TABLE DES MATIERES

Introduction				Page1
Chapitre 1 : Fouille de donnée (DATA MINING)				
1.1			Introduction	Page3
	1.1.1		Qu'est ce que la fouille de données	Page3
1.2			Les tâches du Data-Mining :	Page5
1.3			Les techniques de Data-Mining	Page5
1.4			Qu'est ce qu'une donnée	Page6
	1.4.1		Notations	Page6
	1.4.2		Les différentes natures d'attributs	Page7
	1.4.3		Les différentes natures de valeur d'attribut	Page7
	1.4.4		Le bruit	Page8
	1.4.5		Différentes tâches d'extraction d'information	Page8
	1.4.6		Les étapes d'un processus d'Extraction de Connaissances à partir des données (ECD)	Page10
		1.4.6.1	Nettoyage et intégration des données	Page10
		1.4.6.2	Prétraitement des données	Page12
		1.4.6.3	Fouille de données (Data-Mining)	Page12
		1.4.6.4	Evaluation et présentation	Page12
1.5			Conclusion	Page14
Chapitre 2 : Règles d'Association				
2.1			Introduction	Page15
2.2			Cadre informel	Page15
2.3			Cadre formel	Page16
2.4			La découverte des règles d'association	Page18
	2.4.1		Extraction des itemsets fréquents	Page18
2.5			Génération des règles d'association	Page23
2.6			Les améliorations de l'algorithme Apriori	Page26
2.7			Réduction de l'ensemble de règles d'association	Page27
	2.7.1		Approche orientée données	Page27
	2.7.2		Approche orientée utilisateur	Page31
2.8			Domaine d'applications	Page33
2.9			Conclusion	Page33
Chapitre 3 : Les algorithmes de génération des Règles d'Association				
3.1			Introduction	Page34
3.2			Génération des Itemset fréquents	Page35
	3.2.1		Algorithme FP-GROWTH	Page35
	3.2.2		Algorithme ECLAT	Page43
	3.2.3		Algorithme FIN	Page48
	3.2.4		Algorithme PrePost+	Page54
3.3			Génération des Règles d'association	Page59
	3.3.1		Génération des règles d'association par deux étapes	Page60
		3.3.1.1	L'algorithme Apriori pour génération des règles d'association	Page60
	3.3.2		Génération des règles d'association par	Page63

			approche d'expansion de règles	
		3.3.2.1	Génération des Règles d'Association de Top-K	Page63
		3.3.2.2	Génération des Règles d'Association de Top-K Non-Redondantes	Page68
3.4			Conclusion	Page72
Chapitre 4 : Modélisation et Implémentation				
4.1			Introduction	Page73
4.2			Spécification des besoins	Page73
	4.2.1		Spécification des besoins fonctionnels	Page73
	4.2.2		Spécification des besoins Non fonctionnels	Page73
4.3			Méthodologie et Approche Adoptée	Page74
	4.3.1		Présentation d'UML	Page74
	4.3.2		Les avantages d'UML	Page75
	4.3.3		Les Diagrammes de Cas D'utilisation	Page75
		4.3.3.1	Identification des acteurs	Page75
		4.3.3.2	Diagramme de cas d'utilisation globale	Page76
		4.3.3.2.a	Description du cas d'utilisation « Gérer les transactions »	Page76
		4.3.3.2.b	Description du cas d'utilisation « Générer des fichiers d'entrée »	Page77
		4.3.3.2.c	Description du cas d'utilisation « Choisir et paramétrer l'algorithme à exécuter »	Page77
		4.3.3.2.d	Description du cas d'utilisation « Analyser les résultats » :	Page78
		4.3.3.2.e	Description du cas d'utilisation « Comparer les résultats » :	Page78
		4.3.3.2.f	Description du cas d'utilisation « S'authentifier»	Page79
4.4			La Conception Générale	Page80
	4.4.1		Le cycle de développement en v	Page80
4.5			La Conception Détaillée	Page81
	4.5.1		Le diagramme de déploiement	Page81
	4.5.2		les diagrammes de séquence	Page81
		4.5.2.1	le diagramme de séquence « s'authentifier »	Page82
		4.5.2.2	le diagramme de séquence « Gérer les transactions »	Page82
		4.5.2.3	le diagramme de séquence « Générer des fichiers d'entrée »	Page83
		4.5.2.4	le diagramme de séquence «Choisir et paramétrer l'algorithme à exécuter »	Page83
		4.5.2.5	le diagramme de séquence «Analyser les résultats »	Page84
		4.5.2.6	le diagramme de séquence «Comparer les résultats»	Page84
	4.5.3		Les diagrammes d'activité	Page84
	4.5.4		Diagramme de Classe	Page86
4.6			Implémentation	Page86
	4.6.1		Choix technique	Page86

		4.6.1.1	Choix du langage de programmation	Page86
		4.6.1.2	Choix de l'architecture de l'application	Page87
	4.6.2		Environnement Logistique	Page88
		4.6.2.1	Environnement de développement	Page88
		4.6.2.2	système de gestion de base de données	Page90
		4.6.2.3	Les données et structure de données	Page90
4.7			Travail réalisée	Page95
	4.7.1		Description de l'application	Page95
	4.7.2		Description des tables de la base de données	Page101
4.8			Discussion des Résultats	Page102
	4.8.1		Algorithme TopKRules pour confiacne = 0,8	Page102
	4.8.2		Algorithme TopKRules pour confiacne = 0,5	Page102
	4.8.3		Algorithme TopKRules Non redunadante pour confiance = 0.5 et delta=200	Page103
	4.8.4		Algorithme TopKRules Non redunadante pour confiance = 0.5 et delta=3600	Page104
	4.8.5		Fichier de sortie	Page105
4.9			Comparaison entre les deux algorithmes (TopKRules et TNR)	Page106
4.10			Conclusion	Page107
Conclusion Générale				Page108