UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierias

Automatic Seismic Event Classification Using Metaheuristic and Machine Learning Techniques

Randall Mateo Mencias Madero

Ingenieria en Ciencias de la Computacion

Trabajo de fin de carrera presentado como requisito
para la obtención del título de
Ingeniero en Ciencias de la Computacion

Quito, 19 de diciembre de 2024

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierias

HOJA DE CALIFICACIÓN
DE TRABAJO DE FIN DE CARRERA

Automatic Seismic Event Classification Using Metaheuristic and Machine Learning Techniques

Randall Mateo Mencias Madero

Nombre del profesor, Título académico Noel Pérez Pérez PH.D.

Quito, 19 de diciembre de 2024

© DERECHOS DE AUTOR

Nombres y apellidos: Randall Mateo Mencias Madero

Código: 00321469

Cédula de identidad: 1720164761

Lugar y fecha: Quito, 19 de diciembre de 2024

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en http://bit.ly/COPETheses.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on http://bit.ly/COPETheses.

RESUMEN

La detección rápida y temprana de eventos sísmicos volcánicos es fundamental para la respuesta rápida y la mitigación de desastres naturales. Los sistemas de clasificación de señales de acción rápida permiten que los mecanismos de alerta en áreas amenazadas por volcanes proporcionen advertencias tempranas, minimizando así los daños materiales y la pérdida de vidas. Este estudio propone un método de clasificación de eventos sísmicos volcánicos de dos pasos basado en la combinación de métodos metaheurísticos para encontrar el subconjunto más relevante de características y modelos de aprendizaje automático para maximizar la clasificación de eventos sísmicos. El método propuesto fue entrenado y validado en el conjunto de datos SeisBenchV1 utilizando una estrategia de división de datos de 75%´25%. El esquema de clasificación formado por el algoritmo genético con la función estadística x-cuadrado y la red neuronal de propagación hacia adelante-hacia atrás superó a los esquemas restantes, alcanzando un puntaje de área bajo la curva característica del receptor de 0,979. Este resultado destacó el desempeño exitoso de la clasificación del método propuesto y su aplicación práctica en sistemas de monitoreo sísmico volcánico.

Palabras clave: Metaheurísticas, Clasificación sísmica, Eventos volcánicos, Selección de características, Funciones de fitness, Redes neuronales, Algoritmo genético, Optimización.

ABSTRACT

The quick and early detection of volcanic seismic events is critical for the rapid response and mitigation of natural disasters. Fast-acting signal classification systems enable alert mechanisms in volcanic-threatened areas to provide early warnings, thereby minimizing material damage and loss of life. This study proposes a two-step volcano seismic event classification method based on the combination of metaheuristic methods for finding the most relevant subset of features and machine learning models for maximizing the seismic event classification. The proposed method was trained and validated on the SeisBenchV1 dataset using a data split strategy of 75%/25%. The classification scheme formed by the genetic algorithm with the x-squared statistical function and the feedforward-backpropagation neural network outperformed the remaining schemes, reaching an area under the receiver characteristic curve score of 0.979. This result highlighted the successful classification performance of the proposed method and its practical application in volcanic seismic monitoring systems.

Keywords: Metaheuristics, Seismic classification, Volcanic events, Feature selection, Fitness functions, Neural networks, Genetic algorithm, Optimization.

TABLE OF CONTENTS

Table Index

Figure Index

INTRODUCTION

Early detection of volcanic activity is crucial to mitigate natural disasters and save lives, particularly in regions threatened by active volcanoes, such as Cotopaxi in Ecuador. Cotopaxi poses a significant risk due to its proximity to densely populated areas and its history of large-scale eruptions. Early warning systems are essential to provide time for evacuations and protective measures. Recent research has focused on automatic recognition of seismic events, such as long-period events, which are key indicators of volcanic unrest.As shown by Lara-Cueva et al. [1], the automatic classification of these events in Cotopaxi has greatly improved volcanic monitoring through advanced signal processing techniques. This underscores the need for continuous improvements in these systems to ensure life-saving applications. Beyond signal processing, feature selection is another critical element in optimizing machine learning models for seismic event classification. As Venegas et al. [2] have highlighted, high-dimensional data gathered from volcanoes like Cotopaxi can make real-time processing computationally expensive. Reducing the feature space helps improve model performance while ensuring that real-time systems are efficient. These improvements are vital in enabling faster decision-making in crises, where time is of the essence. Seismic event classification is crucial in understanding tectonic and volcanic activity and is essential for improving early warning systems. Given the exponential increase in data from advanced acquisition technologies, manual seismic data analysis is no longer feasible, necessitating the use of automated systems [3]. While advancements have been made, significant challenges remain, particularly in the automatic and accurate classification of seismic events. Seismic events classification using metaheuristic techniques is a promising area of research for optimizing classification models, which could improve the current monitoring systems, ensuring faster data processing. Thus, we propose a two-step volcano seismic event classification method based on the combination of metaheuristic methods for finding the most relevant subset of features and machine learning models for maximizing the seismic event classification..

STATE OF THE ART

Cotopaxi Volcano's eruption potential is particularly concerning due to its history and the possible scale of future events. Anzieta Reyes and Ortiz Erazo explored the 2015 Cotopaxi eruption, identifying potential precursors using unsupervised machine learning techniques. Their work highlighted the importance of these tools in early detection systems, particularly through applying K-means clustering to classify spectral density and dynamic time warping. They were able to identify low-frequency events that may signal the initiation of volcanic activity [4].

Various machine learning methodologies have been applied to classify seismic events, yielding promising results across different datasets. Support Vector Machines (SVM) [5], Random Forests [3], K-means [4], and Convolutional Neural Networks (CFFBPs) [6] have been implemented to analyze seismic data from various volcanoes and monitoring networks. For example, SVM applied to seismic data from Finland achieved a remarkable 97% accuracy in classifying events [5]. Similarly, in classifying volcanic earthquakes at the Ubinas volcano in Peru, both Random Forests and SVMs achieved over 90% accuracy, even when utilizing 112 features—an unusually high number compared to other studies [3]. However, K-means, when applied to classify waveforms emitted by Cotopaxi Volcano, achieved only 23% accuracy in identifying events with high correlation, indicating the challenges of applying certain techniques in complex volcanic settings [4]. CFFBPs, used on data from Nevado del Ruiz and Laima volcanoes, yielded approximately 50% accuracy. Comparisons between active learning and random learning, in this case, showed a slight advantage for active learning [6]. Given the variety of datasets and classification methods, it is essential to consider each study's objectives and potential applications carefully.

Volcano monitoring is one of the most critical applications of seismic event classification due to the potentially devastating consequences of eruptions. In countries like Ecuador, monitoring active volcanoes such as Cotopaxi is crucial to mitigating material, human, and economic losses. Anzieta et al. [4] demonstrated using K-means clustering on frequency measurements from Cotopaxi to identify patterns preceding an eruption. This analysis lays the foundation for early warning systems, offering greater efficiency and speed than more complex models.

Currently, many studies rely on datasets with a limited number of features, often selected manually based on predefined characteristics. Malfante et al. [3] categorized their 112 features into three groups: statistical, entropic, and descriptive. However, a more scientific approach to feature selection, such as using metaheuristic methods, has not yet been fully explored in the field of seismic event classification. Given the growing volume of seismic data, it is increasingly important to identify the most suitable methodologies for real-time monitoring. Methods such as Bayesian dynamic networks [7] have been proposed but require careful consideration of time and feature selection constraints. It is crucial to identify the smallest number of relevant features to develop effective early warning systems while maximizing the performance of the model. This balance is essential to ensure that such systems respond quickly and accurately to potential volcanic activity.

MATERIALS AND METHODS

A. Experimental dataset

We used the publicly available seismic benchmark SeisBenchV1 dataset, which provides a set of feature vectors obtained from the original microseism signal and its corresponding class label, which could be long-period or volcano-tectonic seismic events. Each vector of features contains a total of 84 descriptors from the time (13 features), frequency (21 features), and scale domains (50 features), which were computed from each microseism signal recorded between January and March 2009. The frequency features were calculated using the fast Fourier transform, while the scale features were extracted from the application of the Wavelets transform. Time features, on the other hand, belong to the statistical calculation of the signal. Complete information about their calculation can be found in reference [1]. The SeisBenchV1 dataset contains 1044 and 101 samples of long-period and volcano-tectonic event classes, respectively.

B. Metaheuristic Models

Metaheuristic algorithms are generic framework algorithms that can be used in optimization problems with adequate adapting. The strategies aim to find acceptable solutions within a reasonable timeframe rather than seeking guaranteed optimal results. The fundamental principle behind these algorithms is to navigate the search space effectively by balancing exploration, which aims to discover diverse regions of potential solutions, and exploitation, which refines promising solutions for improved outcomes [8].

Metaheuristic algorithms are classified into taxonomies of diverse criteria that can be categorized in various ways. Some classifications focus on the search strategy, distinguishing between local search methods, which iteratively improve a single solution, and population-based methods, which simultaneously evolve a set of solutions. Others differentiate based on inspiration, dividing metaheuristics into nature-inspired, physics-inspired, and miscellaneous algorithms. Nonetheless, metaheuristic algorithms focus on balancing computational effort between exploration and exploitation. This balance is key to their success in solving diverse optimization problems across all research domains [9].

C. Fitness Functions

A fitness function is a mathematical expression used to evaluate how well a selected solution meets a given problem's objective. The function's problem-specific design is a measure to compare different solutions, helping the algorithm prioritize better-performing candidates [10].

D. Machine Learning Classifiers

Machine learning classifiers (MLCs) are computational algorithms designed to identify patterns in data and make predictions or decisions without human intervention [11]. These models are widely used for tasks such as classification, regression, and clustering, making them essential tools in data-driven research. Their effectiveness lies in their ability to generalize from training data to unseen data, enabling prediction in diverse fields such as healthcare, finance, engineering, etc. They are broadly categorized into three main types: supervised learning, where the model learns from labeled data; unsupervised learning, where the model identifies patterns in unlabeled data; and reinforcement learning, where the model learns by interacting with an environment to maximize cumulative rewards. [11]

E. Proposed Method

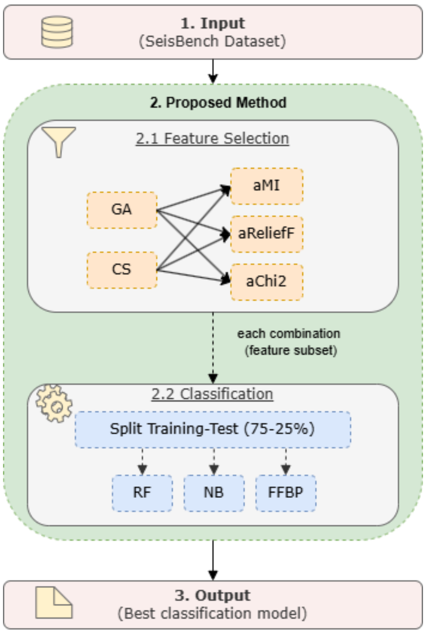

Figure 1: Proposed Method

We proposed a two-step volcano seismic event classification method based on the combination of metaheuristic methods for finding the most relevant subset of features and a set of three MLCs for maximizing the seismic event classification, as shown in Fig. 1.

The feature selection stage combined two metaheuristic methods: the genetic algorithm (GA) and the Cuckoo search model for exploring the whole feature space in conjunction with three different fitness functions, such as the average of mutual information (aMI), Relieff, and chi-squared (aChi2), to measure the merit (quality) of each provided feature subset as shown in Fig. 1, step 2.1. The selected search algorithms and fitness functions were based on different taxonomies to assess the relevance of features from different perspectives, enabling a more comprehensive exploration of the feature space. A brief description of the selected methods in this stage is next:

- Genetic Algorithm (GA): An evolutionary algorithm inspired by the process of natural selection [12]. Operates by iteratively generating and refining populations of candidate solutions through operations such as selection, crossover, and mutation. Known for its ability to explore diverse regions of the solution space effectively, making it suitable for complex optimization problems.

- Cuckoo Search (CS): A swarm intelligence algorithm inspired by the brood parasitism behavior of cuckoo birds [13]. Combines local search mechanisms with global randomization based on Lévy flights to explore the solution space efficiently. Particularly effective for identifying global optima in high-dimensional and non-linear problems.
- Mutual Information (MI): It measures the dependency between features and the target variable, quantifying how much information a feature provides about the output class [14]. Features with higher mutual information scores are more likely to contribute to improving classification accuracy.
- Chi-Square: is a statistical test used to evaluate the independence of features and the target class. Higher Chi-square scores indicate stronger associations between features and the output, making them more relevant for classification tasks [2].
- ReliefF: is an algorithm that evaluates feature importance by considering how well features differentiate between instances of different classes [15]. It is particularly effective in handling interactions between features, making it a robust choice for feature selection.

It should be noted that each fitness metric computes the merit of a single feature against the output class, and the search algorithms provide a set of features per candidate solution. Thus, we modified each metric by introducing the average function to estimate the relevance of a feature subset instead of a singular feature.

On the other hand, the classification stage considered three multi-label classifiers (MLCs): random forest (RF), naive Bayes (NB), and feedforward-backpropagation (FFBP) neural network to analyze each feature subset provided by the previous stage and maximize the seismic event classification, as shown in Fig. 1, step 2.2. These classifiers belong to different shallow learning branches to gather classification results from diverse viewpoints. An overview of the employed MLCs is next:

- NB is a probabilistic classifier based on Bayes' theorem, assuming conditional independence between features [16]. It is computationally efficient and provides a baseline for evaluating more complex classifiers. This study used the Gaussian Naive Bayes model from the scikit-learn library.
- RF is an ensemble learning method that constructs multiple decision trees during training and outputs the majority class as the final prediction [16]. Each tree is a hierarchical model that makes predictions by recursively splitting the data based on feature values. At each split, the model selects the feature and threshold that best separates the data into distinct classes. By aggregating the predictions of multiple trees, the forest can capture more complex patterns in the data while maintaining robustness to noise and outliers.
- FFBP is an artificial neural network designed to capture non-linear patterns in data by iteratively learning from errors [16]. In this model, data flows through layers of neurons in a forward direction—from the input layer to the output layer. Each neuron applies a mathematical function to its inputs and passes the result to the next layer. It later calculates the error (difference between predicted and actual output) and propagates it backward through the network to adjust the weights in the neurons.

EXPERIMENTAL SETUP

This section describes the experimental setup used to implement and evaluate the proposed method. The process is divided into distinct phases: data processing, dataset partitioning, model configuration, and performance assessment, as detailed in the following subsections.

A. Data Processing

All feature vectors were normalized using the Min-Max scaler method, which scales each feature to a range between 0 and 1. This normalization process avoids critical data dispersion while enhancing the MLCs' performance.

B. Training and Test Sets

The dataset was divided into training and testing subsets using a 75%-25% split, following the methodology proposed by Vanegas et al. [2] for a similar dataset. This partitioning ensured that a significant portion of the data was reserved for model training while maintaining a separate set for unbiased performance evaluation.

C. Metaheuristic Optimization

An optimization process was conducted for each metaheuristic to determine the best-performing metaheuristic model and its optimal parameters. The goal was to identify parameter configurations that maximized the average fitness across three distinct fitness functions: Mutual Information (MI), Chi-Square (Chi2), and ReliefF. For each metaheuristic, the parameters were tuned systematically, and the configurations yielding the highest average fitness per fitness function were selected.

The parameters of the Cuckoo Search algorithm were optimized to enhance its performance during feature selection. The following parameters were adjusted based on previous work: the number of generations was tested between 100 and 500 in steps of 50, while the number of features (dim) was fixed at 84, matching the total number of features in the dataset. The discovery rate (pa) was set at 0.25 following standard conventions, and Lévy flight parameters used a flight scale of 1.5 and a step size of 1, adhering to common implementation practices. The number of nests was fixed at 50 to ensure sufficient diversity in the search space. The optimization results indicated that the best performance for aMI and aReliefF was achieved with 500 generations, while for aChi2, 400 generations provided the best results.

The Genetic Algorithm was similarly optimized by varying its generation parameter while keeping other parameters fixed based on established conventions. The number of generations was tested between 500 and 1000 in steps of 50. The mutation rate was fixed at 0.1, and the crossover rate was set at 0.9. The population size was set to 42, representing half the number of features (dim/2), and the number of parents was fixed at 28, half the population size, following conventional practices. The optimization process revealed that the best results for aMI were obtained with 950 generations, while 800 generations were optimal for both aChi2 and aReliefF.

D. Model Configuration

The proposed method evaluated the feature subsets generated by the metaheuristic algorithms using three classifiers: Naive Bayes, Random Forest, and a Feedforward Neural Network (FFBP). The Gaussian Naive Bayes model was employed, assuming that features are conditionally independent and follow a Gaussian (normal) distribution. This classifier was implemented with the default settings provided by the scikit-learn library, serving as a baseline for comparison. The Random Forest classifier, constructed using the Classification and Regression Trees (CART) algorithm, was configured with 100 decision trees. The parameter max_features="sqrt" was utilized, allowing each tree to consider a random subset of features at each split, thereby ensuring diversity among the trees and improving generalization. The Feedforward Neural Network (FFBP) consisted of an input layer, two hidden layers, and an output layer. The input layer size corresponded to the number of features selected by the metaheuristic algorithms. The first hidden layer contained 64 neurons, and the second contained 32 neurons, both using the ReLU (Rectified Linear Unit) activation function to model non-linear relationships. The output layer was configured with a single neuron and a sigmoid activation function for binary classification. The network was trained using the Adam optimizer with the default learning rate of 0.001 and binary cross-entropy as the loss function. The training process consisted of 100 epochs and a batch size of 32, with no additional regularization techniques such as dropout applied.

E. Assessment Metrics

The performance of the classification models was evaluated using five metrics to provide a comprehensive analysis of their predictive capabilities [17]. Accuracy (ACC) measures the overall proportion of correctly classified instances, offering a general performance overview. Precision (PRE) assesses the classifier's ability to minimize false positives, which is critical in applications where false alarms are costly. Recall (REC) evaluates the proportion of true positive instances correctly identified, making it essential in scenarios where missing positive cases is highly undesirable. The F1-score provides a balanced evaluation by combining precision and recall, making it particularly valuable when both false positives and false negatives carry significant implications. Lastly, the area under the ROC curve (AUC) measures the model's ability to distinguish between classes across various decision thresholds, offering robust performance insights. By utilizing these metrics, we ensure a thorough assessment of the models' strengths and weaknesses across different aspects of classification performance.

F. Selection Criteria

The selection of the optimal metaheuristic model was based on finding the model achieving the highest Average Fitness across all the fitness functions during the optimization process. Using results from the optimization phase with parameters validated in prior studies, the chosen metaheuristic model ensured both high performance and adherence to established best practices [18], [19].

The selection and evaluation of the best models were primarily based on the AUC (Area Under the ROC Curve) metric, chosen as the main performance indicator due to its ability to assess the overall classification effectiveness across varying decision thresholds [20]. Additional metrics, such as accuracy, precision, recall, and F1-score, were used as supporting evidence to provide a comprehensive evaluation. The selection criteria focused on identifying the highest AUC score across all combinations of metaheuristics and fitness functions. For each classifier algorithm, a single combination of metaheuristic and fitness function was selected, ensuring a broad analysis of the results. This approach was designed to facilitate drawing conclusions that enhance the generalization of the experiment by considering diverse configurations and their respective performances.

G. Platform Implementation

To ensure reproducibility, all experiments were conducted using fixed random seeds for data splitting and algorithm initialization. The pandas library was used for data manipulation, while scikit-learn was employed for classifier implementation and for computing the Mutual Information and Chi-square fitness functions. The ReliefF fitness function was implemented using a nearest-neighbor approach default to the skrebate library, and the Keras library was utilized to build and train the neural network classifier. All metaheuristic algorithms, including the feature selection process, were manually implemented. The full codebase, including the scripts and configurations used in this study, is available in the following repository: [https://github.com/RandallMencias/Tesis].

RESULTS AND DISCUSSION

A. Performance evaluation of the proposed method

In evaluating the performance of the proposed method, we compared the fitness functions calculated using the full feature set without any feature selection to those obtained through metaheuristic optimization. As shown in Table II, all configurations demonstrated notable improvements with feature selection. Specifically, the average Mutual Information (aMI) values achieved with the Genetic Algorithm (GA) and Cuckoo Search (CS) algorithms were 2.656 and 2.901, respectively, significantly surpassing the baseline value of 0.039 obtained without optimization. Even the smallest gain, a 0.1 improvement with the CS algorithm using the aReliefF fitness function, underscores the positive contribution of the metaheuristic approach. This indicates that metaheuristic-driven feature selection effectively enhances the evaluation metrics, validating the effectiveness of the proposed method.

A key objective of metaheuristic algorithms is to select the most relevant features for the classification task. Each fitness function across both metaheuristics yielded similar cardinalities, with aMI resulting in 55 features for both GA and CS. In contrast, aChi2 yielded 30 and 24 features for GA and CS, respectively, and aReliefF resulted in 29 and 23 features for GA and CS, respectively. The minimal difference of six features between the metaheuristics suggests a stable selection process. While aMI was less effective at identifying relevant features compared to aChi2 and aReliefF, the latter two fitness functions demonstrated similar performance, indicating their reliability in feature selection for seismic event classification.

The primary evaluation metric, AUC, revealed that the proposed method achieved a highest value of approximately 97%, marking a significant improvement over baseline results as presented in Table I. Additionally, the accuracy metric closely approached 96%, further corroborating the model's robust classification performance. However, a more detailed analysis of other metrics, such as Recall, which did not exceed 0.78, suggests that further refinement may be necessary to achieve a more balanced performance across all evaluation criteria. This highlights the need for optimizing the model to enhance sensitivity in identifying true positive cases while maintaining high precision.

Among the evaluated configurations, two combinations exhibited the highest AUC values, identifying them as the best-performing models—one for each metaheuristic. The GA + aChi2 + FFBP combination achieved an AUC of 0.979 and an accuracy of 0.968, with a precision of 0.81 and recall of 0.684. In contrast, the CS + aReliefF + FFBP combination demonstrated a precision of 0.70 and a higher recall of 0.736, along with an AUC of 0.972 and accuracy of 0.96. These combinations effectively maximized their respective metrics while maintaining values close to the baseline and optimizing their fitness functions. Notably, the Feedforward Neural Network (FFBP) classifier consistently outperformed other classifiers in this study, likely due to its ability to model complex non-linear relationships inherent in seismic data. The slight disparities between precision and recall metrics suggest areas for further model optimization to enhance the consistency of identifying true positive cases while minimizing false positives.

B. Feature Relevance

The subset cardinality was reduced to approximately one-third of the original feature space of 84 features, with the GA + *aChi* configuration utilizing 30 features and the CS + *aReliefF* configuration employing 23 features. Despite utilizing significantly fewer features, the classifiers achieved similar or superior performance compared to those using the full feature set. This reduction in dimensionality is particularly beneficial in real-world applications, where computational resources and processing time are often constrained.

Analyzing the distribution of selected features across different domains revealed that features from the time domain consistently had the lowest number selected, followed by the frequency domain, with the majority of selected features belonging to the scale domain. Specifically, the GA combination selected 4 features from the time domain, 7 from the frequency domain, and 19 from the scale domain. In contrast, the CS combination selected 5 features from the time domain, 6 from the frequency domain, and 12 from the scale domain. Both metaheuristics selected features fairly evenly across domains, with a slight preference for the scale domain, suggesting its greater relevance to seismic events and eruptions in volcanic studies.

Additionally, certain features—f13 Time Domain Density of Peaks above RMS, f19 Frequency Domain Variance, and f59 Scale Domain Percentage of Energy D3—were consistently selected regardless of the metaheuristic used, highlighting their critical role in classification. f13 Time Domain Density of Peaks above RMS, a time domain feature, tends to exhibit higher values in VT events due to their sharp and energetic bursts, which result in numerous high-amplitude peaks, whereas LP events display lower values reflecting their more prolonged and sustained energy releases [21]. f19 Frequency Domain Variance, a frequency domain feature, shows greater variance in VT events, capturing the concentrated high-frequency energy typical of tectonic origins, compared to the more uniform frequency distribution observed in LP events. Lastly, f59 Scale Domain Percentage of Energy D3, a scale domain feature, records higher energy distribution across multiple temporal scales in LP events, indicative of their sustained and complex energy patterns, while VT events demonstrate more limited scale energy distribution. These distinctions in feature values between VT and LP events enhance the classification model's ability to accurately differentiate between the two types of seismic activities, thereby improving the overall performance and applicability of the proposed methodology in real-world seismic monitoring and analysis.

B. State of art based comparison

Establishing direct comparisons between experiments is inherently challenging due to variations in methodologies, datasets, and evaluation metrics. However, by analyzing specific components and metrics in isolation, it is possible to draw meaningful conclusions.

When comparing the results with those reported by [21], the proposed method demonstrates an improvement in outcomes for the features and metrics evaluated in common. While methodological differences exist, the classification results were sufficiently similar to permit a meaningful comparison, validating the effectiveness of the proposed approach.

In contrast, comparisons with neural network-based methods reveal that the results obtained were less favorable compared to [20]. Although the differences in methodology complicate direct comparisons, analyzing specific metrics such as the AUC allows some conclusions to be drawn. For instance, aspects like classification accuracy and feature importance suggest potential areas where the proposed method could be refined to achieve results closer to those of neural networks.

TABLE I: Performance results of the proposed method

Metaheuristic	Fitness Function	L	Subset	MLC	AUC	ACC	PRE	REC	F1
GA	aMI (iter=600)	55	$f^{2}, f^{3}, f^{5}, f^{6}, f^{7}, f^{9}, f^{11}, f^{12}, f^{14}, f^{15}, f^{17}, f^{18}, f^{19}, f^{20}, f^{21}, f^{22}, f^{23}, f^{25}, f^{27}, f^{30}, f^{32}, f^{34}, f^{36}, f^{37}, f^{38}, f^{40}, f^{42}, f^{43}, f^{44}, f^{46}, f^{48}, f^{50}, f^{51}, f^{52}, f^{55}, f^{58}, f^{60}, f^{61}, f^{62}, f^{63}, f^{64}, f^{65}, f^{67}, f^{68}, f^{71}, f^{72} f^{73}, f^{74}, f^{76}, f^{78}, f^{79}, f^{80}, f^{81}, f^{82}, f^{83}$	NB	0.849	0.947	0.583	0.736	0.651
				RF	0.832	0.961	0.722	0.684	0.702
				FFBP	0.973	0.961	0.785	0.578	0.666
	aChi2 (iter = 500)	30	$f^{9}, f^{11}, f^{12}, f^{13}, f^{16}, f^{19}, f^{21}, f^{23}, f^{28}, f^{30}, f^{34}, f^{38}, f^{39}, f^{40}, f^{48}, f^{55}, f^{58}, f^{59}, f^{60}, f^{61}, f^{65}, f^{66}, f^{70}, f^{71}, f^{72}, f^{74}, f^{75}, f^{76}, f^{77}, f^{84}$	NB	0.827	0.951	0.600	0.684	0.650
				RF	0.834	0.965	0.764	0.684	0.722
				FFBP	**0.979**	**0.968**	**0.812**	**0.684**	**0.742**
	aReliefF (iter=550)	29	$f^{2}, f^{4}, f^{9}, f^{11}, f^{14}, f^{16}, f^{17}, f^{19}, f^{21}, f^{23}, f^{25}, f^{27}, f^{38}, f^{39}, f^{42}, f^{49}, f^{55}, f^{58}, f^{62}, f^{64}, f^{65}, f^{66}, f^{67}, f^{69}, f^{71}, f^{72}, f^{76}, f^{78}, f^{81}$	NB	0.876	0.951	0.619	0.789	0.681
				RF	0.808	0.961	0.750	0.631	0.685
				FFBP	0.971	0.961	0.750	0.631	0.685
CS	aMI (iter = 500)	55	$f^{3}, f^{5}, f^{6}, f^{10}, f^{11}, f^{12}, f^{13}, f^{14}, f^{15}, f^{17}, f^{18}, f^{21}, f^{23}, f^{24}, f^{25}, f^{26}, f^{28}, f^{30}, f^{31}, f^{32}, f^{33}, f^{34}, f^{35}, f^{36}, f^{37}, f^{38}, f^{39}, f^{40}, f^{41}, f^{44}, f^{45}, f^{46}, f^{49}, f^{53}, f^{56}, f^{57}, f^{58}, f^{59}, f^{60}, f^{61}, f^{62}, f^{63}, f^{64}, f^{65}, f^{66}, f^{84}$	NB	0.874	0.947	0.576	0.789	0.666
				RF	0.834	0.965	0.764	0.684	0.722
				FFBP	0.961	0.947	0.600	0.631	0.615
	aChi2 (iter = 400)	24	$f^{1}, f^{7}, f^{10}, f^{19}, f^{20}, f^{24}, f^{26}, f^{27}, f^{30}, f^{32}, f^{34}, f^{37}, f^{38}, f^{42}, f^{58}, f^{60}, f^{61}, f^{64}, f^{65}, f^{67}, f^{68}, f^{69}, f^{72}, f^{77}$	NB	0.855	0.958	0.666	0.736	0.700
				RF	0.832	0.961	0.722	0.684	0.702
				FFBP	0.968	0.958	0.705	0.631	0.666
	aReliefF (iter = 500)	23	$f^{4}, f^{6}, f^{7}, f^{10}, f^{13}, f^{15}, f^{19}, f^{22}, f^{23}, f^{29}, f^{33}, f^{39}, f^{42}, f^{50}, f^{55}, f^{57}, f^{58}, f^{59}, f^{62}, f^{63}, f^{64}, f^{65}, f^{67}$	NB	0.855	0.958	0.666	0.736	0.700
				RF	0.812	0.968	0.857	0.631	0.727
				FFBP	**0.971**	**0.961**	**0.700**	**0.736**	**0.717**
			All Features	NB	0.946	0.943	0.608	0.777	0.682
				RF	0.981	0.965	0.800	0.631	0.705
				FFBP	0.968	0.954	0.687	0.578	0.628

L - Subset cardinality; iter - number of iterations; MLC - Machine Learning Classifier.

TABLE II: Performance Comparison of Metaheuristics and Fitness Functions

Metaheuristic	Fitness Function	Subset Features	Fitness Value
GA	*aMI*	$f^{2}, f^{3}, f^{5}, f^{6}, f^{7}, f^{9}, f^{11}, f^{12}, f^{14}, f^{15}, f^{17}, f^{18}, f^{19}, f^{20}, f^{21}, f^{22}, f^{23}, f^{25}, f^{27}, f^{30}, f^{32}, f^{34}, f^{36}, f^{37}, f^{38}, f^{40}, f^{42}, f^{43}, f^{44}, f^{46}, f^{48}, f^{50}, f^{51}, f^{52}, f^{55}, f^{58}, f^{60}, f^{61}, f^{62}, f^{63}, f^{64}, f^{65}, f^{67}, f^{68}, f^{71}, f^{72}, f^{73}, f^{74}, f^{76}, f^{78}, f^{79}, f^{80}, f^{81}, f^{82}, f^{83}$	2.656
	aChi2	$f^{9}, f^{11}, f^{12}, f^{13}, f^{16}, f^{19}, f^{21}, f^{23}, f^{28}, f^{30}, f^{34}, f^{38}, f^{39}, f^{40}, f^{48}, f^{55}, f^{58}, f^{59}, f^{60}, f^{61}, f^{65}, f^{66}, f^{70}, f^{71}, f^{72}, f^{74}, f^{75}, f^{76}, f^{77}, f^{84}$	12.119
	aReliefF	$f^{2}, f^{4}, f^{9}, f^{11}, f^{14}, f^{16}, f^{17}, f^{19}, f^{21}, f^{23}, f^{25}, f^{27}, f^{38}, f^{39}, f^{42}, f^{49}, f^{55}, f^{58}, f^{62}, f^{64}, f^{65}, f^{66}, f^{67}, f^{69}, f^{71}, f^{72}, f^{76}, f^{78}, f^{81}$	2.30
CS	*aMI*	$f^{3}, f^{5}, f^{6}, f^{10}, f^{11}, f^{12}, f^{13}, f^{14}, f^{15}, f^{17}, f^{18}, f^{21}, f^{23}, f^{24}, f^{25}, f^{26}, f^{28}, f^{30}, f^{31}, f^{32}, f^{33}, f^{34}, f^{35}, f^{36}, f^{37}, f^{38}, f^{39}, f^{40}, f^{41}, f^{44}, f^{45}, f^{46}, f^{49}, f^{53}, f^{56}, f^{57}, f^{58}, f^{59}, f^{60}, f^{61}, f^{62}, f^{63}, f^{64}, f^{65}, f^{66}, f^{67}, f^{69}, f^{70}, f^{71}, f^{72}, f^{73}, f^{74}, f^{75}, f^{77}, f^{79}, f^{82}, f^{84}$	2.901
	aChi2	$f^{1}, f^{7}, f^{10}, f^{19}, f^{20}, f^{24}, f^{26}, f^{27}, f^{30}, f^{32}, f^{34}, f^{37}, f^{38}, f^{42}, f^{58}, f^{60}, f^{61}, f^{64}, f^{65}, f^{67}, f^{68}, f^{69}, f^{72}, f^{77}$	14.451
	aReliefF	$f^{2}, f^{3}, f^{6}, f^{7}, f^{8}, f^{10}, f^{11}, f^{12}, f^{13}, f^{14}, f^{16}, f^{18}, f^{21}, f^{23}, f^{35}, f^{36}, f^{49}, f^{55}, f^{56}, f^{58}, f^{64}, f^{65}, f^{70}, f^{71}, f^{76}, f^{79}, f^{84}$	0.3099
No Metaheuristic	*aMI*	All Features	0.039
	aChi2		6.439
	aReliefF		0.224

CONCLUSIONS AND FUTURE WORK

This study successfully proposed and evaluated a two-step method that integrates metaheuristic algorithms with supervised classification to enhance the accuracy and efficiency of seismic event classification. By optimizing feature subsets using Genetic Algorithm (GA) and Cuckoo Search (CS) metaheuristics, guided by fitness functions such as *aMI*, *aChi2*, and *aReliefF*, the methodology effectively reduced the feature space while maintaining or improving classification performance. Key findings include the GA + *aChi2* + FFBP combination achieving an AUC of 0.979 and an accuracy of 0.968, and the CS + *aReliefF* + FFBP combination achieving an AUC of 0.972 and an accuracy of 0.96. These configurations not only maximized their respective metrics but also reduced the feature subset to approximately one-third of the original space, from 84 to 30 features in the GA + *aChi2* configuration. The Feedforward Neural Network (FFBP) consistently outperformed other classifiers, demonstrating its capability to model complex non-linear relationships in seismic data. Additionally, the selected features predominantly belonged to the scale domain, highlighting their significance in seismic event classification.

For future work, it is proposed to expand the use of metaheuristic algorithms and fitness functions to verify whether the optimal feature subsets obtained remain consistent across various search mechanisms. This will involve incorporating additional metaheuristic algorithms and developing new fitness functions that capture different aspects of the features. Furthermore, this expansion will enhance the analysis of feature consistency and provide a deeper understanding of their contributions to classification accuracy. By refining and broadening this methodology, its applicability and effectiveness in the fields of volcanology and seismic event classification can be further solidified.

ACKNOWLEDGEMENTS

This paper was made possible thanks to the many people who supported me in developing the idea and generously shared their time and expertise to help push this research forward. I am especially grateful to the Computer Science professors, with special recognition to Noel Pérez and Ricardo Flores, whose patience, constant questioning, and suggestions not only made this thesis better but also made it possible. Their guidance, support, knowledge, and effort were invaluable; alongside with the resources they provided, played a crucial role in the completion of this work. A heartfelt thanks also goes to the Computer Laboratory where much of the time, effort, joys, and friendships were spent and made in the finalization of this capstone project.

REFERENCES

[1] R. A. Lara-Cueva, D. S. Benítez, E. V. Carrera, M. Ruiz, and J. L. Rojo-Álvarez, “Automatic recognition of long period events from volcano tectonic earthquakes at Cotopaxi volcano,” IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 9, pp. 5247–5257, 2016.

[2] P. Venegas, N. Pérez, D. Benítez, R. Lara-Cueva, and M. Ruiz, “Combining filter-based feature selection methods and Gaussian mixture model for the classification of seismic events from Cotopaxi volcano,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 6, pp. 1991–2003, 2019.

[3] M. Malfante, M. Dalla Mura, J. I. Mars, J.-P. Métaxian, O. Macedo, and A. Inza, “Automatic classification of volcano seismic signatures,” Journal of Geophysical Research: Solid Earth, vol. 123, no. 12, pp. 10–645, 2018.

[4] J. C. Anzieta Reyes and H. D. Ortiz Erazo, “Finding possible precursors for the 2015 Cotopaxi volcano eruption using unsupervised machine learning techniques,” 2019.

[5] J. Kortström, M. Uski, and T. Tiira, “Automatic classification of seismic events within a regional seismograph network,” Computers & Geosciences, vol. 87, pp. 22–30, 2016.

[6] G. F. Manley, T. A. Mather, D. M. Pyle, D. A. Clifton, M. Rodgers, G. Thompson, and J. M. Londono, “A deep active learning approach to the automatic classification of volcano-seismic events,” Frontiers in Earth Science, vol. 10, p. 807926, 2022.

[7] C. Riggelsen, M. Ohrnberger, and F. Scherbaum, “Dynamic Bayesian networks for real-time classification of seismic signals,” in Knowledge Discovery in Databases: PKDD 2007: 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, September 17-21, 2007. Proceedings 11. Springer, 2007, pp. 565–572.

[8] M. Abdel-Basset, L. Abdel-Fatah, and A. K. Sangaiah, “Chapter 10 - metaheuristic algorithms: A comprehensive review,” in Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications, ser. Intelligent Data-Centric Systems, A. K. Sangaiah, M. Sheng, and Z. Zhang, Eds. Academic Press, 2018, pp. 185–231. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128133149000104

[9] K. Rajwar, K. Deep, and S. Das, “An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges,” Artificial Intelligence Review, vol. 56, no. 11, pp. 13 187–13 257, 2023.

[10] A. E. Eiben and J. E. Smith, Introduction to evolutionary computing. Springerisc, 2015.

[11] B. Mahesh, “Machine learning algorithms-a review,” International Journal of Science and Research (IJSR).[Internet], vol. 9, no. 1, pp. 381–386, 2020.

[12] M. Mitchell, “Genetic algorithms: An overview.” in Complex., vol. 1, no. 1. Citeseer, 1995, pp. 31–39.

[13] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, “Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems,” Engineering with computers, vol. 29, pp. 17–35, 2013.

[14] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, “Normalized mutual information feature selection,” IEEE Transactions on neural networks, vol. 20, no. 2, pp. 189–201, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128133149000104

[15] N. Spolaôr, E. A. Cherman, M. C. Monard, and H. D. Lee, “Relieff for multi-label feature selection,” in 2013 Brazilian Conference on Intelligent Systems, 2013, pp. 6–11.

[16] B. Mahesh, “Machine learning algorithms-a review,” International Journal of Science and Research (IJSR).[Internet], vol. 9, no. 1, pp. 381–386, 2020.

[17] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” Scientific Reports, vol. 14, no. 1, p. 6086, 2024.

[18] X.-N. Bui, H. Nguyen, Q.-H. Tran, D.-A. Nguyen, and H.-B. Bui, “Predicting ground vibrations due to mine blasting using a novel artificial neural network-based cuckoo search optimization,” Natural Resources Research, vol. 30, pp. 2663–2685, 2021.

[19] G. Curilem, J. Vergara, G. Fuentealba, G. Acuña, and M. Chacón, “Classification of seismic signals at Villarrica volcano (Chile) using neural networks and genetic algorithms,” Journal of Volcanology and Geothermal Research, vol. 180, no. 1, pp. 1–8, 2009.

[20] A. Salazar, R. Arroyo, N. Pérez, and D. Benítez, “Deep-learning for volcanic seismic events classification,” in 2020 IEEE Colombian Conference on Applications of Computational Intelligence (IEEE ColCACI 2020). IEEE, 2020, pp. 1–6.

[21] M. Titos, A. Bueno, L. Garc´ıa, M. C. Ben´ıtez, and J. Iba˜nez, “Detection and classification of continuous volcano-seismic signals with recurrent neural networks,” IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 4, pp. 1936–1948, 2018.

[22] A. Duque, K. Gonzalez, N. Perez, D. Benitez, F. Grijalva, R. Lara-Cueva, and M. Ruiz, “Exploring the unsupervised classification of seismic events of Cotopaxi volcano,” Journal of Volcanology and Geothermal Research, vol. 403, p. 107009, 2020.