Republic of Iraq

Ministry of Higher Education

And Scientific Research

University of Diyala

College of Science

Prediction Of Reservoirs Porosity Based On Resulting Seismic Data Attributes Using Deep Learning Approach

A Thesis

Submitted to the Department of Computer Science\ College of Science\ University of Diyala

In a Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science

By

Mohammed Wahab Raheem

Supervised By

Assist.Prof. Dr. Abdulbasit Kadhim Shukur

2022 A.D. 1443 A.H.

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

﴿ يَرۡفَعِ ٱللَّهُ ٱلَّذِينَ ءَامَنُواْ مِنكُمۡ وَٱلَّذِينَ أُوتُواْ ٱلۡعِلۡمَ دَرَجَٰتٖۚ
وَٱللَّهُ بِمَا تَعۡمَلُونَ خَبِيرٞ ١١ ﴾

صدق الله العظيم

سورة المجادلة – الآية (١١)

DEDICATION

To…

- *The souls of my father and mother, my God have mercy on them…*
- *My beloved wife …*
- *My precious children Abdullah, Gana, and Humam*
- *My beloved brothers and sisters…*
- *My dear friend Dr. Hassan*
- *My friends and classmates… and everyone who want goodness and success for me.*

Mohammed

2022

Linguistic Certification

*This is to certify that the thesis entitled "**Prediction of Reservoirs Porosity Based on Resulting Seismic Data Attributes Using Deep Learning Approach**" was prepared by "**Mohammed Wahab Raheem**" was prepared under my linguistic supervision. It was amended to meet the style of English language.*

Signature:

Name: **Dr.** Ghazwan Mohammed Jaafar

Date: / /2022

Scientific Amendment

I certify that the thesis entitled ***"Prediction of Reservoirs Porosity Based on Resulting Seismic Data Attributes Using Deep Learning Approach"*** *was prepared by* ***Mohammed Wahab Raheem*** *has been evaluated scientifically ,therefore, it is suitable for debate by the examining committee.*

Signature:

Name: **Assist. Prof. Dr.** Abbas Abdulazeez Abdulhameed

Date: / / 2022

Signature:

Name: Assist. Prof. Dr. Ruaa Adeeb Abdulmunem

Date: / / 2022

Supervisor Certification

I certify that this research entitled ***"Prediction of Reservoirs Porosity Based on Resulting Seismic Data Attributes Using Deep Learning Approach"*** *was prepared under my supervision at Department of Computer Science\ College of Sciences\ the University of Diyala by* ***"Mohmmed Wahab Raheem"*** *as a partial fulfillment of the requirements for the degree of* ***Master of Science in Computer Science***

(Supervisor)

Signature:

Name: **Assist. Prof. Dr. Abdulbasit Kadhim Shukur**

Date: / /2022

Approved by Department of Computer Science\ College of Science\ University of Diyala

Signature:

Name: **Assist. Prof. Dr.Bashar Talib Hamed**

Date: / /2022

Head of Department of Computer Science\ College of Science\ University of Diyala

Examination Committee Certification

*We certify that we have read the thesis entitled **"Prediction of Reservoirs Porosity Based on Resulting Seismic Data Attributes Using Deep Learning Approach"** and as an examination committee, examined the student **"Mohammed Wahab Raheem"** in the thesis content and that in our opinion, it is adequate as fulfill the requirement for the Degree of Master in Computer Science at the Computer Science Department, University of Diyala.*

(Chairman)
Signature:
Name: **Prof. Dr.Taha Mohammed Hasan**
Date: / / 2022

(member)
Signature:
Name: **Assist. Prof. Dr. Maki Mahdi Abdulhasan**
Date: / / 2022

(member)
Signature:
Name: **Assist. Prof. Ali Abdulrahman Mahmood**
Date: / / 2022

(member) and (Supervisor)
Signature:
Name: **Assist. Prof. Dr. Abdulbasit Kadhim Shukur**
Date: / / 2022

Approved by the Dean of College of Science, University of Diyala

(The Dean)

Signature:

Name: **Prof. Dr. Tahseen Hussein Mubarak**

Date: **/ / 2022**

ABSTRACT

The massive development of shale oil formations has changed the rules of the game. On the other hand, Machine Learning (ML) and Deep Learning (DL) play an important role in the rapid development of all industries by automating most of the routine processes. The oil industry also gets equal benefits from ML and DL for reservoir development planning and operational accuracy through a series of automated systems. To develop the field, computational static and dynamic simulation models are generated based on various petrophysical properties collected through various resources that are time-consuming and expensive. This study aims to present a comprehensive model in the field of application of ML and DL to model the petrophysical properties using different methods and algorithms. Finally, the multiple ML and DL techniques that are tested in this study are discussed in detail in order to achieve more accuracy in the petrophysical simulation models. Machine learning models were used to support vector regression (SVR) and nearest neighbor regression (KNR), for further improvement, using deep learning algorithms. Use long-term memory (LSTM) and prepare the output by an artificial neural network (ANN). Also, to improve deep learning by recurrent neural networks (RNN) a hybrid method (LSTM) with a recurrent gates unit (GRU) and an artificial neural network (ANN) is used. The best decisions obtained in forecasting oil reservoirs and reducing uncertainty in exploration and drilling is if the data set is divided as follows, the prediction model using machine learning is 90% training and 10% testing. The best results were MAE = 0.238 and RMSE = 0.255 with SVR, while the KNR algorithms achieved results of MAE = 0.276 and RMSE = 0.301. While in deep learning algorithms when splitting the data into 80% training and 20% testing. The best performing result in LSTM had values of MAE = 0.023 and RMSE = 0.029, meaning the best performance for deep learning.

LIST OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

Abbreviation	***DESCRIPTION***
ACE	*Alternative Conditional Expectation*
AI	*Artificial Intelligence*
ANN	*Artificial Neural Network*
CC	*Correlation Coefficient*
CM	*Committee Machine*
DLNN	*Deep Learning Neural Network*
DT	*Decision Tree*
FN	*Functional Network*
GRU	*Gated Recurrent Units*
IDDM	*Insulin-Dependent Diabetes Mellitus*
KNN	*K-Nearest Neighbor*
LR	*Logistic Regression*
LSTM	*Long-Short Term Memory*
MAA	*Multi-Attribute Analysis*
MAE	*Mean Absolute Error*
MDR	*Multifactor Dimensionality Reduction*
ML	*Machine Learning*
NN	*Neural Network*
OFL	*Optimized Fuzzy Logic*
ONN	*Optimized Neural Network*
PNN	*Probabilistic Neural Network*

RMAE	*Root Mean Absolute Error*
SAA	*Single Attribute Analysis*
SVM	*Support Vector Machine*
SVR	*Support Vector Regression*
TS	*Time Series*
TSA	*Time Series Analysis*

Chapter One

Introduction

Chapter One

Introduction

1.1 Overview

Oil is one of the most important natural resources on which developed countries depend on developing their economies. Its exploration and production stages are among the most important priorities of many countries. Oil prospecting and exploration are among the priorities of the mission that require the use of advanced techniques and methods. These methods and techniques differ from place to place according to the nature of the land and the geophysical formations of the fluid reservoirs. Exploration of oil wells involves obtaining information called the well log, which is a set of basic information on wells [1].

The different nature of the Earth's geophysical data is a problem facing drilling and exploration, which is unreliability and the increased risk of drilling. Therefore, Which is why seismic survey technology was discovered, which produced data called seismic data, It is a set of features by which the porosity and permeability of the well and some of the features of the reservoir formations can be known. With the help of machine learning models, which contributed to reducing drilling risks, increasing reliability, and predicting the porosity and permeability of the well [2].

Porosity is an important factor to determine the capacity of reservoirs of liquids and to give an understanding of the liquid and gaseous formations in them. Therefore, the basic standard requirements in tanks are porosity and permeability [3].

Data acquisition and analysis is very time-consuming and expensive, requires significant human and technical efforts, and the reservoir may not be adequately described. As a result, a less expensive and faster method for porosity quantification is required. Porosity can be estimated using the well log and seismic data, but many of these logs are difficult to obtain accurately [4].

An oil and gas reservoir is a rock formation in which petroleum and natural gas have accumulated. The oil and gas inside the reservoir are held by adjacent and accumulating layers of rock. Using available field and laboratory data, ML can describe different reservoir properties. The process of developing a reservoir, usually between the discovery and management phases of a reservoir, incorporates certain characteristics related to its ability to store and produce petroleum [5].

In recent years, deep learning techniques have been developed to process different types of data. One type of deep learning is a recurrent neural network (RNN), which is used for sequential or time-series data, such as text, audio, and video [6]. One associated technique is Long-Short-Term Memory (LSTM) which has processed time-series for a variety of data, and almost all of the excellent results have been achieved through deep learning [7].

This thesis presents a porosity prediction model using machine learning algorithms based on (SVR) Support Vector Regression and K Neighbors Regression (KNR), LSTM-based deep learning algorithms, and Gated Recurring Units (GRU). In addition, a hybrid algorithm was proposed using LSTM and GRU, and (ANN) was used with deep learning algorithms to adjust the output weights.

1.2 Related Works

The following are some recent studies on the relationship between machine learning and porosity prediction:

- **S.R. Na'imi. et al. (2014)** [8]**.** In this study, an SVR approach is represented by ML as a functional regression method in regression problems. Use the principle of structural risk reduction. Where appropriate seismic characteristics are extracted, which mainly depend on the porosity of the tank and the water saturation. Then, a quantitative formula for the relationship between porosity parameters. It is obtained by using a nonlinear vector regression algorithm in water saturation and selected seismic features. In the proposed SVR model, the results showed that it is suitable for implementation to predict porosity in small data and solve complex problems, compared to other methods that require more challenges.

- **Amin Gholami. et al (2017)** [9]. In this study. A mixed model is proposed to determine the articulation between porosity and seismic features by machine learning in three steps. In addition, the appropriate seismic features that have a prominent effect on porosity are extracted using the reorientation variable method and used as model input parameters. In addition, when compared to the non-parametric method known as alternative conditional expectation (ACE), the input variables are shifted to larger data space. In the next step, the correlation between the input parameters and porosity is quantitatively transformed through the optimized intelligence model, including optimized neural network (ONN), optimized support vector regression (OSVR), and optimized fuzzy logic (OFL) to achieve the predictive validity. In the final step, through the Committee Machine (CM), the integrated outputs of the optimized models to improve prediction accuracy are embedded in the modeling intelligence. The Committee Machine (CM) model error distribution is very close to the normal distribution. The CM predictions are very compatible with reality because the errors (0.0068) from samplings show the range degree as to be in ± (0.0067) and (0.0301).

- ❖ **S. P. Maurya, et al (2018)[10].** The study goal is to discover an effective mix of seismic reflection techniques and geostatistical approaches for predicting porosity and identifying potential areas in 3D seismic data spaces. In this study, three geostatistical methods were used to predict porosity: single-attribute analysis (SAA), multi-attribute analysis (MAA), and the probabilistic neural network (PNN) algorithm. In a time interval of 1060-1075 ms, the result obtains a very high porosity (N 15%). These techniques make use of the seismic features generated by model-based reflection and color reflection techniques. The results demonstrated that all three statistical methods used to predict porosity are effective and reliable, but multi-feature and probabilistic neural network analysis provides more accurate and high-resolution porosity sections.

- ❖ **Xu Zhou, et al (2019)** [11]**.** This paper shows how to use big data analysis to verify the statistical correlations between seismic attributes parameters from three-dimensional seismic surveys and petro-physical properties from (well logs). Using Deep Learning Neural Network (DLNN) approach. The system used in this study consists of four different states with different types of seismic properties designed. To analyze the effect of each seismic property on approach execution. In addition, predict the porosity estimation of each case special features apply cases with the features applied. The cases approach has higher accuracy in predicting the porosity estimation, and the prediction accuracy may change due to the added features to increase seismic quality.

- ❖ **Anifowose, et al (2019**) [3]. This study used four types of (ML), which are Artificial Neural Network (ANN), Functional Network (FN), (SVM), and Decision Tree (DT). Demonstrate the effectiveness of these techniques in handling large amounts of seismic data. , which aims to estimate the porosity and predict the permeability of the reservoir. Therefore, from the point of view of the study, comparing the results with implementation criteria such as correlation coefficient (CC), root mean absolute error (RMAE) and mean absolute error (MAE) gives better results, it was discovered that SVM, when applied to seismic data, has high accuracy and depth matching. This leads to a significant difference in the results compared to other technologies, it positively affects the efficiency and quality of exploration and production. The study also showed that ANN has more smoothing power than FN with SVM performance. No heterogeneity was found with FN and DT. Porosity estimation and prediction of reservoir permeability were not very effective because five or more traits were used.

- ❖ **Qitao Zhang, et al (2019) [12]**. This study presents a method for predicting the spatial distribution of reservoir saturation using machine learning. This study used (LSTM) to predict the water saturation distribution. In addition, using data from actual and simulated monitoring of reservoirs. To get a better prediction of water saturation in rocks, the study compared RNN and (GRU), which are popular machine learning algorithms, with LSTM. The results showed that the LSTM method improved other machine learning methods and the fluid crowding prediction pattern. This study presented an alternative method to predict the water saturation distribution in reservoirs quickly and reliably. The LSTM can deal with questions location prediction problems.

❖ **Wei Liu, et al (2020) [13].** This study uses a numerical simulation method to predict oil production. Three prediction values have been proposed using the empirical ensemble decomposition method EEMD in LSTM, ANN, and SVM. The oil production chain in Chinese oil fields was selected as an experimental study. In base petroleum production, the data set must first be divided into training and testing. Then, the test set data is gradually added to the training set and analyzed by (EEMD) to obtain multiple intrinsic mode functions (IMFs). Then an appropriate number of constants (IMFs) are chosen as predictive variables for machine learning. In two real oil fields, the proposed evaluation and verification model was applied to the three values. The experimental results show that the proposed method can provide near-perfect predictions using LSTM over other algorithms.

❖**A. Ogbamikhumi et al (2021)**[14]. In this study conducted to predict reservoir properties, seismic reflection was combined with an artificial neural network (ANN), to predict fluid saturation and improve porosity. Using neural network techniques (NN) and multilayer feed neural networks (MLFN) and probabilistic neural networks (PNNs) computed from target characteristics where reservoir properties performance for porosity prediction predicted from seismic reflection. The expected attributes of the seismic data are related to the characteristics of the reservoir to test the accuracy of the process. The results gave good correlations for MLFN and PNN per well with a mean (CC) of 0.69 and 0.96, respectively, which indicates the evolution of PNN over MLFN.

1.3 Related Work Analyzing

Through the analysis of related works that there are similarities and differences between the previous studies and the current study as follows:

1. The studies are similar to the current study in terms of the following sides:

    - Dealing with porosity and permeability and subjecting them to experiments using a machine learning approach.
    - The type of data usage is the same in the current study, which are seismic data as well as similar in that they are a reliable source in predicting reservoir porosity.

    - The previous studies dealt with many of the experiences that the researcher benefited from in our current study.

2. The studies differ from the current study in terms of the following sides:

    - Selection of experimental characteristics. Where the studies used the design of experimental characteristics according to the well log, while the researcher in the current study used the characteristics of seismic data according to the sequence of the time signal.
    - Dealing with the geological diversity of the Earth and subjecting it to experiments through the proposed model, which was achieved by the proposed system in the current thesis.

1.4 Problem Statement

The main problem facing oil exploration is how to add big data for prediction, exploration, and production. These studies used ML and DL techniques to determine drilling accuracy and reliability, reduce uncertainty and reduce costs, and this is called the use of smart systems and machine learning algorithms in research development, drilling, and production. This thesis will discuss two issues.

✓ Data is the number one problem: Iraq still has flaws in complex calculations. Description of seismic survey data.

✓ Porosity is the second issue: the porosity of the oil tank is very important. Porosity estimation should be very good in tanks and oil tanks should have high reliability before drilling.

1.5 Aim of Thesis

The main objective of this thesis is to design and implement an efficient and effective approach to reservoir porosity prediction based on temporal sequential data processing, ML, and DL techniques to achieve a high degree of accuracy, as well as to compare these techniques to determine the best among them.

1.6 Contribution

The main contribution of this thesis is the application of the oil reservoir porosity prediction system. However, the new contribution to this thesis uses an intelligent system based on seismic data. Another contribution to this thesis is the use of seismic data from a well under exploration and drilling.

1.7 Outline of the Thesis

In this work, the Thesis **" *Prediction Of Reservoirs Porosity Based On Resulting Seismic Data Attributes Using Deep Learning Approach* "** is structured in five chapters; here is a brief description of their contents is given:

Chapter 2: This chapter provides theoretical backgrounds and an overview of reservoir engineering and seismic data. ML and DL are explained with their respective sections. In addition, how the proposed systems can be used through the ML and DL approaches, with an explanation of all the algorithms used in the proposed approach with all the examples and detailed equations.

Chapter 3: This chapter details the proposed approach introduces the proposed main system and design objectives and covers seismic data features that predict porosity.

Chapter 4: This chapter gives presents the results and tests of the proposed system. and experimental results obtained from the implementation of the proposed system.

Chapter 5: This chapter includes conclusions and future work for the development of use seismic data attributes to predict porosity approaches with lists several suggestions for future studies.