *Harnessing the future of machine learning*. *October*.

Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications*, *197*, 116659. https://doi.org/https://doi.org/10.1016/j.eswa.2022.116659

Kwelwa, S. D. (2017). *Geological controls on gold mineralization in the Kukuluma Terrain, Geita Greenstone Belt, NW Tanzania.* 1–206.

Lazzaro, S. (2021). *Machine learning ’s rise , applications , and challenges*. 1–6.

Macdonald, E. H. (2007). *HANDBOOK OF GOLD EXPLORATION AND EVALUATION*.

Matters, W. I. (2021). *Machine learning, explained*. 1–12.

McGregor, M. (2020). 8 Clustering Algorithms in Machine Learning that All Data Scientists Should Know. *Freecodecamp Machine Learning*, 8. https://www.freecodecamp.org/news/8-clustering-algorithms-in-machine-learning-that-all-data-scientists-should-know/

Michaud. (2016). *Basic Gold Prospecting & Exploration Methods - Mineral Processing / Metallurgy*. 1–22. https://www.911metallurgist.com/blog/basic-gold-prospecting-exploration-methods

Miller, A. (2022). *10 new machine learning techniques for business*. *April*, 1–17.

Monolith. (2023). *Data Quality and Quantity for Machine Learning*. 1–10.

Myers, I. A., Smith, E. I., & Wyman, R. V. (1986). Control of gold mineralization at the Cyclopic Mine, Gold Basin District, Mohave County, Arizona. *Economic Geology*, *81*(6), 1553–1557. https://doi.org/10.2113/gsecongeo.81.6.1553

Neil Phillips, G. (2011). Gold exploration success. *Applied Earth Science*, *120*(1), 7–20.

Okada, K. (2021). A Historical Overview of the Past Three Decades of Mineral Exploration Technology. *Natural Resources Research*, *30*(4), 2839–2860. https://doi.org/10.1007/s11053-020-09721-4

Planner, T., Us, C., & Nilotics, T. (2013). *Pre-colonial History*. 1–4.

Porwal, A., & Carranza, E. J. M. (2015). Introduction to the Special Issue: GIS-based mineral potential modelling and geological data analyses for mineral exploration. *Ore Geology Reviews*, *71*, 477–483. https://doi.org/10.1016/j.oregeorev.2015.04.017

REID R. KEA YS, W. R. H. RAMSAY, D. I. G. (2007). *ECONOMIC GEOLOGY MONOGRAPH 6 THE GEOLOGY OF GOLD DEPOSITS : The Perspective in 1988*.

Robert, F., Brommecker, R., Bourne, B. T., Dobak, P. J., McEwan, C. .J., Rowe, R. R., Zhou, X. (2007). *Models and Exploration Methods for Major Gold Deposit Types*. 691–711.

Robert, F., Poulsen, K. H., & Dube, B. (1997). Gold Deposits and Their Geological Classification. *Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration*, 209–220.

Rodriguez, M. Z., Comin, C. H., Casanova, D., & Bruno, O. M. (2019). *Clustering Algorithms : A Comparative Approach Description of the clustering algorithms ' parameters*. 4–7.

Scikit-learn. (2023). *1.10. Decision Trees*. 1–9.

Self, G. (2018). The 5 Clustering Algorithms Data Scientists Need to Know. *Towards Data Science*, 1–7.

Sessions, V., & Valtorta, M. (2006). *The Effects of Data Quality on Machine Learning*

Algorithms.

Shahi, H., & Kamkar Rouhani, A. (2014). A GIS-based weights-of-evidence model for mineral potential mapping of hydrothermal gold deposits in Torbat-e-Heydarieh area. *Journal of Mining and Environment*, *5*(2), 79–89.

Sillitoe, R., & Hedenquist, J. (2010). Gold exploration: Deposits and methodology. *Mining Journal, Special Supplement*, *37*, 3–5.

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*(xxxx), 105524. https://doi.org/10.1016/j.asoc.2019.105524

Soares, L. (2023). *BigData Clustering Algorithms*. *21*, 21–24.

Sun, T., Chen, F., Zhong, L., Liu, W., & Wang, Y. (2019). GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China. *Ore Geology Reviews*, *109*(March), 26–49. https://doi.org/10.1016/j.oregeorev.2019.04.003

Sun, T., Li, H., Wu, K., Chen, F., Zhu, Z., & Hu, Z. (2020). Data-driven predictive modelling of mineral prospectivity using machine learning and deep learning methods: A case study from Southern Jiangxi Province, China. *Minerals*, *10*(2). https://doi.org/10.3390/min10020102

Talend. (2017). *Data quality and machine learning: what's the connection?* 1–11. https://www.talend.com/resources/machine-learning-data-quality/

Team Asana. (2021). What is decision tree analysis? 5 steps to make better decisions. *Asana*, 1–16. https://asana.com/resources/decision-tree-analysis

Torvela, T., Chapman, R., & Lambert-Smith, J. (2022). An introduction to Recent Advances in Understanding Gold Deposits: from Orogeny to Alluvium: the importance of multi-method approaches and developing a characterization. *Geological Society, London, Special Publications*, *516*(1), 1–14.

Vanneschi, L., & Silva, S. (2023). Introduction to Machine Learning. *Natural Computing Series*, 115–148. https://doi.org/10.1007/978-3-031-17922-8_5

Vogt, W. (2015). Decision Tree. *Dictionary of Statistics & Methodology*, 1–11. https://doi.org/10.4135/9781412983907.n503

Wakefield, K. (2017). Predictive Modeling Analytics and Machine Learning. *SAS Institute Inc.*, 2–5. https://www.sas.com/en_gb/insights/articles/analytics/a-guide-to-predictive-analytics-and-machine-learning.html

Wikipedia, F. (2023). *Orogenic Gold Deposit at Bombana - Indonesia*. 1–16.

Wolff, R. (2020). What Is Training Data in Machine Learning? *Monkeylearn.Com*. https://monkeylearn.com/blog/training-data/

Yousefi, M., & Nykänen, V. (2017). Introduction to the special issue: GIS-based mineral potential targeting. *Journal of African Earth Sciences*, *128*, 1–4. https://doi.org/10.1016/j.jafrearsci.2017.02.023

Yumul, G. P., Dimalanta, C. B., Maglambayan, V. B., & Tamayo, R. A. (2003). Mineralization controls in island arc settings: Insights from Philippine metallic deposits. *Gondwana Research*, *6*(4), 767–776. https://doi.org/10.1016/s1342-937x(05)71023-6