EXPLOITATION DE RAPPORTS D'INCIDENT POUR L'ANALYSE DU RISQUE CYBER

Justin Kher, Olivier Lopez, Hugo Rapior

AGENDA

1- Techniques d'embedding

2- Réseaux de neurones

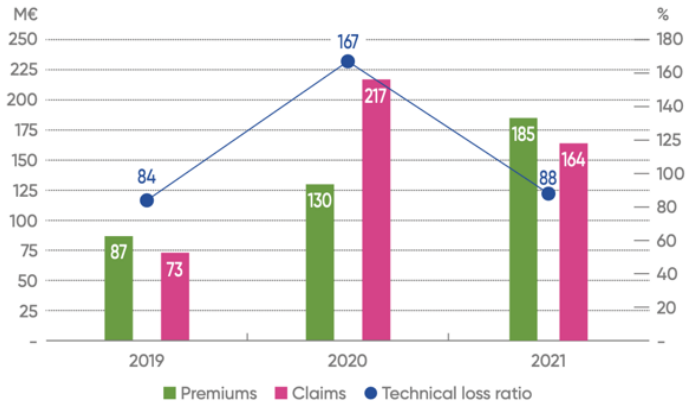
3- Enrichissement du modèle

4- Applications et perspectives

INTRODUCTION

- Rapport LUCY de l'AMRAE (données de courtiers) :
  - Loss ratios : 84% en 2019, 167% en 2020, 88% en 2021
  - Premium : +44,4% collectés, pour un effectif en croissance de 27,5%

- Rapport de la Direction Générale du Trésor sur le risque cyber, septembre 2022 :
  - soulève la question des données
  - pointe la nécessité de méthodes innovantes (ex: bayésien)



Source: 2022 AMRAE LUCY Study.

INTRODUCTION

- **Question :** peut-on utiliser les multiples ressources textuelles disponibles (par exemple rapports d'incidents) pour mieux apprendre sur le risque ?

- **Objectifs :**
  - « data augmentation »
  - transformer de l'expertise « littéraire » en expertise quantitative nourrissant des modèles bayésiens
  - mieux anticiper la gestion d'incident

AGENDA

1- Techniques d'embedding

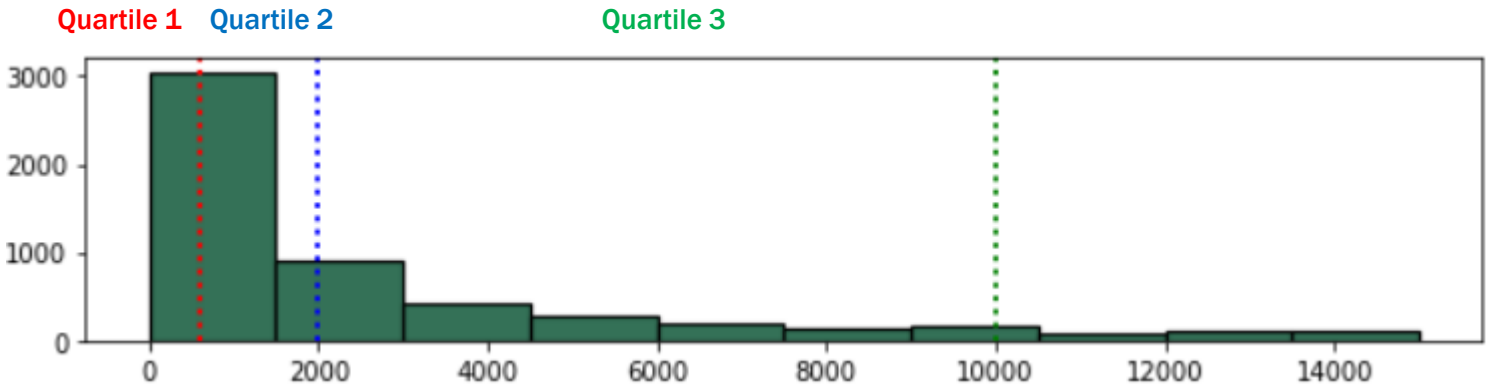
2- Réseaux de neurones

3- Enrichissement du modèle

4- Applications et perspectives

BASE PRC: PRIVACY RIGHTS CLEARINGHOUSE (US)

- Présence d'un marqueur de sévérité (number of records) et d'une description du sinistre



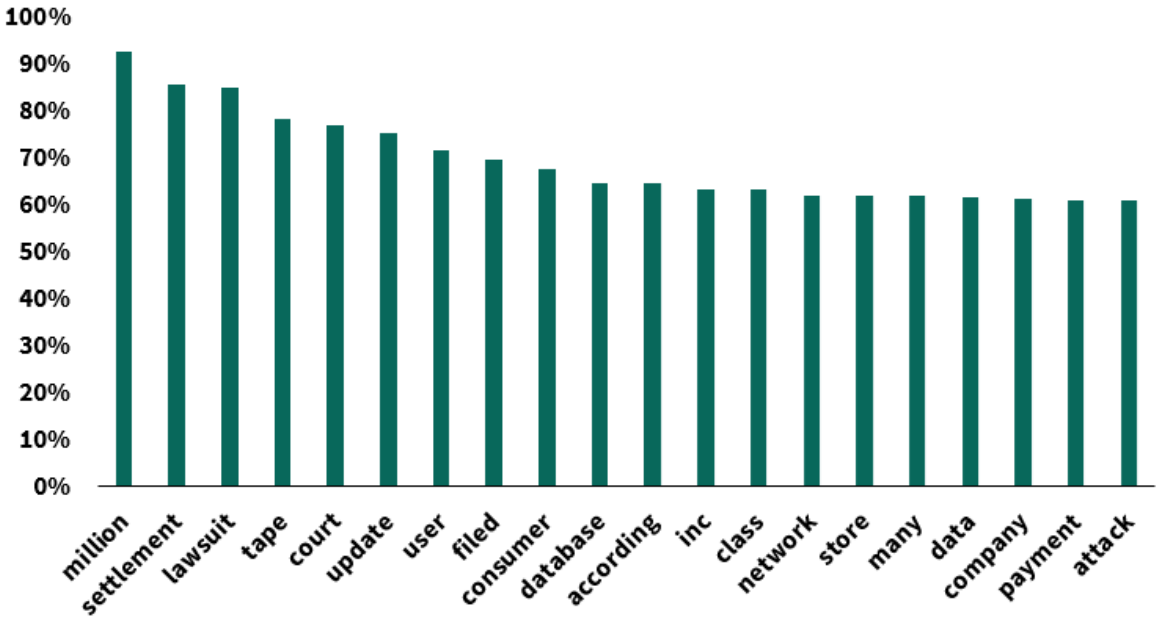
- Segmentation de la base et identification des sinistres les plus sévères selon le number of record.

NB OF RECORDS < 4767 ?	NB OF RECORDS > 4767 ?
Sinistre attritionnel	Sinistre grave

BASE PRC: PRIVACY RIGHTS CLEARINGHOUSE (US)

- Les mots utilisés dans la description du sinistre sont un indicateur de la sévérité

Sinistres graves



Proportion de sinistre grave lorsque le mot apparait dans la description

Sinistres attritionnels

Les mots suivants sont associés aux descriptions de sinistres attritionnels :

- Paper
- Document
- Dishonest
- Accidentally
- School

PRE-TRAITEMENT DES DONNES TEXTUELLES

Traitement sur la description du sinistre

Description of the incident	Description clean
Union Hospital suffered an inadvertent disclosure on approximately 1/18/16 that resulted in 1 record being exposed, which included social security numbers.	union hospital suffered inadvertent disclosure approximately resulted record exposed included social security number

Information parasitaire

- Stopwords
- Dates
- Ponctuation
- Nombres

PRE-TRAITEMENT DES DONNES TEXTUELLES

Traitement sur la description du sinistre

Description of the incident	Description clean
Union Hospital suffered an inadvertent disclosure on approximately 1/18/16 that resulted in 1 record being exposed, which included social security numbers.	union hospital suffered inadvertent disclosure approximately resulted record exposed included social security number

Information parasitaire

- Stopwords
- Dates
- Ponctuation
- Nombres

Dictionnaire sur le corpus

- Chaque mot présent dans le corpus est un « **token** »
- Une analyse du texte permet d'identifier de nouveaux tokens, composés de 2 ou 3 mots à ajouter au dictionnaire

PRE-TRAITEMENT DES DONNES TEXTUELLES

Traitement sur la description du sinistre

Description of the incident	Description clean
Union Hospital suffered an inadvertent disclosure on approximately 1/18/16 that resulted in 1 record being exposed, which included social security numbers.	union hospital suffered inadvertent disclosure approximately resulted record exposed included social security number

Information parasitaire

- Stopwords
- Dates
- Ponctuation
- Nombres

Dictionnaire sur le corpus

- Chaque mot présent dans le corpus est un « **token** »
- Une analyse du texte permet d'identifier de nouveaux tokens, composés de 2 ou 3 mots à ajouter au dictionnaire

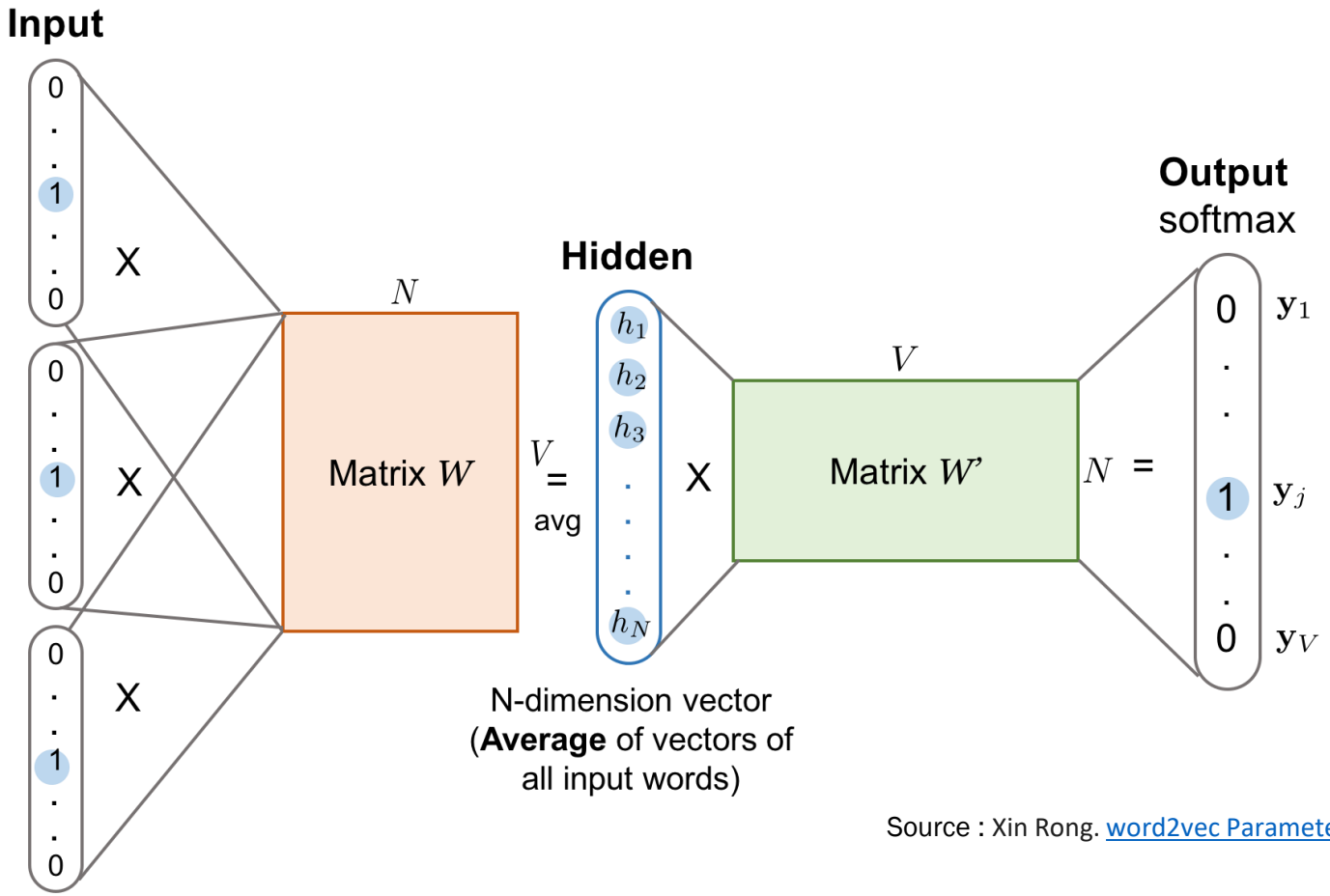
Exemple de suite de mots

« social security number »

« personal information »

« email adress »

WORD EMBEDDING – WORD2VEC



Source : Xin Rong. word2vec Parameter Learning Explained

WORD EMBEDDING – WORD2VEC

Un mot est encodé dans un **espace** de **dimension N**

Les mots avec une **signification** ou une **influence** similaire sont **proches**

	Man (5182)	Woman (9742)	King (4815)	Queen (7464)	Apple (421)	Orange (6151)
Gender	-1	1	-0,95	0,97	0,00	0,01
Royal	0,01	0,02	0,93	0,95	-0,01	0,00
Age	0,03	0,02	0,7	0,69	0,03	-0,02
Food	0,04	0,01	0,02	0,01	0,95	0,97
…	…	…	…	…	…	…

WORD EMBEDDING – WORD2VEC

Un mot est encodé dans un **espace** de **dimension N**

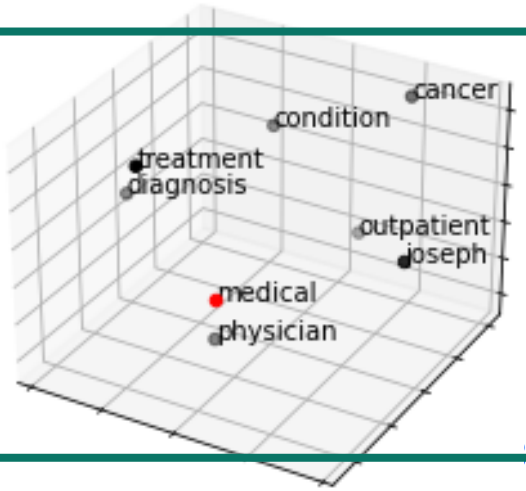
Les mots avec une **signification** ou une **influence** similaire sont **proches**

	Man (5182)	Woman (9742)	King (4815)	Queen (7464)	Apple (421)	Orange (6151)
Gender	-1	1	-0,95	0,97	0,00	0,01
Royal	0,01	0,02	0,93	0,95	-0,01	0,00
Age	0,03	0,02	0,7	0,69	0,03	-0,02
Food	0,04	0,01	0,02	0,01	0,95	0,97
…	…	…	…	…	…	…

Mots proches : insurance
Life
Insurer
Coverage
Enrollee
Plan
Guarantor
Aflac

Représentation 3D

Les mots proches de « **medical** » dans notre corpus de description peuvent être représentés en **3D**



AGENDA

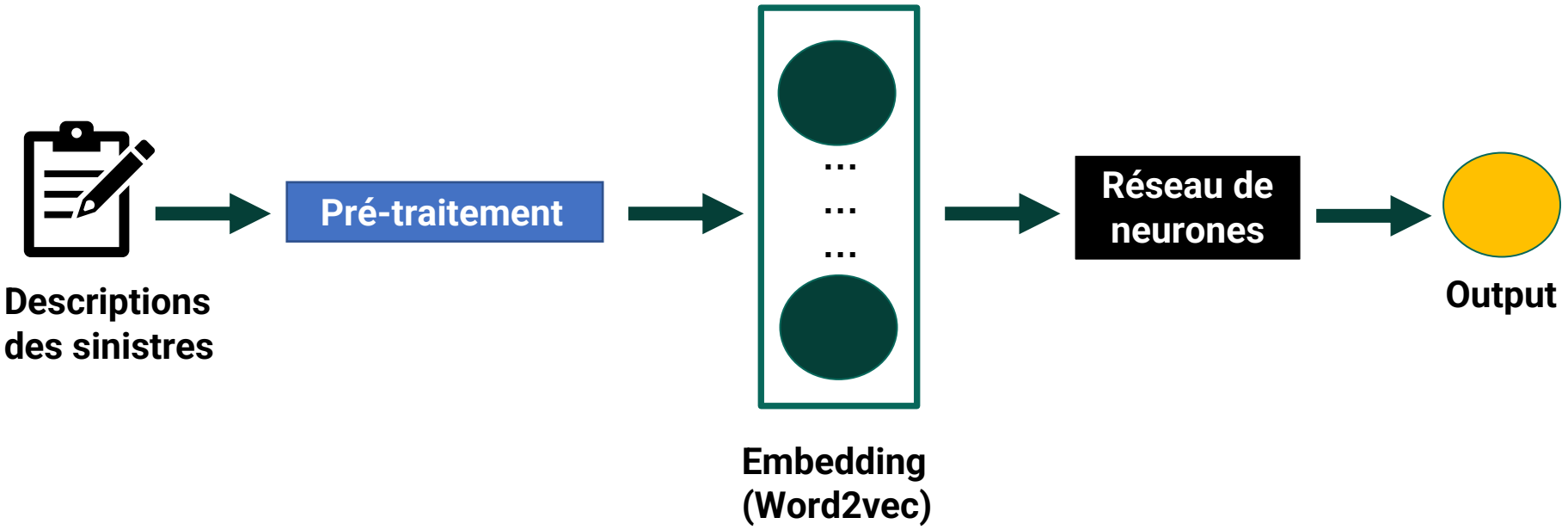
1- Techniques d'embedding

2- Réseaux de neurones

3- Enrichissement du modèle

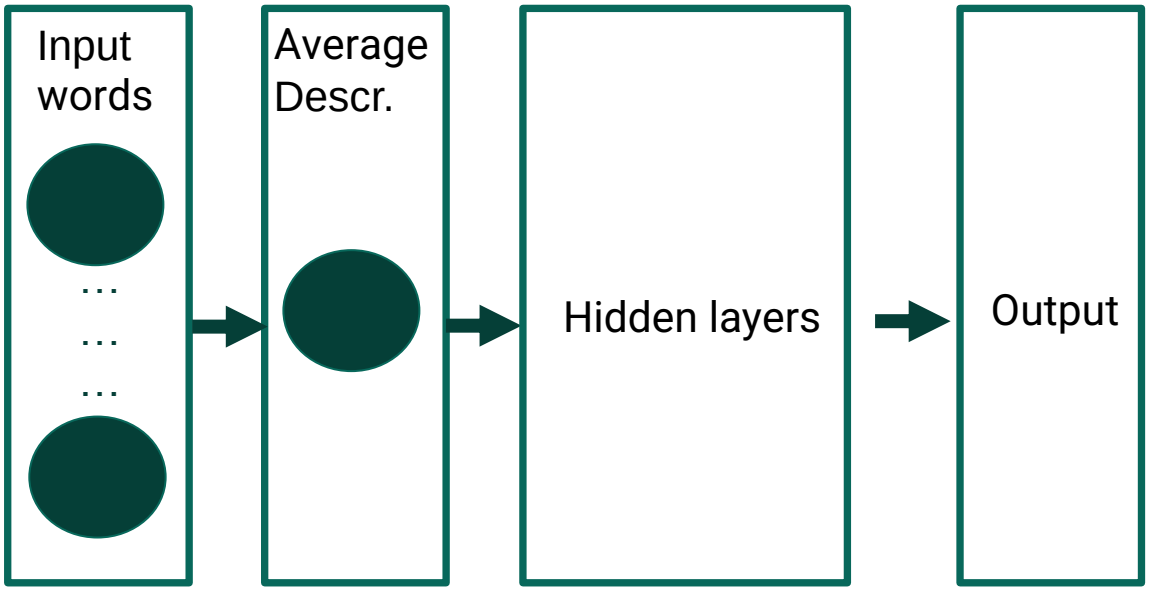
4- Applications et perspectives

Méthode


Descriptions des sinistres
Pré-traitement
Embedding (Word2vec)
Réseau de neurones
Output

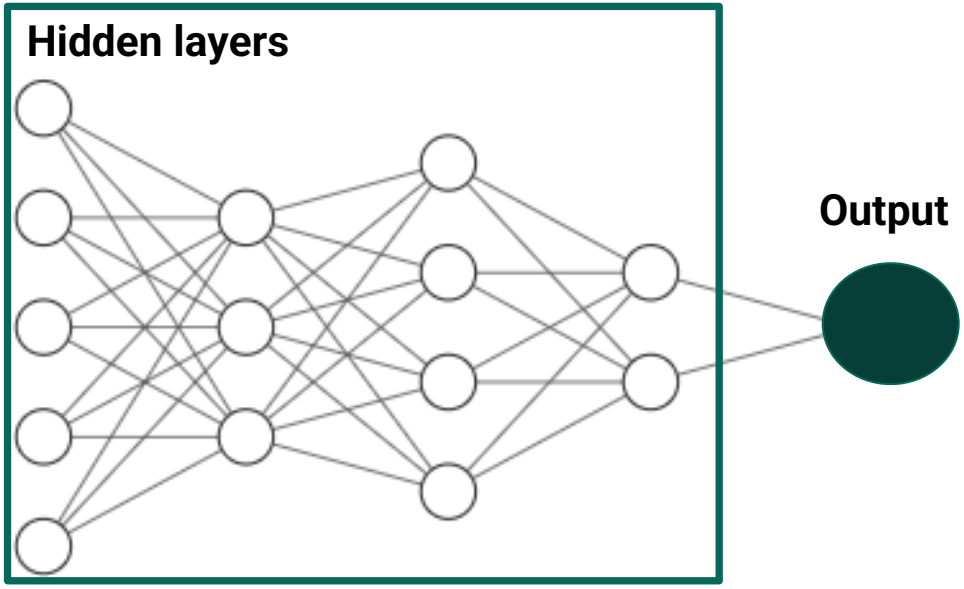
MODÈLE CHOISI : PERCEPTRON MULTICOUCHE

- Zoom sur la partie **Réseau de neurones** -> Perceptron multicouche



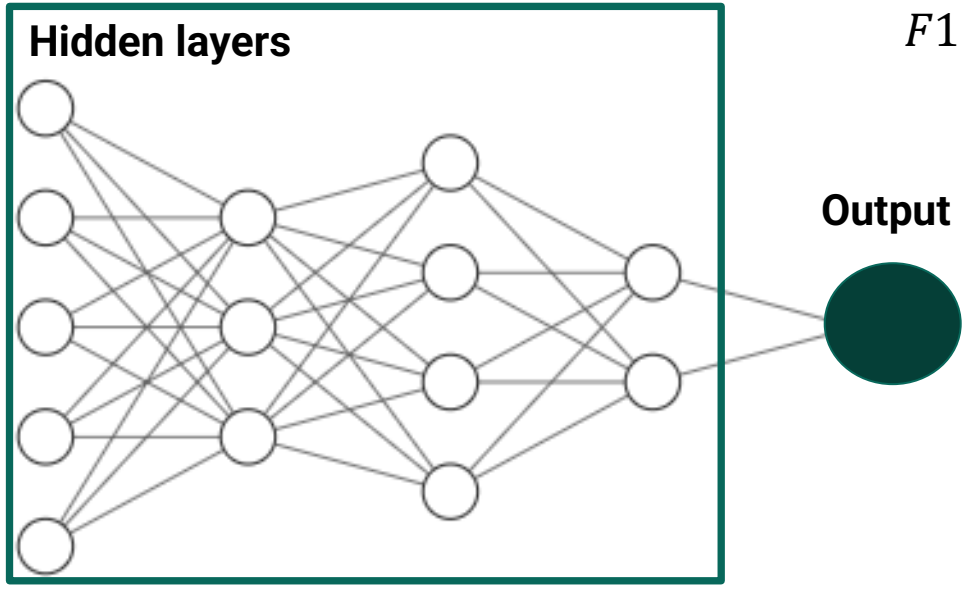
MODÈLE CHOISI : PERCEPTRON MULTICOUCHE

- Gridsearch sur les couches cachées du réseau de neurone (F1 score)



MODÈLE CHOISI : PERCEPTRON MULTICOUCHE

- Gridsearch sur les couches cachées du réseau de neurone (F1 score)



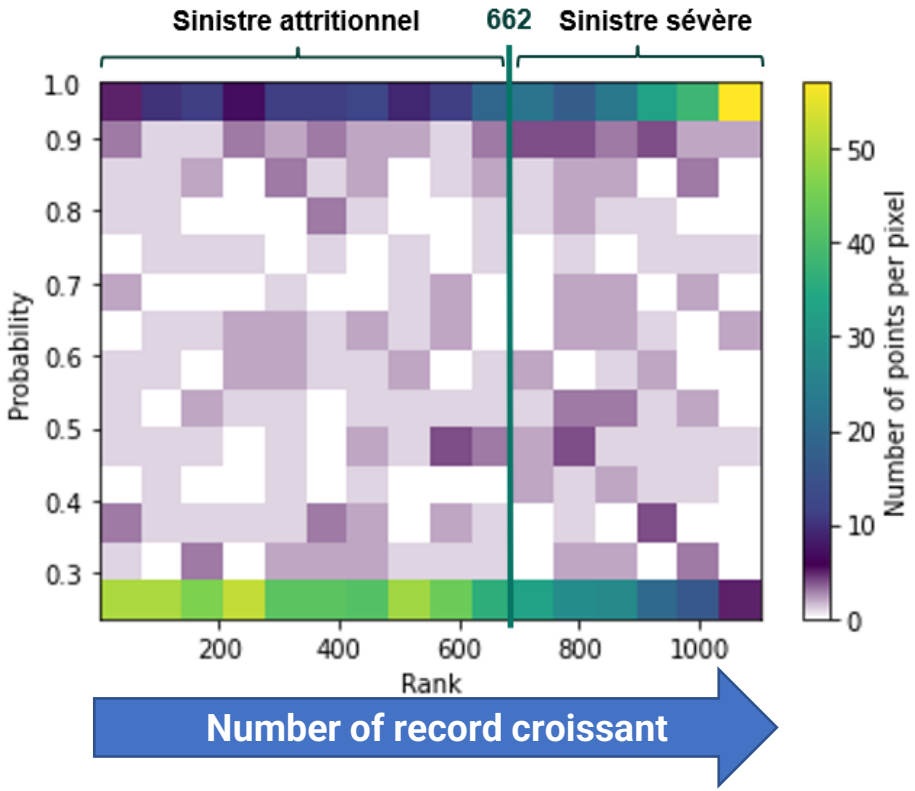
$$F1\ score = \frac{VP}{VP + MOYENNE(FP; FN)}$$

VP : Vrai positifs
FP : Faux positifs
FN : Faux négatifs

		Severity?	
		0	1
Pred	0	479	176
	1	182	265

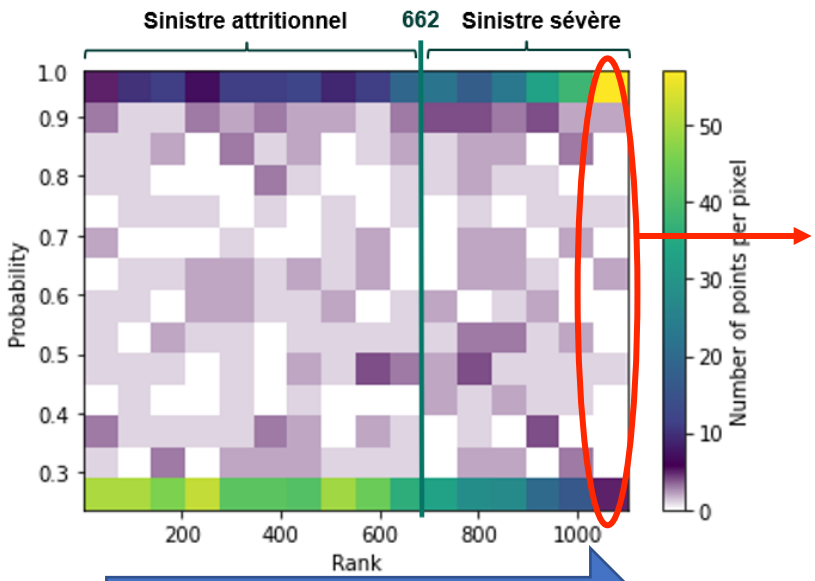
F1-Score = 60%

RÉSULTATS : DISTRIBUTION DES PRÉDICTIONS

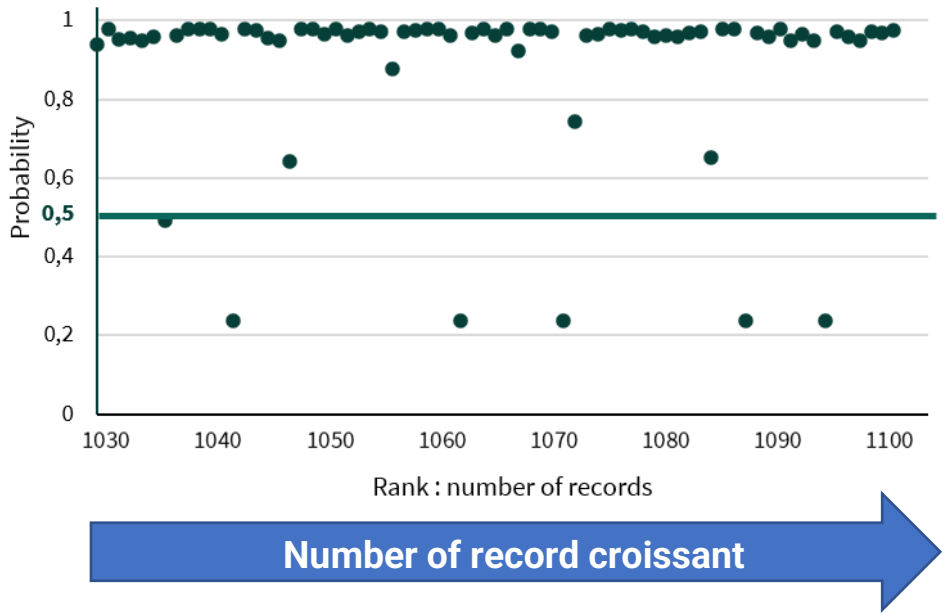


RÉSULTATS : DISTRIBUTION DES PRÉDICTIONS

Distribution des prédictions

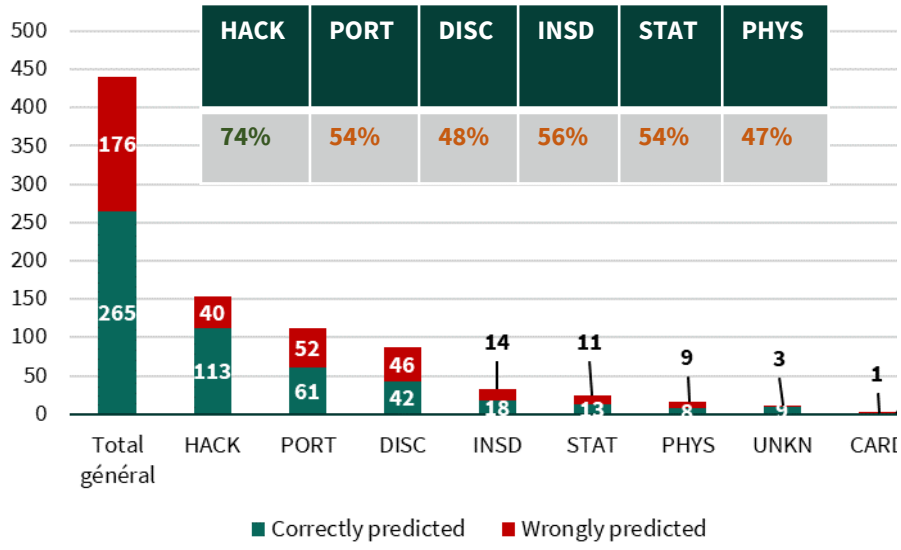


Zoom: sinistres particulièrement importants (>300,000 records)

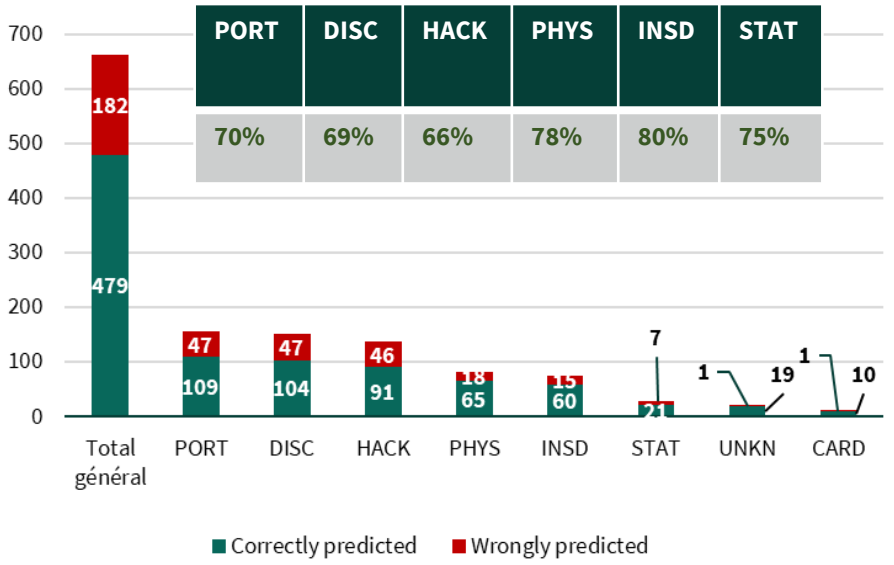


QUALITÉ DES PRÉDICTIONS PAR TYPE D'INTRUSION

Prédiction des sinistres graves

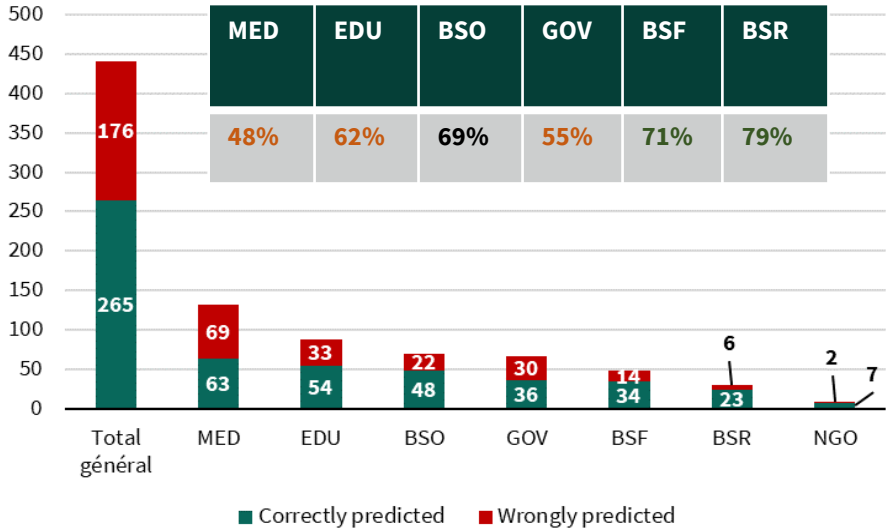


Prédiction des sinistres attritionnels

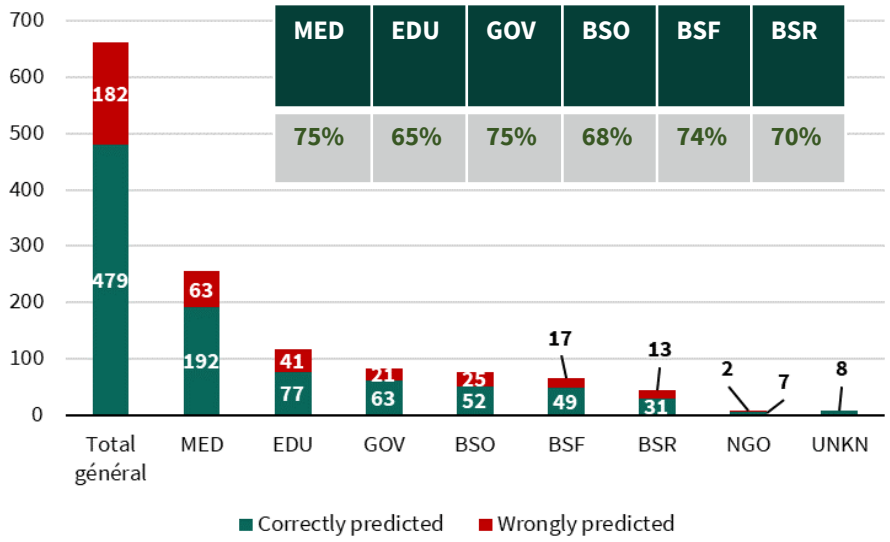


QUALITÉ DES PRÉDICTIONS PAR TYPE D'ORGANISATION

Prédiction des sinistres graves



Prédiction des sinistres attritionnels



AGENDA

1- Techniques d'embedding

2- Réseaux de neurones

3- Enrichissement du modèle

4- Applications et perspectives

TEXT MINING : EXPRESSIONS RÉGULIÈRES

Un réseau de neurone n'interprète pas la relation entre les nombres et les mots.

12 social security numbers ≠ 12 companies ≠ 12 million

Comment interpréter les données numériques dans les descriptions de sinistre ?

TEXT MINING : EXPRESSIONS RÉGULIÈRES

Un réseau de neurone n'interprète pas la relation entre les nombres et les mots.

12 social security numbers ≠ 12 companies ≠ 12 million

Comment interpréter les données numériques dans les descriptions de sinistre ?

- Dans 50% des descriptions, nous observons des formes récurrentes

nombre × mot

- **Ces informations sont directement en lien avec le « number of records »**

=> Les données numériques sont un indicateur de la sévérité du sinistre.

TEXT MINING : EXPRESSIONS RÉGULIÈRES

Un réseau de neurone n'interprète pas la relation entre les nombres et les mots.

12 social security numbers ≠ 12 companies ≠ 12 million

Approximation :

*A hacker […] has potentially revealed the names, Social Security numbers, and, in some cases, the birth dates and bank accounts of **27,000** employees working at **1,900** companies nationwide.*

Comment interpréter les données numériques dans les descriptions de sinistre ?

- Dans 50% des descriptions, nous observons des formes récurrentes

nombre × mot

- **Ces informations sont directement en lien avec le « number of records »**

=> Les données numériques sont un indicateur de la sévérité du sinistre.

EXPRESSIONS RÉGULIÈRES

Dictionnaire

MOT	nb. Pred KO	nb. Pred OK	Total	Pouvoir prédictif?
patient	16	55	71	77%
people	15	52	67	78%
million	43	5	48	10%
record	2	27	29	93%
student	4	24	28	86%
employee	2	22	24	92%
current	7	14	21	67%
individual	2	19	21	90%
year	12	7	19	37%
customer	3	15	18	83%

MOT	nb. Pred KO	nb. Pred OK	Total	Pouvoir prédictif?
companies	5	0	5	0%

Approximation :

*A hacker […] has potentially revealed the names, Social Security numbers, and, in some cases, the birth dates and bank accounts of **27,000** employees working at **1,900** companies nationwide.*

Dictionnaire $\boldsymbol{V}$ et paires $(\boldsymbol{N_u}, \boldsymbol{M_u})$

$$\sum_{k=0}^{n} \boldsymbol{N_u 1_{M_u \in V}} < \boldsymbol{Seuil}$$

REVENONS À NOTRE EXEMPLE

*A hacker […] has potentially revealed the names, Social Security numbers, and, in some cases, the birth dates and bank accounts of **27,000** employees working at **1,900** companies nationwide.*

Dictionnaire $\boldsymbol{V}$ et paires $(\boldsymbol{N_u}, \boldsymbol{M_u})$

$$\sum_{k=0}^{n} \boldsymbol{N_u 1_{M_u \in V}} < \boldsymbol{Seuil}$$

27000*1 + 1900*0 = 27000 > 4700

Le sinistre est grave d'après cet estimateur !

REVENONS À NOTRE EXEMPLE

*A hacker […] has potentially revealed the names, Social Security numbers, and, in some cases, the birth dates and bank accounts of **27,000** employees working at **1,900** companies nationwide.*

Dictionnaire $\boldsymbol{V}$ et paires $(\boldsymbol{N_u}, \boldsymbol{M_u})$

$$\sum_{k=0}^{n} \boldsymbol{N_u 1_{M_u \in V}} < \boldsymbol{Seuil}$$

27000*1 + 1900*0 = 27000 > 4700

Le sinistre est grave d'après cet estimateur !

Type d'organisation	Perceptron	Expressions régulières
MED	63	+ 10
EDU	54	+ 11
BSO	48	+ 7
GOV	36	+ 10
BSF	34	+ 1
BSR	23	+ 0
NGO	7	+ 1

Nombre de sinistre grave selon la méthode

AGENDA

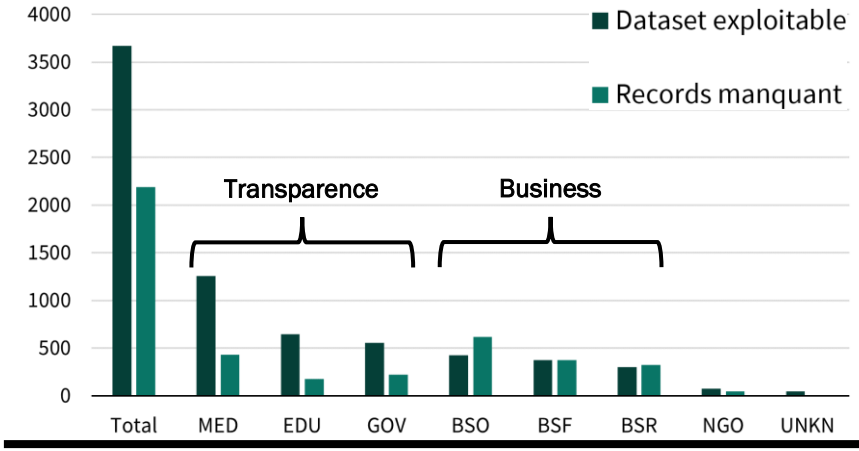
1- Techniques d'embedding

2- Réseaux de neurones

3- Enrichissement du modèle

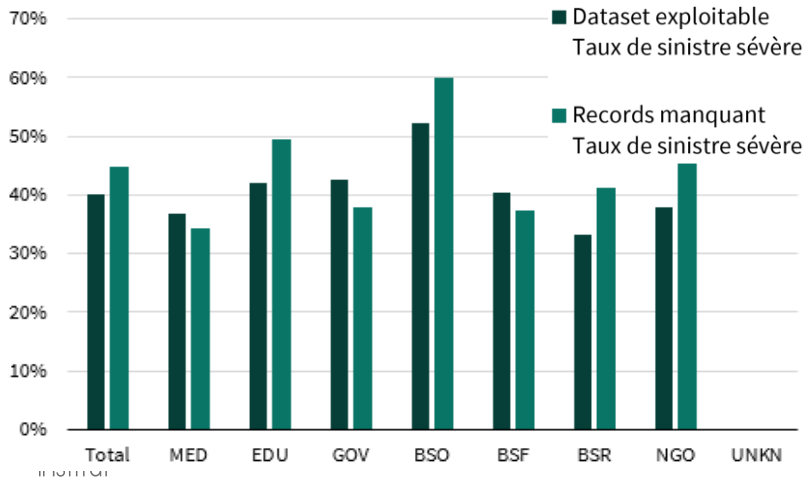
4- Applications et perspectives

DATABASE : MISSING NUMBER OF RECORDS



BSO, BSF, BSR

- Les catégories business est particulièrement représenté dans cette base
- Le diagnostic (nb. of records) semble donc moins transparent



EDU & business non liés au système bancaire et financier (BSO, BSR)

- Taux de sinistre sévère plus élevé

Hypothèse :

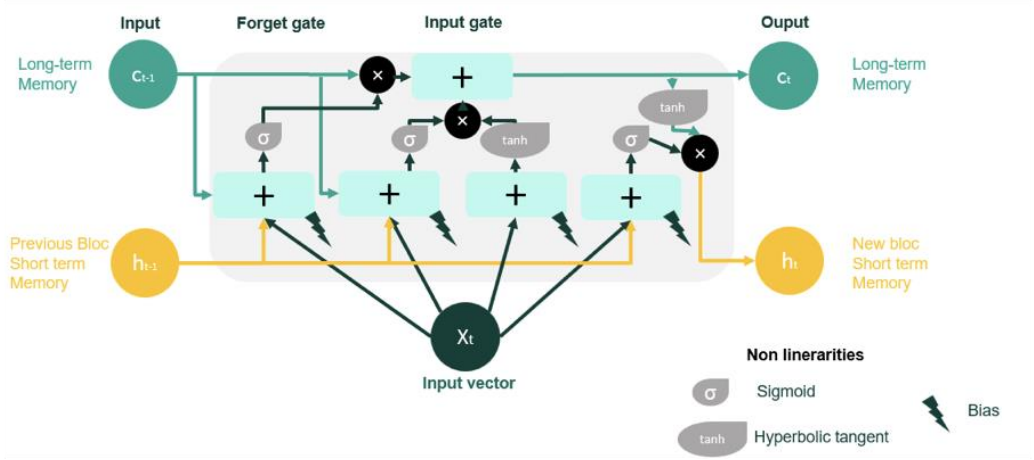
- Ces organisations sont moins bonnes sur le diagnostic
- Elles ne souhaitent ou ne savent pas quantifier les données perdues

LE PROBLÈME DE LA GESTION D'INCIDENT

- Assurance cyber : comporte une part d'**assistance à la victime**.
- Si l'activité d'assistance n'est pas nouvelle en assurance, l'assureur dispose d'une faible expertise du cyber.
- Une utilisation des méthodes précédentes : **comment détecter les sinistres qui nécessitent une réponse particulière, afin d'améliorer leur gestion** (et de minimiser leurs conséquences néfastes) ?

- **Input** : rapports d'incidents, expertises préliminaires.
- **Output** : diagnostic et recommandations en termes d'assistance.
- **Extension** : suivi de l'évolution des sinistres au cours du temps

RÉSEAUX RÉCURRENTS POUR L'ANALYSE DU TEXTE



- Exemple : LSTM
- Autre référence sur problématique connexe : Cohen-Sabban, I., Lopez, O., Mercuzot, Y. (2021) *Automatic analysis of insurance reports through deep neural networks to identify severe claims*, **Annals of Actuarial Science**.

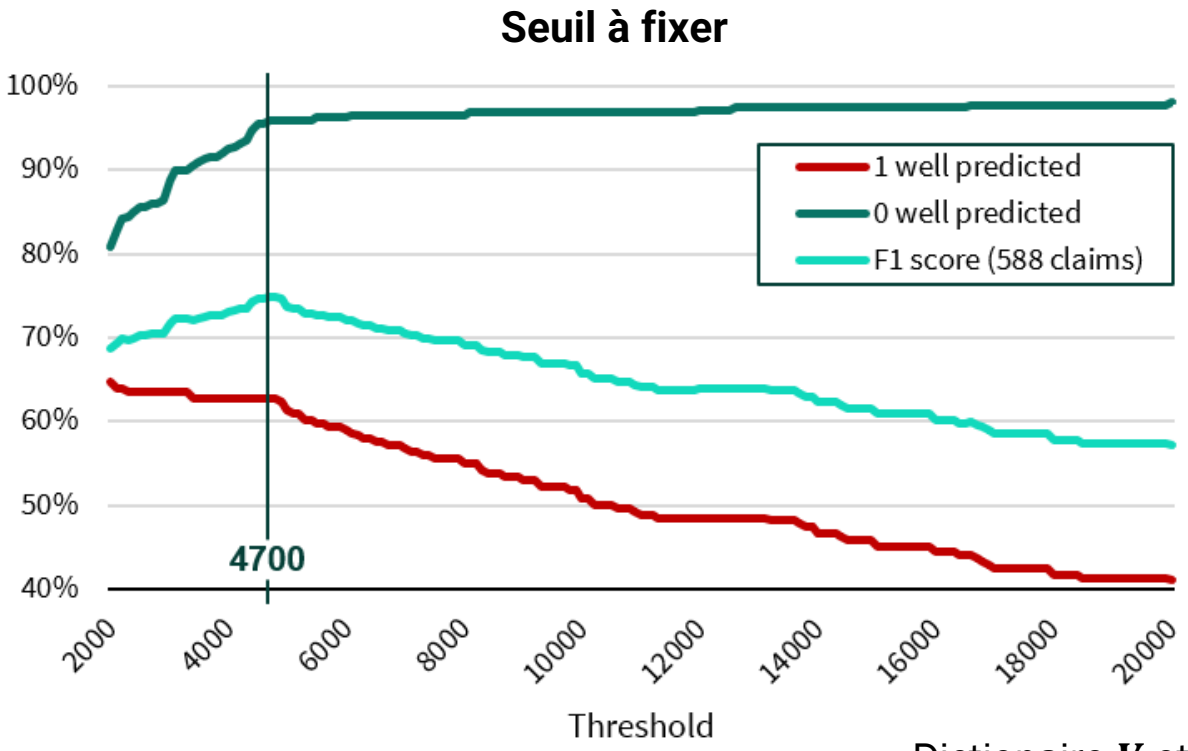
PERSPECTIVE : LE BAYÉSIEN

- Vision fréquentiste : on dispose $(X_1, \cdots, X_n)$ de loi $\mathbb{P}_{\theta_0}$, et on estime θ_0 à partir de ces seules informations.
- Vision bayésienne : on suppose que θ_0 est aléatoire, de loi a priori π donnée, et on observe $(X_1, \cdots, X_n)$ dont la loi, sachant $\theta_0 = t$, est $\mathbb{P}_t$.

- A priori : expertise préliminaire.
- Question : comment transformer cette expertise en « a priori » au sens mathématique du terme ?

ANNEXES

EXPRESSIONS RÉGULIÈRES



Dictionaire $\boldsymbol{V}$ et paires $(\boldsymbol{N_u}, \boldsymbol{M_u})$

$$\sum_{k=0}^{n} \boldsymbol{N_u 1_{M_u \in V}} < \boldsymbol{Seuil}$$

MISSING NUMBER OF RECORDS