



FACULTY OF ENGINEERING AND TECHNOLOGY

DEPARTMENT OF MINING ENGINEERING

FINAL YEAR PROJECT REPORT

**COMPARATIVE ASSESSMENT OF RANDOM FOREST AND
SUPPORT VECTOR MACHINE IN GOLD VEIN PREDICTION**

CASE STUDY:

TIIRA GOLD FIELDS

BY

NAMPEELA SAIDAH SANYU

BU/UG/2021/0050

A final year project presented to the Department of Mining Engineering as partial fulfillment of the requirements for the award of a Bachelor of Science in Mining Engineering of Busitema University

APRIL 2025

ABSTRACT

Accurate forecasting of ore veins is essential in mining due to the intrinsic variability in vein geometry, grade, and orientation, often complicated by the nugget effect which challenges conventional resource estimation. Traditional geostatistical methods like kriging frequently struggle with the complex geometries and skewed grade distributions typical of vein deposits. This research evaluates the performance of machine learning algorithms, specifically Random Forest (RF) and Support Vector Machines (SVM), for predicting the presence and continuity of gold veins in the Tiira Gold Fields, Uganda. Utilizing geological survey data (gold concentration) combined with geophysical measurements (magnetic intensity, resistivity, chargeability) and geospatial coordinates, predictive models were developed following rigorous data integration, cleaning (using Inverse Distance Weighting for interpolation), and preprocessing. A comparative assessment of RF and SVM was conducted for both binary classification (vein presence > 10 ppm) and regression (gold concentration prediction). Results indicated strong performance for both models in classification, with RF achieving a higher Area Under the Curve (AUC) (0.953) but SVM showing better precision (0.909) on test data. For regression, SVM demonstrated superior performance in predicting gold concentration (Test $R^2 = 0.718$) compared to RF (Test $R^2 = 0.617$). Sensitivity analysis identified chargeability and resistivity as the most influential predictors. This study provides a robust framework for applying machine learning in mineral exploration within complex geological settings, aiding informed decision-making and potentially reducing unnecessary excavation, thereby supporting sustainable resource management.

Key words: Gold Vein Prediction, Random Forest, Support Vector Machine, Machine Learning, Mineral Exploration, Tiira Gold Fields, Geophysical Data Integration.

DECLARATION

I, NAMPEELA SAIDAH SANYU, hereby certify that the content in this report was created by me using my own time and the knowledge and that it has never been shared with a business or organization. I assume complete responsibility for it.

Date

19/05/2025

Signature



Email: saidahsanyu@gmail.com

Tel No: 0786219469/0740647871

APPROVAL

The following supervisors have given their approval for this report to be submitted to the Department of Mining Engineering.

Mr. Maseruka Bendicto

Maseruka Bendicto 20/5/25

Mr. Bagoole Christopher

Bagoole Christopher 20/5/25

ACKNOWLEDGEMENT

I genuinely thank the All-Powerful God for giving me wisdom, health, safety, and strength during the preparation of this work.

To My Parents, thank you for your unwavering encouragement, love, and support have been my driving force throughout my academic journey. Thank you for believing in me and providing a nurturing environment that fostered my growth.

I appreciate the opportunity to work with and access to valuable resources, expertise, and data that significantly contributed to the success of this project. Thank you for your commitment to innovation and research.

To my supervisors, Mr. Maseruka Bendicto and Mr. Bagoole Christopher thank you for your expert guidance, constructive feedback, and patience helped shape this project. Your dedication to mentoring is truly appreciated.

And lastly, to my fellow course mates, your camaraderie, shared experiences, and collaborative spirit made the learning process enriching. Thank you for being an integral part of my academic journey.

DEDICATION

This report is dedicated to my parents, for their excellent support throughout my career, and for instilling in me the values of courage, determination, and hard work that have contributed to my academic success and to the community of Tiira gold field for their unwavering dedication to excellence, environmental stewardship and community development. May the Almighty God lavishly bless you.

LIST OF ACRONYMS

AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC	Area Under Curve
BIF	Banded Iron Formation
CCA	Canonical Correlation Analysis
CNN	Convolutional Neural Network
CNV	Copy Number Variation
DAE	Differential Algebraic Equations
EDA	Exploratory Data Analysis
FPR	False Positive Rate
IP	Induced Polarization
KKT	Kharush Kuhn-Tucker
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multilayer Perception
MPM	Material Point Method
MSE	Mean Squared Error
NN	Neural Network

NPV Net Present Value

OK Ordinary Kriging

OOB Out-Of-Bag

PCA Principal Component Analysis

Ppm part per million

RBF Radial Basis Function

RC Reversed Circulation

RF Random Forest

ROC Receiver Operating Characteristic

ROC Receiver Operating Characteristic

SLT Statistical Learning Theory

SRM Structural Risk Minimization

SV Support Vector

SVM Support Vector Machines

SVR Support Vector Regression

TPR True Positive Rate

WOE Weights of Evidence

NS Number of Samples

NT Number of Trees

SPSS Statistical Package for the Social Sciences.

Table of Contents

ABSTRACT	ii
DECLARATION	Error! Bookmark not defined.
APPROVAL.....	Error! Bookmark not defined.
ACKNOWLEDGEMENT.....	iv
DEDICATION	v
LIST OF ACRONYMS	vi
Table of Contents	ix
CHAPTER ONE: INTRODUCTION.....	1
1.1 Background	1
1.2 Problem statement.....	2
1.3 Justification	3
1.4 Main objective.....	3
1.5 Specific objectives	3
1.6 Research Questions	4
1.7 Significance.....	4
1.8 Scope.....	4
CHAPTER TWO: LITERATURE REVIEW	6
2.1 Data collection methods.....	6
2.1.1. Data collection approaches	7
2.1.2. Data Quality and Integrity.....	7
2.1.3. Data analysis and management.....	8
2.2 Modelling of the vein	9
2.2.1. Geological structures and vein formation.	9
2.2.2. Vein geometry and orientation	10
2.2.3. Modelling approaches	11

2.2.3.1 Random Forest Model	13
2.2.3.2 Support Vector Machine	15
2.2.4. Research Gaps in modelling of the vein	18
2.2.5. Model Evaluation and comparison	18
2.2.6. Research gaps in validation	21
CHAPTER THREE: METHODOLOGY	22
3.1 Geological setting of the study area	22
3.1.1. Data Sources and Procurement	22
3.1.2. Data collected	23
3.1.3. Data Integration and Cleaning	24
3.1.4. Exploratory Data Analysis	25
3.1.5. Model prediction of the geophysical parameters	26
3.1.6. Performance Evaluation	26
3.2 Prediction of the Gold Vein Using Random Forest (RF) and Support Vector Machine (SVM) Models	27
3.2.1 Binary classification setup	27
3.2.2 Sensitivity Analysis on the prediction of gold presence	28
3.3 To evaluate and compare the accuracy of RF and SVM models.	28
3.3.1. Data preparation	28
3.3.2. Model Development	29
3.3.3. Sensitivity Analysis	30
CHAPTER FOUR: RESULTS AND DISCUSSION	32
4.1 Characterization of the existing gold vein	32
4.1.1 Descriptive analysis	32
4.1.2 Correlation Analysis of Spatial and Geophysical Data	33
4.1.2.1 Spatial Coordinates	33

4.1.2.2	Gold vs. Geophysical Variables	34
4.1.2.3	Relationships Among Geophysical Variables	34
4.1.2.4	Spatial Coordinates vs. Geophysical Variables	34
4.1.3	Model prediction of the geophysical parameters	35
4.2	Prediction of the gold vein presence using RF and SVM	37
4.2.1	Performance metrics for SVM and RF	37
4.2.2	Confusion Matrix Analysis	38
4.2.3	ROC Curves and AUC Analysis	39
4.2.4	Sensitivity Analysis of RF vs SVM for gold presence	40
4.3	Comparative Analysis of Random Forest and SVM Regression Models	43
4.3.1	Model prediction on gold concentration	43
4.3.2	Visual Prediction versus Truth	44
4.3.3	Sensitivity Analysis on Gold concentration	45
4.4	Discussion of results	47
4.4.1	Descriptive analysis	47
4.4.2	Correlation Analysis of Spatial and Geophysical Data	48
4.4.3	Model prediction of geophysical parameters	49
4.4.4	Prediction of the gold vein presence using RF and SVM	50
4.4.5	ROC Curves and AUC analysis	51
4.4.6	Sensitivity Analysis of RF VS SVM for gold presence	51
4.4.7	Model prediction on gold concentration	53
4.4.8	Sensitivity Analysis on Gold concentration	54
CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS.....		55
5.1	Conclusion	55
5.2	Recommendations	55
REFERENCES.....		57

APPENDIX: CODE REPOSITORY	60
---------------------------------	----

REFERENCES

- Abuntori, C. A., Al-Hassan, S., & Mireku-Gyimah, D. (2021). Assessment of Ore Grade Estimation Methods for Structurally Controlled Vein Deposits - A Review. *Ghana Mining Journal*, 21(1), 31–44. <https://doi.org/10.4314/gm.v21i1.4>
- Aiyelokun, O. O., & Agbede, O. A. (2021). Development of random forest model as decision support tool in water resources management of Ogun headwater catchments. *Applied Water Science*, 11(7). <https://doi.org/10.1007/s13201-021-01461-x>
- Aldrich, C., Peña-Graf, F., Órdenes, J., Wilson, R., & Navarra, A. (2022). *metals Discrete Event Simulation for Machine-Learning Enabled Mine Production Control with Application to Gold Processing*. <https://doi.org/10.3390/met>
- Cheng, Q., Chunhong, Z., & Qianglin, L. (2023). Development and application of random forest regression soft sensor model for treating domestic wastewater in a sequencing batch reactor. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-36333-8>
- Daviran, M., Maghsoudi, A., Ghezelbash, R., & Pradhan, B. (n.d.). *A new strategy for spatial predictive mapping of mineral prospectivity: Automated hyperparameter tuning of random forest 2 approach 3*.
- Deng, H., Zheng, Y., Chen, J., Yu, S., Xiao, K., & Mao, X. (n.d.). *Learning 3D Mineral Prospectivity from 3D Geological Models Using Convolutional Neural Networks: Application to a Structure-controlled Hydrothermal Gold Deposit Authorship contribution statement*.
- Dutta, S., Bandopadhyay, S., Ganguli, R., & 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*, 02(02), 86–96. <https://doi.org/10.4236/jilsa.2010.22012>
- Eichstaedt, H., Ho, C. Y. J., Kutzke, A., & Kahnt, R. (2022). Performance measurements of machine learning and different neural network designs for prediction of geochemical properties based on hyperspectral core scans. *Australian Journal of Earth Sciences*, 69(5), 733–741. <https://doi.org/10.1080/08120099.2022.2017344>

Eichstaedt, H., Ho, C. Y. J., Kutzke, A., & Kahnt, R. (2022). Performance measurements of machine learning and different neural network designs for prediction of geochemical properties based on hyperspectral core scans. *Australian Journal of Earth Sciences*, 69(5), 733–741. <https://doi.org/10.1080/08120099.2022.2017344>

Gao, L., Wang, K., Zhang, X., & Wang, C. (2023). Intelligent Identification and Prediction Mineral Resources Deposit Based on Deep Learning. *Sustainability (Switzerland)*, 15(13). <https://doi.org/10.3390/su151310269>

Gao, L., Wang, K., Zhang, X., & Wang, C. (2023). Intelligent Identification and Prediction Mineral Resources Deposit Based on Deep Learning. *Sustainability (Switzerland)*, 15(13). <https://doi.org/10.3390/su151310269>

Iglesias, C., Antunes, I. M. H. R., Albuquerque, M. T. D., Martínez, J., & Taboada, J. (2020). Predicting ore content throughout a machine learning procedure – An Sn-W enrichment case study. *Journal of Geochemical Exploration*, 208. <https://doi.org/10.1016/j.gexplo.2019.106405>

Judah, A., & Muwanguzi, B. (n.d.). *Characterisation of Muko Iron Ores (Uganda) for the Different Routes of Iron Production*.

Kong, Y., Chen, G., Liu, B., Xie, M., Yu, Z., Li, C., Wu, Y., Gao, Y., Zha, S., Zhang, H., Wang, L., & Tang, R. (2022). 3D Mineral Prospectivity Mapping of Zaozigou Gold Deposit, West Qinling, China: Machine Learning-Based Mineral Prediction. *Minerals*, 12(11). <https://doi.org/10.3390/min12111361>

Kraipeerapun, P., Fung, C. C., Brown, W., & Wong, K.-W. (n.d.). *Neural Network Ensembles using Interval Neutrosophic Sets and Bagging for Mineral Prospectivity Prediction and Quantification of Uncertainty*.

Lee, S., & Oh, H.-J. (n.d.). *5 Application of Artificial Neural Network for Mineral Potential*

Mapping. www.intechopen.com

Lee, S., & Oh, H.-J. (n.d.). *5 Application of Artificial Neural Network for Mineral Potential Mapping*. www.intechopen.com

Li, S., Liu, C., & Chen, J. (2023). Mineral Prospecting Prediction via Transfer Learning Based on Geological Big Data: A Case Study of Huayuan, Hunan, China. *Minerals*, 13(4). <https://doi.org/10.3390/min13040504>

Machuca-Mory, D. F., Munroe, M. J., & Deutsch, C. v. (2009). Tonnage Uncertainty Assessment of Vein-Type Deposits Using Distance Functions and Location-Dependent Correlograms. In *CCG Annual Report* (Vol. 11).

Mehdi Tahernejad, M., Khalokakaie, R., & Ataei, M. (n.d.). *Analyzing the effect of ore grade uncertainty in open pit mine planning; a case study of the Rezvan iron mine, Iran*.

Ndhlovu, J., & Chileshe, P. R. K. (n.d.). *Mine Productivity Issues: A Case Study of Mopani Mindola*

Subvertical Mine, Zambia. www.ijert.org

Tetteh, G. M. (2016). *Blast Induced Ore Movement Prediction Using Rock Strength Parameters- A Case Study*. 6(3). www.iiste.org

Ullah, J., Li, H., Ashraf, U., Ehsan, M., & Asad, M. (2023). A multidisciplinary approach to facies evaluation at regional level using well log analysis, machine learning, and statistical methods.

Geomechanics and Geophysics for Geo-Energy and Geo-Resources, 9(1). <https://doi.org/10.1007/s40948-023-00689-y>