



**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**  
**Université Mohamed Khider – BISKRA**  
**Faculté des Sciences Exactes, des Sciences de la Nature et de la Vie**  
**Département d'informatique**

N° d'ordre : SIOD 26 /M2/2018

## **Mémoire**

Présenté pour l'obtention du diplôme de master académique en

# **Informatique**

Parcours SIOD (Informatique Décisionnelle et Optimisation)

---

## **La classification associative de modèles du non recouvrement par la CNAS par la technique des motifs fréquents**

---

**Par :**

**Othmani Marabout Amima**

**Soutenu le 26/06/2018, devant le jury composé de :**

- |                             |             |                   |
|-----------------------------|-------------|-------------------|
| - Mr Meadi Med Nadjib       | Grade M C B | <b>Président</b>  |
| - Mr Zerarka Mohamed Faouzi | Grade M C A | <b>Rapporteur</b> |
| - Mr Telli Moutia           | Grade M C B | <b>Examineur</b>  |

## ملخص

من بين المهام المسندة لصندوق الضمان الاجتماعي للعمال الأجراء (CNAS) تحصيل الاشتراكات، و يعتبر هذا الأخير هام و ضروري من أجل تمويل الصناديق الاجتماعية الأخرى. و مشكل عدم تحصيل الاشتراكات يمثل عائق و خطر على سير و استمرارية هذه الهيئة. ويرجع هذا الإشكال لعدة أسباب من أهمها: الصعوبة المالية التي تواجه أرباب العمل، إهمال تسديد المستحق من طرف رب العمل، صكوك بدون رصيد، حجم المعلومات التي يتعين معالجتها. و يهدف مساعدة متخذي القرار للصندوق الضمان الاجتماعي و تقليل مخاطر عدم تحصيل الاشتراكات التي تؤثر على السير الغير حسن لهذه الهيئة ، وضعنا هدفا يتمثل في وضع نماذج و تحقيق نظام التصنيف الترابطي للنماذج بواسطة واحدة من أكثر الطرق تكييفها المعروفة باسم " تقنية الربط عن طريق الدوافع المتكررة "

كلمات البحث : CNAS، عدم تحصيل اشتراكات الضمان الاجتماعي، استخراج البيانات، التصنيف الترابطي بدوافع متكررة

## Résumé

Parmi les fonctions assignées à la CNAS, le recouvrement de la cotisation est considéré comme important et essentiel pour le financement de différentes caisses sociales. Le problème du non recouvrement (crédit non payé) représente un véritable obstacle et danger pour la gestion et la continuité de cet organisme. Ce problème est due à plusieurs causes dont les principales sont : la difficulté financière des employeurs, les négligences de paiement des dues d'un employeur, les chèques impayés et le volume d'information à traiter.

Dans le but d'aider les décideurs de la CNAS et de **diminuer les risques du non recouvrement** qui perturbe la gestion de cet organisme pivot de la sécurité sociale, Nous nous fixons un objectif de la modélisation et la réalisation d'un système de **classification associative des modèles** par l'une des méthodes la plus adaptée dite « **la technique d'association par les motifs fréquents** ».

**Mots Clés** : CNAS, non recouvrement social, data mining, la classification associative par les motifs fréquents.

## Abstract

Among the fonctions assigned to the CNAS, the collection of the contribution is considered as important and essential for the financing of the other insurance agency. The problem of *the evaded contributions* ( unpaid credit) represents a real obstacle and a danger for the management and the continuity of *social security body*. This problem is due to a number of causes, the most important of which are: *the financial difficulties faced* by employers, the negligence of payment of the due of an employer, the unpaid checks and the volume of information to be treated.

With the aim of helping the decision-makers of the CNAS and decreasing the risks of the *the evaded contributions* which perturbs the management of this body pivot of the Social Security, We set a goal of the modelling and the realization of a system of associative classification of the models by one of the most adapted methods known as « the technique of association by the frequent motives ».

**Keywords:** CNAS, the evaded contributions, data mining, the associative classification by the frequent motives.

## Remerciements

Avant tous, je remercie mon dieu **Allah** qui m'a donné la volonté, le courage et la patience d'achever ce modeste travail.

En premier lieu, je tiens à remercier mon encadreur du mémoire, **Mr Zerarka Mohamed Faouzi** de m'avoir proposé ce sujet, pour les conseils prodigués et la confiance qu'il m'a témoigné.

De même je remercie les membres de jury qui m'ont fait l'honneur de juger ce travail.

Je n'oublie pas mes parents pour leur contribution, leur soutien et leur patience.

Enfin, j'adresse mes plus sincères remerciements à tous mes proches et ami(e)s, qui m'ont toujours encouragée au cours de la réalisation de ce mémoire.

Merci à tous et à toutes.

## Tables des matières

<b>Introduction général</b> .....	1
<b>Chapitre I : La Présentation du cadre d'étude</b>	
1 Présentation de système de la sécurité sociale en Algérie.....	3
2 La Présentation de la CNAS .....	4
2.1 Le statut de la CNAS .....	4
2.2 La structure et l'organisation de la CNAS .....	4
2.3 Les attributs de la CNAS .....	5
2.4 Le financement de la CNAS .....	6
3 Recouvrement des cotisations sociales.....	6
3.1 Cadre organisationnel .....	6
3.2 Cadre juridique.....	7
3.2.1 Les obligations des assujettis.....	7
3.2.2 Le contrôle employeur .....	9
3.2.3 Le décompte des cotisations.....	9
3.2.4 Le mécanisme du recouvrement .....	9
3.3 Cadre fonctionnel .....	10
3.3.1 Immatriculation Employeur et assuré .....	11
3.3.1.1 Immatriculation Employeur.....	11
3.3.1.2 Immatriculation des assurés sociaux .....	11
3.3.2 Recouvrement des cotisations .....	12
3.3.2.1 Facturation.....	12
3.3.2.2 Encaissement .....	13
3.3.3 Contentieux du recouvrement.....	13
4 Conclusion .....	14
<b>Chapitre II : La Classification</b>	
1 Le processus « ECD » .....	19
1.1 Définition.....	19
1.2 Les étapes du processus « ECD ».....	19
2 L'acquisition des données.....	20
3 Le prétraitement des données.....	21
4 Le Data mining.....	21
4.1 Les modèles du data mining.....	21
4.1.1 Modèle de data mining supervisé .....	21
4.1.2 Modèle de data mining non supervisé.....	22

5	Les règles d'association.....	23
5.1	Définitions.....	24
5.2	Les étapes d'extraction des règles d'association .....	25
5.2.1	Préparation des données .....	25
5.2.2	Recherche des Itemsets fréquents .....	25
5.2.3	Production des règles d'association.....	26
5.3	L'algorithme Apriori .....	26
5.3.1	Définition .....	26
5.3.2	Le principe de l'algorithme .....	26
5.4	Définition d'une règle solide.....	27
6	Evaluation et interprétation du modèle.....	28
7	Conclusion .....	28
<b>Chapitre III : la Conception du système</b>		
1	L'objectif .....	30
2	La conception du système.....	30
2.1	Architecture générale du système.....	30
2.2	Architecture Détaillée .....	32
2.2.1	Sous système d'acquisition des données cibles .....	32
2.2.1.1	Les entrées du sous-système de l'acquisition des données cibles.....	33
2.2.1.1.1	<b>Acquisition des données des Employeurs</b> .....	33
2.2.1.1.2	<b>Acquisition des données des déclarations et encaissements</b> .....	34
2.2.1.2	Les sorties du composant .....	37
2.2.2	Le sous système Prétraitement .....	37
2.2.2.1	Les entrées du Sous système.....	38
2.2.2.2	Composants interne du sous système .....	38
2.2.2.3	La sortie du Sous système .....	41
2.2.3	La classification.....	41
2.2.4	Le sous-système de la visualisation .....	45
3	Conclusion .....	45
<b>Chapitre IV : l'implémentation et la validation du système</b>		
1	L'environnement matériel/logiciel .....	53
1.1	Matériel.....	53
1.2	Logiciel.....	53
1.2.1	Système d'exploitation utilisé .....	53
1.2.2	Langage de Programmation utilisé .....	53

1.2.3	L'environnement utilisé .....	54
1.2.4	Les SGBDS utilisé .....	54
2	L'implémentation .....	55
2.1	Sous-système d'acquisition des données cibles .....	55
2.2	Sous-système de prétraitement .....	57
2.2.1	Les composants internes .....	58
2.2.2	La sortie du sous-système .....	61
2.3	Sous-système de la classification .....	64
3	Les tests et la visualisation des résultats .....	65
4	Conclusion .....	72
	<b>Conclusion Générale</b> .....	<b>73</b>
	<b>Références bibliographiques</b> .....	<b>74</b>



## Liste des tableaux

**Tableau 01** : Répartition du taux de cotisation

**Tableau 02**: les avantages et les inconvénients de quelques algorithmes de la classification.

**Tableau 03** : Définition d'une règle solide.

## Listes des figures

- Figure01** : Le processus « ECD »
- Figure02**: Architecture générale
- Figure03** : Le Sous-système de l'acquisition des données cibles
- Figure04** : Architecture de l'acquisition des données Employeurs
- Figure05** : Architecture d'acquisition des données des déclarations
- Figure06** : Architecture de l'encaissement des déclarations (obligations)
- Figure07**: Diagramme de classe de « Acteurs »
- Figure08**: Architecture du sous-système Prétraitement
- Figure09**: Architecture du Recouvrement
- Figure10**: Architecture du Contentieux
- Figure11**: Diagramme de classe de « Débiteurs »
- Figure12**: Architecture du sous-système de la Classification
- Figure13** : Le processus de l'extraction des Itemset fréquent
- Figure14** : L'architecture du sous-système de la visualisation
- Figure15**:Exemple d'un Code Java.
- Figure16**:NETBEANS.
- Figure17**: ConnexionBase.java
- Figure18**: Connexion\_Postgresql.java
- Figure19**: « Acteurs » avant mise à jour
- Figure20**:Code source de la méthode Mettre\_a\_jour\_acteurs ()
- Figure21** : « Acteurs » après mise à jour
- Figure22**: Recouvrement avant calcul\_data
- Figure23**:le code source de la méthode calcul\_data () pour le recouvrement
- Figure24** : Le « Contentieux » avant calcul\_data
- Figure25**: le code source de la méthode calcul\_data() pour le « Contentieux »
- Figure26**: Le « Contentieux » après calcul\_data
- Figure27**: Le « Débiteurs » avant mise à jour
- Figure28**: le code source de la méthode mettre\_à\_jour\_débiteurs() du « Débiteurs » Première Partie
- Figure29**: le code source de la méthode mettre\_à\_jour\_débiteurs () du « Débiteurs » deuxième Partie
- Figure30**: « Débiteurs » après mise à jour
- Figure31**: Consultation des individus sélectionnés
- Figure32**: l'entrée des valeurs des seuils minimaux
- Figure33**: L'extraction des Itemsets fréquent et les Règles Solides
- Figure34**: L'interface principale du système proposé :La classification associative de Modèles du non recouvrement par la CNAS par la technique des motifs fréquents.
- Figure35**: sous système d'acquisition des données cibles
- Figure36**: les Acteurs
- Figure37**: sous système de prétraitement
- Figure38**: Le Recouvrement

**Figure39:** Le Contentieux

**Figure40:** les valeurs entrées de seuils minimaux

**Figure41:** l'extraction Item fréquent et les règles solide

**Figure42:** le sous système de la visualisation

**Figure43:** Histogramme des règles

**Figure44:** Courbe des règles

## Liste des abréviations

<b>AT/MP</b>	-Accidents du travail et Maladie professionnel
<b>BTPH</b>	-Secteur du bâtiment, des travaux publics et de l'hydraulique
<b>CACOBATPH</b>	-Caisse nationale des congés payés et du chômage intempéries des Secteurs BTPH
<b>CASNOS</b>	-Caisse nationale de sécurité sociale des Non-salariés
<b>CNAC</b>	-Caisse nationale de l'assurance chômage
<b>CNAS</b>	-Caisse nationale de sécurité sociale des salariés
<b>CNASAT</b>	-Caisse nationale des assurances sociale, des AT/MP
<b>CNR</b>	-Caisse nationale de retraite
<b>DAS</b>	-Déclaration annuelle des salaires et des salaries
<b>FNPOS</b>	-Fonds national de péréquation des œuvres sociales

## Introduction générale

Habituellement, le système de sécurité sociale Algérien représente des mesures et organisation pour garantir l'assurance des individus contre les risques sociaux. Pour assurer une grande couverture de différentes branches de risques sociaux, le système de sécurité sociale repose sur cinq (05) caisses (la CNAS, la CNR, la CNAC, la CASNOS et la CACOBATPH).

La Caisse Nationale des Assurances Sociale des Travailleurs Salaries (CNAS) est considérée comme l'organisme pilier de ce système. En réalité, elle protège presque 80% de la population contre la majorité des risques (maladie, AT/MP, invalidité, maternité et décès). Elle comptabilise des revenus substantiels sous forme des prestations diverses. Pour que la CNAS puisse assurer son financement et celui de la CNR et la CNAC, le recouvrement des cotisations constitue la source principale du financement des risques qu'elles couvrent.

Le recouvrement consiste en la gestion de la perception de sommes dues à la CNAS par une personne physique ou morale dit « Employeur ». Le suivi de cette opération importante, vitale et délicate d'un volume important de données (amas de données) exige un traitement nécessitant beaucoup d'habileté, de compétence, de précision et **un temps de réponse rapide**. La mise en place d'un système informatique d'aide à la décision pour les fonctions assignées à la structure du recouvrement s'avère nécessaire pour faciliter le travail en termes de métier « **Le suivi du non recouvrement** ».

Notre système proposé est l'élaboration et la réalisation d'un nouveau modèle du non recouvrement par l'extraction des connaissances par les données se basant sur une classification associative.

Pour pouvoir exprimer ce projet de fin d'études, nous avons utilisé quatre (04) chapitres. Dans le premier chapitre, nous avons étudié et définie la structure et le fonctionnement de l'organisme CNAS ainsi que le concept du recouvrement des cotisations considéré comme besoins essentiel du projet. Dans le chapitre deux nous avons présenté le processus de l'extraction des connaissances à partir des données « ECD », ses étapes, essentiellement la fouille de données « Data mining » et particulièrement la classification associative « Règles d'association et les motifs fréquents ». Dans le troisième chapitre, nous avons proposé une étude conceptuelle de notre système. Le quatrième chapitre est réservé à la description de l'implémentation du système conçu, des tests et de la validation du modèle obtenue.

Enfin, nous terminons notre expression par une conclusion générale ainsi que des perspectives futures qui seront dégagés de ce travail.

# *Première partie*

---

# Chapitre I :

## La Présentation du cadre d'étude

---

Ce chapitre, parlant de l'historique et l'organisation du système de la sécurité sociale en résumé, La CNAS est considéré comme une caisse mère de ce système, sa structure, ses objectifs et son financement sont présenté consécutivement.

Pour conclure ce chapitre, en abordant l'organisation, la structure, les aspects juridiques et fonctionnels du cœur de métier du recouvrement des cotisations de la sécurité sociale.



## **1 Présentation de système de la sécurité sociale en Algérie**

Depuis l'indépendance de l'Algérie (1962), un système de sécurité social fragmenté était existé, composé de plus de 11 diverses régimes en environ 20 d'organisme chargées de la gestion du régime général du secteur non agricole, des caisses de régimes spéciaux, 29 caisses du régime agricole, 13 caisses de secours minier qui étaient différents dans le financement, la nature et le niveau des prestations ainsi que le mode de gestion.

Dans les années 1970 la réforme du système de sécurité sociale est posée dans les buts de l'unification des régimes, l'uniformisation des avantages et l'extension des bénéficiaires, une commission nationale de réforme s'installa en février 1975 mais ce projet ne pouvait pas s'adapter à cause des résistances des ministères de tutelle des régimes particuliers.

C'est avec les lois de «**juillet 1983** » (lois 83-11/12/13/14/15) que le système de l'assurance sociale est devenu un facteur de progrès social et un instrument privilégié de la politique de la solidarité nationale et ce système est basé sur Unification des régimes, Uniformisation des avantages, Unicité du financement et Unification de l'administration du système. En 1985 les organismes de la sécurité sociale ont été regroupés en deux caisses la Caisse Nationale des Assurances Sociales des Accidents de Travail et Maladies Professionnelles (**C.N.A.S.A.T.**) et la Caisse Nationale de Retraite (**C.N.R.**).

Au cour de l'année 1992 la CNASAT changeait appellation par Caisse Nationale des Assurances Sociales des Travailleurs Salariés (**C.N.A.S.**), et il y avait aussi la Caisse des Assurances Sociales des Non-salariés (**C.A.S.N.O.S.**)

En 1994, la Caisse Nationale d'Assurance Chômage (**C.N.A.C.**) a été constitué et cet organisme chargé de l'indemnisation des salariés ayant perdu leur emploi pour raison économique ; elle gère un fonds de crédit pour la création des entreprises, et en 1997 la création de la Caisse Nationale des Congés Payés des Travailleurs du Bâtiment (**C.A.C.O.B.A.T.P.H.**) qui répond aux nécessités d'organiser une gestion spécifique des congés payés et des indemnités de chômage-intempéries des secteurs du BTPH.

Depuis, le système de la sécurité sociale en Algérie s'articule les cinq (05) caisses qui assurent une couverture contre les risques liés à la maladie, la maternité, l'invalidité, le décès, la vieillesse, les accidents du travail, les maladies professionnelles, les prestations familiales, et le chômage.

Le régime de la protection n'est pas seulement applicable aux personnes exerçant une activité professionnelle mais aussi les apprentis, les bénéficiaires des emplois d'attente, les étudiants, les stagiaires de la formation professionnelle, les handicapés, les moudjahidines et les titulaires d'avantages de sécurité sociale (pensionnés et rentiers). Pour le remboursement de tous les frais sont donc essentiellement des cotisations à la charge des employeurs et des travailleurs[1]. Le financement de la sécurité sociale en Algérie est assuré par ces propres ressources, et pour que ce système survive dépend à l'efficacité du recouvrement des cotisations.

## **2 La Présentation de la CNAS**

### **2.1 Le statut de la CNAS**

La CNAS est un établissement public à gestion spécifique en application de l'article 49 de la loi n°88-01 du 12 janvier 1988, elle est dotée de la personnalité morale et de l'autonomie financière, et réputée commerçante dans ses relations avec les tiers.[site [www.cnas.dz](http://www.cnas.dz)], Conformément aux dispositions de l'article 78 de la loi n° 83-11 du 2 juillet 1983 susvisée, la caisse est placée sous la tutelle du ministre chargé de la sécurité sociale[4].

### **2.2 La structure et l'organisation de la CNAS**

En termes d'organisation administrative, la CNAS obéit aux dispositions du Décret exécutif n° 92-07 du 4 janvier 1992, portant statut juridique des caisses de sécurité sociale, organisation administrative et financière de la sécurité sociale [4].

La Direction générale de La CNAS située à Alger (Ben Aknoun), elle dispose des services centraux, des agences locales dont la compétence territoriale et le nombre sont fixés par arrêté du ministre chargé de la sécurité sociale[4].

La CNAS dispose 49 Agences Wilaya qui ne sont pas autonomes financièrement, ils sont mis sous l'autorité d'agents de direction auxquels le directeur général et l'agent chargé des opérations financières peuvent déléguer sous leurs responsabilités une partie de leurs pouvoirs. Il existe trois catégories d'agence (Agence de 1ère catégorie compte plus de 200 000 assurés, 2ème catégorie compte entre 100 000 et 200 000 assurés et la 3ème catégorie compte moins de 100 000 assurés). En plus La CNAS possède 826 structures de paiement réparties en : 356 centres de paiement, 401 antennes de paiement, 69 correspondances locales [www.cnas.dz].

L'article 34 du Décret exécutif n° 92-07 du 4 janvier 1992 détermine le conseil d'administration et les agents de direction, ils comprennent le directeur général, le directeur général adjoint, l'agent chargé des opérations financières, les directeurs centraux ainsi que les directeurs d'agences.

### **2.3 Les attributs de la CNAS**

- Gérer les prestations des assurances sociales (maladie, maternité, invalidité, décès) ainsi que des accidents du travail et maladies professionnelles (AT/MP).
- Gérer les allocations familiales pour le compte de l'Etat.
- Assurer le recouvrement, le contrôle et le contentieux du recouvrement des cotisations destinées au financement des prestations.
- Le contrôle et la gestion du volet contentieux relatif à la collecte des souscriptions visant à financer des rendus
- Attribuer un numéro d'immatriculation national aux assurés sociaux et aux employeurs.
- Contribuer à promouvoir la politique de prévention des AT/MP et gérer le fonds de prévention AT/MP.
- Gérer les prestations dues aux personnes bénéficiaires des accords bilatéraux de sécurité sociale.
- Exercer le contrôle médical des bénéficiaires.
- Entreprendre des actions en vue de faire bénéficier les travailleurs et leurs ayants droit de prestations collectives, sous forme de réalisations à caractère sanitaire et social.
- Gérer le fonds d'aide et de secours.
- Conclure des conventions avec les prestataires de soins.
- Assurer l'information des bénéficiaires et des employeurs. [[www.cnas.dz](http://www.cnas.dz)]

## 2.4 Le financement de la CNAS

Le recouvrement des cotisations destinées au financement des prestations de toutes les branches soit celles versées par la CNAS, par la CNR et par la CNAC. La CNAS assure également le recouvrement des cotisations dédiées au financement du fonds national de péréquation des œuvres sociales (FNPOS), chargé de la promotion du logement social. Le montant de cotisation est opéré à partir de taux fixés par voie réglementaire et sont à la charge des employeurs et des salariés.

Depuis le 1<sup>er</sup> Janvier 1999, « le taux plein » est fixé à 35% pour l'ensemble des salariés des secteurs d'activité car il représenté le taux le plus élevé et couvre l'ensemble des risque de sécurité sociale et il est réparti comme suit : 25% sont à la charge de l'employeur ,9% sont à la charge du salarié, et 1% sont à la charge des œuvres sociale.[2][3][5]

Partis	Employeur	Salarié	œuvres sociales	Total
Branches				
Assurances sociales	11,50%	1,5 %	-	<b>13 %</b>
AT/MP	1,25 %	-	-	<b>1,25 %</b>
Retraite	11 %	6,75 %	<b>0,50 %</b>	<b>18,25 %</b>
Assurance chômage	1 %	0,5 %	-	<b>1,5 %</b>
Retraite anticipée	0,25 %	0,25 %	-	<b>0,5 %</b>
Logement social	-	-	0,50 %	<b>0,50 %</b>
Total	<b>25 %</b>	<b>9 %</b>	<b>1 %</b>	35%

**Tableau01 : Répartition du taux de cotisation**

## 3 Recouvrement des cotisations sociales

### 3.1 Cadre organisationnel

La CNAS est l'une des organismes de la sécurité sociale qui se charge de fonction du recouvrement des cotisations pour son propre compte et pour celui des autres caisses à l'effet de mettre à leurs disposition les fonds au financement des risques qu'elles couvrent. Cette activité est organisé au niveau de la direction centrale qui est chargé par les fonctions du pilotage et sont représenté par l'orientation et l'accompagnement des sous-directions au niveau des agences wilaya dans l'exercice de leurs missions.

Chaque sous directions de recouvrement (agence wilaya) est constitué d'un ensemble des services et chaque service représente une sous-structure de la sous-direction, ces services assurent un ensemble des fonctions qui sont:

- Immatriculation des employeurs et ainsi leurs salariés.
- la tenue et la mise à jour des fichiers des assujettis avec le suivi et le contrôle de l'exécution des obligations des employeurs.
- la mise à disposition de chaque caisse de sécurité sociale des fonds nécessaire pour le paiement des prestations et les frais de fonctionnement dans la limite de quote-part.
- le recouvrement des cotisations de la sécurité sociale destinée au financement de la sécurité sociale des travailleurs salariés, le contentieux est relatif au recouvrement des cotisations de sécurité sociale.

### **3.2 Cadre juridique**

Le système du recouvrement de la sécurité sociale est défini par un dispositif juridique constitué de plusieurs textes réglementaires, notamment la loi 83-14 du 2 juillet 1983, modifiée et complétée relative aux obligations des assujettis en matière de sécurité sociale et la loi 08-08 relative au contentieux en matière de sécurité sociale.

Le recouvrement repose sur la notion de déclaration avec un droit de contrôle postérieur à l'intention de l'organisme chargé du recouvrement. Les articles 3 et 4 de la loi 83-14 définissent les employeurs comme assujettis : « les personnes physiques ou morales occupant un ou plusieurs travailleurs, quelles que soient la nature juridique, la durée et la forme de la relation de travail » et aussi « Sont également considérés comme employeurs, les particuliers qui emploient des personnes pour leurs propres comptes, en qualité de gens de maison, chauffeurs, femmes de ménage, couturières, lingères et infirmières ».[9]

#### **3.2.1 Les obligations des assujettis**

L'employeur joue un rôle important s'agissant d'assujettissement et de recouvrement des cotisations. Selon la loi 83-14, il doit dans des délais prescrits :

- La déclaration d'activité : la déclaration de l'activité de l'employeur doit être effectuée dans les 10 jours qui suivent la date de recrutement du premier salarié. « Art 6 de la loi 83-14 »[9]
- la déclaration des salariés : c'est obligatoire d'affilier à la sécurité sociale toute personne quel que soit sa nationalité exerçant une activité salariée ou assimilé dans les jours 10 qui suivent sa date de recrutement. « Art 8 de la loi 83-14 » [9].
- la déclaration et le versement des cotisations : Le versement des cotisations de sécurité sociale incombe à l'employeur, ce dernier est tenu effectuer ses déclarations et versements dans les 30 jours qui suit soit la fin de chaque trimestre pour les employeurs occupants moins de (10) salariés soit de chaque mois pour ceux occupant plus de (9) salariés. « Art 21 de la loi 83-14 » [9].
- la déclaration annuelle des salaires et des salariés (DAS) : chaque employeur est tenu d'envoyer ou présenter à l'organisme de la sécurité sociale, un fichier contenant la liste nominative des salariés et les salaires, dans les trente (30) jours qui suivent la fin de chaque année civile. « Art 14 de la loi 83-14 » [9]
- la déclaration d'accident de travail : L'employeur est tenu de déclarer tout accident de travail dans les quarante-huit(48) heures de son événement. « Art 13 de la loi 83-13 du 2 juillet 1983 ». [8]

Le Non-respect de ses obligations conduit à l'application de sanctions administratives et pénales :

- les sanctions administratives : ce sont des amendes prononcées par la CNAS suivant la loi 83-14 et elles représentent :
  - ✓ la majoration de retard .
  - ✓ la pénalité de retard .
- Les sanctions pénales : on a trois cas
  - ✓ Défaut de déclaration de salariés.
  - ✓ Défaut de versement de la quote-part ouvrière.
  - ✓ Pour chèque impayé.

Suite à la loi 83-14 et la loi 83-13, le tableau suivant condense l'ensemble des obligations, les délais et les sanctions. [9]

### 3.2.2 Le contrôle employeur

Les articles de 28 à 38 de la loi 83-14 disposent, Le contrôle employeur s'effectue par des agents de l'organisme de la sécurité sociale agréés par le ministre chargé de la sécurité sociale et dûment assermentés le contrôle des assujettis se fait dans un cadre législatif et de la réglementation de la sécurité sociale, il s'agit de vérifier que l'ensemble des déclarations de l'employeur sont exactes et qu'il accomplit tous ces obligations .

### 3.2.3 Le décompte des cotisations

L'assiette des cotisations de la sécurité sociale est à la charge de l'employeur qui est constitué de l'ensemble des éléments des salaires ou de revenus proportionnelles aux résultats du travail à l'exception des prestations à caractère familiale, des primes de départ, des indemnités pour conditions de vies particulières [5].

Les salaires soumis aux cotisations ne doivent pas être en aucun cas inférieure au SNMG et pour les catégories particulières, l'assiette de cotisation est le SNMG en vigueur.

$$\text{assiette} = \sum \text{des salaires de poste}$$

$$\text{Montant des cotisations} = \text{assiette} * \text{taux des cotisations}$$

### 3.2.4 Le mécanisme du recouvrement

Le recouvrement est une activité exécutée réglementée pour que la sécurité sociale puisse la pratiquer elle doit respecter un certain nombre d'étapes obligatoires pour pouvoir procéder au recouvrement. Selon les articles de 57 à 73 de la loi 83-15, le recouvrement peut se distinguer sous trois(3) formes : recouvrement réglementaire, recouvrement amiable et forcé.

- **recouvrement réglementaire** : c'est le règlement des montants des cotisations déclaré par l'employeur dans les délais prescrits.

- **recouvrement amiable** : c'est le règlement des montants des cotisations par calendrier de paiement, pour les employeurs de bonne foi connaissant des difficultés financières.
- **recouvrement forcé** : pour recouvrer les cotisations impayées, la CNAS recourt à des procédures particulières à l'encontre des assujettis débiteurs. L'organisme de la sécurité sociale créancier doit avertir préalablement l'assujettit par une mise en demeure en l'invitant de régler sa situation (paiement des cotisations principale, majoration de retard et pénalité de retard), avant de procéder une des procédures du recouvrement forcé : l'opposition sur compte courant postaux/bancaires, contrainte et rôle.
  - ✓ **l'opposition sur compte** : l'organisme de sécurité sociale créancier peut faire une opposition sur les comptes postaux et bancaires de ses débiteurs. ainsi sur les biens, meubles ou les liquidités appartenant au débiteur de l'organisme dans la limite des sommes qui lui sont dues.
  - ✓ **La procédure de la contrainte** : le directeur de la sécurité sociale créancier signe l'état des sommes dues et puis le visé et rendu exécutoire par le président du tribunal. L'employeur est averti officiellement par un agent de contrôle agréé de la sécurité social à travers un procès verbale de réception ou par un huissier de justice.
  - ✓ **La procédure du rôle** : notifié conformément aux dispositions du code des procédures fiscal et il est exécuté par les services des impôts territorialement compétents, le directeur de l'organisme de sécurité sociale créancier signe l'état des sommes dues qui sera visé par le Wali et le rendre exécutoire.[10]

### 3.3 Cadre fonctionnel

La structure du recouvrement comprend quatre (4) fonctions principales :

- ✓ Immatriculation
- ✓ Recouvrement des cotisations.
- ✓ Contentieux du recouvrement
- ✓ Contrôle employeur.



### 3.3.1 Immatriculation Employeur et assuré

#### 3.3.1.1 Immatriculation Employeur

L'immatriculation de l'employeur est faite selon la demande de l'employeur ou par voie de contrôle. Dans le cas où l'employeur met une demande d'affiliation, l'agent de l'immatriculation vérifie le dossier physique s'il n'est pas complet, il doit notifier le document manquant afin que l'employeur complète son dossier sinon l'agent d'immatriculation procède à la recherche ou bien l'identification de l'employeur par dénomination ou bien raison social dans la base de donnée, dont le but est d'éviter le double immatriculation pour ceux qui sont déjà existant dans la base donnée. En mettant à jour les informations d'un numéro employeur déjà existant, sinon :

- l'attribution d'un nouveau numéro employeur qui est composé de 10 chiffres dont son format est unique, la structure de ce dernier est comme suit :

W	W	S	S	S	S	C	C
Code Wilaya		Numéro séquentiel				clé de contrôle	

- La création d'un compte cotisant pour permettre aux services recouvrement d'enregistrer et de suivre toute action effectuée par l'employeur (ses obligations et ses sanctions).

#### 3.3.1.2 Immatriculation des assurés sociaux

L'immatriculation d'un assuré indique son déclaration au niveau de la CNAS en attribuant un numéro d'assurance national permettant de l'identifier facilement et rapidement, ce numéro est composé de 12 chiffres dont sa structure est comme suit :

A	A	AC	AC	AC	AC	S	S	S	S	C	C
Année de naissance		Numéro d'acte de naissance				Numéro séquentiel				clé de contrôle	

Le dépôt des déclarations de salariés s'effectue sous forme de dossier qui doit comprendre le formulaire « SECU01 » appelé aussi « Déclaration et demande d'affiliation de l'assuré social » fourni par la CNAS et un extrait d'acte de naissance, auprès des services

concernées de la CNAS , soit par ses employeurs soit par le biais des contrôleurs dont le traitement est fait sur plusieurs étapes :

- La vérification du dossier, s'il n'est pas complet l'agent d'immatriculation doit notifier le document manquant ou bien les informations manquantes sur les formulaires (comme le numéro employeur, date recrutement, .....etc.) , sinon il procède à la recherche de ces salariés :
  - soit mettre à jour à ceux déjà immatriculés, au niveau du service affiliation pour effectuer les modifications nécessaires.
  - soit attribuer des nouveaux numéros à ceux qui sont inexistant dans la base de donnée, ces informations doivent être saisie de manière identique aux données contenues dans la déclaration et l'acte de naissance, puis vérifier et enregistrer les données afin de générer un numéro d'assure social national . Après la phase de la création, l'envoi du dossier au service affiliation pour faire le nécessaire.

Sachant que le dépôt tardif de la déclaration du salarié entraîne l'application d'une pénalité de retard. « Art 13 de la loi 83-14 » [9] .

### **3.3.2 Recouvrement des cotisations**

Cette fonction représente le noyau du métier recouvrement, elle prend en charge toute la gestion des opérations du recouvrement (réglementaire, amiable ou forcé), en d'autre terme elle est porté sur le suivi et le traitement de la facturation et l'encaissement des assujettis en prenant en compte l'impact de la législation en la matière ainsi que l'exploitation de la comptabilité du recouvrement.

#### **3.3.2.1 Facturation**

La facturation est l'opération de la saisie des données de la déclaration sur le compte cotisant de l'employeur concerné. Cette déclaration soit déposée au niveau des services de la CNAS, soit envoyée via le portail télé-déclaration.

Une exécution d'un traitement mensuel est effectuée pour positionner et déterminer les périodes de cotisations sur les comptes cotisants, en fonction du nombre de salariés occupés par chaque employeur.

En ce qui concerne les comptes employeur suspendus, aucune modification s'effectue jusqu'à la reprise de l'activité qui permet la réactivation du compte cotisant.

Le calcul des pénalités et des majorations de retard est généré et actualisé automatiquement sur les comptes employeurs concernés suite aux manquements d'une ou plusieurs obligations en matière de sécurité sociale.

### **3.3.2.2 Encaissement**

Cette fonction consiste à prendre en charge les règlements effectués par différents moyens de paiement (chèque, virements,.....etc.).

Les versements effectués par l'employeur peuvent être, des cotisations de la période considérée dont les délais exigés n'ont pas pris fin, ou des cotisations des périodes antérieures dont les délais sont dépassés et des majorations de retard sont appliquées conformément à « Art24 de la loi 83-14 » [9]

Suite aux dispositions de la loi 08-08 du 23 février 2008 relative au contentieux en matière de sécurité sociale, l'employeur a la possibilité de faire un recours pour dégrever ou annuler le montant des pénalités et des majorations de retard appliqué, auprès de la commission de recours préalable qualifiée.

Les pénalités et les majorations de retard contestées peuvent être dégrévées de 50% du montant initial. La commission de recours peut décider de l'annulation totale des pénalités et des majorations de retard seulement dans le cas de force majeure dûment constatée. Et cette commission traite seulement les contestations en matière des montants des pénalités et des majorations de retard mais les cotisations principales ne sont pas incluses, elles doivent être payées sinon l'employeur ne peut pas bénéficier du dégrèvement des pénalités et des majorations de retard accordé par la commission.[11]

### **3.3.3 Contentieux du recouvrement**

Le métier contentieux du recouvrement de la sécurité sociale obéit à des procédures et des règles spécifiques, son rôle est d'assurer le recouvrement des créances litigieuses en participant à la maîtrise des risques financiers.

En d'autres termes, le service contentieux permet de définir la stratégie de recouvrement amiable qui va être suivie dans la première phase de recouvrement avant de procéder au recouvrement forcé, comme il a été indiqué précédemment.

## **4 Conclusion**

Dans ce chapitre, la présentation de l'organisme de la sécurité sociale CNAS a été entreprise ainsi que sa structure, son organisation, ses missions et son financement. La description de l'activité de recouvrement des cotisations suivie par ses objectifs et ses différents mécanismes faisant une partie majeure de cette présentation.

---

# Chapitre II :

## La Classification

---

Aujourd'hui, le Data mining est devenu un domaine très en vogue. Il est adapté et utilisé dans un large nombre de domaines ayant de bases de données volumineuses (banques, assurances, médecine, Etc.). L'intérêt principal est l'extraction des connaissances cachées dans ce tas volumineux de données, la détection des dépendances et des corrélations utiles entre les données afin d'offrir des outils de prédiction (nouveaux modèles) au service d'aide à l'organe décisionnel. Nous commençons ce chapitre, par la notion du processus de l'extraction des connaissances à partir des données « ECD » ou en anglais « Knowledge Discovery in Databases »(KDD). Nous enchainons par les définitions de ses différentes étapes : l'acquisition des données cibles, le prétraitement, l'étape de data mining. Cette dernière n'est qu'un ensemble de techniques et méthodes qui sont utilisés afin de construire un modèle ou un concept pré à être visualiser et évaluer. Suivi par, une étude détaillée de la technique des règles d'association qui est la technique ciblée dans notre sujet. Nous terminerons ce chapitre par une étape d'évaluation et d'interprétation de la nouvelle estimation.

## 1 Le processus « ECD »

### 1.1 Définition

Le terme « extraction de connaissances à partir de données (ECD) » (ou en anglais Knowledge Discovery in Databases (KDD)) est utilisée depuis 1989 [26], est le résultat de la convergence de recherches en apprentissage automatique, reconnaissance de formes, bases de données, statistique, intelligence artificielles, visualisation de données, etc.

D'après Han et Kamber [26], le processus ECD est l'analyse de **BDs** (souvent très grandes) afin de découvrir de relations insoupçonnées et de résumer les données d'une manière à la fois compréhensible et utile.

Fayyad et al [26], ECD désigne l'ensemble de processus d'extraction de connaissances utiles à partir de données, et souligne le « high-level » dans l'application de Data mining.

Ce processus vise à transformer des données (volumineuses, multiformes) stockées dans des entrepôts de données sous différents formats sur des supports pouvant être distribués en connaissances. Ces derniers peuvent s'exprimer sous forme d'un concept général ou un modèle : un rapport ou un graphique. Aussi, elles peuvent s'exprimer comme un modèle mathématique ou logique pour la prise de décision.

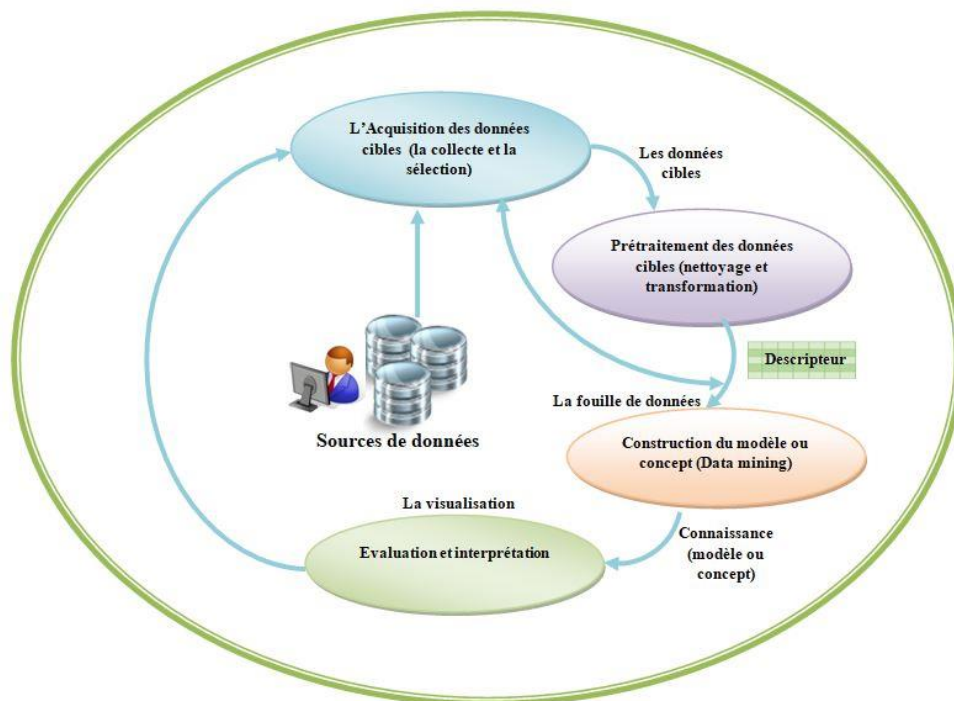
### 1.2 Les étapes du processus « ECD »

L'ECD « extraction de connaissances à partir de données » est un processus complexe qui se déroule suivant une suite de phases à effectuer pour extraire des informations cruciales, pertinentes, cachées et inconnus auparavant à savoir :

- ✓ **L'acquisition des données** : c'est la première phase permettant la collecte et la sélection des données pertinentes pour la tâche de **data mining** à accomplir. Ces données n'ont pas toujours le même format, le même type et la même structure.
- ✓ **Le prétraitement des données** : Cette phase traite l'élimination des bruits, d'erreurs et prétraitement de données manquantes, afin de les consolider ou de les transformer dans une structure de description de format approprié à la tâche de **data mining** choisie.

- ✓ **Le data mining** : c'est la phase d'utilisation des différentes méthodes et techniques intelligentes pour l'extraction ou l'exploration ou la découverte des nouveaux modèles ou nouveaux concepts.
- ✓ **Evaluation et interprétation du modèle** : cette dernière phase réalise l'identification des modèles ou concepts intéressants représentant les connaissances, en se basant non seulement sur des mesures d'intérêt mais aussi sur l'avis de l'expert.

Le schéma suivant représente le processus « ECD »



**Figure01 : Le processus « ECD »**

## 2 L'acquisition des données

Cette phase comprend deux étapes:

### ➤ La collecte

Après la définition du problème et les objectifs du data mining qu'on veut les atteindre, on peut avoir une idée sur les données qui doivent être utilisées. Ces données ne sont pas toujours unifiées. On peut avoir types

d'informations tels que des textes, des images, des bases de données, des pages web, ...etc.

➤ **La sélection**

Cette étape va nous permettre de sélectionner parmi toutes les données de la base celles qui concernent plus particulièrement le problème qu'on souhaite étudier ou traiter.

### **3 Le prétraitement des données**

Cette phase est réalisée en deux étapes :

➤ **Le nettoyage**

Dans l'étape de nettoyage, nous nous intéressons à l'examen de la qualité des données collectées et sélectionnées. On traite des problèmes courants tels que les doublons, les erreurs de saisie, l'intégrité de données et le problème des valeurs manquantes.

➤ **La transformation**

Permet de rassembler et d'unifier dans un cadre conceptuel commun les différentes données cibles. A la fin de cette étape, on obtient généralement un entrepôt de données, qui servira de référentiel de source de données cibles pour la suite du processus.

### **4 Le Data mining**

C'est l'une des étapes du processus désignant le cycle de la découverte d'informations ou de connaissances dans les bases de données (ECD).

#### **4.1 Les modèles du data mining**

La phase du data mining regroupe un ensemble des techniques et méthodes supervisées et non supervisées destinées à l'exploration et l'analyse de grandes bases de données informatiques, dans le but de détecter des règles, des associations, des tendances nouvelles et imprévisibles (modèles ou concepts), dont l'objectif est l'aide à la décision.

##### **4.1.1 Modèle de data mining supervisé**

Le processus du modèle supervisé peut se représenter en deux étapes :

➤ **Apprentissage supervisé**



Construction du modèle à partir de l'ensemble d'apprentissage (training data set) sachant que les classes sont **prédéterminées**.

➤ **Classification**

Utilisation du modèle d'apprentissage pour la classification de nouvelles données (modèles ou concepts).

Il existe plusieurs techniques et méthodes du modèle de data mining supervisés dont on cite les plus populaires :

- ✓ les arbres de décision
- ✓ les réseaux de neurones
- ✓ autres.

#### 4.1.2 Modèle de data mining non supervisé

Le processus du modèle non supervisé est la séquence de deux étapes :

➤ **Apprentissage non supervisé**

Dans cette phase nous construisons l'ensemble d'entraînement (training set). Dans ce type de modèle, les classes ou les valeurs des éléments ne sont pas prédéterminées, donc il vise à identifier des ensembles d'éléments qui partagent certaines similarités (base d'apprentissage).

➤ **Classification**

Cette phase consiste à examiner des caractéristiques d'un élément nouvellement présenté afin de l'affecter à une classe d'un ensemble défini par le modèle d'apprentissage non supervisé.

On peut citer par exemple les méthodes de la technique non supervisé les plus connues:

- ✓ K-means.
- ✓ les règles d'association.
- ✓ Autres.

Le tableau suivant représente les avantages et les inconvénients des algorithmes mentionnés ci-dessus [24] :

Type d'apprentissage	Méthodes	avantages	Inconvénients
Supervisé	arbre de décision	<ul style="list-style-type: none"> <li>- Simple et facile à comprendre.</li> <li>- peut gérer à la fois des valeurs numériques et des catégories.</li> </ul>	<ul style="list-style-type: none"> <li>-très sensible aux points aberrants et au bruit</li> <li>-sa sensibilité au changement des données</li> <li>-Une détection difficile des interactions entre les variables</li> </ul>
	réseaux de neurones	<ul style="list-style-type: none"> <li>-La possibilité de combiner ce type d'algorithmes avec d'autres méthodes de classification.</li> <li>- On peut simuler son comportement uniquement à partir d'un certain nombre d'exemples observés</li> </ul>	<ul style="list-style-type: none"> <li>-les résultats obtenus sont difficiles à comprendre et à les interpréter.</li> <li>-utilisables que sur des variables numériques</li> <li>- Difficulté de choisir la structure (type, nombre de nœuds, organisation, connexions,...etc.) la mieux adaptée au Problème.</li> </ul>
Non supervisé	K-means	<ul style="list-style-type: none"> <li>-La facilité de mise en œuvre de cet algorithme.</li> <li>-Son efficacité pour des classes réparties de manière irrégulière</li> </ul>	<ul style="list-style-type: none"> <li>-Le nombre de classe doit être fixé au départ.</li> <li>-Les clusters sont construits par rapports à des objets inexistantes (les milieux).</li> </ul>
	les règles d'association	<ul style="list-style-type: none"> <li>-Leur application dans plusieurs domaines de la vie quotidienne, comme l'analyse du panier de la ménagère.</li> <li>-La découverte de connaissances utiles, cachées dans les grandes bases des données</li> <li>-Leur simplicité, efficacité et facilité de compréhension.</li> <li>-Leur formalisme non supervisé et général. Leurs résultats clairs et faciles à interpréter.</li> </ul>	<ul style="list-style-type: none"> <li>-Le temps énorme consacré à la recherche des ItemSets fréquents.</li> <li>-La grande quantité des règles d'association générées.</li> <li>-La production des règles triviales et inutiles qui n'apportent pas de nouvelles informations.</li> </ul>

**Tableau 03:** les avantages et les inconvénients de quelques algorithmes de classification.

## 5 Les règles d'association

La recherche de règle d'association consiste à extraire des informations à partir de la coïncidence dans les données permettant d'identifier des liens implicites qui sont cachés par l'important volume des bases de donnée opérationnelle.

Agrawal et Al (1993) ont été les premiers à évoquer le problème des règles d'association dans la littérature scientifique, en introduisant l'algorithme AIS. Les auteurs ont également introduit les concepts de base de l'extraction des règles à savoir la notion de support minimum et de confiance minimum. Ces de mesures statistiques sont utilisées comme critères pour éviter de considérer les règles non significatives.

En 1994, Agrawal et Al ont présenté l'algorithme Apriori, un algorithme d'extraction de règles d'association qui permet de trouver plusieurs items dans la conséquence de la règle d'association trouvée. « Apriori » sera traité en détail par la suite [25].

## 5.1 Définitions

Plusieurs termes utilisés dans la recherche et l'extraction des règles d'association, donc on définit :

### 1. Transaction

On considère un ensemble  $E = \{E_1, \dots, E_n\}$  de  $n$  éléments (items) distincts. On appelle une transaction  $T$  le sous ensemble  $E'$  inclus dans  $E$ . Dans une base de données  $D$ , chaque transaction est identifiée par une clé unique [15][25].

### 2. Une règle d'association

Soit  $A$  et  $B$  deux sous-ensembles d'éléments qui appartiennent à un ensemble  $M$ . On appelle  $A \Rightarrow B$  la règle d'association ou la relation entre  $A$  et  $B$ , signifie que si  $A$  se trouve dans une transaction,  $B$  peut apparaître dans cette même transaction. L'évaluation de cette relation se base principalement sur le calcul de deux paramètres qu'on appelle le support et la confiance [15][25].

3. **Itemset** Un Itemset désigne un ensemble d'objets ou d'articles [15][25].

4. **K-Itemset** Un K-Itemset est un Itemset de  $k$  Items [15][25].

### 5. Support d'un Itemset

Soit  $A$  un Itemset de  $n$  éléments. Dans une base de données transactionnelle  $D$ , le support de  $A$  est le nombre de transactions dans  $D$  incluant  $A$  divisé par le nombre total des transactions de  $D$  [15][25] :

$$\text{Support}(A) = \frac{|\{t \in D / A \subseteq t\}|}{|D|}$$

### 6. Support d'une règle d'association

Dans une base de données  $D$ , le support d'une règle d'association  $A \Rightarrow B$  est le nombre de transactions qui contiennent  $A$  et  $B$  divisé par le nombre total des transactions [15][25] :

$$Support(A \Rightarrow B) = \frac{|\{t \in D / (A \cup B) \subseteq t\}|}{|D|}$$

## 7. Confiance

La confiance d'une règle d'association  $A \Rightarrow B$  est le rapport entre le nombre de transactions de  $D$  contenant  $A \cup B$ , et le nombre de transactions de  $D$  contenant  $A$ . C'est-à-dire que [15][25]:

$$Conf(A \Rightarrow B) = \frac{|\{t \in D / (A \cup B) \subseteq t\}|}{|\{t \in D / A \subseteq t\}|}$$

Cette formule est équivalente à :

$$Conf(A \Rightarrow B) = \frac{Support(A \cup B)}{support(A)}$$

## 8. Itemset Fréquent

Un Itemset  $A$  est fréquent si et seulement si son support est supérieur à un support minimum défini par l'utilisateur [15][25].

### 5.2 Les étapes d'extraction des règles d'association

Le processus d'extraction des règles d'association se déroule en trois étapes :

#### 5.2.1 Préparation des données

Cette étape consiste à réduire la quantité des données en gardant seulement celles les plus pertinentes, et en transformant par la suite, ces derniers en un contexte d'extraction, c'est-à-dire une transformation en un triplet constitué: d'un ensemble d'objets, d'un ensemble d'Itemsets ainsi qu'une relation binaire entre les deux. Cette transformation des données en données binaires permettra d'améliorer la qualité des règles d'association [15],[25].

#### 5.2.2 Recherche des Itemsets fréquents

Un Itemset fréquent est un ensemble d'éléments dont le support est supérieur ou égal à un certain support minimal spécifié par l'utilisateur. Cette étape est très coûteuse en temps d'exécution. Pour un ensemble de  $n$  items par exemple, le nombre d'Itemsets fréquents qui peut être générés est de  $2^N$  [15],[25].

### 5.2.3 Production des règles d'association

La génération des règles d'association consiste à déterminer les règles d'association dont le support et la confiance sont supérieurs ou égaux à un certain support et confiance minimaux définis par l'utilisateur [15],[25].

## 5.3 L'algorithme Apriori

### 5.3.1 Définition

Apriori (Agrawal et Al 1994) est un algorithme itératif de recherche des itemsets les plus fréquents par niveaux. Dans le but de limiter le nombre d'itemsets candidats, de sorte à traquer les itemsets les plus fréquents répondant au seuil de support minimum.

### 5.3.2 Le principe de l'algorithme

Cet algorithme se base principalement sur les deux règles suivantes :

- Tout sous-motif d'un motif fréquent est fréquent.
- Tout sur-motif d'un motif non fréquent est non fréquent.

Le déroulement de l'algorithme Apriori peut être décomposé en quatre étapes:

- a) Recherche de tous les nouveaux candidats.
- b) Pour chaque candidat trouvé, on calcule son support.
- c) Évaluation du support calculé par l'algorithme, par rapport au support minimum défini par l'utilisateur.
- d) On supprime les candidats dont le support est inférieur au support minimum.

La description complète de l'algorithme Apriori se résume dans le pseudo code suivant [15]:

---

### Algorithme Apriori

---

Entrées : Base de données de transactions D, Seuil de support minimum  $\sigma$

Sorties : Ensemble des motifs fréquents

$i \leftarrow 1$

$C_i \leftarrow$  Ensemble des motifs de taille 1 (un seul motif)

**Tantque  $C_i \neq \emptyset$  faire**

Calculer le Support de chaque motif  $m \in C_i$  dans la base

$$F_i \leftarrow \{m \in C_i \mid \text{support}(m) \geq \sigma\}$$

$C_{i+1}$  Toutes les combinaisons possibles des motifs de  $F_i$  de taille  $i + 1$

$$i \leftarrow i + 1$$

**Fin Tantque**

retourner  $\bigcup_{i \geq 1} F_i$

**5.4 Définition d'une règle solide**

Les règles qui dépassent un minimum de support et un minimum de confiance sont appelées règles solides qui seront utilisées pour la classification.

Une fois les motifs fréquents dans une base de données sont extraits, il devient simple de générer les règles d'association qui vérifient un minimum de support et un minimum de confiance, comme suit :

- Pour chaque motif fréquent  $l$ , générer tous les sous-ensembles non vides de  $l$ ,
- Pour chaque sous-ensemble non vide  $s$  de  $l$ , enregistrer la règle  $(s \rightarrow l - s)$  si :

$$\text{Confidence}(s \rightarrow l - s) \geq \text{Min\_confiance}$$

Où  $\text{Min\_confiance}$  est un seuil minimum de confiance.

Puisque les règles sont générées des motifs fréquents, chacune vérifie automatiquement le support minimum [15], [25].

<b><math>m1 \Rightarrow m2</math></b>
<b><math>\text{sup}(m1 \Rightarrow m2) \geq \text{min\_support}</math></b> <b><math>\text{conf}(m1 \Rightarrow m2) \geq \text{min\_confiance}</math></b>

**Tableau 04 : Définition d'une règle solide.**

## **6 Evaluation et interprétation du modèle**

Cette étape consiste à estimer la qualité, mesurer l'utilité, l'originalité et l'intelligibilité du modèle ou concept prédit. Un modèle est considéré comme une connaissance s'il est utile, inconnu auparavant, et dépasse un certain pourcentage de précision. L'évaluation prend généralement une forme qualitative ou une forme quantitative.

La forme qualitative consiste à restituer la connaissance sous forme graphique ou textuelle. Cette contribution permet d'améliorer la compréhension des résultats aux décideurs et faciliter le partage de la connaissance.

La forme quantitative nécessite une certaine expertise fonctionnelle tels que l'intervalle de confiance, la validation par test, etc. pour interpréter les résultats aux décideurs.

## **7 Conclusion**

Dans ce chapitre, on a étudié le processus de l'extraction des connaissances à partir des données « ECD ». Cette analyse, nous a permis de définir les différentes étapes qui le composent en particulier le data mining, considéré comme l'étape d'estimation de l'ECD. On a noté une importance particulière à l'une des méthodes d'exploration très attractive par sa simplicité et efficacité celle de la segmentation par les règles d'association. Cette méthode associative utilise l'algorithme populaire Apriori [15], [25] pour l'extraction des motifs fréquents.

# Deuxième partie



---

# Chapitre III :

## La Conception

---

La conception est un processus important dans le cycle de vie du développement d'un logiciel, ce processus consiste à décrire le système de manière globale et détaillé.

Notre système décisionnel proposé, se base sur l'une des techniques de modélisation importante de l'extraction de connaissances par les données. Elle sert à la prédiction d'un nouveau modèle, dans notre cas celui du «non recouvrement de la sous-direction du recouvrement et du contentieux de la CNAS ».

La conception du système proposé utilise et répond aux définitions des besoins du non recouvrement ainsi que les limites de ce problème. Elle assure le prétraitement des données acquises et estime le nouveau modèle. La méthode associative est celle utilisée pour la classification par les règles d'association et les motifs fréquents [15].

## **1 L'objectif**

L'objectif de notre analyse est de modéliser un système de classification qui peut fournir un ensemble des règles d'association par les motifs fréquents. Ce système permettant l'aide à la minimisation du non recouvrement au niveau de la CNAS.

## **2 La conception du système**

Notre système est globalement composé de cinq composants au niveau du back office.

### **2.1 Architecture générale du système**

Le schéma suivant représente l'architecture du système proposé contenant quatre sous-systèmes à savoir (Figure06) :

- Acquisition des données cibles.
- Prétraitement.
- Classification.
- Visualisation.

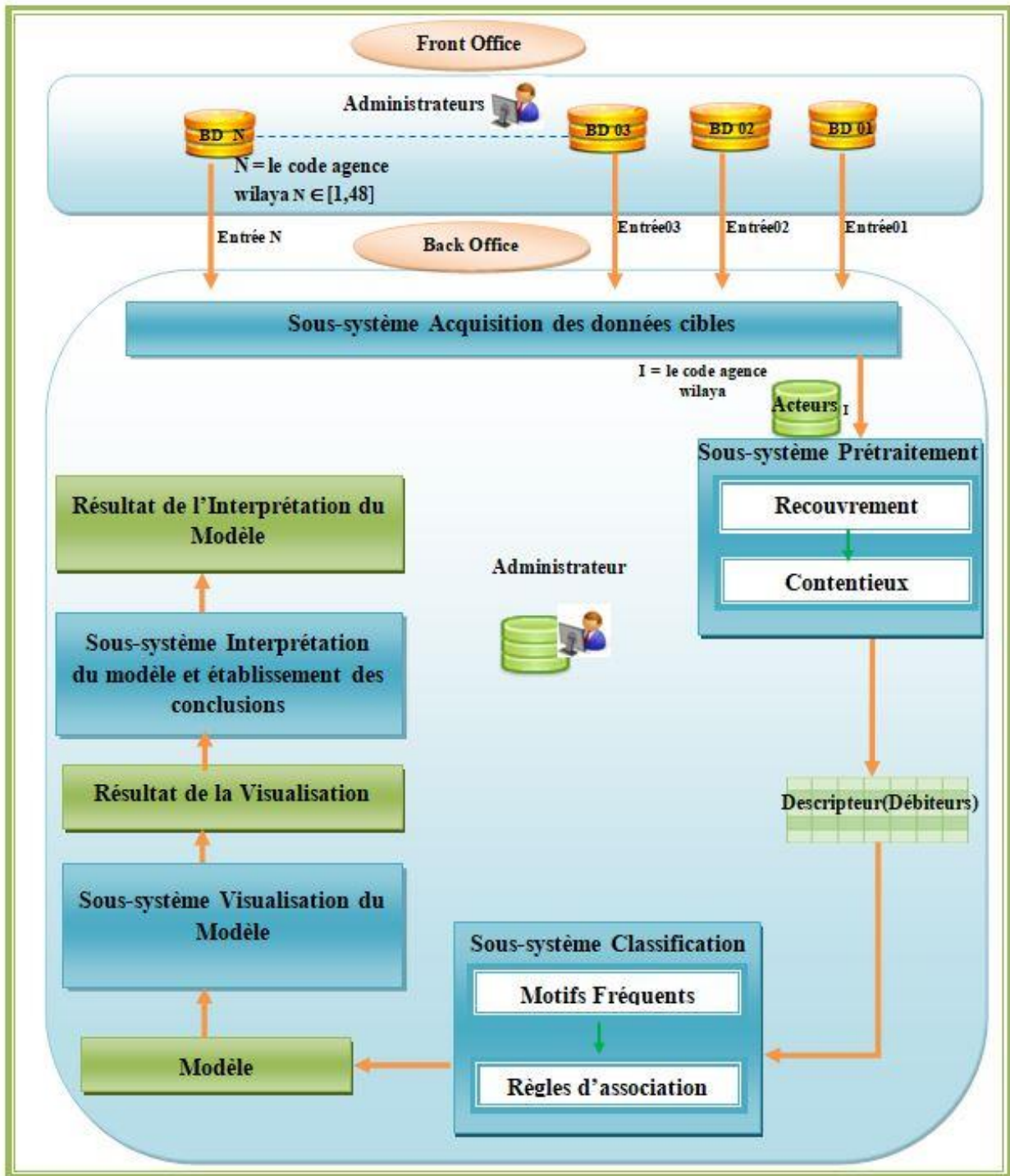


Figure02: Architecture générale

## 2.2 Architecture Détaillée

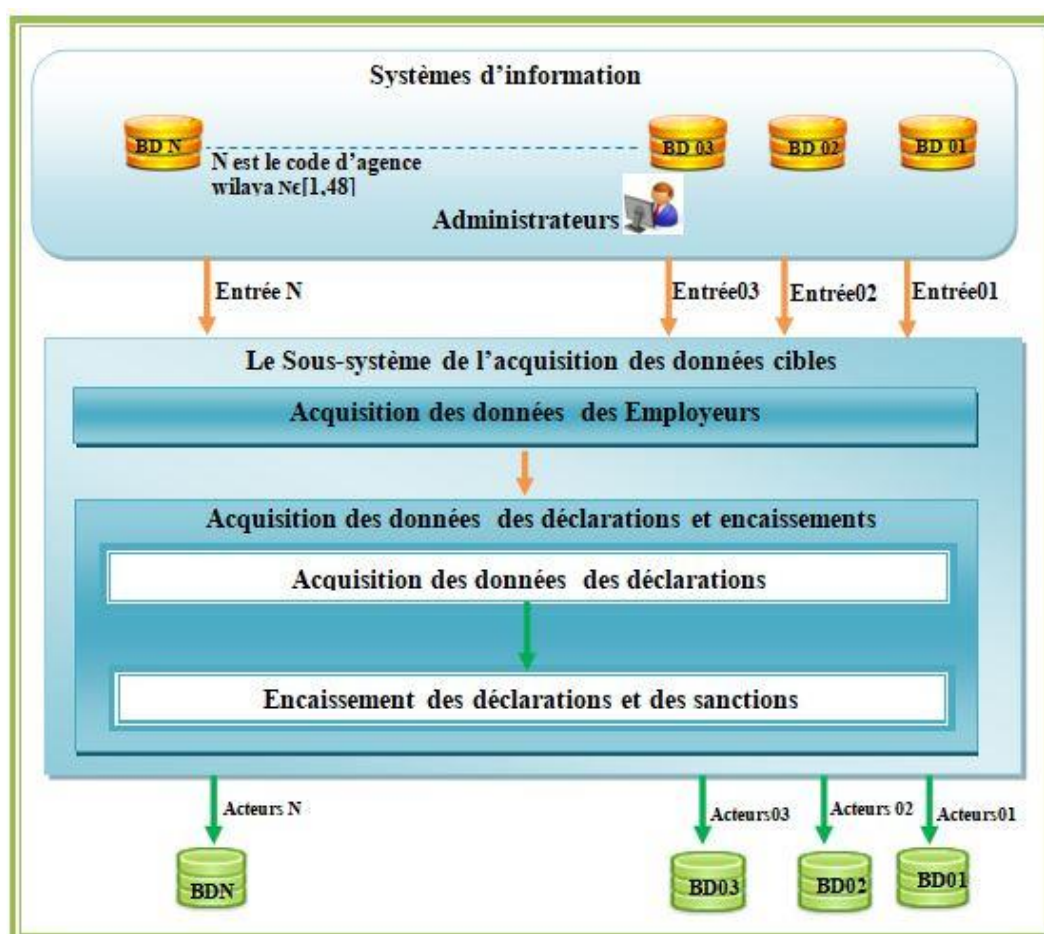
Cette partie a pour but de donner les détails précis des éléments produits dans la conception globale et préparer au mieux l'implémentation, en précisant les entrées, les sorties et les traitements de chaque élément.

### 2.2.1 Sous système d'acquisition des données cibles

D'après la définition du problème et les objectifs qu'on veut atteindre, on peut avoir une idée sur les données qui doivent être utilisées, sachant qu'elles n'ont pas toujours le même format et la même structure. Ces données seront extraites et déterminées dans le sous système d'acquisition des données cibles, afin qu'elles seront traités et manipulés dans les prochains sous systèmes.

Le sous système d'acquisition des données cibles contient deux composants :

- Cotisants
- La gestion des cotisations



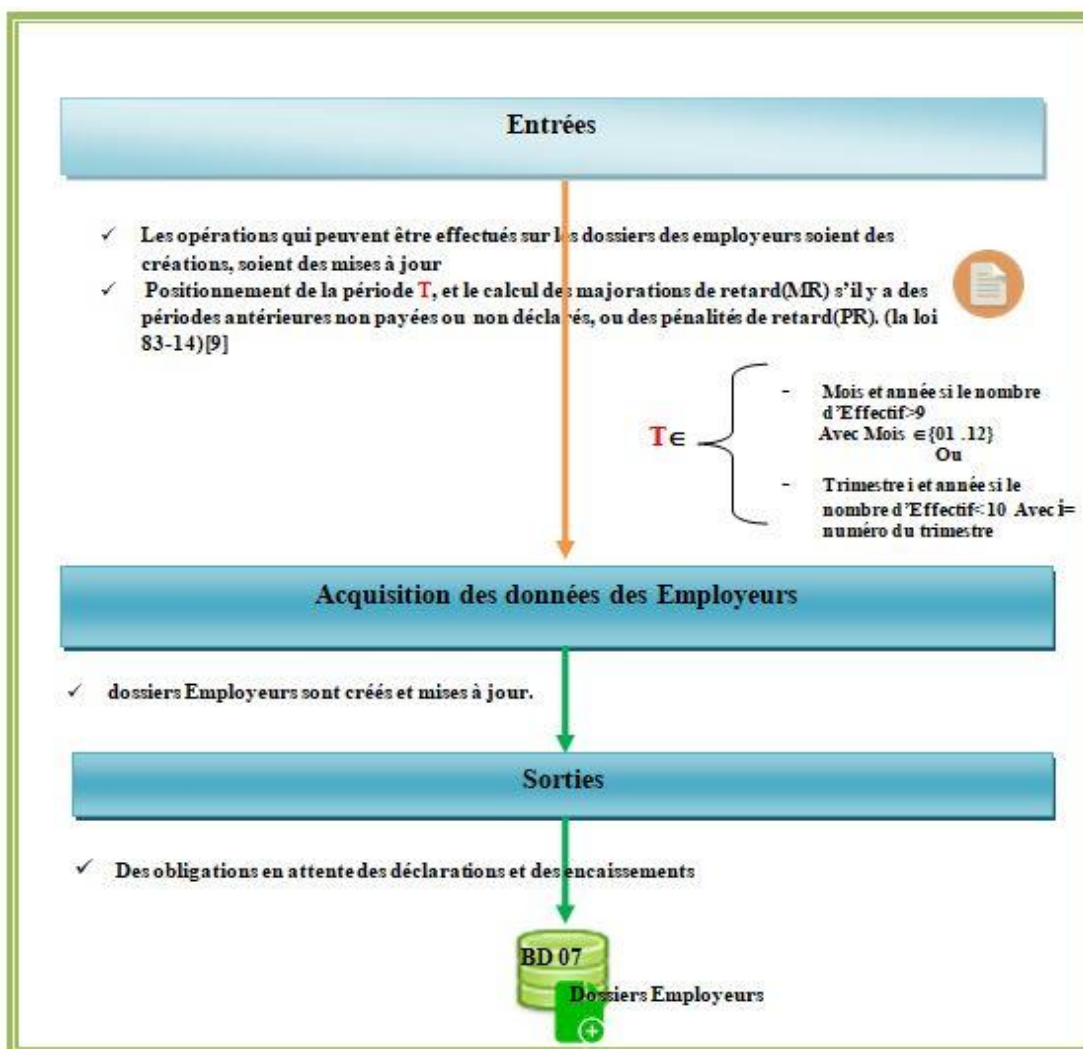
**Figure03** : Le Sous-système de l'acquisition des données cibles

### 2.2.1.1 Les entrées du sous-système de l'acquisition des données cibles

Le système d'information de la sécurité sociale, est développé par les ingénieurs de la CNAS, ce système permet l'automatisation des règles de gestion relatives au recouvrement des cotisations, l'immatriculation des assurés et les employeurs. En prenant la **base d'entraînement simulé** à celle de l'Agence CNAS de Biskra dont les données utilisées sont **virtuelle**(ils ne sont pas réels juste pour le cas d'étude).

#### 2.2.1.1.1 Acquisition des données des Employeurs

L'employeur, représente le noyau du métier recouvrement. La figure suivante représente l'architecture globale de l'acquisition des données Employeurs.



**Figure04 :** Architecture de l'acquisition des données Employeurs  
Le diagramme suivant représente le diagramme de classe « Employeur »

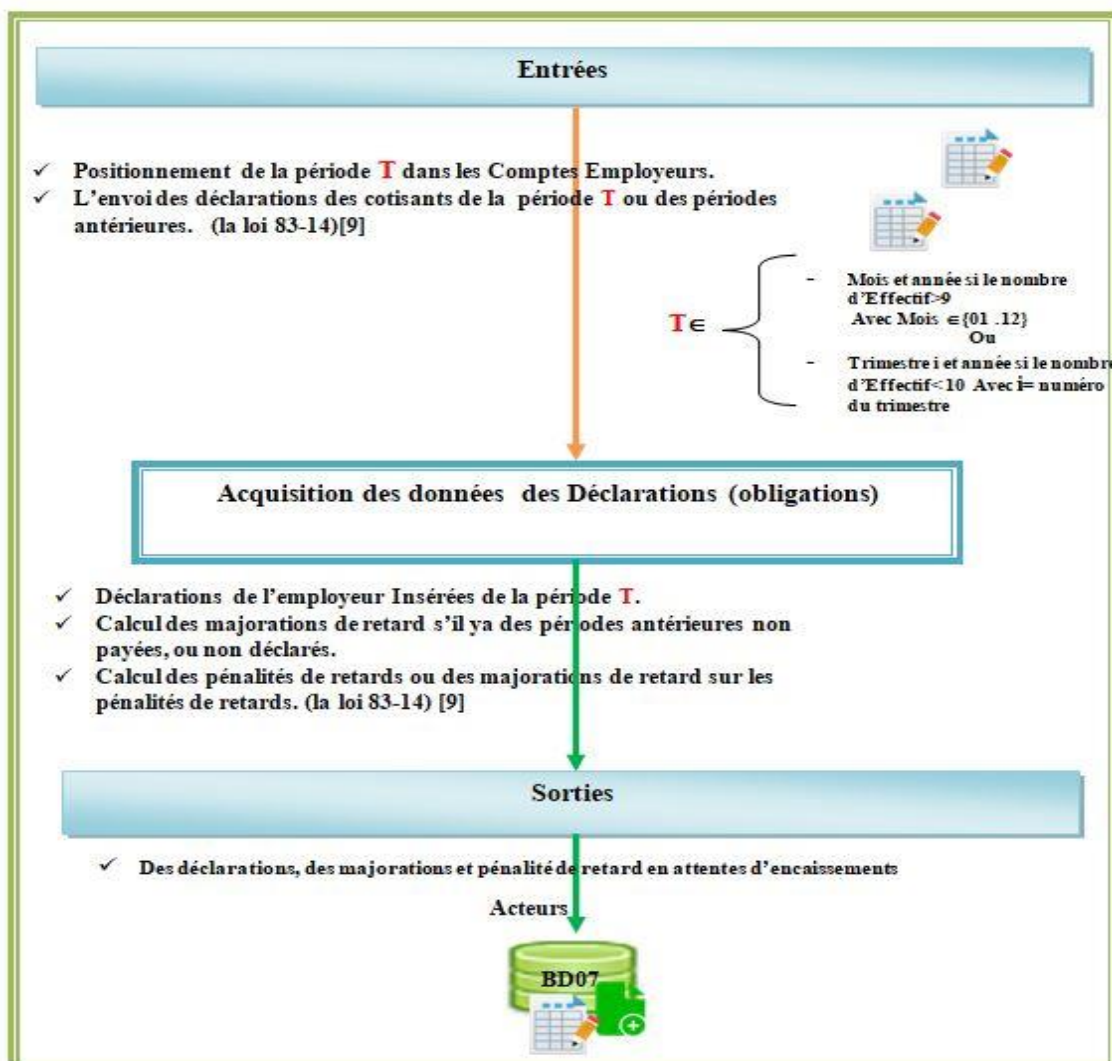
### 2.2.1.1.2 Acquisition des données des déclarations et encaissements

Cette partie, gère l'acquisition des données qui représentent les situations des employeurs à travers leurs déclarations et encaissements.

#### ➤ Acquisition des données des Déclarations

Les déclarations d'employeurs représentent toutes les obligations visé à vis la CNAS.

La figure suivante représente l'architecture de la déclaration.

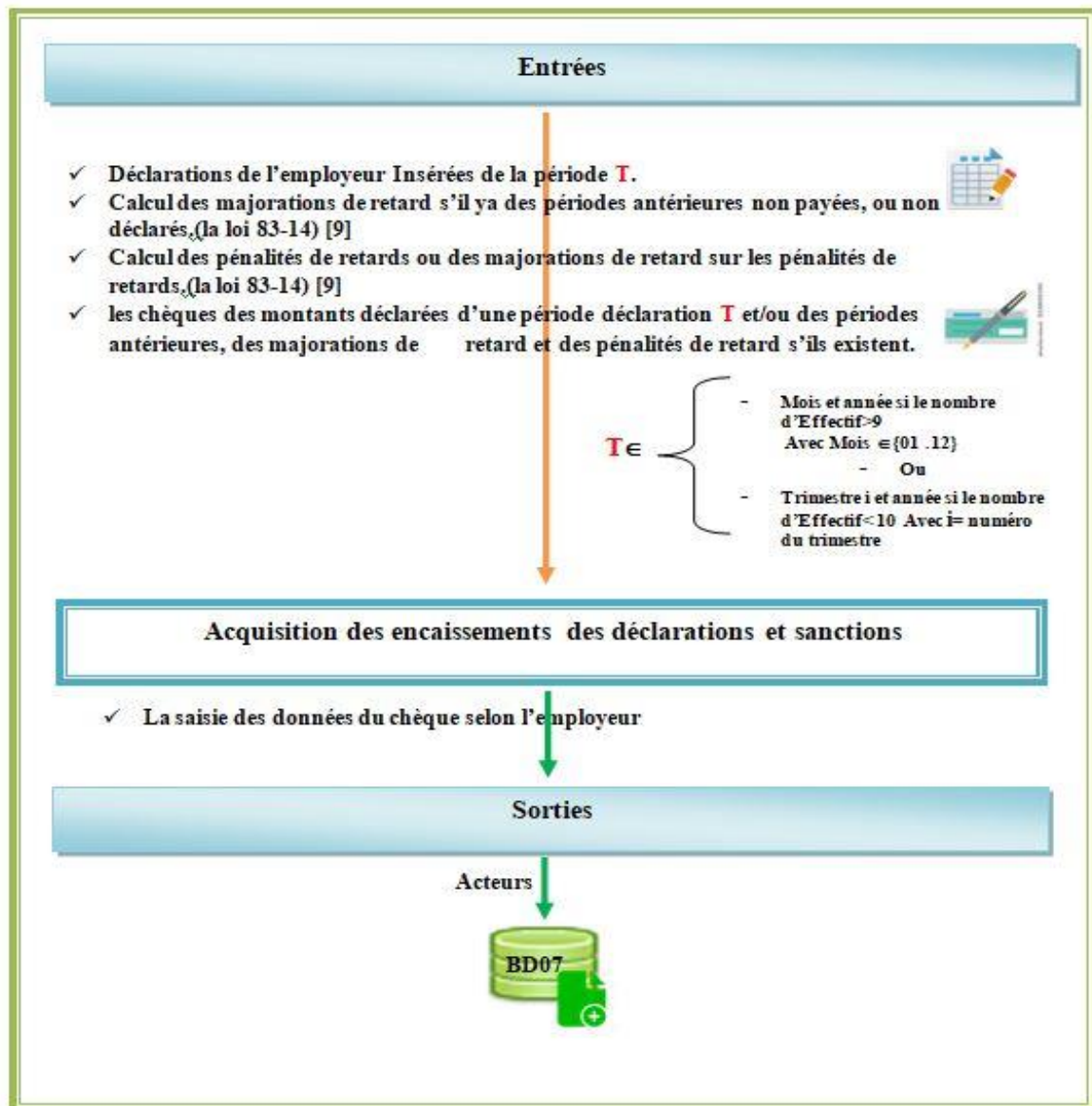


**Figure05** : Architecture d'acquisition des données des déclarations

➤ **Encaissements des déclarations**

L'encaissement est une opération par laquelle le cotisant demande à la CNAS de procéder au traitement d'un chèque du montant déclaré.

La figure suivante représente l'architecture de l'encaissement

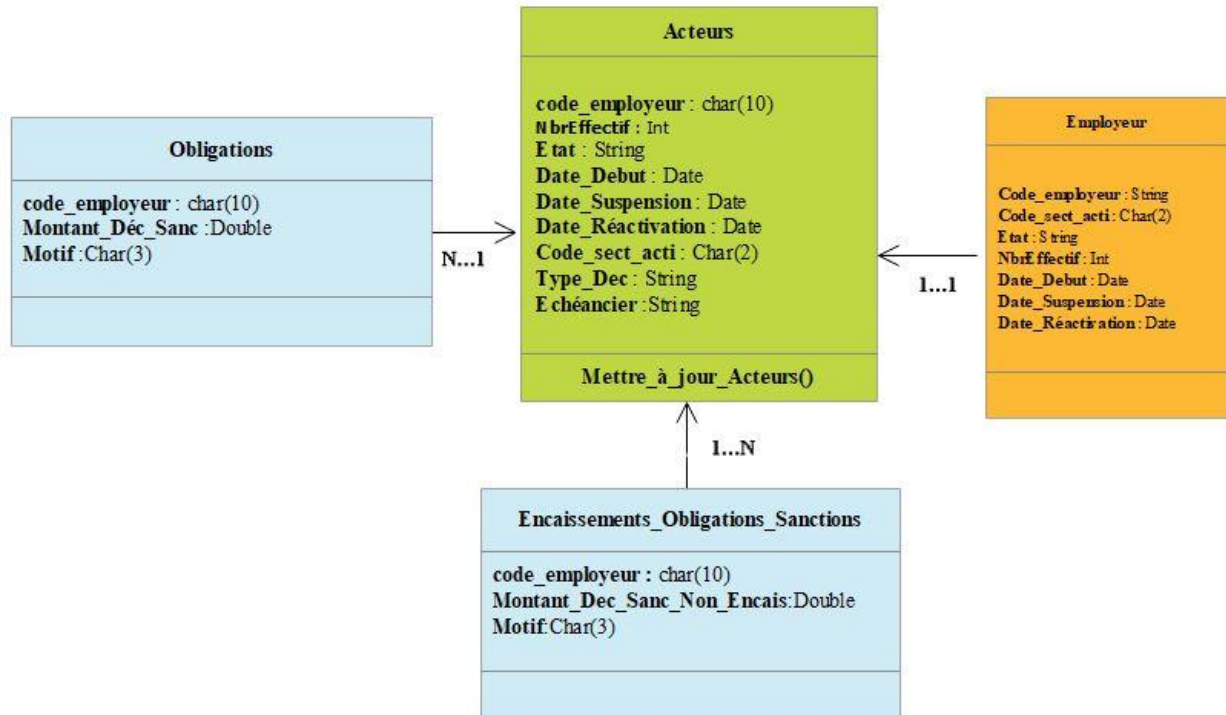


**Figure06** : Architecture de l'encaissement des déclarations (obligations)



### 2.2.1.2 Les sorties du composant

Les données sortantes du sous-système de l'acquisition des données cibles, représentent des acteurs pré à traiter par le sous-système prétraitement.

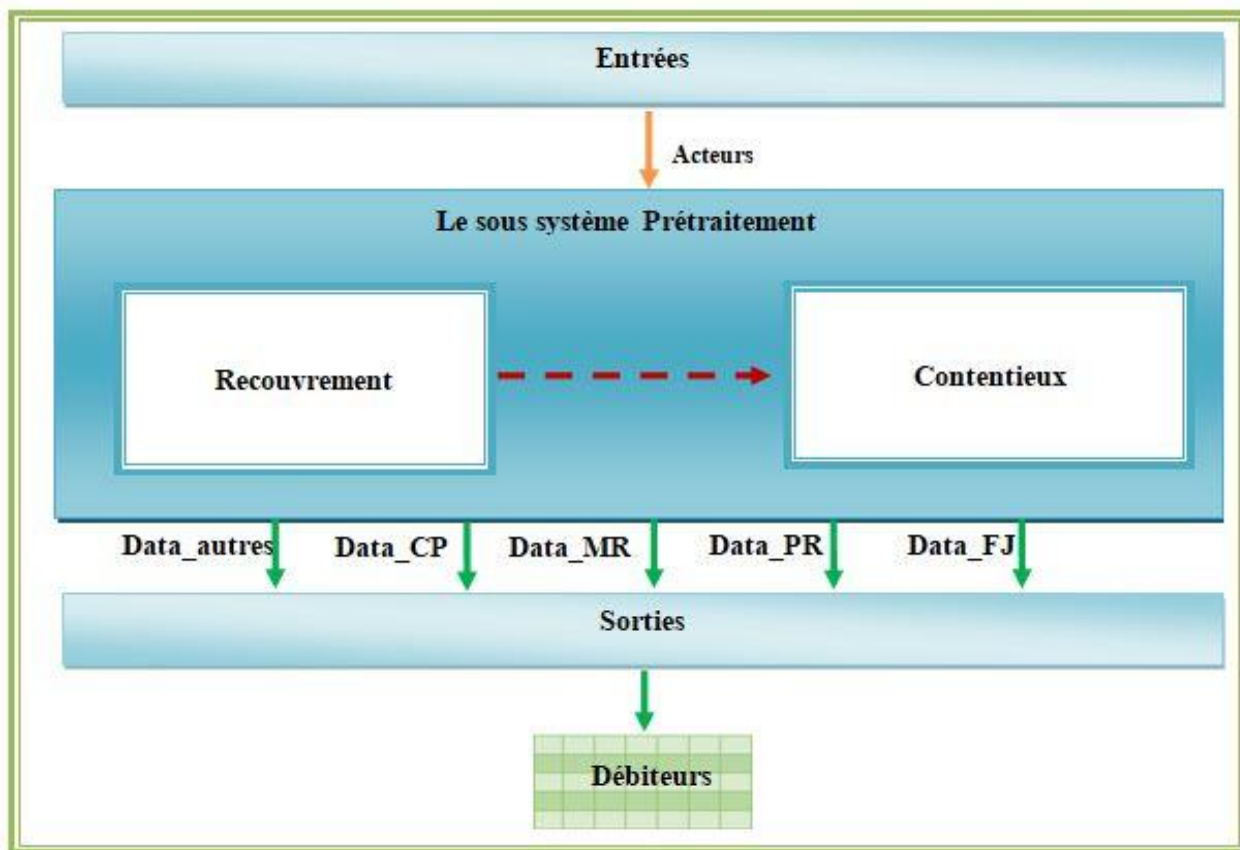


**Figure07:** Diagramme de classe des Acteurs

### 2.2.2 Le sous système Prétraitement

Dans l'étape de prétraitement on fait l'extraction des motifs qui caractérisent chaque employeur en terme non recouvrement, et élimine tous bruits ou informations inutiles. Ce sous-système contient deux composants

- Recouvrement
- Contentieux



**Figure08:** Architecture du sous-système Prétraitement

### 2.2.2.1 Les entrées du Sous système

L'ensemble des entrées de ce sous-système représente seulement les motifs fréquents « Acteurs ».

### 2.2.2.2 Composants interne du sous système

#### ➤ Recouvrement

Ce composant est chargé de sélectionner l'ensemble des employeurs qui ont une dette vis-à-vis la CNAS, exécutée par le service des cotisations.

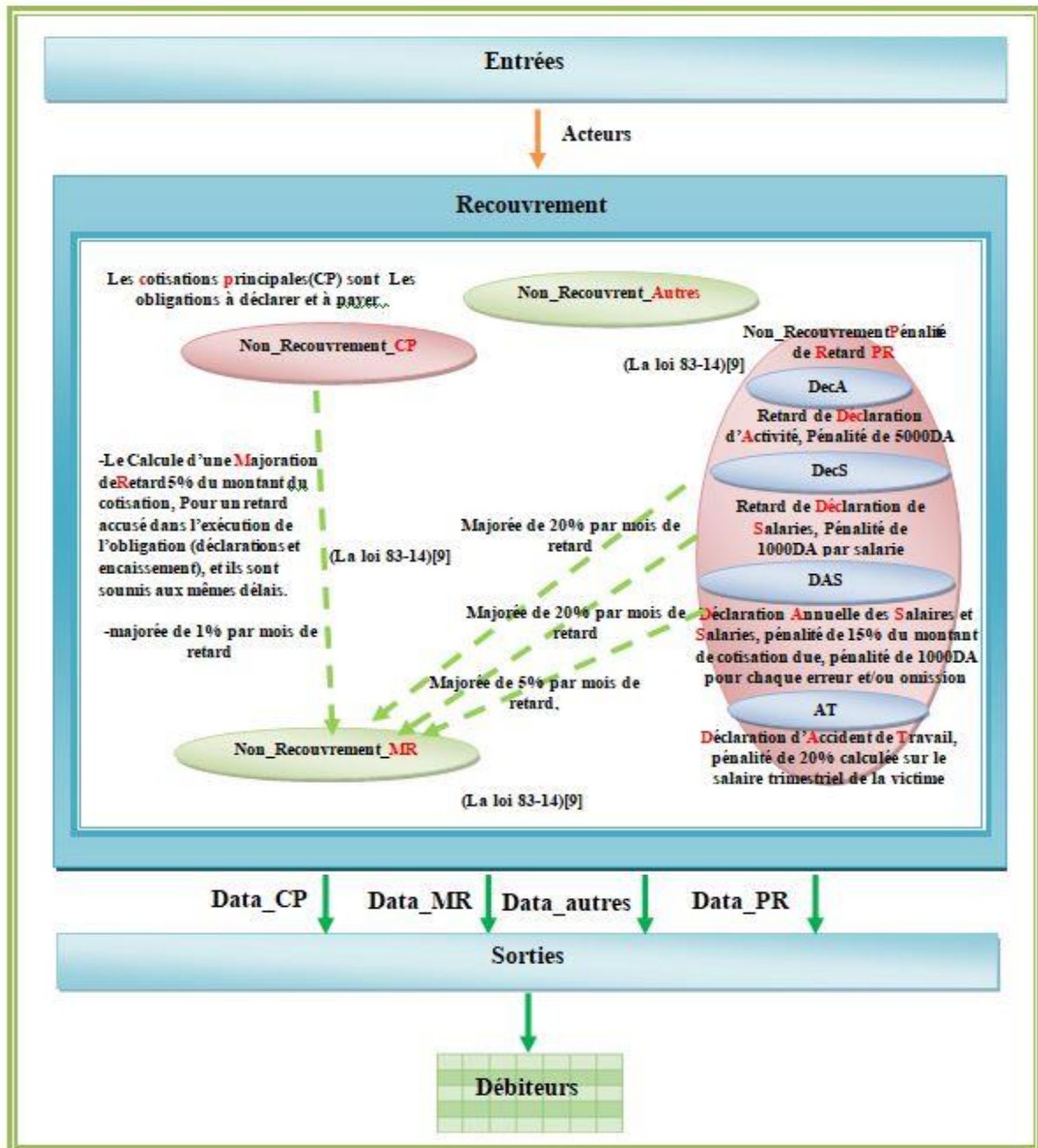


Figure09 : Architecture du Recouvrement

➤ **Contentieux**

Ce composant est chargé de sélectionner l'ensemble des débiteurs qui ont une dette vis-à-vis la CNAS, exécutée par le service du contentieux.

La figure suivante représente l'architecture du Contentieux :

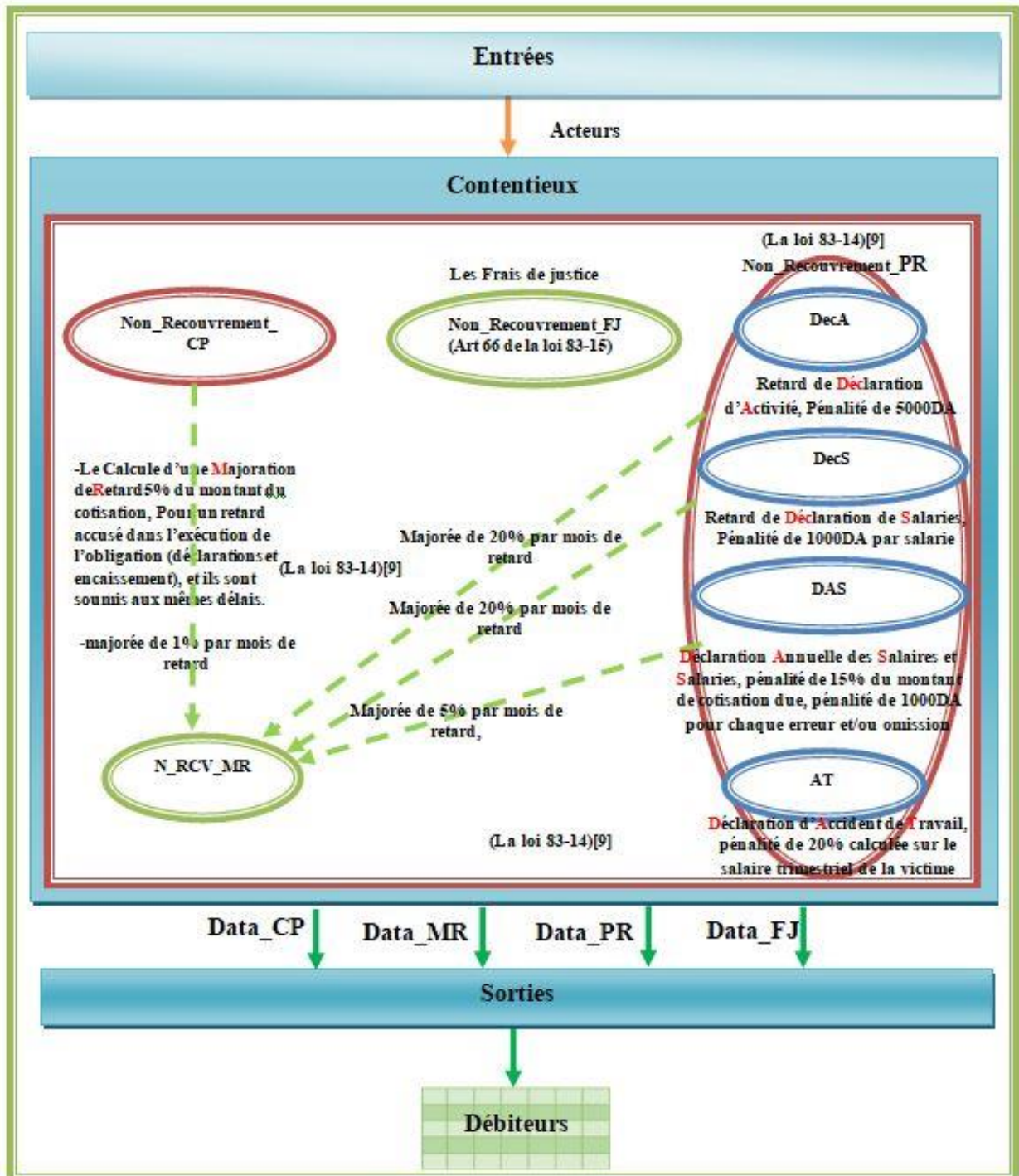


Figure10: Architecture du Contentieux

### 2.2.2.3 La sortie du Sous système

La donnée sortante du sous-système Prétraitement représente un Descripteur. Notre descripteur « débiteurs » est une table bidimensionnelle dont ces valeurs sont binaires. La ligne du débiteur représente l'individu dit cotisant et la colonne représente un motif du non recouvrement.

La figure suivante représente le diagramme de classe d'un Débiteur.

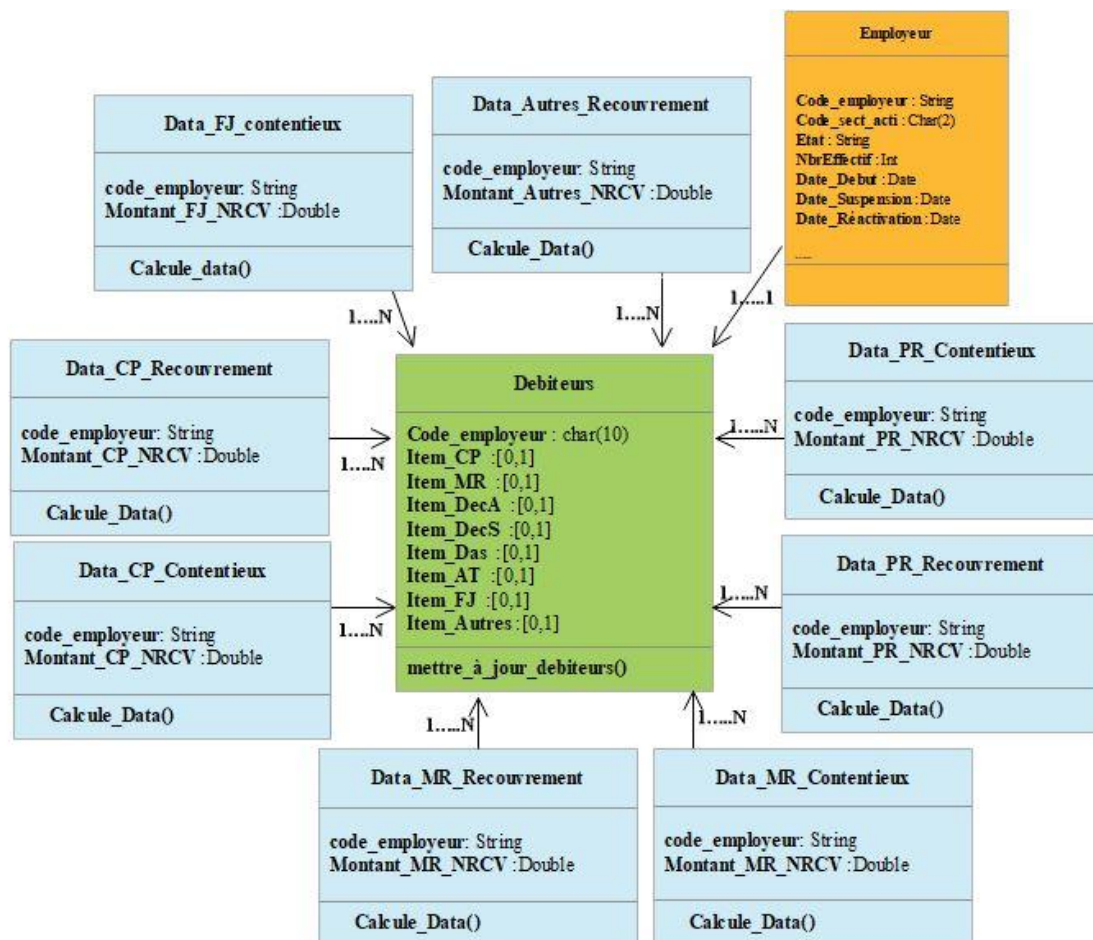


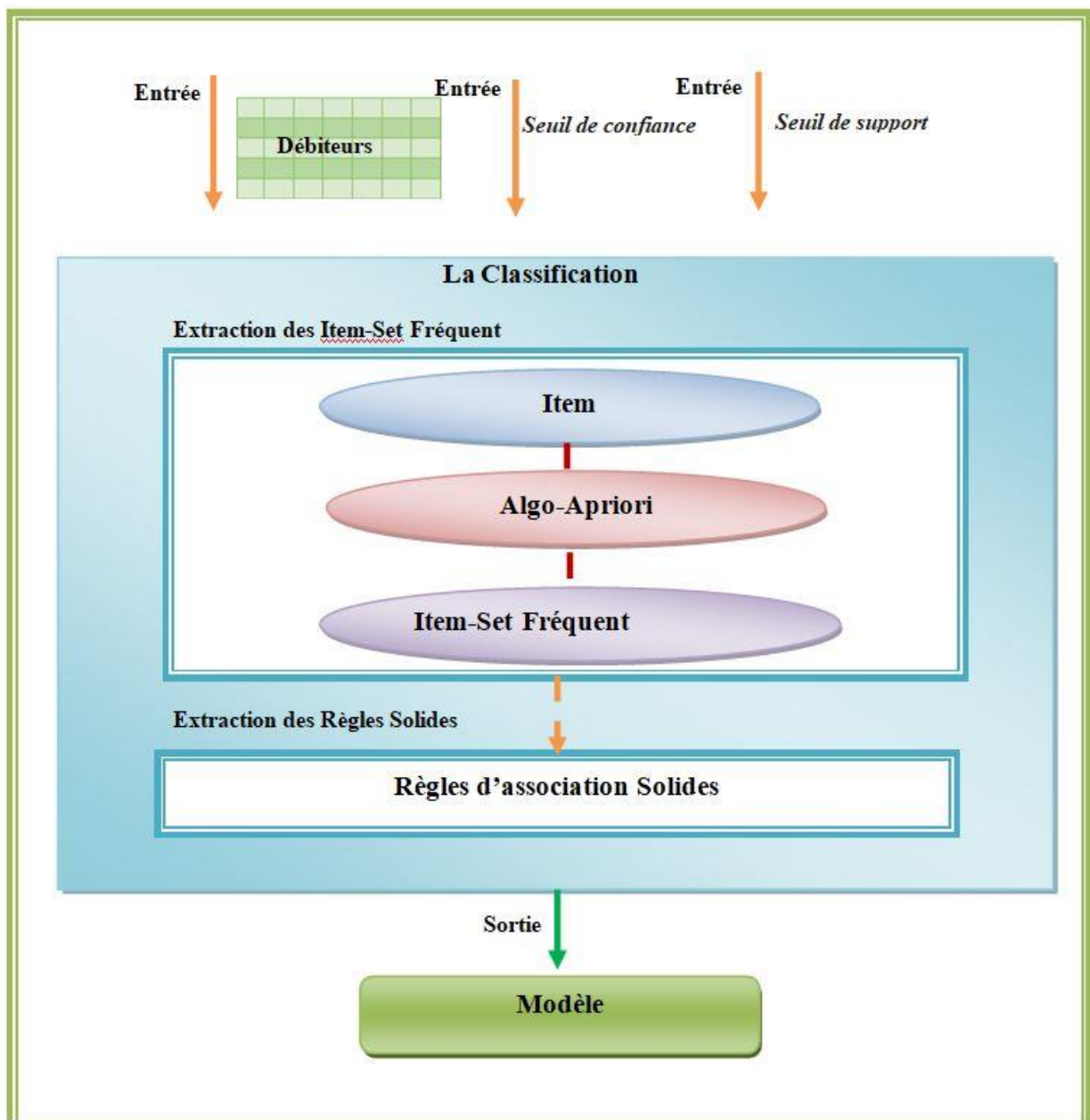
Figure11: Diagramme de classe de « Débiteurs »

### 2.2.3 La classification

Dans ce module on utilise l'algorithme Apriori, ce dernier sert à reconnaître des propriétés qui reviennent fréquemment dans l'ensemble de données de notre descripteur qui représente l'un des

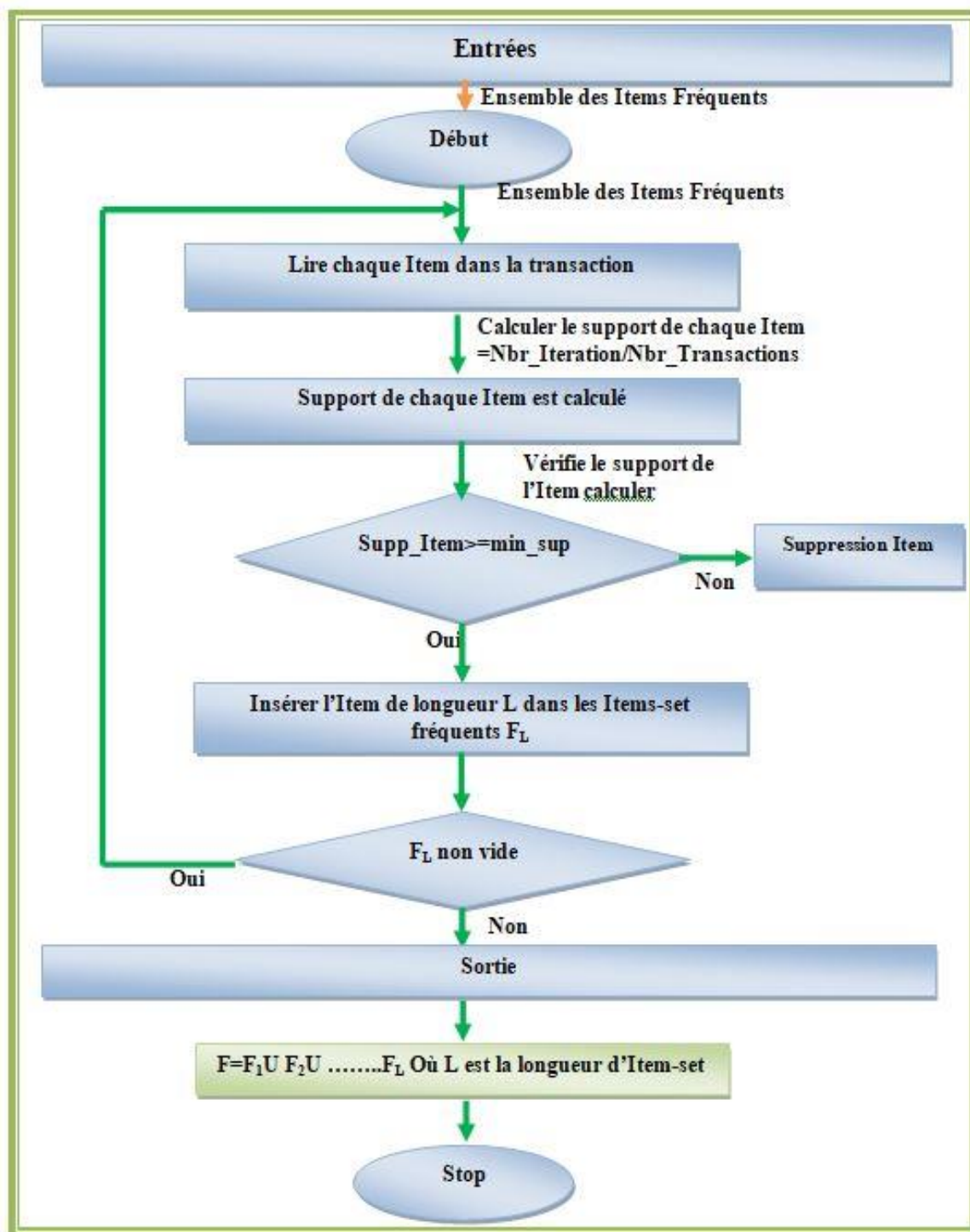


composants d'entrée de ce module, et d'en déduire une catégorisation. Apriori permet aussi de déterminer les *règles d'association* présentes dans un jeu de données, pour un *seuil de support* et un *seuil de confiance* fixés au préalable par les décideurs.



**Figure12:** Architecture du sous-système de la Classification

Le schéma suivant représente le processus de l'extraction des Itemset fréquent.



**Figure13** : Le processus de l'extraction des Itemset fréquent

Le schéma suivant représente le processus de l'extraction des règles solides

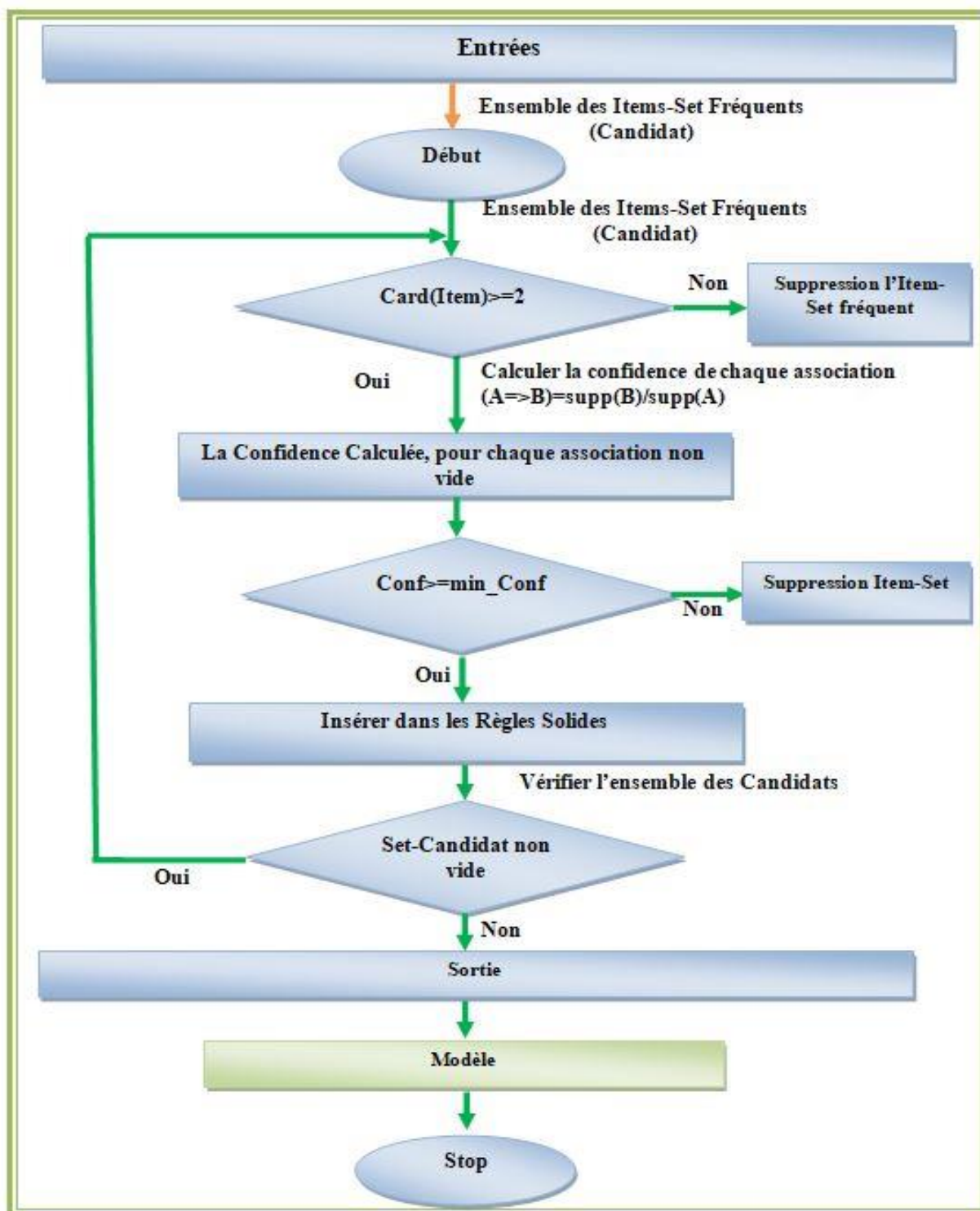


Figure14 : Le processus de l'extraction des Règles solides



### 2.2.4 Le sous-système de la visualisation

Ce module consiste à estimer la qualité, mesurer l'utilité, l'originalité et l'intelligibilité du modèle ou concept prédit. L'estimation généralement représenté par des graphiques, car plus faciles à interpréter les résultats aux décideurs de la CNAS et ils représentent toutes les connaissances qui caractérisent le facteur du Non recouvrement.

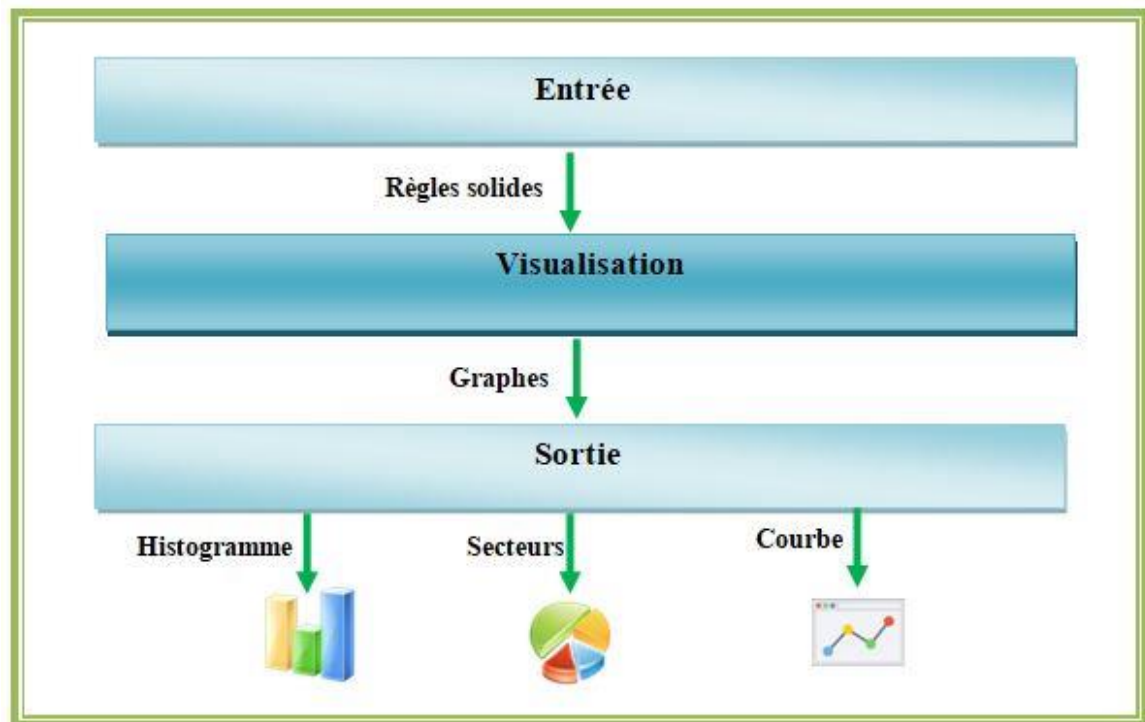


Figure15 : L'architecture du sous-système de la visualisation

## 3 Conclusion

La conception nous a permis de modéliser notre système de la classification des modèles de non recouvrement par les motifs fréquents comme étant un ensemble des sous-systèmes en interaction. Nous avons commencé par une architecture globale du système interactif proposé et nous l'avons suivi par une conception détaillée de ses différents sous-systèmes. Notre système est composé d'une séquence de sous-systèmes à savoir : sous-système d'acquisition des données cibles, sous-système de prétraitement et de filtrage des données permettant la génération d'un descripteur de modèle du non recouvrement, un sous-système d'estimation associative par les motifs

fréquents et un dernier sous-système de visualisation des données et des règles d'association du nouveau modèle du non recouvrement.

---

# Chapitre IV :

## L'implémentation et la validation du système

---

Dans ce chapitre, on traite la mise en œuvre des différents sous-systèmes de notre système proposé. On commence par la description de l'environnement matériel/logiciel, suivi du choix du langage et des outils de programmation, le code des différents sous-systèmes et on termine par les tests et la visualisation des résultats. L'environnement de développement choisit est le NETBEANS.

## 1 L'environnement matériel/logiciel

### 1.1 Matériel

On a développé notre système sur un micro portable Dell avec les caractéristiques suivantes :

- Modèle : Inspiron 15-3000 series
- Processeur : Intel Core I7
- Mémoire : 8 Go RAM et 1024 Go Disque
- Ecran/Poids : Ecran 15.6 pouces et le Poids 2.4 Kg

### 1.2 Logiciel

#### 1.2.1 Système d'exploitation utilisé

On a installé un Windows server 2003 sur une machine virtuelle en utilisant VMware Workstation (Simulation d'un serveur pour la base d'entraînement) et un Windows 7 32Bits sur la machine physique (client).

#### 1.2.2 Langage de Programmation utilisé

Pour l'implantation du notre Système on a choisi le langage de programmation orienté objet « **JAVA** » développée par Sun Microsystems. Dans un des premiers papiers sur le langage **JAVA**, SUN le décrit comme suit : « Java : a simple, object-oriented, distributed, robust, secure, architecture neutral, portable, high-performance, multithreaded, and dynamic language ». [28]

```

int j = 0; int pass = 0;
for( j = 0; j < data.length(); j++){
    /* if(data.charAt(j)=='[') ;
    else if(data.charAt(j)=='|') ;
    else if(data.charAt(j)=='\n') ;
    else */
    if(!(Character.isDigit(data.charAt(j)))) pass = 0;

    else if((Character.isDigit(data.charAt(j)))
    && ((data.charAt(j+1)=='\n') || (data.charAt(j+1)=='|'))){
        pass = 1;
        String c = ""+data.charAt(j);
        int colm = Integer.parseInt(c);

        fileWriterR.write(""+(String)jtbl.getModel().getColumnName(colm-1)+ " ");

        int[] value = new int[61];
        double somme = 0;

        int l=0;
        for(l = 0; l<jtbl.getRowCount();l++){

            value[l] = (int)jtbl.getValueAt(l, colm-1);
            fileWriterR.write(" "+value[l]);

            somme+=value[l];

```

**Figure16:**Exemple d'un Code Java.

On utilise le NetBeans comme environnement de développement.

### 1.2.3 L'environnement utilisé

Le **NetBeans** est un environnement de développement intégré (EDI), placé en open source. En plus de Java, **NetBeans** permet la prise en charge native de divers langages tels le C, le C++, le JavaScript, le XML, le Groovy, le PHP et le HTML. Il offre toutes les facilités d'un IDE moderne (éditeur en couleurs, projets multi-langage, éditeur graphique d'interfaces et de pages Web). **NetBeans** est disponible sous Windows, Linux, Solaris

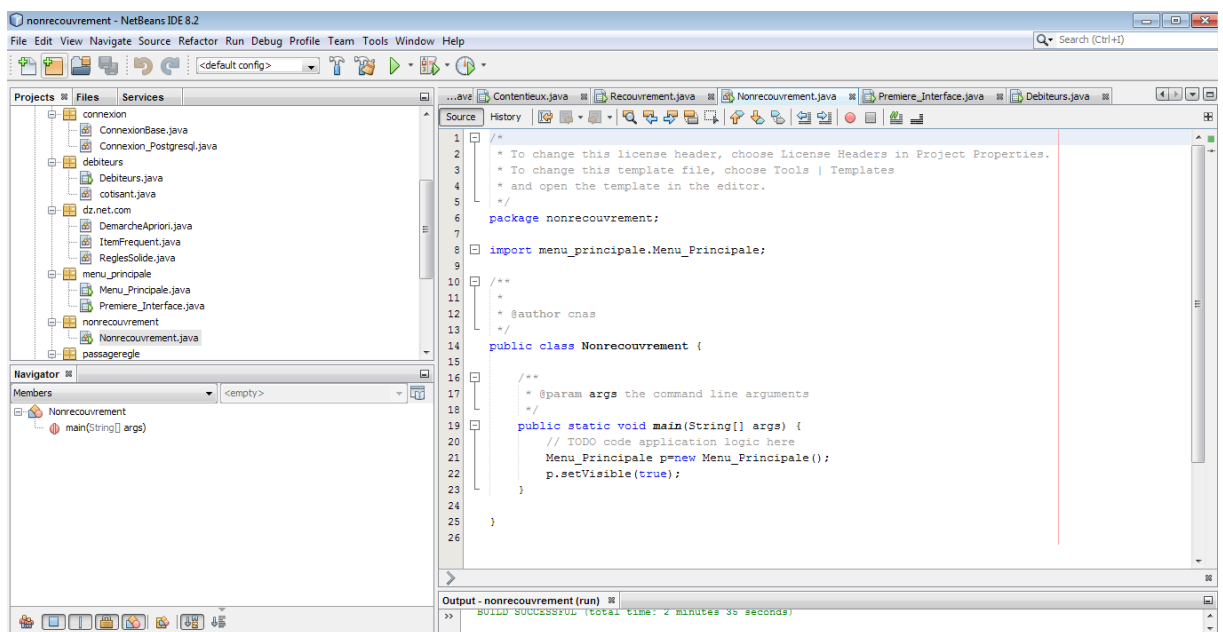


Figure17:NETBEANS.

### 1.2.4 Les SGBDS utilisé

Pour gérer les différentes bases de données utilisées dans notre système on utilise :

➤ **Oracle8i**

L'oracle8i est un système de gestion de base de données relationnelle (SGBDR) qui propose de nombreux outils à l'administrateur pour que celui-ci puisse proposer à son entreprise une gestion efficace et performante de ses applications informatiques. Oracle accepte de nombreuses plates-formes UNIX et fonctionne également dans

l'environnement Windows ou NOVELL [28]. On l'utilise Pour **la simulation de la base d'entraînement** dont les données utilisées sont **virtuelles** (ne sont pas réelle juste pour le cas d'étude).

- **PostgreSQL** : ce dernier est un des principaux SGBD-R (systèmes de gestion des bases de données relationnelles) du marché, il est libre et gratuit, parmi de ses avantages : déploiement illimité, excellent support, fiabilité et stabilité légendaires, conçu pour une grande capacité, outils graphiques de modélisation et d'administration très bonne compatibilité SQL. le **PostgreSQL** est compatible avec presque toutes les plates-formes (linux, Windows, Solaris.....etc.). En l'utilisant pour la gestion de la base de données résultat (le descripteur et les Règles Solides) [29].

## 2 L'implémentation

Notre système interactif comprend :

- Sous-système d'acquisition des données cibles
- Sous-système de prétraitement
- Sous-système de la classification par les motifs fréquents.
- Sous-système de la visualisation (test et validation)

### 2.1 Sous-système d'acquisition des données cibles

Ce sous-système permet l'acquisition de toutes les données nécessaires pour l'ensemble des acteurs.

- ✓ « **ConnexionBase.java** » et « **ConnexionPostgreSQL.java** »

Ces deux classes java permettant la connexion dynamique à la base de données d'entraînement « Employeur » et la base de données résultat « RegleSolide ». Les figures suivantes représentent le code de ces deux classes :

```

public void Connection ConnexionBase()
try {
    String url = "jdbc:oracle:thin:@192.168.1.9:1521:EMPLOYEUR";
    String user = "base";
    String psw = "base";
    Class.forName("oracle.jdbc.driver.OracleDriver");
    c = DriverManager.getConnection(url, user, psw);
    } catch (ClassNotFoundException | SQLException ex) {
    JOptionPane.showMessageDialog(null, "Erreur:" + ex.toString());
    c=null;
    }
return c;
}
    
```

**Figure18:** ConnexionBase.java

```

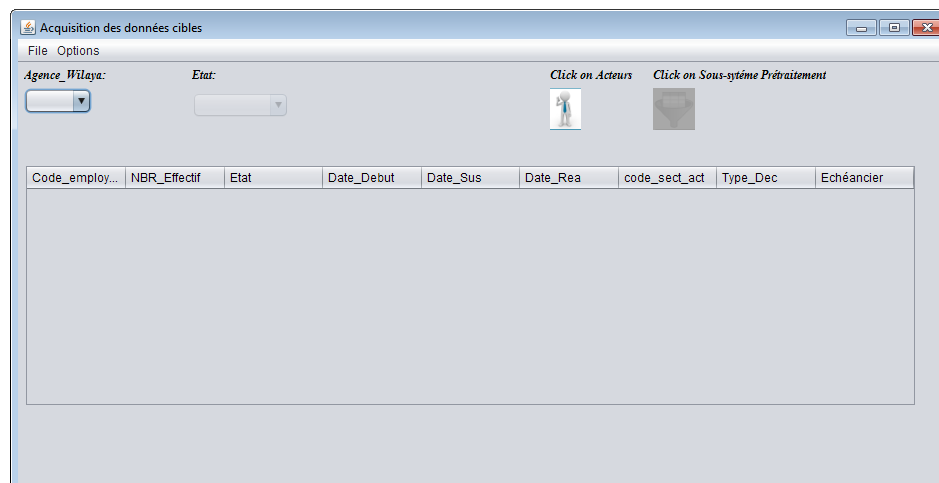
public void Connection Connexion_Postgresql()
try {
    String url = "jdbc:postgresql://localhost:5432/RegleSolide";
    String user = "postgres";
    String psw = "base";
    Class.forName("org.postgresql.Driver");
    c2 = DriverManager.getConnection(url, user, psw);
    } catch (ClassNotFoundException | SQLException ex) {
    JOptionPane.showMessageDialog(null, "Erreur:" + ex.toString());
    c2=null;
    }
return c2;
}
    
```

**Figure19:** Connexion\_Postgresql.java

✓ **La liaison entre les données**

Cette dernière est géré par les requêtes SQL .elles nous ont permis de développer l’interface « Acteurs ». Ce dernier représente l’élément d’entré du sous système prétraitement.

Les figures suivantes représentent : « Acteurs » et le code source de la méthode « Mettre\_à\_jour\_Acteurs() ».



**Figure20:** « Acteurs » avant mise à jour

```

public void Mettre_a_jour_acteurs()

String choix = jComboBox1.getSelectedItem().toString();
String code = jComboBox2.getSelectedItem().toString();
if(choix.equalsIgnoreCase("Globale")) {
    try {
        String de1 = "delete from acteurs ";
        srb.execute(de1);

        String req = "select c.code_employeur as num, h(c.code_employeur) as c, h(c.date_debut) as debut, c.date_susp as
d_sus, c.date_rea as d_rea, c.code_sect_act as act, c.nbr_effectif as eff, c.etat as etat, sum(e.mont_decla-e.mont_encais) as solde in"
        + "from obligation_encais e, employeur c "
        + "where c.agence_wilaya=" + code + " and e.code_employeur=c.code_employeur"
        + "group by c.code_employeur, c.date_debut, c.date_susp, c.date_rea, c.code_sect_act, c.nbr_effectif, c.etat order by
c.code_employeur ";
        rt = st.executeQuery(req);
        d.setRowCount(0);
        while (rt.next()) {
            String n = "";
            String ef = rt.getString("eff");
            int e = Integer.parseInt(ef);
            if(e > 9) {
                n = "Mensuelle";
            } else {
                n = "Trimestrielle";
            }
            String sum1 = rt.getString("solde");
            double sum = Double.parseDouble(sum1);
            String echeancier = "";
            if(sum > 0) {
                echeancier = "Oui";
            } else {
                echeancier = "Non";
            }

            String num = rt.getString("num");
            String cl = rt.getString("cl");
            String etat1 = rt.getString("etat");
            String etat = position(etat1);
            String debut = rt.getString("debut");
            String susp = rt.getString("d_sus");
            String rea = rt.getString("d_rea");
            String sect_act = rt.getString("act");
            String req_in = " insert into acteurs
(code_employeur, nbr_effectif, etat, date_debut, date_rea, date_susp, code_sect_act, type_dec, echeancier) values (" + num + "," + cl + "," +
ef + "," + etat1 + "," + debut + "," + susp + "," + rea + "," + sect_act + "," + n + "," + echeancier + "); commit ";
            srb.executeUpdate(req_in);
        }
    }
}

```

Figure21: Code source de la méthode Mettre\_a\_jour\_acteurs ()

Code_employ...	NBR_Effectif	Etat	Date_Debut	Date_Sus	Date_Rea	code_sect_act	Type_Dec	Echancier
0700000101	1	Actif	01/01/1998			001	Trimestrielle	OUI
0700000201	4	Actif	01/03/2001	01/01/2003	01/03/2004	04	Trimestrielle	NON
0700000301	100	Suspendu	01/08/1976	01/01/1990		15	Mensuelle	OUI
0700000401	11	Actif	01/06/1980			79	Mensuelle	NON
0700000501	1	Actif	01/01/2011			80	Trimestrielle	OUI
0700000601	10	Décédé	01/01/2010	01/01/2008		03	Mensuelle	OUI
0700000701	11	Actif	01/01/1996			17	Mensuelle	OUI
0700000801	3	Actif	01/03/2005			09	Trimestrielle	OUI
0700000901	21	Actif	01/08/1992			69	Mensuelle	OUI
0700001001	16	Actif	01/06/1999	01/07/2006	01/01/2008	02	Mensuelle	NON
0700001101	2	Actif	01/01/2016			08	Trimestrielle	OUI
0700001201	13	Actif	01/01/2010			05	Mensuelle	OUI
07000013	4	Actif	01/01/2016			54	Trimestrielle	NON

Figure22: « Acteurs » après mise à jour

## 2.2 Sous-système de prétraitement

L'implémentation de ce 2eme sous-système de notre conception se fait par la programmation de trois classes java, dont « Recouvrement .java » et « Contentieux.java »



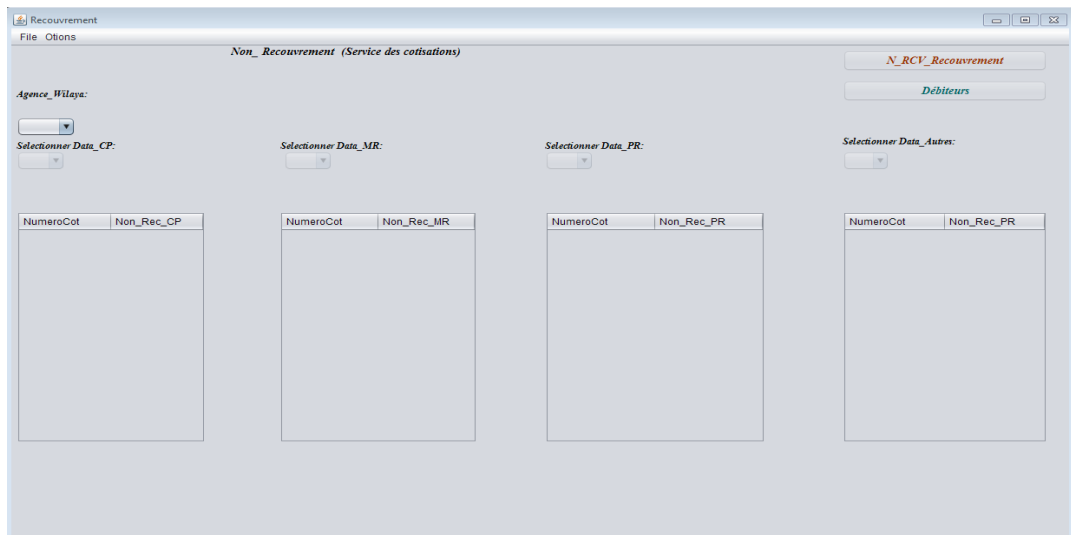
représentent les composants internes du sous système et « Débiteurs.java » qui est notre descripteur représente la sortie du sous-système de prétraitement.

### 2.2.1 Les composants internes

#### ➤ Recouvrement.java

Cette classe nous permet d'extraire des attributs associant les données du recouvrement, qui représentent un non recouvrement pour la CNAS (données cotisation principale « Data\_CP », données majoration de retard « Data\_MR », données pénalité de retard « Data\_PR », et données Autres « Data\_Autres »).

Les figures suivantes décrivent le composant « **Recouvrement** »



**Figure23:** Recouvrement avant calcul\_data

Le code source suivant représente le calcul de Data\_Cp en terme recouvrement. (La même chose pour Data\_MR, DAT\_PR, DATA\_Autres).

```

public void Calcul_Data() //CP Recouvrement
{
    jButton1.setCursor(Cursor.getPredefinedCursor(Cursor.WAIT_CURSOR));
    jLabel12.setText("");
    jLabel11.setText("");
    jLabel10.setText("");
    jLabel17.setText("");
    //Affichage au niveau du tableau Cotisation principale
    String cp = JComboBox1.getSelectedItem().toString();
    String code = JComboBox3.getSelectedItem().toString();
    if (!cp.equalsIgnoreCase("")) {
        try {
            String req2 = "select t.code_employeur as code_employeur,sum(t.mont_deda - t.mont_encaiss) as solde"
            + " from obligation_encaiss t, employeur c \n"
            + " where c.agence_wilaya=" + code + " and t.motif=NR" + cp + " and t.code_employeur=c.code_employeur \n"
            + " having sum(t.mont_deda-t.mont_encaiss)>0\n"
            + " group by t.code_employeur order by t.code_employeur";
            rt2 = st.executeQuery(req2);
            d.setRowCount(0);
            while (rt2.next()) {
                String num = rt2.getString("code_employeur");
                String α = rt2.getString("solde");
                String t[] = new String[]{num, α};
                d.addRow(t);
            } catch (SQLException ex) {
                JOptionPane.showMessageDialog(null, ex.toString());
            }
        }
    }
}

```

Figure24:le code source de la méthode calcul\_data () pour le recouvrement

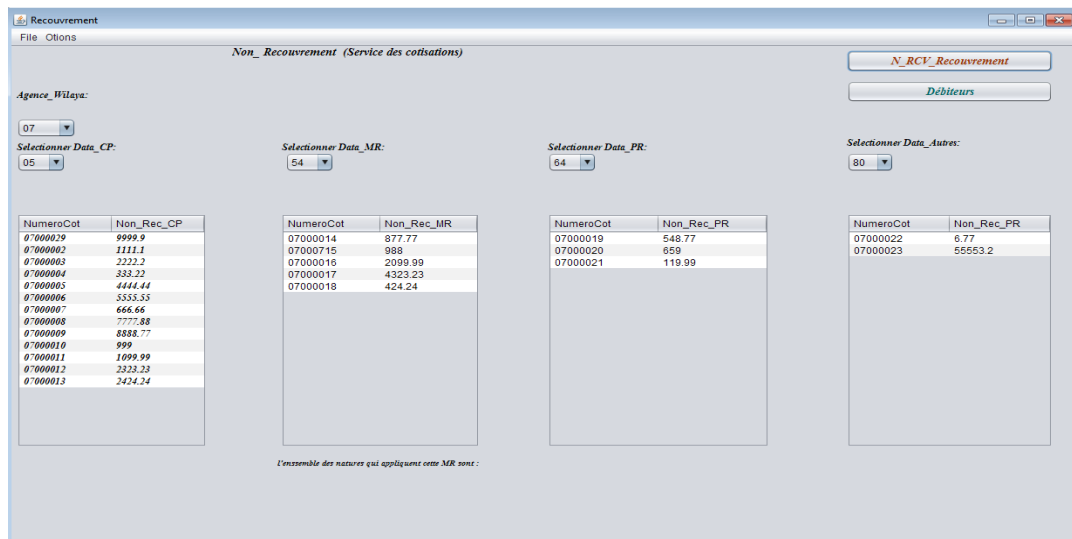
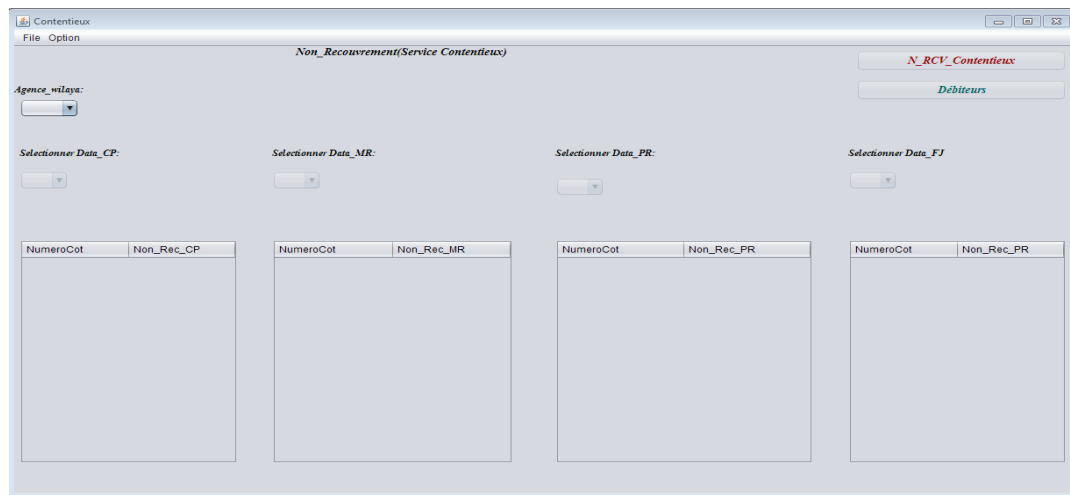


Figure25: Le Recouvrement après calcul\_data

### ➤ Contentieux.java

Cette classe nous permet de traiter par extraction des attributs associant les données du contentieux représentant un non recouvrement (Données Cotisation principale « Data\_CP », données majoration de retard « Data\_MR », données pénalité de retard « Data\_PR » et celles des frais de justice « Data\_FJ »).

Les figures suivantes décrivent le composant « Contentieux »



**Figure26 :** Le « Contentieux » avant calcul\_data

Le code source suivant représente le calcul de Data\_Cp en terme contentieux.(la même chose pour Data\_MR, DAT\_PR, DATA\_FJ).

```

public void Calcul_Data() //CP Contentieux
{
    jButton1.setCursor(Cursor.getPredefinedCursor(Cursor.WAIT_CURSOR));
    jLabel12.setText("");
    jLabel11.setText("");
    jLabel10.setText("");
    jLabel17.setText("");
    //Affichage au niveau du tableau Cotisation principale
    String cp = JComboBox1.getSelectedItem().toString();
    String code = JComboBox3.getSelectedItem().toString();
    if (!cp.equalsIgnoreCase("")) {
        try {
            String req2 = "select t.code_employeur as code_employeur,sum(t.mont_deda - t.mont_encaiss) as solde"
+ " from obligation_encaiss t, employeur c \n"
+ " where c.agence_wilaya=" + code + " and t.motif='N' + cp + " and t.code_employeur=c.code_employeur \n"
+ " having sum(t.mont_deda-t.mont_encaiss)>0\n"
+ " group by t.code_employeur order by t.code_employeur";
            rt2 = st.executeQuery(req2);
            d.setRowCount(0);
            while (rt2.next()) {
                String num = rt2.getString("code_employeur");
                String cr = rt2.getString("solde");
                String t[] = new String[]{num, cr};
                d.addRow(t);
            }
        } catch (SQLException ex) {
            JOptionPane.showMessageDialog(null, ex.toString());
        }
    }
}

```

**Figure28:** le code source de la méthode calcul\_data() pour le « Contentieux »

The screenshot shows a software window titled 'Contentieux' with a menu bar (File, Option) and a title bar 'Non\_Recouvrement(Service Contentieux)'. It features several dropdown menus for 'Agence\_Wilaya' (07), 'Selectionner Data\_CP' (05), 'Selectionner Data\_MR' (09), 'Selectionner Data\_PR' (10), and 'Selectionner Data\_FJ' (81). There are two buttons: 'N\_RCV\_Contentieux' and 'Débiteurs'. Below are four data tables:

NumeroCot	Non_Rec_CP
07000101	9999.9
07000030	1111.1
07000033	2222.2
07000004	333.22
07000036	5555.55
07000037	666.66
07000039	777.88
07000039	8888.77
07000040	999
07000060	1099.99
07000062	2323.23
07000013	2424.24

NumeroCot	Non_Rec_MR
07000014	877.77
07000015	988
07000001	2099.99
07000007	4323.23
07000008	424.24
07000005	4444.44

NumeroCot	Non_Rec_PR
07000021	548.77
07000022	659
07000008	119.99

NumeroCot	Non_Rec_PR
07000025	6.77
07000028	58553.2
07000005	66.44

**Figure29:** Le « Contentieux » après calcul\_data

## 2.2.2 La sortie du sous-système

### ➤ Débiteurs.java

Ce descripteur se présente sous forme d'une table binaire dont les lignes représentent l'ensemble des individus dit « Débiteurs » du non recouvrement de la CNAS, et les colonnes représentent les Itemsets fréquent du non recouvrement.

Le résultat des Itemsets extraits, sauvegardés dans une base de données « Btransac.txt » et le détail des Itemsets de chaque individu dans la table det\_debiteurs de la base de données « RégleSolide » qui a été créée dans le PostgreSQL. Les figures suivantes représentent le composant « débiteurs », ainsi le code source de la méthode mettre\_à\_jour\_débiteurs ().

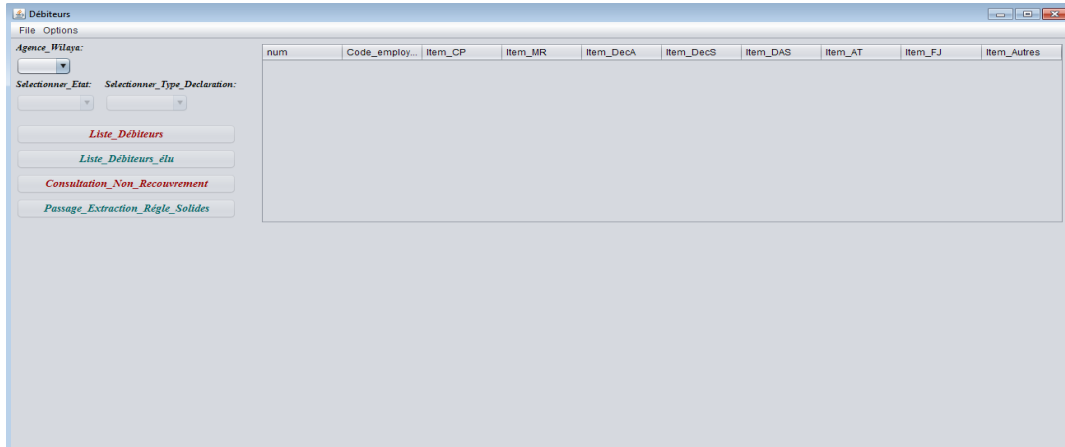


Figure30: Le « Débiteurs » avant mise à jour

```

public void Debiteurs()
{
    String po = jComboBox2.getSelectedItemAt(0).toString(); String c = jComboBox1.getSelectedItemAt(0).toString();
    String del = "delete fromdebiteur "; String del1 = "delete from det_debitteur ";
    if (ssq.equalsIgnoreCase("")) { stb.execute(del); stb.execute(del1);
    } else { del = del + " where " + ssq; del1 = del1 + " where " + ssq;
    stb.execute(del);
    stb.execute(del1);
    }
    try {
        String choix = jComboBox1.getSelectedItemAt(0).toString();
        String code = jComboBox3.getSelectedItemAt(0).toString();
        write = new BufferedWriter(new FileWriter(new File("BTransac.txt")));
        try {
            String req = "Select a.code_employeur as code_emp, de(a.code_employeur) as de ,b.etat as etat,b.nbr_effectifas
            eff \n"
            + "sum(decode(a.motif,'NRCP',a.mont_deda- a.mont_encaiss,'NCCP',a.mont_deda- a.mont_encaiss)) as
            CP,\n"
            + "sum(decode(a.motif,'NRMR', a.mont_deda- a.mont_encaiss,'NQMR', a.mont_deda- a.mont_encaiss)) as
            MR,\n"
            + "sum(decode(a.motif,'NCFJ',a.mont_deda- a.mont_encaiss)) as FJ ,\n"
            + "sum(decode(a.motif,'NRCP',null,'NRMR',null,'NRPR',null,'FJ',null,a.mont_deda- a.mont_encaiss)) as
            Autres , \n"
            + "sum(decode(a.motif,'NRART56',a.mont_deda- a.mont_encaiss,'NCART56',a.mont_deda-
            a.mont_encaiss)) as decS,\n"
            + "sum(decode(a.motif,'NRART55',a.mont_deda- a.mont_encaiss,'NCART55',a.mont_deda-
            a.mont_encaiss)) as decA,\n"
            + "sum(decode(a.motif,'NRART41',a.mont_deda- a.mont_encaiss,' NCART41',a.mont_deda-
            a.mont_encaiss)) as das,\n"
            + "sum(decode(a.motif,'NRART59',a.mont_deda- a.mont_encaiss,'NCART59',a.mont_deda-
            a.mont_encaiss)) as AT \n"
            + " from obligation_encaissa , employeurb \n" + " where b.agence_wilaya=" + code + ""\n" + " and
            b.code_employeur=a.code_employeur\n" + " and b.etat=" + s + ""\n" + " group by a.code_employeur \n" + " having
            sum(a.mont_deda- a.mont_encaiss)> 0\n" + " order by a.code_employeur";
            cotisant p = new cotisant();
            rt = st.executeQuery(req);d.setRowCount(0); compt = 0;
            while (rt.next()) {
                String num = rt.getString("code_emp");
                compt++; String cle = rt.getString("de"); String etat = rt.getString("etat");String effec = rt.getString("eff");
                String ligne = ""; String CP1 = "";
                String CP = p.calculer_cp(num);
                if (CP == null || CP == "") { CP1 = "0";
                } else { ligne = ligne.concat(CP); CP1 = "1"; }
                String MR = "0",mr="0"; double v = 0; m r = rt.getString("MR");
            }
        }
    }
}

```

Figure31: le code source de la méthode mettre\_à\_jour\_débiteurs() du « Débiteurs » Première Partie

```

public void Debiteur() ( suite)

if (mr == null || ans.equalsIgnoreCase("") || mr.equalsIgnoreCase("0")) { MR = "0";
} else { if (lans.equalsIgnoreCase("0") && lans.equalsIgnoreCase("") && lans.equalsIgnoreCase(" ")) {
ligne = ligne + "MR "; MR = "1"; } }
String decs="0",deca="0", das1="0", at="0";
String decS = rt.getString("decS");String decA = rt.getString("decA");
String das = rt.getString("das");String AT = rt.getString("AT");
if (decS == null || decS.equalsIgnoreCase("") || decS.equalsIgnoreCase("0")) {
decs= "0"; } if (ldecS.equalsIgnoreCase("0") && ldecS.equalsIgnoreCase("") && ldecS.equalsIgnoreCase(" ")) {
ligne = ligne + "PR ";decs = "1";
}
}
if (deca == null || decA.equalsIgnoreCase("") || decA.equalsIgnoreCase("0")) {
deca = "0"; }
if (ldeca.equalsIgnoreCase("0") && ldeca.equalsIgnoreCase("") && ldeca.equalsIgnoreCase(" ")) {
ligne = ligne + "PR ";deca = "1"; }
if (das == null || das.equalsIgnoreCase("") || das.equalsIgnoreCase("0")) {
das1 = "0"; }
if (ldas.equalsIgnoreCase("0") && ldas.equalsIgnoreCase("") && ldas.equalsIgnoreCase(" ")) {
ligne = ligne + "PR "; das1= "1"; }
if (AT == null || AT.equalsIgnoreCase("") || AT.equalsIgnoreCase("0")) { at = "0";
} if (!AT.equalsIgnoreCase("0") && !AT.equalsIgnoreCase("") && !AT.equalsIgnoreCase(" ")) {
ligne = ligne + "PR_ART26 "; at = "1"; } String fj = "", fj1 = "0";
fj = rt.getString("FJ");
if (fj == null || fj.equalsIgnoreCase("") || fj.equalsIgnoreCase("0")) { fj1= "0";
} else { if (lfj.equalsIgnoreCase("0") && lfj.equalsIgnoreCase("") && lfj.equalsIgnoreCase(" ")) {
ligne = ligne + "FJ "; fj1 = "1"; }
String au = "", au1 = "0"; au = rt.getString("Autres");
if (au == null || au.equalsIgnoreCase("") || au.equalsIgnoreCase("0")) { au1 = "0";
} else { if (lau.equalsIgnoreCase("0") && lau.equalsIgnoreCase("") && lau.equalsIgnoreCase(" ")) {
ligne = ligne + "Autres"; au1= "1"; } }
String reqin = "insertintodebiteur (num erocot,de,position1,effectif,cp,mr,pr,fj,autres) values (" +code_emp
+ "," + de + "," + posit + "," + effec + "," + CP + "," + mr + "," + PR + "," + fj + "," + au + ")"; commit;";
sib.executeUpdate(reqin); write.write(ligne); write.newLine(); String compt1 = String.valueOf(compt);
String t[] = new String[] {compt1, rt.getString("code_emp"), CP1, MR, deca, decs, das1, at, fj1, au 1};
d.addRow(t); } } catch (SQLException ex) {
JOptionPane.showMessageDialog(null, ex.toString()); } }

```

Figure32: le code source de la méthode mettre\_à\_jour\_débiteurs () du « Débiteurs » deuxième Partie

num	Code_employ...	Item_CP	Item_MR	Item_DecA	Item_DecS	Item_DAS	Item_AT	Item_FJ	Item_Autres
1	07000001	1	1	0	0	0	0	0	1
2	07000002	1	1	0	1	0	0	0	1
3	07000003	1	1	0	1	1	0	0	1
4	07000004	1	0	0	0	0	0	0	1
5	07000005	0	1	0	0	0	0	0	1
6	07000006	1	1	0	1	0	0	0	1
7	07000007	1	0	0	1	0	0	0	1
8	07000008	0	1	0	1	1	0	0	0
9	07000009	0	1	0	1	0	0	0	1
10	07000010	1	0	0	0	0	0	0	1
11	07000011	1	1	0	0	0	0	0	1
12	07000012	1	1	0	0	1	0	0	0
13	07000013	0	1	0	0	0	0	0	1
14	07000014	1	0	0	0	0	0	0	0

Figure33: « Débiteurs » après mise à jour

code_employeur	Effectif	Etat	Datedebu	Datesusp	Datereac	CP	MR	PR	FJ	Autres
07000005	1	Actif	01.01.2011			0	444.44	0	0	66.44
07000001	1	Actif	01.01.1998			999.9	24.06	0	0	6.66
07000008	3	Actif	01.03.2005				424.24	119.99	0	6.66

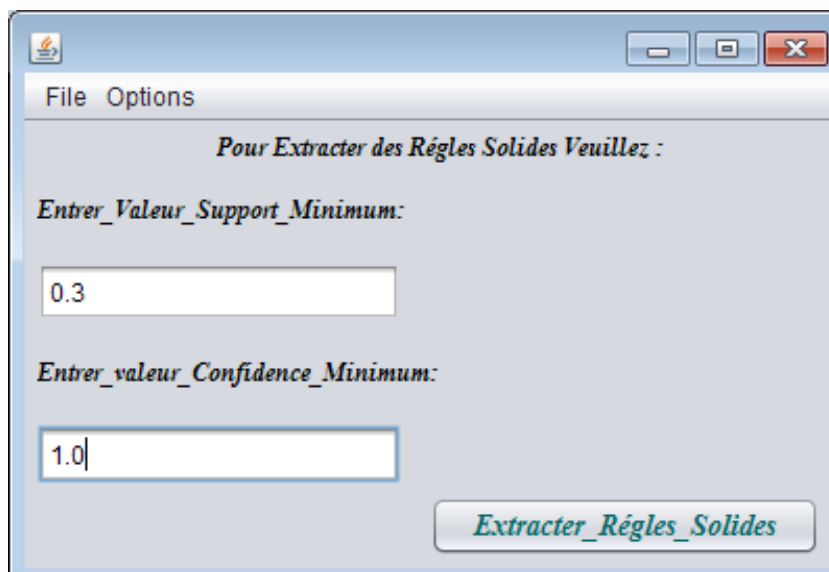
Figure34: Consultation des individus sélectionnés

### 2.3 Sous-système de la classification

Ce 3eme sous-système de notre conception (La classification) a été implémenté par la programmation d'une seule classe java.

➤ **AlgorithmeApriori.java**

Dans Cette classe on applique l'algorithme Apriori sur la base de données « Btransac.txt » et on extrait l'ensemble des règles solides, ces derniers sont sauvegardés dans une table « reglesol » dans la base de données « RègleSolide ». les figures suivante représentent le sous système de la « Classification», ainsi le code source de l'Algorithme *inspiré* dit Apriori.[15]



**Figure35:** l'entrée des valeurs des seuils minimaux

ItemFréquent:				Les Règles Solides	
ItemSet	Fréquence	Support	Longueur	Les Règles Solides	confiance
Motif1	28	0.93	1	Motif5=>Motif16	1.88
Motif2	20	0.66	1	Motif3=>Motif17	1.06
Motif3	27	0.9	1	Motif1=>Motif20	1.0
Motif4	10	0.33	1	Motif20=>Motif1	1.0
Motif5	16	0.53	1	Motif2=>Motif18	1.51
Motif6	13	0.43	1		
Motif7	19	0.63	2		
Motif8	25	0.83	2		
Motif9	10	0.33	2		
Motif10	15	0.5	2		
Motif11	12	0.4	2		
Motif12	20	0.66	2		
Motif13	13	0.43	2		
Motif14	15	0.5	2		
Motif15	11	0.36	2		

[Visualisation Graphique](#)

**Figure36:** L'extraction des Itemsets fréquent et les Règles Solides

### 3 Les tests et la visualisation des résultats

Après avoir présenté les principaux sous-systèmes de notre système de la Classification de modèles de problème de non recouvrement par la C.N.A.S par la technique d'association et motifs fréquent, nous essayons dans cette phase de présenter les résultats obtenues.

La fenêtre principale de notre système s'affiche comme suit :





**Figure37:** L'interface principale du système proposé : La classification associative de Modèles du non recouvrement par la CNAS par la technique des motifs fréquents.

Notre système cyclique et diviser en trois parties :

- 1) La présentation du sous système d'acquisition des données cibles



**Figure50:** sous système d'acquisition des données cibles

Acquisition des données cibles

File Options

Agence\_Wilaya: 07 Agence Bisra

Etat: Globale

Click on Acteurs

Click on Sous-système Prétraitement

Code_employ...	NBR_Effectif	Etat	Date_Debut	Date_Sus	Date_Rea	code_sect_act	Type_Dec	Echéancier
0700000101	1	Actif	01/01/1998			001	Trimestrielle	OUI
0700000201	4	Actif	01/03/2001	01/01/2003	01/03/2004	04	Trimestrielle	NON
0700000301	100	Suspendu	01/08/1976	01/01/1990		15	Mensuelle	OUI
0700000401	11	Actif	01/06/1980			79	Mensuelle	NON
0700000501	1	Actif	01/01/2011			80	Trimestrielle	OUI
0700000601	10	Décédé	01/01/2010	01/01/2008		03	Mensuelle	OUI
0700000701	11	Actif	01/01/1996			17	Mensuelle	OUI
0700000801	3	Actif	01/03/2005			09	Trimestrielle	OUI
0700000901	21	Actif	01/08/1992			69	Mensuelle	OUI
0700001001	16	Actif	01/06/1999	01/07/2006	01/01/2008	02	Mensuelle	NON
0700001101	2	Actif	01/01/2016			08	Trimestrielle	OUI
0700001201	13	Actif	01/01/2010			05	Mensuelle	OUI
07000013	4	Actif	01/01/2016			54	Trimestrielle	NON

Figure38: les Acteurs

2) La présentation du sous système Prétraitement



Figure39: sous système de prétraitement

- Recouvrement

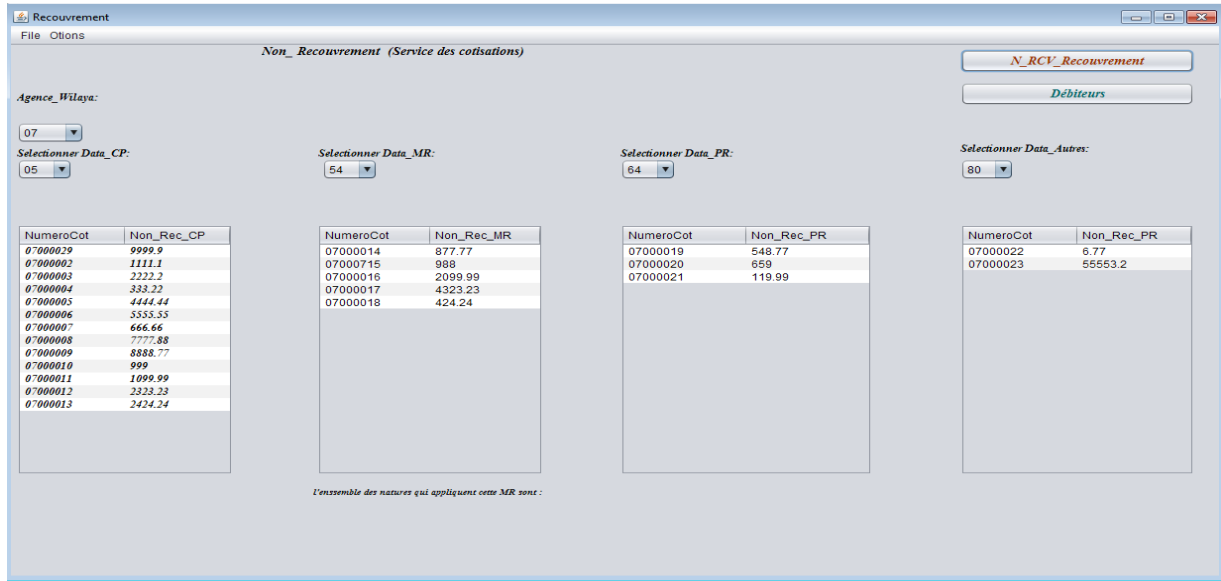


Figure40: Le Recouvrement

- Contentieux

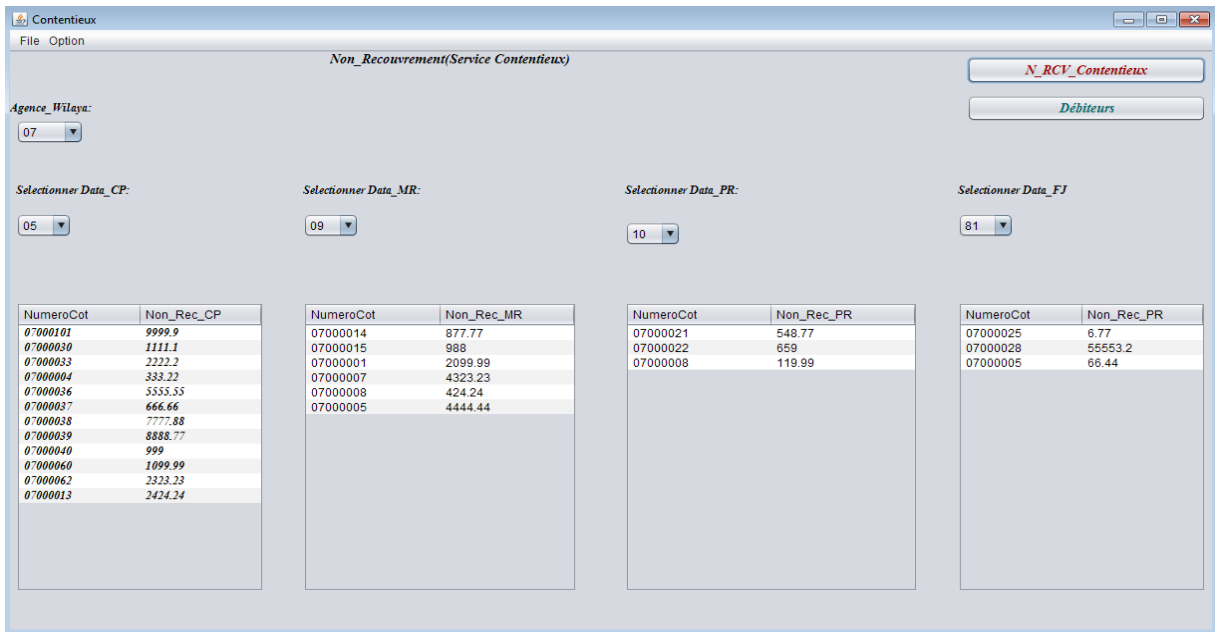


Figure41: Le Contentieux

- Débiteurs

The screenshot shows a software window titled 'Débiteurs'. On the left, there are controls for 'Agence\_Wilaya' (07), 'Selectionner\_Etat' (Actifs), and 'Selectionner\_Type\_Declaration' (Trimestrielle). Below these are buttons for 'Liste\_Débiteurs', 'Liste\_Débiteurs\_élu', 'Consultation\_Non\_Recouvrement', and 'Passage\_Extraction\_Règle\_Solides'. The main area contains a table with 10 columns: num, Code\_employ..., Item\_CP, Item\_MR, Item\_DecA, Item\_DecS, Item\_DAS, Item\_AT, Item\_FJ, and Item\_Autres. Below this is a summary table with 11 columns: code\_employeur, Effectif, Etat, Datedebu, Datesusp, Datereac, CP, MR, PR, FJ, and Autres.

num	Code_employ...	Item_CP	Item_MR	Item_DecA	Item_DecS	Item_DAS	Item_AT	Item_FJ	Item_Autres
1	07000001	1	1	0	0	0	0	0	1
2	07000002	1	1	0	1	0	0	0	1
3	07000003	1	1	0	1	1	0	0	1
4	07000004	1	0	0	0	0	0	0	1
5	07000005	0	1	0	0	0	0	0	1
6	07000006	1	1	0	1	0	0	0	1
7	07000007	1	0	0	1	0	0	0	1
8	07000008	0	1	0	1	1	0	0	0
9	07000009	0	1	0	1	0	0	0	1
10	07000010	1	0	0	0	0	0	0	1
11	07000011	1	1	0	0	0	0	0	1
12	07000012	1	1	0	0	1	0	0	0
13	07000013	0	1	0	0	0	0	0	1
14	07000014	1	0	0	0	0	0	0	0

code_employeur	Effectif	Etat	Datedebu	Datesusp	Datereac	CP	MR	PR	FJ	Autres
07000005	1	Actif	01/01/2011			0	4444.44	0	0	66.44
07000001	1	Actif	01/01/1998			999.9	24.06	0	0	6.66
07000008	3	Actif	01/03/2005			0	424.24	119.99	0	6.66

Figure42: les Débiteurs

3) Le sous système de la classification

The screenshot shows a dialog box titled 'Pour Extraire des Règles Solides Veuillez :'. It contains two input fields: 'Entrer\_Valeur\_Support\_Minimum:' with the value '0.3' and 'Entrer\_valeur\_Confidence\_Minimum:' with the value '1.0'. A button labeled 'Extraire\_Règles\_Solides' is at the bottom right.

Figure43: les valeurs entrées de seuils minimaux

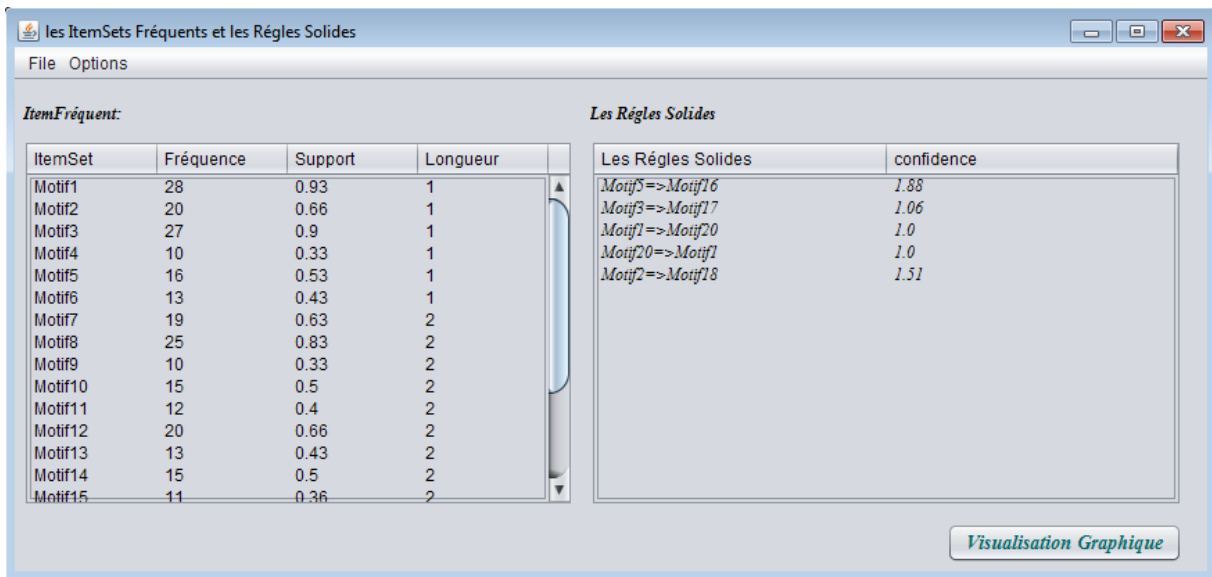


Figure44: l'extraction Item fréquent et les règles solide

4) Le sous système de la visualisation du modèle

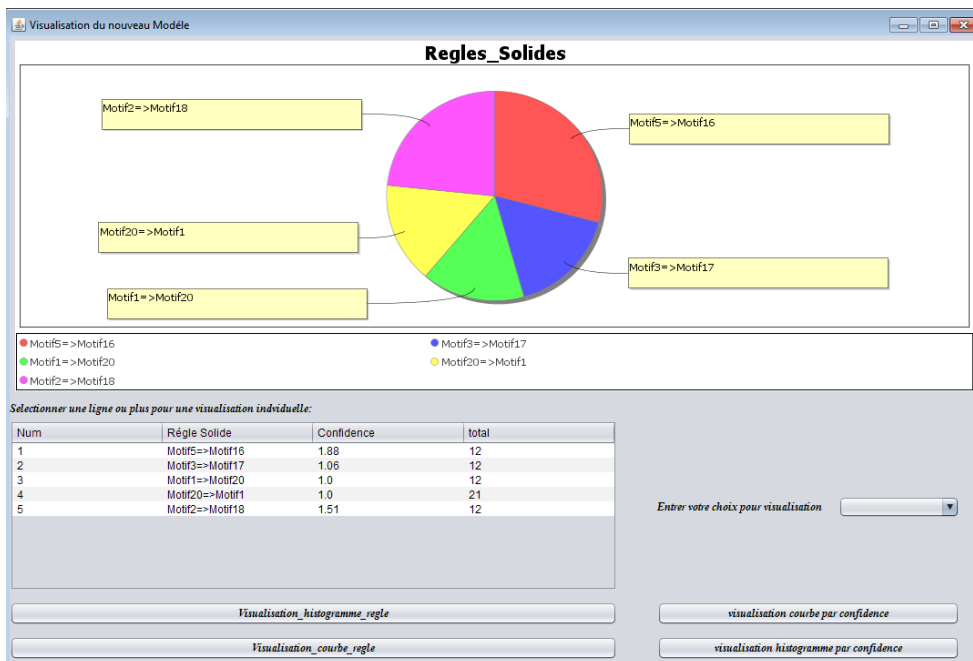


Figure45: le sous système de la visualisation

Les figures suivantes représentent les visualisations par histogramme et courbe, suite aux choix de la sélection des règles ou intervalle de confiance.

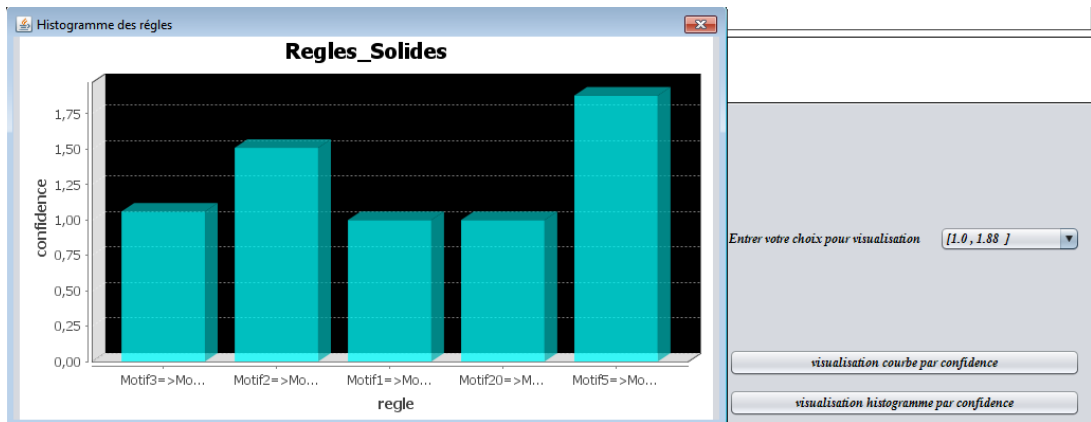


Figure46: Histogramme des règles

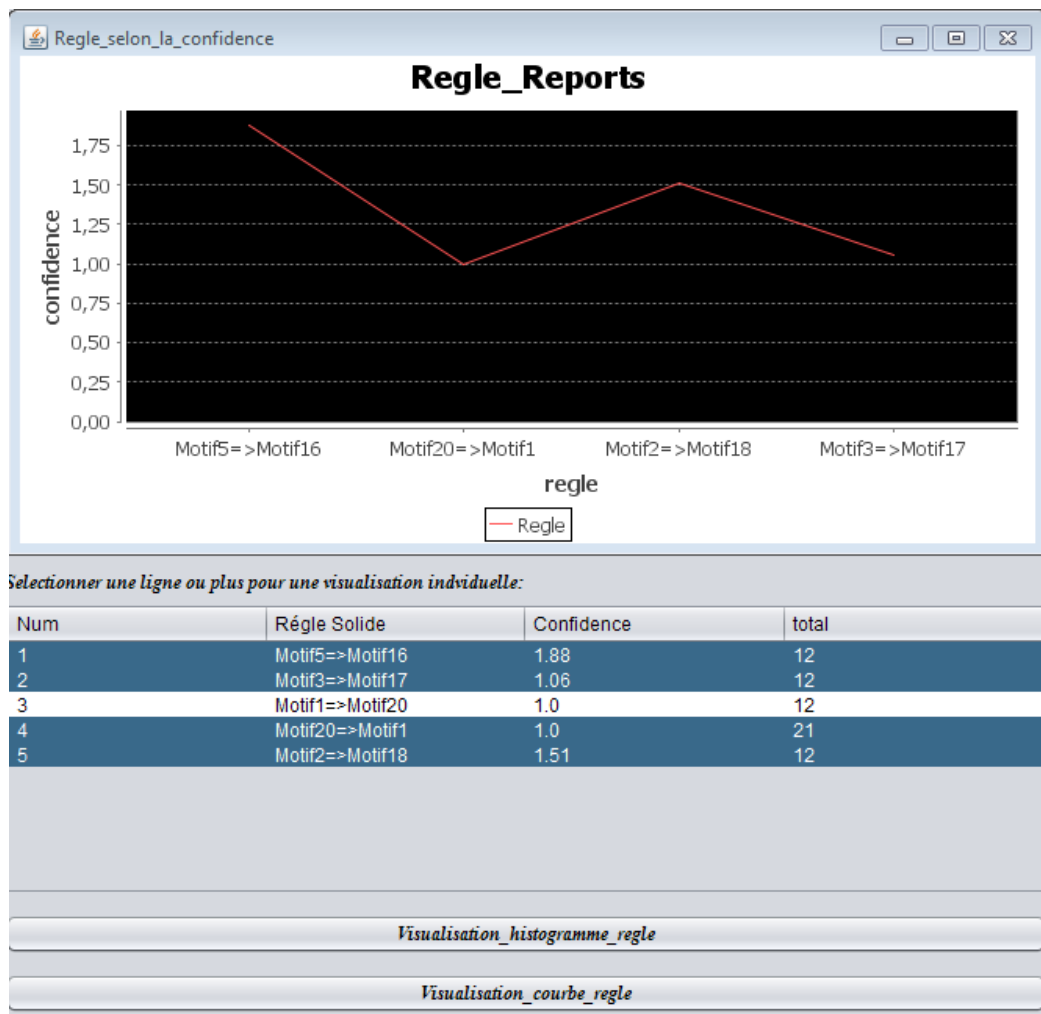


Figure47: Courbe des règles

#### **4 Conclusion**

Dans ce chapitre, Nous avons représenté l'implémentation de notre système de La classification du Modèle de non recouvrement par la CNAS par la technique d'association et motif fréquent. Premièrement, on a décrit l'implémentation des différents modules composant notre système proposé de la classification par les motifs fréquents, deuxièmes, on a présenté un test de notre réalisation dans le but de validation par simulation de notre système et Enfin nous avons essayé de visualiser les résultats obtenus dans notre système de classification des modèles du non recouvrement par la technique d'association et motif fréquent.

# **Conclusion Générale**



## Conclusion générale

Ce projet de fin d'étude consiste à l'étude de la classification associative des modèles du non recouvrement par la CNAS par les motifs fréquents avec un cas d'étude : la CNAS du Biskra. on a présenté une étude de la structure, l'organisation, le fonctionnement, et le financement de la CNAS. On a focalisé et accordé une importance particulière au processus du recouvrement des cotisations par l'organisme CNAS. on a enchainé notre étude par des préliminaires sur les différentes définitions et concepts du processus de l'extraction des connaissances à partir des données « ECD » et ses étapes, en particulier l'étape de datamining où on a noté une importance particulière à la classification associative par les motifs fréquents, en utilisant l'algorithme Apriori [15] afin d'avoir un classifieur.

Pour pouvoir répondre aux besoins du non recouvrement des cotisations, nous avons procédé à une analyse du processus de **l'extraction des connaissances** à partir des données de la base d'entraînement **simulé** à celle de la CNAS de Biskra (**Données ne sont pas réelles**). Le système proposé comprend quatre sous système à savoir : le sous-système d'acquisition des données cibles, le sous-système du prétraitement, le sous-système de la classification par les motifs fréquents et enfin le sous-système de visualisation de résultats des nouveaux modèles. Le premier sous système répond aux besoins de l'acquisition des données cibles du recouvrement financier des cotisations à explorer, le deuxième prévoit un **prétraitement** sur les données cibles pour avoir uniquement les données utiles et nécessaire qui décrivent le non recouvrement tel que les données de la cotisations principale, la majoration de retard, la pénalité de retard, le frais de justice et Autres données(data\_CP, data\_MR, data\_PR, data\_FJ, data\_Autres) et à homogénéiser ces données pour pouvoir les disposer dans un **descripteur dit « Débiteurs »**, la classification associative par les motifs fréquent a été adapté pour l'estimation d'un nouveau modèle du non recouvrement dans le but d'aider le décideur de la CNAS et de diminuer le taux du non recouvrement des cotisations.

Le système proposé a été implémenté dans un environnement informatique universel et tester sur des données d'entraînement simulé dans les résultats sont très appréciable .

On propose des recommandations pour l'utilité de notre système de la classification associative du modèle de non recouvrement par les motifs fréquents tels que :

- ✓ L'utilisation des données réelles pour la validation des modèles
- ✓ La prise en considération du système de classification par les motifs fréquent par l'organe décisionnel de la CNAS.

# Références Bibliographiques

## Références bibliographiques

- [1] Histoire de la sécurité sociale en Algérie :  
<https://abdelkadirremal.wordpress.com/2014/06/22/histoire-se-securite-social-en-algerie/>
- [2] Le système actuel de la sécurité sociale en Algérie :  
<https://www.djazair.com/fr/lesoirdalgerie/89524>
- [3] La présentation du système de sécurité sociale Algérien :  
[http://www.conselho.saude.gov.br/cm/docs/presentation\\_securite\\_sociale\\_algerie.pdf](http://www.conselho.saude.gov.br/cm/docs/presentation_securite_sociale_algerie.pdf)
- [4] Décret exécutif n° 92-07 du 4 janvier 1992 portant statut juridique des caisses de sécurité sociale et organisation administrative et financière de la sécurité sociale :  
[http://casnos.com.dz/doc/decret\\_executif\\_n92-07.pdf](http://casnos.com.dz/doc/decret_executif_n92-07.pdf)
- [5] Le régime algérien de sécurité sociale (salariés) :  
[http://www.cleiss.fr/docs/regimes/regime\\_algerie\\_salaries.html](http://www.cleiss.fr/docs/regimes/regime_algerie_salaries.html)
- [6] Loi n° 83-11 du 2 juillet 1983 relative aux assurances sociales :  
[http://casnos.com.dz/doc/loi\\_n83-11.pdf](http://casnos.com.dz/doc/loi_n83-11.pdf)
- [7] Loi n° 83-12 du 2 juillet 1983 relative à la retraite :  
[http://www.elmouwatin.dz/IMG/pdf/decret\\_83-12.pdf](http://www.elmouwatin.dz/IMG/pdf/decret_83-12.pdf)
- [8] Loi n° 83-13 du 2 juillet 1983 relative aux accidents du travail et aux maladies professionnelles :  
[http://www.elmouwatin.dz/IMG/pdf/loi\\_83-13.pdf](http://www.elmouwatin.dz/IMG/pdf/loi_83-13.pdf)
- [9] Loi n° 83-14 du 2 juillet 1983 relative obligations des assujettis en matière de sécurité sociale :  
<http://www.mtp.gov.dz/GUIDE%20JURIDIQUE/accidents-de-travail-maladies-professionnelles/4-Loi-n83-14.pdf>
- [10] Loi n° 83-15 du 2 juillet 1983 relative au contentieux en matière de sécurité sociale :  
[http://casnos.com.dz/doc/loi\\_n83-15.pdf](http://casnos.com.dz/doc/loi_n83-15.pdf)
- [11] la loi 08-08 du 23 février 2008 : relative au contentieux en matière de sécurité sociale :  
[http://casnos.com.dz/doc/loi\\_08-08.pdf](http://casnos.com.dz/doc/loi_08-08.pdf)
- [12] Guillaume CALAS . [SCIA] EPITA 2009 Études des principaux algorithmes de data mining :  
<http://guillaume.calas.free.fr/data/Publications/DM-Algos.pdf>

- [13] Gilbert Saporta.« DATA MINING » ou FOUILLE DE DONNÉES :  
<https://cedric.cnam.fr/fichiers/RC1034.pdf>
- [14] Yves Lechevallier.INRIA-Rocquencourt .Master ISI 2010-2011  
[https://www.lri.fr/~antoine/Courses/Master-ISI/ISI\\_apriori.pdf](https://www.lri.fr/~antoine/Courses/Master-ISI/ISI_apriori.pdf)
- [15] La recherche de motifs fréquents :une méthode de fouille de données :  
<http://bruno.maitresdumonde.com/optinfo/ads-scie/2007.pdf>
- [16] François RIOULT .GREYC - Équipe Données-Documents-Langues. Université de  
 Caen Basse-Normandie France : [https://www.apmep.fr/IMG/pdf/Conf\\_RIOULT.pdf](https://www.apmep.fr/IMG/pdf/Conf_RIOULT.pdf)
- [17] Nicolas Pasquier. La boratoire d'Informatique (LIMOS)Université Clermont-Ferrand  
 II : Extraction de Bases pour les Règles d'Association à partir des Itemsets Fermés  
 Fréquents :<https://liris.cnrs.fr/inforsid/sites/default/files/a308c1YGpkbqaZQEU.pdf>
- [18] Marc Plantevit.1'Université Montpellier II. Juillet 2008 :Extraction De Motifs  
 Séquentiels Dans Des DonnéesMultidimensionnelles :<https://tel.archives-ouvertes.fr/tel-00319242/document>
- [19] Règles d'association :<https://perso.univ-rennes1.fr/valerie.monbet/doc/cours/IntroDM/Chapitre5.pdf>
- [20] Règles d'associations : [https://eric.univ-lyon2.fr/~ricco/cours/slides/regles\\_association.pdf](https://eric.univ-lyon2.fr/~ricco/cours/slides/regles_association.pdf)
- [21] Pascal Poncelet .LIRMM. Extraction de motifs :Règles d'association et  
 motifsséquentiels :  
[http://textmining.biz/Staff/Roche/ECD\\_M2/Cours/DataMiningRAMS.pdf](http://textmining.biz/Staff/Roche/ECD_M2/Cours/DataMiningRAMS.pdf)
- [22] Pierre-Antoine Papon. Université Blaise Pascal - Clermont-Ferrand II,  
 2016 :Extraction optimisée de règles d'association positives et négatives  
 intéressantes :<https://tel.archives-ouvertes.fr/tel-01412054/>
- [23] Manuel Bouillon : INSA de Rennes .26 janvier 2012 :Apprentissage incrémental et  
 décrémental : [ftp://ftp.irisa.fr/local/caps/DEPOTS/BIBLIO2012/Bouillon\\_Manuel.pdf](ftp://ftp.irisa.fr/local/caps/DEPOTS/BIBLIO2012/Bouillon_Manuel.pdf)
- [24] <http://depot-e.uqtr.ca/1201/1/030110265.pdf>

- [25] <http://biblos.hec.ca/biblio/memoires/m2006no62.pdf>
- [26] <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&cad=rja&uact=8&ved=0ahUKEwjswJf0gYvbAhUDnRQKHSUPBjIQFghEMAU&url=http%3A%2F%2Ftelecharger-cours.com%2Fdoc%2F1271.doc&usg=AOvVaw1eu8G92S1QKG6x57IOjDRB>
- [27] <https://eric.univ-lyon2.fr/~ricco/cours/slides/IntroDMDraft2002.pdf>
- [28] <http://www.cs.armstrong.edu/liang/intro6e/JavaCharacteristics.pdf>
- [28] définition de l'oracle 8i :
- [https://books.google.dz/books?id=1P6MbaXSNYwC&pg=PT219&lpg=PT219&dq=definition+de+oracle+8&source=bl&ots=6VNMAPs0uS&sig=ke4EGVOGQbhKHFix62AgzFFW1Xs&hl=fr&sa=X&ved=0ahUKEwjJ5oeV\\_sTbAhVQkRQKHxeMCdcQ6AEIjQEwCA#v=onepage&q=definition%20de%20oracle%20&f=false](https://books.google.dz/books?id=1P6MbaXSNYwC&pg=PT219&lpg=PT219&dq=definition+de+oracle+8&source=bl&ots=6VNMAPs0uS&sig=ke4EGVOGQbhKHFix62AgzFFW1Xs&hl=fr&sa=X&ved=0ahUKEwjJ5oeV_sTbAhVQkRQKHxeMCdcQ6AEIjQEwCA#v=onepage&q=definition%20de%20oracle%20&f=false)
- [29] définition de postgresql : [https://www.postgresql.org/message-id/attachment/6011/presentation\\_erwan\\_fs.html](https://www.postgresql.org/message-id/attachment/6011/presentation_erwan_fs.html)