People’s Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Kasdi Mebah Univesity, Ouargla

Faculty of new information and communication technologies

Computer science department

Dissertation:
ACADEMIC MASTER
Domain:
Computer Science
Major:
Industrial Computer Science

THEME

Deep learning approach for predicting the lithofacies using standards and measurements taken during the drilling process

Evaluation Date:24/06/2024

Study carried out by:

Bennouna Samah
Melouah Roumaissa Nour-Elhouda

Befor the Jury :

Dr. Zerdoumi .Oussama	President	UKMOuargla
Dr.Bekkari . Fouad	Supervisor	UKM Ouargla
Dr.Kaoudja .Zineb	Examiner	UKM Ouargla

Academic Year:2023/2024

الاهداء

" بسم خالقي وميسر اموري وعصمت أمري ، لك كل الحمد والامتنان "

اهدي هذا النجاح لنفسي اولاً ثم الى كل من سعى معي لإتمام هذه المسيرة، دمتم لي سنداً لا عُمر له

من وكلله الله بالهيبة والوقار. إلى من أحمل اسمه بكل فخر. إلى من حصد الأشواك عن دربي ليمهد لي طريق العلم بعد ما أنا فيه يعود إلى - أبي - الرجل الذي سعى طوال حياته لكي نكون أفضل منه **- أبي الغالي -**

إلى ملاكي في الحياة إلى معنى الحب وإلى معنى الحنان والتفاني.. إلى بسمة الحياة وسر الوجود

إلى من كان دعاؤها سر نجاحي وحنانها بلسم جراحي التي كانت لي الأم والأخت والصديقة وداعمي الأول ووجهتي التي استمد منها القوة **- أمي الحبيبة** -.

إلى مصدر قوتي، الداعمين **الساندين اخواني،** الى اختي وحبيبتي من كانت دوما موضع اتكاء عثرات حياتي

صفاء

الى كل الاهل والعائلة الكريمة كل باسمه ومقامه... **وبالأخص طاطا سعيدة**.

إلى من هونوا تعب الطريق إلى من شجعوني على المثابرة وإكمال المسيرة إلى

رفقتي السنين وسندي فالدراسة ... **سميحة ونورالهدى** انا ممتنة لكم.

إلى كل من ساعدنا في إنجاز هذا العمل ولو بالنصيحة.

لله الشكر كله ان وفقني لهذه اللحظة، فالحمد لله رب العالمين والصلاة والسلام-

على نبيه الكريم.

-- خريجتكم سماح --

الاهداء

الحمد لله وحده والصلاة والسلام على من لا نبي بعده..ها نحن بصدد انتهاء مسيرتنا الدراسية الجامعية وبمناسبة هذه المناسبة التي تهب إلي بأطيب نسائم الفرح يسعدني أن اهدي هذا العمل المتواضع

إلـــى:

* تلك الروح الطاهرة الراحلة التي بين السماء...التي سقتني بالحب والحنان وكل الرجاء علمتني الصبر والجهاد وتحمل الصعاب..التي دعواتها رافقتني طيلة المشوار...هي غائبة في الحياة ولكنها في قلبي حاضرة **أمي الغالية رحمها الله برحمته الواسعة**

*من دعمني بلا حدود وأعطاني بلا مقابل وقادني للمجد إلى الذي جد وبذل كل جهده ودعمني ماديا ومعنويا من اجل أن اعتلي سلالم النجاح **أبي الحنون**

* من قيل فيه سنشد عضدك بأخيك والذي زرع لديَ روح المثابرة والاجتهاد لأصل إلى ما أصبو إليه أخي العزيز **محمد يونس**

* عزيزات ومن أتكئ عليهم في حزني وفرحي، هن شموع حياتي وكن خير عون لي في دربي وتعلمت منهم حب الحياة أخواتي : **مريم، أنيسة، ليلى وهاجر**

* الطفولة البريئة التي استمد منهم قوتي وأبصر في عيونهم الأمل وأستشعر من نظراتهم حب الخير سكراتي: **رتاج، انتصار ملاك، إلياس، جنى وأحمد سيراج**

* من طابت الأوقات بصحبتها صديقة في المذكرة **سماح بن نونة**

* الأيادي التي لم تبخل عليَ بالعطاء وساندتني في انجاز هذا العمل خاصة **عبد الحميد وعبد الرزاق**

* كل صديقاتي التي جمعني بهم منبر العلم والصداقة وكل من اعرفهم من قريب أو بعيد

ملوح رميصاء نور الهدى

Acknowledgement

This work was undertaken at the Department of Computer Science and Information Technology, Kasdi Merbah Ouargla University. We would like to take this opportunity to express our appreciation.

Firstly, we would like to thank God for giving us the strength and patience to carry out this humble work.

We are immensely grateful to **Dr.Bekkari Fouad** for his continuous guidance, unwavering support, and invaluable advice throughout the realization of this work.

We also extend our gratitude to all the individuals who contributed directly or indirectly to the successful completion of this work.

Abstract:

Drilling is crucial for oil and gas exploration but is costly and time-consuming. Enhancing productivity, reducing hazards, and cutting costs are ongoing challenges. Deep learning has emerged as a promising solution for improving decision-making in drilling operations, requiring less human intervention. Key parameters like surface pressure (SPP), rotary speed (RPM), and rate of penetration (ROP) are essential for creating models that predict subsurface conditions.

Accurate identification of oil reservoirs is vital due to the risk of destruction from mud density. Engineers often rotate the drill bit at the same location, which is inefficient and costly. Real-time identification of reservoir access using drilling data can significantly reduce costs and improve productivity. Machine learning, especially deep learning, is effective in predicting geological layers and lithofacies, enhancing drilling safety and efficiency.

Key words: drilling, drilling parameters, lithofacies, prediction, lithofacies predicting, RNN, LSTM, GRU.

Résumé :

Le forage est crucial pour l’exploration pétrolière et gazière, mais il est coûteux et prend du temps. Améliorer la productivité, réduire les risques et réduire les coûts sont des défis permanents. L’apprentissage profond s’est imposé comme une solution prometteuse pour améliorer la prise de décision dans les opérations de forage, nécessitant moins d’intervention humaine. Des paramètres clés tels que la pression de surface (SPP), la vitesse de rotation (RPM) et le taux de pénétration (ROP) sont essentiels pour créer des modèles qui prédisent les conditions souterraines.

Une identification précise des réservoirs de pétrole est vitale en raison du risque de destruction dû à la densité de la boue. Les ingénieurs font souvent tourner le foret au même endroit, ce qui est inefficace et coûteux. L'identification en temps réel de l'accès au réservoir à l'aide des données de forage peut réduire considérablement les coûts et améliorer la productivité. L’apprentissage automatique, en particulier l’apprentissage profond, est efficace pour prédire les couches géologiques et les litho facies, améliorant ainsi la sécurité et l’efficacité des forages.

Mots clés : forage, paramètres de forage, litho faciès, prédire de litho faciès , prédiction ,RNN,LSTM,RGU.

ملخص:

يعد الحفر أمرًا بالغ الأهمية للتنقيب عن النفط والغاز ولكنه مكلف ويستغرق وقتًا طويلاً. يعد تعزيز الإنتاجية وتقليل المخاطر وخفض التكاليف من التحديات المستمرة. لقد ظهر التعلم العميق كحل واعد لتحسين عملية صنع القرار في عمليات الحفر، الأمر ، ومعدل الاختراق(RPM) ، وسرعة الدوران(SPP) الذي يتطلب تدخلًا بشريًا أقل. تعد المعلمات الرئيسية مثل الضغط السطحي .ضرورية لإنشاء نماذج تتنبأ بالظروف تحت السطح (ROP)

يعد التحديد الدقيق لخزانات النفط أمرًا حيويًا نظرًا لخطر التدمير الناتج عن كثافة الطين. غالبًا ما يقوم المهندسون بتدوير لقمة الحفر في نفس الموقع، وهو أمر غير فعال ومكلف. يمكن أن يؤدي تحديد الوصول إلى الخزان في الوقت الفعلي باستخدام بيانات الحفر إلى تقليل التكاليف بشكل كبير وتحسين الإنتاجية. يعد التعلم الآلي، وخاصة التعلم العميق، فعالاً في التنبؤ بالطبقات الجيولوجية والسحايا، مما يعزز سلامة وكفاءة الحفر

الكلمات المفتاحية: الحفر، معلمات الحفر، السحنات، التنبؤ بالسحنات، التنبؤ

Table of Contents

List of figures

List of tables

List of Abbreviations

RP : Rotation Per Minute (Drill string rotary speed)

ROP : Rate of penetration

SPP: Standpipe pressure

TOB : Torque-on-bit

USD : United States dollar

WOB : Weight-on-bit

NN : Neural Network

ANN : Artificial Neural Networks

RNN : Recurrent Neural Networks

GRU : Gated Recurrent Unit

LSTM : Long Short-Term Memory

CNN : Convolutional Neural Networks

AI : Artificial Intelligence

DL : Deep Learning

ML : Machine Learning

MAE : Mean Absolute Error

MSE : Mean Squared Error

General introduction

The exploration for oil and gas is a costly and time-consuming process. Improving efficiency, reducing risks, and cutting costs are crucial challenges facing the industry. Recently, deep learning techniques have shown promising results in enhancing decision-making in drilling operations by reducing the need for extensive human intervention.

This study includes 5 wells oilfield.

Location: BASSIN AMGUID MESSAOUD

PERIMETRE : TOUGGOURT EAST

Objectives of the dissertation

This study investigates the use of Recurrent Neural Networks (RNN) to predict rock layers meters away using data collected during the drilling process.

Organization of the work

This work was structured as follows:

Chapter I: Definitions and presentations of drilling parameters and

focuses on reservoir properties , detailing geological formations, rock and fluid properties, thereby enhancing understanding of the drilling environment and reservoir conditions.

Chapter II: explores the fundamentals of artificial intelligence, including neural network structure, training methods, and practical applications, with a particular emphasis on RNN and its variations such as Simple RNN, LSTM, and GRU, highlighting their ability to process and predict time-series data and continuous changes.

Chapter II: Applies deep learning techniques to drilling data to predict rock layers meters away. Various models were created and tested, with LSTM models showing exceptional performance. The results confirm the effectiveness of LSTM in predicting rock layers, emphasizing their value in improvement.

The general conclusion will summarize all the results obtained .

Chapter I: General informations on drilling and oil reservoirs.

I.1 Introduction

The petroleum industry plays a crucial role in powering global economies and meeting the world's energy needs. Within this industry, the most critical aspect lies in exploration and drilling, as these processes are essential for discovering and extracting valuable oil and gas reserves that drive economic growth and development worldwide. Exploration and drilling require advanced technologies and equipment to analyze complex geological data Additionally, the use of sophisticated machinery and tools is necessary to ensure the efficiency and safety of drilling operations, adding to the overall costliness of the process. Add to the need for skilled human resources, and effective planning and management and our work focuses on attempting to reduce the overall costs of the drilling process. cost price.

I.2 Definition of drilling

Drilling, in its most expansive sense, refers to any method employing a cutting tool to bore a round aperture in a solid material or resistant surface. Widely considered one of humanity's most remarkable innovations, drilling boasts myriad applications across numerous industries.[14][10]

I.3 The drilling processes

Upstream operations represent a virtual part of exploration and production operations in industries such as oil, gas, and mining. These processes involve the use of a variety of restraints and tools to stimulate and penetrate the ground in an efficient and safe manner.

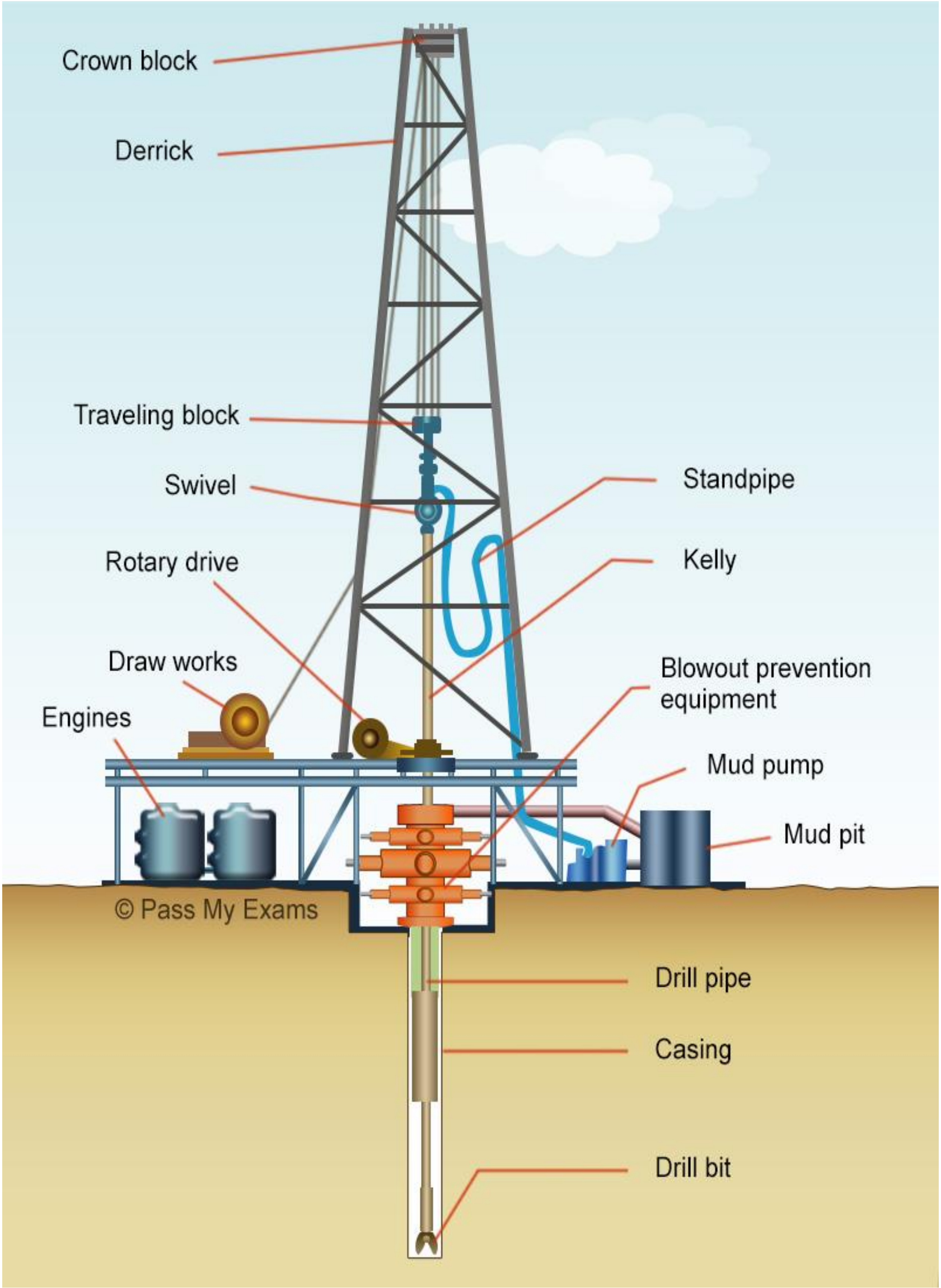


Figure 1: Drilling Rig

1. Planning and Exploration:

Based on geological and geophysical evaluations, possible drilling locations are identified to start the drilling process.

- A drilling plan is created, detailing the tools to be utilized and the anticipated results. [2]

2. Preparation:

- Equipment and materials are transported to the drilling location.

- The wells are ready for drilling operations, and the drilling platform or rig is installed (rig up).[2]

3. Primary Drilling:

- To start primary drilling and make a first well, a drill bit is utilized.

- Mud pumps are used to remove soil and rocks (to pump fluid). [9]

4. Evaluation and Monitoring :

- To locate possible oil and gas deposits, the excavated rocks and soil are examined.

- To guarantee operational safety and prevent technical problems, drilling operations are constantly observed. [2]

5. Secondary Drilling:

- Secondary drilling is done to extend the well and look for additional resources if possible places for oil and gas are discovered. [2]

6. Extraction and Production:

- Extraction pipelines are used to extract oil and gas once sufficient quantities have been established.

- To be used and transported, the oil and gas are processed. [2]

7. Evaluation and Analysis :

- To examine the outcomes and pinpoint areas for future development, a thorough evaluation of the entire process is conducted. [9]

I.4 A drilling rig system

Drilling rigs vary widely in their external design and how they are used, however all rotary rigs have the same basic drilling equipment. The following are the principal parts of a rotary rig:

• Power system (Generators, VFD, MCC)

• Hoisting system

• Circulating system

• Rotary system

• Well control system [15]

The derrick, the draw works with its drilling line, crown block, and moving block, and the drilling fluid circulation system with its standpipe, rotating hose, drilling fluid pits, and pumps make up the majority of the rig. Together, these parts enable rotary rigs to perform the three primary tasks.[11]

Drilling methods:

There are two main types of drilling methods in mining engineering: rotary and percussive. Rotary drilling involves rotating a drill bit at a high speed to cut through the rock. Percussive drilling involves hammering a drill bit into the rock with a rapid succession of blows. Each method has its own advantages and disadvantages, depending on the rock characteristics, the hole diameter, the depth, and the purpose of drilling. [10]

I.5 Drilling well problems

1. Pipe sticking (stuck Pipe)

Pipe sticking is one of the most common problems faced during drilling that causes a lot of nonproductive time (NPT). Pipe is considered stuck if it cannot be freed and dragged out of hole without damaging the pipe or surpassing the maximum allowed hook. [10] [2]

2. Loss of circulation (mud loss)

Loss of circulation is the undesirable flow of portion or whole mud into the formation. There are two types of circulation loss:

-Partial loss, at which only portion of the mud flow into the formation and the rest flows to the surface.

-Total loss, at which the whole mud flows into the formation with no return to the surface. [15]

3. Hole deviation

Hole deviation is defined as the unintentional departure of the path of the drill bit from the preselected one Several factors are responsible for the deviation including:

• Heterogeneity of the formation

• Bottom hole assembly characteristics

• Stabilizers

• Weight on bit

• Hole inclination angle

• Drill bit type and design

• Hydraulics of the bit

• Improper hole cleaning [4]

4. Borehole instability

Discussed borehole instability and its parameters indicating that it is geo-mechanical issue that is related to hydraulic and chemical factors.

5. Hydrogen sulfide bearing zones (H2S)

Hydrogen-sulfide anticipated during drilling operations can involve a serious consequence for personnel and equipment. Exposure to a relatively low concentrations of H_2S for a short period of time may cause health problems or even death to personnel in site, in addition to be very corrosive to equipment in presence of oxygen. [4]

I.6 Geological Rocks

In the field of geology, a rock is defined as a naturally formed and cohesive assemblage of one or more minerals. These assemblages serve as the fundamental structural and compositional components of the lithosphere, comprising the solid outer shell of the Earth. [8]

Rock consists of grains bound together by a bunch of material such:

Clay: is a naturally occurring rock material composed primarily of hydrated silicates or aluminosilicates with a layered structure.[5]

Sandy clay: a type of clay soil that contains a significant proportion of sand particles.

Sandy Ochre Clay: Is similar to sandy clay but has an ochre coloration, which is often due to the presence of iron oxide minerals.[3]

Silt: is a fine-grained sedimentary material consisting of particles smaller than sand but larger than clay. It is often deposited by moving water and can be found in riverbeds, floodplains, and coastal areas. Silt soils are fertile and well-draining, making them suitable for agriculture.[5]

Fine Sandstone: is a sedimentary rock made up of compacted sand grains. It has a fine-grained texture and may contain various minerals, giving it different colors and patterns. Fine sandstone is often used in building materials and decorative applications. [3]

Medium Sandstone: is similar to fine sandstone but with slightly larger sand grains. It is commonly found in sedimentary rock formations and can be used for construction, paving, and landscaping. [3]

Dolomitic Limestone: is a limestone containing a high proportion of magnesium carbonate (dolomite). It forms through the replacement of calcium carbonate by magnesium-rich fluids. [3]

Dolostone: dolostone, also known as dolomite rock, is a sedimentary rock composed mainly of dolomite. It forms through the chemical alteration of limestone by magnesium-rich fluids. [3]

I.7 Reservoir Rock

These rocks provide appropriate storativity and conductivity for accumulating and flowing hydrocarbon. To evaluate and understand reservoir behavior and also improvement of reservoir performance, studying reservoir rock properties is vital. [1]

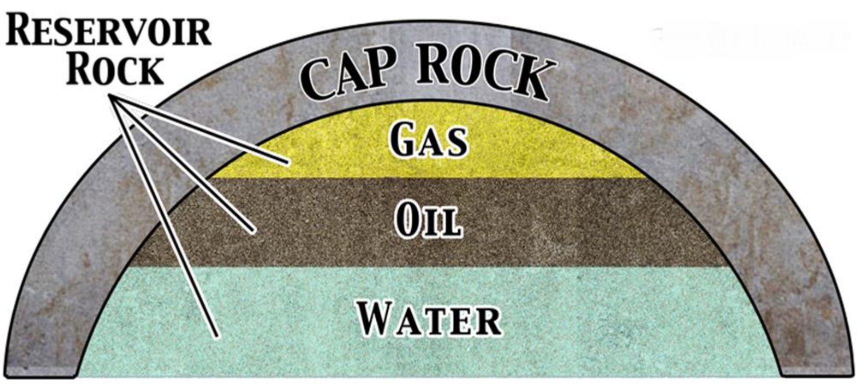


Figure 2: Reservoir Rock

Most reservoir rock properties are determined by lab-based works. In order to perform experimental tests, the reservoir rock should be sampled.

The special sample of reservoir rocks is called the core. The lengths of cores are varied, from a few inches in core plugs to several meters in whole cores. [3]

I.8 Conclusion:

In this chapter, it has been outlined the fundamentals of drilling operations, discussed the essential components of drilling systems, and explained their roles and functions. examined various systems such as the hoisting system, circulation system, and draw works, emphasizing their significance in the drilling process. Additionally, the challenges associated with drilling,was highlighted including efficiency, risk mitigation, and cost reduction. The next chapter will focus on deep learning and artificial intelligence, exploring the advanced techniques and methodologies that will be applied in our project. It will delve into the ways these technologies can be leveraged to predict lithofacies, optimize drilling operations, and enhance overall drilling efficien

Chapter II: The basics of artificial intelligenc

II.1 Introduction

In the era of modern digital technology, artificial intelligence and machine learning technologies are considered one of the most important fields witnessing rapid development and significant impact on our daily lives and various industries. Among these technologies, deep learning stands out as an important and advanced subfield in the field of machine learning.

Due to the recent trend of intelligent systems and their ability to adapt with varying conditions, deep learning becomes very attractive for many researchers. In general, neural network is used to implement different stages of processing systems based on learning algorithms by controlling their weights and biases.

II.2 Machine learning

II.2.1 The main components of machine learning

Any system using machine learning will need three main components, which are:

Data, features and Algorithm[6]

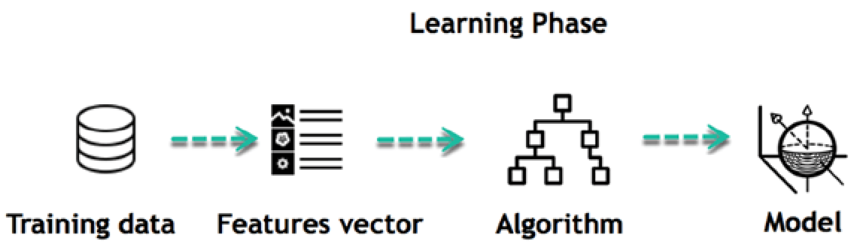


Figure 3:The main components of ml

II.2.2 Types of machine learning algorithms

There are four types of machine learning algorithms: supervised, unsupervised, semi-supervised, and reinforcement.

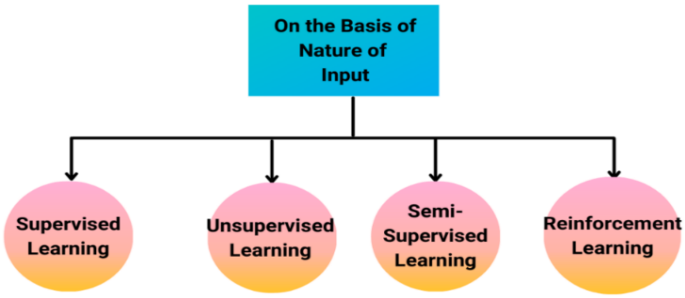


Figure 4:Types of ML algorithme

2.2.1 Supervised learning algorithms

Supervised learning can be separated into two types of problems when data mining: classification and regression. Some popular Regression algorithms are given below:

- Artificial neural networks
- Decision tree algorithms
- K-nearest neighbor
- Linear regression
- Logistic regression
- Naïve Bayes
- Random forests.
- Support vector machines (SVM). [16]

2.2.2 Unsupervised learning algorithms

Unlike supervised learning, unsupervised learning uses unlabeled data. From that data, the algorithm discovers patterns that help solve clustering or association problems. Some popular Regression algorithms are given below:

- Clustering.
- Hierarchical clustering.
- K-means clustering.
- Applications of Unsupervised Learning.
- Network Analysis.
- Recommendation Systems.
- Anomaly Detection.
- Singular Value Decomposition. [6]

2.2.3 Semi-supervised learning algorithms

Semi-supervised learning algorithms operate when only a portion of the input data is labeled, giving the algorithm a slight "head start." This approach combines the advantages of both supervised and unsupervised machine learning. It offers improved accuracy associated with supervised learning while also utilizing cost-effective, unlabeled data typical of unsupervised learning. [16]

2.2.4 Reinforcement algorithms

In this context, the algorithms are trained in a manner similar to human learning-through rewards and penalties. These are measured and tracked by a reinforcement learning agent, which has a general understanding of the probability of successfully increasing the score versus decreasing it. Through trial and error, the agent learns to take actions that lead to the most favorable outcomes over time. Reinforcement learning is often used in resource management, robotics, and video games. [16]

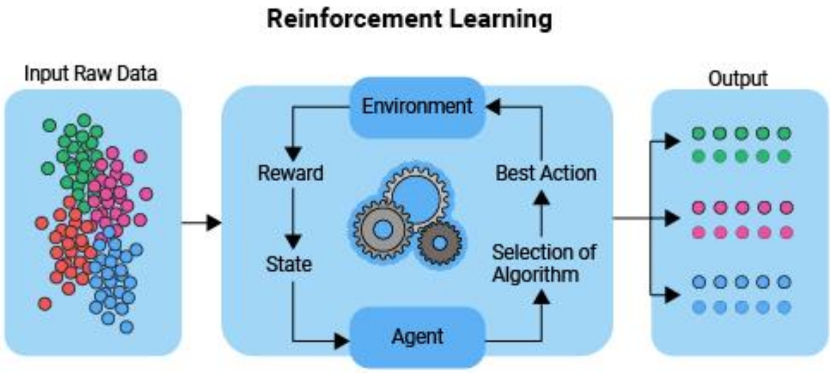


Figure 5:Reinforcement algorithms

II.3 Deep learning

Deep learning is based on the branch of machine learning, which is a subset of artificial intelligence. Since neural networks imitate the human brain and so deep learning will do. Nothing is explicitly programmed in deep learning. In essence, it is a machine learning class that does feature extraction and transformation by utilizing several nonlinear processing units. All subsequent levels accept as input the output from the layer that comes before it. With minimal assistance from the programmer, deep learning models can concentrate on the accurate features on their own and are highly beneficial in resolving the dimensionality issue.[17] In particular, deep learning algorithms are employed when there are a large number of inputs and outputs. Artificial neurons, sometimes referred to as nodes, make up a neural network, which is organized similarly to the human brain. [7]

These nodes are stacked next to each other in three layers:

- The input layers.
- The hidden layers.
- The output layers.

II.3.1 Deep learning applications

There are many uses for deep learning, such as the following:

- Image recognition: The process of recognizing features and objects in pictures, including people, animals, locations, etc.
- Natural language processing: Used in spam filters and chatbots for customer support to assist in deciphering textual content.
- Finance: To assist in the analysis of financial data and the forecasting of market trends
- Text to image: Use the Google Translate app, for example, to convert text to images. [20]

II.4 Artificial neural networks ANN

II.4.1 Definition of artificial neural networks

Artificial Neural Networks (ANNs) are computing systems inspired by the biological neural networks of animal brains. They consist of interconnected nodes, or "neurons," that process and analyze complex data. [25] ANNs recognize patterns, learn from data, and make decisions or predictions, making them widely used in fields such as image and speech recognition, medical diagnosis, and financial forecasting. Their ability to model complex, nonlinear problems without prior knowledge of data distribution enhances their applicability across various domains.[19]

II.4.2 Neural network structure

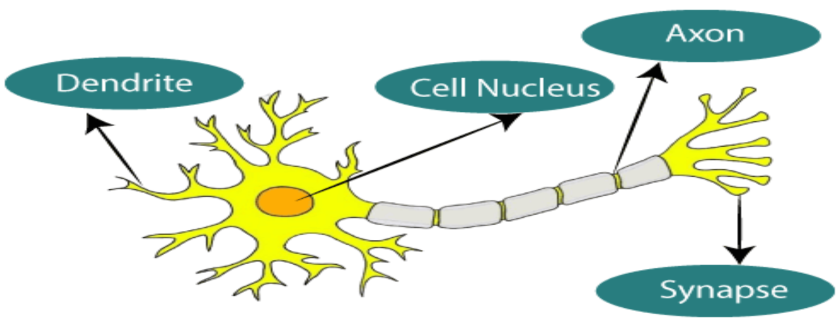


Figure 6: diagram of Biological Neural Network.

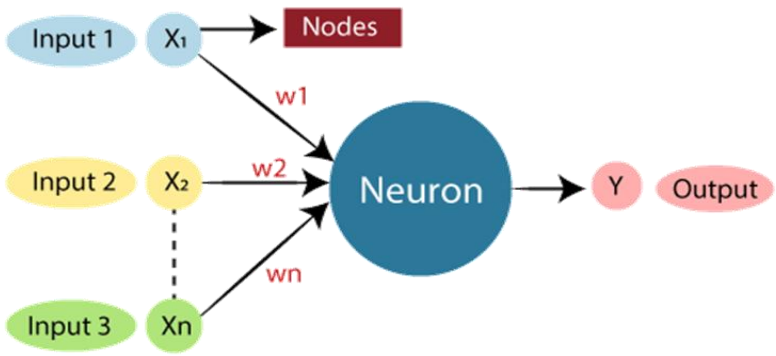


Figure 7: The typical Artificial Neural Network .

Artificial neural networks, dendrites from biological neural networks stand in for inputs, cell nuclei for nodes, synapses for weights, and axes for outputs.

Relationship between Biological neural network and artificial neural network:

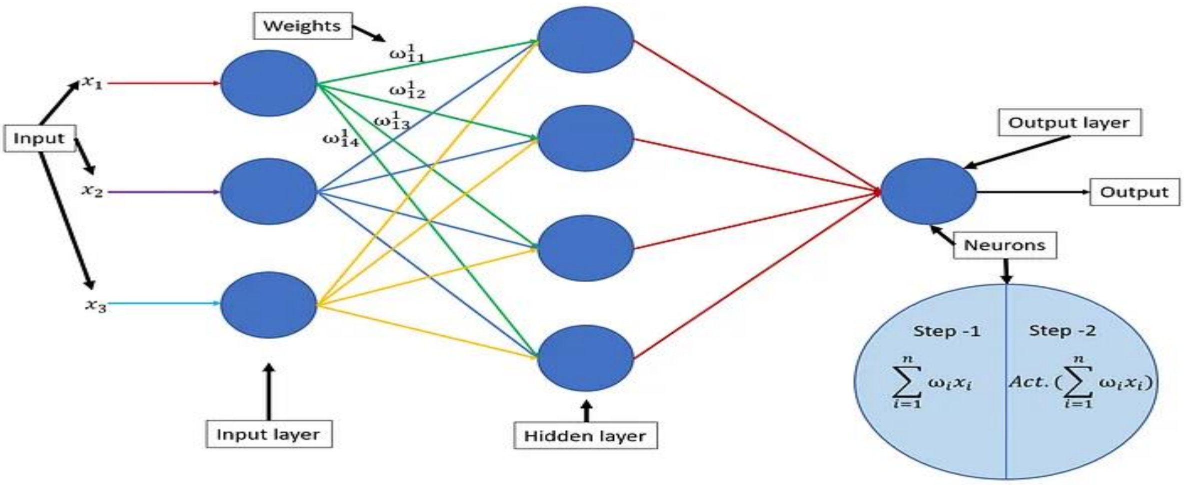


Figure 8: Artificial Neural Network

In the definitions of neuron and artificial neural network, several key terms emerge, including input, input layer, weights, hidden layer, output layer, output, and activation functions.

Activation functions add an extra computational step at each layer during forward propagation. Despite this additional computation, they significantly improve the network's performance, making the effort worthwhile [25]

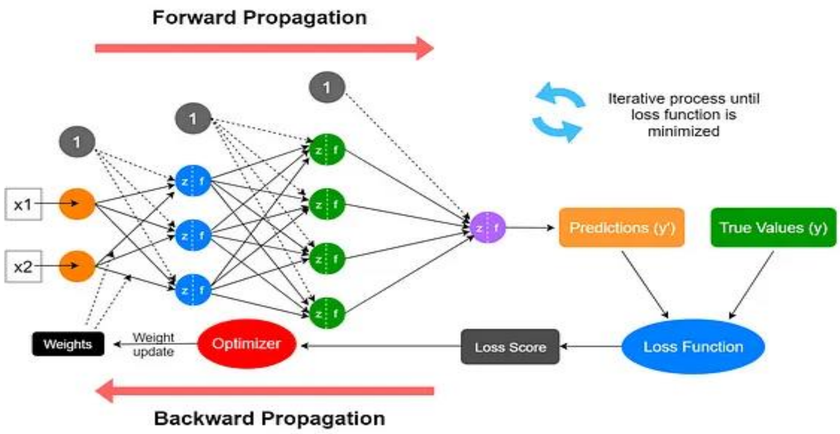


Figure 9:Neural Networks-Forward pass and Backpropagation

There are three primary components to the overall learning process:

1. Forward Passage, or Forward Propagation

2. The loss function calculation

3. Backpropagation, also known as backward pass or backpropagation

The network receives the input data and feeds it forward. After receiving the input data and processing it in accordance with the activation function, each hidden layer moves on to the next layer. [19]

II.4.3 Backpropagation

The networks error is assessed when the forward pass is finished, and it should ideally be reduced.

A high mistake rate in the present process suggests that the network has not properly learned from the data. Stated differently, the existing set of weights lacks the necessary precision to minimize error and produce accurate predictions. As a result, in order to lower the error, the neural network weights must be updated.

Finding the gradients of Loss in relation to the aspects of our neural network that we can modify is the goal of backpropagation.

To put it simply, backpropagation makes a backward pass through a network after each forward run while modifying the models parameters (weights and biases). [13]

II.4.4 Loss Function

The loss function, also referred to as the error function, is a crucial component in machine learning that quantifies the difference between the predicted outputs of a machine learning algorithm and the actual target values.

II.5 Types of Loss Functions

Machine learning loss functions can be divided into groups according to the types of tasks that they might be used for. Machine learning problems involving regression and classification can be addressed by most loss functions. [18]

II.5.1 Loss Functions for Regression

a) Mean Square Error (MSE):

$$MSE = (1/n) * \Sigma (y_i - \bar{y})^2 \qquad (II\text{-}1)$$

Where:

- n : the number of samples in the dataset.
- y_i ; the predicted value for the i-th sample.
- $\bar{y}$: the target value for the i-th sample.

b) Mean Absolute Error (MAE):

$$MAE = (1/n) * \Sigma|y_i - \bar{y}| \qquad (II\text{-}2)$$

Where:

- n : the number of samples in the dataset.
- y_i : the predicted value for the i-th sample.
- $\bar{y}$: the target value for the i-th sample.

II.5.2 Loss Functions for Classification

a) Binary Cross-Entropy Loss / Log Loss:

$$L (y, f(x)) = -[y * log(f(x)) + (1 - y) * log(1 - f(x))]II\text{-}3)$$

Where:

- L: represents the Binary Cross-Entropy Loss function.
- y : true binary label (0 or 1).
- f(x) : predicted probability of the positive class (between 0 and 1).

- log(.): represents the natural logarithm.

b) Categorical Cross-Entropy (CCE) :

For multi-class classification

$$CCE = - (1/n) * \Sigma\Sigma\, y_{ij} * \log(\hat{y}_{ij}) \qquad (II\text{-}4)$$

Where:

- n : number of samples in the dataset.
- y_{ij} : actual (ground truth) label of sample i for class j, where y_{ij} = 0 or 1.
- $\hat{y}_{ij}$:predicted probability of sample i for class j, where the probabilities for all classes sum up to 1.
- log(.): represents the natural logarithm.
- Σ : represents the sum over all i samples in the dataset.
- ΣΣ : represents the sum over all classes j. [18]

II.6 Convolutional Neural Network CNN

Convolutional Neural Networks (CNNs) are specialized neural architectures are primarily employed for a variety of computer vision tasks, including object detection and image classification. These neural networks use convolution operations to use the strength of linear algebra to find patterns in photos. [20]

CNN components in general it contains the following layers:

- Input layer.
- convolutional (Conv) layer.
- Pooling layer.
- Fully connected (FC) layer.
- Softmax/logistic layer.
- Output layer.

II.6.1 Architecture of CNN

All CNN models follow a similar architecture, as shown in the figure below.

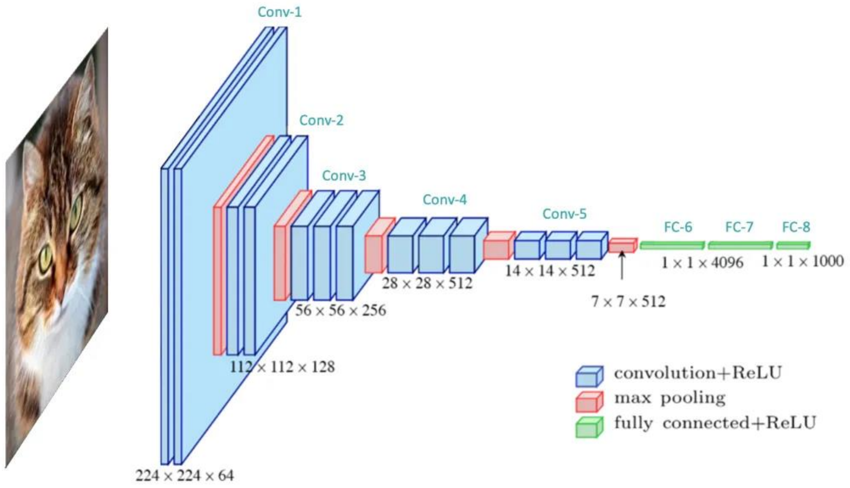


Figure 10:Architecture of CNN

From linear algebra, particularly convolution operations, to extract features and identify patterns within images. Although CNNs are predominantly used to process images, they can also be adapted to work with audio and other signal data.[26]

II.6.2 CNN components in general It contains the following layers

- Input layer
- Convolutional (Conv) layer
- Pooling layer
- Fully connected(FC) layer
- Softmax/logistic layer
- Output layer

II.6.3 Training CNN

CNN is trained via gradient descent and backpropagation, just like ANNs. This is outside the purview of this study because it involves more mathematics due to the convolution technique.

II.7 Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNNs) are specialized for processing sequential data by maintaining a cyclic connection structure that captures temporal dependencies. Unlike feedforward networks, RNNs use a hidden state to retain information from previous time steps, enabling contextual understanding during current input processing. Traditional RNNs struggle with the vanishing gradient problem, where gradients diminish over time, impairing their ability to learn long-term dependencies. [21] To address this, Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) were developed. LSTMs introduce gating mechanisms that manage information flow within memory cells, enhancing the network's capability to handle long-range dependencies effectively. GRUs offer a simpler architecture yet achieve comparable performance in various tasks, making both LSTM and GRU vital tools in sequential data processing tasks like language modeling and time series prediction.[30]

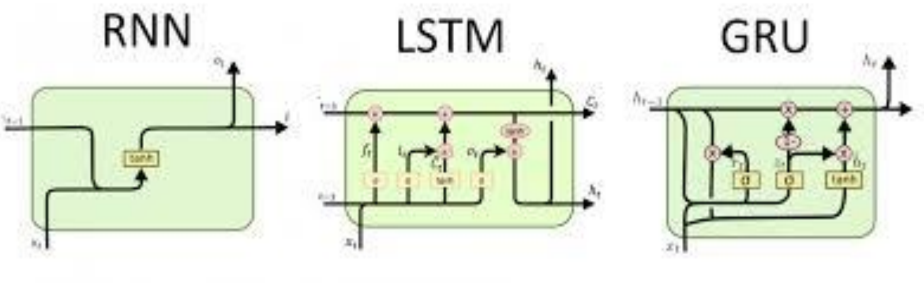


Figure 11: Recurrent Neural Networks (RNNs)

II.7.1 The importance of return neural networks:

Recurrent Neural networks stand out in artificial intelligence and machine learning for their efficient pattern recognition and dynamic capabilities. However, conventional neural networks struggle with connecting events over time. While they excel in tasks like fruit classification, predicting stock prices or auto-completing texts challenges them. Recurrent Neural Networks (RNNs) address this limitation by incorporating memory, making them adept at predicting sequences. This feature unlocks new avenues for advancing machine learning approaches.

II.7.2 Architecture and working of RNN:

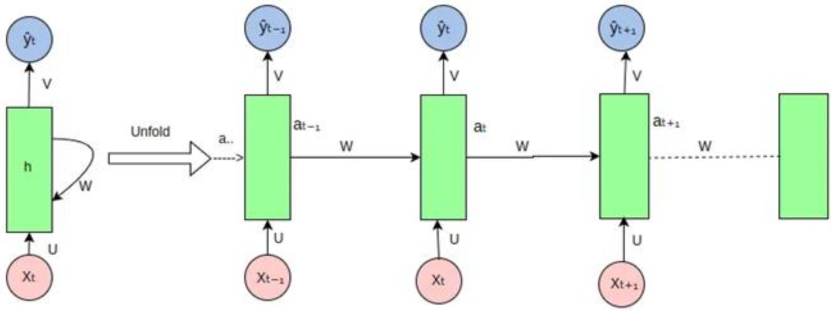


Figure 12:Architecture of RNN

The RNN takes an input vector X and the network generates an output vector y by scanning the data sequentially from left to right, with each time step updating the hidden state and producing an output. It shares the same parameters across all time steps. This means that, the same set of parameters, represented by U, V, W is used consistently throughout the network. U represents the weight parameter governing the connection from input layer X to the hidden layer h, W represents the weight associated with the connection between hidden layers, and V for the connection from hidden layer h to output layer y. This sharing of parameters allows the RNN to effectively capture temporal dependencies and process sequential data more efficiently by retaining the information from previous input in its current hidden state.[30]

The RNN processes an input vector X to generate an output vector y by sequentially scanning the data from left to right. At each time step, the network updates its hidden state and produces an output using the same set of parameters U, V, and W across all time steps. Here's what each parameter represents:

$$a_t = f\,(U * X_t + W * a_{t-1} + b\,) \qquad (\,II\text{-}5\,)$$

where:

- f : the activation function.
- U : the weight matrix governing the connections from the input to the hidden layer; U ∈ θ
- X_t : he input at time step t.
- W : the weight matrix governing the connections from the hidden layer to itself (recurrent connections); W∈ θ
- a_{t-1} : the output from hidden layer at time t-1.
- b : bias vector for the hidden layer; b ∈ θ

The output at each time step t, denoted as y_t is computed based on the hidden state output a_t using the following formula.

$$\hat{y}_t = f\,(V * a_t + c) \qquad (\text{II-6})$$

where:

- $\hat{y}_t$: the output predicted at time step t.V is the weight matrix governing the connections from the hidden layer to the output layer.
- c : the bias vector for the output layer.

II.7.3 Backpropagation Through Time (BPTT)

Backpropagation is the process of updating a model's parameters (weights and biases) to minimize the difference between predicted outputs and actual target values. Its objective is to enhance the model's performance by minimizing the loss function. Specifically for training Recurrent Neural Networks (RNNs), Backpropagation Through Time (BPTT) is employed. BPTT involves propagating the error backward through time, starting from the final time step back to the initial time step (t = 1). The process of backpropagation comprises two essential phases: the forward pass, where inputs are processed through the network to produce predictions, and the backward pass, where errors are propagated backward to adjust the parameters accordingly.[21]

1-Forward Pass:

During forward pass, the RNN processes the input sequence through time, from t=1 to t=n, where n is the length of input sequence. In each forward propagation, the following calculation takes place.[30]

$$a_t = U * X_t + W * a\_(t\text{-}1) + b \qquad (\text{II-7})$$

$$\hat{y}_t = softmax\,(V * a_t + c) \qquad (\text{II-8})$$

After processing the entire sequence, RNN generates a sequence of predicted outputs $\hat{y} = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_t]$. Loss is then computed by comparing predicted output $\hat{y}$ at each time step with actual target output y. Loss function given by:

$$L\,(y, \hat{y}) = (1/t) * \Sigma(y_t - \hat{y}_t)^2 \text{ —— ——— —} \rightarrow MSE \qquad \text{II-9)}$$

2- Backward Pass:

The backward pass in BPTT involves computing the gradients of the loss function with respect to the network's parameters (U, W, V and biases) over each time step.

Let's explore the concept of backpropagation through time by computing the gradients of loss at time step t=4. The figure below also serves as an illustration of backpropagation for time step 4. [30]

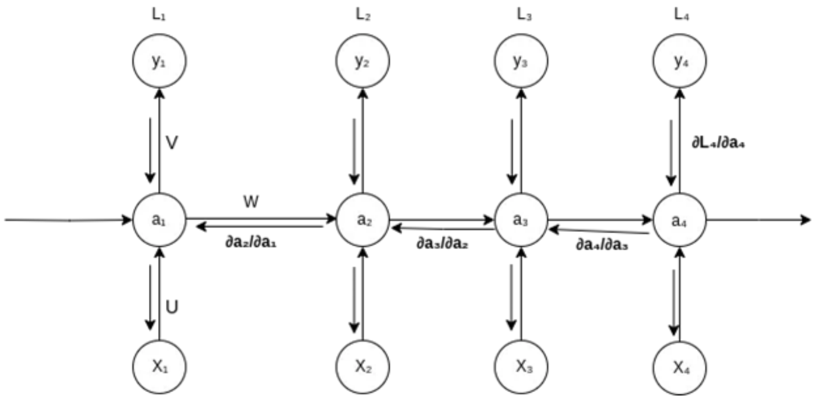


Figure 13:BPTT

3- Derivative of loss L w.r.t V:

Loss L is a function of predicted value ŷ, so using the chain rule ∂L/∂V can be written as:

$$\partial L/\partial V = (\partial L/\partial \hat{y}) * (\partial \hat{y}/\partial V) \qquad (II\text{-}10)$$

4- Derivative of loss L w.r.t U:

Applying the chain rule of derivatives, ∂L/∂U can be expressed as follows: The loss at the 4th time step depends on ŷ, which is determined by the current time step's hidden state a_4. This state a_4 is influenced by both U and a_3. In turn, a_3 is connected to both a_2 and U, and a_2 depends on a_1 as well as on W.

$$\partial L_4/\partial U = (\partial L_4/\partial \hat{y}_4 * \partial \hat{y}_4/\partial a_4 * \partial a_4/\partial U) + (\partial L_4/\partial \hat{y}_4 * \partial \hat{y}_4/\partial a_4 * \partial a_4/\partial a_3 * \partial a_3/\partial U) + (\partial L_4/\partial \hat{y}_4 * \partial \hat{y}_4/\partial a_4 * \partial a_4/\partial a_3 * \partial a_3/\partial a_2 * \partial a_2/\partial U) + (\partial L_4/\partial \hat{y}_4 * \partial \hat{y}_4/\partial a_4 * \partial a_4/\partial a_3 * \partial a_3/\partial a_2 * \partial a_2/\partial a_1 * \partial a_1/\partial U) \qquad (II\text{-}11)$$

Here we're summing up the gradients of loss across all time steps which represents the key difference between BPTT and regular backpropagation approach.[27]

II.7.4 Problems of traditional return networks

1.vanishing gradient problem:

During the training of deep neural networks, the vanishing gradient problem refers to a phenomenon where the gradients used for updating the network diminish significantly or "vanish" as they propagate backward from the output layers to the earlier layers. This issue hinders effective learning of long-term dependencies in the network. [24]

2.Exploding Gradient problem:

The Exploding Gradient problem occurs during neural network training when error gradients accumulate excessively, leading to exponentially large updates in model weights. This issue causes prolonged training times and results in poor performance and accuracy. [24]

II.8 Long Short-Term Memory (LSTM)

LSTM, an advanced type of recurrent neural network developed by Hochreiter and Schmidhuber, is specifically tailored for sequence prediction tasks. It excels in capturing long-term dependencies within data sequences [27]. LSTM is highly effective in applications requiring the modeling of temporal relationships, such as time series analysis, machine translation, and speech recognition. Its robustness lies in its capability to handle complex problems by efficiently preserving and utilizing sequential information. [22]

The main goal of its design was to avoid the vanishing gradient problem that can occur when training traditional RNNs was addressed with the development of LSTMs.

II.8.1 Architecture and Working of LSTM:

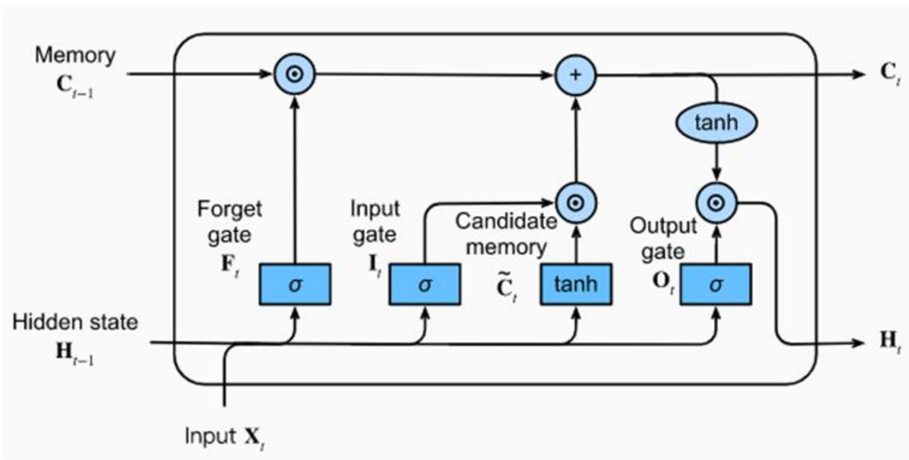


Figure 14:Architecture LSTM

Information is retained by the cells and the memory manipulations are done by the gates. There are three gates

a) Forget Gate

The forget gate in LSTM networks plays a crucial role in managing the cell state by deciding which information to retain or discard. It receives inputs xt (current input at time t) and ht-1 (previous cell output) which are processed through weighted matrices and bias addition. An activation function then produces a binary output: 0 indicates forgetting the information, while 1 signifies retaining it for future use within the cell state. This mechanism enables LSTM to selectively update and maintain relevant information across sequential data. [23]

The equation for the forget gate is:

$$f_t = (w_f \; [h_{t-1}, x_t] + b_f) \qquad (II\text{-}12)$$

where:

- w_f: represents the weight matrix associated with the forget gate.
- $[h_{t-1}, x_t]$: denotes the concatenation of the current input and the previous hidden state.
- b_f: bias with the forget gate.
- σ :sigmoid activation function

b) Input gate

The input gate in LSTM updates the cell state by filtering and integrating new information. It uses sigmoid activation to decide what to retain based on current and previous inputs (ht-1 and xt). A tanh function creates a vector of potential values, which is then multiplied with the regulated information to incorporate useful updates into the cell state.[23]

The equation for the input gate is:

$$i_t = (Wi[h_{t-1}, x_t] + bi) \qquad (II\text{-}13)$$

$$\hat{c}_t = tanh\,(w_c[h_{t-1}, x_t] + b_c) \qquad (II\text{-}14)$$

We multiply the previous state by ft, disregarding the information we had previously chosen to ignore. Next, we include it∗Ct. This represents the updated candidate values, adjusted for the amount that we chose to update each state value.

$$c_t = f_t\, C_{t-1} + i_t \hat{c}_t \qquad (II\text{-}15)$$

where

- $\odot$ denotes element-wise multiplication
- tanh is tanh activation function

c) Output gate

The LSTM output gate extracts essential information from the current cell state. It uses tanh on the cell state to create a vector. This vector is filtered and regulated by the sigmoid function based on inputs ht-1 and xt to decide what to retain. Finally, the resulting values are multiplied to produce the output sent to the next LSTM cell. [23]

The equation for the output gate is:

$$o_t = (w_0[h_{t-1}, x_t] + b_0) \qquad (II\text{-}16)$$

II.8.2 Advantages and Disadvantages of LSTM

a) Advantages:

- Handling Long Sequences.
- Avoiding Vanishing Gradient Problem.
- Handling Variable-Length Sequences.
- Memory Cell.
- Gradient Flow Control. [27]

b) Disadvantages:

- Computational Complexity.
- to overfitting.
- Hyperparameter Tuning.
- Limited Interpretability.
- Long Training Times. [27]

II.8.3 Applications of LSTM

- Some of the famous applications of LSTM includes:
- Language Modeling
- Speech Recognition
- Time Series Forecasting
- Anomaly Detection
- Recommender Systems
- Video Analysis[29]

II.9 Gated Recurrent Unit or (GRU)

GRU, or Gated Recurrent Unit, represents an evolution of the traditional RNN (Recurrent Neural Network). It shares similarities with LSTM in utilizing gates to manage information flow. Compared to LSTM, GRU is relatively newer and offers some enhancements while maintaining a simpler architecture. [28]

II.9.1 Architecture and Working of GRU:

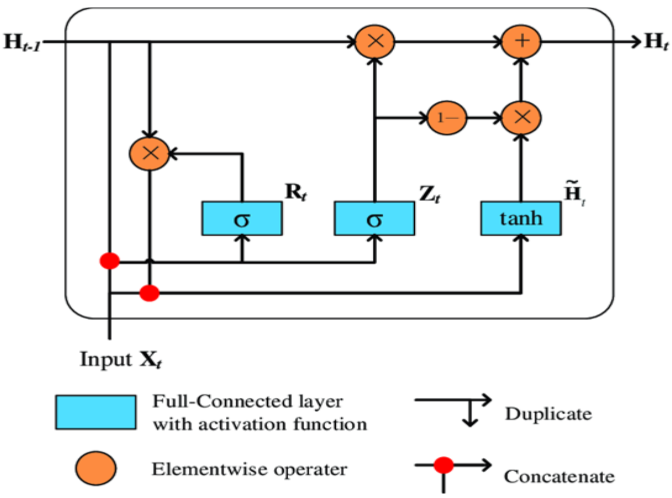


Figure 15:Architecture of GRU

At time $t = 0$, the output vector is $h0 = 0$:

$$z_t = \sigma(w_z x_t + U_z h_{t-1} + b_z) \tag{II-17}$$

$$r_t = \sigma(w_r x_t + U_r h_{t-1} + b_r) \tag{II-18}$$

$$\hat{h}_t = \phi(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \tag{II-19}$$

Where:

- x_t : the input vector at time t .
- ht : the output vector at time t.
- Wz :weight matrices for the input-to-hidden connections.
- Uz : weight matrices for the hidden-to-hidden connections.
- bz : are bias vectors.
- σ : denotes the sigmoid activation function.
- ϕ : denotes the hyperbolic tangent activation function.
- ⨀: denotes the element-wise (Hadamard) product. [28]

II.9.2 Advantages and Disadvantages of GRU

a) Advantages of GRU:

- Faster Training and Efficiency
- Effective for Sequential Tasks
- Less Prone to Gradient Problems[27].

b) Disadvantages of GRU:

- Less Powerful Gating Mechanism
- Potential for Overfitting
- Limited Interpretability[27]

II.9.3 Applications of Gated Recurrent Unit

Here are some applications of GRUs where their ability to handle sequential data shines:

- Natural Language Processing (NLP)
- Machine translation
- Text summarization
- Chatbots
- Sentiment Analysis[27]

II.10 LSTM Vs GRU

Here are the key distinctions between GRU and LSTM:

- GRU incorporates two gates, while LSTM employs three gates.
- Unlike LSTM, GRU lacks an output gate.
- In LSTM, the input and forget gates are linked via an update gate, whereas in GRU, the reset gate directly affects the previous hidden state without this linkage.

- GRU has fewer parameters than LSTM, resulting in lower memory usage and faster execution. However, LSTM tends to perform better on larger datasets where accuracy is critical.
- LSTM is preferable for tasks involving extensive sequences and prioritizing accuracy, whereas GRU is chosen for scenarios requiring less memory consumption and faster processing times.

II.11 Types of Recurrent Neural Networks

There are four types of Recurrent Neural Networks:

1. One to One :

This type of RNN behaves the same as any simple Neural network it is also known as Vanilla Neural Network. In this Neural network, there is only one input and one output. [31]

2. One to Many

In this type of RNN, there is one input and many outputs associated with it. One of the most used examples of this network is Image captioning where given an image we predict a sentence having Multiple words. [31]

3. Many to One

In this type of network, Many inputs are fed to the network at several states of the network generating only one output. This type of network is used in the problems like sentimental analysis. Where we give multiple words as input and predict only the sentiment of the sentence as output. [31]

4.Many to Many

In this type of neural network, there are multiple inputs and multiple outputs corresponding to a problem. One Example of this Problem will be language translation. In language translation, we provide multiple words from one language as input and predict multiple words from the second language as output. [31]

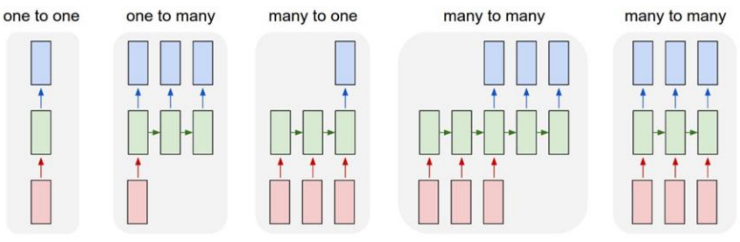


Figure 16:Types of Recurrent Neural Networks

II.12 Types of Neural Networks Activation Functions

The most popular neural networks activation functions listed below:

II.12.1 Binary Step Function

If the node input value is less than 0, it returns 0 as output. Else, it returns 1.

Mathematically, it can be defined as:

$$f(x) = \begin{cases} 0 & for \;\; x < 0 \\ 1 & for \;\; x \geq 0 \end{cases}$$

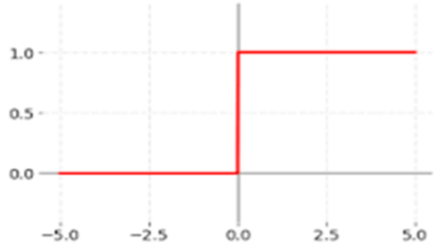


Figure 17: Binary Step Function graph

II.12.2 Linear Activation Function

The linear activation function, also known as "no activation," or "identity function" (multiplied x1.0), is where the activation is proportional to the input.

$$f(x) = x$$

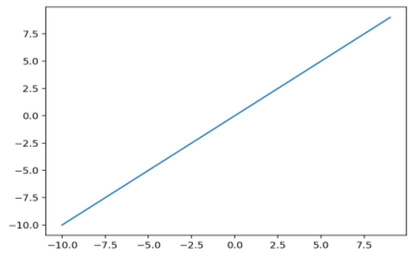


Figure 18: Linear Activation Function graph

II.12.3 Non-Linear Activation Function Sigmoid / Logistic Activation Function

It is commonly used for models where we have to predict the probability as an output. Since probability of anything exists only between the range of 0 and 1, sigmoid is the right choice because of it

sigmoid / logistic

$$f(x) = \frac{1}{1 + e^{-x}}$$

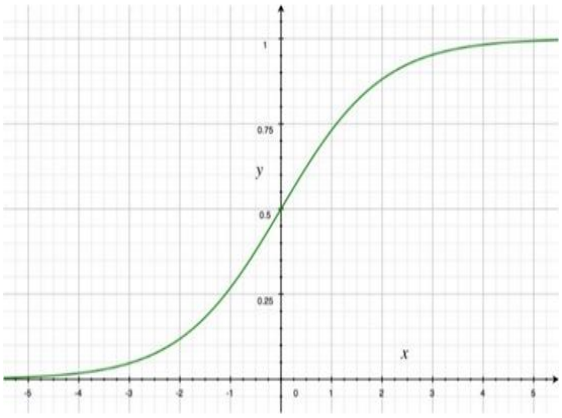


Figure 19:Sigmoid /Logistic Activation Function graph

II.12.4 Non-Linear Activation Function Tanh Function (Hyperbolic Tangent)

Tanh function is very similar to the sigmoid/logistic activation function, and even has the same S-shape with the difference in output range of -1 to 1. In Tanh, the larger the input (more positive), the closer the output value will be to 1.0, whereas the smaller the input (more negative), the closer the output will be to -1.0.

$$Tanh$$

$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$

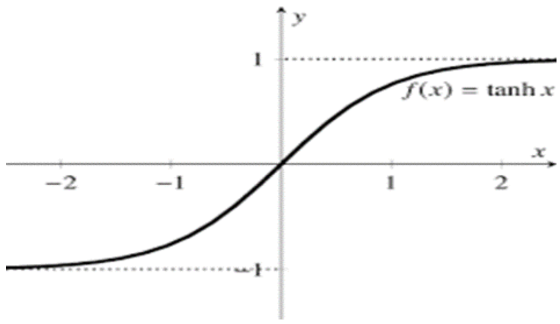


Figure 20 :Activation Function Tanh Function (Hyperbolic Tangent) graph

II.12.5 Activation Function ReLU Function (Rectified Linear Unit

Since only a certain number of neurons are activated, the ReLU function is far more computationally efficient when compared to the sigmoid and tanh functions.ReLU accelerates the convergence of gradient descent towards the global minimum of the loss function due to its linear, non-saturating property

Mathematically it can be represented as:

$$ReLU$$

$$f(x) = max\,(0, x)$$

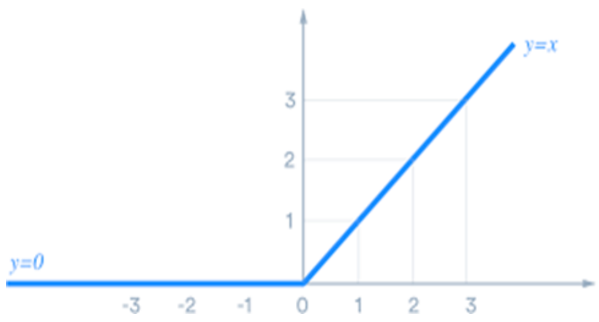


Figure 21:ReLU Activation Function (Rectified Linear Unit) graph

II.12.6 Softmax Function

It calculates the relative probabilities. Similar to the sigmoid/logistic activation function, the SoftMax function returns the probability of each class. It is most commonly used as an activation function for the last layer of the neural network in the case of multi-class classification.

Softmax

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)}$$

V7 Labs

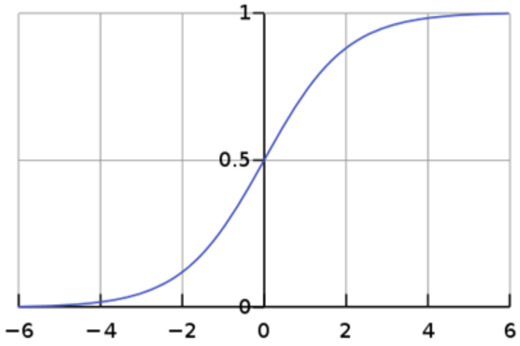


Figure 22:Softmax Activation Function graph

II.13 Conclusion

In this chapter, we explored the basic fundamentals of artificial intelligence, focusing in particular on recurrent neural networks (RNNs). We began by introducing artificial intelligence and how it is evolving as a branch of computing that seeks to create systems traditional to human intelligence, which includes the ability to learn and adapt to changing environments.

Then, we detail recursive neural networks (RNNs), a type of neural network that handles time-series data. We looked at the structure of RNNs, which include recursive units that enable them to retain memory and use past information to improve their predictions of future data. We also discussed the challenges facing RNNs, such as the problem of vanishing gradients, and how to solve them using techniques such as Long Memory Neural Networks (LSTM) and Variational Recursive NeuralNetworks (GRU).

Chapter III: Predicting the proximity of reaching an oil reservoir using deep learning and drilling data

III.1 Introduction

In this chapter, we will delve into the architecture of three recurrent neural network (RNN) variants: SimpleRNN, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). These models, implemented using TensorFlow and Keras, are essential tools for learning and classification tasks. To optimize their performance, we will employ a simple yet effective technique known as dropout.

Firstly, let's discuss the significance of comparing these three models. While all three belong to the family of RNNs and share the ability to model sequential data, they exhibit distinct architectural variances that can impact their performance in different scenarios. By comparing their architectures and evaluating their performance, we aim to identify the model that best suits our lithofacies prediction task, ensuring accurate predictions even at depths as shallow as 5 meters.

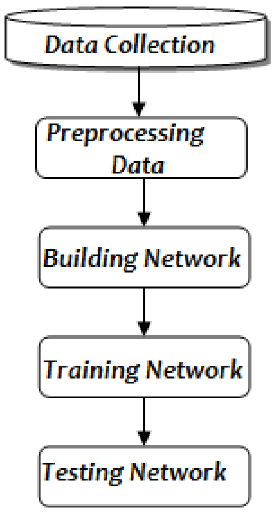


Figure 23 :Basic flow for designing artificial neural network model

III.2 Software and libraries Used in the implementation python

Libraries used

- Pandas

pandas is a Python package that provides fast, flexible, and expressive data structures designed to make working with "relational" or "labeled" data both easy and intuitive.

Figure 24: logo of Pandas

- TensorFlow

TensorFlow is an open-source machine learning framework developed by the Google Brain team. It is designed to facilitate the development and deployment of machine learning models by providing a comprehensive ecosystem for building, training, and deploying neural networks and other machine learning algorithms. [32]

Figure 25:logo of TensorFlow

- Keras

Keras is an open-source, high-level neural networks API written in Python. It is designed to be user-friendly, modular, and extensible, enabling fast experimentation and easy prototyping of deep learning models.[32]

Figure 26: logo of Keras

- Scikit-learn

Scikit-learn is a widely-used open-source machine learning library for Python. It is built on top of NumPy, SciPy, and matplotlib, and provides simple and efficient tools for data mining and data analysis.

Figure 27: logo of Scikit-learn

- Seaborn

Seaborn is a popular Python data visualization library built on top of matplotlib. It provides a high-level interface for creating attractive and informative statistical graphics.

Figure 28: logo of Seaborn

III.3 DATASET

III.3.1 Data collection

Sufficient data is collected constitute a basis necessary for modeling.

This study includes 5 wells oilfield.

BASSIN : AMGUID MESSAOUD

PERIMETRE : TOUGGOURT EST

III.3.2 Description and pre-processing of data

NAN values are a big problem for the deep learning because it will harmfully influence the results, but fortunately our data set is empty of NAN values. We used the pandas library to load the data into a data frame.

III.3.3 Data Selection

The final selection was made after analyzing all wells data without any missing (Nan). Data was selected where the Depth was in the same range in each well section from 3000 m depth until bottom of well.

III.3.4 Data Split

selecting 80% percent of data as training set and remaining 20% as test Dataset stored in xlsx format

(excel) contains 5 wireline log measurements, in meter interval, with corresponding facies label.

Table 1: Data Split

Well Name	Total Feautures	Input data		Output data		Labels
		Train	Test	Train	Test	
RAA-2	1201	960	241	960	241	7
N_RAA	1201	960	241	960	241	7
E_RAA	1232	985	247	985	247	9
W_RAA	1149	918	231	918	231	9
BMTG-1	1201	960	241	960	241	9
Total 5 Wells	5984	4787	1197	4787	1197	13

III.3.5 Input data

Drilling parameters used as input are:

1- <u>ROP (Rate of Penetration):</u>

In the drilling industry, the rate of penetration (ROP), is the speed at which a drill bit breaks the rock under it to deepen the borehole.

This speed is usually reported in units of **feet per hour** (ft/h) or **meters per hour(m/h)**. [33]

2- <u>RPM (Rotation per minute):</u>

Determines the speed of the drill bit, which can affect the quality and speed of the drilling process.

A lower RPM may not drill as quickly, but it can be gentler on the drill bits and may cause less wear and tear. It ultimately depends on the specific task you will be using the drill for and what type of rock will be drilling into This speed is usually reported in units of **rpm.** [33]

3- <u>WOB (Weight On Bit):</u>

Weight on bit (WOB) is the amount of downward force exerted by a drill bit during drilling operations. Weight on bit is generally measured in thousands of pounds, or tons. [33]

4- <u>SPP (Stand pipe Pressure):</u>

The standpipe pressure monitoring is the most effective conventional tool for detecting downhole problems such as drill string washout.

By the continuous monitoring of the SPP, the driller can detect a sudden change in SPP, which is in the most cases an indication of a downhole problem. SPP is Measured on **psi.** [33]

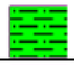
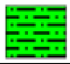
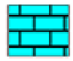
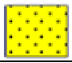
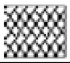
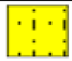
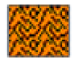
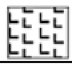
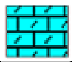
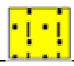
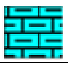
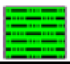
5- Pump flow rate:

Volume of liquid or mud which moves through a pipe in one unit of time, the volumetric flow rate of a fluid in a drilling rig return line. Measured on **l/mn.** [33]

III.3.6 Output data

The facies (classes of rocks) used as output are shown in the table below:

Table 2 :Labels codification (facies)

Facie Name	Label Code	Labels
Argile	0	
Argile silteuse	1	
Dolomie	2	
Calcaire	3	
Silt	4	
Anhydrite	5	
Gres fin	6	
Roches eruptives	7	
Sel	8	
Calcaire dolomitiq	9	
Gres moyen	10	
Calcaire argileux	11	
Argile Sableuse	12	

III.4 Correlation matrix visualization for Drilling data Well BMTG-1

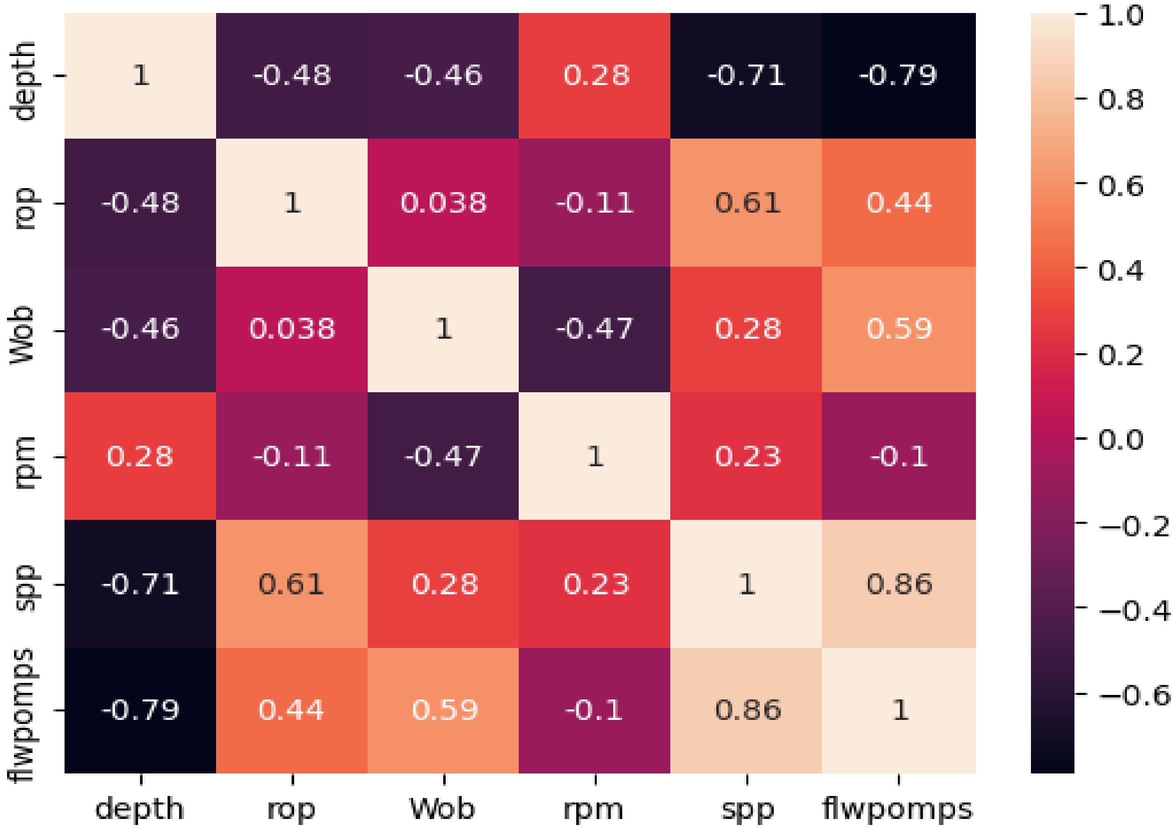


Figure 29:Correlation matrix for drilling data, Well BMTG-1

The correlation matrix is used to evaluate the dependence between several variables at the same time. The result is a table containing the correlation coefficients between each variable and the others.

A correlation matrix helps us to determine the relationship and also the strength and the direction between the variables.

For our case above Correlation analysis result shows that there is a significant relationship between spp and flwpopms (86%)

Negative correlation is a relationship between two variables in which one variable increases as the other decreases, and vice versa.

III.5 Visualization of “Facies” dataset using boxplot

The boxplot with a box plot visualization provides a detailed summary of the distribution of various "Facies" within the dataset. The primary function of this visualization is to offer a clear representation of the statistical distribution, including the median, quartiles, and potential outliers for each category of facies.

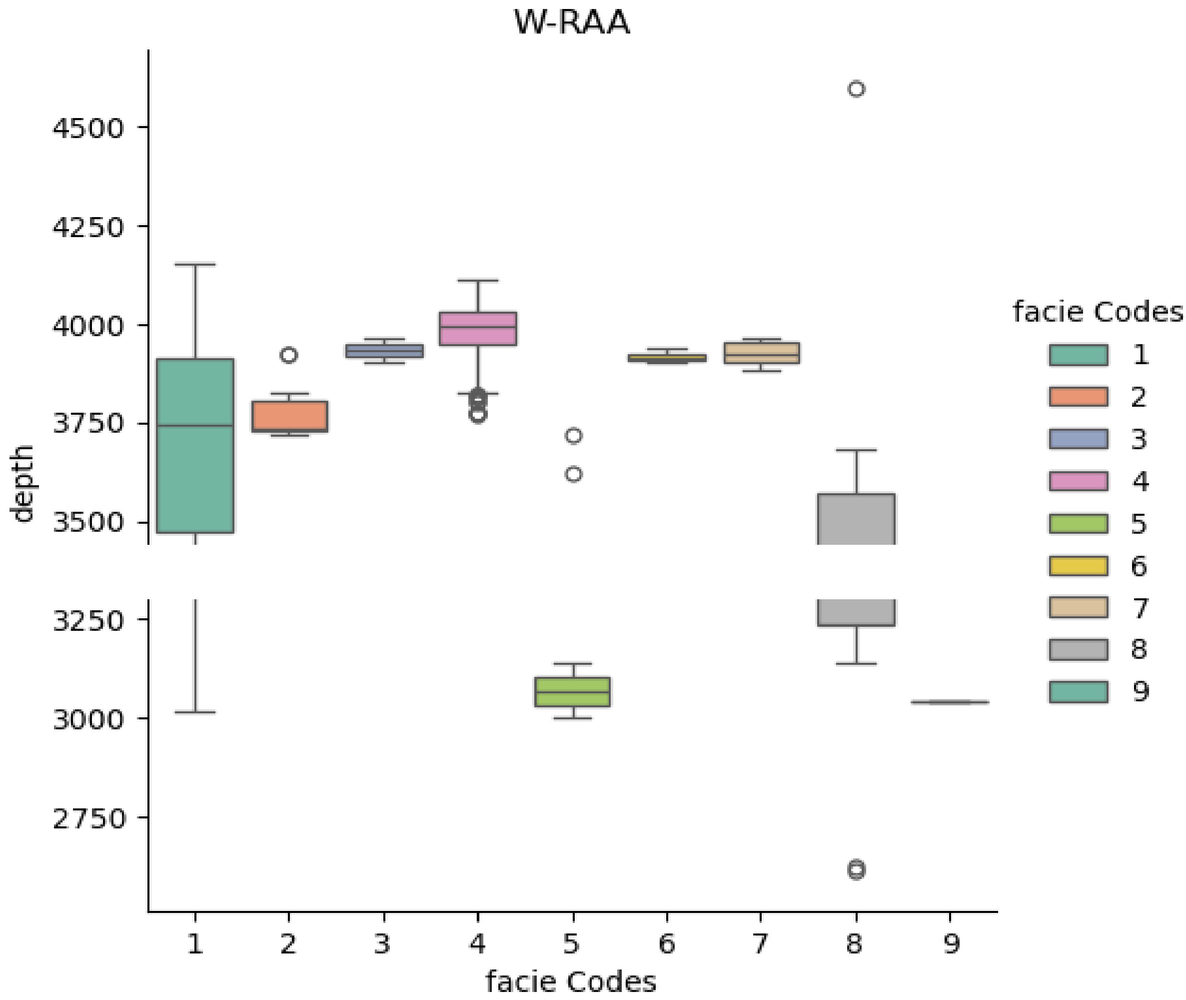


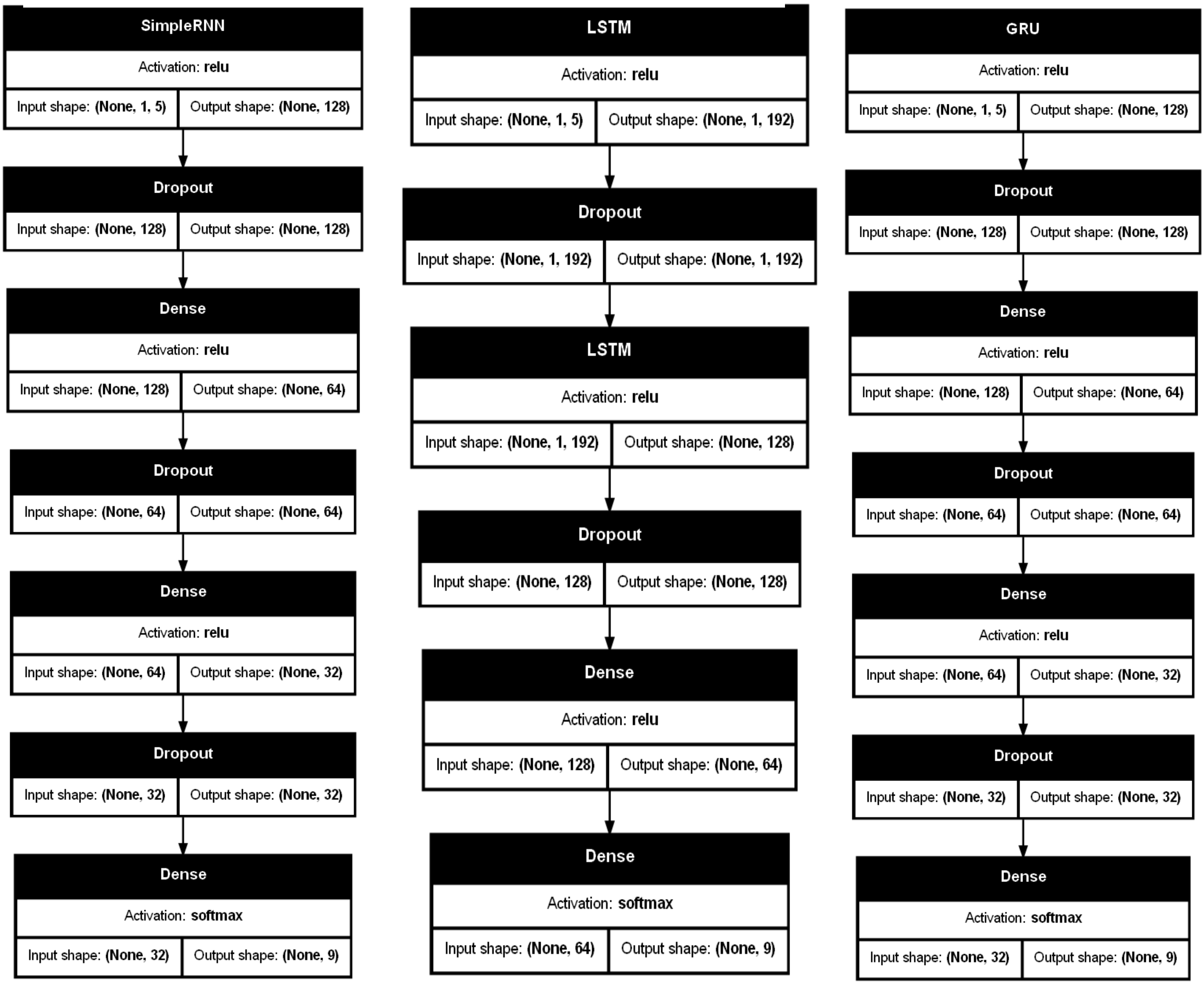
Figure 30: Visualization of “Facies” using boxplot for W-RAA

III.6 Architecture for developed models

For each well, we developed three models: Model 1 (RNN), Model 2 (LSTM), and Model 3 (GRU). Below, we outline the architecture of these models:

<u>Architecture Models for W RAA-WELL</u>

Table 3:Architecture models RNN,LSTM,GRU for W-RAA



III.7 Results obtained and discussion

Here, it presents a comprehensive analysis of the developed models. This includes accuracy and error rates, as well as graphical representations such as confusion matrices and classification reports, providing a detailed overview of the performance of each model.

III.7.1 Performance of developed models, by well

Table 4:Performance of developed models ,Well BMGT-1

Well Name	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
BMGT-1	RNN	72.91	0.62	64.58	0.98
	LSTM	71.65	0.63	67.50	0.97
	GRU	71.65	0.63	65.00	0.97

Table 5:Performance of developed models ,Well RAA-2

Well Name	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
RAA-2	RNN	78.90	0.50	74.80	0.67
	LSTM	78.49	0.53	70.33	0.73
	GRU	78.49	0.53	71.95	0.73

Table 6:Performance of developed models, Well NRAA

Well Name	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
NRAA	RNN	74.90	0.57	70.90	0.78
	LSTM	80.97	0.47	70.49	0.91
	GRU	79.12	0.45	70.08	0.83

Table 7:Performance of developed models, Well WRAA

Well Name	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
WRAA	RNN	80.35	0.51	74.35	0.69
	LSTM	82.10	0.41	76.09	0.74
	GRU	79.59	0.49	73.48	0.70

Table 8:Performance of developed models, Well ERAA

Well Name	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
ERAA	RNN	75.42	0.57	62.08	0.88
	LSTM	78.56	0.48	61.67	0.99
	GRU	77.30	0.48	65.00	0.99

Table 9:Performance of developed models, 4 Wells

Well Name	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
4WELLS	RNN	68.39	0.74	69.66	0.78
	LSTM	75.29	0.57	72.17	0.72
	GRU	75.29	0.58	72.51	0.73

In this study, we evaluated the performance of three Recurrent Neural Network (RNN) models—Simple RNN, LSTM, and GRU—using drilling data from five different wells. The models were evaluated on both training and test data based on their accuracy and loss metrics.

The Simple RNN model showed moderate performance overall, with training accuracy ranging from 68.39% to 80.35% and test accuracy from 62.08% to 74.80%.

The higher loss values compared to other models suggest potential overfitting issues in some cases.

The LSTM model exhibited excellent training performance with accuracy ranging from 71.65% to 82.10%.

Test accuracy was generally good, ranging from 61.67% to 76.09%. However, some gaps between training and test accuracy indicate potential overfitting.

The GRU model's performance was very similar to that of the LSTM, with training accuracy between 71.65% and 79.59% and test accuracy between 65.00% and 73.48%.

Loss values were also comparable to those of the LSTM, indicating that both models perform similarly in many cases.

All models demonstrated good training accuracy, but there was variability in test accuracy, highlighting the need for model refinement to reduce the training-test performance gap.

III.7.2 Graphic curve of accuracy and Loss

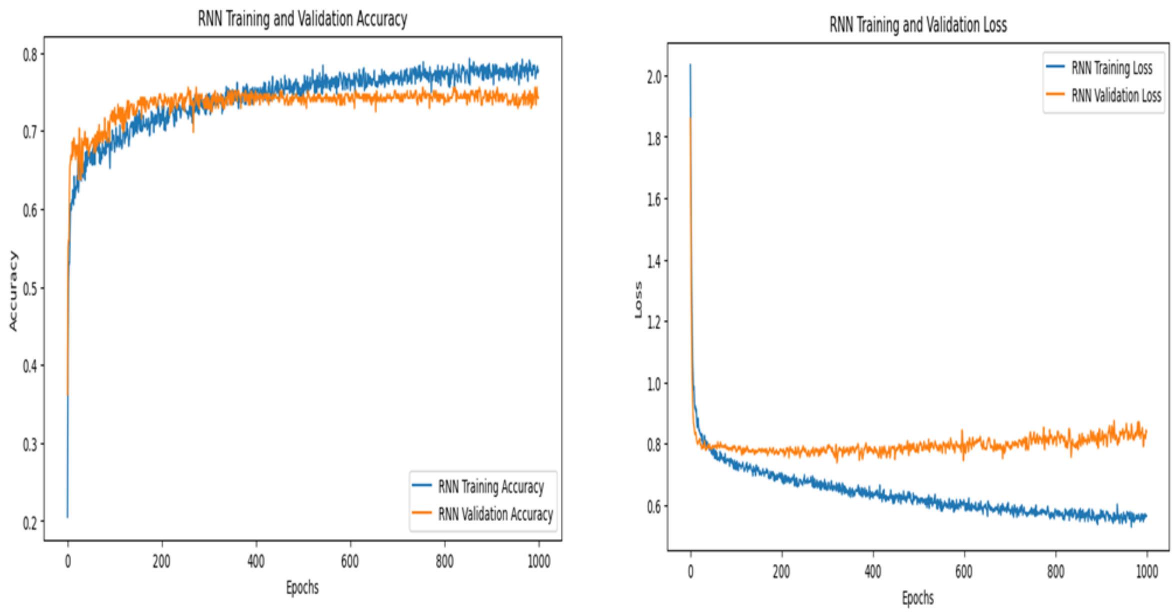


Figure 31: *accuracy and loss curves for the RNN,WRAA well*

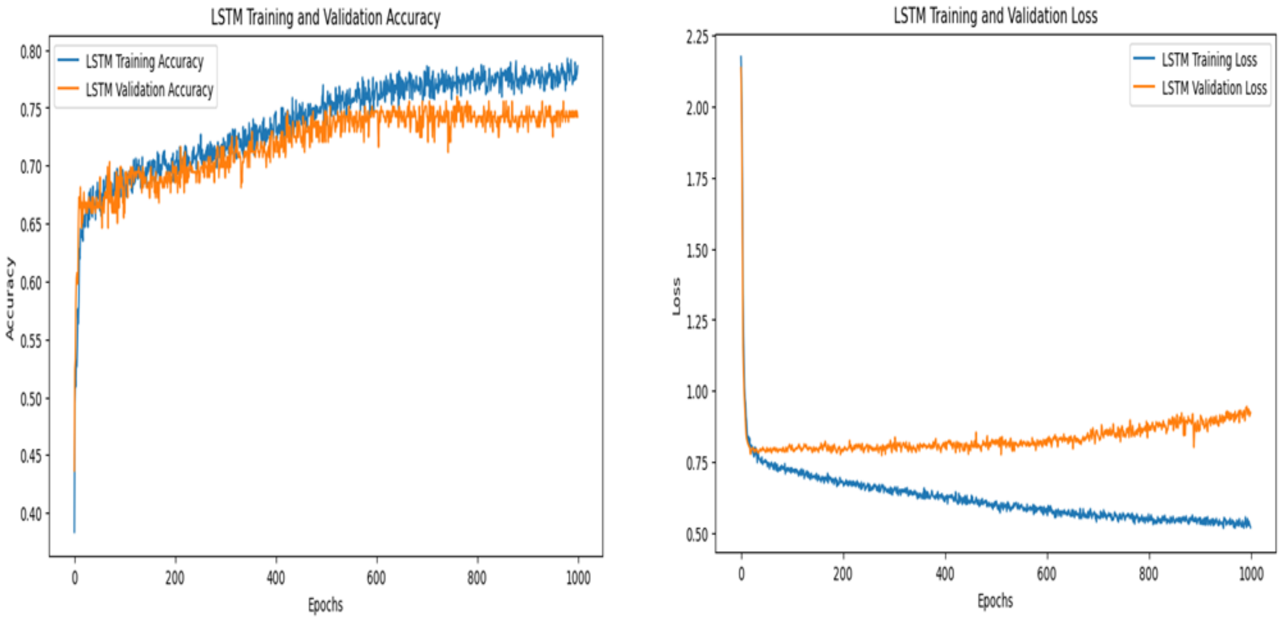


Figure 32: accuracy and loss curves for the LSTM,WRAA well

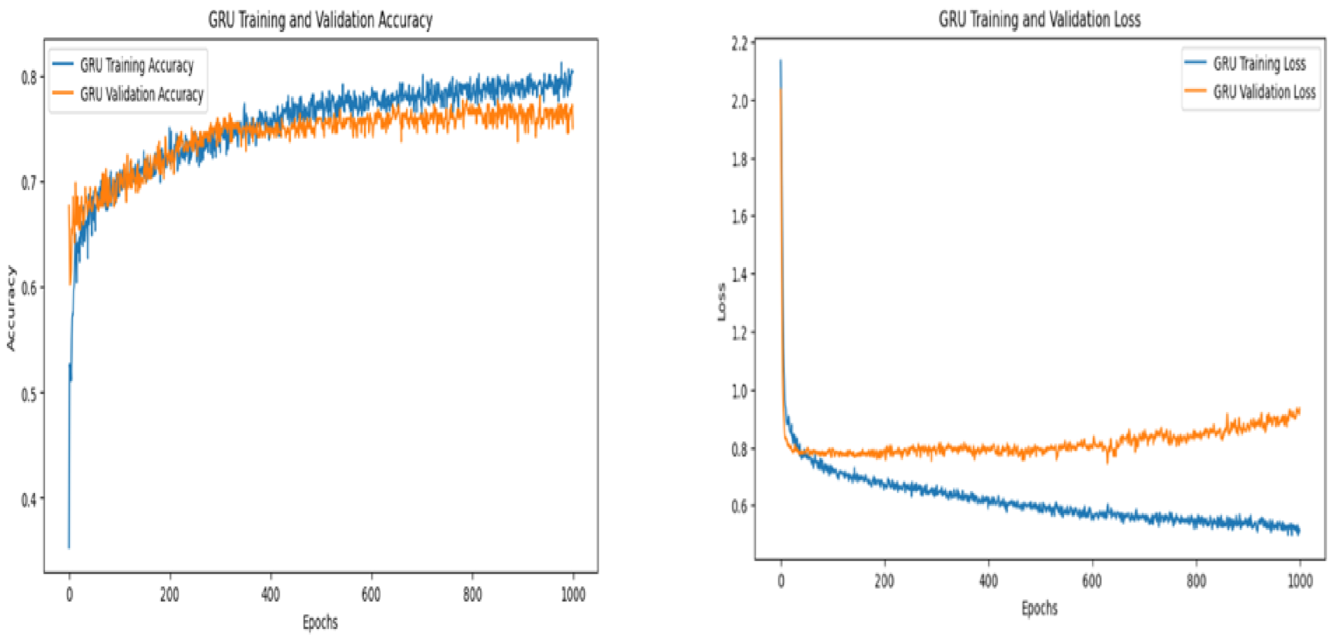


Figure 33:accuracy and loss curves for the GRU,WRAA well

Comparing RNN, LSTM, and GRU models across five different wells: BMGT-1, RAA-2, NRAA, WRAA, and ERAA, the results showed variations in performance in terms of accuracy and loss for each model during training and testing stages.

In well WRAA, the LSTM model was the best in both training and testing with accuracies of 82.10% and 76.09% respectively, indicating strong and consistent performance compared to the other models.

III.8 Confusion matrix, predicted vs trues facies

In this section, we present and analyze confusion matrices for Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) neural networks, and Gated Recurrent Unit (GRU) across five different wells: BMGT-1, RAA-2, N-RAA, W-RAA, and ERAA, in addition to analyzing confusion matrixes for model results across all wells together.

In general, confusion matrices provide a powerful tool for analyzing model performance by highlighting points where the models perform well and those that need improvement. Through this analysis, we can make informed decisions to enhance the models and increase their accuracy and effectiveness in lithofacies classification using well logging data.

Confusion matrix visualization for W-RAA- WELL

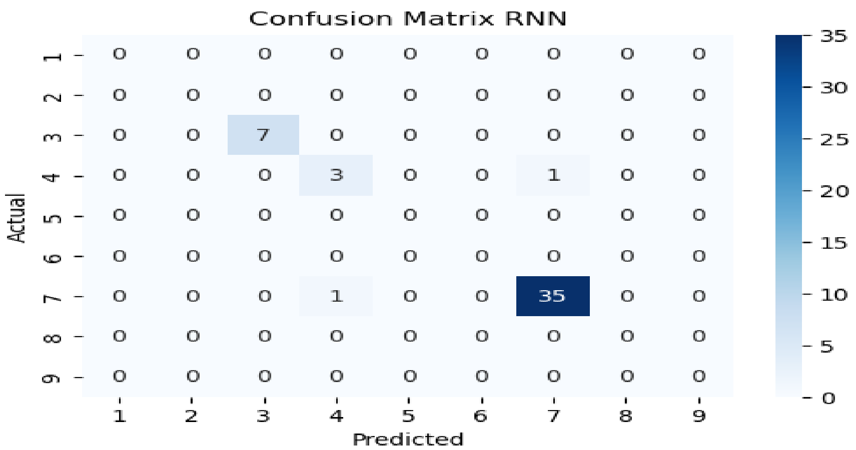


Figure 32 : Confusion matrix for RNN facies class predictions for W-RAA- WELL

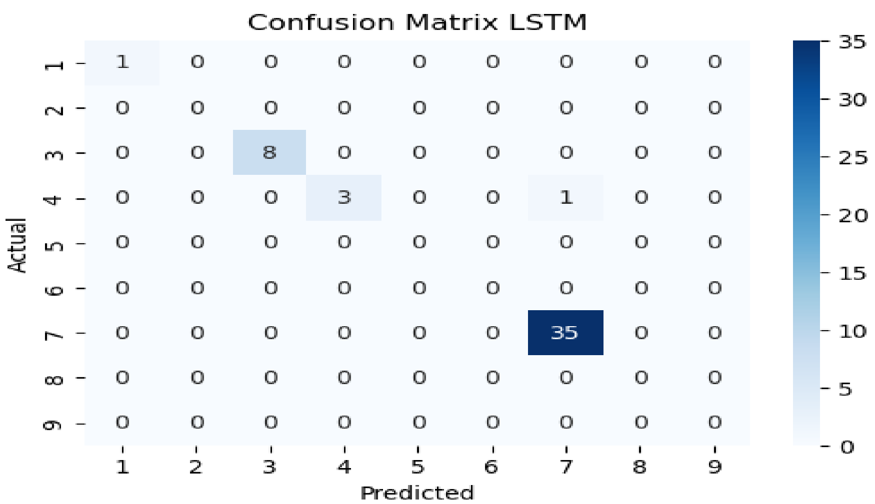


Figure 33 : Confusion matrix for LSTM facies class predictions for W-RAA- WELL

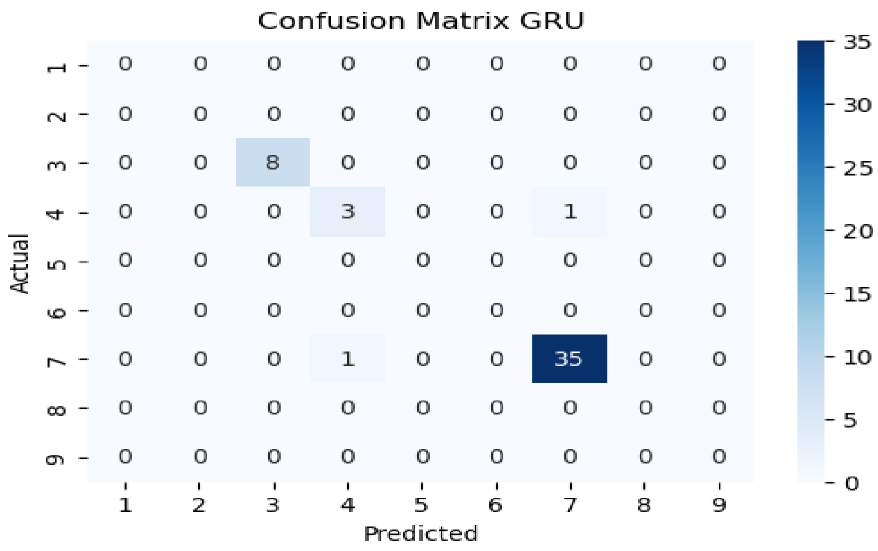


Figure 34: Confusion matrix for GRU facies class predictions for W-RAA- WELL

Confusion matrixes based on the actual lithology and classification results by RNN, LSTM and GRU are shown in above Figures (Fig.32, Fig.33 and Fig.34)

Along the diagonal space of the matrixes, the lithology is classified correctly, LSTM and GRU classified the largest number of data samples

We have a total of 46 labels classified by LSTM and the same for GRU, but a total of 45 classified by RNN

It is observed that 1 is no classified by LSTM and the same for GRU, but 2 not classified by RNN

III.9 Classification Reports

1. **Model Performance Metrics**:
   - **Precision**: The ability of the classifier to not label a negative sample as positive.
   - **Recall**: The ability of the classifier to find all the positive samples.
   - **F1-score**: A weighted harmonic mean of precision and recall, providing a single metric that balances both.
   - **Support**: The number of actual occurrences of the class in the dataset.
2. **Macro Average**: Calculated as the average of precision, recall, and F1-score for each class, treating all classes equally.
3. **Weighted Average**: Takes into account the support (number of true instances for each label), balancing the metric according to the class distribution.

Classification Report for W-RAA well

Classification Report RNN:

	precision	recall	f1-score	support
0	0.79	0.85	0.81	39
1	0.00	0.00	0.00	1
3	1.00	0.58	0.74	12
4	0.50	0.75	0.60	4
7	0.88	0.90	0.89	39
accuracy			0.82	95
macro avg	0.63	0.62	0.61	95
weighted avg	0.83	0.82	0.82	95

Classification Report LSTM:

	precision	recall	f1-score	support	
0	0.81	0.87	0.84	39	
1	1.00	1.00	1.00	1	
3	1.00	0.67	0.80	12	
4	0.75	0.75	0.75	4	
7	0.88	0.90	0.89	39	
accuracy			0.85	95	Best score
macro avg	0.89	0.84	0.86	95	
weighted avg	0.86	0.85	0.85	95	

Classification Report GRU:

	precision	recall	f1-score	support
0	0.80	0.85	0.83	39
1	0.00	0.00	0.00	1
3	1.00	0.67	0.80	12
4	0.50	0.75	0.60	4
7	0.88	0.90	0.89	39
accuracy			0.83	95
macro avg	0.64	0.63	0.62	95
weighted avg	0.84	0.83	0.83	95

Based on the results of classification of wells BMTG-1, RAA-2, N-RAA, W-RAA, E-RAA, it appears that the performance of the models varies depending on the type and architecture. The LSTM models appear to show overall superior performance compared to RNN and GRU in most cases, with high accuracy and f1 for many categories.

Due to the results showing that the W-RAA well provided excellent performance using the LSTM model, as the accuracy reached **85%,** which indicates its ability to accurately predict rock types using geological data and deep learning techniques. This result can be considered positive because it shows the efficiency of the LSTM model in predicting the rock types in this well better than other models. This outstanding performance could have a positive impact on drilling and oil exploration operations in the region in question, as it could direct new research to improve the use of LSTM models in geological and engineering risk predictions in upcoming drilling projects.

.

III.10 Facies Prediction for blind well RAA-1

The best Trained LSTM model contributes to predict Facies types for the well RAA-1 as blin well, facies predicted between 3000 to 3250 m depth. Types in the well RAA-1 as blin well, facies predected between 3000 to 3250 m depth.

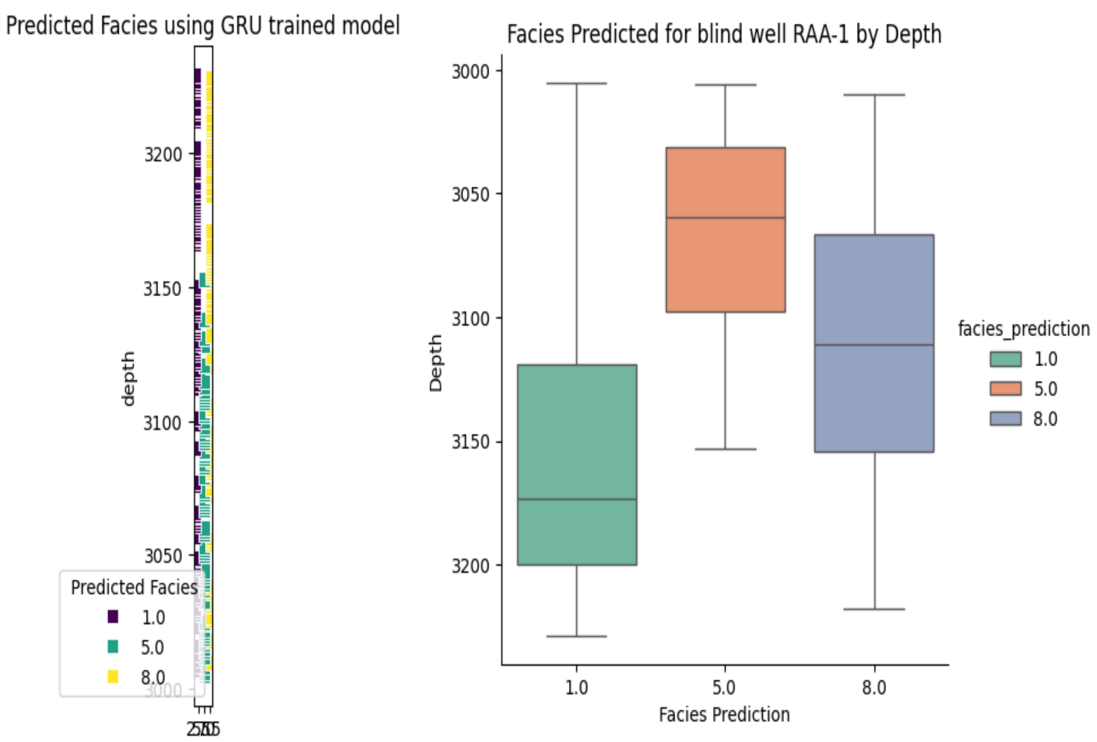
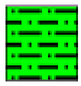
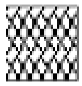
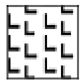


Figure 35 :Log for predicted facies RAA-1 well

Results of the blind well lithology prediction

As shown in Figure 35 and table 10 bellow, it was observed that the LSTM model exhibited lithological classification of the blind well RAA-1 between 3000 to 3250 m depth.

Table 10 : Predicted Facies

Facie Name	Label Code	Labels
Argile silteuse	1	
Anhydrite	5	
Sel	8	

III.11 conclusion

In this chapter, we developed three different recursive neural network models (RNNs), SimpleRNN, LSTM, and GRU, and test their performance in predicting lithostratigraphy using drilling data and using TensorFlow and Keras for implementation and evaluate the models on the test data and visualize the training loss, training _accuracy, test_loss and test_accuracy.The results of each model are presented in detail

General conclusion

In our study, a dataset of oil wells from the Hassi Messaoud basin located in Touggourt, east Algeria. These models were developed to predict geological facies during the drilling phase with a depth offset of 5 meters.

Based on classification reports, the results showed that the LSTM model achieved an accuracy of 85%. This indicates that the LSTM model has demonstrated superior performance in predicting lithofacies using drilling data. The high accuracy suggests that the LSTM model effectively captures the temporal dependencies within the data, enabling it to make accurate predictions.

The success of the LSTM model underscores the effectiveness of recurrent neural networks, particularly in tasks involving sequential data like lithofacies prediction. By leveraging the LSTM architecture's ability to retain information over time, the model can effectively learn and represent complex patterns present in drilling data, leading to more accurate predictions.

Furthermore, the high accuracy achieved by the LSTM model has significant implications for drilling operations. Accurate lithofacies prediction is crucial for optimizing drilling processes, reducing costs, and minimizing risks associated with drilling operations. The ability to predict lithofacies with high accuracy using deep learning models like LSTM can greatly enhance the efficiency and effectiveness of oil exploration and production activities.

In summary, the impressive performance of the LSTM model, with an accuracy of 85%, highlights its potential as a valuable tool for lithofacies prediction in drilling operations. This underscores the importance of leveraging advanced machine learning techniques, such as LSTM, to improve decision-making and optimize processes in the oil and gas industry. In summary, the impressive performance of the LSTM model, with an accuracy of 85%, highlights its potential as a valuable tool for lithofacies prediction in drilling operations. This underscores the importance of leveraging advanced machine learning techniques, such as LSTM, to improve decision-making and optimize processes in the oil and gas industry.

References

[1] A. Y. Dandekar, Petroleum reservoir rock and fluid properties. CRC press, 2006.

[2] Z. Guan, T. Chen, and H. Liao, Theory and Technology of Drilling Engineering. Singapore: Springer Singapore, 2021.

[3] N. Alyafei, Fundamentals of reservoir rock properties. QScience. com, 2021.

[4] E. H. Austin, Drilling engineering handbook. Springer Science & Business Media, 2012

[5] R. Pusch, Rock mechanics on a geological base. Elsevier, 1995.

[6] M. E. Celebi and K. Aydin, Eds., Unsupervised Learning Algorithms. Cham: Springer International Publishing, 2016.

[7] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.

[8] A. Gudmundsson, Rock fractures in geological processes. Cambridge University Press, 2011.

[9] M. Tolouei-Rad, Drilling Technology. BoD–Books on Demand, 2021.

[10] A. S. S. A. S. S. Committee, Drilling fluids processing handbook. Elsevier, 2011.

[11] I. Bowman, ... Well-drilling Methods, vol. 257. US Government Printing Office, 1911.

[12] C. Szepesvári, Algorithms for reinforcement learning. Springer nature, 2022.

[13] Y. Chauvin and D. E. Rumelhart, Backpropagation: theory, architectures, and applications. Psychology press, 2013.

[14] T. Inglis, Directional drilling, vol. 2. Springer Science & Business Media, 2013.

[15] E. H. Austin, Drilling engineering handbook. Springer Science & Business Media, 2012.

[16] M. W. Berry, A. Mohamed, and B. W. Yap, Eds., Supervised and Unsupervised Learning for Data Science. in Unsupervised and Semi-Supervised Learning. Cham: Springer International Publishing, 2020.

[17] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, Advances in Deep Learning, vol. 57. in Studies in Big Data, vol. 57. Singapore: Springer Singapore, 2020.

[18] S. A. Klugman, H. H. Panjer, and G. E. Willmot, Loss Models: From Data to Decisions. John Wiley & Sons, 2012.

[19] K. Suzuki, Artificial neural networks: Architectures and applications. BoD–Books on Demand, 2013.

[20] M. A. Wani, B. Raj, F. Luo, and D. Dou, Eds., Deep Learning Applications, Volume 3, vol. 1395. in Advances in Intelligent Systems and Computing, vol. 1395. Singapore: Springer Singapore, 2022.

[21] F. M. Salem, Recurrent Neural Networks: From Simple to Gated Architectures. Cham: Springer International Publishing, 2022.

[22] H. Sak, A. Senior, and F. Beaufays, “Long short-term memory recurrent neural network architectures for large scale acoustic modeling,” in Interspeech 2014, ISCA, Sep. 2014.

[23] H. Sak, A. Senior, and F. Beaufays, “Long short-term memory recurrent neural network architectures for large scale acoustic modeling,” in Interspeech 2014.

[24] J. F. Kolen and S. C. Kremer, A Field Guide to Dynamical Recurrent Networks. John Wiley & Sons, 2001.

[25] B. Yegnanarayana, Artificial neural networks. PHI Learning Pvt. Ltd., 2009.

[26] K. Huang, A. Hussain, Q.-F. Wang, and R. Zhang, Eds., Deep Learning: Fundamentals, Theory and Applications, vol. 2. in Cognitive Computation Trends, vol. 2. Cham: Springer International Publishing, 2019.

[27] L. Deng and Y. Liu, Deep learning in natural language processing. Springer, 2018

[28] A. Shewalkar, D. Nyavanandi, and S. A. Ludwig, “Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU,” Journal of Artificial Intelligence and Soft Computing Research.

[29] J. Brownlee, Long Short-Term Memory Networks With Python: Develop Sequence Prediction Models with Deep Learning. Machine Learning Mastery, 2017.

.[30] K. K. Hiran, R. K. Jain, K. Lakhwani, and R. Doshi, Machine Learning: Master Supervised and Unsupervised Learning Algorithms with Real Examples (English Edition). BPB Publications, 2021

[31] Gulli and S. Pal, Deep learning with Keras. Packt Publishing Ltd, 2017.

[32] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, and V. Zocca, Python Deep Learning: Exploring deep learning techniques and neural network architectures with PyTorch, Keras, and TensorFlow. Packt Publishing Ltd, 2019.

[33] Z. Guan, T. Chen, and H. Liao, Theory and Technology of Drilling Engineering. Singapore: Springer Singapore, 2021.