SnT 2023
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE
HOFBURG PALACE - Vienna and Online
19 TO 23 JUNE

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches

S.H.Elkhouly(1) and Ghada Ali (1)
National Research Institute of Astronomy and Geophysics

INTRODUCTION

Machine learning plays a critical role in classifying between an earthquake and a nuclear explosion by providing accurate, fast, consistent, and adaptable analysis of data feature. all these techniques had been used in this work first on the base model with the data collecting. On the other hand, SMOTE (Synthetic Minority Oversampling Technique) to balanced data had been utilized in this work.

METHODS/DATA

We examine and assurance their generalizability of data by collecting from five regions (China, Pakistan, India, North Korea and USA). 80 % of data collected are utilized for training and 20 % for testing by using Machine Learning Techniques.

START

RESULTS

The best model classify between earthquakes and nuclear explosions is Random Forest Classifier at SMOTE model on the other hand, SMOTE model give low accuracy than Base models.

CONCLUSION

Finally, machine learning models that use extracted features have demonstrated considerable promise in differentiating between earthquake and nuclear explosion occurrences. However, accuracy could be hampered by imbalanced datasets. Consequently, adding SMOTE approach has greatly enhanced the classification, producing findings that are more accurate.

Please do not use this space, a QR code will be automatically overlayed

P3.5-524

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches

Introduction

- Machine learning plays a critical role in classifying between an earthquake and a nuclear explosion by providing accurate, fast, consistent, and adaptable analysis of data feature. Recent successful applications of machine lear ning to many fields of seismology (Karpatne et al., 2019; Kong et al., 2019) imply that a data-driven strategy may be appropriate for source classification challenges.
- Machine learning models with manually selected characteristics have been used in research for discrimination (Mousavi et al., 2016). Combining features gathered through machine learning with those based on physics might enhance the ability to distinguish between earthquakes and explosions (Kong et al., 2022). These investigations ar e frequently based on characteristics picked by feature and employ a machine learning algorithm to determine the categorization border that best classify between earthquakes and nuclear explosions. There is a wide range of tech niques for Machine Learning (ML) classifiers, which include decision trees, logistic regression, support vector m achines, k-nearest neighbors, XGB Classifier, Naive Bayes and random forests.
- Decision trees use a tree-like model of decisions and their potential consequences to classify instances. Logistic r egression utilizes a sigmoid function to map input features to categorical output labels. Support vector machines u se a hyperplane to separate classes in high-dimensional space (Kim et al., 2020). K-nearest neighbors classify bas ed on the similarity of an instance to its nearest neighbors. Random forests employ an ensemble of decision trees t o improve the accuracy of classification.
- Each of these techniques is advantageous in different data sets, and their effectiveness in classification tasks depe nd on various factors such as the amount and quality of data, the complexity of the problem, and the features used . So, all these techniques had been used in this work first on the base model with the data collecting. On the other hand, SMOTE (Synthetic Minority Oversampling Technique) to balanced data had been utilized in this work.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches

Objectives

- Different kernel functions of SVM are applied in the testing method to determine the difference between earthquakes and explosions. The results of the efficiency range are 70% to 95% (Kim et al., 2020).

- Based on our previous results on the methods of discrimination between earthquakes and nuclear explosions, the percentage of discrimination of complexity and spectral ratio parameters are ranging is 70%.

- The percentage of discrimination of corner frequency of P and S waves is 78% and the percentage of discrimination of body wave magnitude and surface wave magnitude is 84%.

- The above features are used to automatically discriminate using various machine learning techniques in different data sets, and their effectiveness in classification tasks depending on various factors such as the amount and quality of data, and the features used to improve the classification between earthquakes and nuclear explosions.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches

Methodology/ Data

Data Collection

For earthquakes and nuclear explosions from Five regions (China, Pakistan, India, North KOREA and USA)

Extract 3 main feature from data

Feature extract

Calculate:

Body and surface wave magnitudes (M_b & M_s)

Complexity for time domain and spectral ratio for frequency domain

Corner frequency from frequency domain

Methodology

Machine Learning for Feature

7 techniques and voting had been used in this work

1-Logistic Regression
2- SVM support vector Machine
3-KNN algorithm Classifier
4-Decision Tree Classifier
5-Random Forest Classifier
6-XGB Classifier
7- Naive Bayes
8-Voting Classifier

Events of earthquake and nuclear explosion have imbalanced Classification. So, we using SMOTE (Synthetic Minority Oversampling Technique) to balanced data

SMOTE Technique has been used for all the same models. Therefore, this investigation compares between base and SMOTE accuracy.

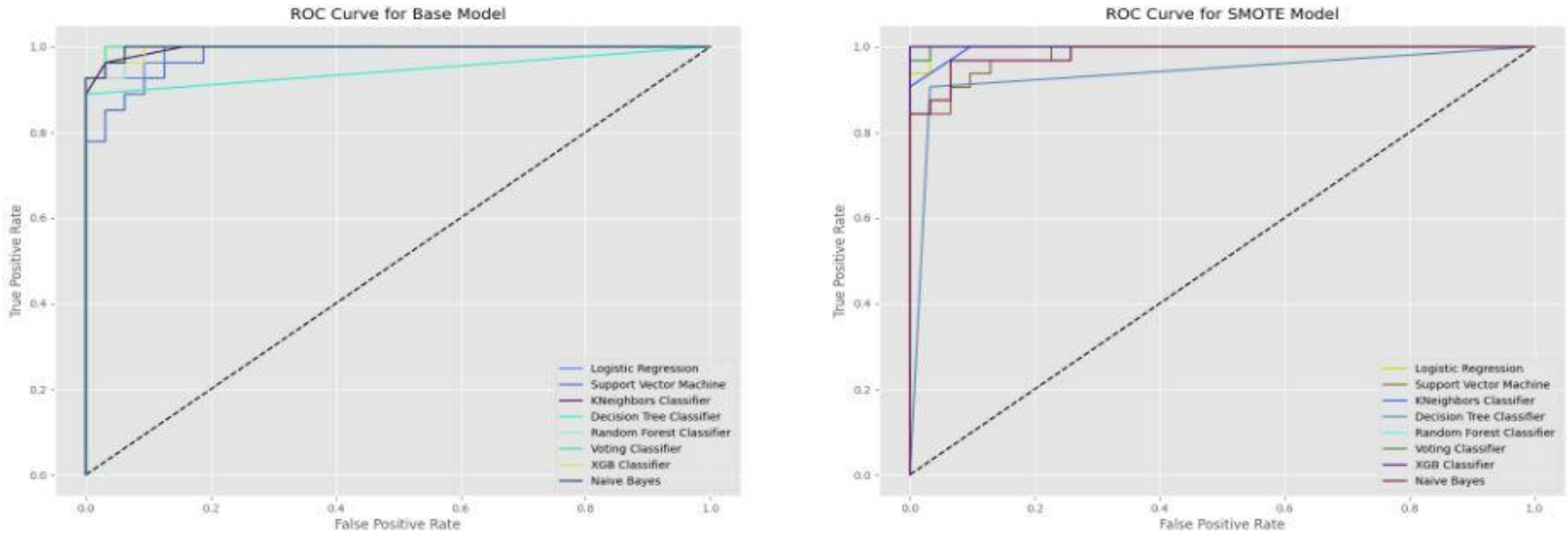
Validation and test

We examine and assurance their generalizability of data by collecting from five regions. 80 % of data collected are utilized for training and 20 % for testing.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches
Results

- We judge 16 models performance using Receiver Operating Characteristic (ROC) curves using testing data (the following Figure), which take into account both true positive and false positive rates. In our situation, explosions are good, but earthquakes are bad. The Area Under the Curve (AUC) is used to assess the ROC curve's quality, with 1 being the best.



	AUC (Base Model)	AUC (SMOTE Model)
Logistic Regression	0.9918981481481481	0.997983870967742
Support Vector Machine	0.9803240740740741	0.9818548387096774
KNeighbors Classifier	0.9953703703703703	0.9954637096774195
Decision Tree Classifier	0.9444444444444444	0.9369959677419355
Random Forest Classifier	0.9953703703703703	1.0
Voting Classifier	0.9976851851851852	0.998991935483871
XGB Classifier	0.9953703703703703	0.9999999999999999
Naive Bayes	0.9965277777777778	0.9848790322580645

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches
Conclusion

- Finally, machine learning models that use extracted features have demonstrated considerable promise in differentiating between earthquake and nuclear explosion occurrences. However, accuracy could be hampered by unbalanced datasets. Consequently, adding SMOTE approach has greatly enhanced the classification, producing findings that are more accurate specially with Random Forest Technique.

- It is clear that predictive models play a key role in identifying abnormal events and that they must be continually enhanced with an emphasis on creating high-quality training data and honed algorithms.

- Machine learning will definitely become increasingly important in identifying and classifying a wide range of environmental and social phenomena as technologies develop.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Seismic Discrimination between Nuclear explosions and natural earthquakes using multi-Machine learning Approaches

References

- Kong, Q., Wang, R., Walter, W.R., Pyle, M., Koper, K. and Schmandt, B., (2022). Combining Deep Learning With Physics Based Features in Explosion-Earthquake Discrimination. *Geophysical Research Letters*, *49*(13), p.e2022GL098645.

- Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2019). Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, **31**(08), 1544– 1554. https://doi.org/10.1109/TKDE.2018.2861006.

- Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, **90**(1), 3– 14. https://doi.org/10.1785/0220180259.

- Mousavi, S. M., Horton, S. P., Langston, C. A., & Samei, B. (2016). Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression. *Geophysical Journal International*, **207**(1), 29– 46. https://doi.org/10.1093/gji/ggw258.

- Kim, S., Lee, K., & You, K. (2020). Seismic discrimination between earthquakes and explosions using support vector machine. *Sensors*, **20**(7), 1879. https://doi.org/10.3390/s20071879.