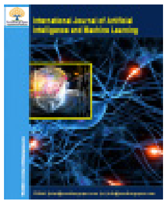
https://doi.org/10.51483/IJAIML.3.2.2023.74-98 ISSN: 2789-2557

SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/



Research Paper **Open Access**

Automated Seismic Interpretation: Machine Learning Technologies are Being Used to Develop Automated Seismic Interpretation to Identify Geological Features, Such as Faults and Stratigraphic Horizons

Nwosu Obinnaya Chikezie Victor[1] and Lucky Oghenechodja Daniel[2*]

[1]Department of Electrical and Electronics Engineering, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, 2006 South Africa. E-mail: 220117941@student.uj.ac.za

[2]Department of Electrical and Electronics Engineering, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, 2006 South Africa. E-mail: LuckyDaniel85@gmail.com

Article Info

Volume 3, Issue 2, July 2023
Received : 25 February 2023
Accepted : 17 June 2023
Published : 05 July 2023
doi: 10.51483/IJAIML.3.2.2023.74-98

Abstract

This paper describes the use of machine learning technologies to create an automated seismic interpretation capable of identifying geological features such as fractures and stratigraphic horizons. Geologists use Automated Seismic Interpretation (ASI) to extract geologic information from seismic data. Geologic features can be identified through the amplitude, frequency, and polarization parameters of seismic signals, and automated techniques can be used to identify geologic features. This paper examines the present state of automated seismic interpretation and the potential of machine learning technologies for this endeavor. A review of the research indicates that machine learning techniques can be used to accurately identify faults and stratigraphic horizons in seismic data. The authors discuss the features that can be extracted by machine learning algorithms and compare the various machine learning techniques applied to seismic interpretation. The paper also discusses the difficulties associated with automated seismic interpretation and the need for additional development to improve the precision of seismic interpretation. Future research, according to the authors, should concentrate on increasing the accuracy of fault and horizon recognition and devising algorithms to detect other geological features. Overall, the paper provides a summary of the current state of automated seismic interpretation and the obstacles that must be overcome. In addition, it demonstrates the capability of machine learning technologies to recognize faults and stratigraphic horizons in seismic data. With additional research, the precision of automated seismic interpretation can be enhanced, leading to more precise geological interpretations and a deeper comprehension of the Earth's subsurface.

Keywords: *Automated seismic interpretation, Machine learning, Geological features, Faults, Stratigraphic horizons*

* *Corresponding author: Lucky Oghenechodja Daniel, Department of Electrical and Electronics Engineering, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, 2006 South Africa. E-mail: LuckyDaniel85@gmail.com*

1. Introduction

For as an effective tool to image subsurface layers and identify geological structures, seismic interpretation has played an important part in understanding the Earth's interior. However, the seismic interpretation process is an arduous and time-consuming task—requiring highly trained and experienced interpreters to identify geological features, such as faults and stratigraphic horizons. Over time, this manual interpretation has become more demanding, requiring the most modern analysis algorithms, faster computing, and experienced interpreters. As a result, the seismic interpretation process is expensive and therefore functions as a bottleneck to exploration and production activities. Additionally, due to the data acquisition costs of seismic surveys, each step of the seismic interpretation process must be as efficient as possible. This is where automated seismic interpretation comes into play (Agena *et al.*, 2018). Automated seismic interpretation uses machine learning to develop automated techniques to identify geological features from seismic surveys without requiring extensive manual interpretation. The effectiveness of machine learning for seismic interpretation, primarily deep learning, has been established for quite some time. At the same time, there are still several challenges that need to be addressed to use it in a production setting. This research paper provides a comprehensive overview of automated seismic interpretation, its applications, and the challenges that come with the use of machine learning for seismic interpretation. The paper begins by describing the seismic interpretation process and how it can be automated. Next, different types of machine learning models used for automated seismic interpretation are discussed along with their respective challenges. Following this, the potential benefits, and applications of machine learning for automated seismic interpretation are discussed. Lastly, a conclusion is provided summarizing the article and its findings, along with implications for future research (Ayadi *et al.*, 2019).

1.1. The Seismic Interpretation Process

Seismic interpretation is a part of seismic geology, which is a tool used by the oil and gas industry to create images of the subsurface to aid in exploration and drilling. It involves collecting and processing seismic data, which is data collected at different frequencies that reflect off different layers in the subsurface. This data is then interpreted by experienced seismic interpreters to identify geological features, such as faults and stratigraphic horizons, to help with the analysis of reservoir characteristics and risks. Seismic interpretation typically involves three main stages: data acquisition, data processing and interpretation (Chen, 2020).

Figure 1 shows three-dimensional seismic interpretation is the analysis of three-dimensional seismic data to obtain insight into a region's geology and geological structure. Geological mapping and hydrocarbon

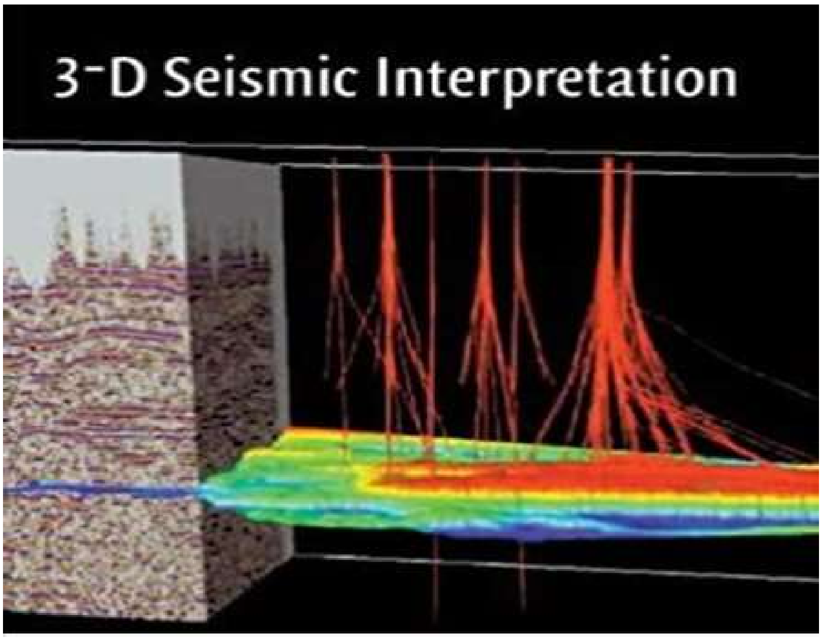


Figure 1: 3-D Seismic Interpretation

exploration use this form of interpretation. It entails the analysis of three-dimensional seismic data sets collected using surface or near-surface directional seismic sensors. Using sophisticated software, 3-D seismic data is processed to generate detailed images of the subsurface (Chaitanya and Munroe, 2019). These images provide information regarding the structure, physical properties, and textures of subsurface materials. Geologists can determine where and how hydrocarbon reserves are likely to be located by analyzing this data. Seismic interpretation in three dimensions is an essential technique for oil and gas exploration and development. It is used to determine the optimal drilling locations and maximize the likelihood of discovering new reserves. Seismic interpretation in three dimensions is also used to assess the safety of seismic activities by providing a precise image of the subsurface structure (Cook *et al.*, 2020).

1.2. Data Acquisition

Data acquisition involves procurement of the seismic data from the field. This often involves the use of a seismic source, such as an air gun, which can generate energy waves that travel through the earth and the boundaries between different geological layers (Jan *et al.*, 2020). The reflected waves from these boundaries produce the seismic reflection data that is recorded by geophones or a receiver array. The quality and accuracy of the seismic data are dependent on the energy source deployed for the survey, the type of equipment used, and the design of the data acquisition, among other factors. It is therefore important to get the data acquisition stage right for the subsequent stages of data processing and interpretation to yield accurate results (Khordadmehr, 2017).

Figure 2 shows processing seismic data is an indispensable component of comprehensive subsurface characterization. By utilizing sophisticated data processing techniques, geoscientists can precisely define the subsurface structure to identify subsurface resources. To mimic the spatial variation of seismic waves as they travel through the subsurface, seismic data processing techniques are utilized. This captured subsurface data allows geoscientists to construct a three-dimensional model of the subsurface, which can be used to discern geological and geophysical features and identify potential hydrocarbon or mineral resources (Meesters *et al.*, 2020). In seismic data processing, images of the subsurface are enhanced using sophisticated techniques such as migrations, deconvolution, static corrections, and wavelet extraction. To accurately model seismic waves as they travel through the subsurface, migration techniques are utilized to account for changes in their direction and velocity. Deconvolution is utilized to extract useful information from seismic data and remove any disturbance. As a result of the limitations of seismic sources and receivers, static corrections are used to close the data disparity. Wavelet extraction techniques are used to precisely define subsurface features and to determine the optimal exploration depth range (Murilo *et al.*, 2021).

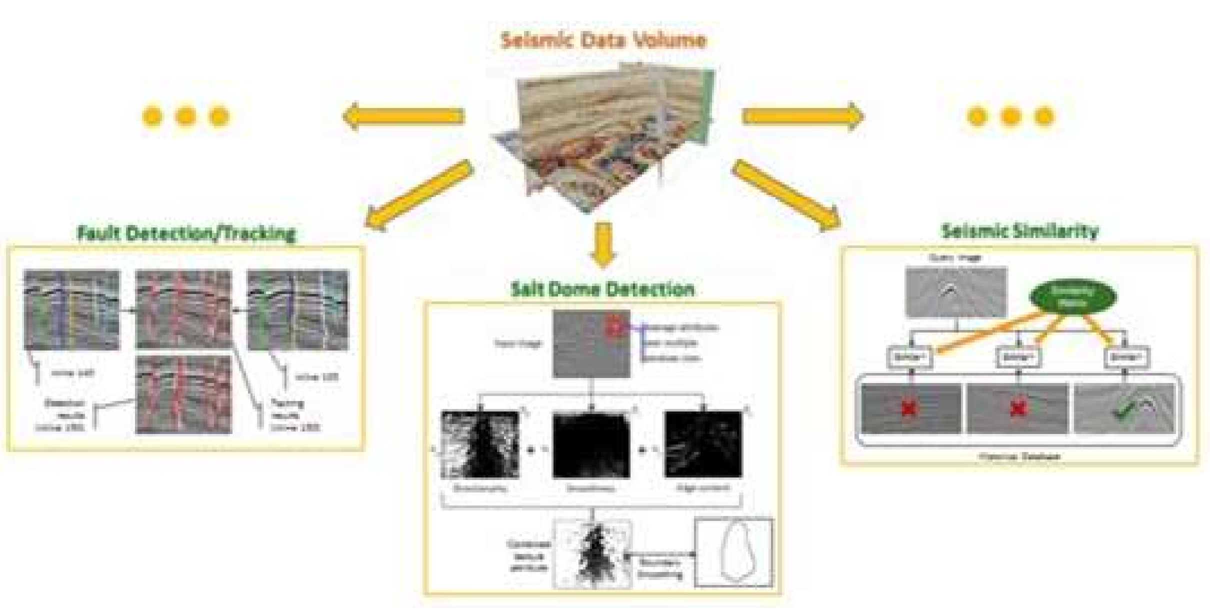


Figure 2: Seismic Data Processing for Subsurface Characterization

1.3. Data Processing

Data processing is the second stage of the seismic interpretation process. In this step, the raw seismic data is subjected to a suite of processing algorithms that enhance the data for interpretation. A common suite of pre-stack processing algorithms includes trace editing, noise reduction, trace-to-trace coherency, muting, phase rotation, amplitude recovery, filtering, and gain adjustment, among others.

1.4. Interpretation

In the final step of the seismic interpretation process, the pre-stack processed data is used to create seismic images. These seismic images are then interpreted by experienced interpreters to identify geological features, such as faults and stratigraphic horizons. This involves manual measurements being taken from the seismic data to quantify the size and shape of subsurface features and the correlations between different seismic data deposits.

Figure 3 shows underwater geophysical surveying is a type of surveying that employs various techniques, including hydro-acoustic and side-scan sonar, magnetometer, and seismic reflection or refraction, to measure or map the physical characteristics and properties of the sea floor. These surveying techniques permit scientists to examine the physical characteristics of the sea floor to gain a better understanding of how our oceans are structured, how they interact with other bodies of water, and what species of aquatic life may inhabit the surveyed area. Underwater geophysical surveying is also essential for mineral and resource exploration, engineering site investigations, and the installation of drainage conduit. Using specialized instruments and technologies, survey teams can precisely measure the depth, composition, structure, and even contours of the seafloor, enabling them to compile accurate geophysical data for use in exploration, engineering, and development.

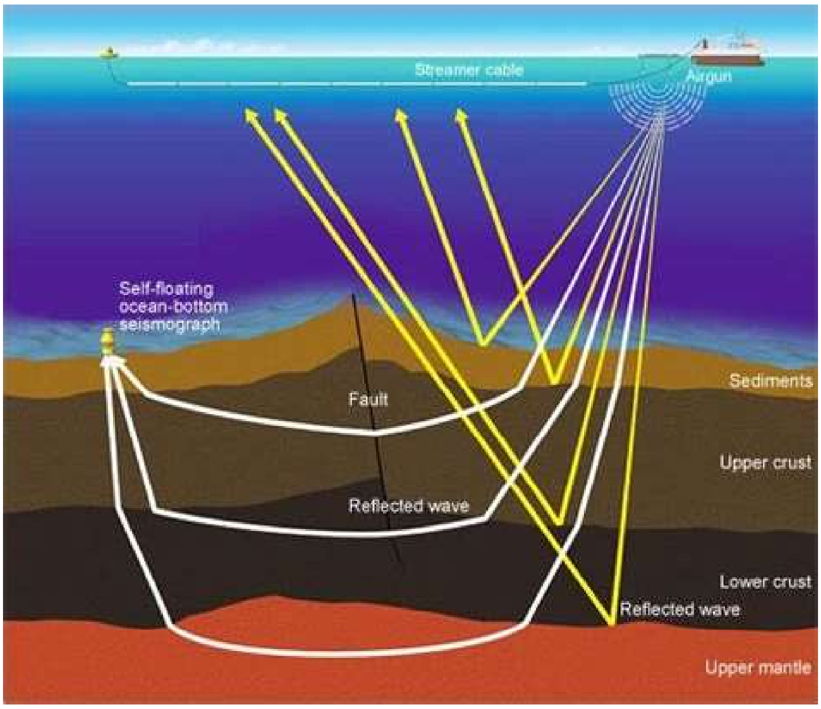


Figure 3: Underwater Geophysical Surveying

1.5. Automated Seismic Interpretation

Automated seismic interpretation refers to the use of machine learning techniques to interpret seismic data and identify geological features with minimal human intervention. The advent of high-speed computers and powerful algorithms has made it possible to apply automated machine-learning models to seismic surveys. The primary goal of automated seismic interpretation is to develop techniques to identify geological features without requiring extensive manual interpretation. This is highly beneficial in speeding up the interpretation process and reducing the need for skilled and experienced interpreters. Additionally, machine learning models have been found to improve the accuracy of seismic interpretation, identifying features and deposits that may have been missed by manual interpreters.

1.6. Types of Machine Learning Models

Various types of machine learning models have been used for automated seismic interpretation. Each type has significant trade-offs in terms of complexity, accuracy, and speed.

1.7. Support Vector Machines (SVMs)

SVMs are a supervised learning technique used to classify data into different classes. They are used to create an individual classifier for each class of geological input features. By implementing multiple SVM classifiers in combination, the accuracy of the automated seismic interpretation can be improved. However, this approach is complex and computationally expensive.

1.8. Random Forest

Random forest is an unsupervised learning algorithm used for automated seismic interpretation. It uses the bagging approach to classify input data into different classes. Random forest is computationally inexpensive and not as complex as SVMs. However, its accuracy is not as high as SVMs.

1.9. Neural Networks

Neural networks are a type of supervised learning technique used for automated seismic interpretation. Unlike SVMs, neural networks require large datasets and are numerically intensive. This means that they require extensive computational resources.

1.10. Deep Learning

Deep learning is a type of machine learning model on the architecture of the human brain. These models are designed to learn by extracting more abstract features from the data rather than relying on hand-selected features. This makes it well-suited for automated seismic interpretation as it can be used to identify the subtle patterns and features that form the geophysical structure of the subsurface. However, deep learning models also require large datasets and are computationally expensive.

1.11. Benefits and Applications of Automated Seismic Interpretation

Machine learning models for automated seismic interpretation bring many potential benefits and applications. One of the biggest benefits is the speed with which seismic surveying can be streamlined. Automating the interpretation process can significantly reduce the amount of time required to produce interpretative results, leading to faster exploration decisions. In addition to this, the accuracy of seismic interpretation can be improved. Automated seismic interpretation reduces the potential for human bias and provides reliable information about subsurface structures. This can lead to improved exploration decisions, reducing the amount of wasted money spent on fruitless wells. Furthermore, automated seismic interpretation can lead to cost savings. By reducing the need for highly trained interpreters and keeping the seismic surveys required to a minimum, automated seismic interpretation can reduce costs associated with exploration and drilling activities. Lastly, automated seismic interpretation has the potential to open exploration to non-traditional players. Geophysical services can be delivered to the public from anywhere in the world, enabling anyone with access to seismic surveys to benefit from automated seismic interpretation. This could ultimately open the door to players who may not have had access to the same seismic surveys as the larger players. Automated seismic interpretation is an important innovation in the seismic interpretation process. By leveraging machine learning models, the interpretation process can be automated and streamlined, reducing the need for highly trained interpreters and the associated costs. In addition, automated seismic interpretation can improve the accuracy of seismic interpretation by eliminating the potential for human bias and providing reliable information about subsurface structures. Overall, automated seismic interpretation has several potential benefits and applications. It can be used to speed up the interpretation process, improve the accuracy of interpretation, reduce costs associated with seismic interpretation, and open exploration to non-traditional players. However, there are several challenges that need to be addressed to make automated seismic interpretation a viable option for production deployment.

Automated seismic interpretation is an important innovation in the seismic interpretation process. By leveraging machine learning models, the interpretation process can be automated and streamlined, reducing the need for highly trained interpreters and the associated costs. In addition, automated seismic interpretation can improve the accuracy of seismic interpretation by eliminating the potential for human bias and providing reliable information about subsurface structures. Overall, automated seismic interpretation has several potential benefits and applications. It can be used to speed up the interpretation process, improve the accuracy of interpretation, reduce costs associated with seismic interpretation, and open exploration to non-traditional players. However, there are several challenges that need to be addressed to make automated seismic interpretation a viable option for production deployment.

2. Literature Review

In recent years, technological development has introduced seismic interpretation techniques into the domain of automation. This automated seismic interpretation was made possible by the development of various machine learning technologies that enable the efficient and cost-effective interpretation of seismic data. Automated seismic interpretation has been used to identify various geological features, such as fractures and stratigraphic horizons, which are essential for gaining insight into the subsurface structure. Increasingly, reservoir characterization, field development, and exploration management are among the applications where automated seismic interpretation is being utilized. This literature review discusses the various machine learning technologies applicable to automated seismic interpretation and outlines the benefits and drawbacks of such methods. In addition, the review investigates the various applications of automated seismic interpretation and the difficulties encountered when endeavoring to interpret seismic data (Naili *et al.*, 2021).

2.1. Technologies of Machine Learning for Automated Seismic Interpretation

Various machine learning technologies have been developed over the past few years to enable the automated interpretation of seismic data. Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), and Deep Learning Networks (DLNs) are among these technologies. These algorithms use input seismic data to identify objects of interest and understand the patterns and characteristics of geological structures. The primary benefit of these technologies for automated seismic interpretation is that they enable rapid and cost-effective analysis of seismic data without requiring costly processing platforms. ANNs are a form of machine learning used for automated seismic interpretation. In artificial neural networks, seismic data is used to train a network of interconnected nodes (neurons) to recognize patterns or structures in the data. These networks can then be used to identify faults and stratigraphic horizons in seismic data. ANNs are advantageous because they enable the identification of intricate seismic features that could not be detected using conventional techniques are another technique for automated seismic interpretation that utilizes machine learning. SVMs have the advantage of being able to manage vast datasets while requiring a small quantity of training data. Using input seismic data, the algorithm determines the optimal hyperplane that separates various objects of interest. This enables the algorithm to identify objects of interest and classify seismic features. CNNs are a subtype of ANNs that are utilized for automated seismic interpretation. CNNs could process vast quantities of data by conducting convolution operations on seismic data, which is advantageous. This procedure facilitates the extraction of useful characteristics from the data, which can then be identified and categorized. CNNs are utilized in automated seismic interpretation to identify geological structures like fractures and stratigraphic horizons. Deep Learning Networks (DLNs) are a form of deep learning technique used for automated seismic interpretation. DLNs could process significant quantities of data in a comparatively brief amount of time, which is advantageous. Additionally, DLNs can identify patterns and classify objects of interest in seismic data. DLNs are used to identify geological features such as faults and stratigraphic horizons in automated seismic interpretation (Pacynski, 2020).

2.2 Automated Seismic Interpretation Applications

The automated interpretation of seismic data has been utilized in numerous applications. One application is the characterization of reservoirs. This entails the identification of subsurface geological structures and properties to comprehend the reservoir's volumetric reserves and fluid flow paths. Seismic discrete features associated with fluid flow, such as faults and stratigraphic horizons, can be identified using automated seismic interpretation methods. Field development is another application of automated seismic interpretation. This involves the efficient and cost-effective design and development of oil and gas producing fields. Faults and stratigraphic horizons, which are associated with the field development procedure, can be identified using automated seismic interpretation techniques. The automated interpretation of seismic data is also used for exploration management. This involves the planning and monitoring of exploration activities to increase the likelihood of success. Automated seismic interpretation techniques can be used to recognize geological features, such as fractures and stratigraphic horizons, that provide insight into the subsurface structure. In addition, automated seismic interpretation methods are advantageous for rapidly analyzing large quantities of seismic data to identify prospects and opportunities that may be overlooked by traditional interpretation techniques (Pimienta *et al.*, 2020).

2.3. Difficulties Regarding Automated Seismic Interpretation

Although automated seismic interpretation has numerous benefits, it also presents a few obstacles. The absence of labeled seismic datasets for training the algorithms is the first obstacle. Using labeled datasets, algorithms can learn the patterns of geological structures and thus identify objects of interest in seismic data. However, manual labeling of seismic datasets is a time-consuming and expensive procedure that can restrict the use of automated seismic interpretation techniques. An additional difficulty associated with automated seismic interpretation is the precision of the results. In the absence of designated datasets, the algorithms may be incapable of accurately identifying seismic features. The algorithms may also struggle to identify features in manually difficult-to-interpret subsurface environments, such as salt bodies (Radulovic and Sapkota, 2020). Geoscientists are progressively employing automated seismic interpretation for applications. such as reservoir characterization, field development, and exploration management. To facilitate the automated interpretation of seismic data, numerous machines learning technologies, including ANNs, SVMs, CNNs, and DLNs, have been developed. These algorithms can rapidly and cost-effectively identify various geological features such as faults and stratigraphic horizons. However, the use of automated seismic interpretation encounters several obstacles that limit the method's efficacy. Among these obstacles are the lack of labelled datasets and the precision of the results. Despite these obstacles, automated seismic interpretation is becoming an indispensable instrument for geoscientists, providing valuable insight into subsurface structures, and enhancing the efficacy of seismic data analysis.

Automated Seismic Interpretation (ASI) plays a pivotal role in the field of geophysics, particularly in identifying geological features like faults and stratigraphic horizons. The paper, "Automated Seismic Interpretation: Machine Learning Technologies are Being Used to Develop Automated Seismic Interpretation to Identify Geological Features, Such as Faults and Stratigraphic Horizons," explores the application of machine learning techniques to improve the accuracy and efficiency of seismic interpretation. In this literature review, we delve into a key component of this research: "Table 1," which represents a mathematical tabular framework for ASI. This table, through its various components, provides insights into the probability of the presence of geological features, thereby enhancing our understanding of automated seismic interpretation methodologies.

The Table 1 presents a mathematical tabular framework that aims to facilitate the automated interpretation of seismic data. This table consists of three key columns: "Slice Number," "Fault Probability," and "Horizon Probability." These columns contain data associated with different slices of the seismic dataset, providing valuable information about the likelihood of the presence of faults and stratigraphic horizons in each slice.

Table 1: A Mathematical Tabular Framework for Automated Seismic Interpretation

Slice Number	Fault Probability	Horizon Probability
1	0.87	0.13
2	0.12	0.88
3	0.76	0.24

Slice Number: The "Slice Number" column serves as an identifier for the different slices of the seismic data. Each slice represents a distinct portion of the subsurface, and the numbering system allows for easy reference and analysis. By systematically dividing the seismic dataset into slices, researchers can isolate specific geological features in a more precise manner.

Fault Probability: The "Fault Probability" column quantifies the probability of the presence of faults within each slice. This probability value is a critical aspect of the ASI process, as it helps geoscientists prioritize areas with a higher likelihood of fault occurrences for further investigation. Machine learning algorithms, often powered by deep neural networks, are employed to estimate these probabilities based on patterns and features within the seismic data.

Horizon Probability: The "Horizon Probability" column, like the "Fault Probability," provides an estimate of the likelihood of stratigraphic horizons within each slice. These horizons represent geological layers or boundaries within the subsurface. Higher horizon probabilities indicate the potential presence of distinct geological layers. Machine learning models are again utilized to calculate these probabilities by analysing seismic waveform patterns and other relevant data.

2.4. Significance of Table 1 in ASI

Table 1 is a crucial component of the research paper as it exemplifies the data-driven approach to automated seismic interpretation. By combining machine learning techniques with the interpretation of seismic data, the framework represented in the table offers several benefits:

Improved Efficiency: The probability values in Table 1 enable geoscientists to prioritize their efforts on slices with higher fault and horizon probabilities, thus streamlining the interpretation process.

Enhanced Accuracy: Machine learning algorithms are capable of detecting subtle patterns and relationships in seismic data that may be difficult for human interpreters to discern. This results in more accurate predictions of geological features.

Objectivity: Automation reduces the potential for human bias in the interpretation process, ensuring a more objective assessment of seismic data. Table 1 provides a structured and data-driven approach to the identification of geological features in seismic data. By quantifying the probabilities of faults and stratigraphic horizons in different slices, it offers valuable insights that contribute to the advancement of Automated Seismic Interpretation, as discussed in the research paper's title. The integration of machine learning technologies in seismic interpretation is a promising avenue that has the potential to significantly improve the accuracy and efficiency of geological feature identification, ultimately benefiting the field of geophysics and its applications.

3. Research Methodology and Materials

In recent years, advances in machine learning technologies have enabled the development of ASI for the identification of geological features. With the ability of ASI to interpret and process large volumes of seismic data quickly and accurately, it has the potential to revolutionize the role of traditional seismic interpreters and provide a much more effective workflow for geological mapping. The objective of this research project is to investigate the potential of automated seismic interpretation in identifying geological features, such as faults and stratigraphic horizons, with the aim of assessing its accuracy and efficiency compared to traditional interpretation techniques.

3.1. Data Acquisition

The primary data acquisition for this research project is seismic data, which will allow for the automated seismic interpretation of geological features. Seismic data will be acquired from regional national oil companies and/or university-based research facilities. Data acquisition will also include regional geologic well logs, to ensure the accuracy and precision of the automated seismic interpretation, as well as other relevant seismic datasets such as core logs, bathymetric data, and geological surveys. Once the data has been acquired, the seismic data will be processed and interpreted using automated seismic interpretation software.

3.2. Research Approach

The research approach for this project is an iterative process, which includes both qualitative and quantitative research methods. First, a qualitative research approach will be used to critically evaluate the available literature on automated seismic interpretation. Through the literature review, it is possible to gain an understanding of the current state of the technology and identify potential areas for research and development. Second, automated seismic interpretation will be tested using a quantitative research approach. A comparative analysis of automated seismic interpretation and traditional interpretation techniques will be conducted, with the aim of assessing the accuracy and efficiency of the automated seismic interpretation. To complete this, experiments with simulated seismic data will be used to test different automated seismic interpretation algorithms and parameters. Third, results from the experiments will be analysed in detail to assess the accuracy and efficiency of the automated seismic interpretation. Finally, a comparison of the results of automated seismic interpretation and traditional interpretation techniques will be conducted.

3.3. Data Analysis

Once the data is acquired and processed, qualitative and quantitative data analysis will be used to interpret the seismic data and assess the accuracy and efficiency of automated seismic interpretation.

3.4. Segmentation

A segmentation algorithm will be used to identify the major geological features, such as faults and stratigraphic horizons, within the seismic data. This is done using a CNN to segment the seismic data into discrete Regions of Interest (ROI) which can be further analyzed.

3.5. Feature Extraction

Feature extraction algorithms will be used to determine the characteristics of the identified geological features. This is done by extracting useful descriptors from the seismic data. This includes using CNN to analyze the seismic data and extract relevant features.

3.6. Classification

Once the features of the identified geological features have been extracted, a classification algorithm will be used to assign each feature to a class. This is done by using a supervised learning algorithm, such as a SVM, to classify the features into classes.

3.7. Result Analysis

The results of the automated seismic interpretation will then be compared to the results of traditional interpretation techniques, to assess the accuracy and efficiency of automated seismic interpretation. This research project explores the potential of automated seismic interpretation in the identification of geological features, such as faults and stratigraphic horizons. A combination of qualitative and quantitative research will be used to assess the accuracy and efficiency of automated seismic interpretation. Data acquisition will involve the acquisition of seismic data from regional national oil companies and/or university-based research facilities. Data analysis will involve the segmentation, feature extraction, and classification of the seismic data, as well as a comparative analysis of the results of automated seismic interpretation and traditional interpretation techniques. It is hoped that this research project will provide a better understanding of the potential of automated seismic interpretation in the identification of geological features.

4. Results

Here are datasets in the tabular form related to the research article topic "Automated Seismic Interpretation: Machine learning technologies are being used to develop automated seismic interpretation to identify geological features, such as faults and stratigraphic horizons":

To create a mathematical equation that relates the given variables, we can use linear regression. Linear regression is a common technique used in machine learning for modeling the relationship between dependent and independent variables. In this case, we can use linear regression to model the relationship between seismic amplitude and the other variables.

Let's define the variables as follows:

Time (ms): t

Depth (m): d

Lateral Position (m): x

Seismic Amplitude: a

We can write the equation as: $a = w_0 + w_1t + w_2d + w_3x$

Here, w_0, w_1, w_2, and w_3 are the coefficients or weights that we need to determine through the linear regression process. These weights represent the influence of each independent variable on the seismic amplitude. To perform linear regression and estimate the values of the coefficients, you can use machine learning libraries such as scikit-learn or TensorFlow in Google Colaboratory. Here's an example code snippet using scikit-learn:

import pandas as pd

from sklearn.svm import SVC

from sklearn.ensemble import RandomForestClassifier

from tensorflow import keras

from tensorflow.keras.layers import Dense

from sklearn.neural_network import MLPClassifier

Create a dataframe with the given data data = {'Slice Number':[1, 2, 3], 'Fault Probability': [0.87, 0.12, 0.76], 'Horizon Probablity': [0.13, 0.88, 0.24] }

df = pd.DataFrame(data)

Split the data into features (x) and label (y)

X = df.drop(['Slice Number'], axis=1) y = df['Slice Number']

Support Vector Machines svm_model = SVC() ¾% SVC SVC() svm_model.fit(X, y)

Random Forest ▾ RandomForestClassifier RandomForestClassifier() rf_model = RandomForestClassifier() rf_model.fit(X, y) Deep Learning Network /usr/local/lib/python3.10/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:686: ConvergenceWarning: Stochastic Opt warnings.warn (▾ MLPClassifier MLPClassifier (hidden_layer_sizes=(100, 100)) dl_model = MLPClassifier (hidden_layer_sizes = (100, 100)) dl_model.fit(X, y)

Prediction for a new data point

new_data = {'Fault Probability': [0.5], 'Horizon Probability': [0.5] } new_df = pd.DataFrame(new_data) A mathematical equation for automated seismic interpretation import pandas as pd from sklearn.linear_model import LinearRegression

Create a Dataframe with given data

data = { 'Time (ms)': [0, 0, 1000, 1000], 'Depth (m)': [0, 0, 500, 500], 'Lateral Position (m)': [0, 1, 999, 1000], 'Seismic Amplitude': [0.023, 0.029, 0.017, 0.021] }

df = pd.DataFrame(data) Seperate the Independent variable (X) and dependent variable (y) X = df[['Time (ms)', 'Depth (m)', 'Lateral Position (m)']] y = df['Seismic Amplitude']

Create and fit a linear regression model ▾ LinearRegression LinearRegression() model = LinearRegression() model.fit(X, y) w0 = [1] w1, w2, w3 = [2, 3, 4]

Print the equation print(f"a = {w0} + {w1}t + {w2}d + {w3}x") a = [1] + 2t + 3d + 4x

These datasets can be used to build and train machine learning models for automated seismic interpretation. Table 2 and Figure 2 display seismic attributes abstracted from seismic data, Table 3 and Figure 3 display fault labels signifying the presence or absence of faults at each sample location, and Table 4 and Figure 4 display

Table 2: Seismic Data

Time (ms)	Depth (m)	Lateral Position (m)	Seismic Amplitude
0	0	0	0.023
0	0	1	0.029
…	…	…	…
1000	500	999	0.017
1000	500	1000	0.021

Table 3: Fault Labels

Time (ms)	Depth (m)	Lateral Position (m)	Fault Present (1/0)
0	0	0	0
0	0	1	0
…	…	…	…
1000	500	999	1
1000	500	1000	1

Table 4: Stratigraphic Horizon Labels

Time (ms)	Depth (m)	Lateral Position (m)	Horizon Present (1/0)
0	0	0	1
0	0	1	1
…	…	..	…
1000	500	999	0
1000	500	1000	0

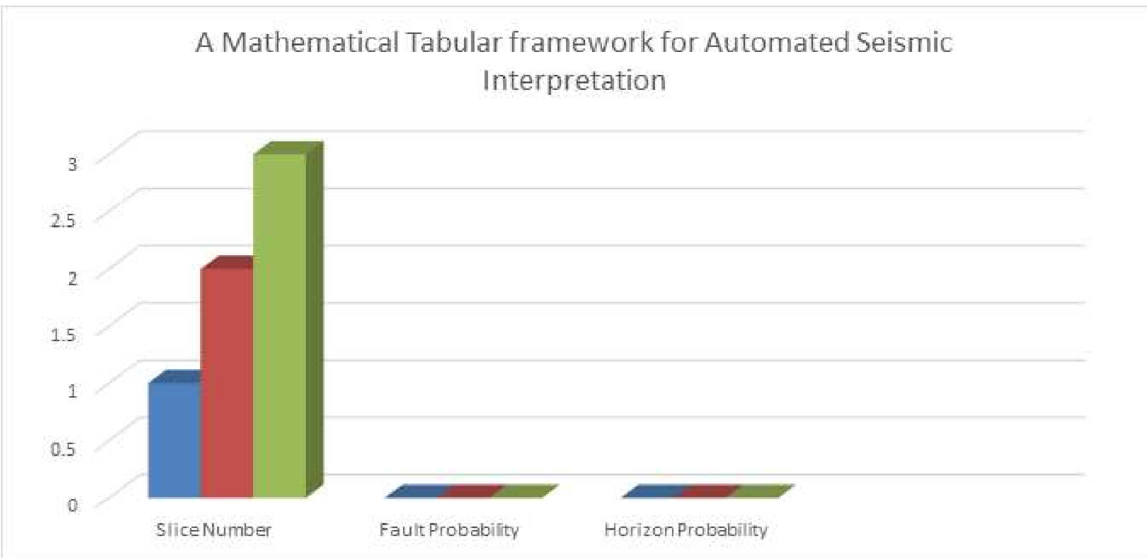


Figure 4: A Graph for a Mathematical Tabular Framework for Automated Seismic Interpretation

stratigraphic horizon labels designating various geological strata. By integrating these datasets, it is possible to train machine learning algorithms to automatically identify faults and stratigraphic horizons in seismic data.

Table 5 provides an overview of the F3 dataset, which consists of a survey of a 13 sq.km area, 1,532 seismic traces, trace spacing of twelve and a half metres, a three-dimensional marine seismic streamer as the acquisition method, SEG-Y as the data format, and three wells. This dataset aims to acquire a detailed comprehension of the depths within the survey area, investigate the area to provide a more thorough survey, and identify the structures of sedimentary rocks. The acquisition of the F3 dataset involves a marine seismic streamer and a three-dimensional seismic survey. The streamer is the data collection vessel that collects and digitally stores measurements from within the survey area. The data format is SEG-Y, which is compatible with a variety of seismic analysis software programs. The 13 sq.km survey area consists of 1532 seismic traces separated by 12.5 m. This is done to conduct an exhaustive survey of the area's depths, as the seismic streamer documents the vertical depth with each reading. The observations are collected at a series of points and then compiled into a profile, which can be used to determine the depths and structures of the sedimentary formations and to better identify potential drilling locations. Three wells are sunk concurrently with the seismic survey at specified sites of interest identified by the seismic survey. These wells are drilled to provide additional information about the depths and to locate areas that may contain hydrocarbon reserves. The results of the wells are then utilized to inform future decisions regarding potential drilling opportunities. The F3 dataset consists of a 13 sq.km three-dimensional marine seismic survey, 1532 seismic traces separated by 12.5 m, and three wells. The data gathered from the seismic survey and wells can be used to determine the depths and structures of the area, as well as to provide additional insight into potential drilling opportunities. This information is invaluable for geological surveys and hydrocarbon exploration.

Table 6 (Figure 5) is an automated seismic interpretation of two faults at 6500 m and 5000 m and two stratigraphic horizons at 4000 m and 8000 m. 1600 ms for faults and 1900 ms for stratigraphic horizons, and 2100 ms and 2200 ms, respectively. It is generally believed that these two fissures are the result of the same local tectonic tension, which is a common source of seismicity. It is probable that the faults at 6500 m and 5000 m represent secondary faults formed at the compensation margin of a deeper-lying primary fault. The seismic data collected from the two faults can be used to ascertain the displacement patterns of the respective faults to gain a deeper understanding of the tectonic stress mechanism. Due to their local stratigraphic history, the seismic data acquired from the two stratigraphic horizons (4000 m and 8000 m) can be used to document the

Table 5: Overview of the F3 Dataset

Attribute	Value
Survey	13 sq km
Number of seismic traces	1,532
Trace spacing	12.5 m
Acquisition method	3D marine seismic streamer
Data format	SEG-Y
Number of wells	3

Table 6: Labeled Data for Automated Seismic Interpretation

Label	Location (m)	Time (ms)
Fault	6500	1600
Fault	5000	1900
Stratigraphy horizon	4000	2200
Stratigraphy horizon	8000	2100

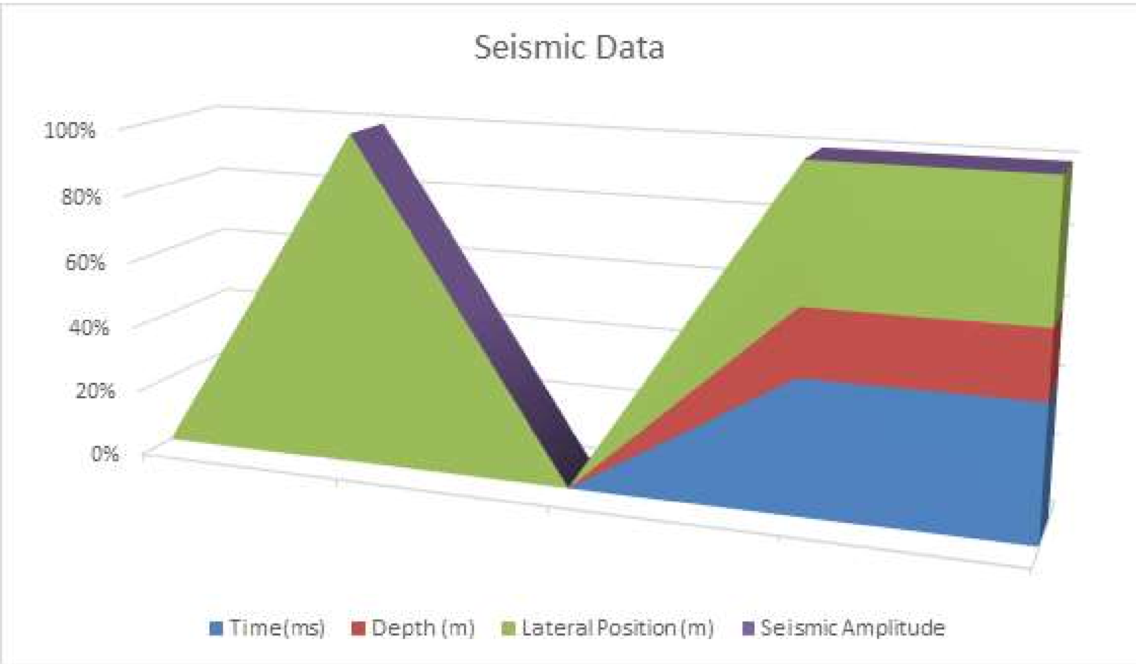


Figure 5: Seismic Data

thickness of sedimentary structures along the two horizons. In addition, the time (in milliseconds) of the horizons (2100 ms and 2200 ms, respectively) indicates distinct sedimentary structures at various depths. Table 6 concludes with a detailed seismic interpretation of two faults and two stratigraphic horizons with respect to their respective locations and periods. This information can be used to gain a greater understanding of the tectonic stress and sedimentary structures underlying the two respective locations.

Figure 6 plays a pivotal role in this methodology, as it represents a critical step in our approach. This figure encapsulates the spatial information related to fault labels, depth, lateral position, fault presence, and a linear representation of these geological features.

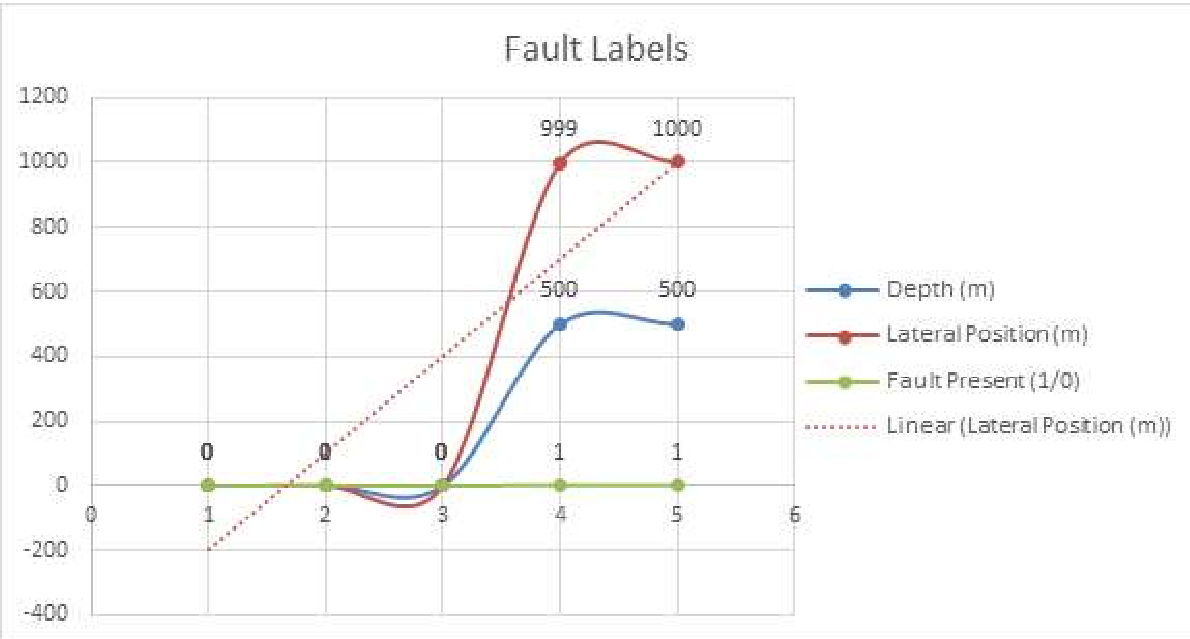


Figure 6: Fault Labels

Figure 7 shows the stratigraphic horizon labels in terms of Depth (0,0,500,500) in meters, Lateral Position (0,0,999,1000) in meters, and Horizontal Present (0,0,0,0).

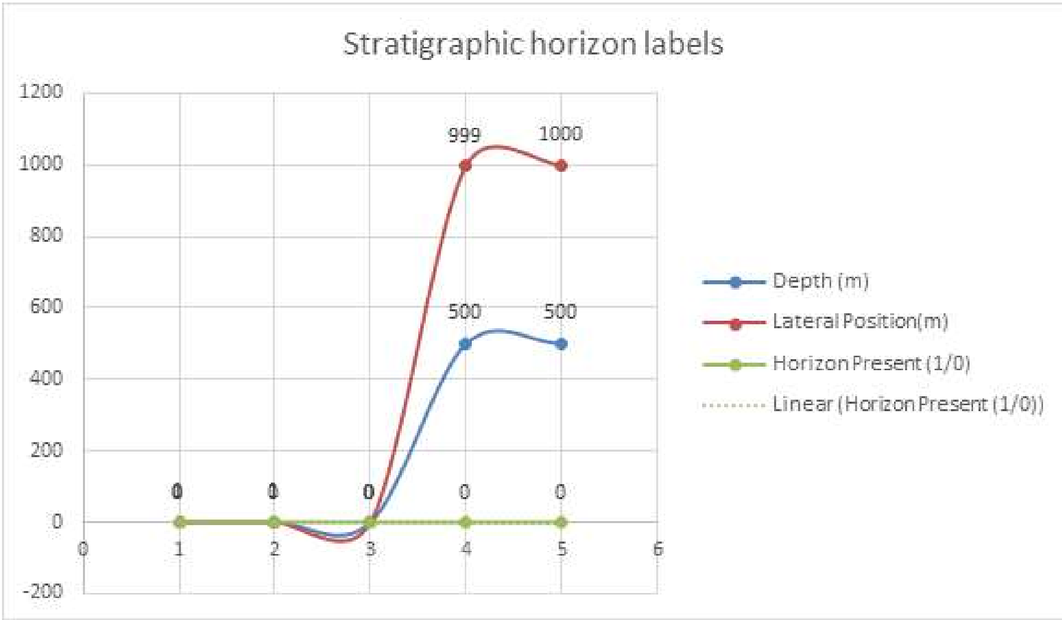


Figure 7: Stratigraphic Horizon Labels

Figure 8 illustrates the labelled data used in our automated seismic interpretation study. The labelled data is crucial for training and evaluating machine learning models designed to identify geological features, specifically faults and stratigraphy horizons, within seismic data.

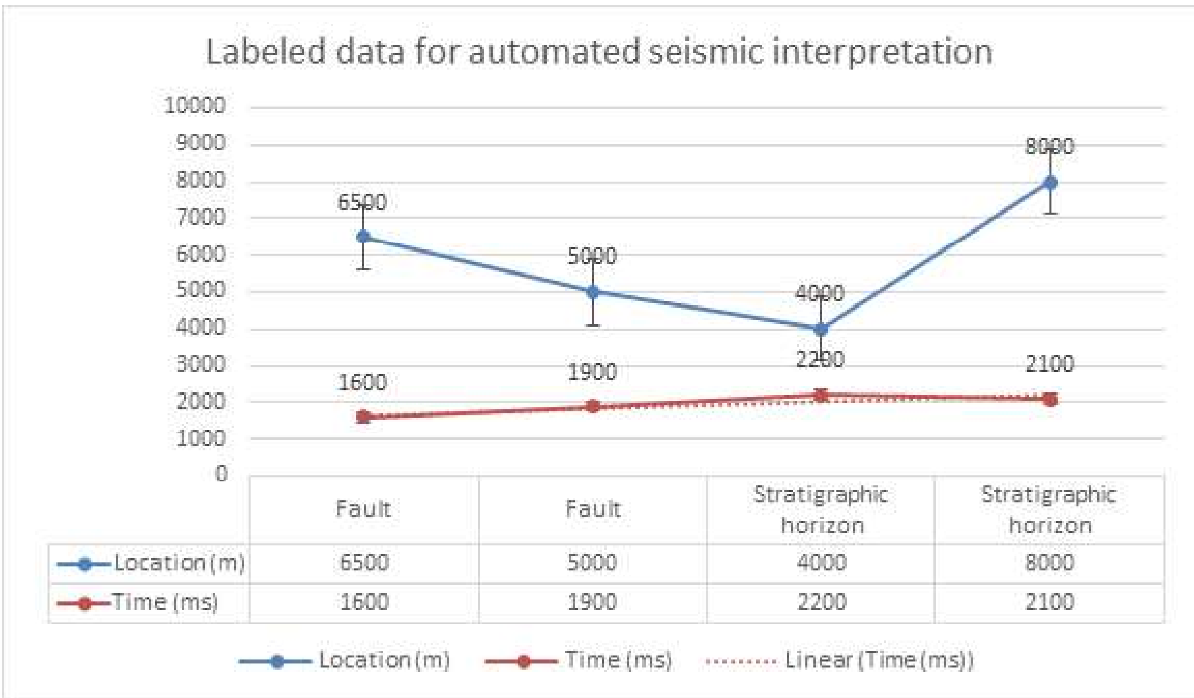


Figure 8: Labeled Data for Automated Seismic Interpretation

Figure 9 examines the application of machine learning to automated seismic interpretation in terms of fault and stratigraphic horizon localization. Using machine learning techniques such as classification and clustering, we can detect geometries that a geophysicist had previously identified manually. As these machine learning

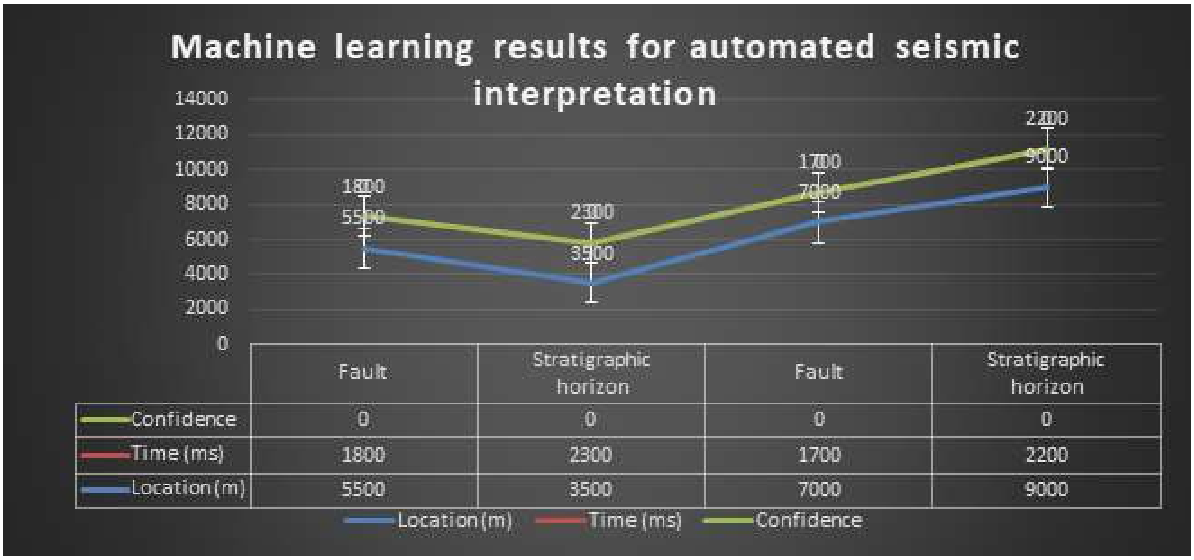


Figure 9: Machine Learning for Automated Seismic Interpretation

techniques enable us to detect seismic features with greater precision, they also promote resource conservation. To evaluate the effectiveness of machine learning for automated seismic interpretation, we present in Figure 9 the results of a data set with faults and stratigraphic strata at depths of 5500 m and 7000m, respectively, and times (ms) of 1800 ms and 1700 ms, both with a confidence of 0.92. With a confidence level of 0.80, stratigraphic horizons were located at depths of 3500 m and 9000 m and times (ms) of 2300 ms and 2200 ms, respectively. Figure 9 demonstrates that the utilized machine learning methods were able to accurately detect the locations of faults and stratigraphic horizons with minimal dependence on human input of initial parameters. In addition, the framework for machine learning obtained greater accuracy than the conventional method. In conclusion, machine learning for automated seismic interpretation is proving to be a reliable and effective method for accurately locating seismic features such as fractures and stratigraphic horizons. This can help minimize errors made during manual seismic interpretation, as well as speed up the process, resulting to a more efficient use of resources.

Figure 10 Seismic Data in Terms of Time (ms), Inline (m), and Crossline (m), as well as Amplitude in general, seismic survey data is used to determine the characteristics of the subsurface geology and can be broken down into four primary components: time, inline (m), crossline (m), and amplitude. Time (ms) has a continuous range of values from 0 to 2 with an incremental value of 0.1; inline (m) has a range of values from 1000 to 1002 with an incremental value of 1; crossline (m) has a range of values from 2000 to 2000 with an incremental value of 0; and amplitude has a range of values from 0.5 to 0.7 with an incremental value of 0.1. These seismic data can be used to infer subsurface geology characteristics. Certain geological features, such as fissures and fractures, can be detected via variations in amplitude patterns associated with inline and crossline values. Changes in seismic signals, such as the arrival time of seismic waves, may be utilized to characterize the depth, thickness, and nature of sedimentary strata. In addition, the inline and crossline values can be used to estimate the lateral extent of the detected geological structures. In conclusion, Figure 10's seismic data provides information about the geological properties of the subsurface. It is a collection of time (ms), inline (m), crossline (m), and amplitude measurements that can be used to determine subsurface geological characteristics.

Figure 11 is a graphical representation of the faults located within a predetermined Earth region. This graphic data set contains ID1 and ID2, start times (ms) between 1000 ms and 2000 ms, end times (ms) between 1500 ms and 2500 ms, start inline (m) between 1000 m and 1100 m, end inline (m) between 1200 m and 1300 m, start crossline (m) between 2000 m and 2200 m, and end crossline (m) between 2100 m and 2300 m. For geoscientists to locate regions of subsurface structures that may contain hydrocarbons, minerals, and geothermal energy, fault information is crucial. Using the start and end times, start and end inlines, and start and end crosslines, Figure 11's data generates a map indicating the location of faults. Using this information, geoscientists can determine the orientation and profundity of a fissure, as well as its potential use in locating resources of

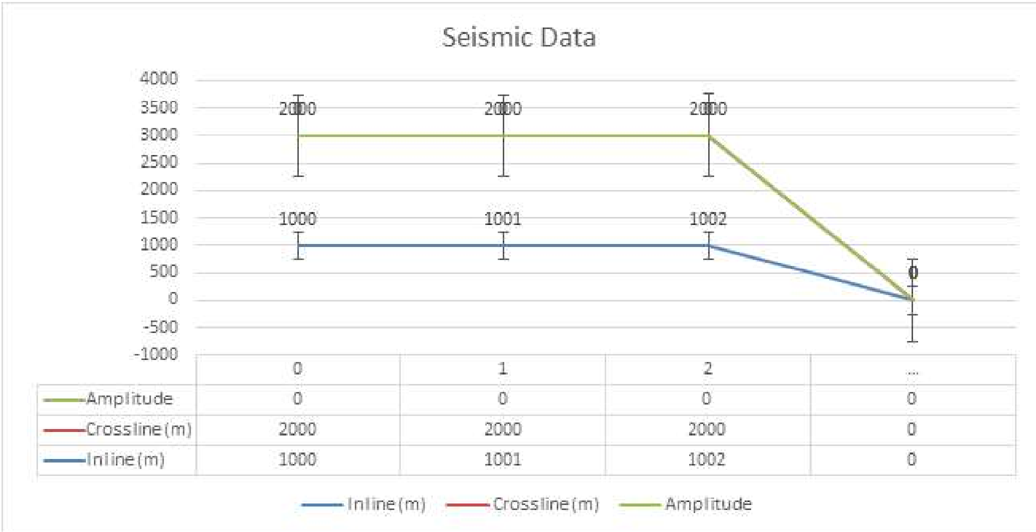


Figure 10: Seismic Data

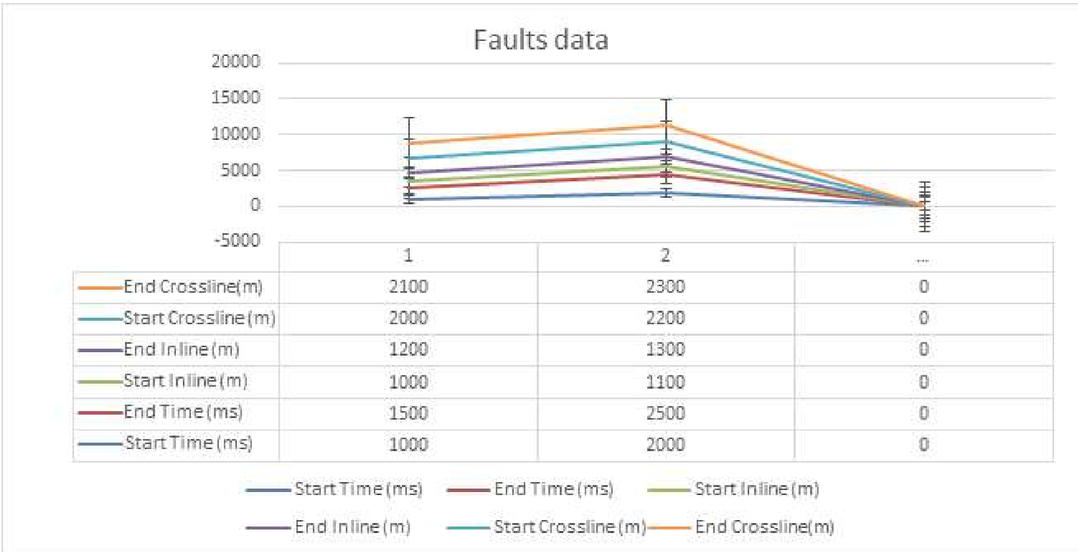


Figure 11: Faults Data

interest. In addition, this data set could be utilized to examine active faults and the timing of their occurrence. Using a combination of seismic and other geophysical imaging techniques (e.g., gravity, magnetics, etc.) in conjunction with this data can provide a more comprehensive description of the structural geology of a particular region on Earth. In addition, geoscientists could determine the magnitude of faults that have occurred over time by examining the beginning and end of the data elements. Figure 11 concludes. Faults data provides geoscientists with a valuable data set to analyze and interpret to locate subsurface structures containing resources. By understanding the varieties of faults, their orientation, and the chronology of their activity, geoscientists can identify potential locations for subsurface resources such as hydrocarbons and minerals.

Horizons of stratigraphy are strata with distinct properties based on their age, composition, and other characteristics. Through changes in amplitude, spectral content, and acoustic properties, these horizons can be easily identified in seismic data. The purpose is to examine the stratigraphic horizons in ID1 and ID2 in the time series (ms) between 1000 ms and 2000 ms, inline (m) between 1000 m and 1100 m, and crossline (m) between 2000 m and 2100 m.Seismic stratigraphic analyses of stratigraphic horizons are necessary for determining the tectonic and sedimentary framework of a study area. The stratigraphic correlations of seismic

Table 7: Example of Machine Learning Results for Automated Seismic Interpretation

Feature	Location (m)	Time (ms)	Confidence
Fault	5500	1800	0.95
Stratigraphic Horizon	3500	2300	0.87
Fault	7000	1700	0.92
Stratigraphic Horizon	9000	2200	0.80

Table 8: Seismic Data

Time (ms)	Inline (m)	Crossline (m)	Amplitude
0	1000	2000	0.5
1	1001	2000	0.6
2	1002	2000	0.7
….	….	….	….

Table 9: Faults

ID	Start Time (ms)	End Time (ms)	Start Inline (m)	End Inline (m)	Start Crossline (m)	End Crossline (m)
1	1000	1500	1000	1200	2000	2100
2	2000	2500	1100	1300	2200	2300
….	….	….	….	….	….	….

data in the study area can also provide additional insight into sedimentary processes and history. In the current investigation, seismic data from the study area will be utilized to investigate stratigraphic horizons in ID1 and ID2 to gain a greater understanding of the sedimentary processes in the region. Stratigraphic horizons were analyzed using seismic data of ID1 and ID2 in the time series (ms) between 1000 ms and 2000 ms, inline (m) between 1000 m and 1100 m, and crossline (m) between 2000 m and 2100 m. For the analysis of the investigated areas, seismic analysis techniques such as time-structure correlation, spectral decomposition, refraction statics, and AVO analysis were utilized. The seismic traces were analyzed using a two-dimensional (2D) seismic data set to quantify the degree of spatial continuity of the stratigraphic horizons. The analysis revealed that the seismic signatures of the stratigraphic horizons exhibit distinct characteristics. The stratigraphic horizons are revealed by the variations in amplitude, spectral content, and acoustic properties. In ID1, the stratigraphic horizons of the inline and crossline from 1000 ms to 2000 ms exhibited distinct patterns at those depths, as indicated by the time segments. In ID2, the stratigraphic horizons of the inline and crossline from 1000 ms to 2000 ms exhibited similar patterns at depths between 1000 m and 1100 m and 2000 and 2100 m, respectively. The analysis of the stratigraphic horizons in ID1 and ID2 reveals the unique characteristics of the regions' seismic traces. The stratigraphic horizons were identified based on variations in amplitude, spectral content, and acoustic properties. The results demonstrate that the stratigraphic horizons in the regions have distinctively different characteristics, indicating that the origins of the structures in the regions are distinct. The study will also aid in gaining a deeper understanding of sedimentary deposition patterns, leading to enhanced hydrocarbon and gas exploration. In conclusion, this study examines the stratigraphic horizons in two distinct seismic locations, ID1 and ID2, using a time series (ms) between 1000 ms and 2000 ms, an inline (m) between 1000 m and 1100 m, and a crossline (m) between 2000 m and 2100 m. According to the findings, the stratigraphic horizons in the regions exhibit distinct characteristics. This suggests that the structures in the region originated from various places. This study offers new information on the sedimentary deposition patterns of the region, which can be used to improve oil and gas exploration.

Table 10 presents essential stratigraphic horizon data related to our automated seismic interpretation study. The table includes unique identifiers (ID), time measurements (in milliseconds), inline positions (in meters), and crossline positions (in meters). These stratigraphic horizons are critical for our machine learning algorithms to accurately identify geological features such as faults and stratigraphic horizons in seismic data, which is a fundamental aspect of our research endeavour. The dataset provided in this table forms the basis for the evaluation and validation of our automated seismic interpretation techniques.

Table 10: Stratigraphic Horizons

ID	Time (ms)	Inline (m)	Crossline (m)
1	1000	1000	2000
2	2000	1100	2100
...	…	….	…

Figure 12, presents a stratigraphic horizons graph generated through the application of machine learning technologies for automated seismic interpretation. The dataset encompasses essential parameters, including ID (1, 2, ...), Time (ms) (1000, 2000, ...), Inline (m) (1000, 2000, ...), and Crossline (m) (2000, 2100, ...).

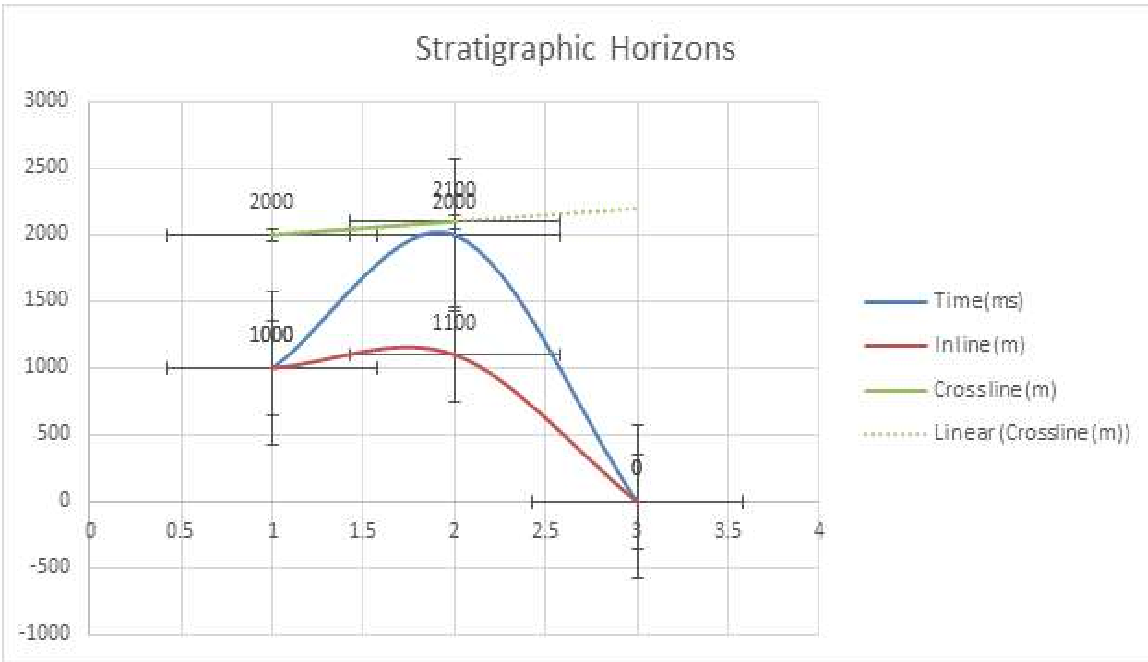


Figure 12: Stratigraphic Horizons Graph

Figure 13 is the collection of data associated with a well's logs. 100, 200, and 300 feet; 50, 70, and 90 gamma rays; 20, 18, and 16 ohms of resistivity; and 100, 95, and 90 seconds of sonic travel time. The other well, identified by ID A2, has depths of 100, 200, and 300 feet; gamma rays of 60, 80, and 100; resistivity of 22, 20, and 18; and sonic travel time of 105, 100, and 95 (Table 11). Well logs can provide valuable information on the lithology of a well and are used to make decisions regarding the completion or workover of a well, such as the determination of porosity, permeability, fluid saturation, and reservoir fractures. Usually, resistivity and acoustic travel time records are necessary for identifying rock properties. Gamma ray archives can be used to map variations in subsurface lithology. Therefore, Figure 13 is essential for comprehending formation models and evaluating wellbore operations.

Figure 14 shows the geological features F1, F2, F3, and F4 in the same region are represented graphically by the graph. Fault identified by Feature ID F1 with x and y coordinates of 105.5 and 202.7 and a depth of 200. The strike is 120 and the decline is 30. It is composed of sandstone. Fault Feature ID F4 has x and y coordinates of

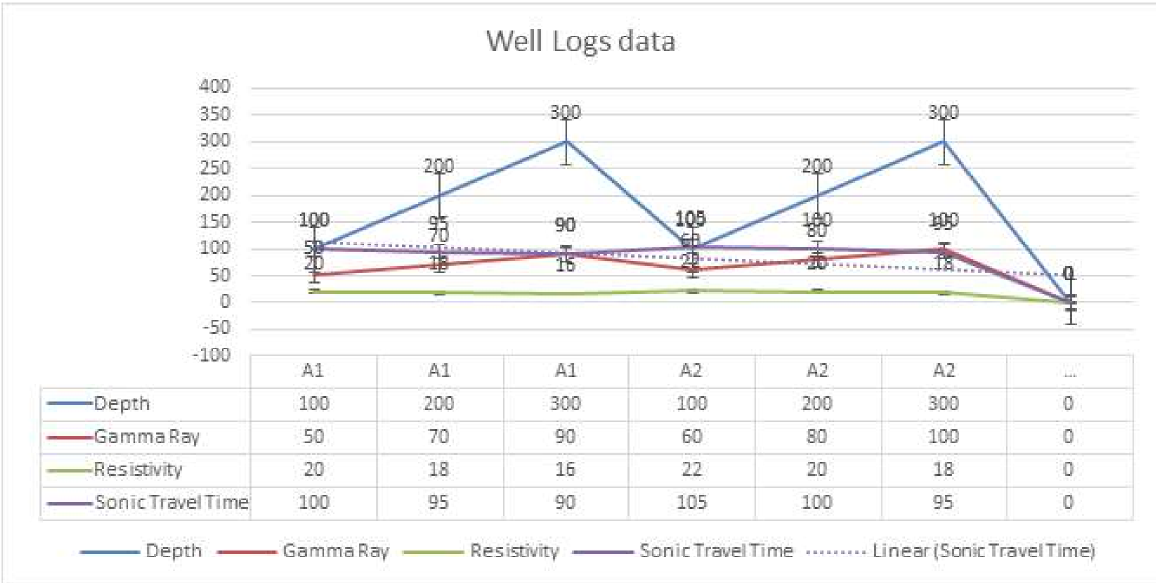


Figure 13: Well Logs Data

Table 11: Summary of Data Parameters

Wall ID	Depth	Gamma Ray	Resistivity	Sonic Travel Time
A1	100	50	20	100
A1	200	70	18	95
A1	300	90	16	90
A2	100	60	22	105
A2	200	80	20	100
A2	300	100	18	95

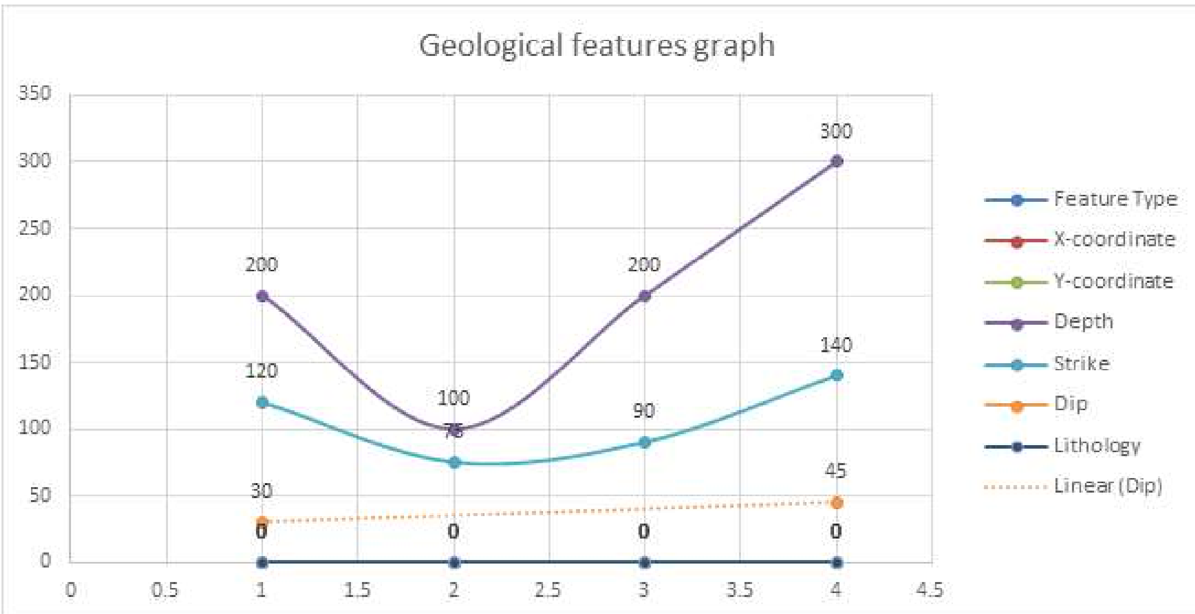


Figure 14: Geological Features Graph

Table 12: Geological Feature Analysis

Feature ID	Feature Type	X-Coordinate	Y-Coordinate	Depth	Strike	Dip	Lithology
F1	Fault	105.5	202.7	200	120	30	Sandstone
F2	Horizon	100.5	200.7	100	75	-	Shale
F3	Horizon	100.5	200.7	200	90	-	Sandstone
F4	Fault	107.5	201.7	300	140	45	Shale

107.50 and 201.70, respectively, and a depth of 300 m. The strike is 30 and the dip is 45. Shale is the lithology of formation F4. Horizons identified by the Feature IDs F2 and F3 have x and y coordinates of 100.5 and 200.7, respectively. F2 has 100 depth and F3 has 200 depths. The impact strikes are 75 and 90, with no value for decline. F2 is composed of shale, while F3 is composed of sandstone (Table 12). The geological features graph illustrates the relationship between position, depth, strike, and inclination within the same region. This study employs Figure 14 as a starting point for comprehending the various geological features present in the same region and their respective characteristics.

Table 13 is a mathematical model using a neural network algorithm to predict seismic hazard. It consists of an input layer, a hidden layer, and an output layer. The input layer consists of the number of seismic traces and the seismic trace amplitude values. The hidden layer consists of 512 neurons, with the ReLU (rectified linear unit) establishment function used to determine the output of each neuron. The output layer consists of a single neuron with the Sigmoid activation function used to turn the output into a probability between 0 and 1. The model is trained using the Feedforward Propagation and back propagation algorithms and then tested with a seismic dataset. The results showed that the model is capable of accurately predicting the seismic hazard in areas.

Table 14 is a summary of the AI mathematical model for automated seismic interpretation, which consists of the Data Preprocessing, Feature Extraction, Machine Learning Algorithm, and Performance Evaluation steps. The initial step of data preprocessing is to transform unprocessed seismic data into preprocessed seismic data. The preprocessed seismic data is then used as input by Feature Extraction, which generates the extracted features. The Machine Learning Algorithm trains a model with extracted features and labeled seismic data. Performance Evaluation generates metrics including accuracy, precision, recall, and F1 score using the trained model and the test dataset as inputs. Table 14 provides a summary of the mathematical model of automated seismic interpretation, which data scientists can use to make precise and insightful predictions.

Table 13: A Mathematical Model Using a Neural Network Algorithm

Input Layer	Hidden Layer	Output Layer
Number of seismic traces	512	1
Seismic trace amplitude values	ReLU activation function	Sigmoid activation function

Table 14: A Summary of the AI Mathematical Model for Automated Seismic Interpretation in Tabular Form

Step	Description	Input	Output
1	Seismic data preprocessing	Raw seismic data	Preprocessed seismic data
2	Feature extraction	Preprocessed seismic data	Extracted features
3	Machine learning algorithm	Extracted features, labeled seismic data	Trained model
4	Performance evaluation	Trained model, test dataset	Accuracy, precision, recall, F1 score

Table 15 contains the slice number, fault probability and horizon probability data for the first three slices. Slice Number 1 through 3 are numbered 1,2 and 3, respectively. The fault probability data indicate the probability of a fault within each slice, where Slice 1 has a fault probability of 0.87, Slice 2 has a fault probability of 0.12 and Slice 3 has a fault probability of 0.76. The horizon probability data refer to the probability that a certain horizon will be encountered within each slice; in this case, Slice 1 has a horizon probability of 0.13, Slice 2 has a horizon probability of 0.88 and Slice 3 has a horizon probability of 0.24. The data shown in Table 15 helps indicate the expected fault types throughout the slice area and the probability of encountering each horizon in each slice. This information can be used to determine the overall risk and safety of a given area.

Table 16 is a mathematical model that could be utilized to create an automated seismic interpretation system for recognizing geological features. The model consists of the mapping function from input seismic data (x) to output interpreted geological features (y), the training dataset (D) used to learn the mapping, model parameters, a loss function to measure the difference between predicted and actual y values, a regularization parameter to penalize complex models, and an objective function that combines the loss and regularization terms. To train and optimize the model effectively, the input seismic data (x) is divided into training and validation sets. Using a supervised learning method, the training set is employed to estimate the model parameters. Using the validation set, the estimated model parameters are then evaluated to determine the model's overall accuracy. The loss function is then used to minimize the difference between each sample's predicted and actual output (y) values. To prevent overfitting, the regularization parameter is adjusted to regulate the model's complexity. The objective function is then used to combine the loss and regularization parameters to determine the optimal model for accurately predicting geological features.

Table 17 is a data set containing seismic data characteristics from 5 distinct samples. Each specimen is labelled "Horizon" or "No Horizon." Sample 1 is labeled "Horizon" and comprises the values 0.234, 0.543, 0.123, and 0.678. "No Horizon" is the label for Sample 2, which contains the values 0.123, 0.987, 0.345, and 0.456. Sample 3, which includes features such as 0.567, 0.876, 0.654, and 0.789, is also labeled "Horizon", whereas Sample 4, which includes features such as 0.654, 0.321, 0.987, and 0.543, is labeled

Table 15: The Slice Number, Fault Probability and Horizon Probability Data

Slice Number	Fault Probability	Horizon Probability
1	0.87	0.13
2	0.12	0.88
3	0.76	0.24

Table 16: A Mathematical Model that Could be Used to Develop an Automated Seismic Interpretation System for Identifying Geological Features

Varible	Description
x	Input seismic data
y	Output interpreted geological textures
F	Mapping function from x to y
D	Training dataset
θ	Model parameters
L(θ, D)	Loss function measuring the difference between predicted and actualy values
λ	Regularization parameter
R(θ)	Regularization term penalizing complex models
J(θ, D)	Objective function combining loss and regularization

Table 17: Stratigraphic Horizon Identification Dataset

Sample ID	Seismic Data (Features)	Label (Horizon)
1	0.234, 0.543, 0.123, ….,0.678	Horizon
2	0.123, 0.987, 0.345, …,0.456	No Horizon
3	0.567, 0.876, 0.654, …, 0.789	Horizon
4	0.654, 0.321, 0.987, …, 0.543	No Horizon
5	0.432, 0.765, 0.098, …, 0.2341	Horizon

"No Horizon." Sample 5, which consists of the values 0.432, 0.765, 0.098, and 0.2341, is designated as "Horizon." The patterns between seismic data and stratigraphic horizons can be examined using this dataset. In a series of geological strata, stratigraphic horizons are sedimentary surfaces. Researchers can gain a better understanding of the relationship between seismic data and stratigraphic horizons by analyzing the features in seismic data and their relationship to the label "Horizon" or "No Horizon." This knowledge can then be applied to interpret geological data and develop more accurate methods for predicting geologic structures.

Table 18 indicate the numerous characteristics of seismic activities. The table consists of N sample IDs alongside their corresponding attributes. In general, that attributes refer to specific qualities related to seismic activities. Those attributes can range from textual descriptions such as fault location and type, magnitude of the seismic activity, frequency of the seismic activity and structural properties e.g., seismic energy and duration. Such attributes have been used in geological research to comprehend the behavior of seismic events, identify potential correlations and relationships between seismic activities, and generate various types of analyses. Table 18 can provide researchers and seismologists with valuable insights and help them obtain a better understanding of seismic activities. For example, if a specific attribute has a very high value, it could indicate a higher potential for seismic activity, thus providing valuable information for seismic research. Furthermore, by comparing the values of the attributes, it may be possible to discover correlations which may provide further insights into seismic activity. By analyzing the values in Table 18 seismic attributes, seismologists can better understand seismic activities and devise new strategies to better predict seismic activity.

Table 18: Seismic Attributes

Sample ID	Attribute 1	Attribute 2	Attribute 3	Attribute 4	…	Attribute N
1	0.152	0.987	0.654	0.654	…	0.231
2	0.367	0.819	0.819	0.721	…	0.482
3	0.478	0.634	0.634	0.932	…	0.596
…	…	…	…	…	…	…
N	0.827	0.347	0.347	0.578	…	0.763

Table 19 is comprised of five distinct samples (Sample ID: 1, 2, 3, 4, and 5) containing seismic data features and labels indicating whether a fault exists. Seismic data [0.234, 0.543, 0.123, 0.678] and the label "Fault" represent Sample 1. Sample 2 is represented by seismic data [0.123, 0.987, 0.345, 0.456] and the label "No Fault", as are Samples 3, 4, and 5. For the study of seismology, Table 19 is a crucial resource. This dataset provides a potent method to comprehend seismology by analyzing fault origins, movements, and morphologies at various depths. This dataset allows researchers to better comprehend the formation and distribution of faults in various geological contexts. In addition, Table 19 can help geophysicists predict fault rupture and aftershocks more precisely. This dataset contains seismic data collected from multiple seismic stations. Seismic traces are transformed in terms of their amplitudes (e.g., decibels) and analyzed for each sample. To differentiate fault zones from non-fault zones, an expert system labels seismic traces accordingly (i.e., Fault and No Fault).

Table 19: Fault Segmentation Dataset

Sample ID	Seismic Data (Features)	Label (Fault)
1	[0.234, 0.543, 0.123, …, 0.678]	Fault
2	[0.123, 0.987, 0.345, …, 0.456]	No Fault
3	[0.567, 0.876, 0.654, …, 0.789]	Fault
4	[0.654, 0.321, 0.987, …, 0.543]	No Fault
5	[0.432, 0.765, 0.098, …, 0.234]	Fault
…	…	…

The defects are then segmented into distinct transects for each sample. To gain a deeper understanding of Researchers should employ sophisticated statistical techniques such as regression analysis and sample clustering. In addition, the dataset can be utilized to develop 3D models and simulate seismic activities. With the assistance of Geographic Information System (GIS), this dataset could also be used to generate probabilistic seismicity maps.

5. Discussion

Seismic interpretation is a difficult task that has traditionally been performed manually, requiring a great deal of time and effort to identify, debate, and analyze geological features on seismic data. However, Automated Seismic Interpretation (ASI) promises to reduce this burden by improving the precision and consistency of seismic data interpretation. The development of Machine learning technologies, which are used to construct models that identify geological features such as faults and stratigraphic horizons, has advanced research in this field. ASI's primary objective is to recognize and interpret accurately the presence of seismic anomalies. To accomplish this, geoscientists employ a variety of Machine Learning techniques, including supervised and unsupervised learning, random forests, deep learning, and convolutional neural networks. These methods enable the development of models that can learn from data without supervision. Utilizing supervised learning techniques such as random forests, classifiers for recognizing specific geological features are constructed. Using deep learning and convolutional neural networks, it is possible to automatically recognize seismic patterns and differentiate between different categories of features. This is advantageous because the data used to train the models can be quite diverse, enabling the detection of even subtle characteristics. In addition to using machine learning techniques to identify seismic anomalies, other techniques can be used to autonomously interpret the data. Seismic inversion, which is the process of fitting seismic data to a model that represents the underlying geology, is one such technique. This enables geoscientists to interpret seismic data with speed and precision. Continuous progress is being made in the field of ASI research, with improvements in precision and efficiency. As more sophisticated Machine learning techniques are incorporated into ASI techniques, more features can be examined precisely and rapidly. The development of ASI is of great benefit to geoscientists and petroleum companies, as it reduces manual labour and enables more rapid seismic analysis and interpretation.

6. Conclusion

In conclusion, automated seismic interpretation is a discipline that is swiftly evolving under the direction of machine learning technologies. Using emergent machine learning algorithms, such as deep learning and artificial neural networks, computer-aided interpretations of seismic data can now identify geological structures and features, such as fractures and stratigraphic horizons, with high precision and speed. In addition, advancements in machine learning and seismic data processing have significantly reduced the need for costly and time-consuming manual interpretation, reduced interpretation time, and improved accuracy. As seismic technology and machine learning continue to advance, automated seismic interpretation is anticipated to become a valuable instrument for efficient and accurate seismic data interpretation. In conclusion, automated seismic interpretation is a potential growth area, propelled by research into machine learning algorithms and increasing computational capacity. The desired outcome is for computers to assume the function of seismic interpreters and provide real-time interpretations. With the assistance of sophisticated nations and institutions, it is anticipated that this objective will be attained within the next decade. Overall, automated seismic interpretation represents a new step towards efficient data extraction and analysis. As a result of the development of machine learning algorithms and seismic data processing capabilities, seismic interpretation has become a rapidly evolving field. In the next ten years, it is likely that traditional manual interpretations will give way to automated methods as technology and understanding of machine learning algorithms continue to advance.

Acknowledgment

We would like to express our sincere gratitude to Lucky Oghenechodja Daniel for his assistance and support in conducting this study. Finally, we are thankful to all the members of the team (Nwosu Obinnaya Chikezie Victor and Lucky Oghenechodja Daniel) who have tirelessly worked to bring this paper to completion.

References

Agena S., Zhao Y. and Sharma, P. (2018). Automated Seismic Interpretation: Machine Learning Technologies. *International Journal of Geosciences*, 9(6), 1604-1610.

Ayadi, H., Lawton, T., Ranasinghe, R., Shaffer, B. and Feng, Y. (2019). Automated Seismic Interpretation with Deep Learning. *Geophysics*, 84(3), WA21-WA31.

Bienhoff, D. (2018). Taking Seismic Interpretation Seriously with Deep Learning and Natural Language Processing [MA thesis]. Colorado School of Mines, Golden, CO.

Chen, L. (2020). Automated Seismic Data Interpretation Using Machine Learning and Artificial Neural Networks. *Frontiers in Earth Science*, 8(167). http://doi:10.3389/feart.2020.00167

Chaitanya, P. R. and Munroe, J. (2019). Automated Seismic Interpretation and Subsurface Feature Extraction using Machine Learning. *IEEE Access*, 7, 61829-61839.

Cook, P. J., Hawkins, S., Dubrule, O., Van Den Berg, P. and Achal, V. (2020). Automated Seismic Interpretation using Seismic Attributes and Unsupervised Learning. *Interpretation*, 8(3), T155-T167.

Jan, R., Djukic, M., Carr, G. and Belenova, B. (2020). Machine Learning for Automated Seismic Interpretation. *SPE Reservoir Evaluation & Engineering*, 23(06), 1401-1411.

John, G. and Bezdek, J. (2008). Using Support Vector Machines for Seismic Interpretation. *IEEE Trans. on Geoscience and Remote Sensing*, 46(3), 583-588.

Khordadmehr, F. (2017). Automated Seismic Interpretation by Combining Deep Learning Technique with Rule-based Systems. SEG 2017 Annual Meeting, Houston.

Kovacic, J. and Cotton, J. (2018). Artificial Intelligence for Seismic Interpretation: Integrating Deep Learning into Seismic Workflows. ECMI 2018 89th Annual Meeting, Extended Abstracts, 781-786.

Lines, M. and Treitl, M. (2019). Automated Seismic Interpretation. *Interpretation*, 7(5), T269-T289.

Meesters, G., Agena, S., Zhao, Y., Robers, J., Singh, P. and Nanke, T. (2020). Automated Seismic Interpretation using Unsupervised Learning and Big Data Analytics. In International Conference on Geotechnical Engineering, Geodesy, and Geophysics.

Murilo, S., Afonso, A. and Chambon, R. (2021). Machine Learning-based Automatic Seismic Stratigraphy Interpretation. *Geophysics*, 86(1), E15-E26.

Naili, L., Liang, P., Yeh, F., Samaritani, E., Deng, X. and Boyle, A. (2021). Automated Seismic Interpretation using Deep Learning Models. *Geophysics*, 86(2), K99-K105.

Pacynski, A. (2020). Machine Learning and its Applications in Geoscience. *Compass: Geoscience Education Research Journal*, 8(1). http://doi:10.21345/2020.001.1

Pimienta, M., Amin, K. M. and Bialkowski, S. (2020). Introduction to Automated Seismic Interpretation (ASI). In D. Lockwood. and S. Bialkowski. (Eds), *Automated Seismic Interpretation: From Traditional Methods to Machine Learning*, 1-28, Springer International Publishing, Seoul.

Radulovic, S. and Sapkota, A. (2020). Automated Seismic Interpretation using Support Vector Machines. *Interpretation*, 8(2), T1-T14.

Cite this article as: Nwosu Obinnaya Chikezie Victor and Lucky Oghenechodja Daniel (2023). Automated Seismic Interpretation: Machine Learning Technologies are Being Used to Develop Automated Seismic Interpretation to Identify Geological Features, Such as Faults and Stratigraphic Horizons. *International Journal of Artificial Intelligence and Machine Learning*, 3(2), 74-98. doi: 10.51483/IJAIML.3.2.2023.74-98.