

Optimisation de l'Extraction de Données des Rapports SFCR

Réalisé par :
Rayen CHENBEH

Plan



Contexte et Objectif



Fonctionnement de ma solution



Performance de la solution apportée



1

Contexte et Objectif

Contexte



Les rapports SFCR sont cruciaux pour évaluer la santé financière des assureurs, mais leur analyse est entravée par la présence de contenu non pertinent, tel que les informations en bas et en haut des pages, ainsi que le contenu des tableaux.



La reconnaissance optique de caractères (OCR) convertit ces rapports en données numériques, mais elle ne distingue pas entre les informations utiles et inutiles, rendant le processus d'analyse difficile et chronophage.

Objectif



-Améliorer la Fonction de Filtrage : `produce_brut()`



-Classification Automatique : classifiez chaque segment de texte comme "Inutile", "Paragraphe" ou "Titre" pour faciliter l'analyse ultérieure et l'extraction d'insights.



-Optimisation de l'Analyse : L'objectif est d'automatiser le processus de distinction entre le contenu pertinent et non pertinent

2

Fonctionnement de ma solution

Améliorer la Fonction de Filtrage : produce_brut()



J'ai ajouté une nouvelle caractéristique nommée 'digit_count' qui représente le nombre de chiffres dans le texte extrait.

```
f["height"] = max(y_list) - min(y_list)
f["area"] = f["width"] * f["height"]
f["chars"] = len(text)

f['digit_count'] = sum(c.isdigit() for c in text)
```

Fonction de Classification Automatique



Titres: Basés sur la taille des caractères, la position sur la page, le nombre de chiffres et la longueur du texte.



Paragraphes: Détectés grâce à une largeur minimum, une longueur de texte supérieure et un nombre de chiffres minoritaire



Contenu Inutile : Tout ce qui ne correspond pas aux critères de 'Titre' ou de 'Paragraphe'

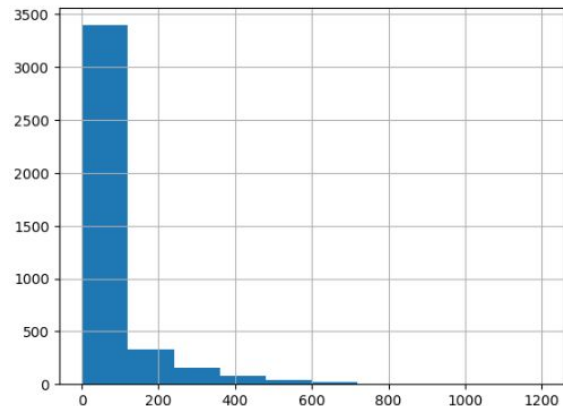
Seuils Clés :

- Taille des caractères pour les titres.
- Position sur la page pour exclure les en-têtes et pieds de page.
- Nombre de caractères pour distinguer les paragraphes.
- Nombre de chiffres pour identifier les tableaux et graphiques.

Visualisation des Caractéristiques Clés : Analyse Graphique

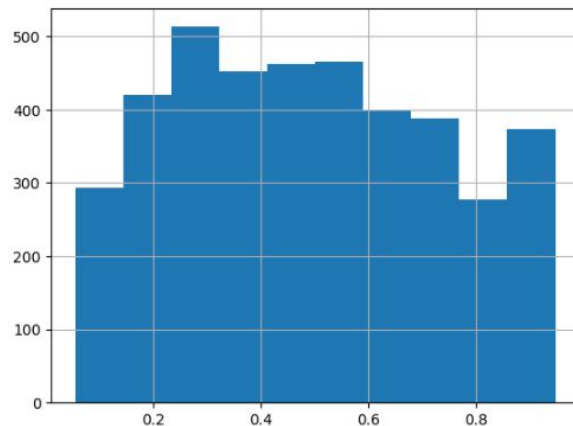
```
df['chars'].hist()
```

<Axes: >



```
df['pos_y'].hist()
```

<Axes: >



3

**performance de la solution
apportée**

Visualisation des Résultats de Classification dans le PDF

- Objectif de la Fonction : Illustrer visuellement la classification des éléments textuels directement dans le fichier PDF.
- Utilisation de PyMuPDF pour dessiner des rectangles sur le PDF.

-Signification des Couleurs :

- Bleu pour les 'Titres'.
- Noir pour les 'Paragraphes'.
- Rouge pour le contenu 'Inutile'.

Exercice 2022

Gestion du capital

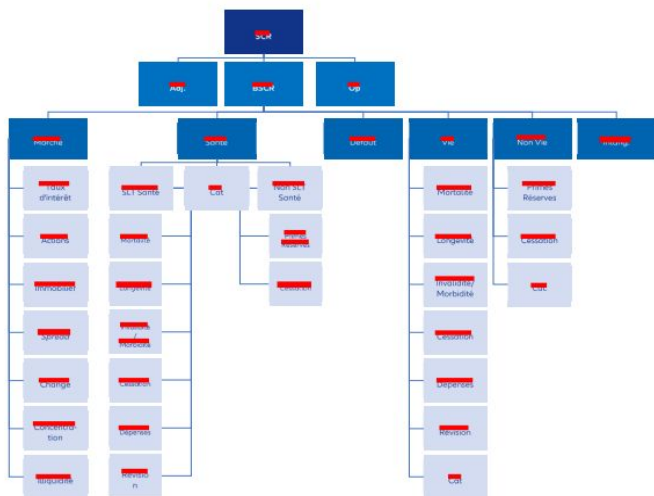


Figure 11 - Structure des risques en Formule Standard

E.4.3. Agregation du Risk Capital et capital add-on

Le processus d'agregation corréle les différents facteurs (actions, taux d'intérêt, etc.) par catégorie de risques (risque de crédit, risque opérationnel, etc.) afin d'obtenir le capital de solvabilité total diversifié. Pour l'agregation des risques, la méthodologie utilisée repose sur une copule de Gauss. La dépendance entre les risques et la copule est gérée par une matrice de corrélation. À l'exception des corrélations basées sur des données historiques, quatre (4) étapes sont requises pour déterminer la corrélation (en se fondant sur l'avis d'experts au niveau du Groupe) :

- la rationalisation y compris les événements extrêmes ;
- l'analyse empirique si possible ;
- la détermination des seuils de tolérance pour obtenir une matrice de corrélation valide ;
- la validation au niveau du Groupe Allianz par le Comité d'Approbation des Paramètres et des Hypothèses (PAAL - Parameters and Assumptions Approval Committee).

Le capital de solvabilité agrégé est ensuite obtenu en calculant selon la méthode Hotelli-Davala la VaR à 99,5% de la distribution

Activité et résultats

Exercice 2022

A.1. Activité

A.1.1. Renseignements généraux

Allianz Vie est une société anonyme de droit français régie par le Code des Assurances.

Elle est placée sous le contrôle de l'Autorité de Contrôle Prudentiel et de Résolution - ACPR, qui est l'organe de supervision français de la banque et de l'assurance. L'ACPR est situé au :

4 Place de Budapest - CS 92459
75 436 Paris Cedex 09.

Le Commissaire aux Comptes titulaire d'Allianz Vie est *PricewaterhouseCoopers Audit*, société par actions simplifiée au capital de 2 510 460 euros dont le siège social est situé 63 rue de Villiers - 92200 Neuilly-sur-Seine. Il a été désigné le 22 juin 2018 pour la durée restant à courir du mandat de son prédécesseur (KPMG SA), soit jusqu'à l'Assemblée Générale appelée à statuer sur les comptes de l'exercice clos au 31 décembre 2022.

Le capital d'Allianz Vie est entièrement libre. Au 31 décembre 2022, il s'élève à 643 054 425 euros et se compose de 42 670 295 actions. La répartition du capital et des droits de vote est présentée dans le tableau ci-dessous :

	2022			2021		
	Nombre d'actions	% du capital	% des droits de vote	Nombre d'actions	% du capital	% des droits de vote
Acronyme						
Allianz France	42 131 176	98,28%	98,28%	42 131 176	98,28%	98,28%
4 cours Actionnaires - CS 92076 Paris La Défense Cedex						
Allianz IARD	739 080	1,72%	1,72%	739 080	1,72%	1,72%
4 cours Actionnaires - CS 92076 Paris La Défense Cedex						
Caisse des dépôts	62	0,00%	0,00%	62	0,00%	0,00%
Administration des Domaines	2	0,00%	0,00%	2	0,00%	0,00%
Autres actionnaires minoritaires	25	0,00%	0,00%	25	0,00%	0,00%
Total	42 670 295	100%	100%	42 670 295	100%	100%

Tableau 1 - Répartition du capital au 31 décembre 2022

Allianz Vie est une entité détenue de façon ultime par le Groupe Allianz, dont la maison mère Allianz SE est une société européenne de droit allemand et dont le siège social est situé à Königinstrasse 28, D-80802 Munich en Allemagne. La maison mère est placée sous le contrôle de la BaFin.

Allianz Vie est consolidée par intégration globale au sein du Groupe Allianz.

Allianz Vie est détenue à hauteur de 98,28% par Allianz France et 1,72% par Allianz IARD comme illustré par l'organigramme simplifié de la Figure 1 :

A.4 Résultats des autres activités

/ Résultat net

(En millions d'euros, sauf pourcentages)	31 décembre 2022	Vont: Activité Holding	Dont: Activités de Réassurance	31 décembre 2021 AXA SA	31 décembre 2021 AXA Global Re
Chiffre d'affaires brut	5 934		5 934		5 983
Ratio de sinistralité courant net	78,3%		78,3%		76,5%
Ratio de sinistralité tous exercices net	86,0%		86,0%		86,5%
Marge technique nette hors frais généraux	73		73		197
Service result	58		58		58
Frais généraux	- 517	- 474	- 43	- 434	- 45
Taux de frais généraux	11,6%		1,0%		1,8%
Résultat financier net	3 035	2 950	85	2 579	45
Autres produits et charges	- 176	- 183	7	10	- 2
Charge d'impôt	336	336		665	- 71
RÉSULTAT NET	2 809	2 629	180	2 819	182
Réallocation revenus nets de placement	-	- 101	101	-	-
RÉSULTAT NET (Comptes sociaux de la Société)	2 809	2 528	281	2 819	182

Le résultat net social après impôt d'AXA SA pour l'exercice clos le 31 décembre 2022 est un bénéfice de 2 809 millions d'euros, dont 2 528 millions d'euros pour l'activité de holding et 281 millions d'euros pour l'activité de réassurance après réallocation de 185 millions d'euros de revenus nets des placements en fonction de la pondération des provisions techniques nettes et des fonds propres. Ce montant réaffecté dépasse de 101 millions d'euros le résultat financier net alloué à l'activité de réassurance. Pour donner une vision plus économique, le résultat financier net alloué à chaque activité dans le tableau ci-dessus et pour l'analyse ci-dessous a été conservé.

Le résultat net au 31 décembre 2021 s'élève à 2 819 millions d'euros pour l'activité de holding et à 182 millions d'euros pour l'activité de réassurance.

Le résultat net de l'activité de holding a diminué de - 190 millions d'euros pour s'établir à 2 629 millions d'euros, principalement en raison de :

- Le résultat net des placements s'élève à 2 950 millions d'euros, en hausse de 371 millions d'euros, dont :
 - 850 millions d'euros de dividendes reçus des filiales, s'élevant à 4 345 millions d'euros par rapport à 3 495 millions d'euros en 2021,
 - partiellement compensés par des provisions supplémentaires pour dépréciation des investissements dans les filiales de 507 millions d'euros.
- Les résultats non techniques affichent un déficit de 606 millions d'euros contre 419 millions d'euros en 2021, dont :
 - 123 millions d'euros de provision sur les moins-values latentes sur options dans le cadre de la protection de la liquidité de la Société,
 - 40 millions d'euros d'augmentation compte tenu de la poursuite des investissements dans le cadre de la transformation informatique.

Geometriae et Algebræ - 4042

Geometriae et Algebræ - 4042

Table 1

892115

Hilary

1800-2000

[illegible]

-Analyse des resultats:

-Nous pouvons observer à partir des exemples précédents que le modèle réussit à distinguer efficacement les paragraphes, les titres et le contenu inutile (tels que les tableaux, graphiques, ainsi que les en-têtes et pieds de page).

-Il est possible d'affiner davantage le modèle pour améliorer sa performance, car certains cas montrent que le modèle ne parvient pas toujours à distinguer clairement les différentes parties du texte.

Nom	Fonction principale
M. Fabien Wathe	Directeur Général d'Aliaz France
M. Pascal Thebe	Head of LEO Office
M. Remi Saucé	Directeur Unité Finance
Mme. Anne-Sophie Groucho	Directrice Unité Transformation Operations et Resilience
M. Julien Martinez	Directeur Unité Data, Engagement, Marketing et Stratégie
M. Matthias Seewald	Directeur Unité Investissements
M. Alexandre du Garreau	Directeur Unité Distribution
Mme Corinne Lapiere	Directrice Unité Service Client
M. Sylvain Loral	Directeur Unité Assurance de Personnes
M. François Needy	Directeur Unité Assurance de Biens et de Responsabilités

(Comme le montre ce tableau, certaines lignes ont été incorrectement classées en tant que paragraphes à cause des conditions spécifiées dans notre modèle. Cela indique le besoin d'ajuster ces conditions pour une classification plus précise.



**MERCI DE VOTRE
ATTENTION!**