

# Innovation en Détection de Fraude en Assurance Automobile au Maroc : Approche Prédicative par Modèles de Machine Learning

Ghita Hajraoui <sup>1</sup>, Ayyoub Saoudi <sup>1</sup>, Kenza Hajraoui <sup>1</sup>, Jamal Zahi <sup>1</sup>

<sup>1</sup> University Hassan 1st, Faculty of Economics and Management, LM2CE, Settat, Morocco

**Abstract.** Cet article se consacre à l'amélioration de la détection et de la prévention de la fraude en assurance automobile au Maroc, en mettant en avant les méthodes innovantes de Machine Learning. À travers l'analyse de schémas récurrents, l'étude propose une approche proactive basée sur des modèles prédictifs avancés. L'accent est particulièrement mis sur l'utilisation efficace du Machine Learning pour détecter les profils de comportement frauduleux, offrant une solution ciblée à ce défi. Une étude de cas concrète vient illustrer l'application pratique de ces méthodes dans le contexte spécifique de l'assurance automobile au Maroc. L'objectif principal est de présenter des solutions avant-gardistes exploitant les avancées en Machine Learning pour optimiser la détection et la prévention de la fraude, contribuant ainsi à la fiabilité des systèmes d'assurance automobile au Maroc.

**Index Terms**— assurance, automobile, compagnies d'assurance, détection, fraude, Machine Learning, risque.

## 1 Introduction

La fraude à l'assurance automobile est un problème majeur pour les compagnies d'assurance, en particulier au Maroc où les réclamations frauduleuses en assurance automobile sont estimées à causer des pertes allant jusqu'à 21 % des pertes du marché en question<sup>1</sup>.

Au Maroc, le secteur de l'assurance est réglementé par le Dahir n° 1-02-238 du 25 rejev 1423 (3 octobre 2002), qui promulgue la loi n° 17-99 portant code des assurances. Ce texte définit les contrats d'assurance, les dispositions générales pour différents types d'assurances (incluant mutuelles et réassurance), ainsi que les règles de gestion, comptables, statistiques, de contrôle, et les sanctions en cas de non-respect des contrats. La loi couvre diverses catégories d'assurance, telles que les assurances de biens, vie, accidents corporels, maladies, etc. Elle établit des dispositions pour protéger tant les assureurs que les assurés, englobant des contrats variés. Dans notre étude, nous avons choisi de nous concentrer sur la détection de la fraude en assurance automobile, en raison de son importance relative par rapport aux autres branches d'assurance et de son caractère obligatoire.

Les compagnies d'assurance sont confrontées à une variété de schémas de comportements frauduleux, tels que la mise en scène d'un incident, la fausse représentation de la réalité ou encore une déformation quant à l'étendue des dommages causés<sup>2</sup>. Les coûts liés à la fraude ont un impact sur la rentabilité de l'assureur et peuvent affecter la pérennité de la compagnie. C'est dans ce sens que les compagnies d'assurance doivent avoir recours à l'analyse prédictive pour anticiper et contrer la fraude. Ces dernières sont, donc, dans le besoin de développer des capacités permettant d'identifier les fraudes potentielles avec un haut degré de précision, afin que les cas identifiés puissent être examinés en détail.

Actuellement, les compagnies d'assurance automobile font face à des défis en matière de contrôle et de détection des fraudes, avec des capacités encore peu développées [1]. En effet, de nombreuses entreprises identifient la fraude

<sup>1</sup> Compte rendu de la 29ème conférence téléphonique A2ii-AICA, Réactions de dans la lutte contre la fraude, Juillet 2018.

<sup>2</sup> [https://www.verisk.com/siteassets/media/campaigns/gated/underwriting/verisk-the-challenge-of-auto-insurance-premium-leakage.pdf?\\_\\_FormGuid=8c509869-699d-4698-9ac3-ada3d271c97&\\_\\_FormLanguage=en-US&\\_\\_FormSubmissionId=65299f67-8c20-408c-97bd-221fc1cee1bf](https://www.verisk.com/siteassets/media/campaigns/gated/underwriting/verisk-the-challenge-of-auto-insurance-premium-leakage.pdf?__FormGuid=8c509869-699d-4698-9ac3-ada3d271c97&__FormLanguage=en-US&__FormSubmissionId=65299f67-8c20-408c-97bd-221fc1cee1bf)

principalement grâce à l'intuition des agents ou à de simples suspicions [2]. Par conséquent, plusieurs dossiers classés comme non frauduleux pourraient contenir des fraudes potentielles, soulignant l'importance pour les compagnies de lancer des enquêtes basées sur des critères objectifs [3].

Face à ce constat, notre étude propose de développer un système de détection de fraude basé sur l'intelligence artificielle. Celui-ci permettra d'analyser les données relatives aux sinistres et d'identifier les cas suspects avec un haut degré de précision. Ce dernier s'appuiera des techniques d'apprentissage automatique pour extraire des caractéristiques prédictives des données et construire des modèles de classification.

La mise en place d'un tel système s'avère cruciale pour lutter contre la fraude à l'assurance automobile au Maroc et pour garantir la pérennité du secteur. Dans la suite de cet article, nous présenterons en détail la méthodologie de développement de nos modèles de détection de fraude, ainsi que les résultats obtenus et les perspectives d'avenir.

## **2 Analyse des tendances de la fraude en assurance automobile**

L'analyse des tendances de la fraude automobile au Maroc révèle des schémas prédominants qui posent des défis significatifs aux compagnies d'assurance. La fraude à l'assurance automobile est en constante augmentation dans le pays, entraînant une hausse notable des paiements d'indemnités par les assureurs. Les dossiers frauduleux représentent au moins 1/5 des indemnisations en assurance automobile, un taux qui augmente avec la hausse prévue du parc automobile. [4]. Parmi les schémas de fraude les plus courants figurent les faux constats, les sinistres inventés et les déclarations intentionnellement fausses lors de la survenance d'un sinistre.

Les faux constats représentent une tactique courante où des parties impliquées dans un accident collaboreraient pour exagérer les dommages ou inventeraient même un accident fictif. Ces manœuvres visent à maximiser les réclamations indemnitaires de manière injuste. Aussi, des sinistres inventés impliquent des incidents qui n'ont jamais eu lieu, mais qui sont déclarés dans le but de percevoir une indemnisation illégitime. Les déclarations intentionnellement fausses lors d'un sinistre ajoutent une couche supplémentaire de complexité, car les personnes impliquées peuvent délibérément fausser les faits pour obtenir des avantages financiers indus. [5]

L'analyse des tendances de la fraude à l'assurance automobile au Maroc met en évidence l'ampleur du problème et ses implications pour les compagnies d'assurance et les assurés. Face à l'évolution des techniques de fraude et à ses impacts négatifs, il devient impératif de mettre en place des solutions proactives pour détecter les profils suspects et endiguer ce phénomène.

## **3 Modélisation prédictive des profils et des dossiers suspects**

### **3.1 Déterminants de la fraude à l'assurance : Vers une détection objective et prédictive**

La détection de la fraude en assurance représente un défi majeur pour les compagnies, exigeant une approche basée sur des indicateurs fiables et objectifs. Dans cette optique, divers chercheurs se sont penchés sur l'étude des déterminants de la fraude, privilégiant des critères concrets et mesurables. L'objectif principal est de mettre en place des processus de prévention efficaces, englobant toute la chaîne, depuis la souscription des contrats jusqu'aux déclarations frauduleuses de sinistres [6]. Cette approche, en contraste avec les critères comportementaux souvent subjectifs, vise à identifier les indicateurs les plus pertinents pour la détection de la fraude post-déclaration.

Au cœur de cette démarche se trouvent les travaux pionniers de Belhadji & Dionne (1997), qui ont jeté les bases dans le contexte de l'assurance automobile au Canada. Leur objectif était de déterminer les indicateurs de fraude impactant la probabilité de fraude d'un sinistre, utilisant un ensemble de données représentatif des compagnies d'assurance canadiennes. À travers une approche de notation, ils ont élaboré une stratégie de « *red flag* » pour identifier les déclarations suspectes. Cette approche a été déterminante pour mettre en évidence les indicateurs les plus significatifs, faisant ainsi des auteurs, les pionniers en matière de prédiction de fraude.

**Tableau 1.** Indicateurs de détection de fraude en assurance automobile les plus significatifs selon Belhadji & Dionne.

1	Une collision mineure a entraîné des coûts de réparation excessifs.
2	Existence de tout dommage non relié à la perte ou incohérent avec les faits de l'accidents.
3	Le véhicule est rapporté volé et retrouvé peu de temps après avec de lourds dommages.
4	L'assuré éprouve des difficultés financières personnelles ou reliées à ses affaires.
5	L'assuré est prêt à accepter un montant de règlement relativement petit plutôt que d'avancer tous les documents reliés à la perte.
6	L'assuré est extraordinairement familier avec le jargon des assurances ou des réparations de véhicules.
7	L'assuré (Ou le réclamant) est trop enthousiaste ou trop franc pour prendre le blâme de l'accident
8	L'accident (ou la perte) a eu lieu peu de temps après l'enregistrement et l'achat d'assurance pour le véhicule, ou dans le mois qui précède la fin de la police (ou de la couverture)
9	Nombreux reçus de taxi, ou bien factures de location de véhicule provenant d'un magasin de débosselage
10	Factures ou preuves semblent fausses ou fabriquées
11	La documentation de l'estimation et de la réparation n'est pas disponible
12	Témoignages contradictoires concernant les circonstances de la perte
13	Accident impliquant un seul véhicule
14	L'achat du véhicule s'est fait au comptant
15	Réclamant est très agressif (menace de faire appel à un avocat, au gouvernement, etc.)
16	Lors de l'enquête, l'assuré est nerveux et semble confus

Source : Résultats de l'étude de Belhadji & Dionne

### 3.2 Adaptation et Évolution : Contextualisation au Maroc :

Reconnaissant les spécificités du marché marocain, une adaptation des indicateurs identifiés par Belhadji & Dionne a été entreprise. Une enquête approfondie auprès d'assureurs marocains a permis de soumettre ces indicateurs à un processus d'ajustement, prenant en compte les nuances locales [7]. Bien que des critiques aient été émises quant à la taille de l'échantillon dans l'étude initiale, la validation de ces indicateurs par la majorité des assureurs interrogés prouve leur pertinence dans le contexte marocain. Cette phase d'adaptation s'avère importante pour la construction d'un modèle de modélisation prédictive des profils et des dossiers suspects en matière de fraude en assurance automobile au Maroc. Celle-ci prend en compte les spécificités locales, assurant ainsi la pertinence des indicateurs dans le paysage marocain de l'assurance automobile.

Les indicateurs jugés plus significatifs dans un contexte marocain à partir de l'enquête et retenus sont : -Une collision mineure a entraîné des coûts de réparation excessifs (CMCE)-Factures ou preuves semblent fausses ou fabriquées (FF)-L'accident (ou la perte) a eu lieu peu de temps après l'enregistrement et l'achat d'assurance pour le véhicule, ou dans le mois qui précède la fin de la police (ou de la couverture) (AAE)-Accident impliquant un seul véhicule (USV).

### 3.3 Innovations méthodologiques dans l'identification des indicateurs de Fraude en assurance automobile :

En quête d'indicateurs de fraude plus pertinents, Benedek & Laszlo (2019) ont opté pour une approche plus moderne en recourant à des algorithmes d'arbre de décision, de Naive Bayes et de réseaux de neurones. Leur démarche méthodologique a consisté en une segmentation des indicateurs en quatre catégories distinctes, à savoir les caractéristiques du véhicule, les caractéristiques de l'accident, les caractéristiques du conducteur, et les caractéristiques du réclamant. Chacune de ces catégories a regroupé divers indicateurs susceptibles d'apporter une information significative en termes de suspicion de fraude.[8]

Cette approche méthodologique novatrice de Benedek & Laszlo met en exergue la constante évolution des techniques de modélisation dans la détection de la fraude en assurance. En élargissant la palette des outils disponibles avec l'utilisation d'algorithmes avancés, elle offre une meilleure capacité d'anticipation et de prévention des cas de fraude. Ces résultats influent directement sur notre démarche d'adaptation des indicateurs de Belhadji & Dionne au contexte marocain, offrant ainsi une perspective contemporaine et enrichissante pour la construction d'un modèle de modélisation prédictive des profils et des dossiers suspects en matière de fraude en assurance au Maroc.

**Tableau 2.** Variables utilisées pour la conception des modèles.

Variables	Description	Modèle	
Fraude	Variable à expliquer : déterminer l'existence d'un cas de fraude ou de non fraude		
CMCE	Une collision mineure a entraîné des coûts de réparation excessifs	Belhaj et Dionne (1994)	Modèle global
FF	Factures ou preuves semblent fausses ou fabriquées		
AAE	L'accident (ou la perte) a eu lieu peu de temps après l'enregistrement et l'achat d'assurance pour le véhicule, ou dans le mois qui précède la fin de la police (ou de la couverture)		
USV	Accident impliquant un seul véhicule		
Incident_type	S'il s'agit d'une collision, d'un vol ou si le véhicule était garé lors de l'accident		
Collision_type	Le type de collision (frontale/ latérale/ arrière ou pas de collision)	Benedek et Laszlo (2019)	
Incident_ville	La ville où l'incident a eu lieu		
Incident_mois	Le mois où l'incident a eu lieu		
Assure_sexe	Le sexe de l'assuré		
Assure_age	L'âge de l'assuré		
Auto_marque	La marque de véhicule		
Auto_catégorie	La catégorie de la voiture : luxe, citadine ou SUV		
Age_véhicule	L'âge de véhicule		
Assurance_type	Le type d'assurance pour lequel l'assureur a souscrit		

Source : Synthèse des variables réalisée par les auteurs

## 4 Élaboration de stratégies de prévention

La gestion du risque opérationnel s'impose comme une préoccupation croissante, entraînant une perte significative de 1,3 milliard de dollars [9]. Ce phénomène a pris une ampleur considérable en raison de l'accroissement des pertes opérationnelles, de la dépendance accrue aux technologies avancées et des impératifs réglementaires. L'accord de Bâle II en 1999, avec ses révisions ultérieures, a marqué un tournant en imposant des mesures prudentielles, incitant les institutions financières à considérer le risque opérationnel de manière distincte et à adopter des pratiques de gestion plus efficaces.

La détection des fraudes dans le domaine de l'assurance revêt une importance capitale. Ceci consiste à repérer les déclarations de sinistres potentiellement frauduleuses et à les signaler afin qu'elles fassent l'objet d'enquêtes approfondies [3].

Les études consacrées à la détection de la fraude en assurance automobile se concentrent sur deux axes principaux. D'une part, la détection par suspicion, impliquant des audits et des vérifications visant à limiter les actions frauduleuses [2, 10]. D'autre part, des approches empiriques utilisant des techniques de datamining et de détection d'anomalies permettent de reconnaître les comportements susceptibles de donner lieu à des déclarations de sinistre frauduleuses.

Historiquement, la détection de la fraude reposait sur des techniques d'échantillonnage et d'audit, avec des listes d'indicateurs pour repérer les dossiers frauduleux. Cependant, des limites ont été détectées, conduisant ainsi à

l'évolution vers des méthodes et les algorithmes de Machine Learning. Les systèmes modernes de détection de fraude, de plus en plus complexes, offrent des avantages tels que le calcul de la probabilité de fraude et le profilage des assurés [11]. Des approches récentes privilégient la détection d'anomalies, basée sur le profilage comportemental. L'apprentissage supervisé demeure la méthode prédominante, générant des profils étiquetés "fraude" ou "non fraude". Cette approche utilise divers algorithmes de classification et de régression, exploitant l'intelligence artificielle et le Machine Learning pour automatiser la détection.

À son tour, l'intégration de systèmes d'information dédiés à la détection de fraude s'avère significativement avantageuse, permettant une identification précoce des activités frauduleuses. Ces systèmes fournissent également des informations pour la création de tableaux de bord, facilitant la prise de décision et la gestion globale des risques [12, 13]. L'automatisation de la détection de fraude est bien plus qu'une simple contribution à la réduction des effets financiers néfastes de la fraude, c'est une solution garantissant la pérennité et la sécurité financière des compagnies d'assurances [14].

## 5 Étude de cas

### 5.1 Méthodologie

Notre démarche de recherche s'inscrit dans une approche hypothético-déductive, visant à modéliser les comportements et les profils liés à la fraude dans le secteur de l'assurance. L'objectif ultime est de mettre en évidence le potentiel de récupération financière résultant de l'adoption d'un tel système.

La première étape consiste à collecter une base de données auprès d'une compagnie d'assurance désirant garder l'anonymat pour l'exercice de 2019, couvrant une période de 6 mois. Le jeu de données comporte 5000 observations et incluent les caractéristiques relatives à l'assuré, au véhicule, à l'incident et à la réclamation. Il est important de noter que la compagnie d'assurance a effectué un rééchantillonnage tentant ainsi de pallier au problème de déséquilibre fréquemment observé lors de la conduite d'études traitant de la fraude.

Ensuite, nous avons chargé les données dans un environnement Python en utilisant la bibliothèque Pandas, et nous avons procédé à un encodage des variables catégorielles. La base de données a été, par la suite, divisée en ensemble d'apprentissage et un autre de test afin d'évaluer la performance du modèle.

Concernant la modélisation, nous avons opté pour l'algorithme XGBoost. Celui-ci a fait ses preuves en matière de détection de fraude. En effet, selon Al Ali *et al.* (2023) [15] XGBoost est largement privilégié pour la détection de fraude en assurance en raison de plusieurs avantages. En termes de performance, XGBoost a prouvé sa supériorité par rapport à d'autres algorithmes tels que la Régression Logistique, les Arbres de Décision, SVM, AdaBoost ou encore les Forêts Aléatoires dans la détection de fraude financière. Ces caractéristiques font de XGBoost un choix judicieux pour la détection de fraude en assurance, capable de relever les défis spécifiques des données financières, notamment le déséquilibre des données et les valeurs manquantes, tout en offrant une précision accrue.

### 5.6 Résultats

Une méthodologie rigoureuse a été pour la collecte, le prétraitement des données, et la modélisation utilisant l'algorithme XGBoost. Ces étapes précédentes constituent les bases de la conception d'un modèle de détection de fraude via les méthodes novatrices de Machine Learning.

#### 5.6.1 Détection des profils suspects

La détection de fraude repose sur la capacité à créer des classes des assurés pour assurer une efficacité maximale. Cela permet de faire ressortir les profils présentant un risque élevé de fraude à l'avenir. Cette approche accorde à la

compagnie d'assurance un avantage stratégique, ce qui lui procure une compréhension approfondie de la clientèle. Les résultats obtenus par le modèle sont les suivants :

```

Accuracy: 0.95
Confusion Matrix:
[[565  7]
 [ 43 385]]
Classification Report:

```

	precision	recall	f1-score	support
0	0.93	0.99	0.96	572
1	0.98	0.90	0.94	428
accuracy			0.95	1000
macro avg	0.96	0.94	0.95	1000
weighted avg	0.95	0.95	0.95	1000

Source : Conçu par les auteurs

**Fig. 1.** Résultats du modèle de classification des profils.

Les performances de notre modèle s'avèrent encourageantes. L'accuracy du modèle, évaluée à 95%, représente la proportion d'instances correctement classées par rapport au total, démontrant ainsi une capacité notable à effectuer des prédictions exactes.

La matrice de confusion présente une analyse approfondie des résultats. Les vrais positifs (385) correspondent aux cas de fraude correctement identifiés, tandis que les vrais négatifs (565) représentent les transactions non frauduleuses correctement classées. Les faux positifs (7) et les faux négatifs (43) signalent respectivement les erreurs de classification des transactions non frauduleuses comme frauduleuses et vice versa.

Quant aux métriques, ces mesures fournissent un équilibre entre la capacité du modèle à identifier correctement les instances positives et à minimiser les erreurs de classification. Une précision élevée suggère une probabilité réduite de fausses alertes, tandis que le rappel met en évidence la capacité du modèle à détecter la majorité des cas positifs.

### 5.6.2 Détection des dossiers suspects

Après avoir examiné les profils susceptibles de commettre une fraude, orientons désormais notre attention vers la classification des dossiers. Si la première phase de notre analyse nous a permis de cerner les profils individuels présentant des risques de fraude, la seconde phase se concentre sur l'application de notre modèle à l'évaluation des dossiers spécifiques. Cette étape permettra d'explorer la capacité de notre modèle à identifier les schémas frauduleux au niveau des réclamations et à contribuer ainsi à une gestion plus fine des risques opérationnels pour la compagnie d'assurance.

```

Accuracy: 0.936
Confusion Matrix:
[[564  8]
 [ 56 372]]
Classification Report:

```

	precision	recall	f1-score	support
0	0.91	0.99	0.95	572
1	0.98	0.87	0.92	428
accuracy			0.94	1000
macro avg	0.94	0.93	0.93	1000
weighted avg	0.94	0.94	0.94	1000

Source : Conçu par les auteurs

**Fig. 2.** Résultats du modèle de classification des dossiers

Les résultats de l'évaluation du modèle de classification démontrent une performance robuste dans la détection de fraudes. Avec une précision de 93,6%, le modèle a réussi à classer correctement une grande proportion d'instances, prouvant ainsi sa capacité fiable à distinguer les cas de fraudes des réclamations régulières.

La vue détaillée des prédictions du modèle, à travers la matrice de confusion, indique les vrais positifs (fraudes correctement détectées), les vrais négatifs (dossiers non frauduleux correctement classés), les faux positifs (dossiers non frauduleux incorrectement considérés comme frauduleux), et les faux négatifs (fraudes manquées).

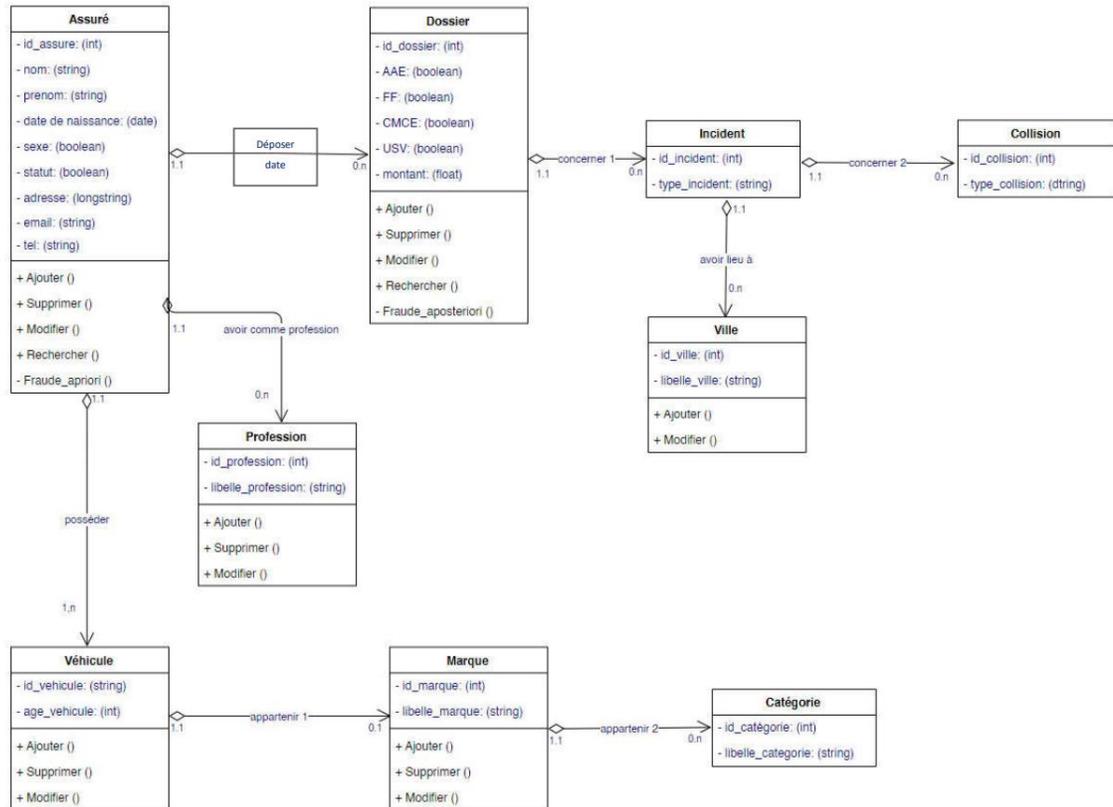
Le rapport de classification approfondit cette analyse en fournissant des métriques telles que la précision, le rappel et le F1-score pour chaque classe. La précision élevée pour les dossiers non susceptibles de comporter une fraude (91%) et les fraudes (98%) démontre la fiabilité du modèle dans ses prédictions. Cependant, le rappel de 87% pour les fraudes pourrait indiquer qu'il existe une marge d'amélioration pour capturer toutes les fraudes potentielles.

### 5.6.3 Système de détection automatique de fraude

Les résultats du modèle de classification ne se limitent pas à la simple détection de fraude. Ils sont, en effet, des éléments essentiels du processus décisionnel de la compagnie d'assurance. Les informations générées par ces modèles fournissent des indications d'une haute importance pour orienter les actions futures, que ce soit en termes de gestion des sinistres, de prévention proactive ou, voire même, d'optimisation des primes.

La transition vers le système d'information de détection automatique marque une évolution remarquable dans l'approche de la lutte contre la fraude. Alors que la détection manuelle se basait souvent sur des méthodes traditionnelles, l'intégration de systèmes d'information automatisés ouvre la voie à une veille plus sophistiquée et proactive. Ces systèmes exploitent les avantages de l'intelligence artificielle et du Machine Learning pour analyser de vastes ensembles de données, offrant ainsi une capacité à repérer les schémas de fraude subtils et émergents.

Pour matérialiser efficacement ce système, le diagramme des classes représente visuellement les relations entre les différentes entités. Il est conçu dans le but spécifique d'automatiser la détection de fraudes et de guider les prises de décision subséquentes. Le diagramme de la figure ci-dessous détaille la structure et les interactions clés entre les composantes fondamentales du système d'information.



Source : Conçu par les auteurs

**Fig. 3.** Diagramme des Classes

Notre système relève plusieurs avantages, tels que : le calcul de la probabilité de fraude pour chaque cas ainsi que le profilage des assurés et l'identification des comportements frauduleux. Ce qui permettra, *in fine*, à la compagnie d'assurance de réduire les pertes financières et d'améliorer la gestion du risque de fraude.

Les actions à entreprendre en cas de détection de fraude dépendent du risque encouru. Selon le système de détection de fraude, l'agent assureur peut vérifier le risque de fraude et décider des actions à entreprendre. Dans certains cas, l'expert assureur peut être appelé pour mener des investigations plus approfondies si besoin. Les actions peuvent inclure l'ouverture d'une enquête ou l'entame du processus usuel d'indemnisation.

Le système contribue, donc, à la prévention et à la détection de la fraude en fournissant des informations précises et des outils pour identifier les comportements frauduleux. Ils permettent aux entreprises de réduire les opportunités de fraude et de renforcer la dissuasion envers les fraudeurs potentiels.

L'adoption d'un système de détection automatique de fraude en assurance automobile permettrait à la compagnie qui le détient d'économiser jusqu'à 20% du montant des remboursements [4]. Ce dernier se trouve en mesure de remédier au problème de détection de fraude en assurance et d'optimiser la gestion du risque associé.

## Conclusion

Le paysage assurantiel contemporain se caractérise par une complexité croissante qui engendre des défis certains, dont la fraude à l'assurance s'avère être une préoccupation majeure. Face à l'augmentation des cas de fraude et aux limites des méthodes traditionnelles de détection basées sur l'intuition, notre étude présente le développement d'un système de détection automatique de fraude en assurance automobile. Ce système s'appuie sur l'intelligence artificielle et les techniques d'apprentissage automatique pour identifier avec précision les cas suspects et garantir la pérennité du secteur.

Le choix de l'algorithme XGBoost s'est fondé sur ses performances avérées dans le domaine de la détection de fraude en assurance. En effet, cet algorithme offre une précision et une efficacité accrues par rapport aux méthodes traditionnelles. De plus, les résultats du modèle ne se limitent pas à la simple identification des cas frauduleux, mais ils fournissent également des informations cruciales pour orienter la gestion des sinistres et la prévention proactive de la fraude.

La mise en place d'un système de détection automatique de fraude représente une avancée significative dans la lutte contre ce phénomène. En exploitant l'intelligence artificielle et le Machine Learning, ce système permet une détection sophistiquée des schémas de fraude et réduit les opportunités pour les fraudeurs potentiels. L'adoption d'un tel système peut, également, générer des économies substantielles pour les compagnies d'assurance en optimisant la gestion du risque associé à la fraude.

Notre travail propose, alors, une solution innovante et efficace pour contrer la fraude à l'assurance automobile. Le système de détection automatique développé s'inscrit comme un outil précieux pour les compagnies d'assurance, leur permettant de garantir la pérennité de leur activité et de protéger leurs intérêts financiers.

## Références

1. Abdallah, M. A. Maarof, A. Zainal, Fraud detection system: A survey, *J. Netw. Comput. Appl.* **68**, 90-113 (2016).
2. N. Belhadji, G. Dionne, Développement d'un système expert de détection automatique de la fraude à l'assurance automobile, *Cahier de recherche* **97**, 04 (1997).
3. S. Behdad et al., Nature-inspired techniques in the context of fraud detection, *IEEE Trans. Syst. Man Cybern. Part C* **42**(6), 1273-1290 (2012).
4. G. Hajraoui, F. El Kassimi, J. Zahi, La gestion du risque de fraude en assurances automobile – cas d'une compagnie d'assurances marocaine, *J. Integr. Stud. Econ. Law Tech. Sci. Commun.* **1**(2) (2023).
5. A2ii, CC 29 | Réactions des contrôleurs dans la lutte contre la fraude, *Compte rendu de la 29e Consultation téléphonique A2ii – AICA Réactions des contrôleurs dans la lutte contre la fraude* (2018).
6. L. Bouzgarne, M. El Alami, M. El Omari, Détection de la fraude à l'assurance automobile : Application d'un modèle de régression logistique, *Rev. Int. Rech. Finance* **17**(2), 1-17 (2019).
7. G. Hajraoui, J. Zahi, Indicateurs pertinents de détection automatique de fraude : Cas des compagnies d'assurance automobile, *Altern. Manag. Econ.* **4**(2), 420-436 (2022).
8. Benedek, Z. Laszlo, A Machine Learning Approach to Insurance Fraud Detection, *Risks* **7**(2), 42 (2019).
9. M. R. Cruz, *Operational risk management: A practical approach* (John Wiley & Sons, 2002).
10. K. J. Crocker, S. Tennyson, *Insurance fraud: A practitioner's guide to detection and prevention* (John Wiley & Sons, 2002).

11. J. Li et al., An integrated data mining approach to insurance fraud detection, *Expert Syst. Appl.* **35**(1), 266-279 (2008).
  12. Al-Jumeily et al., Toward an optimal use of artificial intelligence techniques within a clinical decision support system, in 2015 Science and Information Conference (SAI) (IEEE, 2015), pp. 548-554.
  13. R. Dzomira, Fraud prevention and detection, *Res. J. Finance Account.* **6**(14), 37-43 (2015).
  14. J. Schiller, The impact of insurance fraud detection systems, *J. Risk Insur.* **73**(3), 421-438 (2006).
- A. A. Ali, A. M. Khedr, M. El-Bannany, S. Kanakkayil, A powerful predicting model for financial statement fraud based on optimized XGBoost ensemble Learning technique, *Appl. Sci.* **13**(4), 2272 (2023).