

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

N° Série : /2023

UNIVERSITE KASDI MERBAH OUARGLA

Faculté des hydrocarbures, des énergies renouvelables, science de la terre et de l'univers

Département : Forage et mécanique des chantiers pétroliers



Mémoire

MASTER PROFESIONNEL

Domaine : Sciences et techniques

Filière : Hydrocarbures

Spécialité : FORAGE

Présenté par :

GHERBI Brahim, GOHMES Ahmed, MESSAOUD Khallil el wazir

-Thème-

Elaboration d'un modèle prévisionnel de la vitesse d'avancement (ROP) pour optimiser les outils de forage lors de la phase 6 inch (6") du champ Hassi Messaoud

Soutenue publiquement Le : 12/06/2023

Devant le jury :

President :	Mr. MERABTI Hocine	M.C.A	Univ. OUARGLA
Examineur :	Mr. ABIDI SAAD Lfakour	M.A.A	Univ. OUARGLA
Encadreur :	Mme. HAFSI Fadila	M.A.A	Univ. OUARGLA

Année Universitaire : 2022/2023

REMERCIEMENTS

Toute la gratitude et le merci à Dieu notre créateur qui nous a donné la force pour effectuer et achever ce travail.

Ainsi nos parents qui nous aident.

Nous tenons à remercier en premier lieu et très chaleureusement notre encadreur **Madame HAFSI.F** pour avoir accepté de diriger notre travail, pour ses précieux conseils, pour son esprit d'ouverture et sa disponibilité. Grâce à lui, notre travail s'est déroulé.

Je remercie sincèrement et profondément Mr : **DOUCHANE Sadek** pour sa disponibilité, son engagement et ses conseils judicieux qui ont apportés une valeur supplémentaire à ce projet.

Nous remercions toutes personnes qui nous ont aidés de près ou de loin à la finalisation de ce travail, nous tenons à leur exprimer notre vive gratitude.

Enfin nos remerciements s'adressent aux membres de jury qui nous feront l'honneur de juger notre travail.

DEDICACES



J'ai le grand honneur de dédier ce Modeste travail

A ceux qui ont fait de moi un homme "Ma Mère" et "Mon Père" qui m'ont aidée et encouragée beaucoup durant toutes mes études.

A ma chère épouse, qui a embelli mes jours et les a couronnés de beauté ; je vous dédie ce travail pour vous exprimer ma gratitude pour avoir toujours cru en mon succès.

A mes chère frères et sœurs.

A tous mes amis et mes proches.

A tous ceux qui m'ont aidé de loin ou de près durant les moments difficiles.

AHMED

J'ai toujours pensé faire où offrir quelque chose à mes parents en signe de reconnaissance pour tout ce qu'ils ont consenti, des efforts rien que pour me voir réussir, Et voilà, l'occasion est venue.

A ceux qui m'ont donné la vie, symbole de beauté, et de fierté, de sagesse et de patience.

A ceux qui sont la source de mon inspiration et de mon courage, à qui je dois de l'amour et la reconnaissance.

A mes parents.

A mes frères et mes sœurs, je vous réserve toujours une place dans mon cœur et mes pensées.

A toute ma famille.

A tous mes amis.

BRAHIM

Je dédie ce modeste travail :

Particulièrement à celle qui aurait aimé me voir arriver là, ma très chère mère. A mon très cher père, ma fierté, mon exemple dans la vie qui m'a toujours montré le bon chemin.

Sans oublier tous les membres de ma famille, sœurs et frères, grands et petits, ainsi que mes amis et tous ceux qui me connaissent et m'aident.

KHALLIL EL WAZIR

Résumé :

La phase 6" à Hassi Messaoud est une phase qui peut être très coûteuse en termes de temps et d'argent. Cela est principalement dû à la nature même du réservoir qui est très dure et abrasive. Cependant, le choix des outils de forage pour cette phase peut être complexe et affecter considérablement la performance globale du forage. Ainsi, il est crucial d'identifier les outils de forage les plus adaptés pour optimiser le processus de forage.

Dans cette perspective, l'utilisation de l'intelligence artificielle peut être une solution pertinente pour améliorer le choix des outils de forage. En effet, en utilisant des techniques de Machine Learning, il est possible d'élaborer un modèle de prévision du ROP (rate of penetration) qui permet de prévoir la vitesse d'avancement de l'outil de forage en tenant compte des caractéristiques géo-mécaniques de la formation traversée, des paramètres mécaniques et hydraulique appliqué à l'outil ainsi que du type d'outil utilisé. Ce modèle peut être utilisé pour identifier les outils de forage les plus adaptés à chaque phase du forage. Ainsi, la problématique de cette étude est de savoir comment élaborer un modèle prévisionnel du ROP en utilisant l'intelligence artificielle pour améliorer le choix des outils de forage dans la phase 6" à Hassi Messaoud. Cette problématique est essentielle pour optimiser le processus de forage et améliorer la performance globale de l'activité pétrolière

Mots Clés : Phases de Forage, l'intelligence artificielle, Machine Learning, modèle de prévision du ROP, vitesse d'avancement.

Abstract:

The Phase 6 inch in Hassi Messaoud can be a phase that is very costly in terms of time and money. This is mainly due to the nature of the reservoir itself, which is very hard and abrasive. However, the choice of drilling Bits for this phase can be complex and significantly affect the overall drilling performance. Therefore, it is crucial to identify the most suitable drilling bit to optimize the drilling process.

In this regard, the use of artificial intelligence can be an appropriate solution to improve the choice of drilling bits. Indeed, by using machine-learning techniques, it is possible to develop a prediction model for the rate of penetration (ROP), which predicts the drilling bits rate of penetration while considering the geo-mechanical characteristics of the formation being drilled, the mechanical and hydraulic parameters applied to the bit, and the type of bit used. This model can be used to identify the most suitable drilling bit for each phase of the drilling process, the problem of this study is how to develop a predictive model of ROP using artificial intelligence to improve the choice of drilling Bits in the 6" phase in Hassi Messaoud. This problem is crucial to optimize the drilling process and improve the overall performance of the oil industry.

Keywords: Drilling phases, artificial intelligence, machine learning, ROP prediction model, Rate of penetration.

المخلص:

الحفر على الطور 6" بحقل حاسي مسعود يمكن أن يكون مكلف جداً من حيث الوقت والمال. ويعود ذلك بشكل رئيسي إلى طبيعة الخزان الذي هو صلب وقاسي جداً. ومع ذلك يمكن أن يكون اختيار أدوات الحفر المناسبة لهذه المرحلة أمراً معقداً ويؤثر بشكل كبير على الأداء العام لعملية الحفر. لذلك، من المهم بشكل حاسم تحديد الأدوات الأنسب لتحسين عملية الحفر.

من هذا المنظور، يمكن أن يكون استخدام الذكاء الاصطناعي حلاً ملائماً لتحسين اختيار أدوات الحفر. فباستخدام تقنيات التعلم الآلي يمكن إنشاء نموذج تنبؤي لمعدل الاختراق (ROP)، الذي يتيح التنبؤ بسرعة تقدم أداة الحفر مع مراعاة الخصائص الجيوميكانيكية للتكوين المخترق والخصائص الميكانيكية والهيدروليكية المطبقة على الأداة بالإضافة إلى نوع الأداة المستخدمة. يمكن استخدام هذا النموذج لتحديد الأدوات الأنسب لكل مرحلة من عمليات الحفر. وعليه فإن مشكلة هذه الدراسة هي كيفية إنشاء نموذج تنبؤي لمعدل الاختراق باستخدام الذكاء الاصطناعي لتحسين اختيار أدوات الحفر في هذا الطور بحاسي مسعود. هذه المشكلة أساسية لتحسين عملية الحفر وتحسين الأداء العام للنشاط النفطي.

الكلمات الدالة: أدوات الحفر، الذكاء الاصطناعي، التعلم الآلي، نموذج تنبؤي لمعدل الاختراق، سرعة تقدم أداة الحفر

TABLE DES MATIÈRES

Remerciement.....	I
Dédicaces.....	II
Résumé	III
Liste des figures et des tableaux.....	IV
Liste des abréviations.....	V
Introduction générale	1

CHAPITRE I : PRESENTATION DE LA ZONE D'ETUDE

Introduction.....	3
I.1- Situation géographique.....	3
I.2- Situation géologique.....	4
I.3- Carte des zones et numérotation des puits	4
I.3.1- Carte des zones de production à HMD	5
I.3.2- Zones et numérotation de puits	6
I.4- Situation de la zone d'étude	6
I.5- Description lithologique de la zone d'étude.....	8

CHAPITRE II : LES OUTILS DE FORAGE

Introduction.....	10
I- Les différents types des outils de forage	10
I.1- Les outils à molettes.....	10
I.1.1- Géométrie des cônes.....	11
I.1.2- Caractéristiques des tricônes	12
I.1.3- Les roulements.....	14
I.1.4- L'hydraulique à l'outil.....	15
I.1.5- Mode de travail d'un outil Tricône	16
I.1.6- Classification IADC des outils tricônes.....	17
I.2- Les outils à éléments de coupe fixes.....	18
I.2.1- Les outils PDC	18
I.2.2- Type de diamant	18
I.2.3- Mode de travail d'un outil PDC.....	20
I.2.4- Désignation d'un outil PDC	21
I.2.5- Classification IADC des outils diamants	26
I.2.6- Les outils TSP	27
I.2.7- Les outils imprégnés	28
Conclusion	29

CHAPITRE III : Machine learning (Apprentissage Automatique)

Introduction.....	30
III.1- Définition de l'intelligence artificielle	30
III.2- Définition du machine learning «Apprentissage automatique »	31
III.3- Importance et applications de la machine learning	32
III.4- Les différents types d'apprentissage en machine learning	33
III.5- Les étapes de base du processus de machine learning	34
III.6- L'apprentissage supervisé (Supervised Machine Learning).....	35
III.6.1- Les type de problèmes d'apprentissage supervisé.....	35
III.6.2- Processus d'apprentissage supervisé.....	36
III.6.3- Types d'algorithmes d'apprentissage supervisé	37
III.6.3- Mesures d'évaluation pour les modèles supervisés	38
Conclusion	39

CHAPITRE IV : Conception d'un Modèle ROP

PARTIE I : Description du dataset (Les données).

Introduction.....	40
IV.1 Description du dataset	40
IV.1.1- Données enregistrées	41
IV.1.1.1- Données de mud logging	41
IV.1.1.2- Données Logging (diagraphie).....	43
IV.1.2- Les données calculées	44
IV.1.2.1- Données hydrauliques	44
IV.1.2.2- Données Géo-mécaniques	47
IV.1.2.3- Mécanique spécifique d'énergie (MSE).....	49

PARTIE II : Traitement des données par l'intelligence artificielle et élaboration d'un modèle.

Introduction.....	50
IV.1- Rapid Miner.....	50
IV.1.1- Prétraitement et préparation des données	50
IV.1.2- Elaboration du modèle	53
IV.1.3- Optimisation de modèle	56
IV.1.4- La validation du modèle	58
IV.1.5- Utilisation du modèle comme outil de sélection des outils	59

Conclusion et Recommandation	61
---	-----------

Références Bibliographiques	63
--	-----------

Liste des figures

CHAPITRE I		
Figure	Titre de figures	Page
Figure.I.1	Situation géographique du champ de Hassi Messaoud	Page 04
Figure.I.2	Situation géologique du champ de Hassi Messaoud	Page 05
Figure.I.3	Carte des zones de production	Page 06
Figure.I.4	Carte de position des puits de la zone d'étude	Page 07
Figure.I.5	Fiche stratigraphique de la zone d'étude	Page 09
CHAPITRE II		
Figure.II.1	Les outils à molettes (tricône)	Page 10
Figure.II.2	Nomenclature des tricônes	Page 11
Figure.II.3	Principe de construction des molettes des tricônes	Page 12
Figure.II.4	Les dents pour terrain tendre	Page 12
Figure.II.5	Les dents pour terrain dure	Page 13
Figure.II.6	Outil renforcé sur le talon	Page 13
Figure.II.7	Roulement standard non étanche	Page 14
Figure.II.8	Roulement à rouleaux étanche	Page 14
Figure.II.9	Roulement étanche à Friction	Page 14
Figure.II.10	Outil conventionnel	Page 15
Figure.II.11	Outil à duse	Page 15
Figure.II.12	Mécanisme d'action d'un outil à molettes sur la roche	Page 16
Figure.II.13	Exemple des codes IADC	Page 17
Figure.II.14	Les outils PDC	Page 18
Figure.II.15	Principe de fonctionnement d'un outil PDC	Page 21
Figure.II.16	Terminologie outil PDC - Steel body	Page 22
Figure.II.17	Terminologie outil PDC - Matrix body	Page 23
Figure.II.18	Type des profils PDC	Page 24
Figure.II.19	Angles caractéristiques d'un PDC	Page 25
Figure.II.20	Longueur utile Ld	Page 26
Figure.II.21	Les outils TSP	Page 27
Figure.II.22	Les outils imprégnés	Page 28

CHAPITRE III

Figure.III.1	Processus ML	Page 30
Figure.III.2	Programme traditionnelle vs ML	Page 31
Figure.III.3	Schéma représentatif d'un problème de Classification	Page 31
Figure.III.4	Les différents types d'apprentissage en machine learning	Page 33
Figure.III.5	Apprentissage supervisé	Page 35

CHAPITRE IV

Figure.IV.1	Propriétés dynamiques de la roche	Page 48
Figure.IV.2	Essais de carottage mécanique upside nord à Hassi Messaoud	Page 48
Figure.IV.3	Importation les données dans logiciel RappidMiner	Page 51
Figure.IV.4	Sélectionner les outputs dans logiciel RappidMiner	Page 52
Figure.IV.5	Sélectionner les inputs dans logiciel RappidMiner	Page 53
Figure.IV.6	Schéma descriptif d'une Cross validation contenant l'algorithme 'decision tree'	Page 54
Figure.IV.7	Schéma descriptif d'implémentation des six algorithmes	Page 56
Figure.IV.8	Schéma descriptif du process d'optimisation de Gradient Boosted Trees	Page 57
Figure.IV.9	Schéma descriptif l' d'évaluation des performances	Page 58
Figure.IV.10	Schéma descriptif d'un test de validation de chaque outil	Page 59
Figure.IV.11	Prédiction de ROP pour différents type d'outil	Page 60

Liste des tableaux

CHAPITRE II

Tableau	Titre de Tableau	Page
Tableau.II.1	Caractéristique du diamant naturel	Page 07

CHAPITRE IV

Tableau.IV.1	Résultats des performances des six algorithmes.	Page 56
Tableau.IV.2	Paramètres utilisé pour l'optimisation du modèle	Page 57
Tableau.IV.3	Résultats du ROP Moyen de chaque Outils	Page 59

Liste des abréviations

- RPM** : Révolution Per Minute (Vitesse de rotation de l'outil de forage).
- ROP** : Rate Of Pénétration (La vitesse d'avancement de l'outil de forage).
- WOB**: Weight on Bit (poids appliqué sur l'outil) exprimée en Tonne.
- BHA**: Bottom Hole Assembly.
- PDC**: Polycristalline diamand compact.
- TSP**: Thermally Stable Polycrystalline.
- TVD**: True vertical depth (mètre).
- F_c** : Force de coupe exprimée en Newtons (N) ou en livres-force (lbf).
- F_n** : Force normale exprimée en Newtons (N) ou en livres-force (lbf).
- F_L** : Force latérale exprimée en Newtons (N) ou en livres-force (lbf).
- W_c** : Angle de coupe (° dégrée).
- W_d** : Angle de dépouille (° dégrée).
- W_s** : Angle latérale (° dégrée).
- L_a** : Longueur utile (mm ou inch).
- PV** : Viscosité plastique exprimée en centi-poise (cp) ou en et en (pascal par seconde)
- YP**: Yeild point exprimée en lbf/in² ou en psi.
- PDM**: Positive Displacement Motors.
- MWD**: Measurement While Drilling.
- IA** : Intelligence artificiel.
- ML** : Machine Learning (Apprentissage automatique)
- SVM** : Machines à vecteurs de support.
- KNN** : K-plus proches voisins.
- BIT** : outil de forage
- MW** : Densité de la boue (lb/gal).
- TFA**: Total flow area (sq.in).
- SPP**: Pression de refoulement (Standpipe Pressure en psi).
- GR**: Gamma Rey.
- VS**: Shear Transit Time (Microseconde μ s ou Nanoseconde ns).
- VP**: Compressional Transit Time (Microseconde μ s ou Nanoseconde ns).
- HHP**: Hydraulique horsepower.
- HSI**: Hydraulic horsepower/square inch.

JV : Jet velocity exprimée en m/s.

JIF : Jet impact force exprimée en lbf.

PD : Pressure Drop (psi).

UCS : Résistance à la compression non confinée exprimée en Pa.

CCS : Résistance à la compression triaxiale exprimée en Pa.

Edyn : Module d'Young dynamique exprimée en Pa.

E_{sta} : Module de Young statique exprimée en Pa.

ν_{dyn} : Poisson dynamique.

MSE : Mécanique spécifique d'énergie exprimée en J/cm³.

RMSE: Root Mean Squared Error.



INTRODUCTION GENERALE

INTRODUCTION GENERALE

Le secteur pétrolier et gazier occupe une place centrale dans l'économie nationale, étant donné qu'il représente la principale source d'énergie. L'extraction de ces ressources souterraines nécessite le forage, qui est considéré comme l'étape la plus coûteuse et la plus délicate du processus d'exploitation.

L'objectif d'un forage est d'atteindre la profondeur ciblée avec un coût minimum, dans les meilleures conditions techniques possibles. Le coût principal associé au forage est lié au temps passé à réaliser les travaux d'avancement, Ce temps est influencé par divers facteurs qui impactent la vitesse de forage, notamment l'outil de forage qui joue un rôle crucial dans l'amélioration de la performance. [1]

La phase 6" à Hassi Messaoud peut s'avérer extrêmement coûteuse en termes de temps et d'argent en raison de la nature du réservoir, qui est connu pour être très dur et abrasif. Cette caractéristique peut rendre le forage plus difficile, cependant, Le choix des outils de forage pour cette phase peut s'avérer complexe et avoir un impact significatif sur la performance globale du forage. Il est donc crucial de sélectionner avec soin les outils de forage les mieux adaptés.

Dans cette perspective, L'utilisation de l'intelligence artificielle peut améliorer le choix des outils de forage pour des conditions spécifiques, telles que celles rencontrées dans la phase 6" à Hassi Messaoud. En effet, en utilisant des techniques de Machine Learning, il est possible d'élaborer un modèle de prévision de ROP (rate of penetration) qui permet de prévoir la vitesse d'avancement de l'outil de forage en tenant compte des caractéristiques géo-mécaniques de la formation traversée, des paramètres mécaniques et hydraulique appliqué à l'outil ainsi que du type d'outil utilisé.

Ce qui va nous permettre d'identifier les outils de forage les plus adaptés à chaque phase du processus de forage.

La problématique de cette étude consiste à développer un modèle prévisionnel du ROP en utilisant l'intelligence artificielle afin d'améliorer la sélection des outils de forage pour la phase 6" à Hassi Messaoud, Cette étude cruciale aura un rôle déterminant dans l'amélioration continue du forage.

Ce travail est constitué de deux parties principales, une partie théorique et une étude pratique. La partie théorique a porté essentiellement sur trois principaux chapitres :

Le premier chapitre comporte quelques pages sur la présentation de la zone d'étude (champ de Hassi Messaoud).

Le deuxième chapitre donne des généralités sur les outils de forage et les choix de ces derniers.

Le troisième chapitre on a essayé d'introduire la notion l'intelligence artificiel, Machine learning et ces importances dans l'industrie pétrolières, ainsi Les différents types d'apprentissage en machine learning.

La deuxième partie comporte **le quatrième chapitre** du mémoire pour élaborer d'un modèle prévisionnel de la vitesse d'avancement (ROP) à l'aide de l'intelligence artificielle pour optimiser la sélection des outils de forage de la phase 6".

CHAPITRE

I

**PRESENTATION DE
LA ZONE D'ETUDE**

Introduction :

Le gisement de Hassi Messaoud est l'un des gisements les plus vastes et les plus complexes au monde.

Durant l'histoire géologique, il a subi une évolution tectonique intense, pendant l'enfouissement jusqu'à ce qu'il prenne sa forme actuelle.

Ces incidents sont capables d'améliorer les paramètres pétro-physiques (perméabilité, porosité etc...) comme ils peuvent les réduire (réduction de la porosité, la cimentation des grains, la création de matrices de petits grains, la création des barrières imperméables etc...).

Le gisement de Hassi Messaoud continue de fournir à l'Algérie cette ressource naturelle, le pétrole brut, après plus de 50 ans.

Des investissements importants ont été réalisés et d'autres seront réalisés à l'avenir pour extraire le maximum de pétrole pour une meilleure récupération ultime.

I.1- Situation géographique :

Le champ de Hassi Messaoud est un ensemble des structures formant la partie Nord de la province Triasique, et se situe à environ 850 km au Sud-Est d'Alger, à 280 km au Sud-Est du gisement de gaz à condensat de Hassi Rmel et à 350 km à l'Ouest de la frontière Tunisienne (Figure.I.1). Il s'étend sur une superficie de 2500 km², il a pour coordonnées Lambert :

X = [790.000 - 840.000] Est ;

Y = [110.000 - 150.000] Nord ;

Il est encadré par les latitudes 31°.30' et 32°.00' et les longitudes 5°.40' et 6°.20'.

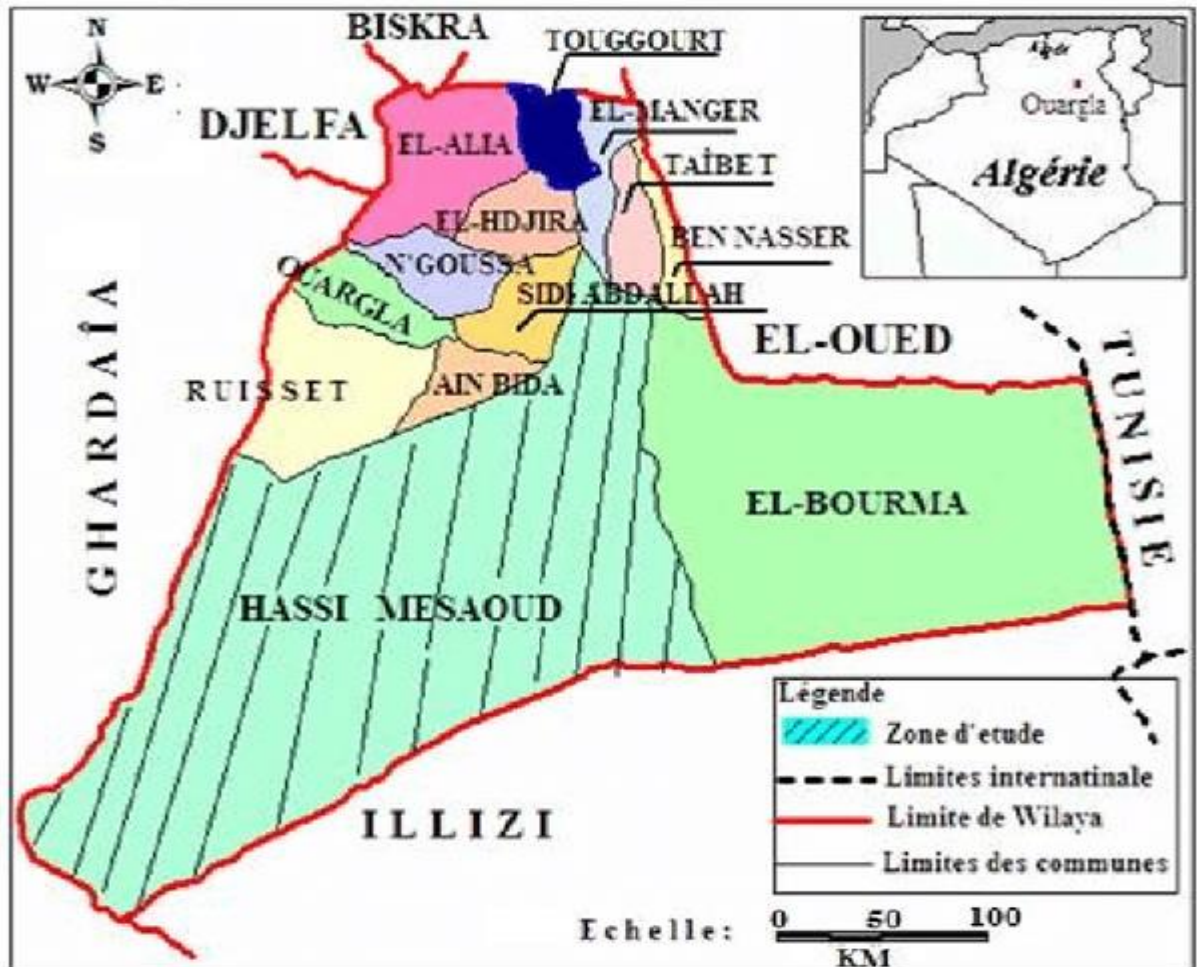


Figure.I.1 : Situation géographique du champ de Hassi Messaoud [2]

I.2- Situation géologique :

Le champ de Hassi Messaoud occupe la partie centrale de la province triasique à l'Est de la dépression d'Oued Mya dans le district IV, qui par sa superficie et ses réserves, est la plus grande province pétrogazeifère.

C'est le plus grand gisement d'Algérie qui s'étend sur 53x44 km de superficie.

Il est limité :

- Au Nord-Ouest par les gisements d'Ouargla [Gellala, Ben Kahla et Haoud Berkaoui]
- Au sud-ouest par les gisements d'El Gassi, Zotti et El Agreb.
- Au sud-est par les gisements ; Rhourde El Baguel et Mesdar.

A plus grande échelle, il est limité géologiquement :

- A l'Ouest par la dépression d'Oued Mya.
- Au Sud par le môle d'Amguid El Biod.

- Au Nord par la structure Djammâa-Touggourt.
- A l'Est par les hauts fonds de Dahar, Rhourde El Baguel et la dépression de Berkine.

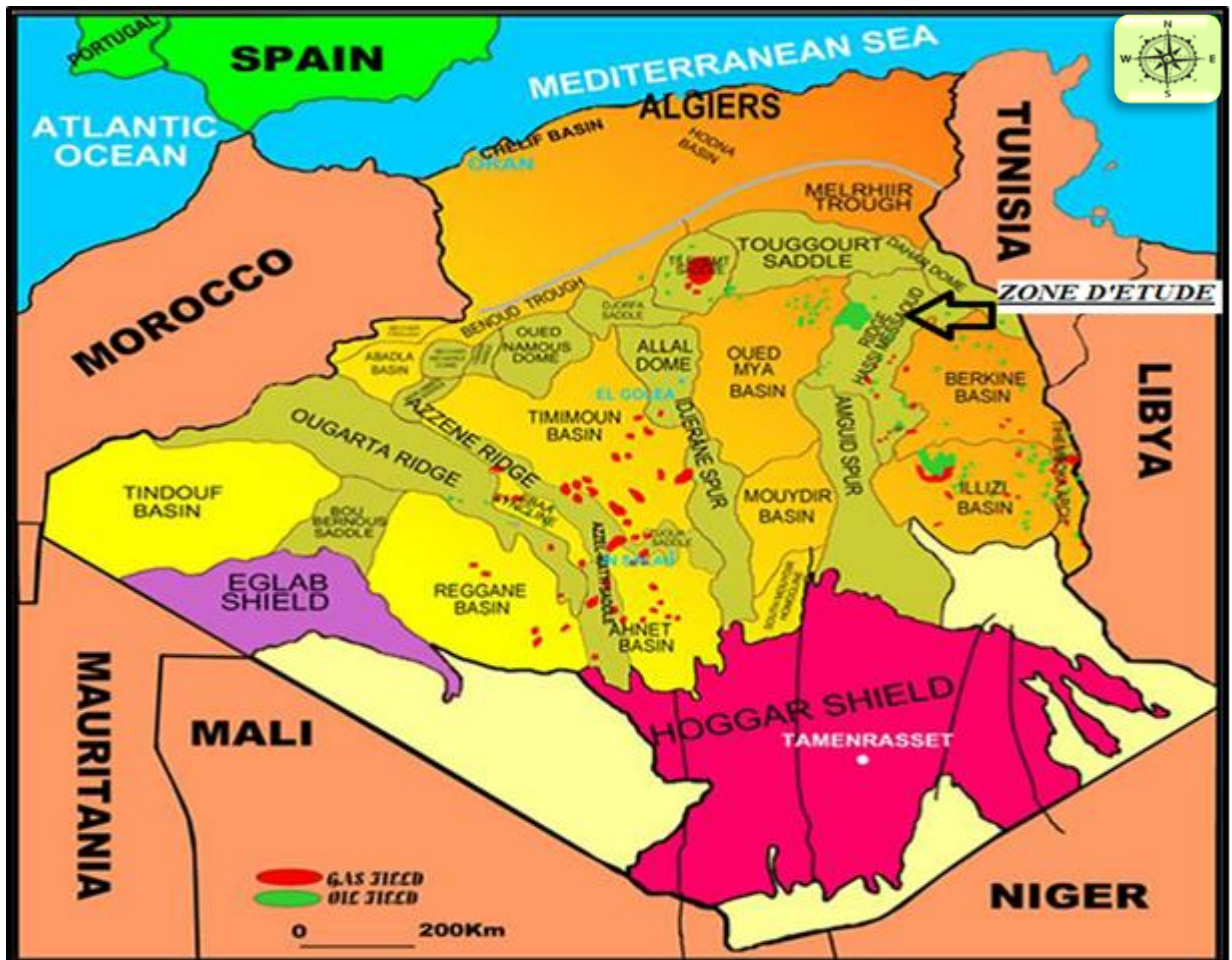


Figure.I.2 : Situation géologique du champ de Hassi Messaoud. [3]

I.3- Carte des zones et numérotation des puits :

I.3.1- Carte des zones de production à HMD :

Le champ pétrolier de Hassi Messaoud (HMD) est subdivisé en 25 zones de production, qui varient en taille. Ces zones sont relativement autonomes et regroupent des puits qui sont interconnectés et qui présentent des comportements similaires en termes de pression du gisement.[4]

Il convient de souligner que la subdivision actuelle du champ n'est pas considérée comme satisfaisante, car une zone donnée peut être divisée en sous-zones.. (Ex : 1a, 1b, 1c).

I.3.2- Zones et numérotation de puits :

Le champ de "Hassi Messaoud" est divisé en deux parties distinctes : le champ Nord et le champ Sud, chacun ayant sa propre numérotation.

- **Champ Nord** : comporte une numérotation géographique complétée par une numérotation chronologique, exemple **OMK_Z**

O : majuscule, permis d'Ouargla.

M : minuscule, carreau de 1600 km².

K : minuscule, carré de 100 km²

Z : puits horizontale.

- **Champ Sud** : La numérotation est chronologique. Ex : Md1, Md2, ..., Md544...

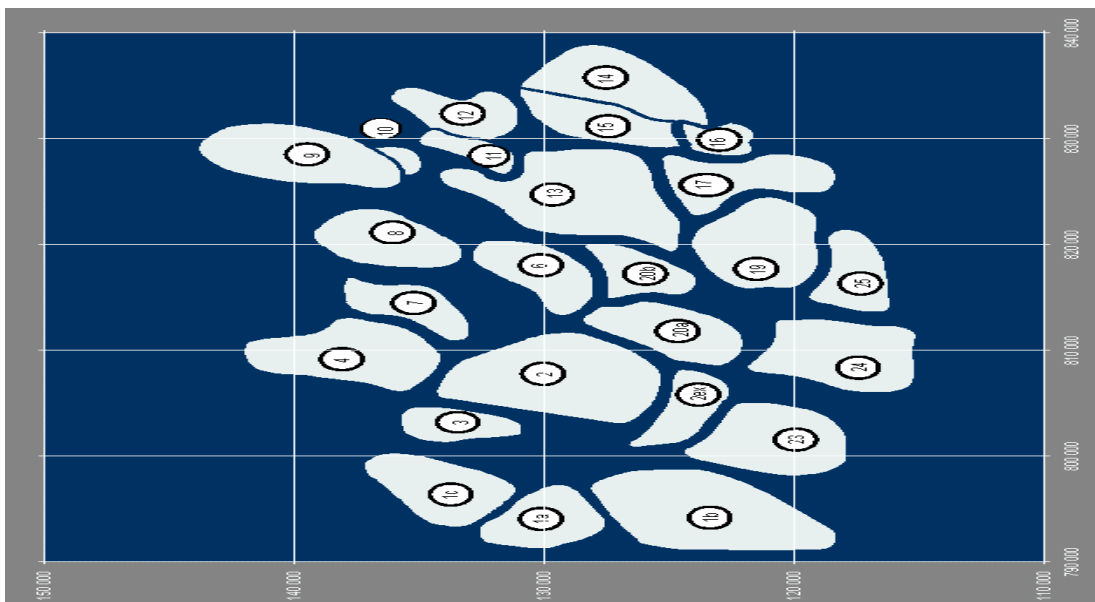


Figure.I.3 : Carte des zones de production. [4]

1.4- Situation de la zone d'étude:

Dans le cadre de réaliser d'un modèle prévisionnel de la vitesse d'avancement (ROP) à l'aide de l'intelligence artificielle, on a fait une étude sur **17** Puits.

Ces puits est défini comme des puits vertical productif d'huile qui doit être foré dans le champ de Hassi-Messaoud et réalisé dans le cadre de développement de la partie Upside Nord du champ de Hassi-Messaoud, (Voir Figure I.4). L'objectif principal l'exploitation du réservoir du Cambrien.

Les puits sont « **Omk572, Omk573, Omk853, Omk872, Omk263, Omj772, Omg302, Omg402, Omg50, Omg501, Omg504, Omg51, Omg601, Omg612, Omg831, Omh102, Oml431** ».

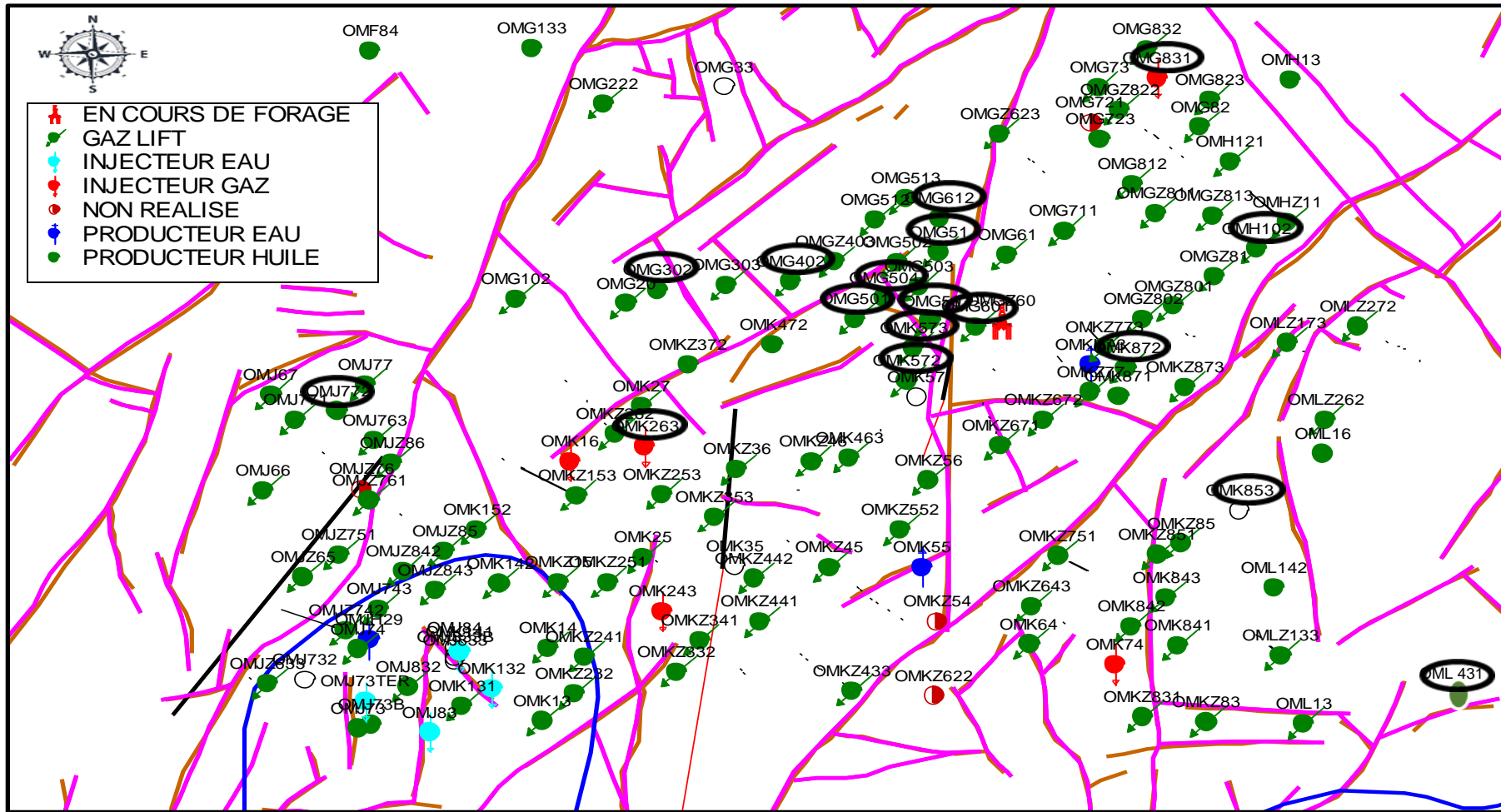


Figure.I.4 : Carte de position des puits de la zone d'étude. [4]

I.5- Description lithologique de la zone d'étude:

Sur la dorsale de Hassi Messaoud une bonne partie de la série stratigraphique est absente, ce sont les dépôts du Paléozoïque reposant sur un socle granitique, qui ont été érodés au centre de la structure au cours de la phase hercynienne.

De ce fait les dépôts mésozoïques reposent en discordance sur le Cambro-Ordovicien. En allant vers la périphérie du champ, la série devient plus complète (Figure.I.5).

a) Le Socle :

Rencontré aux environs de profondeur de 4000 mètres, il est formé essentiellement de granite porphyroïde rose.

b) - Le Paléozoïque :

Sur le socle, les formations paléozoïques reposent en discordance ; c'est la discordance panafricaine. De la base au sommet on distingue :

✓ L'infra-Cambrien :

C'est l'unité lithologique la plus ancienne reconnue au Nord de la structure par le forage OMJ47 à une profondeur de 4092 m, constitué de grès argileux rouge.

✓ Le Cambrien :

Essentiellement constitué de grès hétérogènes, fins à très grossiers entrecoupés de passées de siltstones argilo-micacés. On y distingue trois lithozones R1 (Ra + Ri), R2 et R3.

✓ L'Ordovicien :

On distingue de la base au sommet quatre unités lithologiques.

- Zone des alternances.
- Argile d'El Gassi.
- Les grès d'El Atchane.
- Les quartzites de Hamra.

c) Le Mésozoïque :

Il est subdivisé comme suit :

✓ Le Trias :

Il repose en discordance sur le Cambrien, au centre et sur l'ordovicien aux flancs de la structure. Il est subdivisé en trois unités : (Trias Gréseux, Trias Argileux, Trias Salifère)

✓ Le Jurassique :

Son épaisseur moyenne est de 844 mètres. Le Jurassique est un ensemble argilo-gréseux à intercalations de calcaire au sommet (Malm) et à alternances de faciès lagunaires et marins à la base (Dogger et Lias).

✓ Le Crétacé :

Son épaisseur moyenne est de 1620 mètres. Il est constitué de sept étages,

d) Le Cénozoïque :

Son épaisseur moyenne est de 360mètres, il est constitué de calcaire dolomitique à l'Eocène et d'un recouvrement de type sableux au Mio-Pliocène.

ERE	SYST	ETAGES		Ep moy	DESCRIPTION
CENO-ZOIQUE	NEOGENE	MIO-PLIOCENE discordance alpine		240	Sable, calcaire, marne sableuse
		EOCENE		120	Sable, calcaire à silix
MESOZOIQUE	CRETACE	SENONIEN	CARBONATE	107	Calcaire, dolomie, anhydrite
			ANHYDRITIQUE	219	Anhydrite, marne, dolomie
			SALIFERE	140	Sel massif et traces d'anhydrite
		TURONIEN	90	Calcaire crayeux avec quelques niveaux argileux	
		CENOMANIEN	145	Anhydrite, marne, dolomie	
		ALBIEN	350	Grés, sable avec intercalations d'argile silteuse	
		APTIEN	25	Dolomie cristalline avec niveau argileux, calcaire	
		BARREMIEN	280	Argile, grés, dolomie	
		NEOCOMIEN	180	Argile, marne, dolomie, grés	
	JURASSIQUE	MALM		225	Argile, marne, calcaire, grés et traces d'anhydrite
		DOGGER	ARGILEUX	105	Argile silteuse, marne dolomitique avec fines passées de grés
			LAGUNAIRE	210	Anhydrite, marne dolomitique, marne grise
		LIAS	L.D 1	65	Dolomie, anhydrite, argile
			L.S 1	90	Alternances sel, anhydrite et argile
			L.D 2	55	Anhydrite et dolomie cristalline
			L.S 2	60	Alternances sel et argile
			L.D 3	30	Alternances de dolomie et de marne
		TRIASSIQUE	SALIFERE	TS 1	46
	TS 2			189	Sel massif à intercalations d'anhydrite et argile gypsifère
	TS 3			202	Sel massif et traces d'argile
ARGILEUX	113		Argile rouge dolomitique ou silteuse injectée de sel et d'anhydrite		
GRESEUX	35		Grés, argile		
ERUPTIF discordance hercynienne	0-92		Andésites altérées		
PALEOZOIQUE	ORDOVICIEN	QUARTZITES D'EL HAMRA		75	Quartzites fines avec traces de tigillites
		GRES D'EL ATCHANE		25	Grés fins à ciment argileux, bitumineux
		ARGILES D'EL GASSI		50	Argiles schisteuses, vertes ou noires, glauconieuses à graptolithes
		ZONE DES ALTERNANCES		20	Alternance de grés et argile. Présence de tigillites
	CAMBRIEN	Ri	50	Grés isométriques, fins, silteux	
		Ra	120	Grés à grés quartzitiques anisométriques à niveaux de silts	
		R2	100	Grés moyens à grossiers à ciment argileux illitique	
		R3	300	Grés grossier à ciment argileux, argile silteuse	
	INFRA-CAMBRIEN		45	Grés argileux rouges	
	S O C L E				

Figure.I.5 : Fiche stratigraphique de la zone d'étude.

CHAPITRE

II

LES OUTILS DE FORAGE

PARTIE I : GENERALITES SUR LES OUTILS DE FORAGE**Introduction :**

L'outil de forage est la pièce la moins volumineuse mais la plus importante pour réaliser un puits de forage.

C'est le premier élément à s'attaquer le terrain, donc pour améliorer la performance de l'outil et minimiser le coût de forage, il faut choisir un outil adapté à la formation traversée et compatible à la puissance fournie

Le choix de l'outil de forage doit être effectué en fonction de la nature des terrains à traverser, car chaque outil possède des caractéristiques qui le rendent plus ou moins adapté à un type de terrain spécifique. [5]

I- Les différents types des outils de forage :

Les outils se classent en deux catégories :

I.1- Les outils à molettes :

Les outils à molettes se composent de trois parties principales : les molettes, les roulements et le corps de l'outil, les molettes comportent des rangées circonférentielles de dents interposées entre des rangées de dents de cônes adjacents, les dents sont soit en acier usiné dans le cône, soit des broches en carbure de tungstène sont encastrées dans des trous dans la face du cône, les molettes sont montées sur des arbres porteurs qui font partie intégrante du corps de l'outil.

La dimension ou l'épaisseur des différentes parties de l'outil varient en fonction du type de formation à forer, pour les formations tendres qui requièrent peu de pression, les roulements sont plus petits, les cônes sont moins épais et les jambes ont une section plus fine par rapport aux outils conçus pour les formations dures, cette conception permet de libérer plus d'espace pour des éléments de coupe longs et fins.

Les outils pour formations dures, qui travaillent avec des poids plus importants, ont des éléments de coupe plus trapus, des roulements plus gros et des corps plus solides. [1]



Figure.II.1 : Les outils à molettes (tricône)

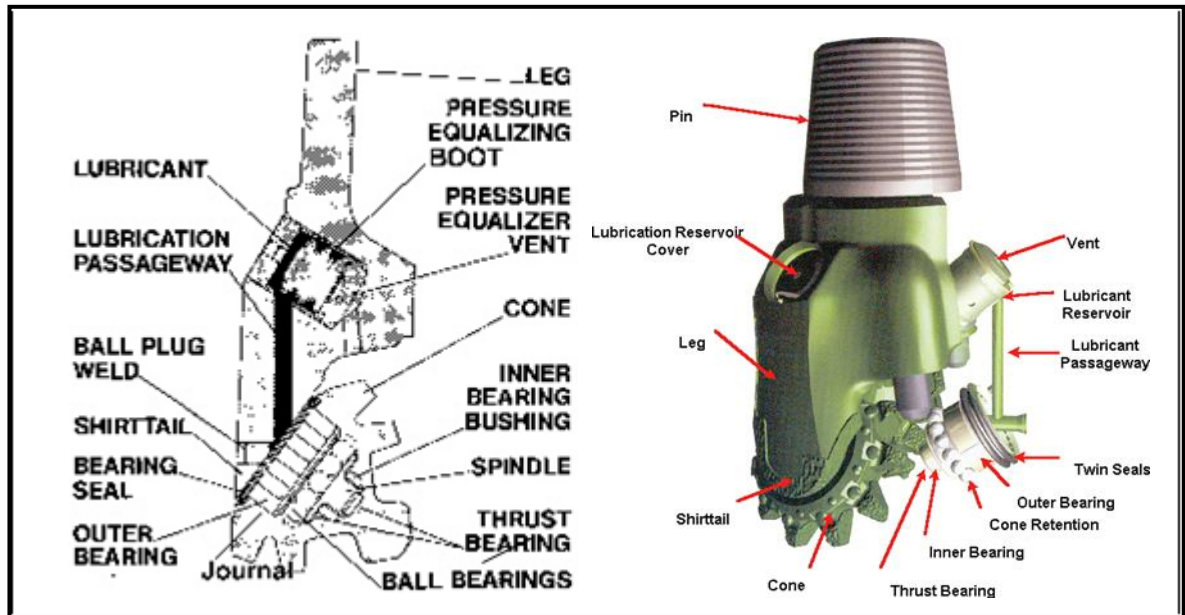


Figure.II.2 : Nomenclature des tricônes. [6]

I.1.1- Géométrie des cônes :

Pour comprendre comment la géométrie des molettes peut affecter la façon dont les dents coupent la roche, considérons le cône pour formations tendre schématisé sur la (Figure.II.3). Il y a deux angles coniques de base, et aucun d'entre eux n'a son sommet au centre de l'outil.

Comme les cônes sont obligés de tourner autour de l'axe de l'outil, ils glissent en même temps qu'ils tournent et ont une action de gouge de cisaillement efficace pour le forage des formations tendres.

L'augmentation de cet effet pour améliorer la vitesse d'avancement dans les formations tendres s'obtient en décalant les axes des cônes par rapport à l'axe de rotation de l'outil

Les cônes des outils pour formations dures ont un mouvement plus proche du roulement vrai, et présentent peu d'offset.

De ce fait, ils brisent la roche surtout par écrasement. [1]

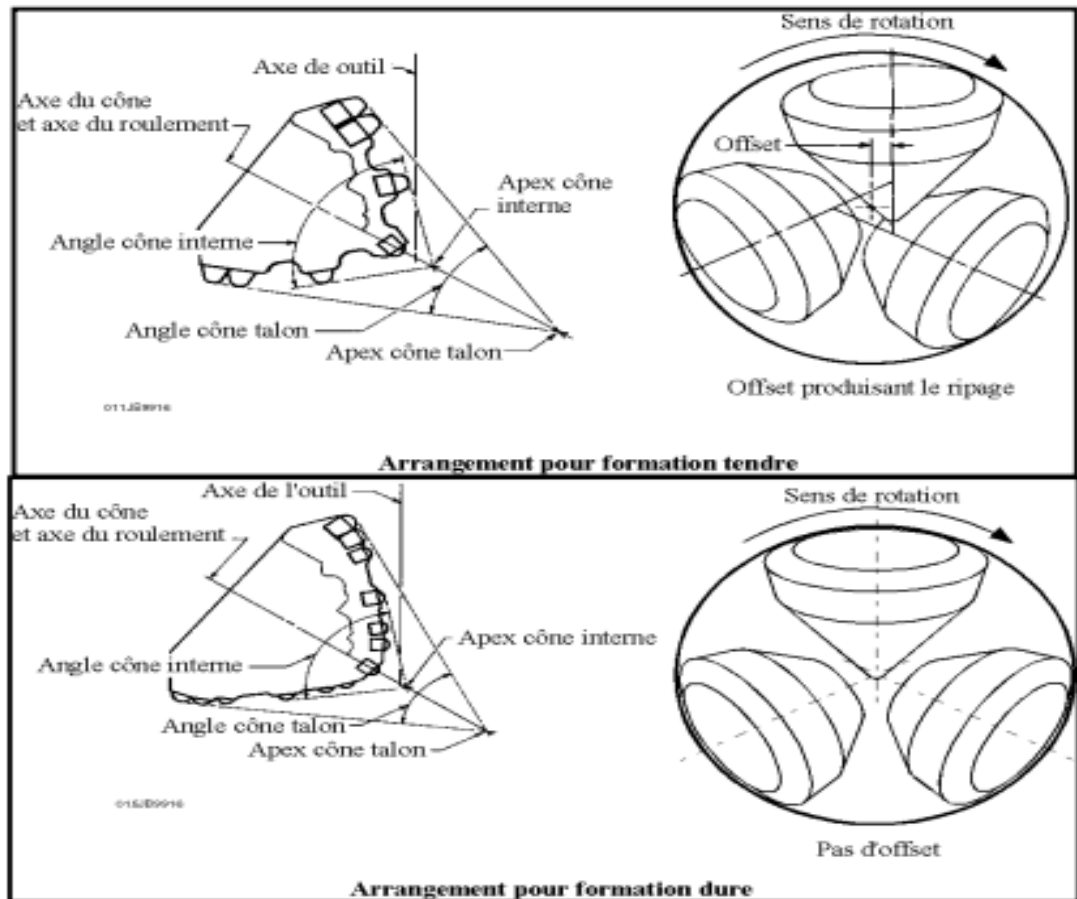


Figure.II.3 : Principe de construction des molettes des tricônes [1]

I.1.2- Caractéristiques des tricônes :

a)- Outils pour terrains tendres :

Les outils conçus pour les terrains tendres ne requièrent pas un poids important, leurs dents sont hautes et espacées pour pénétrer efficacement dans la roche et extraire de gros fragments. Toutefois, la hauteur importante des dents implique une molette plus fine, ce qui entraîne une réduction des chemins de roulement, ainsi que la taille des billes et des galets.

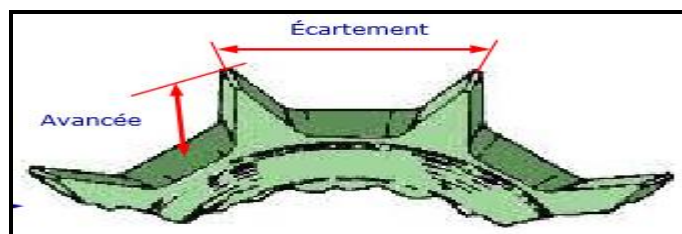


Figure.II.4 : Les dents pour terrain tendre [7]

b)- Outils pour terrains durs :

Dans ces types de terrains, la roche a une résistance à la compression importante, ce qui peut entraîner la rupture des dents de l'outil.

Afin de prévenir cela, des outils comportant un grand nombre de petites dents solides et rapprochées sont utilisés, avec un faible décalage entre les axes des molettes et celui de l'outil.

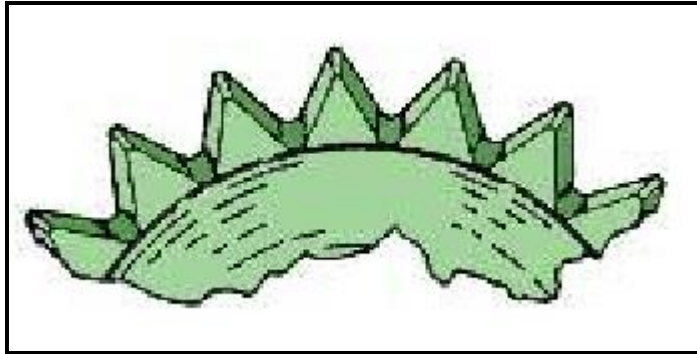


Figure.II.5 : Les dents pour terrain dure. [7]

d)- Les outils pour les terrains abrasifs :

Dans les formations abrasives, les outils sont renforcés au niveau du talon des molettes avec des pastilles plates en carbure de tungstène pour maintenir le diamètre de l'outil.

Certains outils remplacent les dents par des pastilles en carbure de tungstène qui écrasent, fracturent et percutent la roche sans nécessiter une grande pénétration.

Ces outils sont équipés de molettes épaisses pour supporter des poids importants et les chemins de roulement sont également larges. [7]



Figure.II.6 : Outil renforcé sur le talon

I.1.3- Les roulements :

Les roulements supportent la majeure partie du poids exercé sur l'outil. On en distingue deux types :

a)- Les roulements à rouleaux (étanches et non étanches) :

Ces outils utilisent des roulements qui supportent le poids uniquement sur les deux génératrices opposées du rouleau.

Cette conception peut causer une déformation du roulement sous une forte charge, ce qui le rend peu adapté aux poids importants.

Toutefois, ce type de roulement résiste mieux aux températures élevées et supporte des vitesses de rotation importantes. [7]

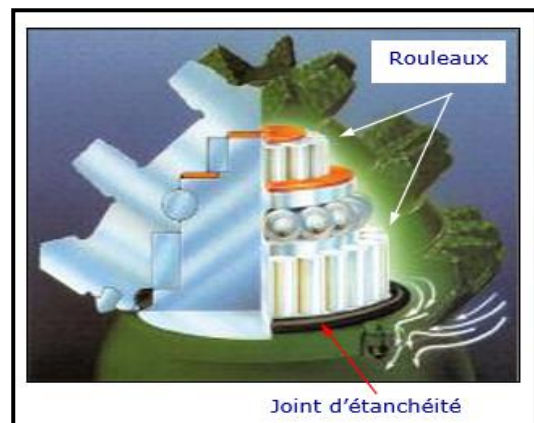
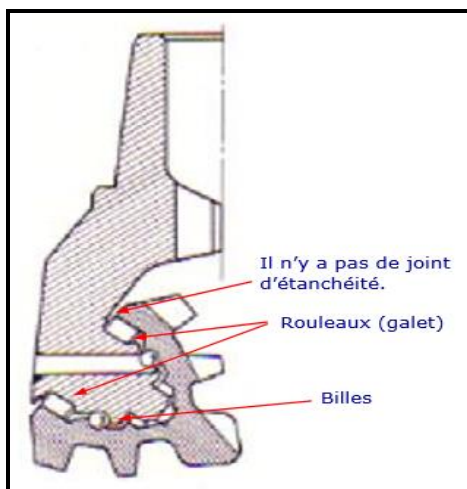


Figure.II.7 : Roulement standard non étanche **Figure.II.8 : Roulement à rouleaux étanche**

b)- Les paliers de friction :

Le palier de friction développé pour faire coïncider sa durée de vie avec celle des éléments de coupe au carbure ne contient aucun élément mobile, mais uniquement un palier mâle ajusté sur l'alésage interne du cône.

Les zones de contact entre l'arbre du palier et l'alésage du cône sont carburées, revêtues de métaux spéciaux et reçoivent un traitement spécial augmentant la résistance à l'usure et protégeant du grippage. [1]

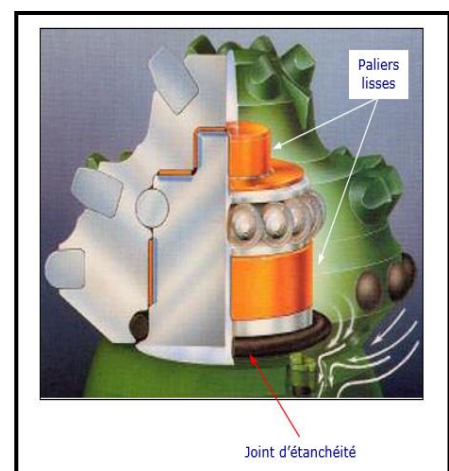


Figure.II.9 : Roulement étanche à Friction

c)- Lubrification des roulements :**• Roulements non étanches :**

Ces roulements ne sont pas étanches et sont graissés directement par la boue de forage. Ce sont en général les outils destinés au début du forage. [7]

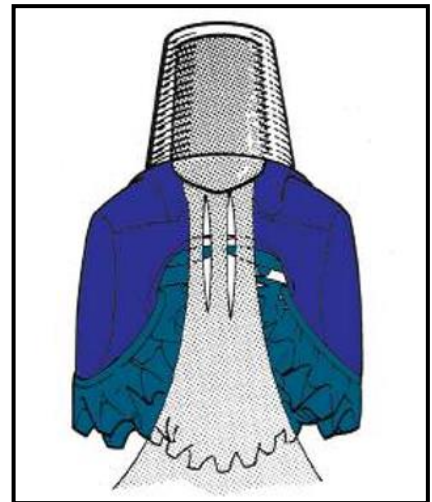
• Roulements étanches :

Les roulements étanches sont graissés grâce à une réserve de graisse située dans le dos du bras et actionnée par la pression hydrostatique dans l'annulaire. Il est isolé de la boue par un joint d'étanchéité.

I.1.4- L'hydraulique à l'outil :**a)- Outil conventionnel :**

Dans l'outil conventionnel, le jet de boue est dirigé sur les molettes à partir d'un évent situé au centre de l'outil. La vitesse à la sortie de cet évent est faible et insuffisante pour nettoyer le fond et favoriser la vitesse de pénétration.

Les déblais ne sont pas remontés immédiatement, mais rebroyés. Ce système hydraulique est utilisé actuellement sur les outils de grands diamètres ou ceux destinés aux formations tendres de faible profondeur. [7]

**Figure.II.10 : Outil conventionnel****b)- Outil à duses :**

L'outil à duses comporte des duses interchangeables situées sur ses côtés, entre les molettes.

Le jet de boue est dirigé sur le front de taille et crée une turbulence qui refoule instantanément les déblais, pour laisser un fond propre devant l'outil. Dans les terrains tendres, le jet de boue sortant des duses aide à la destruction de la roche, ce qui augmente l'avancement.

Ces outils forment des déblais de dimensions plus grandes que ceux obtenus par des outils conventionnels.

**Figure.II.11 : Outil à duse.**

I.1.5- Mode de travail d'un outil Tricône :

L'efficacité d'un outil à molettes dépend de sa capacité à broyer et évacuer les morceaux de roche. Ces outils travaillent par :

- Compression produisant la pénétration de la dent ou du picot dans la formation et l'éclatement de la roche. Plus le terrain est tendre, plus la dent devra être longue pour produire le meilleur avancement. L'outil produit une série de petits cratères résultant de la pression exercée par chaque dent sur le front de taille. La contrainte de compression appliquée sur la formation induit une fracture.

- Arrachage du copeau de terrain par ripage de la molette sur elle-même.

Ces deux effets sont complémentaires. Ils sont privilégiés l'un par rapport à l'autre suivant la dureté de la roche à broyer. Pour un terrain tendre, l'effet de ripage sera prépondérant.

Ces deux effets sont gouvernés par la géométrie des cônes. Pour privilégier l'effet de ripage, la rotation du cône doit s'éloigner de façon significative d'une rotation vraie. [8]

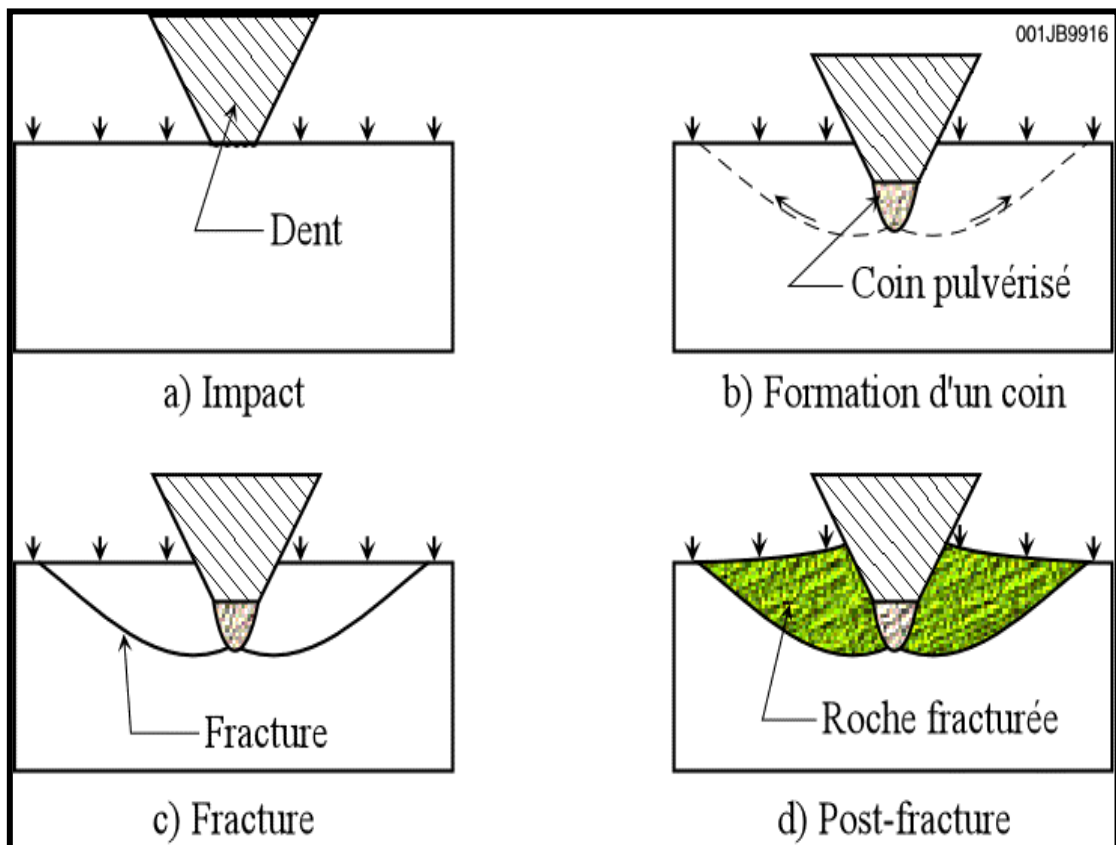


Figure.II.12 : Mécanisme d'action d'un outil à molettes sur la roche. [8]

I.1.6- Classification IADC des outils tricônes :

Le système de classification des outils comprend quatre caractères et permet de classer tous les outils tricônes à denture fraisée et à pastilles. [9]

1er caractère :

- un chiffre de 1 à 3 : tricône à denture fraisée,
- un chiffre de 4 à 8 : tricône à pastilles.

2ème caractère : Un chiffre de 1 à 4 qui désigne la dureté de chaque groupe de formations. C'est une sous-classification de la dureté des formations dans chacune des 8 classes déterminées par le 1er chiffre.

3ème caractère : Un chiffre de 1 à 9 qui désigne les caractéristiques du tricône.

4ème caractère : Une lettre de l'alphabet qui définit des caractéristiques additionnelles des outils tricônes.

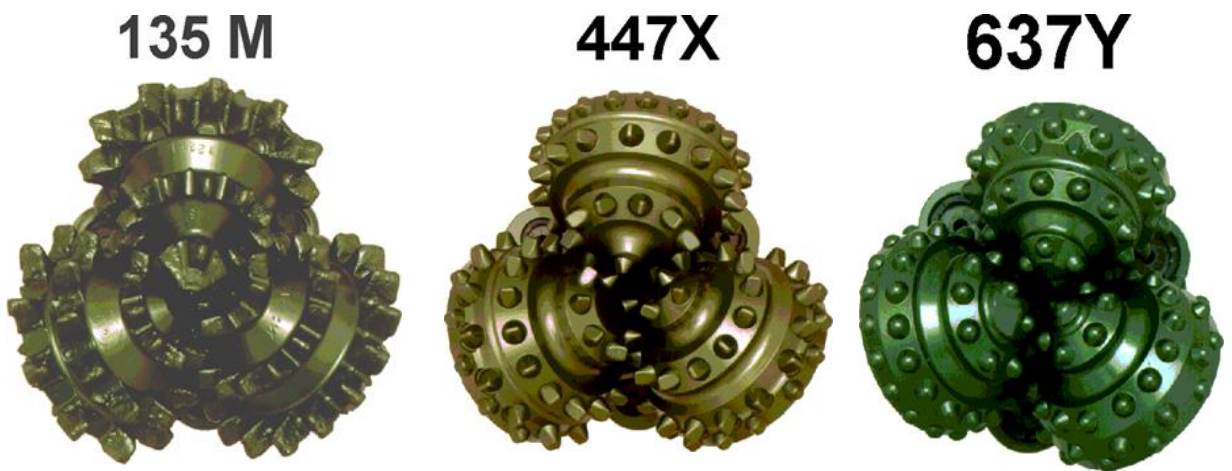


Figure.II.13 : Exemple des codes IADC. [10]

I.2- Les outils à éléments de coupe fixes :

Ils ne possèdent pas de pièces tournantes, ce sont les outils monobloc, des diamants naturels et de synthèse sont utilisés pour leur fabrication.

Les outils à diamants naturels travaillent à la façon d'une lime tandis que les outils à diamants synthétiques travaillent à la façon d'un rabot.

I.2.1- Les outils PDC (*Polycrystalline Diamond Compact*) :

Il existe deux types d'outils : ceux avec un corps en acier et ceux avec une matrice. Les corps en acier sont usinés et ensuite revêtus de carbure de tungstène pour réduire l'usure.

Les corps en matrice sont fabriqués à partir du même matériau au carbure de tungstène que les outils dotés de diamants naturels.

Les outils PDC (polycrystalline à pointe diamant) sont ainsi appelés "outils monobloc" car ils ne possèdent aucune partie mobile, ce qui constitue leur principal avantage.

Cependant, leur inconvénient majeur est leur incapacité à résister à des températures supérieures à 800 °C.



Figure.II.14 : Les outils PDC. [11]

I.2.2- Types de diamant :

Dans le domaine de forage rotary, il y a deux types de diamant :

- a- Diamant naturel.
- b- Diamant synthétique.

a)- Le diamant naturel :

C'est une cristallisation du carbone de densité 3,52. Du point de vue chimique le diamant et le graphite sont identiques. Alors que du point de vue physique, ils sont totalement contradictoires. Le graphite est fragile, noir, faible, souvent utilisé comme lubrifiant.

Le diamant par contre est extrêmement dur meilleur conducteur thermique, hautement abrasif ; la plupart du temps transparent, il est aussi un excellent isolant.

Tableau.II.1 : Caractéristique du diamant naturel

Caractéristiques	Unité	Diamant	PDC	Carbure tungstène	Acier
Dureté	Kg/mm ²	6000-9000	5000-8000	1475	558
Résistance à la compression	10 ³ psi	1260	890	780	238
Conductivité thermique (25°)	W/cm. °C	5.2	5.43	1.00	0.48
Densité	g/cm ³	3.52	3-3.25	14.95	785
Point de fusion	°C	3650	-	2860	1530
Résistance à l'écrasement	Bars	80000	-	50000	15000
Module de Young	10 ⁶ psi	105-152	132	92	29

Le diamant est le matériau le plus dur et a des caractéristiques supérieures parmi les matériaux que l'on connait.

Pour cela, c'est l'élément essentiel dans l'outil de forage, qui résiste aux conditions de fond (haute température et haute pression).

b)- Le diamant synthétique :

On distingue ces deux produits polycristallins :

- ✓ Le diamant synthétique basse température "PDC",
- ✓ Le diamant synthétique haute température "TSD".

➤ **Avantages de diamant synthétique au diamant naturel :**

- Les deux produits (PDC, TSD) sont plus résistants à l'abrasion que les diamants naturels mais ils sont moins durs.

- La raison de l'énorme succès des pierres synthétiques et que le scientifique a la possibilité de fabriquer des diamants qui répondent exactement aux exigences des industriels.

✓ Le diamant synthétique bas température :

Le PDC se présente en général sous la forme d'une fine lamelle de diamant synthétique (épaisseur = 0.5 mm) avec des impuretés de Cobalt, placée sur substrat de carbure de tungstène.

La capacité de coupe de cet outil reste la même durant toute sa durée de vie. En effet, le support en carbure de tungstène s'use plus rapidement que la partie diamantée polycrystalline, ce qui provoque son auto-affûtage d'une façon constante et progressive jusqu'à la base du corps de l'outil. Ayant une stabilité thermique faible (environ 700°C), ces outils sont utilisés dans des formations très tendres et moyennement dures non abrasives.

✓ Le diamant synthétique haut température :

Le TSD (Thermally Stable Diamond) se compose de diamant synthétique sans cobalt et sans le substrat de carbure de tungstène.

Ce type d'outil travaille d'une façon optimale tant qu'il peut conserver son angle d'attaque. Par la suite, son usure augmente et sa surface de contact devient moins efficace. Ayant une bonne stabilité thermique (environ 1200°C), ces outils sont utilisés dans des formations moyennement dures à dures à abrasivité faible.

I.2.3-Mode de travail d'un outil PDC :

Les efforts appliqués sur le cutter :

- La force de coupe F_c (horizontal force),
- La force normale F_n (poids sur l'outil),
- La force latérale F_L .

Lorsque un cutter PDC coupe une roche à une profondeur de passe h , elle est soumise à un effort résultant R , dont la composante normale à la face de la pastille appelée effort de coupe F_c est responsable de la rupture de la roche par cisaillement. Cet effort résultant présente également une composante latérale F_L . Les efforts latéraux de chaque pastille sont équilibrés entre eux de sorte que l'effort latéral est nul.

✓ Principe de fonctionnement :

Le balayage du fond du trou est obtenu par le mouvement de rotation de l'outil autour de l'axe de forage, tous les points de l'outil PDC décrivent dans son mouvement des cercles concentriques, en particulier chacun des éléments de l'outil, qui est maintenu enfoncé dans la roche sous l'effet de la charge verticale, repousse dans ce mouvement de rotation un élément de roche, dont l'épaisseur est égale à la profondeur dont il s'est enfoncé, et à chaque instant il écrase un élément nouveau de roche vierge.

Si la vitesse de rotation de l'outil et la résistance de la roche sont uniformes, chaque élément d'outil, en contact avec le fond du trou, aura une trajectoire hélicoïdale dont le pas sera égal à l'avancement de l'outil par tour.

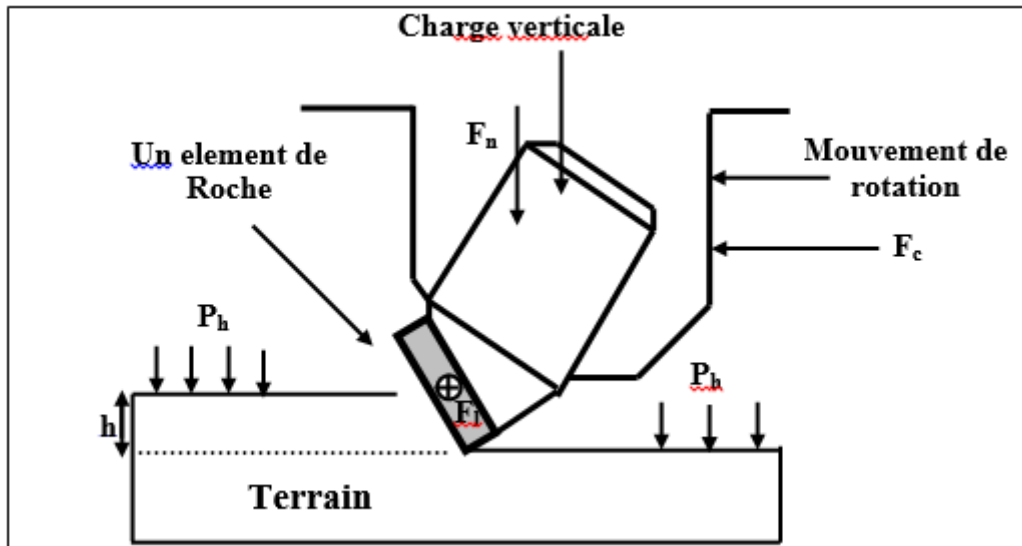


Figure.II.15 : Principe de fonctionnement d'un outil PDC

I.2.4- Désignation d'un outil PDC :

On désigne un outil par :

- 5- Le corps de l'outil,
- 2- Le profil,
- 3- Densité des cutters,
- 4- Angles caractéristiques de parties,
- 5- Longueur utile (L_{edge}).

a)- Le corps de l'outil :

Actuellement, il existe deux types de corps d'outils.

- ✓ Corps en acier (steel body),
- ✓ Corps en carbure de tungstène (Matrix body).

✓ **Steel body :**

Le corps est en acier monobloc obtenu par usinage. Ces outils ont des taillants PDC sertis directement dans le corps sans brasage. Le diamètre est protégé par des barres en carbure de tungstène sur toute la hauteur de la jupe extérieure. Pour réduire le phénomène d'érosion du corps, la surface d'attaque de l'outil est protégée par une métallisation dure. Ce type d'outils offre la possibilité d'utiliser toute la surface du taillant, étant donné qu'il est entièrement exposé, et de remplacer les taillants après utilisation s'il n'y a pas eu érosion du corps. Par contre, les taillants ne peuvent pas supporter un torque important et se cassent à la base du corps. [12]

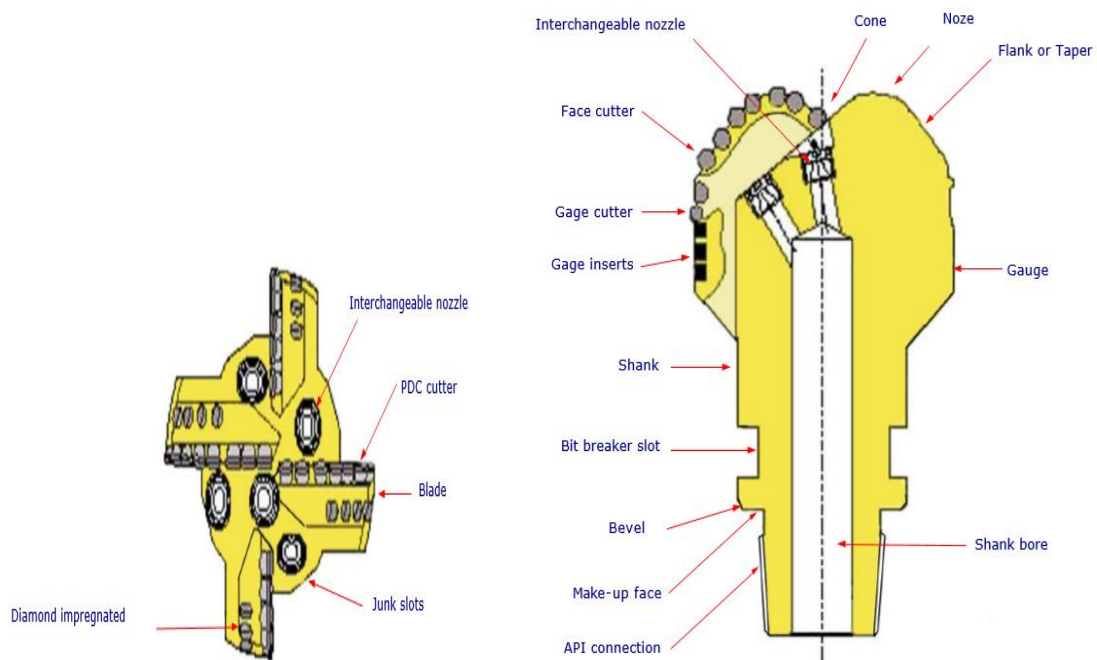


Figure.II.16 : Terminologie outil PDC - Steel body. [7]

➤ **Les avantages de Steel body :**

- possibilité de remplacer les cutters périscope après utilisation de l'outil si toute fois il n'y a pas eu érosion du corps.
- possibilité de pouvoir utiliser toute la surface du cutter étant donné que ce type de cutter est entièrement exposé.

➤ **Les inconvénients du Steel body :**

- Une usure par "Wash out" très rapide du corps acier ce qui limite considérablement sa durée de vie. Cette durée de vie est directement proportionnelle à la quantité de solides de contenus dans la boue, toutefois cette usure par "Wash out" peut être réduite en apportant une métallisation dure sur la surface d'attaque de l'outil.

• **Matrix body :**

Le corps, en deux parties soudées, a une âme en acier et une matrice en carbure de tungstène. Ces outils ont des taillants cylindriques brasés directement dans un évidage dans le corps, qui, en plus, les protège contre les chocs. Ainsi, ils peuvent travailler avec des torques très importants. La durée de vie du corps est supérieure à celle du taillant, du fait de sa très haute résistance à l'abrasion (carbure de tungstène). Dans ce type d'outils, on ne peut pas utiliser toute la surface du taillant. Une fois utilisé, cet outil ne peut plus être réparé. [12]

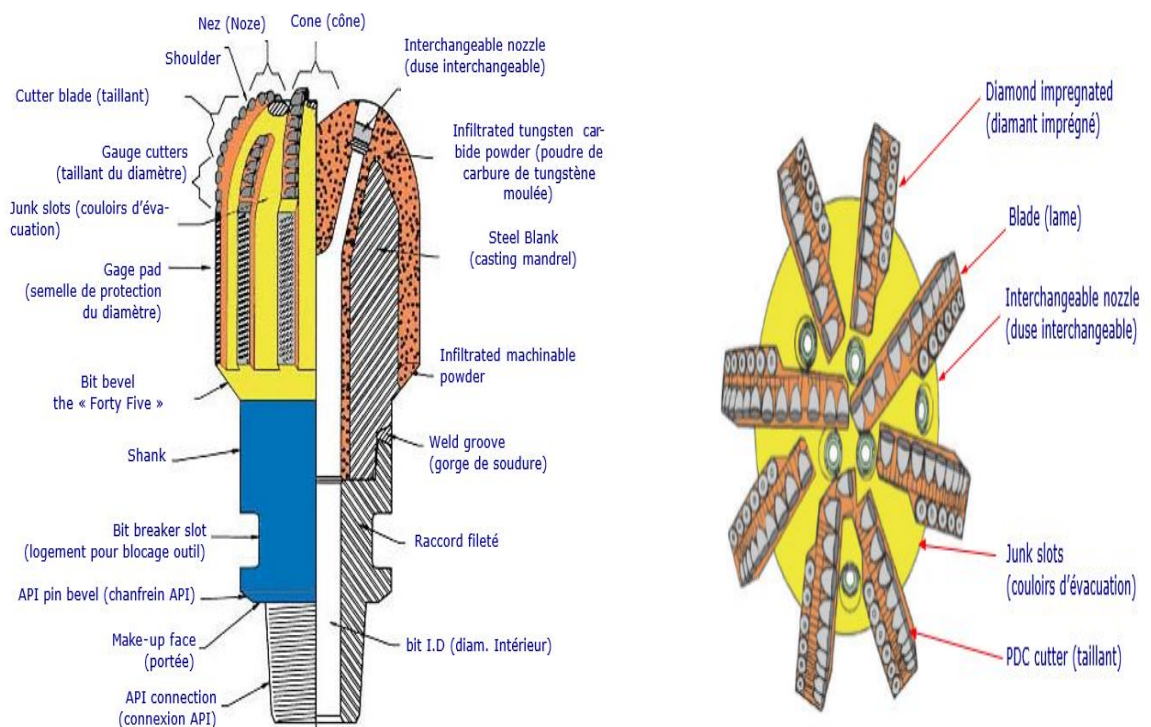


Figure.II.17 : Terminologie outil PDC - Matrix body. [7]

➤ **Les avantages du Matrix body :**

- La très haute résistance à l'abrasion du corps qui est en carbure de tungstène, ce qui donne une durée de vie toujours supérieure à la durée de vie cutter lui-même.
- L'évidage aménagé dans le corps de l'outil pour recevoir le cutter donné.

➤ **Les inconvénients de Matrix body :**

- son prix qui est sensiblement plus élevé que le steel body.
- une fois utilisé, cet outil ne peut pas être réparé.
- toute la surface de cutter ne peut pas être utilisée du fait de sa position dans la matrice.

b)- Le profil :

Il existe principalement trois types de profils :

- Profil plat ou avec un faible cône intérieur.
- Profil double cône (cône intérieur très prononcé).
- Profil parabolique.

✓ Le profil plat :

L'angle du cône est inférieur à 15°. Dans ce type de profil, le poids se répartit uniformément sur les taillants, mais le nombre de ces derniers est limité et leur usure est intégrale du fait de la stabilité de l'outil.

Les outils ayant ce profil sont utilisés pour des terrains tendres et sont peu favorables aux avancements rapides. Ils sont plus économiques dans des terrains consolidés. [6]

✓ Le profil double cône :

Le cône intérieur est très prononcé. Les taillants sont répartis d'une façon croissante vers la périphérie, améliorant ainsi la stabilité et la précision directionnelle

Les outils ayant ce profil sont utilisés pour des terrains durs.

✓ Le profil parabolique :

Court ou long, ce type de profil présente une surface importante sur laquelle un grand nombre de taillants est fixé.

Le profil court ou moyen a l'avantage de réduire le couple résistant durant le forage, ce qui permet d'appliquer davantage de poids sur l'outil et, par conséquent, augmenter la vitesse d'avancement.

En outre, le profil parabolique est plus facile à diriger avec un moteur de fond dans un puits dévié.



Figure.II.18 : Type des profils PDC. [6]

c)- Densité des cutters :

Lorsqu'on augmente la densité des taillants la vitesse de pénétration diminue, mais la charge effective par taillant est réduite, ce qui prolonge la durée de vie de l'outil.

Le nombre et le positionnement des taillants sont déterminés de façon à ce que chaque taillant fournisse la même quantité de travail.

Certains outils ont des lames renforcées d'une autre rangée de taillants disposée juste derrière celle d'attaque.

$$\text{Contrainte par cutter} = \frac{WOB_{total} / \text{nombre de cutter}}{\text{La surface de coupe}}$$

d)- Angles caractéristiques des pastilles PDC :

On définit trois angles caractéristiques sur les pastilles PDC :

- 1- L'angle de coupe W_c (back rake angle).
- 2- L'angle de dépouille W_d (clearance angle).
- 3- L'angle lateral W_s (Side rake angle).

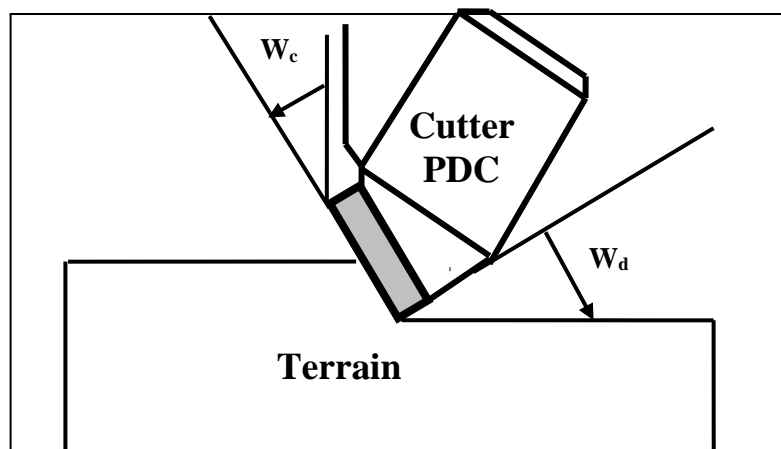


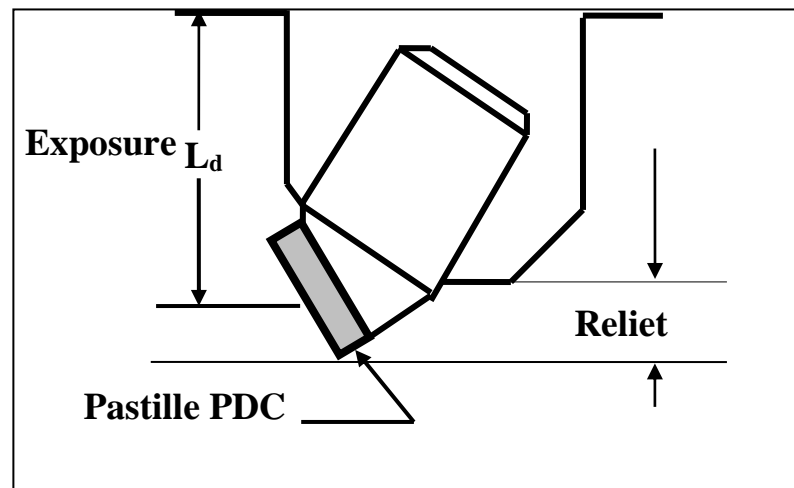
Figure.II.19 : Angles caractéristiques d'un PDC

f)- Longueur utile L_d :

La distance entre la matrice de l'outil et le fond de la saignée creusée par le PDC est appelé "Longueur utile" (exposure).

Si L_d est la distance mesurée suivant la verticale entre la matrice du trépan et le centre de la pastille de diamètre D . Alors le rapport L_d/D est appelé L_{edge} .

Ces deux paramètres sont d'une importance capitale dans le système d'évacuation des copeaux et du nettoyage du puits et sur les efforts latéraux.

Figure.II.20 : Longueur utile L_d

I.2.5- Classification IADC des outils diamants :

Un système de codification à quatre caractères permet de classer tous les outils diamants.

1^{ier} caractère : Type de diamants et de matrice :

D : outil à diamants naturels possédant une matrice en carbure de tungstène

M : outil PDC à matrice en carbure de tungstène

S : outil PDC à matrice en acier

L : outil TSD à matrice en carbure de tungstène

O : autres types d'outils qui seront développés dans le futur.

2^{ème} caractère : Profil d'outils : les profil d'outils sont codifiés à l'aide de neuf chiffres de 1 à 9 qui représentent les formes usuelles des outils diamants.

3^{ème} caractère : Caractéristiques hydrauliques :

R pour sortie radiale du fluide

X pour sortie avec gorge centrale

O pour les autres types de sortie du fluide.

L'hydraulique des outils à diamants synthétiques est codifiée par neuf chiffres (de 1 à 9).

4^{ème} caractère : Taille et densité des diamants : neuf chiffres (1 à 9) symbolisent le type, la taille et la densité des diamants. La taille des diamants naturels est basée sur le nombre de pierres par carat, alors que la taille des diamants synthétiques est basée sur la dimension du taillant. La densité de diamants donne une indication relative de leur nombre afin de distinguer les outils fortement chargés de ceux faiblement chargés.

I.2.6- Les outils TSP (Thermally Stable Polycrystalline):

Ces outils sont fabriqués à partir de diamant synthétique, ce qui les rend extrêmement résistants à la chaleur et à l'usure.

Les outils TSP sont utilisés pour forer des formations géologiques très dures, telles que les formations de basalte, de dolérite et de granit. Ils sont également utilisés pour le forage directionnel et horizontal, où l'usure de l'outil est particulièrement importante.

Les TSP peuvent être regroupés pour former des structures en mosaïque de la dimension des PDC.

Cette structure a permis d'augmenter la vitesse d'avancement et d'allonger la durée de vie de l'outil dans des formations tendres, abrasives et dans les intercalations d'argiles et de roche plus dures, là où les PDC n'étaient pas rentables.

Les outils TSP présentent plusieurs avantages par rapport aux outils traditionnels de coupe utilisés dans l'industrie pétrolière, tels que les outils en carbure de tungstène. Voici quelques-uns de ces avantages :

- Résistance à la chaleur : Les outils TSP sont fabriqués à partir de diamant synthétique, ce qui leur confère une grande résistance à la chaleur. Ils peuvent donc être utilisés pour forer des formations géologiques à haute température, telles que les formations de basalte et de dolérite.
- Durée de vie plus longue : Les outils TSP ont une durée de vie plus longue que les outils en carbure de tungstène. Cela est dû à leur résistance à l'usure et à la chaleur, qui leur permet de fonctionner plus longtemps avant de nécessiter un remplacement.
- Capacité à forer des formations plus dures : Les outils TSP sont capables de forer des formations géologiques plus dures que les outils en carbure de tungstène. Ils peuvent donc être utilisés pour forer des formations géologiques difficiles, telles que les formations de granit.
- Meilleure performance en forage directionnel et horizontal : Les outils TSP sont particulièrement adaptés pour le forage directionnel et horizontal, où l'usure de l'outil est importante. Leur durée de vie plus longue et leur capacité à forer des formations géologiques plus dures les rendent plus efficaces pour ces types d'applications.



Figure.II.21 : Les outils TSP.

I.2.7- Les outils imprégnés :

Ces outils sont fabriqués en imprégnant de petits diamants synthétiques dans une matrice de carbure de tungstène ou de cobalt.

Les outils imprégnés sont utilisés pour forer des formations géologiques tendres à moyennement dures, telles que les formations de calcaire, de grès et de schiste. Ils sont également utilisés pour le forage directionnel et horizontal, où la précision de l'outil est importante.

Les outils imprégnés ont plusieurs avantages par rapport aux outils traditionnels en carbure de tungstène, notamment leur durée de vie plus longue, leur précision de coupe accrue et leur capacité à forer des formations géologiques plus tendres.

Cependant, ils sont moins résistants à la chaleur et à l'usure que les outils TSP, ce qui limite leur utilisation dans certaines applications.

En résumé, les outils imprégnés sont un choix populaire dans l'industrie pétrolière pour le forage de formations géologiques tendres à moyennement dures, où la précision de l'outil est importante. Ils offrent une durée de vie plus longue et une précision de coupe accrue par rapport aux outils traditionnels en carbure de tungstène, mais sont moins résistants à la chaleur et à l'usure que les outils TSP.

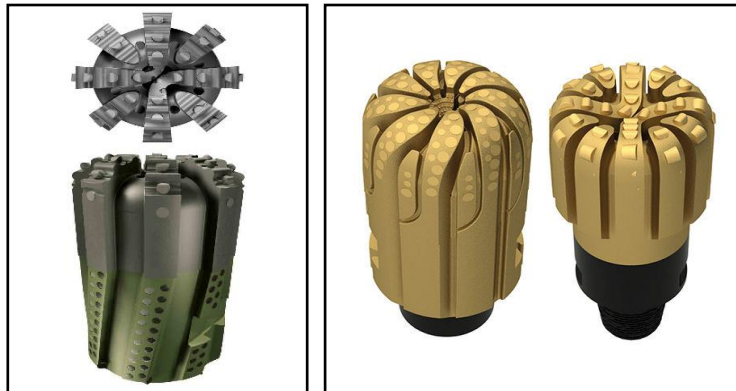


Figure.II.22 : Les outils imprégnés

✓ Fabrication des outils imprégnés :

Pour l'imprégnation, on utilise des diamants de très petite taille (150 pierres par carat, ce qui correspond à un diamètre de l'ordre de 0.8 mm), l'imprégnation se fait soit dans la masse de la matrice sur une épaisseur de l'ordre du cm.

Au fond du moule, on place un mélange de diamants et de poudre servant à la fabrication de la matrice et l'ensemble est passé au four.

Soit, on fabrique des segments de carbure de tungstène imprégnés de diamants. Ces segments sont soit brasés, soit emmanchés en force dans le corps de l'outil.

Conclusion :

Le choix et les types d'outils de forage sont des éléments essentiels pour la réussite d'un projet de forage.

Les outils doivent être choisis en fonction des caractéristiques géologiques et géométriques du terrain à traverser, ainsi que des conditions de forage spécifiques, telles que la profondeur, la pression, la température et la présence de fluides corrosifs.

Il existe de nombreux types d'outils de forage, chacun ayant des caractéristiques techniques spécifiques qui les rendent plus ou moins adaptés à certaines conditions de forage.

Il est important de sélectionner le meilleur outil pour maximiser la productivité du forage et minimiser les coûts associés. En fin de compte, le choix de l'outil de forage doit être évalué sur la base de critères économiques, techniques et de sécurité.

CHAPITRE

III

MACHINE LEARNING

Introduction :

L'utilisation de l'intelligence artificielle (IA) dans l'industrie pétrolière offre plusieurs avantages importants, elle peut aider à améliorer l'efficacité opérationnelle en optimisant les processus de forage, d'exploration, de production et de raffinage de pétrole.

En utilisant des algorithmes avancés et des techniques de modélisation, l'IA peut aider à prévoir les conditions du fond de puits, à identifier les gisements de pétrole et de gaz et à améliorer la précision des forages.

Cela peut aider les entreprises pétrolières à réduire les coûts, à augmenter la productivité et à maximiser leur rendement. [13]

III.1- Définition de l'intelligence artificielle :

L'intelligence artificielle (IA) est une discipline de l'informatique qui vise à créer des machines capables de réaliser des tâches normalement effectuées par des êtres humains, comme la reconnaissance de la parole, la vision, la compréhension du langage naturel, la prise de décision, l'apprentissage, etc...

L'IA utilise des algorithmes et des modèles mathématiques pour traiter l'information et simuler l'intelligence humaine. L'IA englobe un large éventail de techniques et de domaines d'application, tels que l'apprentissage automatique, le traitement du langage naturel, la vision par ordinateur et la robotique. L'IA est considérée comme une technologie qui a le potentiel de transformer de nombreux aspects de la vie humaine, depuis la médecine et l'éducation jusqu'à l'industrie et la gouvernance. [14]

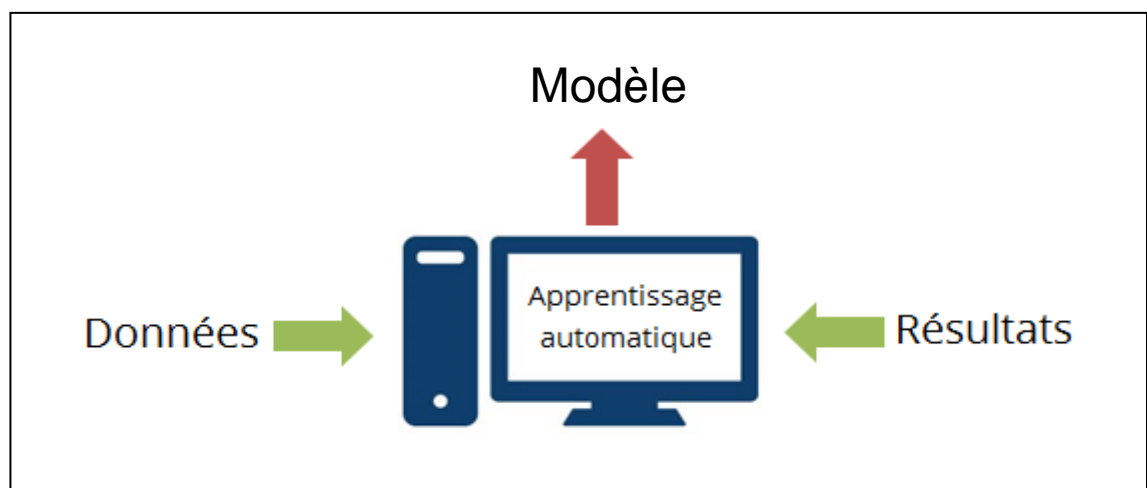


Figure.III.1 : Processus ML

III.2- Définition du machine learning «Apprentissage automatique» :

Qu'est-ce que l'apprentissage automatique ? La définition de l'apprentissage automatique selon **Wikipédia** (octobre 2022) : est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et l'implémentation de telles méthodes. On parle d'apprentissage statistique car l'apprentissage consiste à créer un modèle dont l'erreur statistique moyenne est la plus faible possible. » [15]

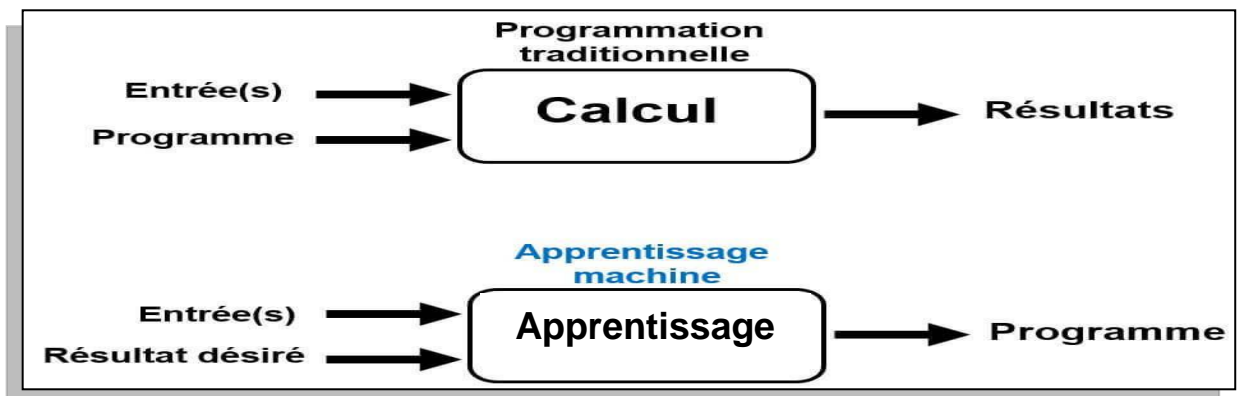


Figure.III.2 : Programme traditionnelle vs ML

Voici trois exemples de problèmes relevant de l'apprentissage automatique

Exemple 1 : Supposons que l'on dispose d'une collection d'articles de journaux. Comment identifier des groupes d'articles portant sur un même sujet ?

Exemple 2 : Supposons que l'on dispose d'un certain nombre d'images représentant des chiens, et d'autres représentants des chats. Comment classer automatiquement une nouvelle image dans une des catégories « chien » ou « chat » ?

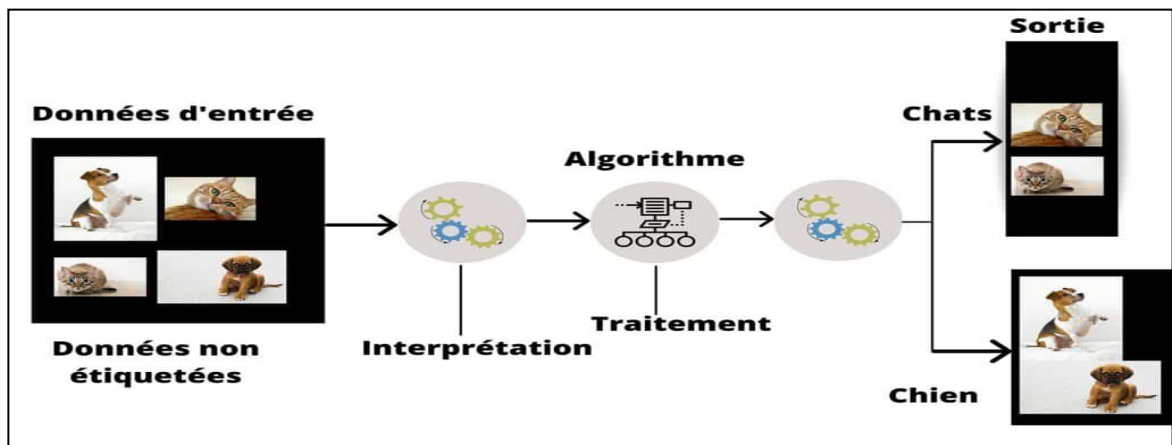


Figure.III.3 : Schéma représentatif d'un problème de Classification

Exemple 3 : Supposons que l'on dispose d'une base de données regroupant les caractéristiques de logements dans une ville : superficie, quartier, étage, prix, année de construction, nombre d'occupants, montant des frais de chauffage. Comment prédire la facture de chauffage à partir des autres caractéristiques pour un logement qui n'appartiendrait pas à cette base ?

III.3- Importance et applications de la machine learning :

Dans diverses industries notamment la santé, la finance, la fabrication et le marketing, l'apprentissage automatique est devenu une technologie nécessaire.

Il aide les entreprises à améliorer leur efficacité, leur précision et leur rapidité via des modèles mathématiques qui déduisent des informations à partir d'ensembles de données robustes.

Dans le domaine de la santé, l'apprentissage automatique est utilisé pour analyser les données médicales, détecter précocement les maladies, conception des traitements personnalisés et prédire les résultats des traitements. [16]

Dans l'industrie, l'apprentissage automatique est utilisé pour la détection des pannes et la prévention des défaillances des machines, l'optimisation des processus de production et la prévision de la demande de produits.

En marketing, l'apprentissage automatique est utilisé pour la segmentation des clients, la personnalisation des offres et l'optimisation des campagnes publicitaires.

L'apprentissage automatique est également utilisé dans les voitures, les drones et les robots autonomes, leur permettant d'apprendre de leur environnement et de prendre des décisions en temps réel.

La machine learning est utilisée dans le domaine du forage pétrolier pour améliorer l'efficacité et la sécurité des opérations.

Les algorithmes de machine learning peuvent aider à prédire la résistance de la roche, surveiller les équipements de forage, améliorer la sécurité des travailleurs et prédire les problèmes de forage. Les entreprises qui utilisent la machine learning dans le domaine du forage peuvent bénéficier d'une meilleure efficacité, rentabilité et sécurité.

En résumé, la machine learning est devenue une technologie clé pour les entreprises et les organisations qui cherchent à tirer le meilleur parti de leurs données et à automatiser des processus complexes pour améliorer l'efficacité et la précision de leurs opérations. Les applications de la machine learning sont vastes et variées, offrant des avantages tels que l'efficacité,

III.4- Les différents types d'apprentissage en machine learning :

En machine learning, il existe trois types d'apprentissage : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement.

- L'apprentissage supervisé est le type d'apprentissage le plus courant. Il implique d'avoir un ensemble de données d'entraînement étiqueté avec des résultats connus. Le modèle de machine learning est entraîné sur ces données pour apprendre à prédire les résultats pour de nouvelles données non étiquetées. Ce type d'apprentissage est utile pour la classification et la régression.

- L'apprentissage non supervisé ne nécessite pas d'ensemble de données étiqueté. Au lieu de cela, il recherche des modèles dans les données et les groupe en fonction de certaines caractéristiques. Les techniques courantes d'apprentissage non supervisé incluent le clustering, la réduction de dimension. Ce type d'apprentissage est utile pour découvrir des informations cachées dans les données, la segmentation de marché et l'analyse des réseaux sociaux.

- L'apprentissage par renforcement consiste à entraîner un modèle à prendre des décisions en interagissant avec un environnement. Le modèle apprend en recevant des récompenses pour les décisions qui mènent à un résultat souhaité et en étant pénalisé pour les décisions qui mènent à un résultat non souhaité. Ce type d'apprentissage est utile pour les jeux, la robotique et la planification de l'itinéraire.

En somme, les différents types d'apprentissage en machine learning offrent une variété d'approches pour résoudre différents types de problèmes.

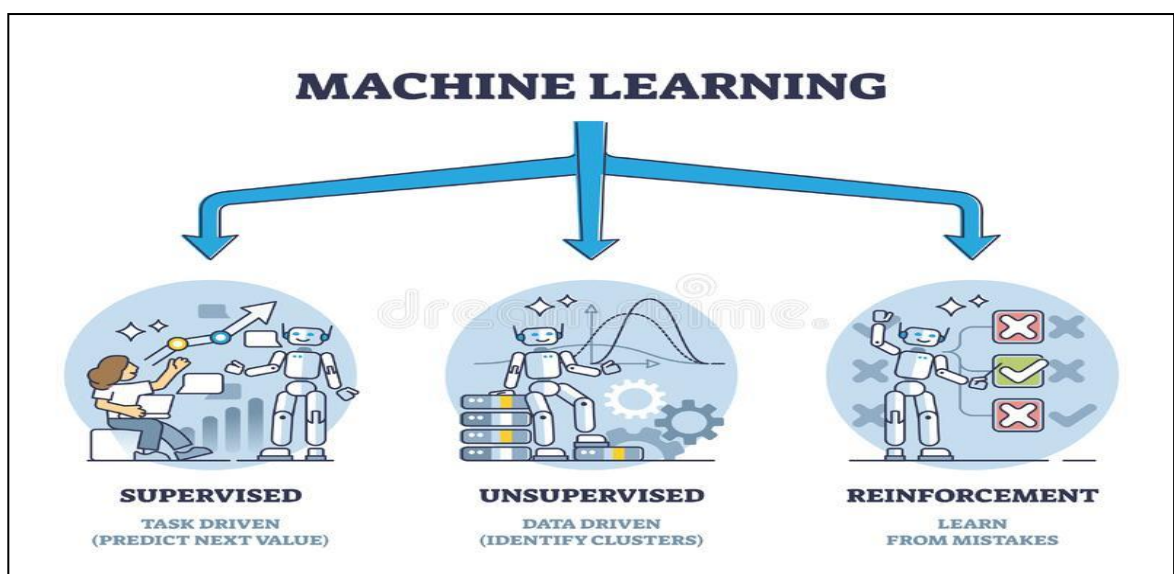


Figure.III.4 : Les différents types d'apprentissage en machine learning

III.5- Les étapes de base du processus de machine learning :

Le processus de machine learning comporte plusieurs étapes, qui peuvent varier légèrement en fonction du type de problème et du type d'apprentissage. En général, les étapes de base sont les suivantes :

- **Collecte et préparation des données** : cette étape consiste à collecter les données nécessaires pour entraîner le modèle.

Les données peuvent être obtenues à partir de différentes sources, telles que des fichiers de données, des bases de données ou des capteurs.

Il est important de nettoyer et de préparer les données pour garantir leur qualité et leur cohérence.

- **Sélection de l'algorithme** : en fonction du type de problème à résoudre, il est nécessaire de sélectionner un algorithme de machine learning approprié.

Il existe de nombreux algorithmes de machine learning, chacun ayant ses propres avantages et inconvénients en fonction du type de données et du problème à résoudre.

- **Entraînement du modèle** : à cette étape, le modèle de machine learning est entraîné à partir des données préparées.

Le modèle apprend à partir des données d'entraînement et ajuste ses paramètres pour minimiser l'erreur de prédiction.

- **Validation du modèle** : après l'entraînement, il est important de tester le modèle sur des données qui n'ont pas été utilisées pour l'entraînement.

Cette étape permet de vérifier la capacité du modèle à généraliser les connaissances acquises pendant l'entraînement.

- **Réglage du modèle** : en fonction des résultats de la validation, le modèle peut nécessiter des ajustements pour améliorer sa performance.

Il peut être nécessaire de réviser les paramètres de l'algorithme, de sélectionner de nouvelles fonctionnalités ou de modifier la stratégie d'entraînement.

- **Test du modèle** : une fois que le modèle a été ajusté, il doit être testé sur des données de test indépendantes.

Cette étape permet d'évaluer la performance du modèle de manière plus fiable.

- **Déploiement du modèle** : lorsque le modèle a été testé et validé, il est prêt à être déployé pour résoudre le problème réel.

Le modèle peut être intégré à une application ou à un système existant, ou utilisé pour fournir des prévisions ou des recommandations.

III.6- L'apprentissage supervisé (Supervised Machine Learning):

Comme vu précédemment, l'apprentissage supervisé est un type d'apprentissage automatique où un modèle est entraîné à partir d'un ensemble de données étiqueté, c'est-à-dire un ensemble de données dans lequel les résultats souhaités sont connus.

Le modèle est entraîné à prédire la sortie en fonction de l'entrée en utilisant les exemples d'apprentissage étiquetés.

L'objectif de l'apprentissage supervisé est de produire un modèle capable de généraliser les relations entre l'entrée et la sortie, afin qu'il puisse prédire de manière fiable la sortie pour de nouvelles entrées.

Les exemples d'apprentissage étiquetés peuvent être fournis à partir de différentes sources, y compris des données historiques, des expériences de laboratoire ou des simulations informatiques.

L'apprentissage supervisé est utilisé dans une variété de domaines notamment la reconnaissance d'images, la classification de texte, la prédiction de la demande, la détection de fraudes, et la détection d'anomalies, etc.... [17]

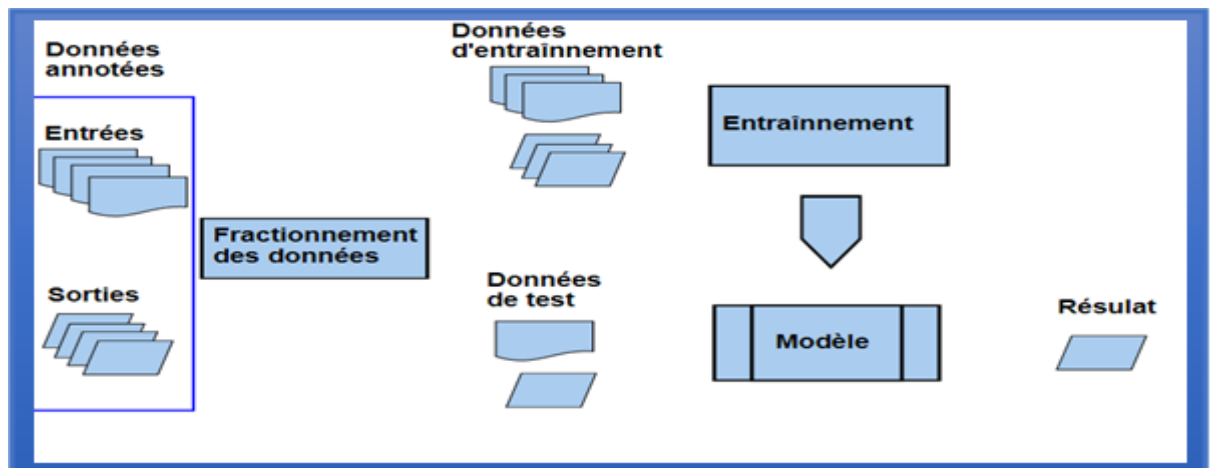


Figure.III.5 : Apprentissage supervisé. [17]

III.6.1- Les type de problèmes d'apprentissage supervisé :

Il existe deux types de problèmes d'apprentissage supervisé : la régression et la classification.

- La régression est utilisée pour prédire une valeur continue, telle que la température ou le prix d'une maison.

Dans ce cas, le modèle est entraîné à prédire une sortie numérique pour une entrée donnée.

▪ La classification, quant à elle, est utilisée pour prédire une étiquette de classe pour une entrée donnée. Les exemples courants incluent la reconnaissance d'images (par exemple, est-ce une image de chat ou de chien ?) ou la classification de spam (par exemple, est-ce un e-mail de spam ou non ?).

Dans ce cas, le modèle est entraîné à prédire une étiquette de classe pour une entrée donnée, en utilisant des exemples d'apprentissage étiquetés.

En fonction de la complexité du modèle et de la quantité de données d'entraînement disponibles, il est possible d'obtenir des résultats précis avec les deux types de problèmes.

Cependant, le choix du type de problème dépend de la nature de la tâche à accomplir et des données disponibles.

III.6.2- Processus d'apprentissage supervisé [17]:

Le processus d'apprentissage supervisé peut être divisé en plusieurs étapes :

▪ **Collecte des données** : cette étape consiste à collecter un ensemble de données étiqueté contenant des exemples d'entrées et de sorties correspondantes.

▪ **Préparation des données** : cette étape consiste à nettoyer et à prétraiter les données en vue de l'entraînement du modèle.

Elle peut inclure des étapes telles que la normalisation des données, la réduction de la dimensionnalité et la gestion des valeurs manquantes.

▪ **Sélection d'un algorithme** : le choix de l'algorithme dépend du type de problème et des caractéristiques des données.

Il existe plusieurs algorithmes pour l'apprentissage supervisé, tels que les arbres de décision, les réseaux de neurones et les machines à vecteurs de support.

▪ **Entraînement du modèle** : à cette étape, le modèle est entraîné sur l'ensemble de données étiqueté.

L'algorithme utilise les exemples d'apprentissage étiquetés pour ajuster les poids et les paramètres du modèle, afin de minimiser l'erreur de prédiction.

▪ **Évaluation du modèle** : une fois que le modèle est entraîné, il est évalué sur un ensemble de données de test distinct, l'objectif de cette étape est de mesurer la précision du modèle sur des données qu'il n'a jamais vues auparavant.

▪ **Utilisation du modèle** : une fois que le modèle est entraîné et évalué avec succès, il peut être utilisé pour prédire les sorties pour de nouvelles entrées. Ce processus peut être itératif, c'est-à-dire qu'il peut être répété plusieurs fois en modifiant les paramètres du modèle ou les caractéristiques des données, l'objectif est d'obtenir un modèle qui généralise bien les relations entre l'entrée et la sortie.

III.6.3- Types d'algorithmes d'apprentissage supervisé [18]:

Il existe plusieurs types d'algorithmes d'apprentissage supervisé utilisés dans le processus de machine learning, on peut citer :

- Régression linéaire : cette technique est utilisée pour prédire des valeurs continues, telles que les prix des maisons ou la température, elle utilise une fonction linéaire pour modéliser la relation entre l'entrée et la sortie.
- Régression logistique : cette technique est utilisée pour prédire des valeurs binaires, telles que oui ou non, vrai ou faux, elle utilise une fonction sigmoïde pour modéliser la probabilité d'une réponse positive.
- Arbres de décision : cette technique est utilisée pour prendre des décisions basées sur des règles logiques en utilisant une hiérarchie de nœuds qui représentent des décisions et des feuilles qui représentent des résultats.
- Réseaux de neurones : cette technique est utilisée pour modéliser des relations complexes entre l'entrée et la sortie en utilisant des couches de neurones interconnectés.
- Machines à vecteurs de support (SVM) : cette technique est utilisée pour séparer des données en utilisant un hyperplan optimal pour séparer les classes.
- K-plus proches voisins (KNN) : cette technique est utilisée pour prédire la classe d'une observation en se basant sur les classes des k observations les plus proches.
- Random forest : cette technique est utilisée pour créer un ensemble de plusieurs arbres de décision en utilisant un sous-ensemble aléatoire des variables à chaque division, les résultats des différents arbres sont combinés pour donner une prédiction finale.
- Gradient boosted tree : cette technique est également utilisée pour créer un ensemble de plusieurs arbres de décision, mais elle utilise une approche de renforcement en entraînant chaque arbre à corriger les erreurs de prédiction du précédent.
- Naive Bayes : cette technique est utilisée pour prédire la probabilité d'une observation appartenant à une classe en utilisant le théorème de Bayes et en supposant que les variables sont indépendantes les unes des autres.

Le choix de l'algorithme dépend des caractéristiques des données et du type de problème à résoudre.

Les différents algorithmes ont des avantages et des inconvénients, et peuvent être utilisés dans différentes situations.

III.6.4- Mesures d'évaluation pour les modèles supervisés ;

Il existe plusieurs mesures d'évaluation pour les modèles supervisés, les plus courantes sont :

- Matrice de confusion : C'est un tableau qui permet de résumer les prédictions d'un modèle en comparant les valeurs prédites avec les valeurs réelles.

Elle permet de calculer d'autres mesures d'évaluation telles que la précision, le rappel, le F-mesure et l'accuracy.

- L'accuracy : également appelée taux de classification correcte, est une mesure de performance couramment utilisée pour évaluer les modèles de classification supervisée, elle représente le nombre de prédictions correctes divisé par le nombre total de prédictions effectuées. Par exemple, si un modèle prédit correctement 90 cas sur un total de 100, l'accuracy sera de 90%. Bien que l'accuracy soit une mesure utile pour évaluer les modèles de classification, elle peut être trompeuse dans certains cas, notamment lorsque les classes sont déséquilibrées ou lorsque les erreurs de classification sont coûteuses. [19]

- Précision (Precision) : C'est le nombre de prédictions positives correctes (TP) divisé par le nombre total de prédictions positives (TP+FP), elle mesure la capacité du modèle à prédire avec précision les cas positifs.

- Rappel (Recall) : C'est le nombre de prédictions positives correctes (TP) divisé par le nombre total de cas positifs (TP+FN), elle mesure la capacité du modèle à identifier tous les cas positifs.

- F-mesure (F1-score) : C'est une moyenne harmonique de la précision et du rappel. Elle est utile pour les classes déséquilibrées, car elle donne un poids égal à la précision et au rappel.

- Aire sous la courbe ROC (AUC-ROC) : C'est une mesure qui permet d'évaluer la performance d'un modèle de classification binaire, elle représente la probabilité que le modèle classe un exemple positif choisi au hasard plus haut qu'un exemple négatif choisi au hasard.

- Log loss : C'est une mesure de la performance d'un modèle de classification binaire qui calcule la perte de log (ou entropie croisée) entre les valeurs réelles et prédites, elle est utile pour les modèles qui fournissent des probabilités de classement.

Ces mesures d'évaluation peuvent être utilisées en fonction du type de problème de classification ou de régression et des objectifs de l'analyse.

Conclusion :

L'utilisation de l'apprentissage automatique dans le processus de forage pétrolier offre un énorme potentiel pour améliorer l'efficacité, la sécurité et la rentabilité des opérations.

Cependant, l'application réussie de cette technologie nécessite une collecte de données appropriée, une expertise dans l'analyse des données et une intégration efficace dans les opérations de forage existantes.

Voici quelques applications de la machine learning dans le processus de forage pétrolier :

- Optimisation de la trajectoire de forage : la machine learning peut être utilisée pour analyser les données géologiques et de forage pour optimiser la trajectoire de forage en temps réel, en fonction des conditions du sol rencontrées.

- Détection de problèmes : la machine learning peut aider à détecter les problèmes de forage avant qu'ils ne se produisent.

Les modèles d'apprentissage automatique peuvent être entraînés à reconnaître les schémas et les signaux d'alarme indiquant un problème imminent, ce qui permet aux équipes de maintenance de prendre les mesures nécessaires pour prévenir les pannes.

- Optimisation des performances : la machine learning peut aider à optimiser les performances du processus de forage en analysant les données de forage historiques pour identifier les tendances et les opportunités d'amélioration.

- Prédiction de la production : la machine learning peut aider à prédire la production future de puits en utilisant des données historiques et en identifiant les facteurs qui ont un impact sur la production.

CHAPITRE

IV

Conception d'un Modèle ROP

PARTIE I : Description du dataset (Les données).

Introduction :

Le chapitre vise à utiliser diverses données de forage et de diagraphie pour concevoir un modèle d'intelligence artificielle capable de prédire la vitesse d'avancement d'un outil donné.

Ce modèle permettra de choisir l'outil de forage le plus adapté en prenant en compte la géologie de la phase à forer, les contraintes géomécaniques, l'utilisation potentielle d'un moteur de fond ou d'une turbine, ainsi que les paramètres mécaniques et hydrauliques à appliquer sur celui-ci.

Tout d'abord, les données nécessaires ont été collectées, calculées et compilées à partir de diverses sources telles que les données de diagraphie, de forage et de mud logging, puis regroupées dans un ensemble de données unique (dataset).

Ce dataset a été soigneusement corrigé, nettoyé, affiné et filtré afin d'éliminer les valeurs manquantes, anormales ou aberrantes.

Ensuite, des simulations ont été effectuées avec plusieurs modèles d'apprentissage automatique en incluant le maximum de variables pour éliminer celles ayant une faible corrélation.

Par la suite, de nouvelles simulations ont été réalisées avec plusieurs modèles d'apprentissage automatique en ne considérant que les variables sélectionnées précédemment, afin de choisir le modèle présentant le degré de corrélation le plus élevé.

Le modèle d'apprentissage automatique a été optimisé en sélectionnant les hyperparamètres ayant le degré de corrélation le plus élevé.

Enfin, le modèle obtenu a été utilisé pour prédire la vitesse d'avancement (ROP) dans une phase 6" pour un puits donné, en utilisant divers outils, afin de choisir l'outil le mieux adapté.

IV.1 Description du dataset (ensemble de données) utilisé :

Le dataset utilisé a été compilé à partir des données géologiques, géophysiques (diagraphies), géomécaniques, hydrauliques et des paramètres du forage de 17 puits, situés dans la partie Upside nord de Hassi Messaoud et qui ont été réalisés entre 2017 et 2022.

. Au départ, il comptait plus de 1211 enregistrements, totalisant 39963 données, avant la phase de prétraitement. Cependant, il est important de souligner que toutes les variables contenues dans notre dataset n'ont pas été directement utilisées dans la modélisation.

Plusieurs variables ont été sélectionnées et filtrées pour des raisons de pertinence. Les données se répartissent en deux catégories distinctes : les données enregistrées ou mesurées et les données calculées.

IV.1.1- Données enregistrées :

Les données enregistrées sont collectées à partir de mesures directes ou indirectes à l'aide d'instruments de mesure et de capteurs divers.

En raison de leur source d'origine, elles sont divisées en deux catégories distinctes :

IV.1.1.1- Données de mud logging :

Les données de mud logging utilisées pour notre étude, sont issues des rapports de fin de puits, des bit records des outils utilisés ainsi que des paramètres de forage enregistrés pendant les opérations de forage par différents capteurs de la cabine du mud Logging. Elles sont notamment :

- Type de l'outil de forage (BIT).
- Poids sur l'outil (WOB).
- Vitesse de rotation (RPM).
- Torque de rotation (TORQUE).
- Début de circulations (FLOW RAT).
- Pression de refoulement (SPP).
- Taux de pénétration (ROP).
- Densité de la boue (MW).
- Profondeur de forage (DEPTH).
- Diamètre de forage (Diamètre).
- Total flow area (TFA).

a) - Type d'outil :

Le type d'outil est une autre variable qui peut influencer le taux de pénétration (ROP), les outils utilisés dans les puits choisis sont de type **imprégné** :

- Outil 6" type *DI820W*
- Outil 6" type *GT64DDHE*
- Outil 6" type *HHD372G8*
- Outil 6" type *TF610DPS*.
- Outil 6" type *K505BCTPX*
- Outil 6" type *K505BEPX*
- Outil 6" type *K505BPX*

b) - Total flow area (TFA):

TFA (Total Flow Area) est une variable qui peut affecter sur le taux de pénétration (ROP) lors de l'opération de forage. C'est une mesure de la surface totale des duses de l'outil de forage qui permettent aux fluides de circuler à travers l'outil.

c) - Poids sur l'outil ou WOB (Weight On Bit) :

C'est le poids appliqué par la BHA sur le trépan. En général, le WOB devrait être appliqué avant une vitesse de rotation excessive afin que la structure de coupe maintienne une profondeur de coupe significative pour stabiliser le trépan et minimiser les vibrations axiales. Généralement, le poids sur l'outil varie entre 1 et 13 tonnes dans la phase 6".

d) - Vitesse de rotation ou RPM (révolution per minute) :

La vitesse de rotation est une autre variable qui peut influencer le taux de pénétration (ROP), une vitesse de rotation totale de l'outil est égale à la vitesse de rotation de surface plus la vitesse de rotation du moteur/turbine de fond de trou.

Elle est variée en fonction de l'outil utilisé ainsi que des formations géologiques traversées afin de réduire les vibrations latérales et éviter de déclencher la résonance de la garniture de forage.

- Certaines vitesses de rotation peuvent déclencher une résonance de la colonne de forage et doivent être évitées. Cela peut être fait en déterminant les RPM critiques

e) - Torque de rotation (TORQUE) :

Le couple rotatif est un indicateur de ce qui se passe au niveau de l'outil de forage, dans les formations tendres, le couple peut indiquer que l'outil est en contact avec le fond avant le WOB, si des mesures de couple de fond de trou sont disponibles, elles peuvent être utilisées en combinaison avec des mesures de surface pour une plus grande précision.

f) - Débit de circulation (FLOW RATE) :

En général, un débit élevé fournit un meilleur nettoyage du trou de forage qu'un débit faible, car il élimine les déblais plus efficacement en raison de la vitesse élevée et de la puissance spécifique élevée (HSI) qui en résultent.

g) - Pression de Refoulement (SPP) :

Elle représente la somme des pertes de charges générées par l'écoulement de la boue à travers la garniture de forage, passant par l'outil et remontant par l'espace annulaire pression.

Elle est mesurée à la surface à travers le Standpipe et fournit des diverses informations sur les conditions du puits, la résistance et la dureté de la formation géologique traversée par l'outil de forage, ainsi que sur la qualité de la boue de forage utilisée.

h) - Vitesse d'avancement ou ROP (Rate Of Penetration):

La ROP (Rate of Penetration), ou vitesse d'avancement, représente la vitesse à laquelle l'outil de forage progresse dans la roche, elle est mesurée en mètres par heure et est influencée par différents facteurs.

Dans cette étude, la ROP est la variable cible à modéliser en utilisant d'autres variables, car elle est le paramètre le plus important pour sélectionner l'outil adéquat afin de forer une phase donnée.

i) - Densité de la boue :

La densité d'une boue de forage, également appelée densité de la boue, fait référence à la masse de la boue par unité de volume.

Elle est exprimée en unités de poids par unité de volume, telles que les livres par gallon (lb/gal) ou les kilogrammes par litre (kg/L).

C'est un paramètre crucial dans les opérations de forage. Elle est contrôlée et ajustée en fonction des besoins spécifiques du puits et des formations géologiques traversées. Une densité de boue adéquate est nécessaire pour contrôler la pression du fluide de forage, maintenir la stabilité du puits, prévenir l'éruption de gaz ou de fluides de formation, et transporter les déblais de forage à la surface.

j) - Profondeur de forage (DEPTH) :

La profondeur de forage est un paramètre important qui peut avoir une influence sur le taux de pénétration (ROP) lors de l'opération de forage.

Généralement, plus la profondeur de forage est importante, plus la résistance à la pénétration est grande, ce qui peut ralentir le ROP.

k) - Diamètre de forage (Diamètre) :

Le diamètre de forage peut avoir un impact sur le taux de pénétration (ROP). En général, plus le diamètre de forage est grand, plus le ROP sera faible en raison de la quantité accrue de matériau à enlever.

IV.1.1.2- Données logging (diagraphie) :

Ce sont des mesures obtenues à partir d'outils de diagraphie pétrophysique, tels que les rayons gamma, les temps de transit de compression et de cisaillement et la densité apparente.

Elles peuvent être utilisées pour optimiser le processus de forage et améliorer le taux de pénétration (ROP).

Par exemple, connaître la lithologie de la formation peut aider à sélectionner les paramètres de forage appropriés, tels que le poids sur l'outil (WOB) et les propriétés du fluide de forage, ce qui peut améliorer le ROP.

De même, les informations sur la porosité et la densité de la formation peuvent être utilisées pour optimiser les propriétés de la boue de forage et réduire le risque de coincement de la garniture de forage.

a)- Gamma Ray (GR) :

La relation entre le rayonnement gamma et le taux de pénétration (ROP) est indirecte mais importante. La diagraphie gamma peut aider à identifier les zones argileuses ou encore des zones de dureté variables, qui peuvent avoir un impact sur le ROP.

b) - Shear Transit Time (DTS):

C'est la mesure de temps de transit pour une onde acoustique de cisaillement, et peut être utilisée pour calculer les paramètres géomécaniques tel que le module de Young ou la résistance à compression de la roche.

c) - Compressional Transit Time (DTC):

C'est la mesure de temps de transit pour une onde acoustique de compression, et combinée à la DTS, elle est utilisée pour calculer les paramètres géomécaniques tel que le module de Young ou la résistance à compression de la roche.

Le temps de transit de compression peut aider à identifier la lithologie et la porosité de la formation, ce qui peut avoir un impact sur le ROP.

Le forage à travers des formations poreuses ou fracturées peut causer de l'instabilité de ROP, en suite le forage à travers des formations plus denses peut nécessiter un WOB plus élevé et ralentir le ROP.

d)- Bulk Density :

La relation entre la densité apparente et le ROP est indirecte mais importante. La densité apparente peut être utilisée pour identifier les zones de formation denses.

En utilisant les mesures de densité apparente en conjonction avec d'autres paramètres de forage et des diagraphies pétrophysiques, les opérateurs peuvent optimiser le processus de forage et améliorer le ROP.

IV.1.2 Les données calculées :

Les données calculées peuvent être divisées en trois catégories principales :

- 1- Donnée hydraulique
- 2- Données géomécaniques.
- 3- Le modèle de l'Energie Spécifique (Es)

IV.1.2.1- Données hydrauliques :**a)- Chute de pression (Pressure Drop PD) :**

La chute de pression ou la pression différentielle à travers l'outil est définie comme la différence entre la pression de la boue sortant des duses et la pression de la boue à l'intérieur de la colonne de forage immédiatement avant d'entrer dans l'outil.

Si la chute de pression de l'outil est extrêmement élevée pour un débit et une densité de boue donnés, le fluide sortant des duses est en conséquence élevée.

En revanche, une chute de pression plus faible, dans les mêmes conditions de débit et de densité de boue, entraîne la sortie de fluide des duses avec une vitesse plus faible.

La chute de pression dépend du débit, de la densité de boue et de la surface d'ouverture du duse (TFA). Elle est calculée par la formule ci-dessous :

$$PD = \frac{\text{Flow Rate}^2 \times \text{Mud Weight}}{10856 \times \text{TFA}^2} \dots\dots\dots (\text{IV.1}).$$

PD : Pressure drop (psi).

- Débit (Flow Rate: gal/min).
- Densité (Mud weight: lb/gal).
- TFA (sq.in.).
- 10,856 facteur de conversation.

b)- Puissance hydraulique (Hydraulique horsepower HHP) :

C'est la puissance hydraulique (HHP) totale, développée au niveau du l'outil, elle dépend du débit (gal/min) et de la chute de pression (PD) selon la formule en unités anglo-saxonnes :

$$HHP = \frac{PD \text{ (psi)} \times \text{Flow rate (gal/min)}}{1714} \dots\dots\dots (\text{IV.2}).$$

c) - Hydraulic horsepower/square inch (Hydraulic horsepower HSI):

L'hydraulique horsepower/square inch (HSI) fournit une mesure de la puissance hydraulique consommée au niveau du l'outil par section de trou, et dépend du débit et de la chute de pression du l'outil, ainsi que du diamètre du trou.

Elle augmente lorsque le débit augmente, mais est à son maximum lorsque la chute de pression à travers le bit est de 65 % de la pression de refoulement.

Elle est calculée par la formule suivante :

$$HSI = \frac{PD \text{ (psi)} \times \text{Flowrate (gal/min)}}{1714 \times \text{Hole Area}} \dots\dots\dots (\text{IV.3}).$$

- Débit (Flow Rate: gal/min).
- Bit pressure drop (psi)
- Hole area (sq.in) = $\pi/4 \times (\text{hole diameter})^2$
- 1714 facteur de conversation

d)- Vitesse du jet (Jet velocity JV) :

C'est la vitesse de jet de la boue à la sortie des duses, elle dépend du débit et de la TFA, si le débit est extrêmement élevé, pour une densité de boue donnée, le fluide sortant des duses à une vitesse élevée, un débit plus faible, en revanche, dans les mêmes conditions de densité de boue, donne lieu à un fluide sortant des duses avec une vitesse plus faible.

Si le TFA est élevée pour un débit et une densité de boue donnés, le fluide sortant des buses a une faible vitesse, l'inverse est également vrai.

$$JV = \frac{0.32086 \times \text{Flow rate(gal/min)}}{\text{TFA}} \dots\dots\dots (\text{IV.4}).$$

- Débit (Flow Rate: gal/min).
- TFA (sq.in).
- 0.32086 est facteur de conversation.

e)- Force d'impact du jet (Jet impact force JIF) :

La force d'impact du jet (JIF) est la force exercée sur le fond du trou par le fluide sortant des duses lorsque l'outil est enfoncé au fond.

Elle dépend de la vitesse du jet, du poids du fluide et du débit, la JIF est maximisée lorsque la chute de pression à travers l'outil représente 49% de la pression de la colonne montante.

$$JIF = \frac{\text{Mud Weight} \times \text{Flow rate(gal/min)} \times JV}{1932} \dots\dots\dots (\text{IV.5}).$$

- Débit (Flow Rate: gal/min).
- Densité (Mud weight: lb/gal).
- TFA (sq.in.)
- 1932 est facteur de conversation

IV.1.2.2- Données Géo-mécaniques :

Les données géo-mécaniques sont également utilisées pour optimiser le taux d'avancement de l'outil de forage d'un puits pétrolier.

Les données géo-mécaniques de cette étude sont :

- Module de Young.
- Résistance à la compression non confinée UCS.
- Résistance à la compression triaxiale CCS.

a)- Module de Young :

En supposant que l'élasticité est isotrope, le module de Young et le coefficient de Poisson sont les principaux facteurs permettant de déterminer les propriétés élastiques de la roche (dynamiques et statiques), les propriétés dynamiques étant calculées en fonction de la densité apparente et du log sonique (compression *DT_c* et de cisaillement *DT_s*).

Le module d'Young dynamique *E_{dyn}* et le coefficient de Poisson dynamique *ν_{dyn}* sont donnés par les équations (IV.6) et (IV.7) respectivement ci-dessous (Gassmann, 1951) :

$$E_{dyn} = 10^6 \frac{\rho_b (3DT_s^2 - 4DT_c^2)}{DT_s^2 - DT_c^2} \dots\dots\dots(IV.6).$$

$$\nu_{dyn} = \frac{0.5(DT_s/DT_c)^2 - 1}{(DT_s/DT_c)^2 - 1} \dots\dots\dots(IV.7).$$

Comme les données de la diagraphie sonique ne fournissent que des propriétés dynamiques de la roche (voir Figure. IV.1), les propriétés statiques peuvent être obtenues à partir d'un essai mécanique sur carotte du puits offset localisé dans la région upside nord à Hassi Messaoud en utilisant la régression linéaire. Le module de Young statique est calculé à partir de l'équation suivante (IV.8) où R² = 0.9073. . [20]

$$E_{sta} = 7.7012 \times e^{0.0317 \times E_{dyn}} \dots\dots\dots(IV.8).$$

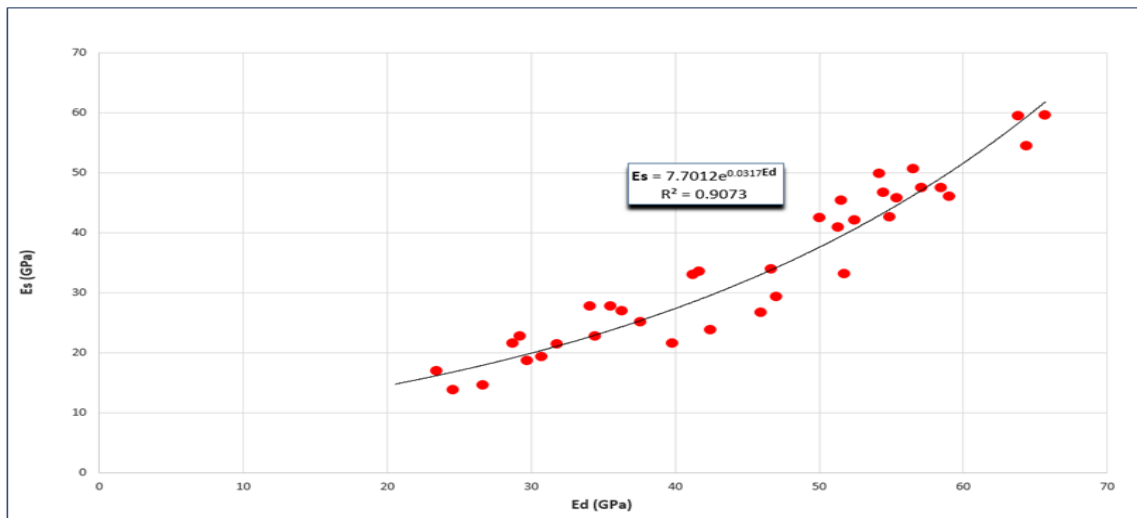


Figure.IV.1 : Propriétés dynamiques de la roche. [20]

b)- La résistance à la compression non confinée (Unconfined compressive strength UCS) :

C'est la contrainte de compression axiale maximale qu'un échantillon de matériau cylindrique droit peut tolérer dans des conditions non confinées.

Dans cette étude, la résistance UCS peut être obtenue à partir des essais de carottage mécanique sur un puits offset localisé dans la région upside nord à Hassi Messaoud en utilisant la régression linéaire (voir Figure. IV.2), ce qui permet d'établir un profil continu du NGC à partir des équations suivantes (IV.9) avec $R^2 = 0.9713$

$$UCS = 2.1595 \times E_{sta} + 6.3235 \dots\dots\dots(IV.9).$$

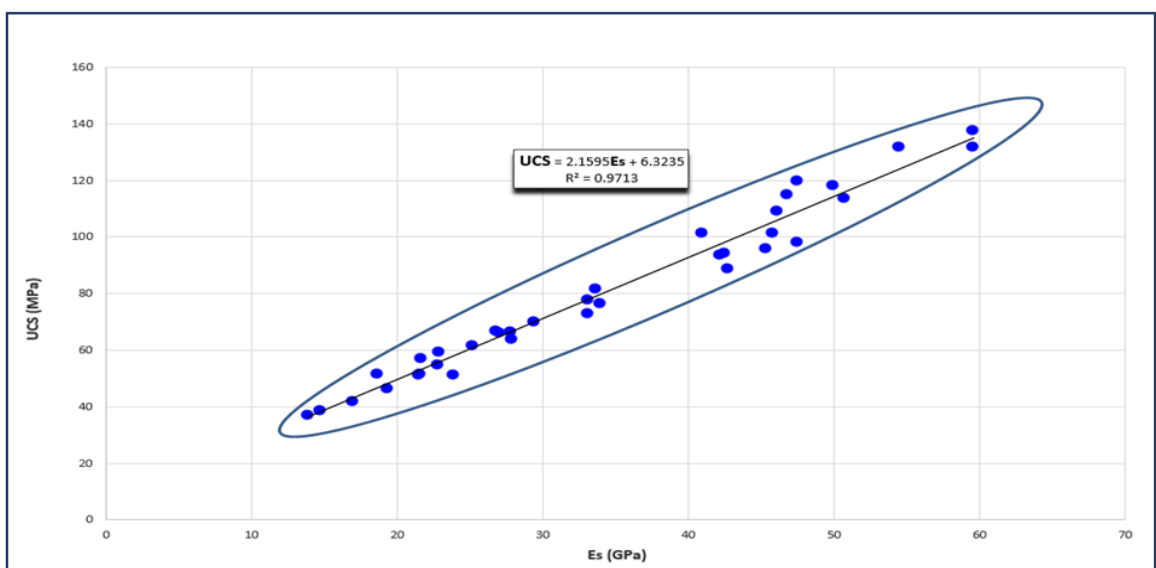


Figure.IV.2 : Essais de carottage mécanique upside nord à Hassi Messaoud. [20]

d)- La résistance à la compression triaxiale (Confined Compressive Strength CCS) :

La résistance à la compression triaxiale (résistance à la compression confinée - CCS) est une mesure de la résistance d'un matériau à la compression lorsqu'il est soumis à une contrainte triaxiale (c'est-à-dire lorsqu'il est contraint dans toutes les directions).

Dans le cas du forage de puits, la résistance à la compression triaxiale est une caractéristique importante de la formation rocheuse traversée. Elle représente la capacité d'une roche à résister à la compression lorsqu'elle est soumise à de fortes contraintes.

VI.1.2.3- Mécanique spécifique d'énergie (MSE) [21] :

L'énergie spécifique est le travail effectué pour arracher un volume unitaire de la roche, ce modèle d'Energie Mécanique Spécifique a été utilisé effectivement dans les laboratoires pour évaluer l'efficacité d'un outil de forage.

Il est évident que pour creuser un volume donné, il faut assurer un minimum de quantité d'énergie, cette quantité dépendra du type d'outil utilisé et de la nature de la roche.

Dans notre étude il est obligatoire de calculer la MSE (mécanique spécifique d'énergie), elle peut aussi être utilisée pour optimiser les données mécaniques en real time, elle est calculée par la formule suivante :

$$MSE = \frac{WOB}{A} + \frac{2 \times 3.14 \times RPM \times TORQUE}{A \times ROP} \dots\dots\dots(IV.10)$$

- WOB : Poids sur l'outil.
- RPM : Vitesse de rotation du l'outil.
- Torque : Couple appliqué au l'outil.
- ROP : Vitesse de pénétration du l'outil.
- A : Section forée par l'outil.

Cela permet d'approximer les valeurs réelles de WOB, RPM, Torque et ROP, en utilisant la MSE comme filtre, il est alors possible d'éliminer les données anormales et aberrantes et de ce fait améliorer la qualité du dataset, en supprimant les valeurs de la MSE anormalement élevée ou faible grandes valeurs. En utilisant ces données mécaniques filtrées au lieu des données globales pour obtenir un modèle de ROP pratique.

PARTIE II : Traitement des données par l'intelligence artificielle et élaboration d'un modèle.

Introduction :

La seconde partie de cette étude a pour but de prédire la vitesse d'avancement (ROP) dans la phase 6" qui caractérisent le champ de Hassi Messaoud.

Et cela, entraînent plusieurs algorithmes d'intelligence artificielle à partir d'une base de données comporte la totalité des paramètres de forage de la phase 6", effectuée lors du forage de 17 puits forés à Hassi Messaoud.

Afin d'implémenter ces algorithmes, on a utilisé la plateforme d'intelligence artificielle RAPIDMINER, ce logiciel a été utilisé pour traiter les données contenues dans notre base de données, d'exécuter plusieurs algorithmes de Machine Learning, de choisir les plus adéquats afin de l'optimiser et au final d'élaborer un simulateur capable de prédire la vitesse d'avancement.

Dans notre étude, on a utilisé l'apprentissage supervisé pour la prédiction du ROP.

VI.1- Rapid Miner [22] :

En tant que plateforme IA, RappidMiner propose un large panel de services et d'outils dédiés à la gestion de projets data science et machine learning. De l'analyse statistique au data mining, sans oublier la modélisation prédictive, elle dispose de fonctionnalités exhaustives pour superviser l'ensemble des étapes de développement.

Notre travail avec RappidMiner se résume comme suit

- Etape 01 : Prétraitement et préparation des données,
- Etape 02 : Elaboration du modèle :
- Etape 03 : Optimisation du modèle choisi,
- Etape 04 : Simulation à l'aide du modèle optimisé obtenu.

VI.1.1- Prétraitement et préparation des données :

- a)- Comprendre les données
- b)- Prétraiter les données
- c)- Diviser les données
- d)- Vérifier la distribution des données
- e)- Évaluer la qualité des données
- f)- Importer les données dans logiciel RappidMiner
- g)- Sélectionner la ROP comme variable cible output
- h)- Sélectionner les variables input (tous les paramètres qui influencent sur le (ROP))

a)- Compréhension des données :

La première étape consiste à comprendre les données compilées et qui seront utilisées pour entraîner le modèle, cela implique de vérifier les descriptions des données, les formats de fichiers, les champs de données et les étiquettes de classe.

b)- Prétraitement des données :

Après avoir exploré le dataset, la prochaine étape consiste à effectuer un prétraitement des données pour les préparer à l'entraînement du modèle, il est également important de vérifier la qualité des données en examinant s'il y a des valeurs manquantes, des valeurs aberrantes ou des erreurs de saisie, cela peut aussi inclure des étapes telles que la normalisation, la transformation de la distribution des données.

c)- Division des données :

La troisième étape consiste à diviser le dataset en deux ensembles distincts, un ensemble de données d'entraînement ou trainset utilisé pour entraîner le modèle, et un autre pour tester la fiabilité de notre modèle et évaluer sa performance ou aussi appelé testset.

d)- Vérification de la distribution des données :

Il est important de vérifier que les ensembles de données d'entraînement, de validation et de test ont des distributions de données similaires pour éviter tout biais potentiel, la distribution utilisée pour le modèle est celle de la loi normale.

e)- Évaluation de la qualité des données :

La dernière étape consiste à évaluer la qualité des données en examinant la performance du modèle sur l'ensemble de test, si le modèle n'a pas une performance satisfaisante, cela peut indiquer que les données utilisées pour l'entraînement du modèle sont de qualité insuffisante ou qu'il y a un biais dans les données.

f)- Importé les données dans logiciel RappidMiner :

Toutes les données portées sur un fichier Excel, ont été intégrées au programme

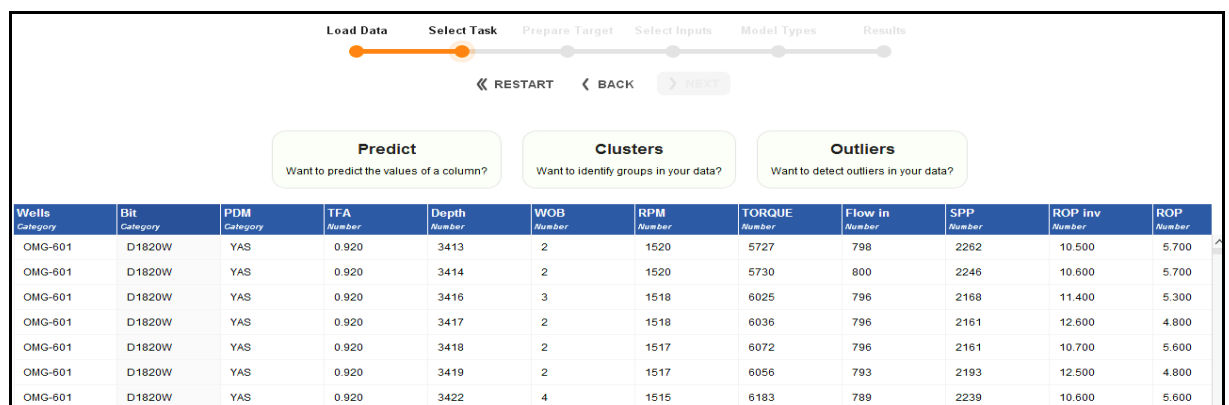


Figure. IV.3 : Importation des données dans logiciel RappidMiner.

g)-Sélectionner output :

Dans le cas d'un modèle du ROP En sélectionne la colonne de ROP

Wells category	Bit Category	PDM Category	TFA Number	Depth Number	WOB Number	RPM Number	TORQUE Number	Flow in Number	SPP Number	ROP inv Number	ROP Number
OMG-601	D1820W	YAS	0.920	3413	2	1520	5727	798	2262	10.500	5.700
OMG-601	D1820W	YAS	0.920	3414	2	1520	5730	800	2246	10.600	5.700
OMG-601	D1820W	YAS	0.920	3416	3	1518	6025	796	2168	11.400	5.300
OMG-601	D1820W	YAS	0.920	3417	2	1518	6036	796	2161	12.600	4.800
OMG-601	D1820W	YAS	0.920	3418	2	1517	6072	796	2161	10.700	5.600

Figure.IV.4 : Sélectionner les outputs dans logiciel RappidMiner.

h)-Sélection des variables entrantes ou inputs :

Le programme, peut déterminer les paramètres les plus influents ou les moins influents sur notre prédiction de la ROP en assignant des couleurs comme suite :

La couleur rouge (n'affecte pas), la couleur jaune (la probabilité affecte), la couleur verte (influencer). Ces couleurs sont attribuées, en fonction de :

- *La corrélation* entre les inputs car lorsque les variables d'entrée sont faiblement corrélées, cela indique qu'elles sont relativement indépendantes les unes des autres. Cela permet de mieux isoler l'effet de chaque variable sur la variable de sortie (output) et facilite l'interprétation des résultats.
- *La stabilité* élevée d'une variable signifie que ses valeurs changent peu, cela rend cette variable peu pertinente.
- Une variable ayant une valeur d'IDness élevée signifie que ces valeurs fluctuent de manière significative ou imprévisible à cause de cela le modèle peut ne pas fournir des prédictions cohérentes et fiables.
- *Les valeurs manquantes*, car cela peut introduire un biais dans les résultats de l'analyse, car les observations incomplètes peuvent ne pas représenter de manière précise la population ou l'échantillon d'origine.
- *Les variables contenant des valeurs textuelles* ou nominales ne sont pas souhaitées car elle peut être difficile à modéliser surtout pour des modèles de régression.

Selected	Status ↑	Quality	Name	Correlation	ID-ness	Stability	Missing	Text-ness
<input type="checkbox"/>	●		section	0.00%	?	100.00%	0.00%	0.00%
<input type="checkbox"/>	●		Wells	0.00%	1.47%	10.01%	0.00%	6.66%
<input checked="" type="checkbox"/>	●		Bit	4.04%	0.60%	39.26%	0.00%	3.67%
<input checked="" type="checkbox"/>	●		PDM	0.14%	0.17%	62.21%	0.00%	1.22%

Figure.IV.5 : Sélectionner les inputs dans logiciel RapidMiner.

Les variables sélectionnées sont :

- Type de l'outil (BIT) - JV
- WOB - JIF
- Torque - TFA
- RPM - HSI
- Densité - YME- STA
- UCS - PD
- CCS - PDM

VI.1.2- Elaboration du modèle :

Après le traitement des données et avant de choisir un ou plusieurs modèles il est nécessaire de passer par l'étape de cross validation.

En effet, comme mentionné précédemment, afin de modéliser un algorithme de machine learning il est nécessaire de fractionner le data set en deux parties.

La première partie est le train set, elle sera utilisée pour entrainer notre modèle afin qu'il puisse apprendre.

La seconde partie est le test set, elle sera par la suite utilisée pour tester la performance du modèle à prédire la valeur d'occurrence d'une vitesse d'avancement (ROP) d'outils.

Pour cela on va utiliser la cross validation.

a)- L'opérateur Cross Validation [23]:

Validation croisée est une technique dans laquelle on entraine le modèle en utilisant le train set, puis on l'évalue en utilisant le test set, mais cela en plusieurs itérations, pour chaque itération, on choisit un test set indépendant du train set ce qui permet d'éviter un sur apprentissage ou overfitting.

Cela passe par les étapes suivantes :

- Réserver une partie de l'échantillon de données pour la validation.
- Utiliser le reste de l'ensemble de données pour entrainer le modèle.
- Tester le modèle en utilisant la partie réservée de l'ensemble de données.
- Le processus est répété pour les n itérations choisies et à la fin une moyenne des performances du modèle sera prise.

L'un des paramètres le plus important du cross validation est le choix du nombre d'itération ainsi que la méthode d'échantillonnage du train et du test set

b)- Implémentation du modèle :

Sur Rapidminer la fenêtre du cross validation est partagée en deux zones. La zone de gauche va contenir le modèle de machine learning choisis et il aura pour input le train set, ce dernier aura pour fonction l'entraînement du modèle

La fenêtre de droite va contenir l'opérateur "Apply model" qui aura pour fonction, le test du modèle avec les données du test set ainsi que l'opérateur "performance" qui aura pour fonction d'évaluer les performances du modèle

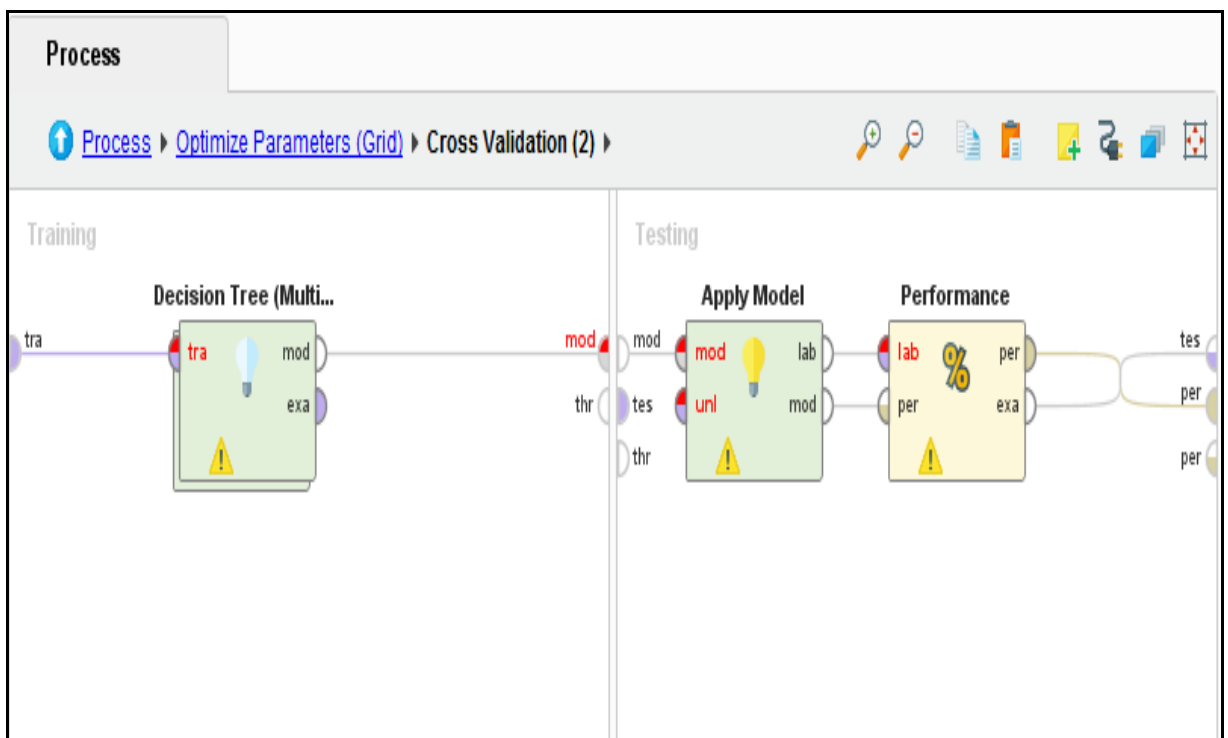


Figure.IV.6 : Schéma descriptif d'une Cross validation contenant l'algorithme 'decision tree'.

c)- La performance du modèle :

La performance d'un algorithme de régression en machine learning est directement liée à sa capacité à prédire les résultats.

Lorsque l'on compare les résultats d'un algorithme à la réalité, on utilise une matrice appelée la matrice de confusion.

Cette matrice met en évidence les prédictions correctes et incorrectes, fournissant ainsi des indications sur les types d'erreurs commises.

Diverses métriques permettent d'évaluer la performance d'un algorithme de régression en machine learning.

Dans notre cas, le coefficient de détermination (R^2) est utilisé pour mesurer la proportion de la variance expliquée par le modèle par rapport à la variance totale des données.

Cette métrique quantifie l'efficacité du modèle dans l'explication des variations des données observées.

d)- Le choix du modèle :

L'objectif de cette modélisation est d'effectuer une régression en utilisant les variables d'entrée fournies.

Cette régression vise à prédire la vitesse d'avancement, plutôt que de classer une opération, la vitesse d'avancement peut prendre différentes valeurs numériques, cette approche de régression fait partie de la branche du machine learning supervisé.

Dans cette étude, six algorithmes différents ont été utilisés : le modèle de régression linéaire généralisée, le gradient boosted tree, les réseaux de neurones, le support vector machine, les arbres de décision et le random forest, bien que les deux derniers soient normalement utilisés pour la classification, ils ont été adaptés pour être utilisés dans le contexte de la régression.

L'objectif est de modéliser la vitesse d'avancement et de déterminer quel algorithme offre les meilleures performances c'est-à-dire le coefficient de détermination le plus élevé.

Afin de choisir le modèle adéquat sur Rapidminer on utilisera l'opérateur "Optimise process grid". Cet opérateur permet d'exécuter plusieurs autres opérateurs et sous-process, de varier leurs paramètres et de les classer en fonction de leur performance.

Il va contenir l'opérateur "Sub process" qui sera divisé en six parties et chaque partie va contenir un des six algorithmes.

On remarque à partir du tableau VI.1 que l'algorithme le plus performant est la Gradient Boosted Trees avec un coefficient de détermination $R^2=0,66$. En effet, cet algorithme bien que complexe à implémenter, il peut se révéler puissant.

L'autre avantage de cet algorithme est le faible temps nécessaire à son exécution, ce qui permettra par la suite, lors de l'optimisation de gagner beaucoup de temps, d'exécuter plusieurs variantes de ce dernier, essayer plusieurs combinaisons de paramètres et par conséquent obtenir une performance élevée.

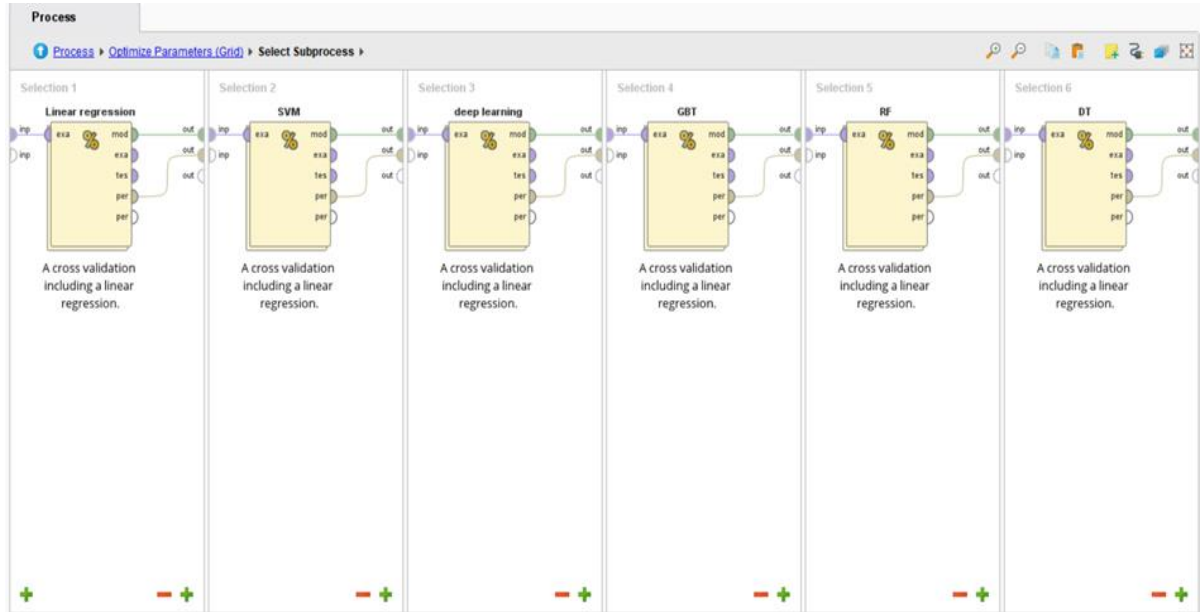


Figure.IV.7 : Schéma descriptif d'implémentation des six algorithmes

Tableau IV.1 : Résultats des performances des six algorithmes.

Algorithme	RMSE	Coef. de corrélation R	Coef. De détermination R ²
GBT	1,66	0,81	0,66
RF	1,59	0,81	0,65
DT	1,61	0,79	0,63
Réseaux de neurones	3,71	0,65	0,44
GLR	12,21	0,54	0,31
SVM	4,55	0,53	0,30

VI.1.3- Optimisation du modèle :

Comme pour trouver l'algorithme le plus performant, on va utiliser le même opérateur à savoir "Optimize process grid" afin de faire varier les paramètres du cross validation et celui de l'algorithme afin de trouver les paramètres qui donnent les meilleures performances, en d'autres mots un coefficient de détermination R² le plus élevé.

Mais dans ce cas, il ne va contenir que le Gradient Boosted Trees qui sera lui-même contenu dans une cross validation.

L'opérateur "Optimize process grid" permet d'exécuter plusieurs itérations de l'algorithme, ces itérations dépendent du nombre des paramètres à varier et du nombre de valeurs qu'un paramètre peut avoir.

De ce fait, plus il y aura de paramètres à faire varier ou de valeur plus aura de combinaisons différentes et plus le temps d'optimisation sera plus long.

Dans notre cas on a varié les paramètres suivants :

- Le nombre d'itération du cross validation, soit de 5 à 10 itérations.
- La méthode d'échantillonnage qui peut être linear, shuffled.
- Le nombre d'arbres de 30 jusqu'à 150 avec un pas de 10.
- La profondeur maximale qu'on fera varier de 5 jusqu'à 10 avec un de 1.

Le model obtenu a un coefficient de détermination $R^2= 0.73$ pour les hyperparamètres suivants :

- Itération du cross validation =**15**.
- Méthode d'échantillonnage =**shuffled**.
- Nombre d'arbres =**110**.
- Profondeur maximale =**7**.

Tableau IV.2 : Paramètres utilisé pour l'optimisation du modèle

Itérat.	Itérations CV	Echantill.	Nbre d'arbres	Profondeur maximale	RMSE	Coef. de corrélation R	Coef. De détermination R ²															
440	15	shuffled	110	7	1,61	0,85	0,73															
922	13	shuffled	90	10	1,58	0,85	0,73															
855	12	shuffled	140	9	1,60	0,85	0,73															
990	15	shuffled	120	10	1,59	0,85	0,73															
900	13	shuffled	80	10	1,60	0,85	0,73															
873	8	shuffled	150	9	1,60	0,85	0,73															
766	11	shuffled	100	9	1,60	0,85	0,73															
789	12	shuffled	110	9	1,60	0,85 </tr <tr> <td>783</td> <td>6</td> <td>shuffled</td> <td>110</td> <td>9</td> <td>1,58</td> <td>0,85</td> <td>0,73</td> </tr> <tr> <td>814</td> <td>15</td> <td>shuffled</td> <td>120</td> <td>9</td> <td>1,60</td> <td>0,85</td> <td>0,73</td> </tr>	783	6	shuffled	110	9	1,58	0,85	0,73	814	15	shuffled	120	9	1,60	0,85	0,73
783	6	shuffled	110	9	1,58	0,85	0,73															
814	15	shuffled	120	9	1,60	0,85	0,73															

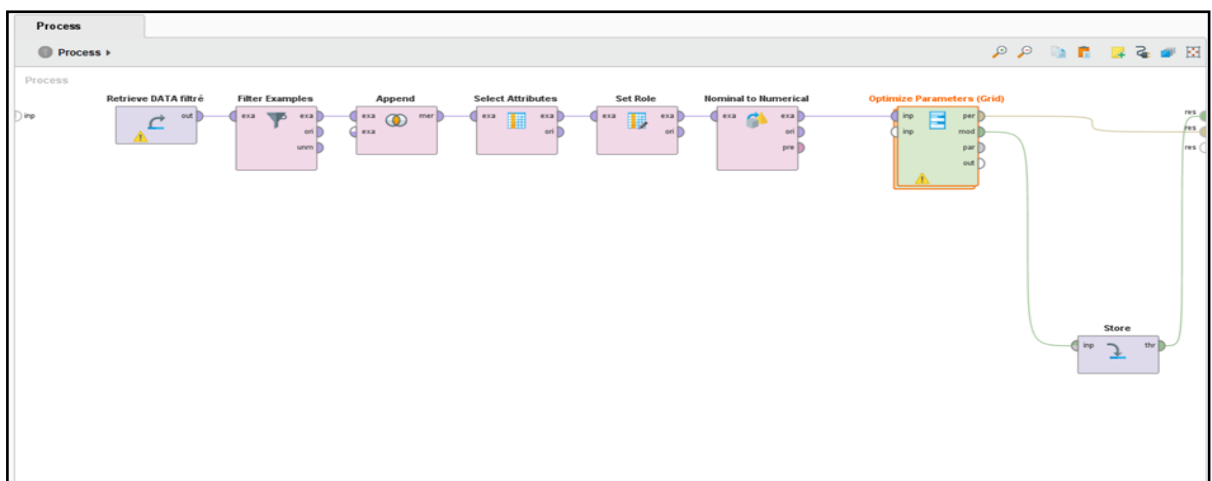


Figure IV.8 : Schéma descriptif du process d'optimisation de Gradient Boosted Trees

IV.1.4- La validation du modèle :

Après le choix de l'algorithme adéquat et son optimisation, l'étape ultime de notre travail est la validation de model en utilisant 4 puits différents qui n'ont pas été utilisé pour l'entrainer afin d'évaluer ces performances (voir la figure.IV.9 ci-dessous).

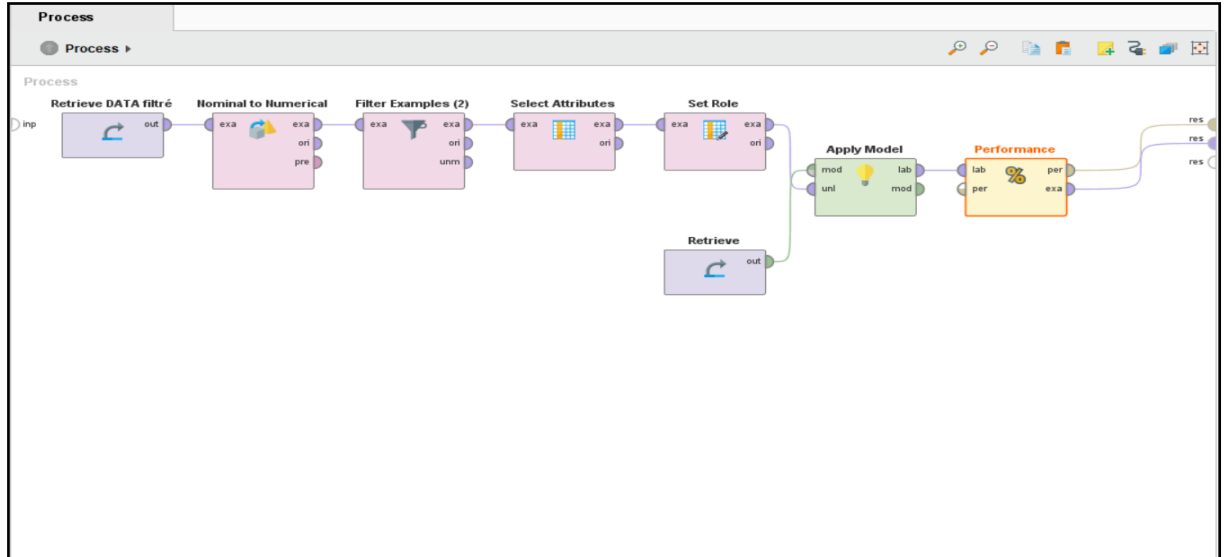
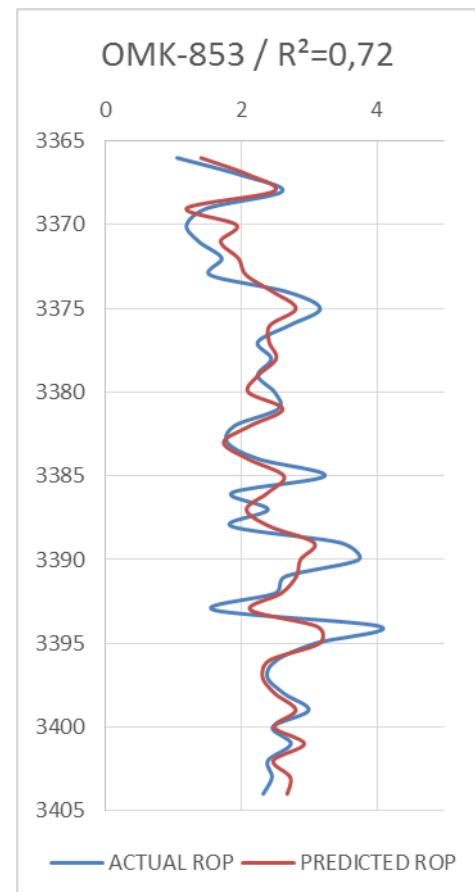
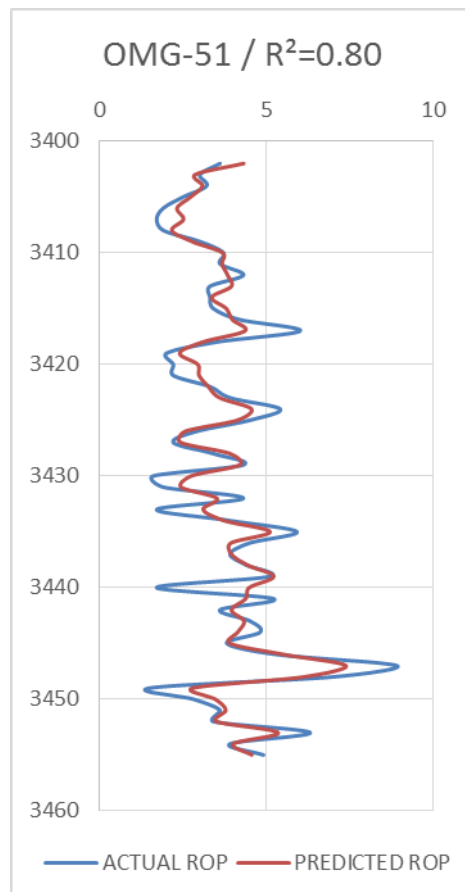


Figure IV.9 : Schéma descriptif l' d'évaluation des performances.

Les résultats du ROP prédit et celui du ROP réel pour ces puits sont représentés dans les figures ci-dessous :



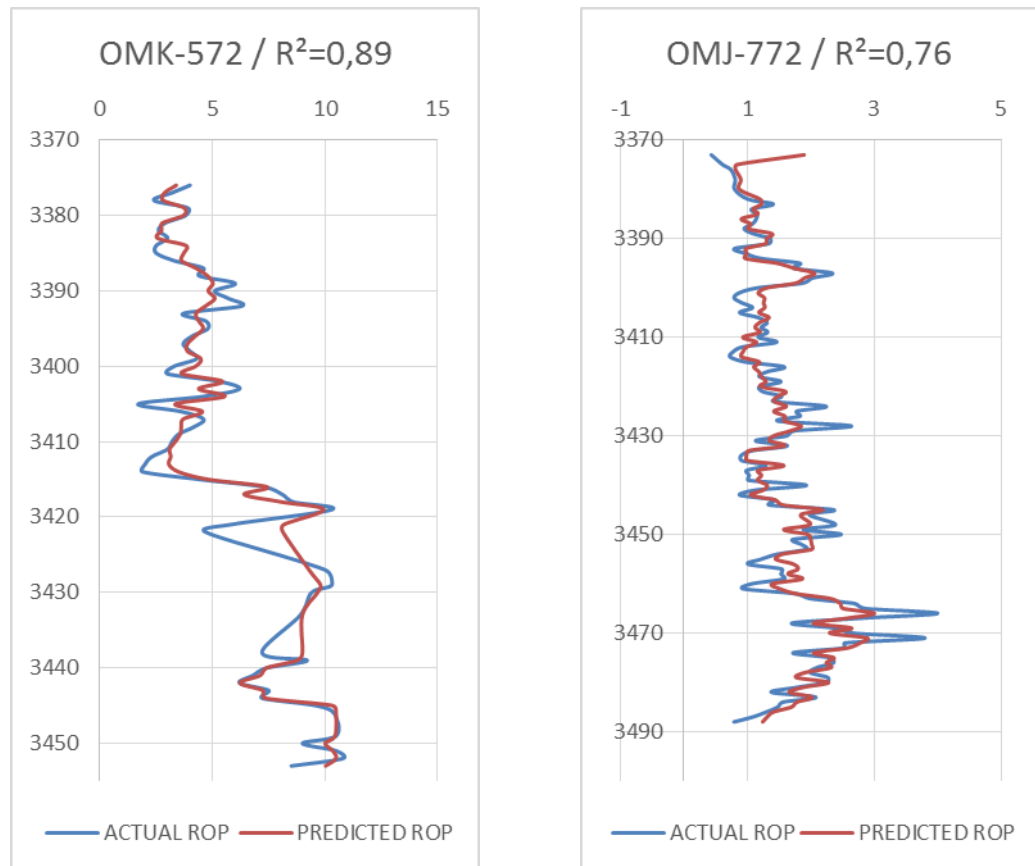


Figure IV.10 : Schéma descriptif d'un test de validation de chaque outil.

On remarque que le modèle permet de prédire la valeur du ROP avec une grande fiabilité avec des coefficients de détermination R^2 variant entre 0,72 jusqu'à 0,89. Néanmoins il pourra être amélioré dans des travaux futurs en incluant d'autres données non disponibles ou bien en améliorant la qualité des données disponibles en incorporant par exemple des données avec une résolution plus élevée.

IV.1.5- Utilisation du modèle comme outil de sélection des outils :

Dans cette ultime étape, le modèle obtenu sera utilisé afin de prédire l'outil pouvant générer le ROP moyen (moyenne harmonique) le plus élevé parmi un éventail de 7 outils différents, les résultats sont résumés dans le tableau suivant :

Tableau IV.3 : Résultats du ROP Moyen de chaque Outils.

Outils	ROP Moyen (moyenne harmonique m/hr)
<i>D1820W</i>	1,60
<i>GT64DDHE</i>	1,91
<i>HHD372G8</i>	1,72
<i>K505BCTPX</i>	1,80
<i>K505BEPX</i>	1,75
<i>K505BPX</i>	1,42
<i>TF610DPS</i>	1,76

Les résultats détaillés des ROP obtenus en fonction de la profondeur sont représentés sur le graphique ci-dessous :

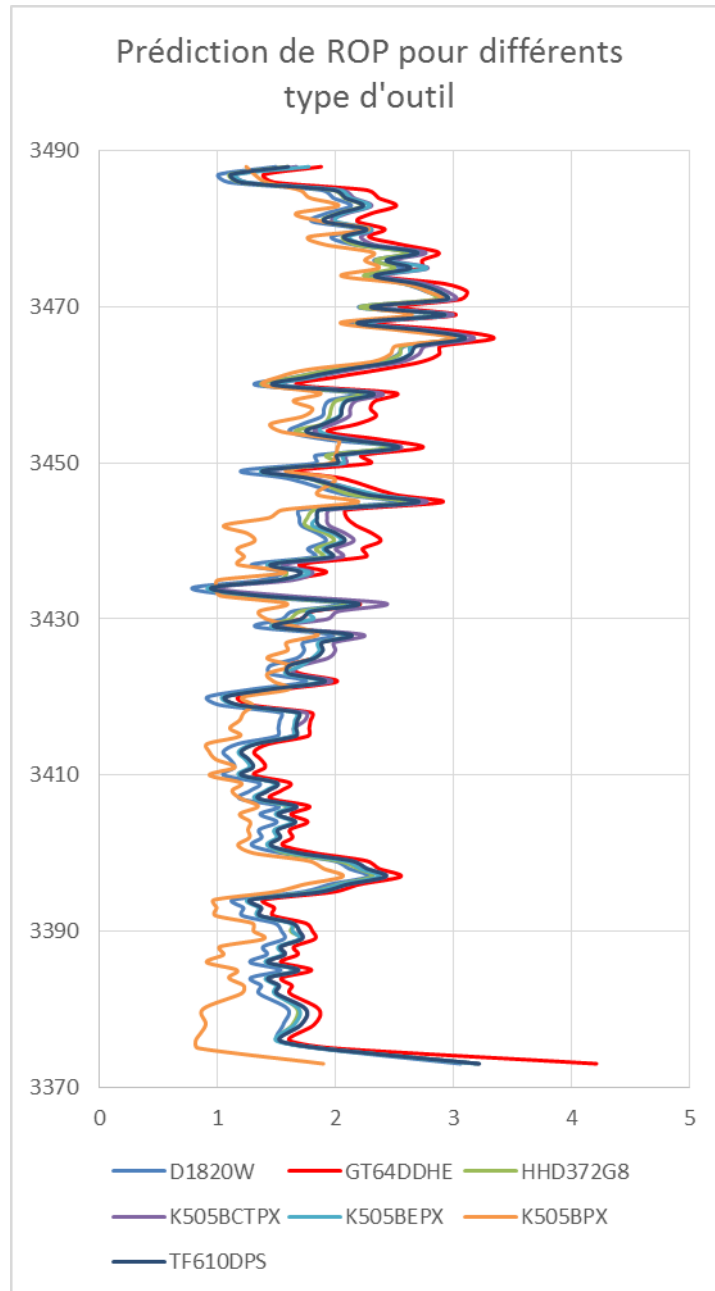


Figure.IV.11 : Prédiction de ROP pour différents type d'outil

On remarque que l'outil de forage pouvant offrir le ROP le plus élevé est le GT64DDHE avec un ROP moyen 1,91 m/hr, suivi du K505BCTPX avec un ROP moyen 1,80 m/hr et du TF610DPS avec un ROP moyen 1,76 m/hr.



**Conclusion
&
Recommandation**

Conclusion & recommandation :

Cette étude est basée sur la création d'un modèle de machine learning supervisé où a été développé et validé en utilisant plusieurs ensembles de données provenant des puits localisés dans la partie upside nord de Hassi Messaoud.

Ce modèle a ensuite été utilisé comme outil décisionnel pour sélectionner de manière préalable l'outil de forage offrant la vitesse d'avancement la plus élevée lors du forage de la phase 6" dans cette région.

Le modèle pratique permet de remplacer la méthode traditionnelle de sélection des outils, qui se base uniquement sur une comparaison basique des taux de pénétration (ROP) des puits voisins (offset wells), sans prendre en compte les caractéristiques géologiques, pétrophysiques et géomécaniques des formations traversées. Ces caractéristiques sont souvent omises dans l'analyse comparative des offset wells. De plus, ce modèle se différencie des modèles analytiques présents dans la littérature scientifique, car il se base sur des données spécifiques au champ d'étude.

Malgré les résultats satisfaisants obtenus, ce modèle peut encore être amélioré afin de pouvoir être étendu à l'ensemble du champ de Hassi Messaoud. Ces améliorations pourraient consister à inclure davantage de puits forés dans l'ensemble du champ, ce qui permettrait d'avoir une approche plus quantitative. Alors, il serait intéressant d'incorporer d'autres variables telles que l'overbalance, la distribution granulométrique de la roche ou la proportion de quartz, afin d'améliorer la qualité des prédictions.

Des variables spécifiques à l'outil de forage ou à l'assemblage de fond de trou (BHA), comme l'usure de l'outil ou les caractéristiques des moteurs de fond, pourraient également être prises en compte, car elles ont une influence significative sur la vitesse d'avancement. Donc, l'utilisation de données mesurées in situ, telles que des mesures downhole, pourrait permettre une meilleure compréhension de l'environnement de travail du trépan.

Il est important de souligner que la qualité des données peut également avoir une influence sur les performances du modèle. Afin d'améliorer la qualité des données, il est recommandé d'inclure des puits forés récemment, car les outils de mesure actuels offrent une résolution et une précision supérieures à celles des instruments utilisés par le passé. De plus, il est essentiel de prendre en compte les caractéristiques spécifiques des capteurs utilisés pour collecter les données, tels que leur précision, leur fréquence d'échantillonnage et leur résistance aux conditions environnementales du champ de forage.

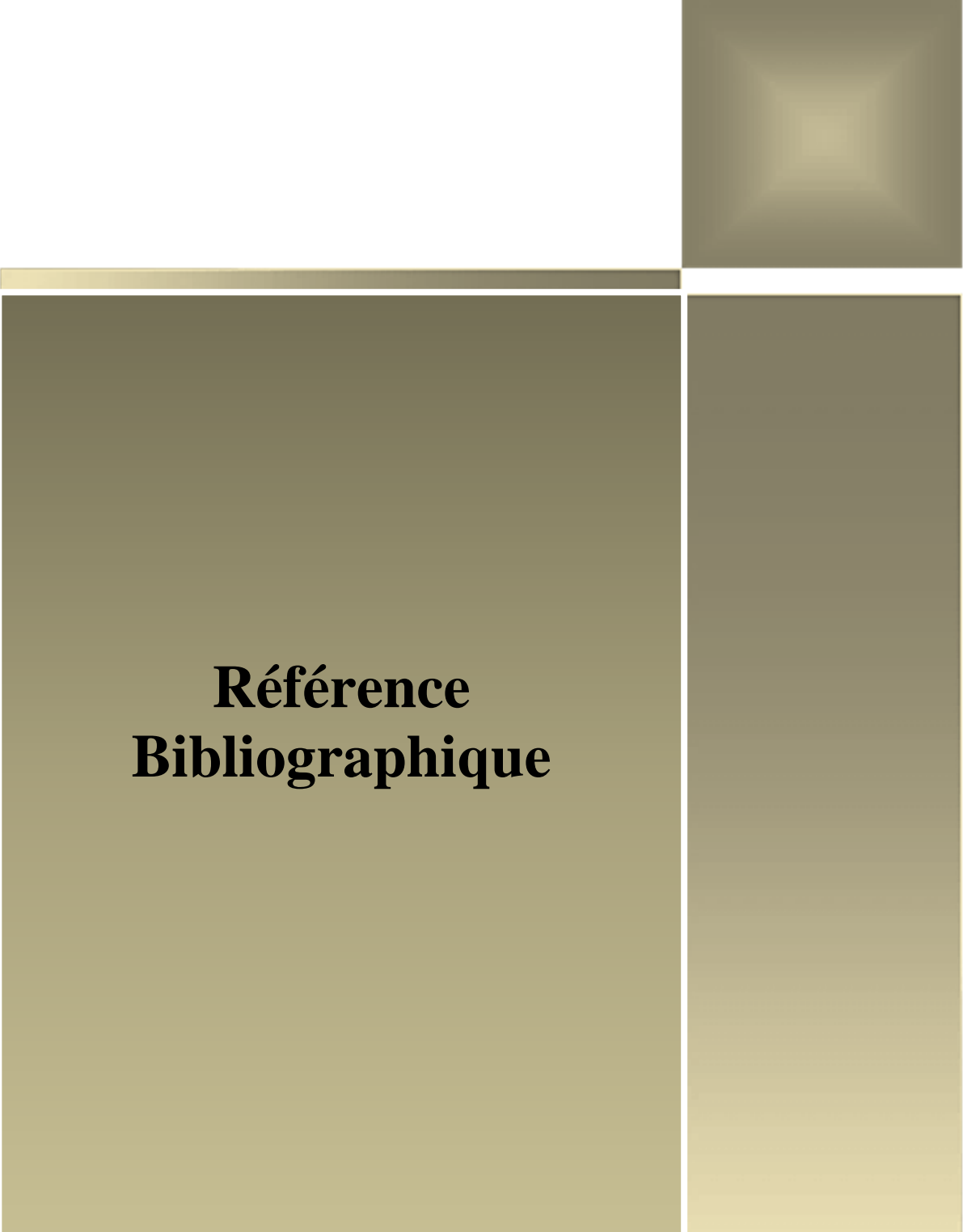
Pour optimiser davantage les performances du modèle, des techniques avancées de prétraitement des données peuvent être appliquées. Par exemple, l'identification et l'élimination des valeurs aberrantes (outliers) peuvent aider à éliminer les mesures incohérentes qui pourraient biaiser les résultats.

Des approches de clustering peuvent être utilisées pour détecter des schémas ou des regroupements de données similaires, ce qui pourrait permettre une meilleure compréhension des différentes conditions de forage, la réduction de dimensionnalité peut être employée pour sélectionner les caractéristiques les plus informatives et réduire la complexité du modèle.

Concernant l'algorithme utilisé (le gradient boosted tree), il est important de noter que l'optimisation des hyperparamètres est un aspect crucial pour améliorer les performances. Outre la méthode d'optimisation en grille mentionnée précédemment, des approches basées sur des algorithmes métaheuristiques, tels que les algorithmes génétiques ou les algorithmes d'optimisation par essaim de particules, peuvent être explorées pour trouver les combinaisons d'hyperparamètres optimales de manière plus efficace.

Dans l'ensemble, l'amélioration continue du modèle de prédiction de la vitesse de pénétration du forage nécessitera une approche globale qui englobe à la fois l'amélioration de la qualité des données, l'utilisation de techniques avancées de prétraitement, l'optimisation des hyperparamètres de l'algorithme et l'exploration de nouvelles approches d'optimisation. En intégrant ces aspects dans les travaux futurs, il sera possible d'améliorer considérablement la précision et la performance du modèle, contribuant ainsi à l'avancement de l'industrie pétrolière dans le domaine de l'optimisation du forage

Enfin, cette étude constitue une étape importante dans le domaine de l'optimisation des modèles de ROP. Les résultats obtenus serviront de base pour des études futures et contribueront à l'amélioration continue des techniques de prédiction de la vitesse de pénétration du forage dans l'industrie pétrolière.



**Référence
Bibliographique**

Références Bibliographiques :

- [1] J.P GUYEN, <<Techniques d'exploitation pétrolière-Le Forage>>, Editions TECHNIP 1993.
- [2] Document, <<Géologie de Hassi Messaoud HMD>>, SONATRACH.
- [3] Region, T. (s. d.). Ministère de l'énergie. [En ligne], disponible sur le site web : <https://www.energy.gov.dz/?rubrique=hydrocarbure>. [Consulter le 01/04/2023].
- [4] Documents SONATRACH <<Carte des zones de production à HMD>>.
- [5] IADC Drilling Manual, 12th Edition 2015.
- [6] Article: Boost your knowledge of Directional Drilling science & art, Wednesday 01/02/2017, [En ligne]. Disponible sur le site web : <https://directionaldrillingart.blogspot.com/2017/02/tcr-bits-cutting-mechanism-twotypes-of.html>. [Consulter le 04/04/2023].
- [7] A. SLIMANI, <<Module M2 >>, SONATRACH DIVISION FORAGE, Juin 2006.
- [8] ENSPM, FORMATION INDUSTRIE FORAGE PRODUCTION ET GISEMENT.
- [9] McGehee, D., Dahlem, J., Gieck, J., Kost, B., Lafuze, D. L., Reinsvold, C, & Steinke, S. (1992). The IADC Roller Bit Classification System. Dans All Days [En ligne] disponible sur le site web: <https://doi.org/10.2118/23937-ms>. [Consulter le 07/04/2023]
- [10] Roller Cone Manual Aug 02 - SMITH BITS TRAINING DOCUMENT Introduction to Roller Cone Bits Copyright 2002 Smith International Inc. All Rights | Course Hero. 10/01/2010). [En ligne], disponible sur le site web: <https://www.coursehero.com/file/18765351/Roller-Cone-Manual-Aug-02/> [Consulter le 09/04/2023].
- [11] ENSPM, <<Les outils de forage>>, Formation Industrie Forage –Production -Gisement Edition 1999.
- [12] F. GRONDIN, <<Module Second Les outils de forage>>, SONATRACH DIVISION FORAGE, Juin 2004.
- [13] Frederic SUR, <<Introduction à l'apprentissage automatique>>.
- [14] S.Russell, P.Norvig, <<Intelligence artificielle, Avec plus de 500 exercices>>,2010.
- [15] C. Bishop, <<Pattern recognition and machine learning>>, Springer 2006.
- [16] Le machine learning au service de la santé – (2020, août 25). [En ligne], disponible sur le site web: <https://techtomed.com/le-machine-learning-au-service-de-la-sante/> [Consulter le 12/04/2023].

- [17] Belaidi, N. (s. d.). L'apprentissage supervisé : définition et exemples. Formation Tech et Data en ligne | Blent.ai. 12 avr. 2022, disponible sur le site web: <https://blent.ai/blog/a/apprentissage-supervise-definition> [Consulter le 13/04/2023]
- [18] Machine Learning : Définition, fonctionnement, utilisations. (2022, 23 décembre). Formation Data Science | DataScientest.com. [En ligne], disponible sur le site web: <https://datascientest.com/machine-learning-tout-savoir> [Consulter le 10/05/2023]
- [19] Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116(44), 22071-22080. [En ligne], disponible sur le site web: <https://doi.org/10.1073/pnas.1900654116> [Consulter le 15/05/2023].
- [20] Document SONATRACH << Centre de recherche et développement Hassi messaoud >>.
- [21] Dr.Ahmad Al-AbdulJabbar <<Rate of penetration prediction for multiple well profiles using Artificial Intelligence techniques>> SPE webinar"Video". Octobre 2020.
- [22] Markus Hofmann, Ralf Klinkenberg, <<RapidMiner Data Mining Use Cases and Business Analytics Applications>>, 19 Avril 2016
- [23] GeeksforGeeks. (2023). Cross Validation in Machine Learning. *GeeksforGeeks*. [En ligne], disponible sur le site web <https://www.geeksforgeeks.org/cross-validation-machine-learning/> [Consulter le 29/05/2023].