

**UNIVERSITE KASDI MERBAH OUARGLA**  
**Faculté des Sciences Appliquées**  
**Département de Génie Civil et d'Hydraulique**



**Mémoire**

Présenté en vue de l'obtention du diplôme de **Master professionnel**

**Domaine : Sciences et technique**

**Filière : Hydraulique**

**Spécialité : Forage d'eau**

Présenté par : HOGGUI Houdaïfa  
MANSOURI Ali

**Thème :**

*Utilisation des techniques de data-mining pour  
La détection de coincement de la garniture de forage*

Soutenu publiquement : le 03 Juin 2 017

**Devant le jury composé de :**

|                     |            |              |                          |
|---------------------|------------|--------------|--------------------------|
| Mr. MAHI Rachid     | M. A – A – | Président    | U. Kasdi Merbah. Ouargla |
| Mme. Rezagui Djihad | M. A – B – | Examinatrice | U. Kasdi Merbah. Ouargla |
| Mr. Djebari Hacene  | M. A – A – | Rapporteur   | U. Kasdi Merbah. Ouargla |

Année Universitaire : **2 016 / 2 017**

# Remerciement

*Nous remercions tout d'abord notre Dieu qui nous a donné la force et la volonté pour élaborer ce travail.*

*Nous adressons nos vifs remerciements à notre encadreur "Mr. Djebari Hacene", qui n'a pas cessé de nous donner les conseils et les bonnes orientations et nous prive pas de son temps et aussi on remercie "Mr. Ghenabzia Ahmed ", qui nous écoute avec grande patience pendant toute la période de préparation malgré ses contraintes d'enseignement.*

*Sans son aide, notre travail n'aurait pas vu la lumière. Notre reconnaissance va aussi à tous ceux qui ont collaboré à notre formation en particulier les enseignants du Département de Génie Civil et d'Hydraulique,*  
**UNIVERSITE KASDI MERBAH**

**OUARGLA**

*Aussi à nos collègues de la promotion 2016-2017 On remercie également tous ceux qui ont participé de près ou de loin à élaborer ce travail*

**H. A**

# Listes des tableaux

|   |    |
|---|----|
| Table.III.1: Les paramètres de notre modèle .....                     | 28 |
| Table.III.2: comparaison entre les algorithmes de classification..... | 30 |

# Listes des figures

## I. Généralités sur le Forage

|   |    |
|---|----|
| Figure I.1 : Système de forage.....                           | 04 |
| Figure I.2: Architecture et e processus de forage.....        | 05 |
| Figure I.3 : Principe du forage Rotar.....                    | 06 |
| Figure I.4: Schéma détaillé d'une sonde de forage rotary..... | 07 |
| Figure I.5 : Différents types d'outils de forage Rotary.....  | 09 |
| Figure I.6 :Stabilisateur.....                                | 10 |
| Figure I.7 : Quelques éléments de la garniture de forage..... | 10 |
| Figure I .8: Exemple de Stuck mécanique.....                  | 14 |
| Figure I.9: Exemple de Stuck différentiel.....                | 15 |

## II. Data-mining

|   |    |
|---|----|
| Figure II. 1: Processus de data – mining..... | 18 |
|---|----|

## III. Conception

|   |    |
|---|----|
| Figure III. 1 :Description de la transmission de données RTOM.....                                | 26 |
| Figure III.2: l'architecture global du système.....   | 26 |
| Figure III.3: l'expérimentation et l'évaluation en utilise l'algorithme de BayesNet ( BN).....    | 29 |
| Figure III.4: l'expérimentation et l'évaluation en utilise l'algorithme de arbre dedécision (DT). | 30 |

## IV. Implémentation et réalisation

|  |    |
|--|----|
| Figure IV.1:La langage de développement.....   | 32 |
| Figure IV.2: L'outil WEKA.....   | 33 |
| Figure IV.3:L'interface principale de l'application.....                             | 34 |
| Figure IV.4:L'interface du comparaison entre les méthodes d'application.....         | 34 |
| Figure IV.5:L'interface de résultat avec le meilleur algorithme (DT) coincement..... | 35 |

# Liste des abreviations

- BN : **B**ayes**N**et (Réseau bayésien)
- KNN : **K** Nearest Neighbor (k-plus proche voisin)
- RPM : Rotation per minute
- RTOM : **R**eal Time **O**perating **M**anagement
- DT : **T**ree of **D**ecision( Arbre de décision)
- WEKA : **W**aikato **E**nvironment for **K**nowledge **A**nalysis
- WOB : **W**eight **O**n **B**it (Poids sur l'outil)

# Sommaire

|                              |            |
|------------------------------|------------|
| <b>Liste des tableaux</b>    | <b>I</b>   |
| <b>Liste des figures</b>     | <b>II</b>  |
| <b>Sommaire</b>              | <b>III</b> |
| <b>Liste d'abréviation</b>   | <b>V</b>   |
| <b>Introduction générale</b> | <b>01</b>  |

## Chapitre I : Généralités sur le Forage

|  |     |
|--|-----|
| Introduction .....   | 03  |
| 1-Qu'est – que le forage?.....                               | 05  |
| 2- Principe de forage rotary.....                            | 06  |
| 3- Description générale d'un appareil de forage.....         | 06  |
| 4- Equipements de lavage.....                                | 06  |
| 4-1. Le mat de forage.....                                   | 06  |
| 4-2. La substructure.....                                    | 08  |
| 4-3. Le mouflage.....  | 08  |
| 4-4. Le moufle fixe.....                                     | 08  |
| 4-5. Le moufle mobile.....                                   | 08  |
| 4-6. Le crochet.....   | 08  |
| 4-7. Le câble de forage.....                                 | 08  |
| 4-8. Le treuil.....  | 08  |
| 5- La garniture de forage.....                               | 09  |
| 5-1. L'outil de forage.....                                  | 09  |
| 5-2. Les stabilisateurs.....                                 | 10  |
| 5-3. Les masses-tiges ( Drill Collars , DC).....             | 10  |
| 5-4. Les tiges (Drill Pipes, DP).....                        | 11  |
| 6- Parameters de forage.....                                 | 11  |
| 6-1. Les paramètres mécaniques.....                          | 12  |
| 7- Types des problèmes de forage.....                        | 13  |
| 7-1. Mud Contamination.....                                  | 13  |
| 7-2. La perte de circulation (Loss circulation).....         | 13  |
| 7-3. Les venues (Kick).....                                  | 13  |
| 7-4. Coincements de la garniture de forage (Stuck Pipe)..... | 13  |
| Conclusion.....  | 161 |

## Chapitre II : Data-mining

|   |    |
|---|----|
| Introduction.....                             | 17 |
| 1- Qu'est ce que la fouille de données ?..... | 17 |
| 2- Le processus de data-mining.....           | 17 |
| 3- Les tâches de data-mining.....             | 18 |
| 3-1. La classification.....                   | 19 |
| 3-2. L'estimation.....                        | 19 |
| 3-3. La prédiction.....                       | 20 |
| 4- Les techniques de data – mining.....       | 20 |
| 4-1. Le Réseau Bayésien (BN).....             | 21 |

|  |    |
|--|----|
| 4-2. L'algorithme des k-Plus proches voisins (k – NN)..... | 21 |
| 4-2-1. Comment cela marche-t-il ?.....                     | 21 |
| 4.2.2 Avantages et inconvénients.....                      | 21 |
| 4-3 l'arbre de décision (DT).....                          | 22 |
| Conclusion.....  | 24 |

### **Chapitre III : Conception**

|   |    |
|---|----|
| Introduction.....   | 25 |
| 1- Description du processus de prévision des coincements..... | 25 |
| 1-1. Le projet RTOM.....                                      | 25 |
| 2- Architecture global de système.....                        | 26 |
| 3- Fonctionnement global.....                                 | 27 |
| 4- Description des données.....                               | 28 |
| 5- Comparaisons entre les méthodes.....                       | 29 |
| Conclusion.....   | 31 |

### **Chapitre IV: Implémentation et réalisation**

|   |    |
|---|----|
| Introduction.....                                 | 32 |
| 1. Environnement et outils de mise en oeuvre..... | 32 |
| 1-1.Java.....                                     | 32 |
| 1-2. Weka (machine Learning).....                 | 33 |
| Conclusion.....                                   | 35 |
| Conclusion Général.....                           | 36 |
| Bibliographié.....                                | 37 |

Le forage représente l'ensemble des opérations permettant de creuser un ouvrage, pour atteindre (en profondeur) de nouvelles zones (aquifère ou réservoirs), susceptibles de contenir soit de l'eau ou des hydrocarbures (gaz ou pétrole).

La compétitivité accrue entre les compagnies internationales (soit dans le domaine hydraulique ou pétrolier) incite à aller vite pour creuser des puits, devenus de plus en plus coûteux.

Aujourd'hui, le forage réalisé généralement par une société de service spécialisée, est devenu une activité très technique. De nos jours, on a tendance à axer nos efforts au développement de méthodes ou de moyens d'aide pour assister les techniciens dans le domaine, notamment dans le choix des différents paramètres de forage, et ce, pour un gain appréciable en temps et une minimisation souhaitée du prix de revient du mètre foré.

Pendant les différentes étapes ou phases du processus de forage, il est fort probable que certaines difficultés surgissent, constituant ainsi des obstacles pouvant engendrer le ralentissement, voire même l'abandon du trou, ce qui se répercute, en grande partie, sur l'augmentation du prix de revient et le coût global de l'investissement.

D'entre les obstacles que nous pouvons rencontrer dans le domaine du forage, tels que les pertes de circulations (soit totales ou partielles), les éboulements des parois, les venues...etc, nous nous sommes intéressés au problème de coincement de la garniture de forage (*Stuck Pipe*).

Ce phénomène est très réputé et très courant dans notre domaine. Il peut être à l'origine du blockage de l'avancement et nous contraint à avoir recours à certaines opérations spéciales, voire l'abandon prématuré du forage, ce qui cause des pertes de temps inutiles et par conséquent, un gaspillage de centaines de millions de dollars chaque année, pour l'industrie, soit hydraulique ou pétrolière.

Dans de telles circonstances, le foreur ne peut plus ni avancer ni retirer la garniture du puits, et le moindre geste, non étudié, peut compliquer la situation d'avantage.

Il est alors indispensable que ce phénomène soit détecté pour se prémunir de complications indésirables, chose qui est possible grâce à certains paramètres hydrauliques et mécaniques.

Néanmoins, il serait beaucoup plus intéressant de pouvoir le prévoir et le prédire, bien avant qu'il ne surgisse, le diagnostiquer et le représenter d'une façon explicite, afin d'en faire bénéficier les responsables.



Pour pallier au problème, cité ci – dessus, nous essayerons, dans cette étude, à travers notre système, de mettre à la dispositions des exploitants un outil qui leur permettra de réaliser plusieurs objectifs :

- ✓ Atteindre la profondeur souhaitée par les géologues ;
- ✓ Avoir la probabilité de détection le type de coincement avant de survenir.

Pour ce faire, le travail qui nous a été confié, est structuré en quatre chapitres principaux:

- Dans le premier chapitre, nous avons essayé de dresser un état de l’art le domaine, qui donne les notions de base sur le forage (hydraulique, pétrolier), et ses divers problèmes dont le coincement de la garniture qui représente notre principale problématique.
- Dans le second chapitre, nous aborderons les méthodes de data – mining qui sont utilisées pour résoudre le problème constituant le sujet de ce travail.
- Dans le troisième chapitre, nous présenterons l’architecture globale du système, et nous décriverons, d’une façon détaillée, l’interaction entre les entités du système, les outils et les méthodes utilisées de son développement.
- Le quatrième chapitre est consacré à la présentation de l’environnement de développement, avec quelques interfaces de notre logiciel, en terminant par une discussion sur les résultats obtenus.

Dans la conclusion générale, nous dresserons la synthèse du travail réalisé dans ce mémoire, suivie des perspectives envisagées.

A red banner with a white border and wavy edges, containing the chapter title in white serif font.

# **Chapitre I : Généralités sur le Forage**

A red banner with a white border and wavy edges, containing the chapter title in white serif font.

**Chapitre II :**  
**Data-mining**

**Chapitre III :**

*Conception*



**Chapitre IV:  
Implémentation et  
réalisation**

## **Introduction**

L'exploitation des ressources souterraines nécessite l'emploi de plusieurs techniques dont le forage. Pour atteindre les objectifs, préalablement désignés, le concours de plusieurs acteurs est conseillé, voire impératif, dont des géologues, des géophysiciens, des experts de forage, des hydrauliciens, des écologistes...etc.

Les techniques existent, telle que le forage par battage, au marteau fond du trou, par havage, au rotary...etc, pouvant effectuer des trous horizontaux, verticaux, dirigés mais, il va falloir d'abord sélectionner la plus adéquate et la plus adaptée.

Dans ce chapitre, on s'intéresse d'une part à introduire et présenter un état de l'art dans le domaine du forage, et d'autre part à exposer le principe de forage au rotary, et d'énumérer les problèmes qu'il est possible de rencontrer, lors de l'exécution des travaux.

### **1. Qu'est – que le forage?**

Le forage, (pétrolier ou hydraulique) est l'ensemble des opérations qui permettent, aux techniciens du domaine, d'atteindre les formations géologiques perméables et poreuses, du sous – sol, susceptibles de se comporter comme aquifère (et contenir de l'eau) ou comme réservoir (et contenir des hydrocarbures liquides ou gazeux).

L'implantation d'un forage (pétrolier ou hydraulique) n'est décidée qu'après réalisation d'une série d'études géologiques, hydrogéologiques et géophysiques, effectuées sur un système aquifère ou un bassin sédimentaire.

Ces études permettent de se faire une idée sur la structure du sous – sol, la profondeur de ses couches constitutives et de leurs possibilités aquifères ou de gisements néanmoins, elles ne sont pas en mesure de prouver, avec précision, l'existence ou l'absence d'eau ou d'hydrocarbure.

Seuls les forages pourront confirmer les hypothèses émises au préalable, et mettre en évidence la nature du (ou des) fluides contenu(s) dans ces roches.

L'incertitude quant à la nature des fluides piégés dans le sous – sol et la complexité des dépôts sédimentaires expliquent le nombre, encore élevé, de forages de reconnaissance et d'exploration négatifs (un sondage d'exploration productif pour une dizaine négatifs) et la part importante de ces forages dans le coût de la recherche.[1]

Donc, la réalisation d'un tel ouvrage vise à atteindre plusieurs objectifs, qui sont presque les mêmes d'ailleurs. Nous citons entre autres:

- ✓ Atteindre la profondeur souhaitée par les géologues ;
- ✓ Réaliser divers tests et mesures ;

- ✓ Récolter le maximum de données (géologiques, hydrogéologiques, géophysiques, etc...) ;
- ✓ Avoir la possibilité d'exploiter le système aquifère et de tester ou de mettre en production le (ou les) réservoir (s).

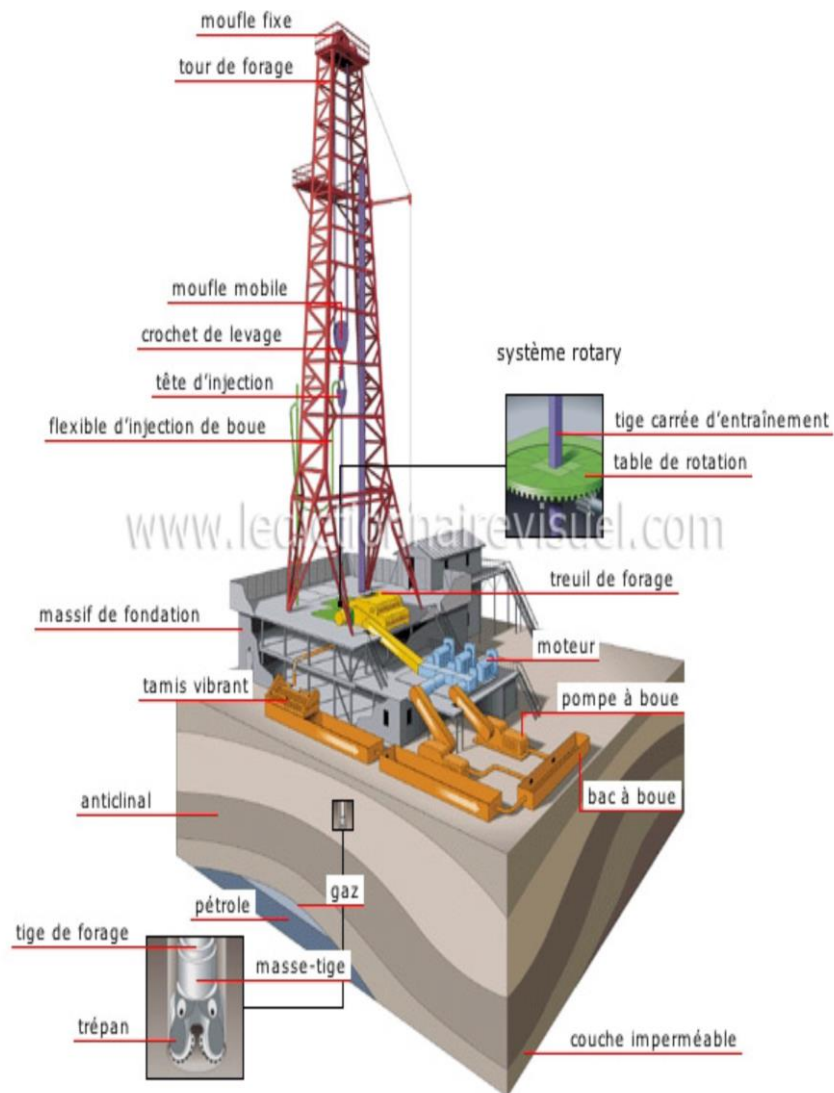


Figure I. 1 : Système de forage

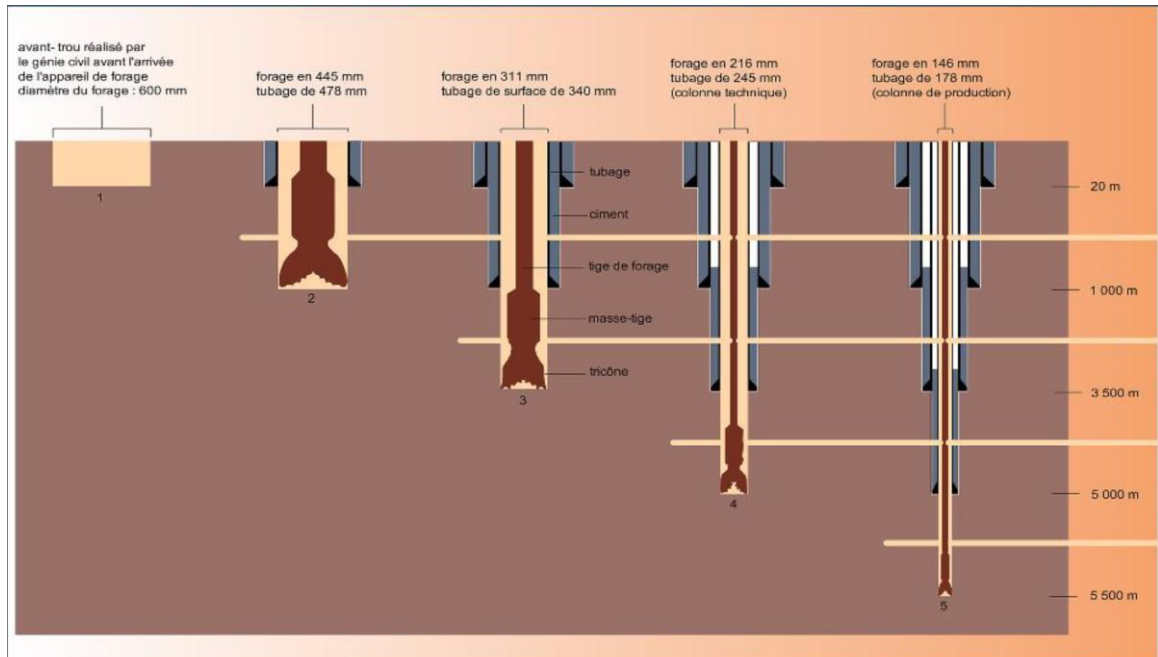


Figure I.2: Architecture et e processus de forage

## 2. Principe de forage au rotary

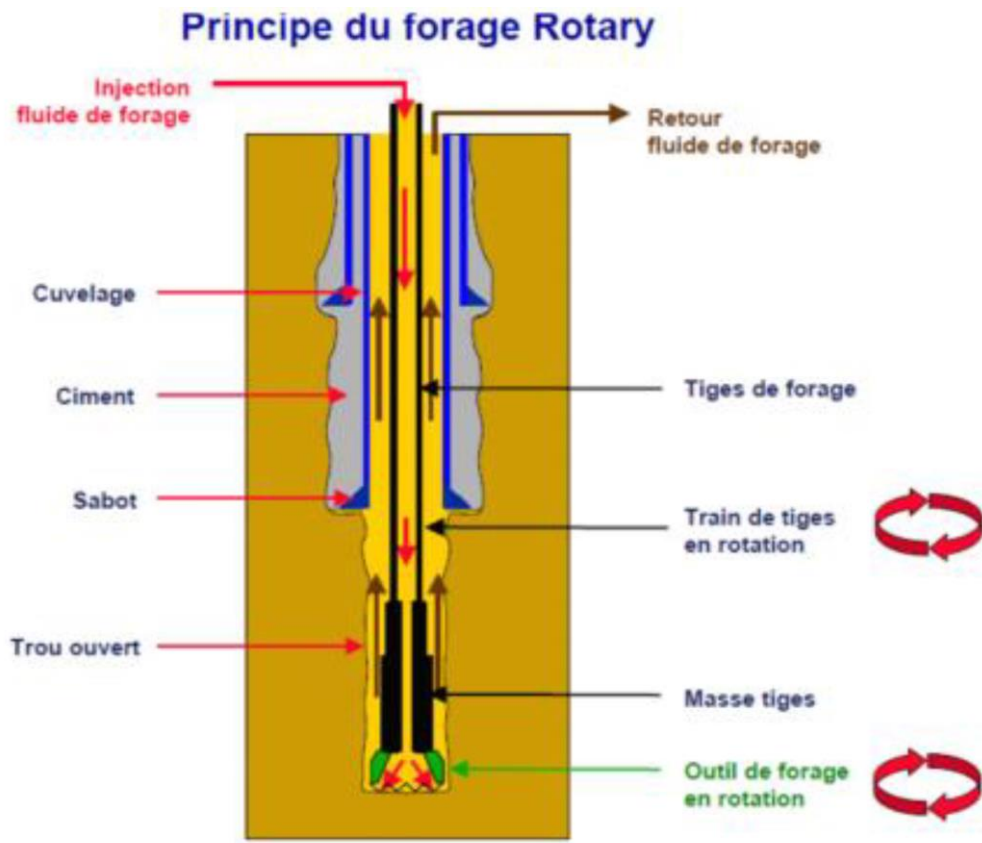
La technique du forage rotary consiste à utiliser un outil, permettant de détruire la roche, sous l'effet du poids et de la rotation du train de sonde, composé d'un assemblage de tiges de forage, dont les tiges proprement dites (DP), de masses – tiges (DC), filetées bout à bout et se terminant par un outil, dit trépan.

Ces tiges doivent être creuses, pour acheminer, sous pression, une boue, dite de forage, au fond du trou, et doivent aussi être lourdes (masses – tiges ou tiges lourdes) pour assurer un poids sur l'outil.

La circulation du fluide de forage permet le maintien des parois, le refroidissement et la lubrification de l'outil ainsi que la garniture et assure le déblaiement du trou, en remontant les cuttings au jour.

Autour des parois du forage, des tubes sont descendus et du ciment est injecté afin de garantir l'ancrage de l'ouvrage, la tenue de ses parois, ainsi qu'une protection contre la corrosion mais aussi la préservation des nappes d'eau souterraines et leur isolation thermique.[1]





**Figure I.3 : Principe du forage Rotary**

### 3. Description générale d'un appareil de forage

Un appareil de forage est constitué d'un ensemble d'équipements, permettant les fonctions de levage, de rotation, de pompage...etc, afin de réaliser un sondage. Il doit aussi comporter d'autres équipements, qui ne sont pour autant négligeables, notamment ceux de fourniture d'énergie et de mise en sécurité du puits, pour pouvoir travailler sans incidents et se prémunir des accidents (cas de venue, de perte, d'éboulement, coincement...etc.). [1]

Deux caractéristiques relativement liées interviennent dans la classification d'un appareil de forage :

- ✓ La capacité de profondeur de forage maximale ;
- ✓ La puissance au treuil.

Les appareils de forage sont classés, en première approche, par la capacité de profondeur maximale de forage (avec des tiges ou  $DP$  de  $4^{n/2}$ ). On distingue :

- **Les appareils légers** : pour des profondeurs inférieures à 1200 m ;
- **Appareil moyen** : pour des profondeurs comprises entre 1200-2500 m ;
- **Appareil lourd** : pour des profondeurs comprises entre 2500-4000 m ;

- **Appareil ultra- lourd** : pour des profondeurs supérieures à 4000 m.

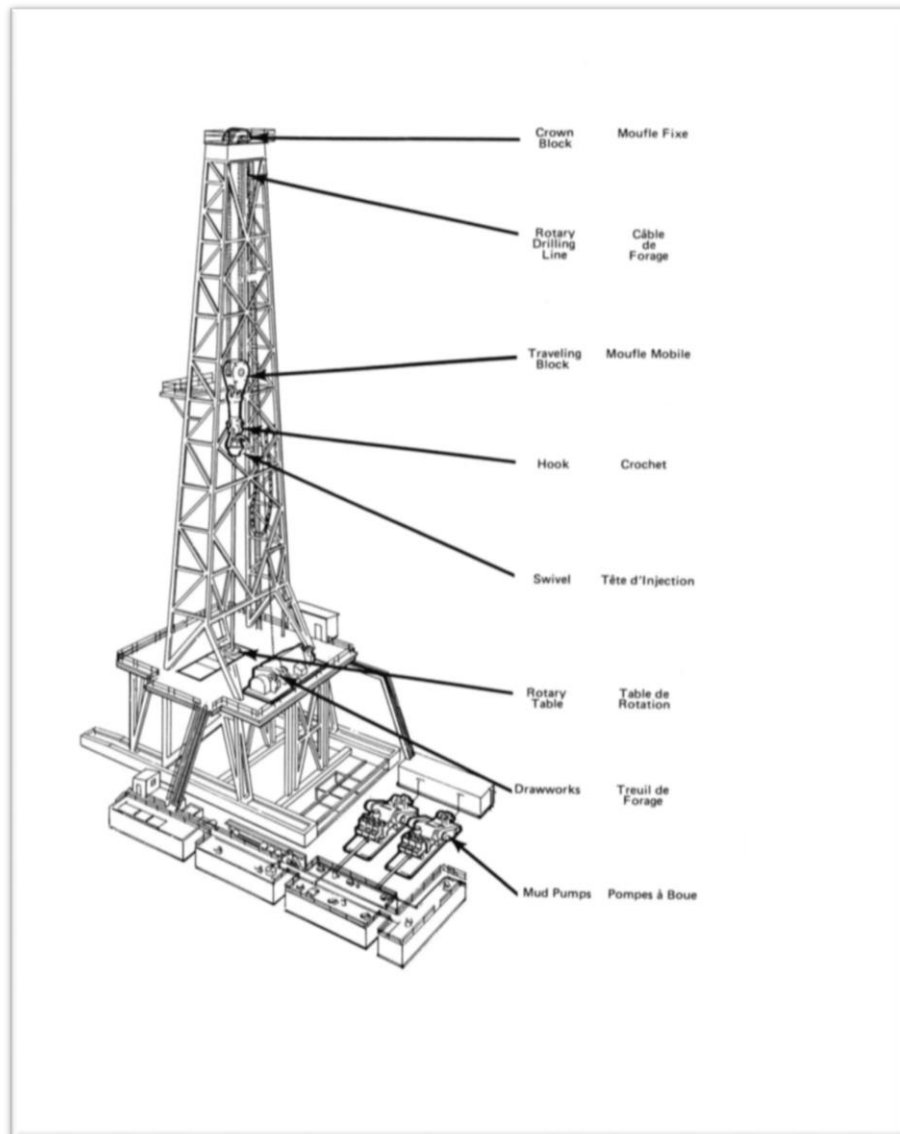


Figure I.4: Schéma détaillé d'une sonde de forage rotative

#### 4. Équipements de levage

Ils servent à suspendre la garniture de forage et la faire monter et descendre dans le puits. Ils permettent aussi d'ajuster le poids sur le trépan qui est nécessaire pour que ce dernier puisse forer. [1]

**4-1. Le mat de forage** : c'est le « trépied » qui supporte le palan. Il a remplacé la tour de forage, pour la rapidité de son montage et démontage. [1]

À son sommet, est placé le moufle fixe. Une passerelle d'accrochage est placée à son milieu (environ 26m), elle sert de lieu de travail pour l'accrocheur, pour l'ajout ou l'enlèvement de longueurs de tiges, lors de la remontée ou la descente de l'outil dans le puits.

Une autre passerelle de hauteur ajustable, placée plus bas (à environ 9m), sert à guider le tubage pour le visser et le descendre dans le puits.

Un plancher de travail est aménagé aux pieds du mât. Il sert d'aire de travail pour l'équipe. Une cabine est aménagée sur ce plancher pour permettre aux ouvriers de se reposer.[1]

**4-2. La substructure :**le plancher est surélevé de quelques mètres, au – dessus du sol, pour permettre l'introduction des éléments de la tête de puits et des obturateurs. [1]

**4-3. Le mouflage :**c'est l'enroulement du câble de forage, entre les poulies des moufles fixe et mobile, en plusieurs brins (jusqu'à 14 brins). Ce système permet de démultiplier le poids de la garniture de forage et de diminuer la vitesse de son déplacement. [1]

En négligeant les frottements, la tension sur le brin actif est divisée par le nombre de brins, et sa vitesse est multipliée par ce nombre. [1]

**4-4. Le moufle fixe :**est formé d'un certain nombre de poulies et placé au sommet du mât.Il possède une poulie de plus que le moufle mobile. [1]

**4-5. Le moufle mobile :**est formé également d'un certain nombre de poulies, par lesquelles passe le câble de forage.Il se déplace sur une certaine hauteur entre le plancher de travail et le moufle fixe. [1]

**4-6. Le crochet :** sert à la suspension de la garniture pendant le forage. Des bras y sont accrochés de part et d'autre servent à supporter l'élévateur, utilisé pour la manoeuvre de la garniture. [1]

**4-7. Le câble de forage :**les câbles utilisés sur les appareils de forage sont composés d'un ensemble de torons, constitués en fil d'acier,enroulés ou tressés autour d'une ame métallique (en acier) ou en chanvre. [1]

**4-8. Le treuil :**c'est un fût (tambour) autour duquel s'enroule le câble de forage. Aux bouts de ce fût sont fixées des jantes qui servent au freinage, au contact de bandes en acier contenant des patins en ferodo, actionnées par un levier. Le treuil est le coeur de l'appareil de forage, donc c'est la capacité du treuil qui caractérise un appareil de forage et indique la classe de profondeur des forages que l'on pourra atteindre ou effectuer. [1]

## 5. La garniture de forage

Il s'agit, comme avancé plus haut, d'un ensemble de tige, de masses – tiges, et d'organes annexes, permettant d'effectuer plusieurs fonctions (acheminne la boue vers le fond, transmet le poids et le mouvement de rotation au rock bit...etc.

De manière générale, une garniture de forage se compose, de bas en haut, des éléments suivants :

**5-1. L'outil de forage :** Extrémité basse de la garniture, l'outil est entraîné en rotation et permet la progression de la garniture par son action directe sur le front de taille. On distingue deux types d'outils, à molettes (tricônes) et monoblocs (diamant, PDC). [1]



Figure I.5 : Différents types d'outils de forage Rotary

## 5-2. Les stabilisateurs :

Ils ont un diamètre presque égal à celui de l'outil et sont intercalés entre les masses – tiges, pour les maintenir droites dans le puits et éviter leur flexion ainsi que pour garder la rectilignité du puits et éviter la déviation. [1]

Nous donnons en figure I. 6, un exemple.

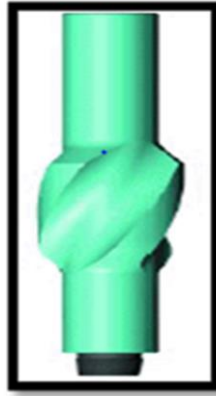


Figure I.6 : Stabilisateur

**5-3. Les masses – tiges ( Drill Collars , DC) :** il s'agit d'une série de tubes massifs en acier, dont le rôle est d'assurer une poussée suffisante sur l'outil. Par conséquent, pendant la phase de forage, le train de masses – tiges est en partie, voire en totalité, en état de compression. [1]



Figure I.7 : Quelques éléments de la garniture de forage

**5-4. Les tiges (Drill Pipes, DP) :** c'est la partie la plus longue de la garniture ; elle est constituée d'une série importante de tubes minces d'acier, qui s'étendent jusqu'à la surface, et dont le rôle est de transmettre le couple vers l'outil et de soutenir, plus ou moins, le train de masses tiges.

Chaque tige est munie, à ses extrémités, de renflements appelés *tool – joint* (TJ) (figure I.7), servant pour le vissage des tiges entre elles (pour protéger leur filetage) et comme points d'appui du train de tiges sur la paroi du puits.

En conditions normales, on s'arrange pour que le train de tiges soit en traction, afin d'éviter son flambement, car celui – ci favorise l'apparition de forces de contacts très élevées entre tiges et parfois le puits.

Ce train de tiges comporte parfois (cas du forage dirigé ou du forage horizontal) une série supplémentaire de tiges lourdes (*Heavy weight Drill Pipe, HWDP*), plus épaisses, placées

souvent juste après les masses – tiges. Ces tiges lourdes ont pour but de réduire les risques de flambement des tiges. [1]

## 6- Les paramètres de forage

### Définition:

Les paramètres de forage consistent en plusieurs grandeurs physiques, mesurées et enregistrées pendant le forage à l'aide de capteurs placés un peu partout dans le RIG et dont les variations dépendent du terrain traversé. Ces paramètres peuvent être classés en deux catégories, mécaniques et hydrauliques. [2]

**6-1. Les paramètres mécaniques:** la valeur des paramètres mécaniques diffère selon la dureté des couches géologiques traversées.

- **Le poids sur l'outil (WOB) :** Le poids sur l'outil « Weight on Bit » représente le poids qu'on doit appliquer sur le trépan pour qu'il puisse détruire la roche et avancer. Il peut être calculé, en utilisant la formule suivante :

$$WOB = WOH L - WOH F$$

Où :

WOH L : C'est le poids de la garniture

WOH F : C'est le poids de la garniture durant le forage.

Ce poids croît de plus en plus dans les formations les plus dures et est de l'ordre de 250 à 1 000 daN par centimètre de diamètre de trépan. Par exemple ce paramètre pourra être de l'ordre de 30 000 daN, pour un trépan de 311 mm de diamètre, pour terrains durs. Cette augmentation de poids provoque la diminution de la durée de vie du trépan. [2]

- **La vitesse de rotation (RPM) :** La vitesse de rotation « Révolution per minute », indique le nombre de tour par minute de la table de rotation sur le rig – floor. Elle est, en général, comprise entre 50 et 300 tours par minute, selon la dureté de la formation. Par exemple, pour le trépan cité précédemment, pour terrains durs de 311 mm de diamètre, la vitesse de rotation pourra être au alentours de 100 tours par minute. Elle est déduite via un capteur, placé au niveau de la table de rotation. [2]
- **Le torque :** il ne s'agit pas ici d'un paramètre à optimiser, c'est juste la force opposée de la rotation, il est principalement fonction du poids sur l'outil (WOB), la vitesse de rotation (RPM), et la dureté de la formation traversée.

Le couple en surface n'est pas transmis intégralement à l'outil de forage, mais la mesure en surface demeure la seule possibilité à notre disposition, actuellement .[2]

- **La vitesse d'avancement (ROP)** : elle est fonction de la nature des terrains, et varie de 1 à 2 m par heure dans les terrains durs, et peut atteindre trente (30) mètres par heure, voire même plus, dans les terrains tendres.

La durée de vie des tricônes à molettes dentées est de l'ordre de 40 heures. Le travail du trépan a pour conséquences l'usure des dents qui peut aller jusqu'à disparition complète de celles – ci, l'usure des roulements aussi, qui peut causer la perte d'une molette au fond du trou, ainsi qu'une perte de diamètre, dans les terrains abrasifs.

Pour améliorer la longévité des tricônes, les paliers des molettes doivent être lubrifiés et éventuellement lisses. Dans ce cas, la durée de vie de l'outil peut dépasser les 150 heures .[2]

## 6-2. Les paramètres hydrauliques

- **Le débit de circulation (Flow in)** : c'est le volume de boue déversé par les pompes à l'intérieur du trou de forage. La quantification de ce paramètre est basée sur le débit unitaire des pompes (POP), selon le nombre de coup par minute, contribué par les pompes (SPM). Le SPM est enregistré à l'aide d'un capteur de fréquence similaire à celle de la rotation. L'unité du débit est le L/mn.[2]
- **Le type de boue** : Le type de boue est choisi en fonction des performances recherchées et désigne les propriétés physico – chimiques du fluide de forage. Trois types de boues sont souvent employés, la boue à base d'eau (Water Based Mud, WBM), la boue à base d'huile (Oil Based Mud, OBM) et la boue synthétique (Synthetic Based Mud, SBM). Une boue synthétique est constituée d'un mélange d'eau et d'additifs chimiques.[2]
- **La densité de la boue**: l'obtention des informations relatives au puits, et particulièrement le contrôle de la pression du fond, s'effectue à travers la densité de la boue. En remontant en surface, le fluide de forage est chargé de déblais, d'huile mais aussi de gaz contenu dans les roches. Cela fournit des indications sur la nature des fluides se situant dans le réservoir et représente un élément important dans le pilotage de la garniture.[2]
- **La pression hydraulique** : le débit et la pression hydraulique représentent les variables physiques qui doivent favoriser une bonne évacuation des déblais et éviter des problèmes d'encrassement du trépan ou du puits.[2]

## 7-Types de problèmes de forage

Lors de la planification, la clé pour atteindre les objectifs avec succès, est de concevoir des programmes de forage sur la base de l'anticipation des problèmes potentiels de trous plutôt que sur la prudence et de confinement.

Les problèmes de forage peuvent être très coûteux. Nous présentons les plus courants à savoir. [3]

### 7-1. Mud Contamination

La boue est dite contaminée, lorsque un matériau étranger s'y introduit et provoque des changements indésirables dans ses propriétés, comme la densité, la viscosité et la filtration.

En général, les systèmes de boue à base d'eau sont les plus sensibles à la contamination. Ce phénomène peut résulter d'un traitement excessif du système de boue avec des additifs ou de matières entrant dans la boue en cours de forage.[2]

### 7-2. La perte de circulation (Loss circulation)

Une perte de circulation est le flux incontrôlé de boue qui pénètre dans une formation, notamment en face des zones fracturées ou fissurées, causant une diminution de volume. On parle alors de zone vierge et de zone envahie. Elle peut être totale ou partielle.[2]

### 7-3. Les venues(Kick)

Une venue est l'entrée de fluide de formation dans le puits, pendant les opérations de forage. Elle se produit en raison de l'insuffisance de la pression hydrostatique exercée par la colonne de boue de forage pour contrebalancer la pression exercée par les fluides des formations forées. L'essence même du contrôle de puits (Well control) est d'empêcher la venue, et si elle arrive, agir avant qu'elle ne se développe en éruption.

Une venue incontrôlée résulte habituellement du mauvais déploiement de l'équipement approprié, en utilisant de mauvaises pratiques, ou un manque de formation des équipes de forage. La perte de contrôle du puits peut conduire en éruption, ce qui représente l'une des menaces les plus graves liées à l'exploration des ressources souterraines, impliquant le risque de vie et les conséquences environnementales et économiques.[2]

### 7-4. Coincements de la garniture de forage (Stuck Pipe) :

**Définition** :c'est le phénomène de collage, de plaquage ou d'immobilisation de la garniture de forage (ou de tubage) contre les parois, et indique le fait que la garniture soit prise au piège dans le puits, on ne peut plus ni avancer ni la retirer.



Il constitue un obstacle qui peut conduire à l'arrêt, plus ou moins long, du forage, ce qui nécessite, en général, des opérations d'instrumentation, et représente 40% du coût total du forage. Chose qui peut avoir plusieurs causes, selon l'état de la garniture, qu'elle soit statique ou en mouvement. Ces coincement sont classé en deux catégorie [2] :

- ✓ Les coincements mécaniques ;
- ✓ Les coincements par pression différentielle ;
- ✓ Les coincements par instabilité de terrains.

### A- Les coincements mécaniques

Les coincements mécaniques se produisent lorsque la garniture est en mouvement dans le puits et est causé par des obstructions ou restrictions physiques. Ces phénomènes sont souvent localisés au niveau des changements de diamètres de la garniture (sur une faible portion).

Leurs causes sont très diverses et sont en général difficiles à identifier. On distingue:

- ✓ Coincements causés par dog – leg ou key – seat ;
- ✓ Chutes d'objets métalliques dans le trou ;
- ✓ Diminution du diamètre nominal du trou ;
- ✓ Coincement liés aux conditions géométriques du puits.

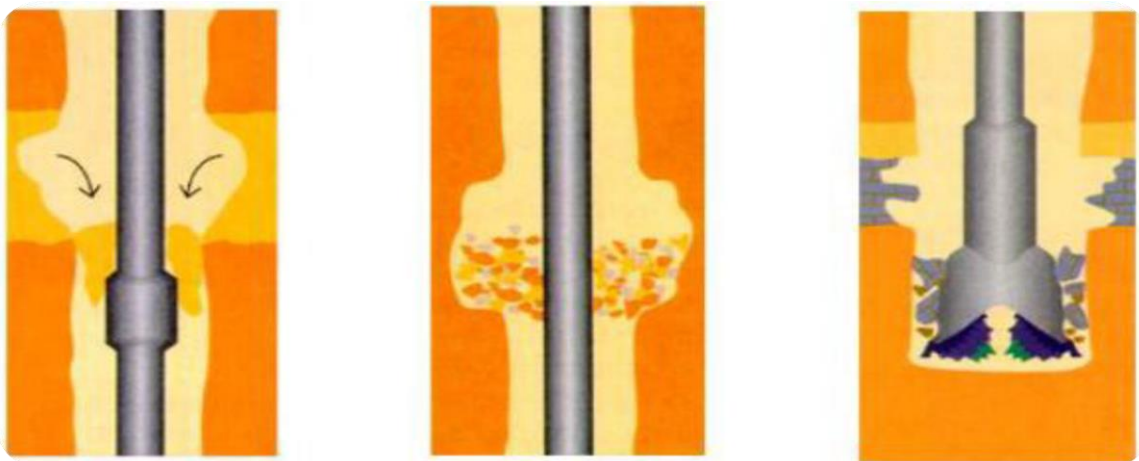


Figure I .8: Exemple de Struck mécanique

### B- Les coincements par pression différentielle

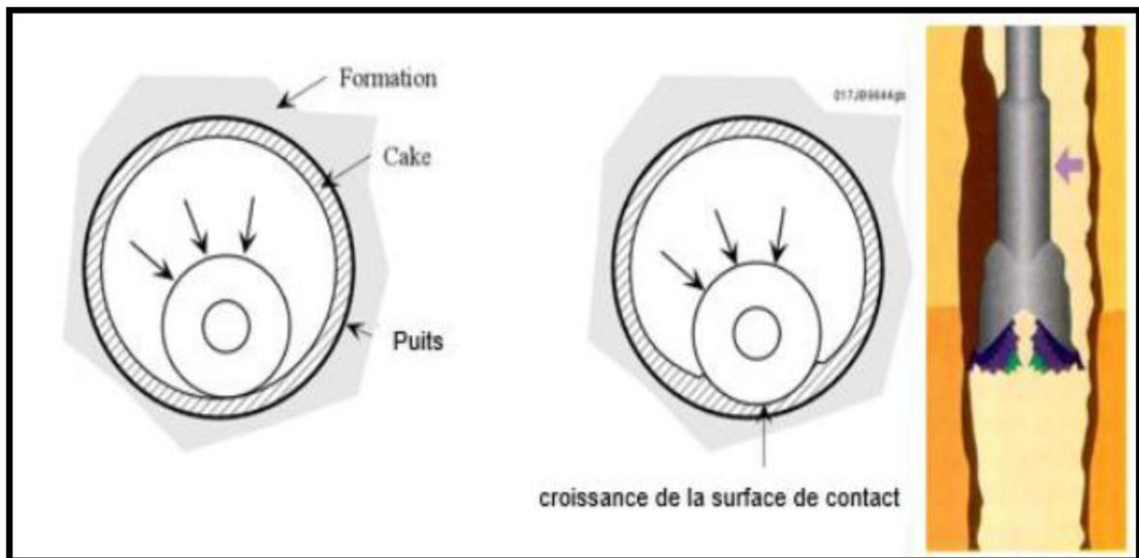
Le coincement par pression différentielle est causé par la différence entre la pression qui règne dans une formation poreuse et perméable et celle exercée par la colonne de boue.

Il est due à une force causée par la pression surpassée de la boue qui agit sur la garniture de forage et la coince dans une formation perméable. La partie de la garniture coincée dans ce que l'on appelle un gâteau de boue « Mud cake» a une pression égale à celle

de la formation qui agit sur elle, alors que l'autre partie de la garniture a une pression égale à la pression hydrostatique de la colonne de boue.

Lorsque cette pression hydrostatique ( $P_h$ ) dans le puits est supérieure à celle de la formation ( $P_f$ ), elle engendre une force qui colle la garniture sur les parois du puits .[2]

Ce type de coincement ne se produit pas dans les schistes et autres formations non perméable, où le cake ne se forme pas habituellement, ils ont lieu généralement quand la garniture ou l'outil de forage est stationnaire (ajout de tige) ou en mouvement lent.[2]



**FigureI.9: Exemple de Stuck différentiel**

Une fois la garniture coincée, tous les efforts devraient être fournis pour la libérer, le plus rapidement possible, car cette probabilité (la libérer) diminue rapidement avec le temps, l'identification prématurée de la cause la plus probable d'un problème de coincement est cruciale, puisque chaque cause doit être remédiée avec des mesures différentes.

Une réaction incorrecte à ce problème pourrait facilement le rendre plus graves. Une évaluation des événements menant jusqu'au coincement indique fréquemment la cause la plus probable et peut mener aux mesures correctives appropriées.[2]

### **C – Les coincements par instabilité des formations**

Ils peuvent avoir lieu en face de certaines formations rendues instables (exécution du trou et présence du fluide de forage). Cet état (instabilité) peut causer la déformation du trou, ce qui entraîne le coincement (partiel ou total) du train de sonde. Il s'agit de :

- ✓ Rétrécissement des parois ;
- ✓ Gonflement et Fluage des Argiles ;

- ✓ Les éboulements ;
- ✓ Coincements par les déblais.

**Conclusion**

Nous avons donné, dans ce chapitre, un aperçu sur les principales notions de base du domaine de forage, les concepts et processus importants pour son déroulement. Cependant, nous avons constaté que malgré les précautions prises et les moyens déboursés dans ce sens, les foreurs peuvent se heurter à plusieurs obstacles, entre autres les coincements de la garniture. Á vrai dire, il y a seulement quelques cas où il n'était pas possible d'éviter ce phénomène, mais il est toujours important d'avoir une réaction rapide et bien adaptée, car seul l'homme peut être en mesure d'agir, mais pour maintes raisons, notamment économiques et sécuritaires, il doit être aidé par la machine.

C'est dans ce sens qu'est orientée notre étude, à travers laquelle nous allons tenter de réaliser un système de prédiction, capable de prévenir à l'avance certains de ces cas de coincements, et de les anticiper, avant même qu'ils ne voient le jour.

## Introduction

Comme avancé plus haut, les problèmes rencontrés au cours du forage doivent être résolus, mais nous estimons que, même si nous réussissons l'opération, les pertes de temps et d'argent qui en résultent sont indésirables. Il est préférable de trouver un moyen de les prospector, si non les prédire avant qu'ils ne surgissent.

D'entre les difficultés que les foreurs peuvent rencontrer sur chantier, le coincement de garniture de forage, qui peut les amener à abandonner le trou.

Comme nous signalé précédemment, dans le présent travail, nous nous sommes fixés pour objectif de proposer une solution, qui doit nous aider à prévoir ce type de problème, mais il est commode de présenter d'abord cette technique, qui est le data – mining ainsi que les méthodes de classification, Réseau Bayésien (BN), l'Arbre de Décision (DT), le k – plus proche voisin (K – NN), qui sont utilisées pour résoudre la problématique objet de ce travail.

### 1- Qu'est ce que la fouille de données ?

Le data – mining est un domaine interdisciplinaire utilisant, dans le même temps, des techniques d'apprentissage automatique de reconnaissance des formes, des statistiques, des bases de données et de visualisation pour déterminer les manières d'extraction des informations de très grandes bases de données. [4] C'est aussi un processus inductif, itératif et interactif dont l'objectif est la découverte de modèles de données valides, nouveaux, utiles et compréhensibles dans de larges bases de données. [5]

### 2- Le processus de data – mining

Le processus de data – mining est d'analyser des données provenant de différentes perspectives et résumant en informations utiles ; information qui peut être utilisée pour augmenter les revenus, réduire les coûts, ou les deux à la fois. [6]

Plus précisément, le data – mining peut se décomposer en 7 étapes [6]:

- ✓ **Le nettoyage de données** : qui gère les données bruitées, erronées, manquantes ou non pertinentes ;
- ✓ **L'intégration de données** : où de multiples sources de données hétérogènes, peuvent être intégrés dans une seule ;
- ✓ **La sélection de données** : où les données relatives à la tâche d'analyse sont récupérées à partir de bases de données ;

- ✓ **Transformation de données** : où les données sont transformées ou regroupées dans des formes appropriées pour l'exploitation, en effectuant des opérations de synthèse ou d'agrégation ;
- ✓ **L'exploration de données** : qui est le processus essentiel, où les méthodes intelligentes sont appliquées, afin d'en extraire les modèles de données ;
- ✓ **L'évaluation de modèle** : ce qui est d'identifier les motifs vraiment intéressants représentant les connaissances basées sur des mesures de qualité ;
- ✓ **Présentation de la connaissance** : où les techniques de visualisation et de représentation des connaissances sont utilisées pour présenter, les connaissances extraites, à l'utilisateur.

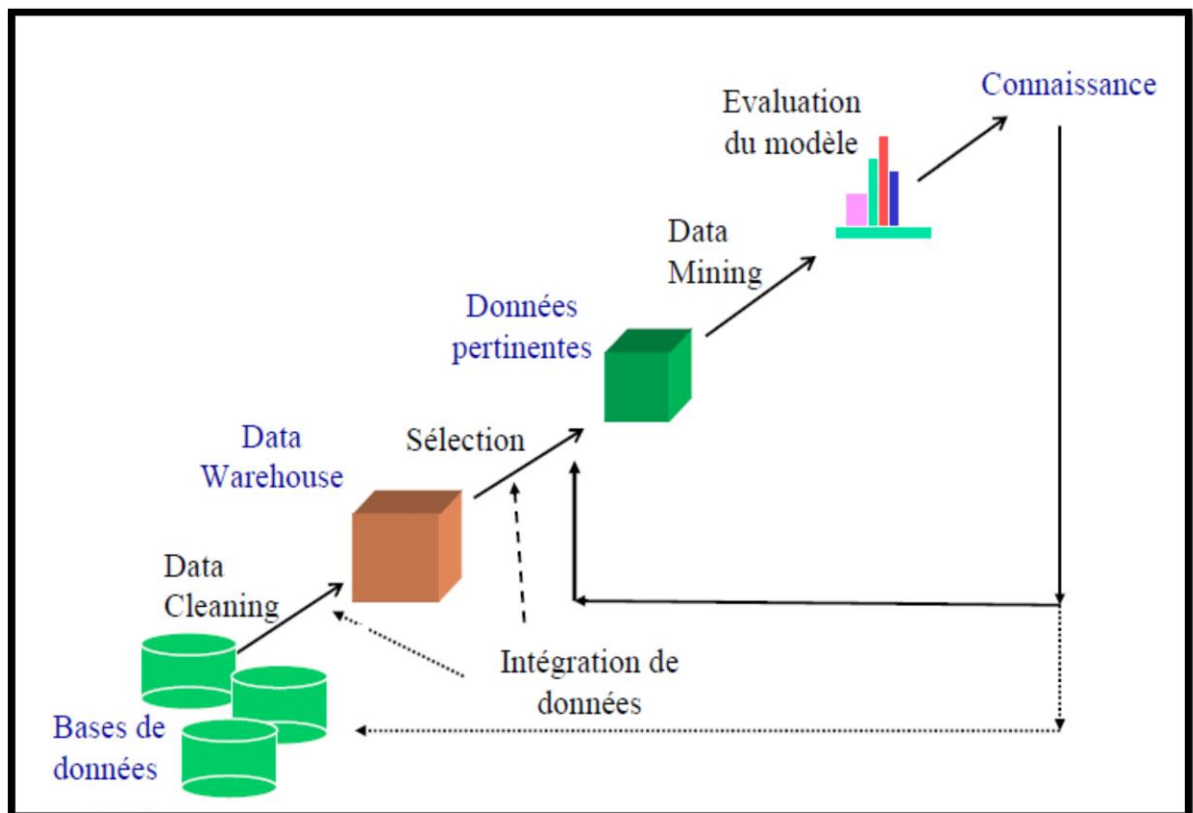


Figure II. 1: Processus de data – mining

### 3 - Les tâches de data – mining

Beaucoup de problèmes intellectuels, économiques ou même commerciaux peuvent être exprimés en termes des six tâches suivantes [7]:

- ✓ La classification ;
- ✓ L'estimation ;

- ✓ La prédiction ;
- ✓ Le groupement par similitude ;
- ✓ L'analyse des clusters ;
- ✓ La description .

Les trois premières tâches sont des exemples du data – mining supervisés, dont le but est d'utiliser les données disponibles pour créer un modèle décrivant une variable particulière, prise comme but, en termes de ces données.

Le groupement par similitude et l'analyse des clusters sont des tâches non-supervisées où le but est d'établir un certain rapport entre toutes les variables [8].

La description appartient à ces deux catégories de tâche, et vue comme une tâche supervisée et non-supervisée en même temps [7].

### 3-1. La classification

La classification est la tâche la plus commune du data – mining et qui semble être une obligation humaine. Afin de comprendre notre vie quotidienne, nous sommes constamment classifiés, catégorisés et évalués [7].

La classification consiste à étudier les caractéristiques d'un nouvel objet, pour lui attribuer une classe prédéfinie. Les objets à classifier sont généralement des enregistrements d'une base de données, et la classification consiste donc à mettre à jour chaque enregistrement, en déterminant un champ de classe.

La tâche de classification est caractérisée par une définition de classes bien précise et un ensemble d'exemples classés auparavant. L'objectif est de créer un modèle qui peut être appliqué aux données non classifiées dans le but de les classifiées [8].

Par exemple, à partir d'un ensemble de patients diagnostiqués, qui servent l'ensemble de la formation, un modèle de classification peut être construit, ce qui met un terme à la maladie d'un patient de son / ses données de diagnostic. Le modèle de classification peut être utilisée pour diagnostiquer la maladie d'un nouveau patient sur la base des données de diagnostic du patient comme l'âge, le sexe, le poids, la température, la tension artérielle...etc.[8]

### 3-2. L'estimation

L'estimation est similaire à la classification, à part que la variable de sortie est numérique plutôt que catégorique. En fonction des autres champs de l'enregistrement, l'estimation consiste à compléter une valeur manquante dans un champ particulier. Par exemple, on cherche à estimer la lecture de tension systolique d'un patient dans un hôpital, en

se basant sur l'âge du patient, son genre, son indice de masse corporelle et le niveau de sodium dans son sang. La relation entre la tension systolique et les autres données vont fournir un modèle d'estimation. Et par la suite nous pouvons appliquer ce modèle dans d'autres cas [7] [9].

### 3-3. La prédiction

La prédiction est la même que la classification et l'estimation, à part que dans la prédiction les enregistrements sont classés suivant des critères (ou des valeurs) prédites (estimées). La principale raison qui différencie la prédiction de la classification et l'estimation est que dans la création du modèle prédictif on prend en charge la relation temporelle entre les variables d'entrée et celles de sortie . [7]

## 4- Les techniques de data – mining

Pour effectuer les tâches du data – mining, il existe plusieurs techniques issues de disciplines scientifiques diverses (statistiques, intelligence artificielle, base de données), afin de faire apparaître des corrélations cachées dans des gisements de données, pour construire des modèles, à partir de ces données.

Dans cette adresse, nous proposons des techniques d'extraction de données qui sont utilisées dans notre étude. [6]

### 4-1. Le Réseau Bayésien (BN)

Un réseau bayésien, en informatique et en statistique, est un modèle graphique probabiliste, représentant des variables aléatoires, sous la forme d'un graphe orienté acyclique. Intuitivement, ils sont à la fois :

- ✓ Des modèles de représentation des connaissances ;
- ✓ Des « machines à calculer » les probabilités conditionnelles ;
- ✓ Une base pour des Système d'aide à la décision.

Pour un domaine donné (par exemple médical), on décrit les relations causales entre variables d'intérêt par un graphe. Dans ce graphe, les relations de cause à effet, entre les variables, ne sont pas déterministes, mais probabilisées.

Ainsi, l'observation d'une cause ou de plusieurs causes n'entraîne pas systématiquement l'effet ou les effets qui en dépendent, mais modifie seulement la probabilité de les observer.

L'intérêt particulier des réseaux bayésiens est de tenir compte, simultanément des connaissances à priori d'experts (dans le graphe) et de l'expérience contenue dans les données.

Les réseaux bayésiens sont surtout utilisés pour le diagnostic (médical et industriel), l'analyse de risques, la détection des spams et le data – mining.

## 4-2. L'algorithme des k-Plus proches voisins (k – NN)

L'algorithme des k plus proches voisins (K-PPV, k nearest neighbor en anglais, ou KNN), est un algorithme de raisonnement, à partir de cas qui est dédié à la classification, qui peut être étendu à des tâches d'estimation.

Le but de cet algorithme est de prendre des décisions, en se basant sur un ou plusieurs cas similaires déjà résolus en mémoire. Dans ce cadre, et contrairement aux autres méthodes de classification (arbres de décision, réseaux de neurones, algorithmes génétiques, ...etc.), l'algorithme de KNN ne construit pas de modèle à partir d'un échantillon d'apprentissage, mais c'est l'échantillon d'apprentissage, la fonction de distance et la fonction de choix de la classe fonction des classes des voisins les plus proches, qui constituent le modèle.[12]

### 4-2-1. Comment cela marche-t-il ?

Nous supposons avoir une base de données d'apprentissage constituée de ( $N$ ) couples « entrée – sortie ». Pour estimer la valeur de sortie d'une nouvelle entrée ( $x$ ), la méthode des  $k$  plus proches voisins consiste à prendre en compte (de façon identique) les  $k$  échantillons d'apprentissage, dont l'entrée est la plus proche de la nouvelle entrée ( $x$ ), selon une distance à définir.[10]

Si nous prenons une base d'apprentissage de cent(100) éléments, dès que nous recevons un nouvel élément, que nous souhaitons classifier, l'algorithme calcule sa distance à tous les éléments de la base. Si cette base comporte cent(100) éléments, alors il va calculer cent (100) distances et donc obtenir cent(100) nombres réels. Si  $k$  est égal à vingt cinq (25) par exemple, il cherche alors les vingt cinq(25) plus petits nombres parmi ces cent (100) nombres qui correspondent donc aux vingt cinq (25) éléments de la base qui sont les plus proches de l'élément que nous souhaitons classifier. La classe attribuée à l'élément à classifier est la classe majoritaire parmi ces vingt cinq(25) éléments.[11]

### 4-2-2. Avantages et inconvénients

#### a – Les avantages

- ✓ La qualité de la méthode s'améliore en introduisant de nouvelles données sans nécessiter la reconstruction d'un modèle. Ce qui représente une différence majeure avec des méthodes telles que les arbres de décision et les réseaux de neurones ;



- ✓ La clarté des résultats : la classe attribuée à un objet peut être expliquée en exhibant les plus proches voisins qui ont amené à ce choix ;
- ✓ La méthode peut s'appliquer à tout type de données, même celles complexes tels que des informations géographiques, des textes, des images, du son. C'est parfois un critère de choix de la méthode PPV, car les autres méthodes traitent difficilement les données complexes. Nous pouvons noter, également, que la méthode est robuste au bruit ;
- ✓ Facile à mettre en oeuvre.[12]

#### **b – Les inconvénients**

- ✓ Temps de classification : la méthode ne nécessite pas d'apprentissage, ce qui implique que tous les calculs sont effectués lors de la classification. Contrairement aux autres méthodes qui nécessitent un apprentissage (éventuellement long) mais qui sont rapides en classification ;
- ✓ Méthode donnera de mauvais résultats, si le nombre d'attributs pertinents est faible, relativement au nombre total d'attributs, car la proximité sur les attributs pertinents sera noyée par les distances sur les attributs non pertinents ;
- ✓ Les performances de la méthode dépendent du choix de la distance, du nombre de voisins et du mode de combinaison des réponses des voisins.[13]

### **4-3 l'arbre de décision (DT)**

L'apprentissage par arbre de décision désigne une méthode basée sur l'utilisation d'un arbre de décision comme modèle prédictif. On l'utilise notamment en fouille de données et en apprentissage automatique.

Dans ces structures d'arbre, les feuilles représentent les valeurs de la variable – cible et les embranchements correspondent à des combinaisons de variables d'entrée, qui mènent à ces valeurs.

En analyse de décision, un arbre de décision peut être utilisé pour représenter, de manière explicite, les décisions réalisées et les processus qui les amènent. En apprentissage et en fouille de données, un arbre de décision décrit les données mais pas les décisions elles – mêmes, l'arbre serait utilisé comme point de départ au processus de décision.

C'est une technique d'apprentissage supervisé où on utilise un ensemble de données pour lesquelles on connaît la valeur de la variable – cible, afin de construire l'arbre (données dites étiquetées), puis on extrapole les résultats à l'ensemble des données de test.

## A – Les avantages

Comparativement à d'autres méthodes de fouille de données, les arbres de décision présentent plusieurs avantages tels que :

- ✓ La simplicité de compréhension et d'interprétation., car c'est un modèle boîte blanche ; si l'on observe une certaine situation sur un modèle, celle – ci peut être facilement expliquée à l'aide de la logique booléenne, au contraire de modèles boîte noire comme les réseaux neuronaux, dont l'explication des résultats est difficile à comprendre ;
- ✓ Peu de préparation des données (pas de normalisation, de valeurs vides à supprimer, ou de variable muette) ;
- ✓ Le modèle peut gérer, à la fois, des valeurs numériques et des catégories. D'autres techniques sont souvent spécialisées sur un certain type de variables (les réseaux neuronaux ne sont utilisables que sur des variables numériques) ;
- ✓ Il est possible de valider un modèle à l'aide de tests statistiques, et ainsi de rendre compte de la fiabilité du modèle ;
- ✓ Performant sur de grands jeux de données ; la méthode est relativement économique en termes de ressources de calcul.

## B – Inconvénients

En revanche, elle présente certains inconvénients :

- ✓ L'apprentissage de l'arbre de décision optimal est NP-complet concernant plusieurs aspects de l'optimalité. En conséquence, les algorithmes d'apprentissage par arbre de décision sont basés sur des heuristiques telles que les algorithmes gloutons, cherchant à optimiser le partage à chaque nœud, et de tels algorithmes ne garantissent pas d'obtenir l'optimum global. Certaines méthodes visent à diminuer l'effet de la recherche gloutonne ;
- ✓ L'apprentissage par arbre de décision peut amener des arbres de décision très complexes, qui généralisent mal l'ensemble d'apprentissage (il s'agit du problème de surapprentissage précédemment évoqué). On utilise des procédures d'élagage pour contourner ce problème, certaines approches comme l'inférence conditionnelle permettent de s'en affranchir ;
- ✓ Certains concepts sont difficiles à exprimer à l'aide d'arbres de décision (comme XOR ou la parité). Dans ces cas, les arbres de décision deviennent extrêmement larges. Pour résoudre ce problème, plusieurs moyens existent, tels que la proportionnalisation, ou l'utilisation d'algorithmes d'apprentissage utilisant des représentations plus expressives (par exemple la programmation logique inductive) ;

- ✓ Lorsque les données incluent des attributs ayant plusieurs niveaux, le gain d'information dans l'arbre est biaisé en faveur de ces attributs. Cependant, le problème de la sélection de prédicteurs biaisés peut être contourné par des méthodes telles que l'inférence conditionnelle.

### **Conclusion**

Dans ce chapitre nous avons présenté, un aperçu sur les techniques de fouille de données, le processus et les tâches (classification, régression ...) de fouille de données, et leurs différents algorithmes (BN,DT,K-PPV), que nous avons utilisés, afin de résoudre la problématique de notre sujet.

Cette première étape sert comme une base permettant d'entamer la prochaine phase de l'étude conceptuelle.

## **Introduction**

Dans les chapitres précédents, nous avons présentés les différents problèmes qu'il est possible de rencontrer pendant le forage, et ensuite une explication détaillée de certaines des techniques utilisées (BN, DT, K-NN), pour la détection de certains d'entre eux, notamment les coincements.

Dans le présent chapitre, nous allons proposer une conception par affinement successif du système, en donnant son architecture générale, puis nous détaillons, en étudiant séparément chacun de ses composants. Ensuite, nous présentons les résultats obtenus.

## **1- Description du processus de prévision des coincements**

### **1-1. Le projet RTOM**

Le projet (Real Time Operating Management) est un projet de la société SONATRACH, en collaboration avec le segment SIS de la société SCHLUMBERGER.

Débuté en 2010, le projet se charge de la supervision, en temps réel, des opérations de forage, grâce à une installation réseau, qui transmet les données à partir de tous les puits connectés (42 et prochainement 84), au centre de contrôle des opérations, Remote Operation Center, (ROC), situé à la base 24 Février de SONATRACH, à Hassi Messaoud, wilaya d'Ouragla.

L'équipe chargée du projet se compose de deux groupes ; (**Front office**) et (**Back office**). [2]

✓ **The Front Office** : ce groupe se charge de tout ce qui est technique ; la liaison entre les sites de forage et le Remote Operation Center (ROC), l'émission et la réception des données, ainsi que la résolution de tous les problèmes relatifs aux matériels et logiciels. Leur mission, est de garantir la disponibilité des données à tout moment. [2]

✓ **The Back Office** : ce groupe se charge du monitoring des opérations de forage et l'élaboration des rapports de suivi. Leur mission consiste à détecter d'éventuelles problèmes, afin de les signaler aux superviseurs sur le RIG et ce, dans le but de les éviter.

L'objectif principal du projet RTOM est d'améliorer le déroulement des opérations de forage, grâce au suivi de leur avancement en temps réel. [2]

Ce suivi révèle les résultats interprétés à partir des données reçues de la base de données, chargées (récupérées) depuis le site de forage, grâce à l'architecture décrite dans le schéma ci – dessous :

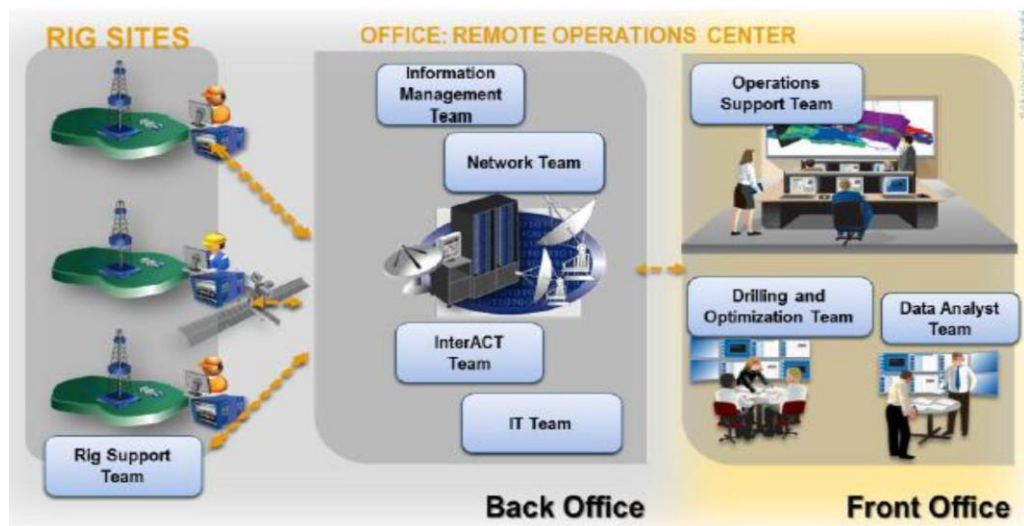


Figure III. 1 :Description de la transmission de données RTOM

À partir de réception et interprétation de ces données, les Operations Support Engineers (OSE ) peuvent détecter, grâce à leur expérience et observations, certains signaux indiquant la possibilité de rencontrer un problème notamment, le coincement de la garniture, ce qui n'est pas une tâche aisée, vu le nombre et la complexité des données reçues, en dépit de la présentation explicite de ces données en forme de graphe et courbe.

Ce travail nécessite une bonne analyse et donne rarement le résultat escompté. [2]

## 2- Architecture global du système

L'architecture globale de notre système est base sur la prédiction du possibilité de l'existence un coincement ou non, et principalement compose par deux partie principale :

(Partie d'apprentissage automatique) et (Partie de test).

Le figure suivante (Fig. III. 2) présente l'architecture global du système:

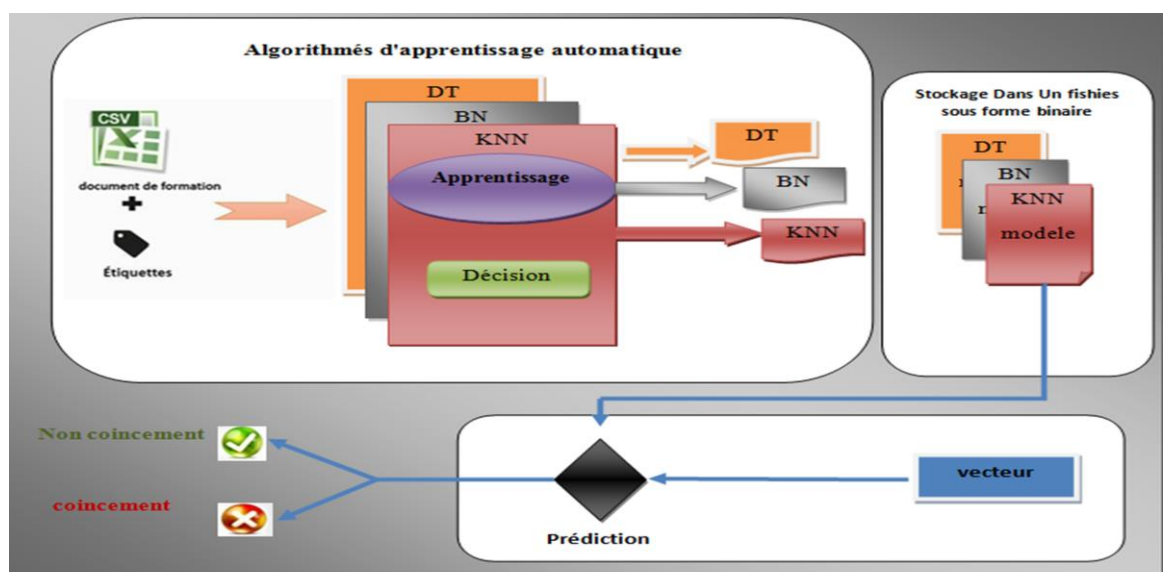


Figure III.2: L'architecture global du système

### **3- Fonctionnement global**

Notre système contient pratiquement deux parties, selon l'architecture globale. Maintenant nous allons voir comment ces composants interagissent entre eux, afin de donner une meilleure efficacité des résultats.

#### **A- La partie d'apprentissage automatique**

Le système prend comme entrée un ensemble de données sous forme de fiches Excel, et consiste, dans cette phase, à utiliser des méthodes de classification (dans notre étude, BN, DT, K-NN).

Elle se divise en deux sous phases ; Apprentissage et Test ou Décision. La première consiste à initialiser la base des modèles, autant que la deuxième consiste à assigner une classe pour chaque nouveau exemple donnée (vecteur).

- ✓ **L'Apprentissage** : cette phase consiste à initialiser ou créer la base des modèles prédictifs, considérant que le système prend les deux tiers de fichiers pour qu'il puisse apprendre sur le problème de coincement et le reste pour la décision ;
- ✓ **La Décision** : elle consiste à utiliser les caractéristiques extraites dans la phase précédente, pour attribuer une classe, en se basant sur les données de la base des modèles prédictifs.

L'objectif de cette phase est obtenu des modèles prédictifs, qu'on doit stocker dans une base de connaissance .

#### **B- La partie teste**

Le but de cette partie est de donner une solution, grâce aux modèles prédictifs, et prévoir si un coincement va survenir ou non. On doit alors tester ces données, qui sont sous forme de vecteurs, selon un modèle prédictif choisi en première partie. Et aussi, pour renforcer et clarifier la meilleure méthode, choisie dans la première partie, pour résoudre ce type de problème.

### **4- Description des données**

Les données de terrain que nous avons obtenues, relatives à un puits vertical, foré dans le champs de Hassi Messaoud ONSHORE ENAFOR 018 , au sud de l'Algérie, plus exactement, à la phase 6" (pouces), ont été utilisés dans cette étude, pour effectuer notre application.

Notre modèle de prédiction du coincement est basé sur les paramètres (mud logging) suivants (Table.III.1) :

Table.III.1 : Les paramètres de notre modèle

| Type de données             | Variable d'entrée |                                     | Unite              |
|-----------------------------|-------------------|-------------------------------------|--------------------|
| Les parameters<br>de forage | <b>RPM</b>        | Rotation per minute                 | <b>c/min</b>       |
|                             | <b>STOR</b>       | avg surface torque                  | <b>1000 ft.lbf</b> |
|                             | <b>DEPT</b>       | bit depth                           | <b>m</b>           |
|                             | <b>BPOS</b>       | block position                      | <b>m</b>           |
|                             | <b>HKLX</b>       | Hklx                                | <b>kgf</b>         |
|                             | <b>HDTH</b>       | hole depth                          | <b>m</b>           |
|                             | <b>TVDE</b>       | hole true vertical depth            | <b>m</b>           |
|                             | <b>HOOKLD</b>     | Hookload                            | <b>t</b>           |
|                             | <b>HKLD</b>       | Hookload                            | <b>kgf</b>         |
|                             | <b>PD_PADS_HY</b> | pressure drop over power drive pads | <b>psi</b>         |
|                             | <b>SPM1</b>       | pump 1 stroke rate                  | <b>1/min</b>       |
|                             | <b>SPM2</b>       | pump 1 stroke rate                  | <b>1/min</b>       |
|                             | <b>SPM3</b>       | pump 1 stroke rate                  | <b>1/min</b>       |
|                             | <b>SPPA</b>       | stand pipe pressure                 | <b>psi</b>         |
|                             | <b>TQX</b>        | Rotary Torque X                     | <b>1000 ft.lbf</b> |
| <b>WOB</b>                  | Weight on bit     | <b>t</b>                            |                    |
| <b>WOBX</b>                 | Weight on bit X   | <b>kgf</b>                          |                    |

## 5- Comparaison entre les méthodes

La prédiction du coincement qui est une valeur bien déterminé (discrète) est considéré comme un problème de classification, l'outil WEKA (explique dans le chapitre suivant) est doté de plusieurs algorithmes pour prendre en charge de ce type de problème, parmi eux nous avons choisi les plus performants d'après les résultats des expérimentations dans l'environnement WEKA et ce que nous avons trouvé dans la littérature. Citons les **BN**, et les **DT** et les **K-NN** qui feront l'objet de notre expérimentation et comparaison.

- ❖ Le critère de comparaison est le taux de corrélation, il existe deux mode d'évaluation :

### A- Division en ensemble d'apprentissage et test

Ce mode consiste à diviser la base de données en deux parties, la première (66%) pour l'apprentissage et le reste pour le test. Le modèle est bâti sur l'échantillon d'apprentissage et validé sur l'échantillon de test. L'erreur est estimée en calculant un test, une mesure ou un score de performance du modèle sur l'échantillon de test, par exemple l'erreur quadratique moyenne . [14]

**B- Validation croisée (cross-validation)**

La validation croisée (« cross-validation ») est une méthode d'estimation de fiabilité d'un modèle fondé sur une technique d'échantillonnage, on divise l'échantillon original en k échantillons, puis on sélectionne un des k échantillons comme ensemble de validation et les (k-1) autres échantillons constitueront l'ensemble d'apprentissage. On calcule comme dans la première méthode l'erreur quadratique moyenne. Puis on répète l'opération en sélectionnant un autre échantillon de validation parmi les (k-1) échantillons qui n'ont pas encore été utilisés pour la validation du modèle. L'opération se répète ainsi k fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. La moyenne des k erreurs quadratiques moyennes est enfin calculée pour estimer l'erreur de prédiction .[14]

- ❖ Dans cette évaluation nous avons sois a les critères de comparaisons suivantes:
  - ✓ Correctly classified instances.
  - ✓ Incorrectly classified instances.
  - ✓ Mean absolute error
  - ✓ Time taken to build model.

Le figure suivant montre bien les paramètres de notre système :

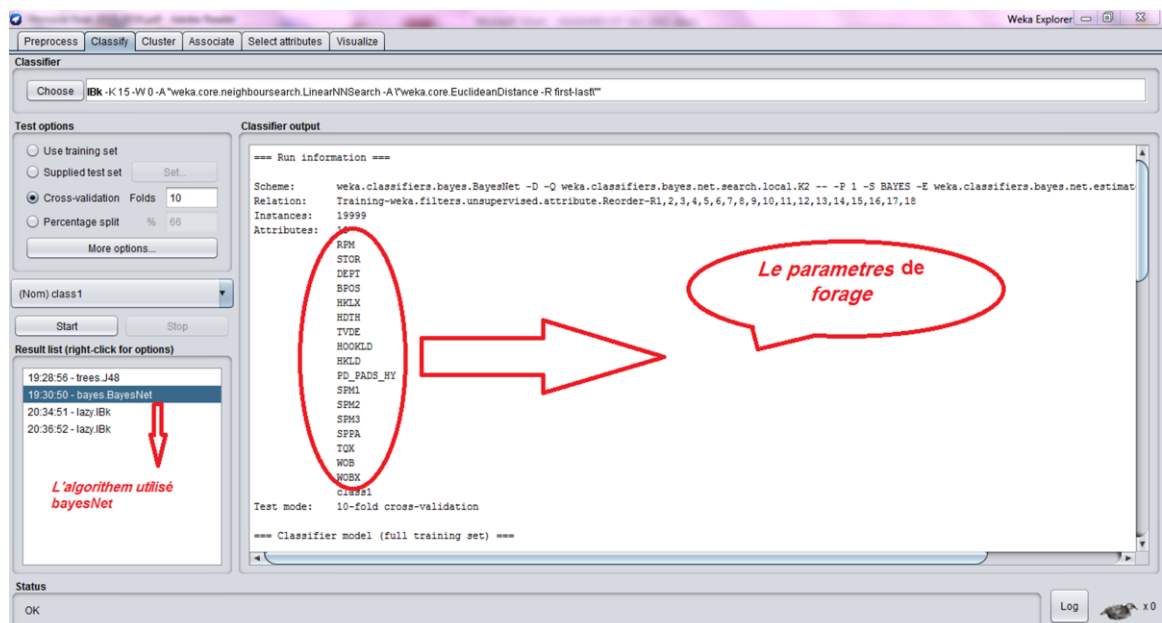


Figure III.3: l'expérimentation et l'évaluation en utilise l'algorithmme de BayesNet(BN).



Le figure suivant présente le temps nécessaire pour construire le modèle par l'algorithme de arbre de décision (DT).

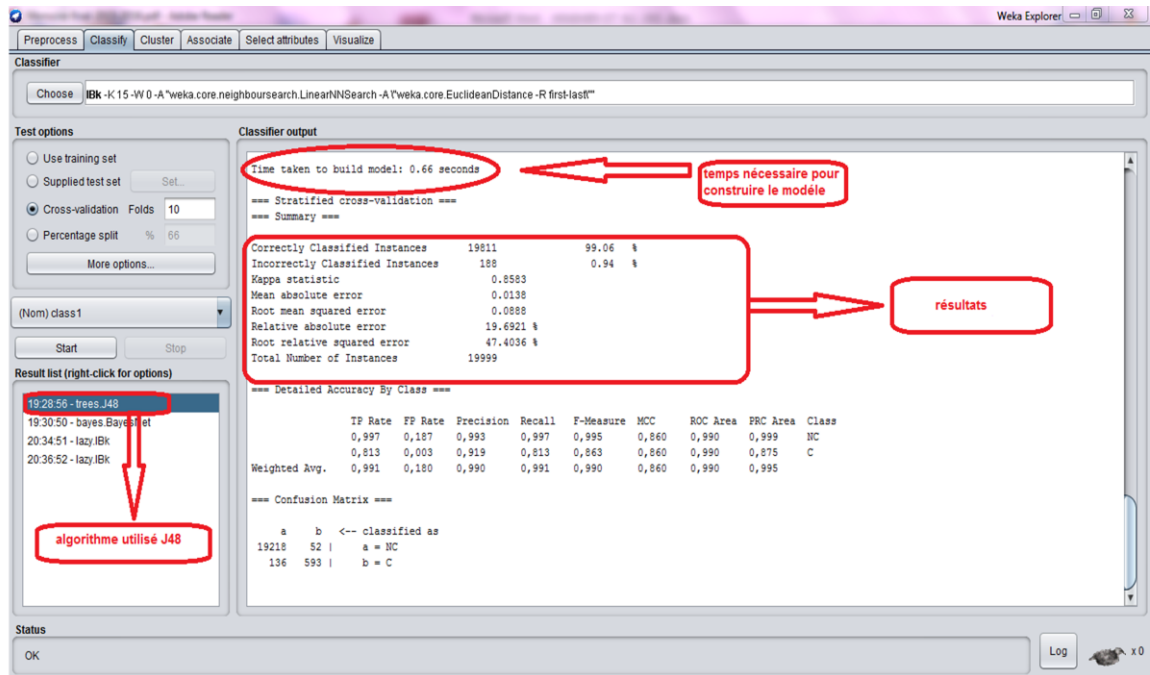


Figure III.4: l'expérimentation et l'évaluation en utilise l'algorithme de Arbre de décision (DT)

Le tableau (Tab.III.2 ) ci-dessous obtenu les résultats de comparaison entre les algorithmes de classification pour donner le modèle prédiction:

| Table.III.2 : comparaison entre les algorithmes de classification |              |               |              |              |
|---|--------------|---------------|--------------|--------------|
| les méthodes<br>les critères                                      | KNN<br>(K=3) | KNN<br>(K=11) | BN           | DT           |
| <i>Correctly classified instances</i>                             | 19745-98.72% | 19762-98.81%  | 19513-97.57% | 19811-99.06% |
| <i>Incorrectly classified instances</i>                           | 254 -1.27%   | 237-1.18%     | 486-2.43%    | 188-0.94%    |
| <i>Kappa statistic</i>  | 0.81         | 0.81          | 0.72         | 0.85         |
| <i>Mean absolute error</i>  | 0.01         | 0.01          | 0.02         | 0.01         |
| <i>Root mean squared error</i>                                    | 0.10         | 0.09          | 0.15         | 0.08         |
| <i>Relative absolute error</i>                                    | 23.03%       | 26.49%        | 36.80        | 19.69%       |
| <i>Root relative squared error</i>                                | 53.45%       | 51.41%        | 81.47        | 47.40%       |
| <i>Total number of Instances</i>                                  | 19999        | 19999         | 19999        | 19999        |
| <i>Time taken to build model</i>                                  | 0.2s         | 0.1s          | 0.3s         | 0.51s        |

Les résultats des tableaux (Table.III.2) montrent que l'algorithme Arbre de décision (DT) donne meilleur correctement classés instances prédiction (Incorrectly classified instances 0.94% ), et le pourcentage d'erreur (Relative absolute error =19.69%) .

Alors l'algorithme Arbre de décision (DT)il est le meilleur choix pour prendre en charge notre problème de prédiction.

**Conclusion**

Nous avons vu dans ce chapitre l'architecture globale du système, et détaillé chaque composant, ainsi que la structure de données qui est entre dans notre système. En fin, la comparaison entre les méthodes des classifications pour choisir la meilleure méthode qui utilise à la fin de l'étude.

Les résultats de ce chapitre seront enrichis par des détails d'implémentation dans le chapitre suivant pour réaliser notre système.

## Introduction

L'implémentation d'un logiciel devient après un enchaînement de plusieurs étapes dans le processus de développement, et son but principal est de réaliser un produit capable de résoudre les problèmes posés en utilisant des outils et des algorithmes.

Nous avons donné dans le précédent chapitre la conception de façon globale et détaillée avec les méthodes utilisées, le présent chapitre donne une vue sur le système et les outils utilisés afin d'arriver à un système fiable.

Dans ce chapitre nous allons présenter en premier temps, l'environnement du développement avec les différentes bibliothèques utilisées, ainsi que les structures de données choisies pour implémenter ce type de système. Ensuite, on va présenter les algorithmes utilisés illustré par quelque résultat obtenu, nous terminons ce chapitre par une conclusion.

### 1. Environnement et outils de mise en oeuvre

Notre système était développé sur un ordinateur de processeur Core I3 avec une RAM de 4 Go sous Windows 7, mais on peut implémenter ce système sur n'importe quel machine grâce au virtuel machine de Java.

#### 1-1. Java

Java est un langage orienté objet développé par la société Sun. La syntaxe générale du langage java est très proche de celle du langage C, et parmi ces avantages [15] :

- Le byte-code assure à Java une portabilité complète vers d'autres systèmes.
- L'importance de l'API de base qui offre tous les services de base, notamment pour la construction des interfaces graphiques.
- La troisième force du Java, c'est son adaptabilité dans des nombreux domaines, autant pour le web que pour les systèmes embarqués.



Figure IV.1:La langage de développement.

## 1-2. Weka (machine Learning)

**Waikato Environment for Knowledge Analysis (Weka)** est un ensemble d'outils permettant de manipuler et d'analyser des fichiers de données, implémentant la plupart des algorithmes d'intelligence artificielle, entre autres, les arbres de décision et les réseaux de neurones .[16]

Il est écrit en java, disponible sur le web

- ❖ Il se compose principalement [14]:
  - De classes Java permettant de charger et de manipuler les données.
  - De classes pour les principaux algorithmes de classification supervisée ou non supervisée.
  - D'outils de sélection d'attributs, de statistiques sur ces attributs.
  - De classes permettant de visualiser les résultats. On peut l'utiliser `a trois niveaux :
    - ✓ Via l'interface graphique, pour charger un fichier de données, lui appliquer un algorithme, vérifier son efficacité.
    - ✓ Invoquer un algorithme sur la ligne de commande.
    - ✓ Utiliser les classes définies dans ses propres programmes pour créer d'autres méthodes, implémenter d'autres algorithmes, comparer ou combiner plusieurs méthodes. C'est cette troisième possibilité qui sera utilisée en travaux pratiques.

Dans notre approche nous avons utilisé l'outil logiciel open source WEKA pour la comparaison des algorithmes de prédiction afin de bien justifier notre choix de l'algorithme des forêts aléatoires, pour l'implémentation le langage JAVA a été choisi avec l'environnement NETBEANS.



**Figure IV.2: L'outil WEKA**

## 2- L'application développée

Nous présentons ici des captures d'écran de l'application développée.

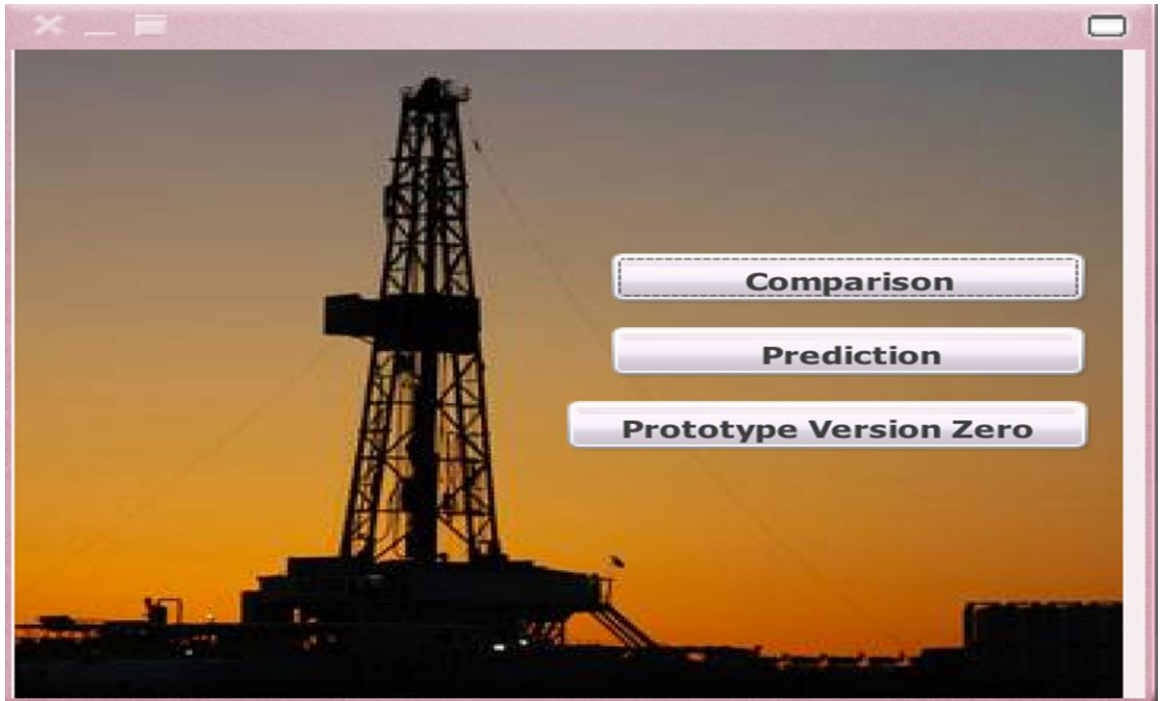


Figure IV.3:L'interface principale de l'application

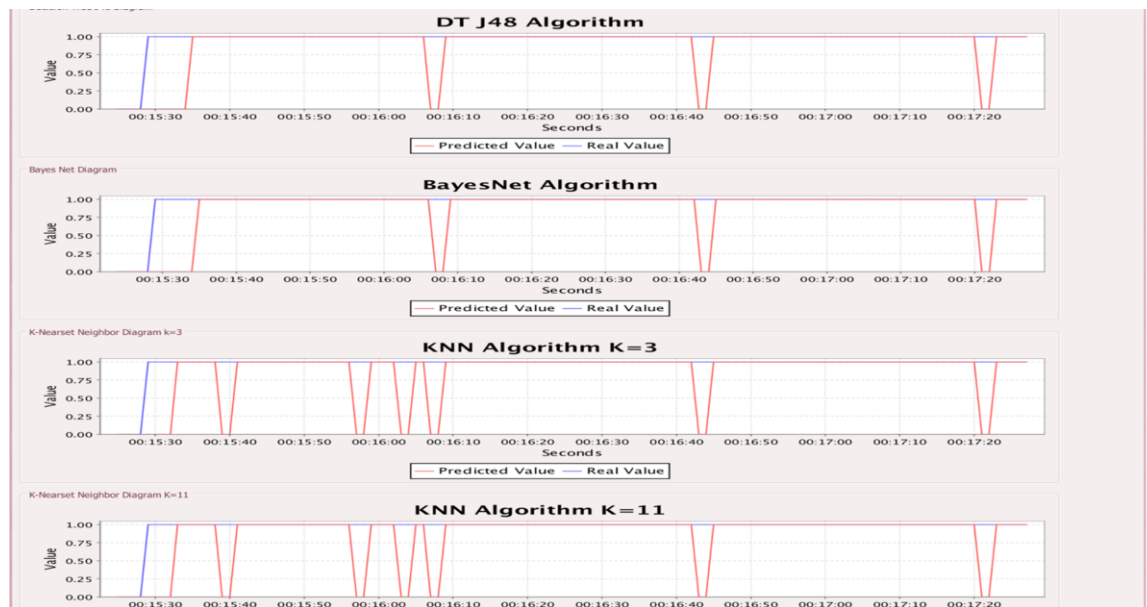


Figure IV.4:L'interface du comparaison entre les méthodes d'application

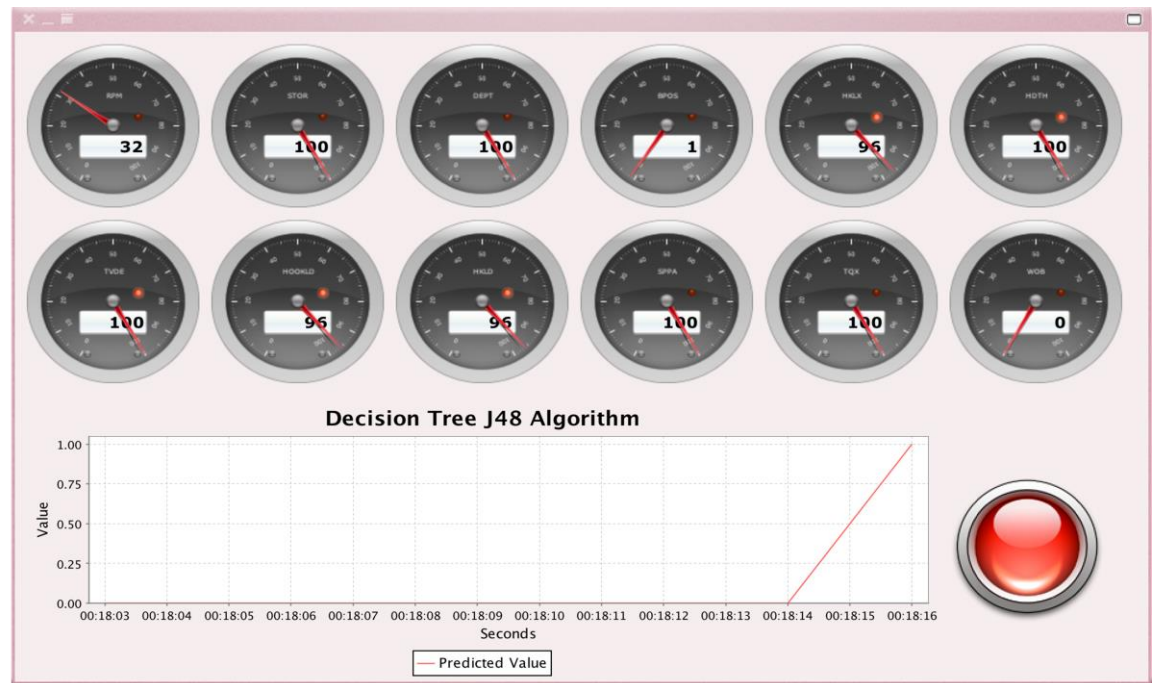


Figure IV.5:L'interface de résultat avec le meilleur algorithme (DT) coincement

### Conclusion :

Arrivé à la fin de ce chapitre, nous avons décrit les outils et environnements utilisés ainsi que l'aspect technique en relation avec le déploiement de la solution, nous avons donc codé et mit en place notre application en suivant les descriptions fournies tout au long du chapitre de conception.

Notre projet de fin d'étude s'inscrit dans le cadre d'un projet de suivi des opérations de forage pétrolier en temps réel. Notre travail consistait à concevoir et réaliser un système de prévention des coincements de la garniture de forage.

Une méthodologie de prédiction et d'optimisation de le coincement de la garniture de forage (stuck pipe) a été développée, démontrée et appliquée afin atteindre des paramètres de forage contrôlables optimales. La première tâche de cette étude est la prédiction de le coincement de la garniture de forage. Les données utilisées dans le cadre de cette étude appartiennent à des puits forés en Sud de l'Algérie.

Le coincement de la garniture de forage (stuck pipe) est prédite à l'aide les algorithme Arbre de décision (DT) pour classification qui est plus performant que les autres algorithmes de classification à savoir les KPP-V et les BN d'après les comparaisons que nous avons effectuées dans l'outil Learning-machine WEKA présentées sur les tableaux III.2 . Les résultats ont indiqué que Le coincement de la garniture de forage peut être prédit avec une haute précision, « Mean absolute error 0.0138 » après le choix des meilleurs paramètres de notre algorithme Arbre de décision (DT).

Travailler sur ce projet durant, nous a permis de capitaliser des concepts techniques en complément aux enseignements théoriques acquis. Cela nous a permis de nous rapprocher et de nous familiariser avec le monde du travail.

En perspective nous pensons exploiter des approches d'intelligence artificielle avancée comme validation croisée, Elle peut servir comme un système d'aide à la décision pour optimiser les opérations de forage en temps réel.

## *Bibliographié*

- [1] :Abdessalem BELAID : Modelisation Tridimensionnelle Du Comportement Mecanique De La Garniture De Forage Dans Les Puits A Trajectoires Complexes : APPLICATION A LA PREDICTION DES FROTTEMENTS GARNITURE-PUIT,These pour obtenir le grade de DOCTEUR DE L'ECOLE NATIONALE SUPERIEURE DES MINES DE PARIS,Anne 2005.
- [2]: MILI Saoussen, KAOUANE Housseyn : Conception et réalisation d'un système de prévention des « Stuck Pipe» HASSI MESSAOUD, Mémoire de Fin d'étude, esi 2014.
- [3] : PetoWiki: <http://petrowiki.org/PetroWiki>, site web pour les information sur le domaine pétrolier .
- [4]: P. CABENA, P. HADJINIAN, R. STADLER, J. VERHEES et A. ZANASI,Discovering Data Mining: From Concept to Implementation, Prentice Hall, Upper Saddle River, NJ, 1998.
- [5]: E-G. TALBI, Fouille de données (Data Mining) : Un tour d'horizon, Laboratoire d'Informatique Fondamentale de Lille.
- [6] : CHAMI Djazia : Une plate forme orientée agent pour le data mining, Mémoire pour l'obtention du diplôme de Magister en informatique , Université HADJ LAKHDAR – BATNA, Anne 2010
- [7] : M. J. BERRY, G. S. LINOFF, Data Mining Techniques For Marketing, Sales, and Customer Relationship, Management, Second Edition, 2004.
- [8] : M. J. BERRY, G. S. LINOFF, Mastering Data Mining: The Art and Science of Customer Relationship Management, 2000.
- [9] : D.T. LAROSE, Discovering Knowledge In Data: An Introduction to Data Mining, Central Connecticut State University, 2005.
- [10] : L'encyclopédie en ligne WIKEPEDIA 2009.
- [11] : Rapporté de [http://interstices.info/encart.jsp?id=c\\_41867&encart=3&size=600,500](http://interstices.info/encart.jsp?id=c_41867&encart=3&size=600,500).
- [12] : R.GILLERON, M. TOMMASI, Découverte de connaissances à partir de données, 2000.
- [13]: C. GROUIN, Les techniques de la fouille de données, INaLCO, 2009/2010.
- [14] : Djamil Rezki :” Système intelligent d'aide à la décision pour le pilotage d'un processus de forage pétrolier ”, memoire de magister Université BATNA 2 ,2016
- [15]: T. Ahmed « Une approche basée agents mobile pour la télésurveillance industrielle », Master en informatique, intelligence artificielle, 2011.
- [16] : weka :<http://www.fil.univ-lille1.fr/~decomite/ue/APE/tp/tp1/weka2009.pdf>, site web pour les information sur l'environnement weka.



# RÉSUMÉ

Les sociétés spécialisées dans le domaine du forage travaillent toujours dans le but de minimiser les dépenses accélérer la tâche du projet en cours afin d'augmenter ses bénéfices le maximum possible.

Le coincement de la garniture de forage (stuck pipe) est le grand problème qui peut rencontrer ces sociétés.

La capacité de prévoir les dangers les problèmes d'extraire les données, les liens et faire l'étude nécessaire afin d'aboutir aux solutions adéquates est le domaine de l'intelligence artificielle.

Notre étude Préparation d'un exemple de prédiction d'un coincement en utilisant les algorithmes de la fouille des données .

Mots clés :

Forage , coincement de la garniture de forage, prédiction, data-mining, algorithme, intelligence artificielle

## المخلص

تعمل الشركات المتخصصة في مجال التنقيب إلى إنقاص تكلفة الحفر وتسريعه .  
الأنابيب العالقة أثناء عملية الحفر هي من أكبر المشاكل المتكررة التي تواجهها الشركات الأمر الذي يترتب عليها عدة خسائر مالية وضياع للوقت.

الهدف الأساسي لمدراء هذه الشركات هو التقليل من علق أنابيب الحفر (الأنابيب العالقة).  
تنقيب واستخراج البيانات هو مجال جديد من مجالات الذكاء الاصطناعي يستعمل في حل مثل هاته المشاكل من خلال القدرة على فحص البيانات وإيجاد علاقات بينها وتعليم الآلة مباشرة من خلال البيانات الميدانية الفعلية.  
عموما، يتم إصلاح اضطرابات الأنابيب العالقة بعد الحوادث وذلك باستخدام التقنيات القياسية .هنا، ونحن نحاول التنبؤ بأسباب حدوث هذه المشاكل لتجنب المخاطر وتكاليف الحفر المفرطة .إذا تم تحديد هذه المخاطر في وقت مبكر، ويمكن توفير أفضل الحلول للحد من العواقب المرتبطة بها.

والغرض من دراستنا هو محاولة اعداد نموذج يتنبأ بلتصاق أنابيب الحفر باستخدام خوارزميات التنقيب عن البيانات.

الكلمات الأساسية:

التنقيب ، التصاق أنبوب حفر الآبار، التنبؤ، استخراج البيانات، معالجة البيانات، الذكاء الصناعي

# ABSTRACT

The factories specialized in oil drilling field make to reduce the cost of digging and its acceleration.

The stopped pipe during the digging operation its the biggest frequent problem wich the factories face make them financial loss and waste of time.

The main aim of this factories' bosses is the reduce of (stock pipe) the drilling and finding out the information is a new field of the artificial intelligent fields used to solve this problem through the ability to control the information and find the relation between learning and machine learning directly through the actual information from the field.

Generally, the trouble of the stopped pipe is fixed after the accident and that's through the use of the limited techniques.

Here we try to predict the problem so that we can avoid the danger and it's financial loss. if we can detect this problem before happening we can provide the best solution to limit the consequence. the purpose of our studies is to make a predicted model that can predict the problem oft the stock pipe by the use of algorithme.

## **Keywords:**

drilling, stuck pipe, prediction, data –mining, algorithme, artificial intelligence