



UFR de Sciences Appliquées et de Technologie (SAT)

UFR de Sciences Economiques et de Gestion (SEG)

Département : Informatique

Département : Gestion

Master II

Filière : Méthodes Informatiques Appliquées à la Gestion (MIAGE)

----- **Mémoire de recherche de master II** -----

Sujet : Développement d'un modèle de Machine Learning pour faire une analyse financière (historique et prédictive) et le développement d'un Chatbot pour interroger les états financiers

Présenté par :

Omar Abd Al Wahab DIASSE

Encadreurs :

Pr Jean Marie DEMBELE

Dr Alioune Badara Mbengue

Présentée le : .. / .. / **2024**

Membres du jury :

... ..

... ..

... ..

Année académique :

2022/2023

Remerciements

Allah

Pr Dembélé et Dr Mbengue

Tous les établissements et professeur

Parent

Frère et sœur

Camarades de classe

Camarade de chambre

Moi même

Sommaire

Remerciements.....	I
Sommaire	II
Liste des figures	IV
Liste des tableaux.....	V
Liste des équations	VI
Liste des sigles et des acronymes.....	VIII
Introduction générale	1
1 Chapitre I : Revue de la littérature scientifique	3
Introduction de chapitre	3
1.1 Généralités l'intelligence artificielle	3
1.2 Etat de la recherche sur l'IA appliquée à la finance.....	9
Conclusion partielle	15
2 Chapitre II : L'analyse financière	16
Introduction de chapitre	16
2.1 Présentation des états financiers.....	16
2.2 L'analyse financière proprement dite.....	30
Conclusion partielle	37
3 Chapitre III : Théories derrières les algorithmes d'intelligence artificielle	38
Introduction de chapitre	38
3.1 Les prérequis	38
3.2 Les algorithmes d'intelligence artificielle.....	43
Conclusion partielle	57
4 Chapitre IV : Implémentation des modèles	58
Introduction de chapitre	58

4.1	Spécifications des besoins	59
4.2	Développement des modèles intelligents	62
4.3	L'interface graphique	77
	Conclusion partielle	83
	Conclusion générale et perspectives	83
	Bibliographies	A
	Webographies	D
	Annexes.....	E
	Table des matières.....	Y

Liste des figures

Figure 1 Le neurone biologique (Reche, 2019).....	6
Figure 2 Zoom sur la dérivée d'une fonction (mathisfun)	41
Figure 3 Les variations de Learning Rate (Cayla, 2021)	45
Figure 4 Simple reseau de neurone (Kumar, 2020).....	46
Figure 5 La fonction sigmoid (Saleem, 2023).....	47
Figure 6 Différents degrés de la régression polynomiale (Graph of polynomial functions, 2020)	48
Figure 7 Structure d'un arbre de decision (Zhao, 2021).....	52
Figure 8 Representation de la fonction XOR (Oman, 2016)	55
Figure 9 Structure de réseau de neurones (Ahmad, 2020)	56
Figure 10 Reseau d'un CNN (Shahriar, 2023).....	56
Figure 11 Réseau d'un RNN (Poudel, 2023)	57
Figure 12 Gestion de projet (Roosj, 2020)	61
Figure 13 Fichier CSV des données.....	66
Figure 14 Conversation 1	75
Figure 15 Conversation 2	76
Figure 16 Conversation 3	76
Figure 17 Conversation 4	77
Figure 18 Architecture globale.....	81
Figure 19 Page d'accueil de l'application.....	81
Figure 20 La page consultation du bilan de l'application	82
Figure 21 La page modification du TFT de l'application	82
Figure 22 La page de l'analyse financiere de l'application	82

Liste des tableaux

<i>Tableau 1 Exemple de bilan</i>	<i>20</i>
<i>Tableau 2 Exemple de compte de résultat.....</i>	<i>25</i>
<i>Tableau 3 Exemple de tableau des flux de trésorerie</i>	<i>29</i>
<i>Tableau 4 Les familles de ratios.....</i>	<i>36</i>
<i>Tableau 5 One hot encoding.....</i>	<i>72</i>
<i>Tableau 6 Bag of word.....</i>	<i>72</i>
<i>Tableau 7 Bag og n-grams.....</i>	<i>73</i>
<i>Tableau 8 Ratios de probabilité de défauts</i>	<i>K</i>
<i>Tableau 9 Zones d'Altman</i>	<i>K</i>
<i>Tableau 10 Resultat XOR</i>	<i>O</i>
<i>Tableau 11 Exemple donnees pour arbre de decision.....</i>	<i>P</i>
<i>Tableau 12 Exemple donnees pour K-means</i>	<i>R</i>

Liste des équations

Équation 1 Actif total.....	17
Équation 2 Actif immobilies.....	17
Équation 3 L'actif circulant.....	17
Équation 4 Le passif total.....	18
Équation 5 Les capitaux propres.....	18
Équation 6 Les dettes financières.....	18
Équation 7 Le passif circulant.....	19
Équation 8 Les charges.....	21
Équation 9 Les produits.....	22
Équation 10 La marge commerciale.....	22
Équation 11 Chiffre d'affaires 1.....	23
Équation 12 Chiffre d'affaires 2.....	23
Équation 13 La valeur ajoutée.....	23
Équation 14 L'excédent brut d'exploitation.....	23
Équation 15 Le résultat d'exploitation.....	24
Équation 16 Le résultat financier.....	24
Équation 17 Le résultat exceptionnel.....	24
Équation 18 Le résultat net.....	25
Équation 19 La trésorerie initiale.....	26
Équation 20 Les flux de trésorerie liés aux activités opérationnels.....	26
Équation 21 Variation du BFR.....	27
Équation 22 Les flux de trésorerie liés aux activités d'investissements.....	27
Équation 23 Les flux de trésorerie provenant des capitaux propres.....	27
Équation 24 Les flux de trésorerie provenant des capitaux étrangers.....	27
Équation 25 Les flux de trésorerie liés aux activités de financement.....	28
Équation 26 La trésorerie finale.....	28
Équation 27 Taux de croissance 1.....	31
Équation 28 Taux de croissance 2.....	31
Équation 29 Mean Square Error.....	44
Équation 30 Mise à jour des poids.....	45
Équation 31 Le chainage des dérivées partielles.....	46
Équation 32 Calcul de la sortie observée.....	46
Équation 33 La fonction sigmoïde.....	47

Équation 34 Separateur en SVM dans un espace 2D.....	50
Équation 35 Fonction de linéarisation	50
Équation 36 Calcul de distance en SVM.....	50
Équation 37 Erreur en SVM.....	50
Équation 38 Fonction de minimisation de l'erreur en SVM.....	51
Équation 39 Entropie	52
Équation 40 Gain d'information	52
Équation 41 Calcul de probabilité.....	53
Équation 42 Probabilité conditionnelle.....	53
Équation 43 Normalisation de probabilité.....	54
Équation 44 Zscore d'Altman.....	K

Liste des sigles et des acronymes

AF	Analyse financière
ANN	Artificial Neural Network
API	Application Programming Interface
ANSD	Agence Nationale des Statistiques et de la Démographie
BAC	Baccalauréat
BFG	Besoin de Financement Global
BFR	Besoin de Fond de Roulement
BRVM	Bourse Régionale des Valeurs Mobilières
CA	Chiffres d'affaires
CNN	Convolutional Neural Network
CSV	Comma Separated Values
CTIC	Incubateur des Startups
DL	Deep Learning
EBE	Excédent Brute d'Exploitation
F CFA	Franc Communauté Financière Africain
FTAF	Flux de Trésorerie des Activités de Financement
FTAI	Flux de Trésorerie des Activités d'Investissement
FTAO	Flux de Trésorerie des Activités Opérationnelles
GAFAM	Google, Amazon, Facebook, Apple, Microsoft
GI	Gain d'Information
GUI	Graphical User Interface
HAO	Hors Activité Ordinaire
IA	Intelligence Artificielle
IEEE	Institute of Electrical and Electronics Engineers
JSON	JavaScript Object Notation
LLM	Large language model
LR	Learning Rate
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MIAGE	Méthodes Information Appliquées à la Gestion
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MSE	Mean Squared Error
MVC	Model View Controller
MySQL	My Structure Query Language
NB	Naïve Bayes
NLP	Natural Language Processing
PDF	Portable Document Format
REP	Résultat Exceptionnel
REST	REpresentational State Transfer
REX	Résultat d'exploitation
RF	Résultat Financier
RN	Résultat net
RNN	Recurrent Neural Network

SGBD	Système de Gestion de Base de Données
SIG	Solde Intermédiaire de Gestion
SOAP	Simple Object Access Protocol
SML	Supervised Machine Learning
SQL	Structure Query Language
SVM	Support Vector Machine
TFT	Tableau des Flux de Trésorerie
UEMOA	Union Economique et Monétaire Ouest-Africaine
UML	Unsupervised Machine Learning
VA	Valeur Ajoutée
XAMPP	Cross-Platform, Apache, MySQL, PHP, and Perl
XML	eXtensible Markup Language
XOR	eXclusive OR

Introduction générale

Dans les années 50, s'est tenue une conférence qui avait rassemblé plusieurs chercheurs de l'époque sur un domaine dont eux-mêmes n'avaient pas conscience de comment cela allait révolutionner le monde. Parmi ce florilège de scientifiques se trouver un mathématicien du nom de John McCarthy, organisateur par ailleurs de cette conférence, qui s'est illustré d'une manière simple : il a tout simplement proposé le terme « *Artificial Intelligence* » pour décrire cette nouvelle science qui était en train d'émerger. C'était la conférence de Dartmouth dans l'Etat du New Hampshire aux Etats-Unis en 1956. Toutes les personnes qui étaient présentes dans cette conférence venaient d'assister non pas à la naissance de l'intelligence artificielle mais au baptême de cette dernière.

Ce domaine qui est l'intelligence artificielle s'applique aujourd'hui dans plusieurs secteurs de nos vies notamment la finance qui va nous intéresser pour ce mémoire. C'est ainsi que nous avons choisi comme sujet de mémoire « **Développement d'un modèle de machine pour faire une analyse financière (historique et prédictive) et le développement un Chatbot pour interroger les états financiers** ». Nous serons amenées à utiliser des termes techniques comme Machine Learning qui est l'apprentissage des machines, Deep Learning qui représente quant à lui l'apprentissage profond des machines et aussi l'analyse financière qui est un sous-domaine de la finance d'entreprise nous permettant de consulter la santé financière d'une entreprise.

Ayant toujours eus une affection particulière pour l'informatique, nous nous sommes naturellement orientés vers ce domaine. Après les premiers cours d'intelligence artificielle, l'affection de l'informatique et de l'IA s'est renforcée puisque nous avons eu la chance de démystifier ce domaine complexe et très intéressant. Etant un étudiant de la MIAGE (Méthodes Informatiques Appliquées à la Gestion), il s'est avéré être pertinent de faire appliquer l'IA au domaine de la finance, c'est ainsi que nous avons choisi avec l'aide de nos professeurs encadreur la finance d'entreprise.

L'intelligence artificielle n'est pas une science nouvelle comme nous l'avons déjà vue même si cette dernière gagne beaucoup de popularité ces derniers temps. Et notre pays le Sénégal n'est pas en reste par rapport à tout cela, déjà en 1998 le professeur Seydina Moussa Ndiaye nous parler de *Reinforcement Learning* dans l'un de ses articles les plus cités. Plus proche de nous, beaucoup

d'initiatives ont été prise dans le sens de l'IA, il y a l'Agence Nationale des Statistiques et de la Démographie (ANSD) qui a ouvert un bureau d'IA pour la prédiction démographique. Cela ne s'arrête pas là, même le gouvernement du Sénégal a lancé un programme appelé « La stratégie IA » à travers le ministère de la communication des télécommunications et du numériques. En plus de tout cela vient s'ajouter un bon nombre de chercheurs et de jeunes passionnés qui essaient tant bien que mal de faire bénéficier ses technologies intelligentes à la population sénégalaise.

La question principale que ce travail de mémoire aura pour but de répondre est : **dans quelle mesure l'intelligence artificielle va pouvoir aider les entreprises cotées à la Bourse régionale des Valeurs Mobilières (BRVM) à faciliter leur analyse financière et tout ce qui tourne au tour de leurs états financiers ?**

Afin de mener cette mission à bien, nous adopterons une démarche bien spécifique, tout d'abord nous nous attèlerons à trouver des données avec lesquelles nous allons travailler. Une fois les données collectées et traitées, nous passerons par la suite à ce qu'on appelle la recherche en grille qui est une méthode utilisée en Machine Learning pour déterminer le meilleur modèle, celui qui sera le plus adapté à nos données. Deux familles de modèles vont être utilisées, il y a les modèles de prédiction et aussi les modèles de NLP (faire comprendre le texte à un ordinateur).

C'est ainsi à la fin de ce travail deux applications vont être produites sous forme de logiciel que les entreprises pourront utiliser pour faire leur analyse financière, prédire leurs états financiers, communiquer avec un Chatbot etc. et tout cela dans un environnement cousu à la taille de leur finance. Ces deux applications vont être déployées dans le réseau local de l'entreprise pour l'intégrité des données.

Dans le but de réaliser ce mémoire de manière efficace, notre document va être divisé en chapitres. Nous allons d'abord voir toute la littérature sur l'intelligence artificielle et la finance, les définitions, l'histoire, les recherches scientifiques (Chapitre I). En plus de cela nous réserverons une partie spéciale pour découvrir l'analyse financière, ses méthodes, son fonctionnement, ses intérêts (Chapitre II). L'une des parties les plus intéressantes de ce document, sera la partie des théories derrière l'intelligence artificielle (Chapitre III). Nous allons terminer avec la présentation des modèles ou nous allons voir de manière pratique comment crée-t-on une intelligence artificielle mais aussi la présentation d'une interface graphique conviviale pour accueillir les clients (Chapitre IV).

1 Chapitre I : Revue de la littérature scientifique

Introduction de chapitre

La revue de la littérature scientifique est obligatoire, car importante, pour n'importe quels documents considérés scientifiques. Ici, il est demandé au rédacteur de faire un tour sur les productions du thème sur lequel il veut travailler.

Pour ce qui nous concerne, notre sujet qui est le **Développement de modèles de Machine Learning pour faire une analyse prédictive des finances d'une entreprise et le développement d'un Chatbot pour communiquer avec ses états financiers**, comporte deux grands domaines. D'abord il y a le Machine Learning (ML) qui est un sous ensemble de l'intelligence artificielle (IA), ensuite l'analyse financière (AF) qui est un domaine de la finance d'entreprise.

Ainsi, nous allons commencer par faire la revue de littérature du côté de l'IA. Dans cette partie nous allons définir une IA et le présenter sur tous ses aspects. Après nous allons faire la revue de littérature sur la finance en général et l'analyse financière en particulier. Nous allons terminer par la revue des deux combinés.

1.1 Généralités l'intelligence artificielle

1.1.1 Définition de l'intelligence artificielle

Avant d'entrer dans les détails, dans les aspects techniques et scientifiques ou dans l'implémentation d'une IA, il serait bien de donner une vue globale de c'est quoi une intelligence artificielle. Cependant, il va être difficile de définir une IA sans savoir c'est quoi l'intelligence, sans savoir c'est quoi artificiel. Une fois s'être bien familiarisé avec ces deux termes, cela va s'avérer bien plus simple de parler de l'IA. Fort heureusement, beaucoup de recherches, études ont été faites par les scientifiques académiciens sur ses termes que nous nous donnons la tâche de définir.

1.1.1.1 C'est quoi l'intelligence

L'intelligence humaine est un concept qui est difficile à définir car on ne sait pas comment il fonctionne, ou est son siège dans le cerveau et on ne peut pas vraiment voir de différence notable entre le cerveau d'une personne intelligente et celle d'une personne qui l'est moins.

Néanmoins, cette difficulté n'a pas empêché, les chercheurs en psychologie d'essayer de définir la chose selon leur entendement et la manière qu'a l'Homme d'interagir avec son environnement. Les avis des scientifiques sont divers et variés. Dans son article de 1993 publié au *British Journal of Psychology* « *On What Intelligence Is* », Robert W. Howard nous fait un inventaire de définition proposées par des psychologues avant lui.

“The word ‘intelligence’ labels three different major concepts: g, the sum of an individual’s knowledge and skills, and the specific mental abilities important in a given culture” (Jensen, 1987).

“Intelligence is not an entity within the organism but a quality of behavior” (Anastasi, 1986).

Mais ce n'est pas seulement les psychologues qui ont tentés de définir l'intelligence, les chercheurs en IA aussi, c'est le cas de James S. Albus qui le définit comme : “. . . the ability of a system to act appropriately in an uncertain environment, where appropriate action is that which increases the probability of success, and success is the achievement of behavioral subgoals that support the system’s ultimate goal”.

“Intelligence constitutes the state of equilibrium towards which tend all the successive adaptations of a sensori-motor and cognitive nature, as well as all assimilatory and accommodatory interactions between the organism and the environment” (Piaget, 2005).

Cette définition nous renvoie à l'individu et son environnement, cet individu prend les données de l'environnement (Inputs) et réagi en conséquence (Outputs).

1.1.1.2 C'est quoi artificielle

La plupart des gens penseraient que l'artificielle est un terme sans ambiguïté et donc facile à définir, et à juste titre d'ailleurs. Mais nous savons qu'il y a des subtilités qui se cache derrière ce terme surtout dans le domaine de l'IA.

Des définitions simples comme « l'artificielle est tout ce qui est créé par l'homme » ou « c'est l'œuvre de l'Homme s'inspirant de la nature » ne sont plus d'actualité.

C'est pourquoi nous allons donner la définition de Francesco Bianchini dans son article « A New Definition of “Artificial” for Two Artificial Sciences » : “The artificial is what is humanly constructed, often in a natural model, also through the manipulation of natural systems and

processes, and maintains existing and acting/operating/behaving in an open-ended context or environment without human control, regardless the substance or materials of its constituent parts”.

Cette définition prend l'ensemble des aspects relatifs à l'IA mais aussi quelque peu à la biologie artificielle. Elle met aussi en avant non seulement la construction de l'Homme mais aussi sa non-intervention dans les opérations dites artificielles.

Mais un problème se pose, selon Francesco Bianchini lui-même avec cette définition peut-on considérer une IA créée par une IA toujours artificielle ?

1.1.1.3 Proposition de définitions de l'intelligence artificielle

Nous y voilà, les termes que nous avons décidé de donner des définitions vont nous permettre de définir ce sur quoi porte notre sujet de mémoire. Dans cette partie nous allons essayer de répondre à la question c'est quoi une IA. Nous allons voir que plusieurs scientifiques ont donné des définitions mais à la fin, c'est plus ou moins les mêmes.

Déjà en 1988 Asa SIMMONS et Steven CHAPPEL avaient publié un article dans le IEEE Journal of Oceanic Engineering sur lequel ils nous rappelaient la définition qu'avait donné Haugeland en 1985 : “The fundamental goal of this research IS not merely to mimic intelligence or produce some clever fate. Not at all. AI wants only the genuine article: machines with minds, in the full and literal sense”.

Vingt-quatre (24) ans plus tard, en 2012 le mathématicien et docteur en IA américain Matt L. Ginsberg donnait, dans son ouvrage « Universal intelligence : A definition of machine intelligence », la définition suivante : “Artificial Intelligence is the enterprise of constructing an artefact that can reliably pass the Turing test”. Dans la suite du document, nous allons parler du teste de Turing.

Maintenant plus proche de nous, au moment où ce mémoire est en train d'être écrit, de nouvelles définitions émergent. En janvier 2023, Haroon Sheikh, Corien Prins & Erik Schrijvers ont copublié un article « Artificial Intelligence : Definition and Background » où ils ont défini la chose comme suit : “Systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals.”

Nous voyons que plusieurs auteurs à travers le temps ont donné leurs définitions de l'IA selon leur entendement de la chose et leurs domaines de recherches. Mais à la fin ils se rejoignent tous.

Au vu de tout cela, nous pouvons conclure que l'AI a pour objectif d'imiter l'intelligence humaine en faisant des tâches qui auraient pu être jugées impossible à faire pour les machines. Cette dernière n'est pas une proposition de définition.

1.1.2 Des autres sciences à l'intelligence artificielle

1.1.2.1 Les mathématiques

“These considerations show that there is a tremendous need for mathematics in the area of artificial intelligence. And, in fact, one can currently witness that numerous mathematicians move to this field, bringing in their own expertise” (Kutyniok, 2022).

Les mathématiques constituent le soubassement de l'IA, comme le vient de rappeler Gitta Kutyniok, c'est-à-dire que tous les algorithmes d'IA reposent derrière des théories mathématiques. D'ailleurs nous allons voir cela dans la suite du document en détails mais pour le moment donnons des exemples : **les dérivées, l'algèbre linéaire, les probabilités et statistiques.**

1.1.2.2 La biologie

Pour parler de l'impact de la biologie dans l'intelligence artificielle, il nous faut forcément parler du neurone biologique.

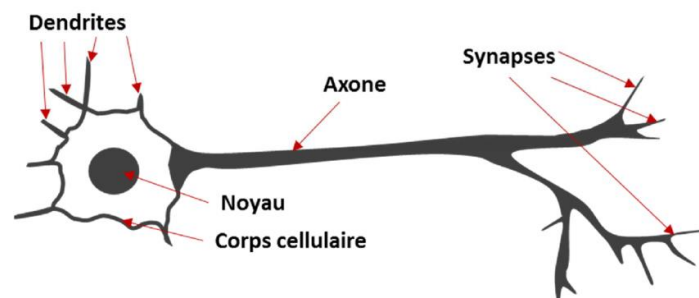


Figure 1 Le neurone biologique (Reche, 2019)

Le neurone est une cellule spécialisée dans le traitement et la transmission de l'information, ce qui est reflété par sa morphologie très particulière. Il se compose d'un corps cellulaire, le soma, et d'expansions : les dendrites d'une part, et l'axone d'autre part. Les dendrites forment des ramifications qui rentrent en contact avec d'autres neurones, typiquement de l'ordre de 10000, et dont le rôle est de recevoir des informations, électriques ou chimiques. L'axone est un

prolongement de la cellule, typiquement long de quelques millimètres, qui conduit un signal électrique jusqu'à son arborisation terminale, où il peut alors entrer en contact avec les dendrites d'autres neurones. La jonction axone-dendrite est appelée synapse (Brette, 2003). C'est cette structure du neurone biologique que les réseaux de neurones artificiels vont essayer de répliquer à travers des calculs mathématiques.

1.1.2.3 L'informatique

Là, un non-initié peut dire que l'IA est un sous-domaine de l'informatique, que nenni ! Nous allons voir qu'il y a des subtilités, on peut dissocier ces termes même s'ils sont bien liés et vont le rester dans le temps.

C'est ainsi que nous allons faire donner un petit aperçu de comment marche un algorithme d'IA pour comprendre cela. Un réseau de neurones prend des données en entrée, les normalise, puis il va les passer à la prochaine couche à travers la multiplication avec les poids. Cette opération va se répéter autant de fois que nécessaire et enfin on va avoir une sortie. Dans ce processus, il n'y a que des calculs mathématiques, ce qui veut dire qu'on peut faire cela sur une feuille (chose qui va certainement prendre beaucoup, beaucoup de temps), on peut le faire aussi avec une calculatrice, ou tout support nous permettant de faire des calculs.

Tout cela pour dire que l'ordinateur nous sert tout simplement à faciliter les calculs et afficher les résultats dans une interface graphique du fait qu'il est la calculatrice la plus puissante. Mais encore, nous pouvons noter que Warren S. McCulloch et Walter Pitts n'avaient pas utilisé d'ordinateur pour réaliser le premier réseau de neurones.

1.1.2.4 La cybernétique

La cybernétique est une science qui est dérivée de plusieurs autres sciences, c'est le mathématicien américain Norbert Wiener que l'on attribue la création de cette science. La cybernétique essaie de répondre à la question de savoir comment les systèmes peuvent se contrôler elles-mêmes. L'exemple qui est très souvent donné c'est le thermostat qui récupère les informations de son environnement c'est-à-dire la température ambiante et se régule lui-même en s'ajustant à la température désirée.

Cette science a beaucoup apporté à l'IA car le fonctionnement de plusieurs modèles intelligents s'inspire des théories de la cybernétique, comme le fait d'utiliser sa propre erreur pour se rectifier

soi-même dans un réseau de neurones ou le fait de donner un bonus ou malus à un agent dans un environnement s'il fait le bon ou le mauvais choix dans un *Reinforcement Learning*.

1.1.3 Domaines d'application de l'intelligence artificielle

1.1.3.1 Les prédictions

L'IA peut être utilisée dans la prédiction, sortir une valeur future à travers des valeurs passées. Cette dimension de l'IA est utilisée dans la prédiction météo, des bourses, la démographie. C'est l'une des applications les plus répandues de l'IA.

1.1.3.2 La robotique

Avant l'avènement des chatbots, la robotique était le domaine auquel tout le monde pensait quand on parle d'IA, car c'est l'une des plus anciennes applications de cette dernière. L'IA va permettre au robot de se déplacer librement dans l'espace 3D où nous évoluons. Avec des algorithmes, les robots (humanoïdes par exemple) peuvent reconnaître des objets et interagir en conséquence. Nous avons déjà donné l'exemple d'ASIMO, mais il y a aussi Titan (Cyberstein Robots) et Atlas (Boston Dynamic) entre autres.

1.1.3.3 La finance

Bien sûr le domaine qui nous intéresse le plus dans ce document, les IA sont utilisées en finance pour prédire les cours des bourses, faire la détection de fraude, prévenir les craques, la faillite etc. l'IA est littéralement en train de révolutionner la finance. Nous allons entrer en détails sur les études, les recherches et les applications de l'IA sur la finance dans ce chapitre.

1.1.3.4 Voitures autonomes

Les voitures autonomes seront l'une des grandes révolutions du XXI siècle. Et l'IA va beaucoup nous aider dans ce processus. Ici, la vision par ordinateur va être beaucoup utilisée pour déplacer d'un point A à un point B et bien sûr la voiture doit être capable de reconnaître les objets qu'elle va rencontrer et faire le bon choix, il y a aussi le Reinforcement Learning.

1.1.3.5 Les Chatbots

Nous pouvons les considérer comme étant une interface de conversation entre homme et machine. L'Homme pose une question et la machine répond, ou l'inverse. Mais maintenant, avec tout ce qu'il y a comme Chatbot, nous pouvons penser à tort que tout Chatbot est intelligent. Nous pouvons

bien créer un programme qui répond à dix (10) questions, et à la onzième de dire je ne sais pas. L'IA va permettre de comprendre les questions et générer des réponses.

1.2 Etat de la recherche sur l'IA appliquée à la finance

Contrairement à ce que certaine personne pourrait penser l'IA n'est pas une science nouvelle, beaucoup de recherches ont été faites sur le sujet depuis la deuxième moitié du XX siècle. Le but des recherches était de reproduire la mémoire humaine à travers les mathématiques.

1.2.1 Les recherche sur l'IA

1.2.1.1 Genèse de l'IA : le premier neurone artificiel

Pour beaucoup de chercheurs dans le domaine de l'IA, la naissance de notre science pourrait remontée en 1943 avec le premier article publié dans le domaine.

Durant cette année, une étude menée par Warren S. McCulloch & Walter Pitts pour une expérimentation mathématique du neurone biologique a vu le jour. L'objectif de cette recherche était de mettre en évidence la simulation du fonctionnement du neurone biologique avec de l'analyse mathématique (*Calculus*). La méthodologie suivie par ses deux scientifiques était, pour chaque étape de la transmission d'information d'un neurone a un autre, de trouver une fonction mathématique qui pourra le répliquée. A la fin, ils ont pu trouver les calculs nécessaires pour reproduire un tant soit peu le fonctionnement du neurone biologique. Les implications de cette étude sont énormes, car elles sont à la base de tous les réseaux de neurones que nous utilisons aujourd'hui du perceptron au ANN (*Artificial Neural Network*).

Cette étude, l'une des toutes premières dans notre domaine va avoir un impact considérable sur les futurs réseaux de neurones. Cependant elle n'est pas exhaustive, plus tard Donald O. Hebb crée l'apprentissage pour les réseaux de neurone, après Frank Rosenblatt va créer le perceptron en s'appuyant sur les travaux de ses prédécesseurs et ainsi de suite.

1.2.1.2 L'IA dans la prédiction

Après les découverts de tous ces gens cités, l'IA a beaucoup évoluée, elle n'est plus seulement à ses balbutiements, mais règle maintenant de vrai problème dans la vie des humains, l'une des domaines auquel elle excelle c'est la prédiction et nous allons voir des études faites dans ce domaine d'application de l'IA.

D'abord en 2019 un article sur la prédiction basée sur l'IA a été rédigé par Ajay Agrawal, Joshua S. Gans et Avi Goldfarb dans le ScienceDirect. Il avait pour objectif de démontrer l'impact qu'aura la prédiction dans la vie de tous les jours des humains mais aussi le jugement de ces derniers. Les étapes précises qu'ils ont suivies étaient d'abord de voir l'impact des prédictions des machines sur le jugement des hommes et aussi la différence entre ses prédictions et ces jugements. Ce qui a découlé de cette étude c'est le jugement des hommes est toujours importants dans la prédiction car la prédiction est faite pour les hommes. Les implications que cette étude peut avoir c'est une considération plus importante de l'avis de l'homme dans la prédiction des machines.

La prédiction peut ainsi être appliquée dans beaucoup de domaine. Dans ce sillage, une étude menée par Gary S Collins et Karel G M Moons en 2019 qui s'intitule "*Reporting of artificial intelligence prediction models*" a été publiée par *The Lancet*. Le but principal de cette étude était d'expliquer la pertinence et l'important de la prédiction par l'IA surtout dans le domaine médical. Dans cet article, les auteurs ont adaptée une démarche théorique en expliquant la manière dont les IA peuvent significativement aider la médecine dans son fonctionnement. Ils ont pu trouver un certain nombre de modèles que l'on peut appliquer dans le domaine de la médecine.

Il y a aussi d'autres domaines sur lesquels la prédiction à base d'IA peut nous être utile : la prédiction des prix des actions dans les marchés financiers. Ainsi quatre chercheurs bangladais (Alam Miah, Zakir Hossain, Amjad Hossain, Muzahidul Islam) ont copublié un article en 2015 sur la prédiction des prix des actions en utilisant un modèle hybride (ANN et *Fuzzy Inference System*,). Ces scientifiques voulaient faire voir la pertinence d'utiliser les modèles hybrides pour la prédiction des valeurs financiers dans le marché bangladais. Ils ont pu obtenir des résultats plus satisfaisants puisque la variation entre leurs valeurs prédites et les valeurs réelles n'est pas vraiment importante.

1.2.1.3 Les Chatbots

Un Chatbot est un programme informatique développé pour être une interface de conversation entre un homme et une machine. Avec tout l'engouement autour des Chatbots aujourd'hui nous pouvons être tenté d'oublier ses origines. Le premier de la sorte était ELIZE (1965) qui est l'ancêtre des LLM (*Large Language Model*) aujourd'hui. Voyons quelques études sur les Chatbots.

Dans une récente étude de 2020 menée par Eleni Adamopoulou et Lefteris Moussiades, ils nous donnent une vue d'ensemble sur la technologie des Chatbots. L'objectif de cette étude est simple, c'est de tout simplement présenter le fonctionnement des Chatbots à un novice. Ce qu'ils ont réussi car ils ont abordé la chose dans tout son ensemble. D'abord ils ont donné l'historique, expliquer les concepts clés, parler des différents types de Chatbots et beaucoup d'autres choses.

Ce travail de recherche dont nous venons de parler nous donne une vue d'ensemble mais il y a d'autres étude qui parle des Chatbots de manière plus spécifique. C'est l'exemple de cet article de Takuma Okudo et Sanae Shoda sur les services d'un Chatbot sur l'industrie financière. Ils s'étaient donnés comme objectif de développer un Chatbot pour les finances d'une industrie, en l'occurrence Sony Bank. La méthodologie suivie était d'abord de rappeler historique des Chatbot puis de constater ce qui se fait déjà dans les entreprises. Après cela ils vont voir le meilleur d'entre eux, qu'ils ont par la suite adapter pour Sony Bank. Les ramifications de cette étude sont qu'il est possible de voir reproduire des modèles qui existe déjà et de les adapter dans une nouvelle industrie (japonaise ici).

Hors du monde académique, nous pouvons dire même que c'est plus intéressant pour les Chatbots car certaines entreprises abattent un travail dans le domaine. Aujourd'hui en 2024, il est très rare de voir des professionnels de tous les corps de métiers qui n'ont jamais utilisé de Chatbot à base de LLM. Des entreprises comme Open AI, Google ou encore Microsoft sont à l'origine de ce succès des LLM qui nous facilitent la vie. Il y a l'exemple de [Chat GPT](#) (Open AI), [Gemini](#) (Google), [Copilot](#) (Microsoft).

Ces derniers sont des Chatbots "*general purpose*" qui ont été développer pour répondre sur plusieurs domaines de connaissance. Mais en revanche, il y a des Chatbot spécialisé dans un domaine, comme la finance par exemple. Parlant de finance, il y a un certain nombre de Chatbots qui ont été conçu seulement pour répondre aux questions d'ordre financiers. C'est notamment le cas de [kakisto](#), [haptik](#), [kore.ai](#) etc.

1.2.2 L'IA appliquée à la finance

Maintenant, analysons ensemble l'application de l'IA sur la finance, ce qui nous intéresse vraiment dans ce document. La rencontre entre l'IA et la finance est naturelle puisque les deux travaillent sur des données. C'est l'un des raisons pour lesquelles beaucoup de recherche sont en train d'être faites sur les deux domaines.

1.2.2.1 La détection de fraude

La détection de fraude est un procédé qui nous permet d'identifier des activités frauduleuses exécutées ou tentées au sein d'une organisation. Les fraude, si ne sont pas identifiées et régler peuvent être un cauchemar pour les organisations. C'est là que l'IA peut intervenir avec ses algorithmes.

Pour mettre en évidence l'aide non négligeable que l'IA peut apporter dans la détection de fraude Muhammad Farman, Muzamil Abbas ont très récemment publié un article sur le sujet "*Artificial Intelligence for fraud detection and prevention*". L'objet de cette étude était de révéler l'apport de l'IA sur la détection de fraude mais aussi sur la prévention. Pour ce faire, ils ont d'abord présenté les algorithmes d'IA et leur fonctionnement spécifique dans la détection de fraude ensuite ils ont parlé de l'IA dans la finance. Les résultats de cette étude ont été conséquents car ils ont pu prouver l'apport considérable de l'IA sur ce domaine sans rien implémenter.

Un peu plus tôt, en 2018 deux chercheurs (Dahee Choi, Kyungho Lee) avaient fait une étude similaire mais plus pratique sur la détection de fraude. Ici, l'étude se voulait non seulement théorique mais aussi pratique en implémentant des modèles qui vont fonctionner pour la détection. Pour ce qui est des modèles ils ont utilisé dans un premier temps des algorithmes de Machine Learning, ensuite de Deep Learning. Les observations faites à la suite des expérimentations sont claires : les modèles de ML ont été plus performant que les modèles DL, comme quoi des fois le simple est plus efficace.

1.2.2.2 Service client

Est-ce que les IA vont remplacer les humains dans les services des entreprises ? Voilà l'un des questions qui se posent le plus dans le domaine. Les services clients ou services après-vente sont très coûteux pour les organisations et demande des ressources humaines et financières. Or l'IA peut aider à améliorer cela d'une manière plus qu'efficace et nous allons voir comment.

Dans un article de 2022 Mengmeng Song, Xinyu Xing, Yucong Duan, Jason Cohen et Jian Mou ont essayé d'apporter des réponses à ces questionnements. Ils nous parlent dans leur article des impacts que l'IA pourrait avoir dans les services clients. Ils ont fait différentes expérimentations, des tests, des comparaisons et des hypothèses pour aboutir aux résultats. Et ses résultats étaient

que les technologies d'IA allaient sans doute prendre une grande place dans les services clients mais il ne faut pas complètement supprimer l'interaction humaines.

Dans le même cadre d'autres études ont été faite sur le domaine comme celle de 2022 de Dimitrios Buhalis et Iuliia Moldavska dans le *Journal of Hospitality and Tourism Technology*. Cet article se donner comme objectif d'investiguer dans les interactions entre les hôtels et ôtes dans le contexte de « l'hospitalité ». La méthodologie que cette étude a suivie de faire des entretiens avec le personnel des hôtels sur l'utilisation des assistantes vocales dans ses structure. Ce qu'ils ont trouvé c'est l'utilisation de ses technologies sont d'un grand apport autant pour eux personnel que pour les clients. Selon les auteurs cette peut les permettre d'explorer domaines du "smart hospitality" et de l'écosystème du tourisme avec l'IA.

1.2.2.3 Analyse financière

Dans notre travail, nous nous sommes donnés comme objectifs de faire appliquer l'IA dans l'analyse financière des entreprises ouest-africaines. Nous allons voir que d'autres choses ont été faites ailleurs qui peuvent ressembler à ce que nous allons faire ici, nous allons les voir.

Ce sujet est tellement d'actualité que des articles très intéressants sont en train d'être rédigée et publiés au moment où j'écris ses lignes. C'est le cas de cet article de Ewerton Alex Avelar et Ricardo Vinícius Dias Jordão publié en juillet 2024. Cet article a l'objectif d'analyser la performance de différents algorithmes d'IA dans la prédiction des mouvements des plus grands marché financiers du monde. L'approche qu'ils ont adaptée c'est de tester différents algorithmes sur des données empiriques, et pour cela ils vont utiliser neuf (9) indicateurs pour y parvenir. Le résultat qu'ils ont trouvé c'est les modèles d'IA sont plus performants que les techniques utilisées par les analystes et aussi de tous les algorithmes le *random forest* est le plus efficace. Les enseignements que l'on peut tirer de cette étude selon les auteurs c'est que l'IA a un vrai rôle dans l'analyse financière et la prise de décision et les managers doivent prendre cela en considération.

D'autres recherches ont aussi été faites sur le domaine comme cette étude de 2000 de Ning Yang sur l'IA et de Data Mining dans le *Financial Big Data*. L'objet de cette recherche était de développer des modèles basés sur le Data Mining en entrant dans les détails du fonctionnement des algorithmes sur des données financières de très grandes tailles. Il a suivi la démarche suivante : d'abord expliquer le fonctionnement des modèles puis de les appliquer sur des données empiriques. Il a réussi à obtenir des résultats satisfaisants puisque l'application parvenait à

répondre dans des tâches importantes dans la finance. Avec le développement rapide des données financières, l'application de ces modèles vont être plus que nécessaire.

1.2.3 Les limites sur les travaux actuels

Dans tous les travaux de recherche, il y a toujours des limites, des défauts ou des points d'améliorations. C'est d'ailleurs pourquoi les chercheurs continuent toujours de faire des recherches. Dans le domaine de l'IA appliquée à la finance, il y a un certain nombre d'endroit où il y a lieu d'amélioration pour parfaire la science et nous allons les voir.

1.2.3.1 Qualités des données financières

Quand on parle de qualité des données, on peut faire référence sur plusieurs aspects des données financières. Pour les données financières nous faisons souvent face à des bruits et aussi des valeurs incomplètes. La conséquence de cette chose c'est qu'il falloir faire un grand travail de *preprocessing*. En outre, les données historiques peuvent être biaisées, la finance change et les méthodes aussi changent, les politiques de finance aussi. Ce qui peut constituer un problème dans l'entraînement. Pour illustrer cela, en 2018 dans la zone UEMOA une nouvelle réglementation a vu le jour, en imposant à toutes les entreprises de la BRVM de présenter leurs états financiers avec les normes IFRS (International Financial Reporting Standards). Mais cela c'est pour les grandes entreprises, pour les petites entreprises c'est encore pire puisque qu'il est très difficile d'accéder à ses données, c'est l'une des raisons pour lesquelles les petites et moyennes entreprises ne bénéficient pas de ses technologies pour le moment. En fin nous pouvons parler du contexte, des données peuvent changer selon contexte géopolitique, économique, sociale etc. ce qui peut biaiser les modèles en prenant en compte des périodes avec des mesures exceptionnelles (la crise des subprimes en 2008 ou le covid 19 en 2020 par exemple).

1.2.3.2 L'éthique

Pour ce qui est de l'éthique des IA dans le contexte financier, elle peut être abordée sur plusieurs angles. Nous allons parler des manières de développement des modèles, le comportement de modèles vont être toujours basé sur les données d'entraînement. Ce qui veut dire que les développeurs peuvent intentionnellement programmer l'IA à se comporter d'une manière ou d'une autre. Mais aussi l'IA ne va faire seulement la chose pour laquelle elle a été développée, à moins de mise à jour, elle ne pourra pas contextualiser pour prendre la meilleure décision. De plus, il y a un autre problème qui subsiste, une IA qui a été programmée d'une manière, donc elle va se

comporter et fonctionner de cette même manière. Pour être plus claire, il y a un risque que l'IA donne le même conseil, si elle le fait et que tout le monde (ou la majorité des gens) suit ses conseils, les gens peuvent prendre les mêmes décisions, ce qui n'est pas bien. Imaginons juste un marché avec seulement des vendeurs mais pas d'acheteurs. Nous pouvons aussi nous questionner sur la confidentialité et la sécurité des données, les délits d'initié, l'IA pourra-t-elle faire preuve de discernement et savoir quelle information donner à qui.

On peut répondre à tous ces questionnements en disant que l'IA est un programme, il est possible de programmer cela dans son développement (chose qui ne va pas être facile). Par exemple il y a des LLM qui ne vont pas répondre sur certains sujets qui fâchent comme la religion, la race, les ethnies etc.

1.2.3.3 L'insuffisance des études dans le contexte africain

Personnellement, la limite qui nous interpelle le plus est l'absence de recherches qui se focalisent sur l'Afrique. Et même pour être plus précis, dans la zone UEMOA, il n'y a pas assez d'études approfondies sur le domaine de la finance des entreprises ouest-africains. Quand nous savons comment certaines entreprises africaines fonctionnent avec toutes leurs difficultés, l'IA leur sera d'une très grande utilité.

C'est d'ailleurs l'une des grandes raisons pour lesquelles nous nous sommes lancés dans ce domaine en plus de notre grand intérêt pour l'IA. Il n'est pas seulement intéressant de constater les dégâts et les manquements, il faut aussi proposer des solutions avec les moyens dont nous disposons.

Conclusion partielle

Faire une revue de la littérature n'a pas été chose aisée car il ne faut rien inventer, il va falloir toujours relater ce que les chercheurs ont déjà trouvé et le mettre en corrélation avec notre sujet. Dans cette revue de la littérature que nous avons proposée pour notre sujet, il a été question de couper le sujet en deux (2) mots clés, en premier lieu pour l'IA, en deuxième lieu pour la finance, et terminer avec une combinaison des deux. Les enseignements de ce travail ont été énormes. C'est deux domaines vraiment très intéressants. L'IA est un domaine en plein essor et il y a tellement de choses qui ont été faites, qui sont en train d'être faites et qui vont être faites. De même que la finance qui est très importante pour les entreprises.

2 Chapitre II : L'analyse financière

Introduction de chapitre

Dans l'environnement ultra concurrentiel dans lequel nous évoluons, chaque entité doit savoir tirer son épingle du jeu. Et le monde des affaires ne fait pas exception. Une entreprise est une personne morale dont le but ultime et fièrement assumé est de faire du profit. Faire du profit veut dire augmenter le plus ses bénéfices pour faire face à ses charges.

Maintenant, il y a tout un processus à suivre qui va nous prendre de la construction du chiffre d'affaires jusqu'aux cash-flows en passant par le résultat net. Tout entreprise qui se respecte doit être en mesure de comprendre ce processus. Justement, comment faire pour comprendre ce processus ? C'est là qu'entre en jeu l'analyse financière.

Tout d'abord, il faut dire que les entreprises ont manifesté le besoin de comprendre les éléments qui composent leurs états financiers afin de prendre les meilleures décisions dans le futur. Non seulement les entreprises mais aussi les investisseurs ont besoin d'une analyse financière avant d'investir dans quelque structure.

Ce chapitre va être divisé en deux grandes parties, premièrement nous allons parler des états financiers : ce sont les outils de l'analyste financier pour faire son travail convenablement. Deuxièmement, nous allons parler de fond en comble de l'analyse financière.

2.1 Présentation des états financiers

Les états financiers d'une entreprise sont des documents comptables qui permettent de juger de la santé d'une entreprise à une période donnée (souvent le 31 décembre de l'année N). Ces derniers sont utilisés par les entreprises pour faire un résumé des activités de l'exercice, pour faire l'analyse financière. Ils sont aussi utilisés par les investisseurs pour prévoir le niveau de risque d'investir dans une entreprise. Ces états financiers sont le bilan, le compte de résultat et le tableau des flux de trésorerie que nous allons détailler dans cette partie.

2.1.1 Le bilan

Il est souvent entendu que le bilan est la photographie d'une entreprise à un instant T. Ce document comptable révèle les actifs et les passifs d'une entreprise. Nous allons voir que pour obtenir ces actifs et ces passifs, il va falloir faire un certain nombre de calculs et à la fin ils vont s'égaliser.

2.1.1.1 Les actifs

Tout d'abord les actifs, ils sont le patrimoine de l'entreprise c'est-à-dire ce que l'entreprise possède durant l'exercice. Ses actifs peuvent être de plusieurs natures. L'actifs total de l'entreprise se calcul en faisant la somme des éléments qui le composent :

$$\text{Actif total} = \text{Actifs immobilisés} + \text{Actifs circulants} + \text{Tresorerie actifs}$$

Équation 1 Actif total

- Les actifs immobilisés

Ils représentent les immobilisations de l'entreprise c'est-à-dire les biens à long terme, ces biens sont destinés à rester de manière durable dans l'entreprise.

Le calcul des actifs immobilisés se fait en faisant la somme des éléments qui le composent.

Actifs Immobilises

$$\begin{aligned} &= \text{charges immobilisées} + \text{immobilisations incorporelles} \\ &+ \text{immobilisations corporelles} + \text{immobilisations financières} \\ &+ \text{amortissements et provisions} \end{aligned}$$

Équation 2 Actif immobilises

- L'actif circulant

Par opposition aux actifs immobilisés, les actifs circulants n'ont pas pour vocation de rester durablement dans l'entreprise. Justement, ces types d'actifs sont utilisées par l'entreprise seulement durant un exercice c'est pour une année N.

Les calculs des actifs circulants peuvent se faire en faisant la sommation de tous les éléments qui le composent.

$$\text{Actifs circulant} = \text{stock} + \text{fournisseurs, avances versées} + \text{clients} + \text{Autres créances}$$

Équation 3 L'actif circulant

- La trésorerie-actif

La trésorerie-actif correspond aux avoirs en liquidité d'une société dite disponible. La trésorerie-actif telle qu'elle figure sur le bilan représente l'argent qu'une entité a dans son compte banque, caisse, établissements financiers et assimilés, instruments de trésorerie, instruments de monnaie

électronique, régie d'avance et accréditifs (disponibilités) et tout ce qui peut être facilement transformé en trésorerie à une brèche échéance (les équivalents de trésorerie). (Landu, 2021)

2.1.1.2 Les passifs

Pour ce qui est des passifs, ce sont les ressources financières que l'entreprise à utilisées pour s'octroyer ses actifs. Tout le patrimoine actif que l'entreprise possèdent durant l'année N est financés par ses passifs.

Le total des passifs se calcul comme suit :

$$\text{Passif total} = \text{Capitaux propres} + \text{dettes financières} + \text{passif circulant} \\ + \text{tresorerie passifs}$$

Équation 4 Le passif total

- Les capitaux propres

Les capitaux propres aussi appelés fonds propres sont des ressources possédées par une entreprise (hors dettes). C'est l'ensemble des éléments financiers que possèdent vraiment l'entreprise.

Pour calculer les capitaux il nous faut faire la sommation des éléments qui le compose.

Capitaux propres

$$= \text{capital social} + \text{primes et réserves} + \text{report à nouveau} + \text{résultat ne} \\ + \text{provisions règlementées}$$

Équation 5 Les capitaux propres

- Les dettes financières

Une dette dans sa définition la plus simple est tout simplement une somme qu'une entité A (emprunteur) doit payer à une entité B (prêteur). Dans notre cas l'emprunteur c'est l'entreprise, et le prêteur peut être n'importe quelle personne (physique ou morale) qui peut donner de l'argent à l'entreprise.

Dettes financières

$$= \text{crédits bancaires} + \text{dettes fournisseur} + \text{provisions financières} \\ + \text{autres dettes financières}$$

Équation 6 Les dettes financières

Une dette permet à une entreprise d'augmenter ses fonds, faire face à certaines dépenses qui sans la dette allait être difficile pour l'entreprise de gérer. La dette n'est pas une mauvaise chose, si elle est bien gérée et utilisé pour apporter plus de revenus à l'entreprise.

- Le passif circulant

Par opposition aux actifs circulants qui représentent les avoirs de l'entreprise à court terme, le passif circulant représente l'ensemble des sommes que l'entreprise doit rembourser sur une échéance de moins d'un an.

passif circulant

= dettes circulantes + clients, avances reçues + fournisseurs d'exploitation
+ dettes fiscales + Dettes sociales + autres dettes

Équation 7 Le passif circulant

- La trésorerie-passif

Selon le site AGICA, on distingue la trésorerie active et la trésorerie passive. Cette dernière regroupe l'ensemble des dettes professionnelles et crédits à court terme. Elle comprend :

- Les soldes créditeurs de banque
- Les concours binares
- Les découverts autorisés et les facilités de caisse
- Les échéances de prêts (moins d'un an)

L'entreprise doit savoir maîtriser la trésorerie passive au risque de mauvaise surprise comme un manque de liquidités ou l'incapacité à faire face à ses dettes.

2.1.1.3 Exemple de bilan

Tableau 1 Exemple de bilan

BILAN		
	N	N-1
ACTIFS	-	-
Charges immobilisées	-	-
Immobilisations incorporelles	-	-
Immobilisations corporelles brutes	-	-
Immobilisations corporelles brutes	-	-
Immobilisations financières	-	-
Amortissements et provisions	-	-
TOTAL ACTIF IMMOBILISE	-	-
Stock	-	-
Fournisseurs, avances versées	-	-
Clients	-	-
Autres créances	-	-
TOTAL ACTIF CIRCULANT	-	-
TOTAL TRESORERIE ACTIF	-	-
TOTAL ACTIF	-	-
PASSIF		
Capital	-	-
Primes et réserves	-	-
Report à nouveau	-	-
Résultat net	-	-
TOTAL CAPITAUX PROPRES	-	-
Emprunts et dettes financières	-	-
Provisions financières	-	-
TOTAL DETTES FINANCIERES	-	-
Dettes circulants	-	-
Clients, avances reçues	-	-
Fournisseurs d'exploitation	-	-
Dettes fiscales	-	-
Dettes sociales	-	-
Autres dettes	-	-
TOTAL PASSIF CIRCULANT	-	-
TOTAL TREORERIE PASSIF	-	-
TOTAL PASSIF	-	-

2.1.2 Le compte de résultat

Le compte de résultat est un document comme le bilan qui permet de juger de la santé des finances d'une entreprise. Cet état financier retrace les charges d'une entreprise et les éléments qui le composent, les produits et les éléments qui le composent et éventuellement les soldes intermédiaires de gestion.

2.1.2.1 Les charges

Les charges d'une entreprise durant un cycle d'exploitation représentent les éléments qui s'imposent à l'entreprise pour produire. Pour qu'une entreprise produise ses biens ou services, elle a besoin des produits de base pour le faire. Les charges sont composées des achats, de la variation de stock, du transport, des assurances, autres charges...

Le calcul des charges varie selon le type d'entreprise, car toutes les entreprises ne font pas les mêmes choses. Mais nous pouvons donner l'exemple d'une entreprise de transformation :

$$\begin{aligned} \text{Total charge} = & \text{Achat marchandise (+/-)variation de stock marchandise} \\ & + \text{Achat matiere premiere (+/-)variation de stock matiere premiere} \\ & + \text{autres charges} \end{aligned}$$

Équation 8 Les charges

Les charges sont très importantes, même s'il faut tout faire pour les réduire au minimum, car elles nous permettent de produire les biens ou services.

2.1.2.2 Les produits

Les produits représentent quant à eux tous ce que l'entreprise a réalisé dans son cycle d'exploitation. C'est-à-dire tout ce que l'entreprise a produit comme biens ou services, ces productions peuvent être vendues ou stockées. C'est pourquoi le CA de l'entreprise entre dans le calcul des produits. Les produits d'une entreprise sont notamment composés des ventes de marchandises, des produits accessoires, des productions stockées, des productions immobilisées, d'autres produits.

La formule pour calculer les produits est la suivante :

$$\text{total produit} = CA + \text{productions stockees} + \text{production immobiliées} \\ + \text{subvention d'exploitation} + \text{autres produits}$$

Équation 9 Les produits

Les produits constituent la principale source de revenue pour les entreprises, il faut bien les maîtriser pour ne pas être en surplus ou en déficit. Les entreprises doivent répondre de manière optimale aux demandes de ses clients.

2.1.2.3 Les soldes intermédiaires de gestion (SIG)

Les SIG sont des éléments calculés à partir du compte de résultat et qui fournissent des éclaircissements sur les performances de l'entreprise. Nous allons les présenter dans cette partie avec leur calcul et interprétation. Les SIG entrent directement dans l'analyse financière, en l'occurrence l'analyse du compte de résultat, car entreprise comme investisseurs les utilisent pour se faire une idée de la situation de l'entreprise.

- La marge commerciale

La marge commerciale ou encore la marge brute est un indicateur de rentabilité de l'entreprise. Elle exprime la différence entre le coût d'achat hors taxes des produits et le prix de vente hors taxes des achats faites par l'entreprise.

Calcul

marge commerciale

$$= \text{ventes de marchandise} - \text{Achats de marchandises} + \text{variation de stocks}$$

Équation 10 La marge commerciale

- Chiffre d'affaires (CA)

D'une manière simple, le chiffre d'affaires représente le montant des ventes, c'est la somme des quantités vendues que l'on multiplie par le prix de vente unitaire. Même avec cela, il y a plusieurs éléments qui entrent en considération dans le CA. Même si ce n'est pas le meilleur des indicateurs, le CA en est le plus populaire car permettant à l'entreprise de se montrer sur son meilleur jour avec ce nombre gigantesque, le plus grand de tous les indicateurs. En outre, il est souvent utilisé pour calculer la croissance d'une entreprise.

Calculs

$$CA = Quantites\ vendues * prix$$

Équation 11 Chiffre d'affaires 1

$$CA = vente\ de\ marchandises + ventes\ produits\ fabriques \\ + travaux\ et\ services\ vendus + produits\ accessoires$$

Équation 12 Chiffre d'affaires 2

- La valeur ajoutée (VA)

La valeur ajoutée est un indicateur qui montre la richesse brute créée par une entreprise dans un cycle d'exploitation d'un exercice. La VA ne prend qu'en compte les produits et les charges d'exploitation. C'est vraiment un indicateur qui montre la création de valeurs seulement sur l'activité de l'entreprise.

Calcul

$$VA = total\ produits - total\ charges$$

Équation 13 La valeur ajoutée

- L'excédent brut d'exploitation (EBE)

L'excédent brut d'exploitation est l'un des indicateurs financiers les plus importants du compte de résultat. Il offre une visibilité sur la rentabilité économique de l'entreprise grâce à l'activité générée sur une période définie. Il constitue l'une des soldes intermédiaires de gestion capitale pour analyser la performance d'un cycle d'exploitation. Selon le site web [Libeo](#).

Calcul

$$EBE = valeurs\ ajoutée - charges\ personnel$$

Équation 14 L'excédent brut d'exploitation

- Le résultat d'exploitation (REX)

Le résultat d'exploitation est un autre indicateur permettant d'apprécier le modèle économique de l'entreprise. Le REX ne prend pas en compte les éléments exceptionnels, ni financiers, c'est pour

cela que ce résultat est dit d'exploitation. Par contre il va prendre en considération les amortissements, dépréciations et provisions : c'est ce qui le différencie de l'EBE.

Calcul

$$REX = EBE - \text{dotations aux amortissements, provisions et depreciations} \\ + \text{reprises amortissements, provisions et depreciations}$$

Équation 15 Le résultat d'exploitation

- Le résultat financier (RF)

Le résultat financier est un solde intermédiaire de gestion qui représente la différence des produits financiers et des charges financières. Il peut être intéressant dans la mesure où il peut faire gonfler le résultat net d'une entreprise sans faire partie de son cycle d'exploitation.

Calcul

$$RF = \text{produits financiers} - \text{charges financières}$$

Équation 16 Le résultat financier

- Le résultat exceptionnel (REP)

Le résultat exceptionnel est un solde intermédiaire de gestion qui représente la différence des produits hors activité ordinaire (HAO) et des charges HAO. Il peut être intéressant comme le résultat financier, il peut faire gonfler le résultat net d'une entreprise sans faire partie de son cycle d'exploitation.

Calcul

$$REP = \text{produits HAO} - \text{charges HAO}$$

Équation 17 Le résultat exceptionnel

- Le résultat net (RN)

Le résultat net est un indicateur qui figure non seulement sur le compte de résultat, mais aussi dans le bilan, dans les capitaux propres. Il représente la somme réellement gagnée par l'entreprise. Le résultat net prend en compte tous les autres indicateurs dans son calcul de manière directe ou

indirecte. Le résultat net sert notamment aux actionnaires car c'est cette que les actionnaires de l'entreprise vont se partager.

Calcul

$$RN = REX + \text{resultat financier} + \text{resultat exceptionnel} \\ - \text{paricipation des travailleurs} - \text{impôts sur la societe}$$

Équation 18 Le resulta net

2.1.2.4 Exemple de compte de résultat

Tableau 2 Exemple de compte de résultat

COMPTE DE RESULTAT		
	N	N-1
Ventes de marchandises	-	-
Production vendue	-	-
Travaux, services vendus	-	-
Produits accessoires	-	-
CHIFFRE D'AFFAIRES	-	-
Production stockée	-	-
Autres produits	-	-
TOTAL PRODUITS	-	-
Achats et frais sur achats	-	-
Variation de stock	-	-
Transports	-	-
Services extérieurs	-	-
Impôts et taxes	-	-
Charges et pertes diverses	-	-
TOTAL CHARGES	-	-
VALEUR AJOUTEE	-	-
Frais de personnel	-	-
EXCEDENT BRUTE D'EXPLOITATION	-	-
Transfert de charges	-	-
Dotations aux amortissements	-	-
Dotations aux provisions	-	-
RESULTAT D'EXPLOITATION	-	-
Produits financiers	-	-
Charges financières	-	-
RESULTAT FINANCIER	-	-
Produits HAO	-	-

<i>Charges HAO</i>	-	-
RESULTAT EXCEPTIONNEL	-	-
<i>Impôt sur la société</i>	-	-
RESULTAT NET	-	-

2.1.3 Les flux de trésorerie (TFT)

Selon le site [L'expert-comptable](https://www.expert-comptable.fr/), le tableau de flux de trésorerie est un outil financier qui permet de déterminer la rentabilité d'un projet, d'évaluer le besoin en fonds de roulement et d'anticiper ses besoins en fonds propres. Il indique les entrées et sorties d'argent de l'entreprise sur une période.

Ce TFT va particulièrement intéresser les investisseurs c'est là que réside le vrai gain auquel ils sont voués à recevoir. Cet état financier n'est pas comme les autres, ses éléments ne proviennent pas directement de l'activité de l'entreprise. Il nous faut faire un certain nombre de calculs. D'ailleurs pour établir un TFT, nous aurons besoin du bilan et du compte de résultat.

2.1.3.1 La trésorerie initiale (au 1 janvier de l'année N)

Avant de parler de la trésorerie initiale, il va falloir définir ce que c'est la trésorerie. Cette dernière est la somme d'argent effectivement disponible sur les comptes bancaires de l'entreprise à un instant T.

Maintenant pour ce qui s'agit de la trésorerie initiale, elle représente tout simplement la différence entre la trésorerie actif de l'année N-1 et sa trésorerie passif qui se trouve dans le bilan.

$$Tresorerie\ initiale = Tresorerie\ actif\ (N - 1) - Tresorerie\ passif\ (N - 1)$$

Équation 19 La trésorerie initiale

2.1.3.2 Les flux de trésorerie liées aux activités opérationnels (FTAO)

Ils représentent les flux de trésorerie provenant de l'activité d'exploitation de l'entreprise. C'est les ressources générées par l'entreprise grâce aux achats et ventes.

Calcul

$$\begin{aligned} FTAO = & \text{capacité d'autofinancement} - \text{variation actif circulant HOA} \\ & - \text{variation des stocks} - \text{variation des créances} \\ & + \text{variation du passif circulant} \end{aligned}$$

Équation 20 Les flux de trésorerie liées aux activités opérationnels

Variation BFR

$$= \text{variation actif circulant } HO - VS - \text{variation des créances} \\ + \text{variation du passif circulant}$$

Équation 21 Variation du BFR

2.1.3.3 Les flux de trésorerie liés aux activités d'investissements (FTAI)

Ils représentent quant à eux les flux de trésorerie provenant de la richesse créée par l'entreprise à travers ses activités d'investissement, en l'occurrence les acquisitions et cession d'actifs.

Les flux de trésorerie liés aux activités d'investissements sont composés des décaissements encaissements.

Calcul

FTAI

$$= \text{décaissements (liés aux acquisitions d'immobilisation incorporelles, corporelles et financières)} \\ - \text{encaissements (liés aux acquisitions d'immobilisation incorporelles, corporelles et financières)}$$

Équation 22 Les flux de trésorerie liés aux activités d'investissements

2.1.3.4 Les flux de trésorerie liés aux activités de financement (FTAF)

Ils représentent les flux de trésorerie issue des activités de financement de l'entreprise, ce sont les dettes, emprunts, subvention... Les flux liés aux activités de financement sont composés des flux de trésorerie provenant des capitaux propres (FTCP) et des flux de trésorerie provenant des capitaux étrangers (FTCE).

Calcul

$$\text{FTCE} = \text{augmentation du capital} + \text{subvention d'investissement} \\ - \text{prélèvements sur le capital} - \text{dividendes versés}$$

Équation 23 Les flux de trésorerie provenant des capitaux propres

$$\text{FTCE} = \text{emprunts} + \text{autres dettes financières} \\ - \text{remboursements des emprunts et autres dettes financières}$$

Équation 24 Les flux de trésorerie provenant des capitaux étrangers

$$FTAF = FTCP + FTCE$$

Équation 25 Les flux de trésorerie liées aux activités de financement

2.1.3.5 La trésorerie finale (au 31 décembre de l'année N)

Elle représente la somme d'argent effectivement présente dans les comptes de l'entreprise en fin d'exercice c'est-à-dire le 31 décembre de l'année N. Elle s'obtient en additionnant la trésorerie initiale et la variation de la trésorerie nette (VTN).

$$VTN = FTAO + FTAI = FATF$$

$$Trésorerie\ finale = Trésorerie\ initiale + VTN$$

Équation 26 La trésorerie finale

2.1.3.6 Exemple de tableau des flux de trésorerie

Tableau 3 Exemple de tableau des flux de trésorerie

TABLEAU DES FLUX DE TRESORERIE		
	N	N-1
TRESORERIE INITIALE	-	-
Capacite d'autofinancement	-	-
Variation actif circulant HAO	-	-
Variation des stocks	-	-
Variation des créances	-	-
Variation du passif circulant	-	-
Variation du BFR	-	-
FLUX DE TRESORERIE DES ACTIVITES OPERATIONNELLES	-	-
Décaissements lies aux acquisitions d'immobilisations incorporelles	-	-
Décaissements lies aux acquisitions d'immobilisations corporelles	-	-
Décaissements lies aux acquisitions d'immobilisations financières	-	-
Encaissements lies aux acquisitions d'immobilisations incorporelles et corporelles	-	-
Encaissements lies aux acquisitions d'immobilisations financières	-	-
FLUX DE TRESORERIE DES ACTIVITES D'INVESTISSEMENT	-	-
Augmentation du capital par apports nouveaux	-	-
Subvention d'exploitation	-	-
Prélèvement sur le capital	-	-
Dividendes verses	-	-
Flux de trésorerie provenant des capitaux propres	-	-
Emprunts	-	-
Autres dettes financières	-	-
Remboursement des emprunts et autres dettes financiers	-	-
Flux de trésorerie provenant des capitaux étrangers	-	-
FLUX DE TRESORERIE DES ACTIVITES DE FINANCEMENT	-	-
Variation de trésorerie nette	-	-
TRESORERIE FINALE	-	-

2.2 L'analyse financière proprement dite

2.2.1 Vérification des états financiers

Avant de pouvoir commencer l'analyse financière en question, il va falloir faire des vérifications pour ne pas avoir une analyse biaisée par certaines valeurs.

- L'actif et le passif : il faut vérifier que le total des actifs est égal au total des passifs
- Le résultat net : il faut vérifier que le résultat au compte de résultat est égal au résultat du bilan
- La trésorerie nette : il va falloir aussi vérifier la différence entre la trésorerie active et la trésorerie passive soit bien égale à la trésorerie nette au 31 décembre de l'année N.

Une fois ses vérifications faites et que tout est OK, on peut faire l'analyse financière proprement dite. Il faut rajouter qu'il y a d'autres types de vérification que l'on peut faire mais ses dernières sont les plus importantes.

2.2.2 Analyse des états financiers

Dans cette partie nous allons faire l'analyse financière des états financiers à savoir le compte de résultat, le bilan et le tableau des flux de trésorerie. Pour chaque état financier, il y aura deux types d'analyse : une analyse verticale et une analyse horizontale.

2.2.2.1 Analyse du compte de résultat

- Analyse verticale

L'analyse verticale du compte de résultat consiste à rapporter tous les éléments du compte de résultat au chiffre d'affaires. C'est-à-dire la part de chaque élément du compte de résultat dans le processus de création de richesse. Qu'allons-nous faire exactement ? nous allons tout simplement diviser chaque rubrique du compte de résultat par le chiffre d'affaires et nous obtiendrons notre tableau pour l'analyse verticale.

- Analyse horizontale

Pour ce qui est de l'analyse horizontale du compte de résultat, et d'ailleurs cela sera pareil pour le bilan et le tableau des flux de trésorerie, nous allons tout simplement calculer la croissance de chaque élément du compte de résultat.

$$Croissance(X) = \frac{X_n - X_{n-1}}{X_{n-1}}$$

Équation 27 Taux de croissance 1

$$Croissance(X) = \frac{X_n}{X_{n-1}} - 1$$

Équation 28 Taux de croissance 2

Ce calcul de croissance va nous permettre de voir l'évolution de chaque élément d'une année à la suivante.

2.2.2.2 Analyse du bilan

- Analyse verticale

Pour ce qui est de l'analyse verticale du bilan, elle va ressembler un tantinet à celle du compte de résultat. Sauf qu'ici, il y aura deux choses. Le premier c'est rapporter tous les éléments qui contribuent à l'actif au total actif, le second c'est de rapporter tous les éléments qui contribuent au passif au total passif.

De ce fait, nous nous retrouvons avec un tableau où nous allons voir apparaître le total actif et la part des éléments qui le composent et de même pour le passif.

- Analyse horizontale

L'analyse horizontale du bilan ne sera en rien différente de celle du compte de résultat comme on l'avait dit. Nous allons calculer la croissance de chaque élément, et plus haut la formule de la croissance est présentée.

2.2.2.3 Analyse du tableau des flux de trésorerie

- Analyse verticale

Pour ce qui est de l'analyse du tableau des flux de trésorerie, l'élément de base va être la trésorerie au 31 décembre, c'est la trésorerie finale. Et après on fait la même chose que les autres états financiers, ce qui veut dire calculer le rapport de tous les éléments sur la trésorerie finale.

- Analyse horizontale

Nous l'avons fait et refaits pour le compte de résultat et le bilan, et nous allons le refaire pour le tableau des flux de trésorerie, par conséquent calculer la croissance de chaque élément.

2.2.3 Analyse de l'activité et des relations de trésorerie

Ce point de l'analyse financière est la partie qui va nous permettre d'apprécier l'évolution des éléments clés d'une entreprise. C'est tous les chiffres que l'entreprise doit maximiser pour rester en bonne santé et durée dans le temps. Il y aura trois (3) d'analyse qui va être faite dans cette partie.

2.2.3.1 Analyse du cycle de vie de l'activité

Le cycle de vie de l'activité peut être jugée en analysant trois éléments fondamentaux des états financiers. A savoir, d'abord le chiffre d'affaires qui représente la création de richesses. Ensuite, le résultat net qui est ce qui reste pour l'entreprise après s'être acquitté de toutes ses charges. Enfin la trésorerie nette qui correspond à l'argent effectivement disponible (caisse ou banque).

Il faut noter que durant cette analyse on ne va pas calculer, l'analyse du cycle de vie va regarder l'évolution de ses valeurs et se faire une opinion du cycle d'activité de l'entreprise.

2.2.3.2 Analyse du comportement des flux de trésorerie

Comme fait précédemment, cette analyse va nous permettre de constater l'évolution des rubriques les plus importants du tableau des flux de trésorerie. Ici, l'analyse va voir et apprécier les flux de trésorerie liés aux activités opérationnelles, les flux de trésorerie liés aux activités d'investissement, les flux de trésorerie liés aux activités de financement, la variation de la trésorerie nette et enfin la trésorerie finale.

2.2.3.3 Analyse des équilibres financiers et la relation de trésorerie

Les équilibres financiers sont composés de trois sous-partie et chacune d'entre elles va relever un tableau qui va nous donner encore plus d'éclaircissement sur les finances de l'entreprise. Nous avons entre autres le bilan économique, les équilibres financiers, la relation de trésorerie.

- Bilan économique

Le bilan économique est un mini-bilan qui va faire ressortir les avoirs et dus de l'entreprise. Nous pouvons notamment y retrouver les actifs immobilisés, les capitaux, les endettements... Cependant, le rôle principal du bilan économique, c'est de calculer l'actif économique et les

capitaux investis. Il faut ajouter que ce mini-bilan est bien fait si l'actif économique est égal aux capitaux investis.

- Les équilibres financiers

L'équilibre n'est plus, ni moins que la trésorerie nette recalculée et vérifiée si c'est conforme à ce qui est aux valeurs présentées au niveau du tableau des flux de trésorerie. Cette trésorerie est obtenue en faisant la différence entre le fonds de roulement nette global et le besoin de financement global.

- La relation de trésorerie

Pour ce qui est de la relation de trésorerie, son but est d'apprécier la création de richesse, les différents besoins de l'entreprise liée à la trésorerie (fonds de roulement et financement) et l'argent qui est effectivement à la disposition de l'entreprise. Eu égard à cela, il est logique de voir apparaître dans son tableau le chiffres d'affaires, le besoin de fonds de roulement global, le besoin de financement global et la trésorerie nette.

2.2.4 Analyse tendancielle et la méthode des ratios

Tout d'abord, l'analyse financière avec la méthode des ratios nous permet de faire le diagnostic de l'entreprise à travers un certain nombre de ratios. Ces ratios sont de tout type et vont faire intervenir les trois (3) états financiers. Ici, nous sommes à un point très important de l'analyse financière c'est pour cela que cette partie va particulièrement intéresser les investisseurs en plus de l'entreprise elle-même.

Ceci étant dit, on appelle ratio le rapport entre deux valeurs par conséquent, nous allons calculer ses ratios, chacun dans sa famille (il a plusieurs familles de ratio comme développée dans la suite). Ensuite, s'en viendra l'interprétation qui va nous permettre de faire parler les chiffres.

Ces calculs de ratios vont nous permettre de noter à la fin l'entreprise de voir si elle en risque de défaut ou risque de faillite. Nous allons utiliser un système de notation développé par Altman en 2005. Finalement, l'évaluation de l'entreprise va mettre fin à cette analyse par les ratios.

2.2.4.1 Les familles de ratios

Si l'analyse a été bien faite, les bonnes décisions vont être prises. C'est là que réside tout l'intérêt de l'analyse financière.

- Les ratios de profitabilité

C'est la famille de ratio de profitabilité pour la plupart d'entre eux calculer à partir du compte de résultat. Cela nous permet de constater la part des SIG dans la création de richesse. C'est ainsi que nous allons tous voir leur part sur le CA.

- Les ratios de rentabilité

Comme son nom l'indique cette famille se veut être une mesure de la rentabilité. Il y en a deux types : la rentabilité financière et la rentabilité économique. C'est une famille qui suscite non seulement l'intérêt de l'entreprise et ses actionnaires mais aussi de potentiels investisseurs.

- Les ratios de politique comptable

La politique comptable est une famille de ratios qui nous permet d'observer l'état des immobilisations. Après calcul, on peut se rendre compte de comment sont amortis les immobilisations et quel politique l'entreprise adopte à cet égard.

- Les ratios de liquidité

La liquidité est l'aptitude qu'a l'entreprise à faire face à ses dépenses par la circulation optimale des flux de trésorerie. Cette famille de ratios va nous permettre de voir comment l'entreprise est liquide sous les trois ratios de liquidité.

- Les ratios de gestion de la dette

La famille de ratios gestion de la dette est très facile à comprendre, puisque son nom en dit tout. Ici, nous allons calculer quelques ratios qui vont nous permettre d'analyser comment endettée l'entreprise est et quelle politique de remboursement elle adapte.

- Les ratios de flux de trésorerie

La famille de ratios des flux de trésorerie nous informe principalement sur les stratégies d'investissement adoptées par l'entreprise et leur efficacité.

- Les ratios d'efficacité des actifs du BFG

C'est une famille qui regroupe les ratios sur l'efficacité du BFG. Nous allons jongler entre le compte de résultat et le bilan pour faire les calculs.

- Les ratios de valorisation

Nous allons terminer les familles de ratios par celle de la valorisation. Mais dans ce cas, nous allons nous concentrer sur la valorisation boursière. Nous serons en mesure d'observer comment l'entreprise se comporte sur les marchés financiers, la bourse régionale des valeurs mobilières (BRVM) en l'occurrence.

- L'analyse de la probabilité de défauts

L'analyse de probabilité de défaut, aussi appelé risque de défaut représente des méthodes pour l'analyste financier de prédire, en lisant des constantes de l'entreprise, si cette dernière va être en défaut, voir la faillite. Il est possible de prédire si une entreprise va être en défaut en utilisant différentes méthodes.

Comme déjà dit, plusieurs chercheurs et structures ont proposé des méthodes évaluation du risque de défauts d'une entreprise. Ces méthodes reposent sur des calculs mathématiques et des interprétations qui selon le cas de figure vont nous permettre de savoir si l'entreprise est un défaut ou pas. En plus des valeurs provenant des états financiers, nous aurons besoin des informations de l'entrepris issues des marchés financiers. Car oui, Cette analyse va se faire seulement pour les entreprises cotées en bourse.

Pour ce document, nous allons voir la méthode proposée par Edward I. Altman en 2005. Cette méthode se veut être une mesure de la probabilité de défauts pour les marchés émergents. Edward I. Altman a développé d'autres méthodes mais c'est celle de 2005 qui va nous intéresser.

2.2.4.2 L'analyse par familles de ratios

Maintenant que nous nous sommes familiarisés avec les familles de ratios, et que nous savons ce qu'elles représentent nous allons passer à leur analyse. Leur analyse est plus ou moins le même pour la plupart d'entre elles. Pour chaque famille, à l'exception de la probabilité de défaut, il y aura trois étapes : le calcul des ratios, l'analyse tendancielle et une représentation graphique. Pour ce qui est de la probabilité de défaut, il sera fait des calculs qui lui sont spécifiques.

2.2.4.2.1 Calcul des ratios

Chaque famille a ses propres ratios et c'est cela d'ailleurs l'intérêt et chaque ratio de chaque famille va nous éclaircir sur un aspect bien particulier de la santé de l'entreprise. Donc tous les ratios sont importants. Ci-après un tableau récapitulatif des familles et certains de leurs ratios.

Tableau 4 Les familles de ratios

Famille	Ratios
Profitabilité	Taux de marge nette, taux de valeur ajoutée, taux bruts d'exploitation ...
Rentabilité	Rentabilité économique et rentabilité financière.
Politique comptable	Ratio de vétusté, taux de provision de stocks, taux de provision de créances.
Liquidité	Liquidité (générale, réduite et immédiate)
Gestion de la dette	Levier, taux d'entêtement, maturité de l'endettement
Flux de trésorerie	Capacité à investir, taux d'investissement net, taux de réinvestissement ...
Efficacité des actifs du BFG	Variation CA, variation VA, taux de valeur ajoutée
Valorisation	Capitalisation boursière, valorisation boursière, capitalisation du CA ...
Probabilité de défaut	Zscore d'Altman, score de la banque de France.

2.2.4.2.2 Analyse tendancielle des ratios

Une analyse tendancielle est la comparaison des ratios sur toute la période d'étude par rapport à une année, cette année peut être n'importe laquelle (généralement une année qui est marquée par un événement fort), mais ici nous choisissons la première par souci de simplicité.

Ce que cette analyse tendancielle des ratios va nous permettre de faire, c'est constater l'évolution des ratios par rapport à l'année de référence. Ici, trois choses vont être intéressantes de noter, d'abord, toutes les valeurs de l'année de référence vont être égales à 100% (la comparaison va se faire en taux de croissance). La deuxième chose c'est, pour toutes les valeurs des autres années ont une croissance inférieure à 100% dans cas ces dernières est inférieure à l'année de référence. Troisièmement, c'est que si une valeur a une croissance de plus 100% c'est cette valeur est supérieur à celle de l'année de référence. Ainsi nous aurons une analyse qui nous montre comment les valeurs d'une famille de ratios évoluent par rapport à une année bien particulier.

2.2.4.2.3 La représentation graphique

Une image vaut mille mots. Il peut être challengeant de lire un tableau de valeur surtout dans de la finance ou les valeurs facilement atteindre des milliards. Ce se fait après avoir fait nos calculs

de ratios et d'analyse tendancielle, il est nécessaire de faire une représentation graphique pour faire parler les valeurs. La représentation graphique va être la dernière étape de l'analyse par les ratios, chose qui va donner une vue d'ensemble du travail fourni par l'entreprise et par l'analyste financier.

Voir annexe 1 pour les détails des calculs et une analyse financière complète sur Excel.

Conclusion partielle

Dans ce chapitre nous nous étions données comme mission de parler et faire comprendre l'analyse financière. Nous avons essayé tant bien que mal d'expliquer de manière claire et concise les aspects techniques de cette discipline de la finance. Il a d'abord été fait la présentation des états financiers. Après cela, nous avons essayé de lever le mystère de l'analyse financière.

De ce fait, la chose à retenir ici c'est comment faire une analyse et comment en tirer des enseignements. Toutes ses étapes sont importantes et l'analyste doit accorder une attention particulière à chacune d'entre elles. A la fin, un rapport doit être produit, rapport retraçant les observations faites par l'analyste.

Au vu de tout cela, nous avons vu que c'est un travail fastidieux, et très exigeant en temps, en effort et en compétence. Et ce travail peut être optimiser avec l'aide de l'IA. Nous allons voir comment l'IA va aider à optimiser l'analyse financière, elle pourra par exemple nous faire l'analyse financière dans quatre (4) avec une bonne précision ou encore prédire la faillite. Et toutes ces bonnes choses vont être abordées dans la suite du document.

3 Chapitre III : Théories derrière les algorithmes d'intelligence artificielle

Introduction de chapitre

Le domaine de l'IA est en train de suivre un boom littéral. Nous voyons de plus en plus des applications utilisant l'IA pour nous faciliter la vie. Mais de l'autre côté, beaucoup d'étudiant et de chercheurs mènent leurs recherches sur ce domaine passionnant. Maintenant, pour tout étudiant, ou ingénieur qui se respecte, il faut comprendre l'algorithme derrière. Car oui, aujourd'hui il y a mille et une outils nous permettant de créer des programmes intelligents, mais c'est toujours intéressant de comprendre ce qu'il y a derrière chaque fonction utilisée.

Pour faire simple les algorithmes d'IA ne sont ni plus ni moins que des mathématiques, nous avons déjà dits que le soubassement de l'IA c'est les mathématiques. Nous allons voir la magie derrière les familles de modèles et comment elles fonctionnent. Avant de commencer il faut préciser que les calculs mathématiques que nous allons voir sont quelque peu complexes et couteux.

Pour ce qui est du plan nous allons voir d'abord les prérequis à savoir les mathématiques, l'informatique et d'autres éléments très intéressants. Après cela, nous allons entrer dans le vif du sujet à savoir comment fonctionnent les modèles intelligents.

3.1 Les prérequis

Avant tout, il va falloir préciser un certain nombre de choses. L'IA n'est pas facile, ce n'est pas un domaine auquel n'importe qui sans le maximum de volonté peut y entrer. Bien évidemment il y a quelques prérequis, deux pour être précis que sont les mathématiques et l'algorithme. Et pour ces deux domaines il va falloir être excellent dans un et avoir un bon niveau dans l'autre ou maîtriser les deux. A part les mathématiques et l'informatique il y a quelques autres prérequis qui ne sont pas nécessaires mais peuvent aider dans notre objectif, et tous ses autres prérequis vont être classés dans le domaine de l'intelligence sociale.

3.1.1 Les mathématiques

Quand on parle de mathématique la plupart des gens vont prendre peur, abandonner voire même fuir. Mais ici, nous allons voir les concepts mathématiques qui nous seront utiles à l'IA mais de manière simple et concise.

Les mathématiques peuvent être compliquées, mais quand on lui trouve une application c'est là que ça devient intéressant et l'une des plus belles applications des mathématiques c'est l'IA. Nous allons vous montrer comment c'est fascinant de résoudre des problèmes mathématiques pour créer des modèles intelligents. Les mathématiques sont plus que nécessaire pour l'IA, elles sont vitales. D'ailleurs mon professeur d'intelligence artificielle nous disait à la fin d'un cours : « l'intelligence artificielle ce n'est ni plus ni moins que des calculs mathématiques ».

En dépit du fait qu'il y a plusieurs domaines mathématiques qui nous seront utiles dans l'IA, pour ce travail de mémoire, nous allons nous concentrer sur seulement trois (3) domaines des mathématiques que sont les statistiques et probabilités, l'algèbre linéaire et l'analyse.

3.1.1.1 Les statistiques et probabilités

Is everything in on this planet determined by randomness? This question is open to philosophy debate. What is certain is that every day thousands and thousands of engineers, scientists, business persons, manufactures, and others are using tools from probability and statistics. (Dekking, Frederik Michel, 2005).

Cette citation de Michel nous renvoie à comment sont important ses domaines dans nos vies de tous les jours et l'IA ne fait pas exception. La statistique est un domaine des mathématiques qui travaillent sur des données en les faisant parler ce qui nous permet de mieux comprendre les valeurs d'une base de données. C'est ce qu'on appelle les statistiques descriptives. Il y a aussi les statistiques inférentielles qui, comme son nom l'indique, vont nous permettre de faire des inférences c'est-à-dire faire des estimations. Et c'est là que réside le lien entre les probabilités et statistique car les statistiques inférentielles vont avoir besoin des probabilités. La probabilité est l'étude de la chance pour qu'un évènement se produise pour faire simple.

Ceci étant dit, comment ses deux sont utiles en Machine Learning et Deep Learning ? Ils interviennent tous les deux avant et après le développement de modèle d'IA.

- **Avant le développement du modèle** : les statistiques nous aident à comprendre les données, car très souvent les données brutes ne sont pas exploitables. Ici, nous vérifions le maximum des valeurs, le minimum, la moyenne, les outliers et l'une partie des plus importantes du « Feature Engineering » la mise à l'échelle etc.

- **Après le développement du modèle** : il va bien falloir calculer la fiabilité du modèle, ce qu'on appelle « accuracy », il faut calculer aussi, la précision, le f1-score, le recall... Ces derniers nous permettent d'apprécier la robustesse du modèle une fois déployer. Nous pouvons aussi faire des graphs comme la matrice de confusion par exemple.

3.1.1.2 L'algèbre linéaire

Au fait, il y a trois grandes parties dans le développement d'un réseau de neurones et à titre illustratif, nous pouvons dire qu'il y a le travail a posteriori, le développement du modèle et travail a priori. Pour le modèle il y a deux parties le Feed-forward et le Back-propagation et L'algèbre linéaire vont intervenir dans ces deux parties.

L'algèbre linéaire est la branche des mathématiques qui s'intéresse à l'étude des espaces vectoriels (ou espaces linéaires), de leurs éléments les vecteurs, des transformations linéaires et des systèmes d'équations linéaires (théorie des matrices). (Algèbre linéaire - Définition, 2024)

Ainsi, la plus grande utilité de l'algèbre linéaire est le calcul de poids, elle va nous permettre d'automatiser les calculs lourds et couteux, de ce fait nous permettant de gagner du temps. Sans elle, nous aurions passé beaucoup de temps sur ces calculs. Je rappelle qu'un réseau de neurones a des milliers de neurones d'inputs, plusieurs couches cachées qui peuvent avoir elle-même des milliers de neurones. C'est juste impossible de calculer tout ceci d'une manière séquentielle.

3.1.1.3 L'analyse

Quand on parle d'apprentissage en IA ou le terme très populaire anglais « Learning », c'est à cause du domaine des mathématiques l'analyse avec le calcul des dérivés. Si l'IA est devenue ce qu'elle est devenue aujourd'hui avec les performances qu'on la connaît, c'est en très grande partie à cause des calculs d'analyse mathématique.

La dérivée d'une fonction nous informe sur la variation de la fonction en un point. Pour être plus claire, elle nous permet de calculer la pente sur n'importe quel point de la fonction. Le calcul de dérivée est très important dans beaucoup de domaines notamment dans de Deep Learning.

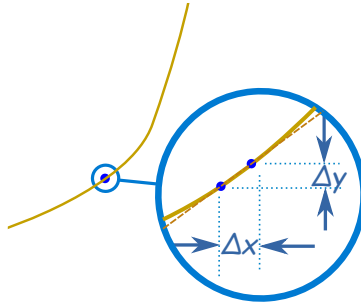


Figure 2 Zoom sur la dérivée d'une fonction (mathisfun)

Maintenant comment se passe l'apprentissage dans un réseau de neurones ? Nous allons rectifier les erreurs commises par l'IA dans son entraînement à travers la dérivée. Si nous répétons cela autant de fois que nécessaire, l'erreur d'assomption va se réduire au minimum et l' « accuracy » va se maximiser. Ce qu'il faut comprendre par-là, c'est que le calcul de la dérivée de la fonction d'erreur va nous permettre de rectifier cette erreur.

De manière pratique, on calcule l'erreur en premier lieu, puis on calcule la dérivée de la fonction d'erreur. La manière dont la rectification va se faire, c'est qu'on va donner à chacun des poids une valeur correspondante à sa responsabilité dans l'erreur et c'est ça le Back-propagation ou la rétropropagation en français.

3.1.2 L'informatique

L'informatique c'est la science de l'automatisation de l'information, d'ailleurs son nom vient de là : une contraction entre information et automatique. Chez les anglo-saxons, ils parlent plutôt de Computer Science qui se traduit littéralement par science de l'ordinateur.

Plus haut, nous avons attesté que l'IA est une science purement mathématique avec que des calculs que l'on pourrait même faire sur feuille. Dès lors, que représente l'informatique pour l'IA : il a le rôle d'une calculatrice géante capable de faire des super-calculs en un temps record et aussi faire office d'interface graphique.

3.1.2.1 L'algorithme

Bien évidemment, la première des choses que l'on va voir c'est l'algorithme. Nous pouvons attester sans prendre beaucoup de risques que l'algorithme est l'informatique, et que l'informatique est l'algorithme. L'algorithme est l'ensemble des étapes auxquelles il va falloir passer pour

résoudre un problème informatique. Il est souvent fait l'analogie de la recette de cuisine pour illustrer l'algorithme et à juste titre.

La raison pour laquelle il faut maîtriser l'algorithme c'est que : pour implémenter un problème mathématique dans un ordinateur, il faut savoir comment s'y prendre et savoir quelles étapes à suivre, sinon beaucoup de frustration nous attend.

Exemple : écrivons un algorithme qui résout un polynôme du second degré :

- Afficher : Donner les valeurs a, b et c.
- Stocker a, b et c dans des variables.
- Calculer delta ($\text{delta} = b^2 - 4 * a * c$)
- Si delta positif alors $x1 = (-b - \text{racine}(\text{delta}) / 2 * a)$ et $x2 = (-b + \text{racine}(\text{delta}) / 2 * a)$
- Si delta nul alors $x = \text{racine}(\text{delta}) / 2 * a$
- Si delta négatif alors il n'y a pas de solution dans R.

Voici ci-dessus un algorithme qui marche pour un polynôme du second degré et cette même manière de réflexion peut nous permettre d'implémenter n'importe quel problème déjà résolu en mathématique en algorithme informatique.

3.1.2.2 Les structure de données

D'abord, les structures de données sont le terme utilisé pour représenter toutes les différentes façons en informatique pour modéliser les données avec lesquelles nous travaillons. Très souvent, pour ne pas dire tout le temps, nous n'avions pas directement la façon optimale de gestion de données.

Ces structures peuvent partir d'un simple tableau dans un langage de programmation jusqu'à atteindre les graphs (structure de données complexe et très puissante). Comme nous l'avons dit et redit l'IA travaille sur des données. Citons quelques exemples de structures de données :

- Les listes chaînées
- Les tables de hachages
- Les arbres
- Les piles et files
- Les graphs

3.1.2.3 Les langages de programmation

Les langages de programmation aussi appelés langages informatiques sont les syntaxes qui traduisent les algorithmes d'une manière compréhensible à l'ordinateur. Il faut préciser que l'ordinateur ne comprend pas le texte, il comprend seulement les chiffres (nombre binaire en l'occurrence). Ce que le langage de programmation fait, c'est de convertir sa syntaxe en langage binaire compréhensible par l'ordinateur et chaque langage a sa propre syntaxe.

L'importance des langages de programmation va être évidente pour tout le monde, nous pouvons citer : C/C++ (important pour l'IA), Python (important pour l'IA), Java, PHP, JavaScript ...

3.2 Les algorithmes d'intelligence artificielle

Nous y voilà, l'une des parties les plus importantes de ce travail de mémoire. Nous avons parlé de l'IA dans ce document mais cette fois nous allons voir comment elle fonctionne en parcourant différents des plus importants algorithmes d'IA, ceux qui sont vraiment utilisés par les grandes entreprises. Donc pour cette partie, je vais vous demander une attention particulière car ce sera très intéressant. Alerte âme sensible !!! il y aura beaucoup de calculs mathématiques dans cette partie.

3.2.1 Machine Learning

Littéralement, Machine Learning veut dire apprentissage des machines. Comme nous êtres humains, nous naissons sans connaissance dans notre tête, mais en regardant notre environnement et en imitant nos parents, nous apprenons. Ce processus peut être répliqué sur un ordinateur, c'est le Machine Learning, il y en a deux : Supervised Machine Learning (SML), Unsupervised Machine Learning (UML).

3.2.1.1 Supervised learning

Si nous reprenons l'analogie de l'enfant, dans sa phase d'apprentissage ses parents vont être derrière et le guider. Si l'enfant commet des erreurs ses parents vont le rectifier s'il fait une bonne chose ses parents vont le récompenser ou l'encenser.

Dans le domaine des ordinateurs, pour faire en sorte qu'une machine apprenne, on aura besoin de données, beaucoup de données. Et chaque ligne de données va être étiquetée, on parle input et d'output. Maintenant, le modèle va essayer de s'adapter à tous les inputs et leurs outputs.

Nous allons voir dans la suite les différents types d'apprentissage supervisé et leurs algorithmes.

3.2.1.2 La régression

La régression est une méthode statistique qui nous permet d'approximer la valeur d'une variable à partir des valeurs déjà présentes et connues. Elle va se faire en traçant une courbe qui représente le mieux la relation des points dans un repaire orthonormal. Il y a plusieurs types de régression mais nous allons en voir trois (3).

3.2.1.2.1 La régression linéaire

La régression linéaire nous permet de tracer une droite qui va au mieux s'adapter aux données d'une courbe. Maintenant si nous voulons tracer une droite qui va au mieux représenter l'évolution de ces points, qu'allons-nous faire. Il y a la méthode des moindres carrés, élaboré par le légendaire Carl Friedrich Gauss, qui est une méthode purement statistique mais nous allons utiliser une méthode d'IA avec la descente des gradients. Cette dernière méthode peut être divisée en trois parties :

- Forward propogation (essaie au hasard)

D'abord la courbe que l'on veut tracer va être de la forme $y = ax + b$, mais dans le jargon on va parler w_0 et w_1 qui vont représenter les poids respectifs, l'équation devient $y = w_0.x + w_1$. Le but du jeu est de trouver les w_0 et w_1 qui vont au mieux correspondre à nos points. Dans un premier temps on va les donner des valeurs aléatoires d'où l'essai au hasard.

- Calculer l'erreur

Très souvent, pour ne pas dire jamais, le premier essaie va être une erreur, de ce fait, il faut calculer l'erreur, nous allons utiliser la fonction suivante :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2$$

Équation 29 Mean Square Error

MSE : Mean Square Error (la moyenne des erreurs au carré)

Y : la sortie attendue

\hat{Y} : la sortie observée

N : le nombre d'élément dans le tableau

Cependant, les plus curieux vont se demander pourquoi élever l'erreur au carré. C'est une bonne question. La raison est simple car une erreur de -1 est égale à une erreur qui vaut 1. Et le fait de l'élever au carré va nous aider dans la mise à jour des poids où nous allons utiliser l'algorithme de la descente des gradients.

- Back-propagation (rétropropagation qui met à jour les poids)

Maintenant que nous avons l'erreur nous pouvons enfin mettre à jour nos poids w_0 et w_1 . Cela veut dire que chacun va prendre une part de l'erreur qui est égale à sa responsabilité de cette dernière et se rectifier lui-même. Pour se faire nous allons calculer la dérivée de toutes les fonctions qui nous ont mené à cette erreur de manière suivante :

$$w_0 = w_0 - lr * \frac{\partial MSE}{\partial w_0}$$

Équation 30 Mise à jour des poids

$$w_1 = w_1 - lr * \frac{\partial MSE}{\partial w_1}$$

Dans la descente des gradients, il y a ce qu'on appelle le pas, il va déterminer à quelle vitesse la descente va se faire. Si le pas est trop petit l'apprentissage va être lent et si le pas est trop grand, nous allons dépasser le point qui minimise l'erreur, ce pas c'est le « LR » dans les deux fonctions cela signifie « *Learning Rate* ».

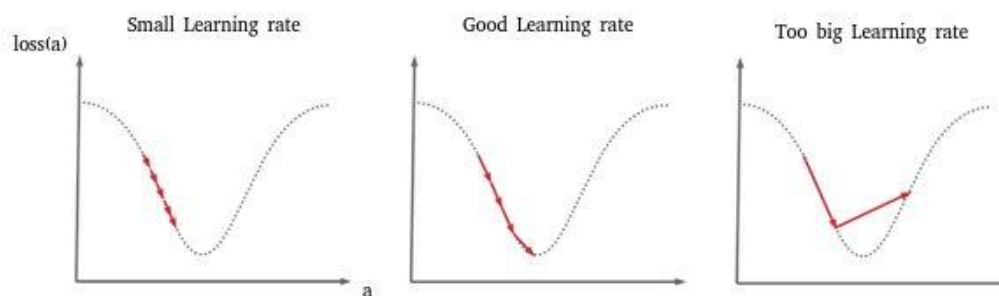


Figure 3 Les variations de Learning Rate (Cayla, 2021)

$\frac{\partial MSE}{\partial w}$ Représente quant à elle, la dérivée de la fonction MSE par rapport au poids concerné, c'est ce qu'on appelle une dérivée partielle. Exemple :

$$f(x, y) = x \cdot y$$

$$\frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x$$

Donc

$$\frac{\partial MSE}{\partial w_0} = \frac{\partial MSE}{\partial Y} * \frac{\partial y}{\partial w_0}$$

Équation 31 Le chainage des dérivées partielles

Voir annexes 2 pour un exemple de calcul de régression linéaire.

3.2.1.2.2 La régression logistique

La régression logistique, contrairement à celle dite linéaire, n'a pas pour vocation de prédire une valeur future. Sa prédiction est du type binaire : oui ou non, bon ou mauvais, 0 ou 1 etc. Ceci va s'avérer être très important dans beaucoup de domaines, nous l'utilisons dans nos vies de tous les jours sans s'en rendre compte. Par exemple, détecter si un email est un spam ou non, si une information est un fake new ou non, si un investissement va être rentable ou pas...

- Forward propagation

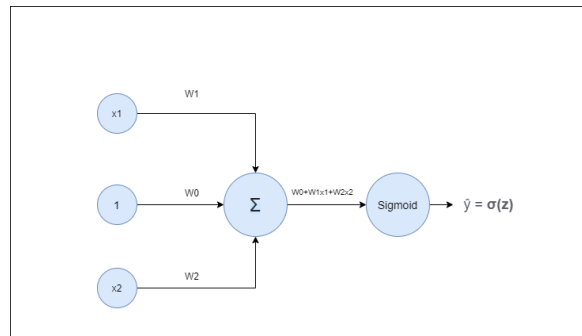


Figure 4 Simple réseau de neurone (Kumar, 2020)

Voici à quoi va ressembler notre réseau de neurones, on va ajouter un autre input en plus x_1 et x_2 , c'est le biais qui va toujours être égale à 1, son utilité est d'éviter que certains neurones ne meurent durant l'entraînement si $x_1 = 0$ et $x_2 = 0$.

$$y = \sum_{i=1}^n w_i \cdot x_i$$

Équation 32 Calcul de la sortie observée

Puisqu'on dit que les valeurs de sortie doivent être 0 ou 1, nous devons trouver un moyen de toujours mettre à l'échelle la sortie observée, c'est là qu'intervient la fonction d'activation. Pour les problèmes de régression logistique il y en a deux très populaires : la fonction à seuil et sigmoid.

Fonction	Formule	Sortie possible
Seuil(x)	$1 \text{ si } x > 1, 0 \text{ sinon}$	0, 1
Sigmoid $\sigma(x)$	$\frac{1}{1 + e^{-x}}$ <i>Équation 33 La fonction sigmoid</i>	Tout réel compris en 0 et 1

Nous allons continuer avec la fonction sigmoid :

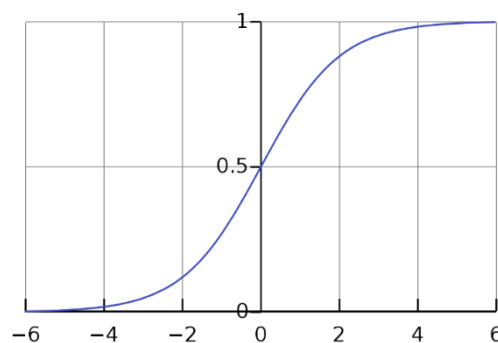


Figure 5 La fonction sigmoid (Saleem, 2023)

- Calculer l'erreur

Pour l'erreur rien ne va changer nous allons utiliser la Mean Square Error :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2$$

- Backpropagation

Nous voici près pour la rétropropagation, seulement ici nous allons mettre à jour trois poids à savoir w_0 , w_1 , w_2 .

$$w_0 = w_0 - lr * \frac{\partial MSE}{\partial w_0}, \quad w_1 = w_1 - lr * \frac{\partial MSE}{\partial w_1}$$

$$w_2 = w_2 - lr * \frac{\partial MSE}{\partial w_2}$$

Même si les formules restent les mêmes, ne prenons encore rien pour acquis, ici la valeur de la dérivée partielle va changer étant donné qu'on a introduit une nouvelle fonction, celle d'activation, nous allons de facto nous retrouver avec trois membres dans le calcul de dérivée partielle.

$$\frac{\partial MSE}{\partial w1} = \frac{\partial MSE}{\partial \sigma} * \frac{\partial \sigma}{\partial Y} * \frac{\partial Y}{\partial w1}$$

Voir annexe 2 pour un exemple de calcul de régression logistique.

3.2.1.2.3 La régression polynomiale

La régression polynomiale nous permet de représenter une courbe de donnée qui adapte une forme exponentielle. Les étapes de régression polynomiale restent les mêmes que les autres algorithmes mais ses calculs vont changer.

- Forward-propagation

Pour le Forward-pass de la régression polynomiale, nous allons utiliser, une fonction quadratique, c'est-à-dire qui admet une puissance dans la variable. On va parler de degré de la fonction. Plus le degré est élevé plus la fonction pourra être en mesure d'aller chercher des variations.

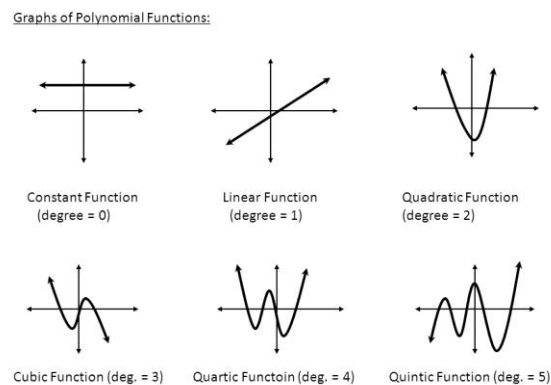


Figure 6 Différents degrés de la régression polynomiale (Graph of polynomial functions, 2020)

Du fait que nous n'avons pas beaucoup de variation dans le tableau, nous allons utiliser le deuxième degré, ainsi notre formule se présente comme suit :

$$y = w0 + w1 * x + w2 * x^2$$

- Calculer l'erreur

La fonction d'erreur ne change toujours pas, c'est le MSE.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2$$

- Backpropagation

$$w0 = w0 - lr * \frac{\partial MSE}{\partial w0}, \quad w1 = w1 - lr * \frac{\partial MSE}{\partial w1}, \quad w2 = w2 - lr * \frac{\partial MSE}{\partial w2}$$

La valeur de la dérivée partielle pour w1 se présente comme suit :

$$\frac{\partial MSE}{\partial w2} = \frac{\partial MSE}{\partial Y} * \frac{\partial Y}{\partial w2}$$

Nous avons décidé de prendre w2 car il a la dérivée partielle la plus compliqué à calculer, avec ce calcul établi, nous pouvons passer à l'étape des mises à jour des poids, ne pas oublier de prendre un Learning Rate.

Voir annexe 2 pour un exemple de calcul de régression polynomiale.

3.2.1.3 La classification

La classification est un problème qui est là depuis longtemps dans le domaine de l'intelligence artificielle. Les académiciens ont fait beaucoup de recherches sur le sujet et nous ont proposé un certain nombre de méthodes.

La classification a pour objectif de déterminer les éléments qui différencient les données dans une base de données, ainsi ranger chacune dans sa classe de prédilection et aussi mais surtout prédire les classes pour des données non encore observées.

Différents algorithmes sont aujourd'hui là pour nous permettre de régler les problèmes de classification mais nous allons en voir trois (3).

3.2.1.3.1 Support Vector Machine (SVM)

Le SVM est un modèle mathématique qui permet de classer des données en utilisant un séparateur. Ce séparateur peut être une ligne dans un espace 2D ou un plan dans un espace 3D, il est possible d'avoir autant de dimension que nécessaire mais le séparateur va avoir une dimension n-1 par

rapport aux données. Comme tout algorithme, il y a un certain nombre d'étapes à suivre pour pouvoir réussir à implémenter un SVM.

- Le séparateur

Pour un problème linéairement séparable dans un espace 2D, nous avons besoin d'un séparateur (une droite) qui doit être aussi loin du point le plus proche d'une classe que du point le plus proche de l'autre classe. Ce séparateur se présente comme suit.

$$y = mx + b = 0$$

Équation 34 Séparateur en SVM dans un espace 2D

Si les données se présentent d'une manière non linéairement séparable, nous allons introduire une fonction de linéarisation appelée kernel (cette fonction ajoute une dimension aux données pour les rendre linéairement séparables). Dans ce cas on aura :

$$y = m'(x) + b = 0$$

Équation 35 Fonction de linéarisation

De là, nous nous retrouvons avec deux classes C1 et C2.

$$C1 = \{x: mx + b \leq 0\}, \quad C2 = \{x: mx + b > 0\}$$

- Calcul de distance

Nous devons maintenant calculer la distance d des droites parallèles qui représentent les frontières.

$$d = \frac{(b2 - b1)}{\sqrt{m^2 + 1}}$$

Équation 36 Calcul de distance en SVM

- Calcul de l'erreur pour chaque point

Ici, la fonction de l'erreur va nous permettre de savoir dans quelle classe un point pris en particulier se trouve-t-il.

$$e = 1 - y(mx + b)$$

Équation 37 Erreur en SVM

- L'apprentissage

Le but de l'apprentissage en SVM, et pour tout modèle de machine Learning d'ailleurs, c'est minimiser l'erreur, en SVM il y a une fonction qui peut minimiser cette erreur.

Il faut faire de telle sorte que l'erreur soit inférieure ou égale à 0, et nous savons que :

$$e = 1 - y(mx + b)$$

Donc

$$1 - y(mx + b) \leq 0, \quad \text{minimiser} = \frac{\|m\|^2}{2}$$

Équation 38 Fonction de minimisation de l'erreur en SVM

$$y(mx + b) \geq 1$$

Et nous y voilà, toutes les informations sur cette partie sur le SVM nous viennent de l'article de S. Suthaharan : Support Vector Machine, chapitre 9.

Voir annexe 3 pour un exemple de calcul du SVM.

3.2.1.3.2 L'arbre de décision

L'arbre de décision ou "decision tree" en anglais est aussi une méthode de classification avec un concept qui lui est bien particulier. Comme son nom l'indique elle prend des décisions en se basant sur l'attribut des données. D'abord l'arbre vérifie l'attribut le plus indicatif et prends la direction d'une de ses valeurs, puis le deuxième attribut le plus significatif et prends la direction d'une de ses valeurs, ainsi de suite jusqu'à classer un nouvel enregistrement.

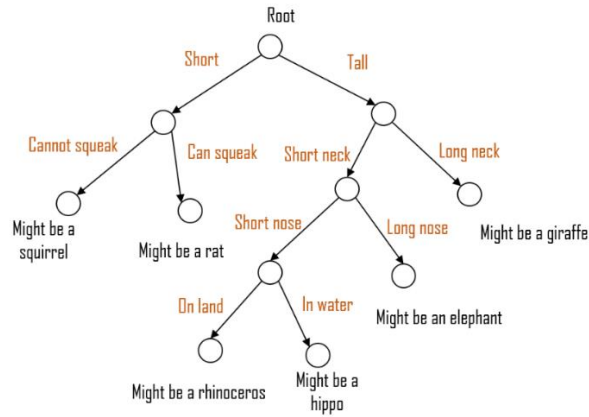


Figure 7 Structure d'un arbre de decision (Zhao, 2021)

En revanche, le fait de distinguer un attribut significatif ne se fait pas arbitrairement, sinon ce ne serait pas une intelligence artificielle. Il y a un certain nombre de calculs (oui encore des maths) à faire pour trouver la bonne structure de l'arbre et nous allons les voir tout de suite.

- Entropie

L'entropie nous renseigne sur la pureté d'un attribut, si deux classes sont équitablement représentées dans un attribut, on dit que le nœud est impur, conséquence l'entropie est maximale (égale ou proche de 1), si une seule classe est représentée le nœud est pure et l'entropie est minimale (égale ou proche de 0).

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i$$

Équation 39 Entropie

- Gain d'information (GI)

La première des choses à faire c'est de calculer le gain d'information c'est-à-dire de tous les attributs, quel est celui qui nous renseigne le plus si l'individu est sénégalais ou pas.

$$GI(S, A) = Entropy(S) - \sum_{i=1}^n p(S_A) * Entropy(S_A)$$

Équation 40 Gain d'information

Voir annexe 3 pour un exemple de calcul d'un arbre de décision.

3.2.1.3.3 Naive Bayes

Le modèle de Naïve Bayes (NB) est un algorithme de ML qui nous vient des statistiques et probabilités. Selon les cas, il peut être très puissant avec un mécanisme simple de calcul de probabilité. Il fonctionne en calculant les probabilités de toutes les valeurs d'attributs avec la variable cible.

- Probabilité des variables cibles

Tout d'abord il faut calculer la probabilité de toutes les variables cibles afin de savoir nos chances de tomber sur l'un ou l'autre (il est possible d'utiliser le NB dans une multi-classe classification aussi).

$$P(C_i) = \frac{\text{nombre } C_i}{N}$$

Équation 41 Calcul de probabilité

- La probabilité conditionnelle des valeurs d'attribut

Pour chaque valeur d'attribut, il nous faut calculer sa probabilité conditionnelle par rapport aux valeurs cibles.

$$P(C_k|x) = \frac{P(C_k) * P(x|C_i)}{P(x)}$$

Équation 42 Probabilité conditionnelle

Cela semble peu mais on a presque tout le travail qui est fait, en pratique il y aura beaucoup de calculs à faire. Maintenant nous pouvons classer un nouvel individu en calculant sa probabilité de se trouver dans une classe ou une autre, ensuite nous allons normaliser les probabilités et classer dans celle qui a la plus grande valeur.

$$P(C_k|N) = P(C_k) * \sum_{i=1}^n P(val\ attri = N\ attrib|Ck)$$

Pour normaliser les probabilités :

$$Pn(C_k) = \frac{P(C_k)}{\sum_{i=1}^n P_{ci}}$$

Équation 43 Normalisation de probabilité

3.2.1.4 Unsupervised learning

Pour ce qui est de l'apprentissage non supervisé, c'est qu'ici nous n'aurons pas d'output pour les inputs. Dans ce cas de figure nous aurons seulement des données d'entrée mais on ne sait pas comment réagir en conséquence. C'est le modèle qui va à lui seul trouver une représentation générale qui correspond le plus aux données qui lui sont présentées.

Pour ce faire il y a ce qu'on appelle le clustering : c'est un modèle dans lequel nous allons essayer de regrouper en cluster les individus qui se ressemblent le plus en utilisant plusieurs variables qui décrivent les données.

3.2.1.4.1 Clustering

Le clustering est une méthode d'apprentissage non supervisé dans lequel le but est de rassembler les individus qui se ressemblent le plus. Le principe est simple, nous avons des données mais qui ne sont pas étiquetées, donc c'est au modèle de trouver la représentation la plus fidèle des données. Il y a plusieurs algorithmes de clustering mais nous allons voir le fameux k-means (le k de k-means représente le nombre de classe ou cluster).

- Définir le nombre de cluster

En premier lieu, il faut définir le nombre de cluster, ce choix peut relever du libre arbitre de l'ingénieur ou peut-être défini en fonction de méthodes.

- Le centre de gravité

Pour chaque cluster il faut calculer son centre de gravité et on affecte chaque point de la base de données à la classe la plus proche. De là, tous les individus appartiennent à une classe et c'est là que le travail commence.

- Calcul de distance

Maintenant, nous allons calculer toutes les distances de tous les individus par rapport à tous les centres de gravité de chaque cluster. Nous allons nous apercevoir que certains individus sont mal classés, car ils sont plus proches d'un autre cluster que celui où ils sont, il suffit de les mettre à

jour. Cette étape va être répétée autant de fois que nécessaire pour avoir des clusters les plus représentatifs des données que possible.

Voir annexe 4 pour un exemple de calcul de clustering.

3.2.1.4.2 Règles d'associations

Les règles d'association ou en anglais "association rules mining" sont des méthodes non supervisées qui nous permettent de trouver la corrélation entre une donnée et les autres. Ses règles permettent de répondre à des questions comme : dans quelle mesure B et C vont apparaître sachant que A est apparu ? Ces calculs vont se faire avec un ensemble de sous-ensemble, nous allons parler ici de itemset pour désigner les sous ensemble. Les règles d'associations sont très fréquentes dans les marchés et supermarché pour déceler les produits qui sont souvent achetés en un temps par les clients. Une fois que nous avons des règles intéressantes les dirigeants peuvent prendre de bonne décision.

3.2.2 Deep Learning

Frank Rosenblatt a créé le perceptron qui nous a permis de résoudre des problèmes, notamment le OU et le ET logique, mais quand ils l'ont essayé pour le XOR, ils se sont rendu compte que le perceptron ne convergeait pas. Le problème était simple, un perceptron traçait des séparateurs linéaires, or ce n'était pas possible pour le problème du XOR.

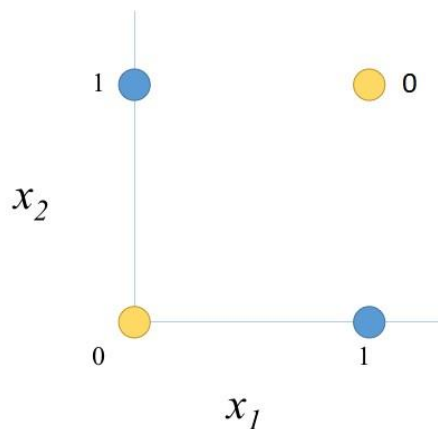


Figure 8 Représentation de la fonction XOR (Oman, 2016)

Allez-y ! essayer de tracer une seule droite qui est capable de séparer les 0 et le 1, une droite ce n'est pas possible. Bienvenue dans le monde du non linéaire, un monde qui fut un casse-tête pour

les chercheurs pendant le longtemps, jusqu'à qu'ils découvrent les solutions qui vont être présentées ici.

3.2.2.1 Artificial neuron network (ANN)

Si vous vous rappelez la partie portant sur la régression logistique, vous avez déjà quelques notions sur les ANN. Là-bas nous faisons un apprentissage, c'est une couche d'entrée et la couche de sortie, mais ici il sera question d'une couche d'entrée, une ou plusieurs couches cachées et la sortie. Plus il y a de couche cachée, plus c'est profond : apprentissage profond : Deep Learning. D'ailleurs nous avons très probablement tous déjà vu la représentation d'un réseau de neurones profonds.

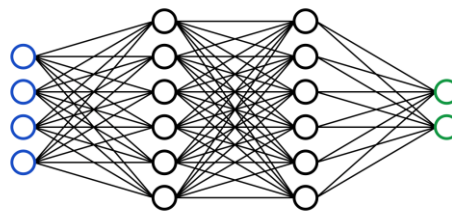


Figure 9 Structure de réseau de neurones (Ahmad, 2020)

3.2.2.2 Convolutional neuron network (CNN)

S'il y a un domaine où l'humain a toujours dépassé la machine, c'est la vision reconnaître des choses, des éléments de la nature et de les classer. Mais depuis quelque temps les scientifiques ont réalisé d'énormes avancées sur le domaine appelé Computer Vision ou vision par ordinateur. Et l'un des premiers algorithmes qui a permis de réaliser cela reste le CNN que l'on va voir tout de suite.

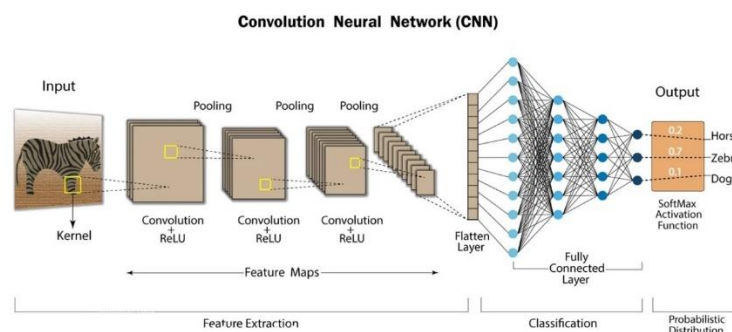


Figure 10 Réseau d'un CNN (Shahriar, 2023)

3.2.2.3 Recurrent neuron network (RNN)

Nous venons juste de parler des ANN et de leurs utilités, mais dans tout domaine, il y a toujours des limites. Le principal reproche que l'on peut faire au ANN, c'est ils n'ont pas de mémoire. Imaginons un jeu de données avec 60 000 inputs, de la première ligne du premier epochs jusqu'à la dernière ligne du dernière epochs, le modèle va oublier tout ce qui s'est passé et se concentre seulement sur les caractéristiques principales. Mais très souvent, il est nécessaire de savoir ce qui s'était passé pour décider de ce que l'on va prédire.

Exemple : le Sénégal est un pays qui se trouve en Afrique et dont l'ethnie principale est composée de ...

Nous voulons prédire ce qui va arriver et nous avons trois propositions : ashantis, masais, wolofs. Et bien évidemment c'est les wolofs. Le mot qui nous a permis de décider c'est Sénégal bien sûr, or ce mot se trouve au début de la phrase et donc ce modèle doit avoir une certaine mémoire pour bien prédire.

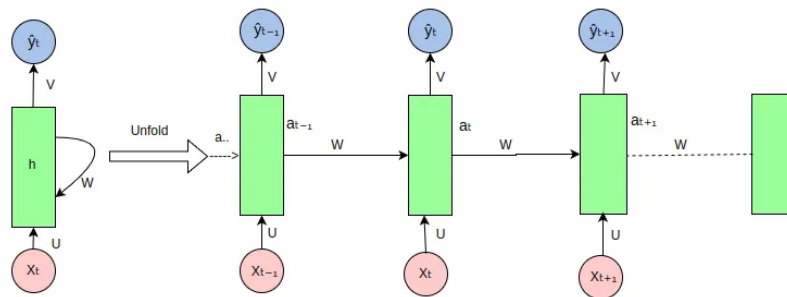


Figure 11 Réseau d'un RNN (Poudel, 2023)

Voici la structure générale d'un RNN, il y a, en fait, une seule couche et il représente évolution dans le temps. Pour ce qui est de l'erreur et du back-propagation, ce sera la même chose que les ANN que nous avons déjà vue.

Conclusion partielle

Nous y voilà, c'est le plus gros du travail dans cette rédaction de mémoire. Comme promis, nous sommes entrés dans les détails des algorithmes et nous espérons que toute personne ayant lu ce chapitre va un tant soit peu comprendre l'IA. Nous avons a priori de cela vu les prérequis pour comprendre les modèles.

La chose la plus importante à retenir dans ce chapitre c'est que l'IA n'est pas facile et demande beaucoup de connaissances dans les mathématiques et l'informatique mais avec le maximum de volonté, de détermination, de discipline et beaucoup de ses bonnes choses, n'importe quelle personne peut le faire.

Avec toutes ses connaissances acquises, nous sommes maintenant fin prêts pour pratiquer tout cela. Nous pouvons à partir de maintenant développer, déployer et intégrer dans une interface graphique nos modèles. Et bonne nouvelle, c'est ce que nous allons faire dans la prochaine partie.

4 Chapitre IV : Implémentation des modèles

Introduction de chapitre

Nous avons décidé de nous lancer dans un trajet très ambitieux, à savoir développer des modèles intelligents qui pourront non seulement prédire des valeurs futures mais aussi nous aider dans le développement d'un Chatbot que les utilisateurs vont pouvoir utiliser pour communiquer facilement avec les états financiers de leurs entreprises respectives.

Il a été constaté qu'il n'est pas pertinent de parler de la théorie seulement sans pour autant réaliser une production issue de ce travail de mémoire. De là, nous allons parler de fond en comble dans le chapitre précédent des théories derrière les algorithmes de l'IA, qui a été plein d'enseignement. Maintenant ici, nous allons mettre en pratique tout cela en développant un logiciel complet d'analyse financière renforcé par l'IA.

Avant tout projet informatique, il y a derrière le cahier des charges qui est un document retraçant les fonctionnalités du produit. C'est pour cela que nous allons commencer par-là pour qu'un non-informaticien puisse être en mesure de comprendre la chose qui va être faite. Ceci étant dit, nous allons continuer avec la présentation des outils que nous allons utiliser pour réaliser cette application, ensuite nous allons montrer les étapes que nous avons suivi pour faire la collecte de données. Après c'est les trois parties les plus importantes, il s'agira de la prédiction des modèles, comprendre comment nous avons fait nos prédictions, le développement du chatbot, expliquer de

manière claire le processus qui a abouti à ce chatbot et enfin nous allons parler de l'interface graphique.

4.1 Spécifications des besoins

Premièrement avant le démarrage de projet informatique, il y a le cahier des charges. C'est un document qui retrace les fonctionnalités de l'application qui doivent être implémentées. Pour cette partie, nous allons nous mettre dans la peau du maître d'ouvrage et rédiger les spécifications de tout le travail que nous allons faire. Il n'y aura rien de complexe, juste une rédaction simplifiée des étapes auxquelles nous devons passer pour réussir à créer un modèle intelligent qui pourra prédire une analyse financière, réussir à développer un chatbot et présenter tout cela dans une interface graphique.

4.1.1 Contexte de l'application

Le domaine de la finance est très vaste, il y a beaucoup de professionnels, d'ingénieurs, de chercheurs qui s'y illustrent. La finance est une science qui vise à une meilleure gestion du portefeuille des entreprises, que cela soit l'investissement, le financement...

C'est ainsi que nous introduisons l'analyse financière qui nage dans cette vaste mer du domaine de la finance. L'analyse financière ou le diagnostic financier nous permet de comprendre le processus de création de richesse, de voir si l'entreprise est dans une santé financière et j'en passe.

Aujourd'hui cette analyse financière se fait par des logiciels généralistes, ou même des fois sur feuille ce qui constitue un travail fastidieux. Malgré le fait que ces logiciels peuvent être d'une grande importance, ils souffrent souvent du trop-plein de fonctionnalités. Il y a toujours des fonctionnalités dont nous n'aurions jamais besoin si nous les utilisons pour faire une analyse financière. Autre chose, ces logiciels sont souvent très coûteux et nécessitent un réinvestissement annuel, chose qui n'est pas à la portée de monsieur tout le monde.

Ceci étant posé, imaginons une application unique à vocation financière, dans cette application, il n'y aura pas de fonctionnalité qui ne sera pas relatif à la finance. Dans cette application, les utilisateurs pourront être en mesure de gérer les finances de l'entreprise, de faire une analyse financière, et aussi faire des prédictions. C'est cela le désir manifesté par les personnes qui travaillent dans la finance. En plus de cela nous en tant que futur ingénieur en IA, nous nous

sommes dit qu'un petit chatbot qui permettrait de faire gagner du temps aux usagers ne serait que bénéfique pour tout le monde.

Finalement, après toutes ces observations nous avons conclu que développer ce genre d'application peut avoir un réel succès autant pour le concepteur que pour les utilisateurs. Cependant, avant de s'y aventurer il faut bien savoir les besoins fonctionnels du logiciel.

4.1.2 L'expression des besoins

L'expression des besoins est une partie indéboulonnable dans la rédaction d'un cahier des charges. Dans sa forme la plus simple (ce que nous allons faire ici), l'expression des besoins se fait en listant les éléments qui entrent dans sa composition. Nous avons deux types de besoins, les fonctionnels et les non fonctionnels :

- Les besoins fonctionnels
 - Création de nouveaux fichiers
 - Sauvegarde d'un fichier
 - Ouverture d'un fichier
 - Consulter le bilan
 - Consulter le compte de résultat
 - Consulter le tableau des flux de trésorerie
 - Modifier le bilan
 - Modifier le compte de résultat
 - Modifier le tableau des flux de trésorerie
 - Prédire le bilan
 - Prédire le compte de résultat
 - Prédire le tableau des flux de trésorerie
 - Faire une analyse financière complète
 - Faire une analyse prédictive financière complète
 - Avoir un chatbot
 - Faire toutes les actions citées avec le chatbot
 - Imprimer des documents
 - Rédiger un rapport d'analyse
- Les besoins non fonctionnels

- Un temps réduit de traitement
- Un prix abordable
- Facile à utiliser
- Sécurité des données
- Assurer la maintenance
- Déploiement seulement dans un réseau local

4.1.3 Les fonctionnalités de l'application

Pour la plupart des projets informatiques, il y a un déphasage entre les besoins du client et ce qu'a compris l'ingénieur, il peut même y avoir un décalage entre ce qu'a compris l'ingénieur et ce qui est vraiment réalisable comme l'illustre cette image.

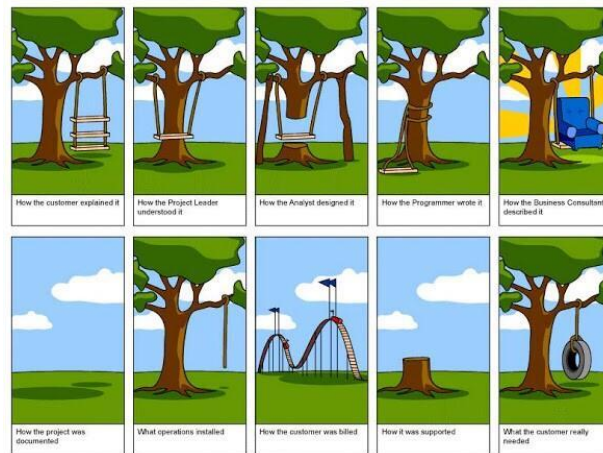


Figure 12 Gestion de projet (Rooij, 2020)

C'est ce qui explique peut-être le taux de succès des projets informatiques mais cela, c'est un autre débat. Bref ! il va falloir toujours être en communication avec le client pour ne pas faire du travail dans le vent. Pour ce qui nous concerne bien que l'application finale qui aura pour vocation d'être vendue dans les entreprises doit répondre à tous les besoins cités, celle que nous allons développer ici qui sera une version bêta va répondre seulement aux besoins indispensables. De ce fait, les fonctionnalités que nous allons développer dans le logiciel pour ce travail de mémoire sont :

- Consulter le bilan
- Consulter le compte de résultat
- Consulter le tableau des flux de trésorerie
- Modifier le bilan

- Modifier le compte de résultat
- Modifier le tableau des flux de trésorerie
- Prédire le bilan
- Prédire le compte de résultat
- Prédire le tableau des flux de trésorerie
- Faire une analyse financière complète
- Faire une analyse prédictive financière complète
- Avoir un chatbot
- Faire toutes les actions citées avec le chatbot
- Rédiger un rapport d'analyse (limitée)

Les besoins non fonctionnels vont être pris au sérieux dans les versions qui vont entrer en production. Maintenant, nous savons exactement ce que nous voulons faire, il va falloir choisir les outils informatiques avec lesquelles nous allons travailler.

4.2 Développement des modèles intelligents

4.2.1 La collecte de données

S'il y a bien une chose sur laquelle nous avons insisté dans ce document, c'est que l'IA travaille sur des données, beaucoup de données : pour apprendre, il lui faut le matériel nécessaire. Les données sont pour l'IA ce que les livres, articles, les enseignements sont pour nous humains.

Cependant, après avoir compris l'importance des données, il faut comprendre qu'il n'est pas chose aisée de les recueillir ses données là. Très souvent c'est la partie la plus chronophage, et des fois demande beaucoup de ressources humaines, financières et matériels.

Pour nos modèles, nous avons la chance que les données soient déjà disponibles gratuitement. La raison est que nous allons travailler avec les données financières des entreprises, et pour les entreprises cotées en bourse c'est une obligation pour eux de les publier par souci de transparence. C'est pour cela que nous pouvons nous servir sur des données des entreprises sans risque. De l'autre part, nous aurons en outre besoin de données texte pour notre chatbot, et là, c'est encore plus facile car du texte, sous n'importe quelle forme, ce n'est pas ce qui manque sur internet. Nous allons être amenés à voir comment il a été fait la collecte de nos données.

4.2.1.1 Les données d'états financiers

Le travail que nous faisons va porter sur des entreprises cotées à la Bourse Régionale des Valeurs Mobilières (BRVM). La BRVM est le marché financier de l'Union Economique et Monétaire Ouest-Africain (UEMOA). C'est ici que l'on peut échanger des actions et obligation pour le compte des entreprises et Etats de la zone UEMOA. La BRVM regroupe les 46 entreprises et banque les plus performantes de la zone.

Eu égard de cela, tout entreprise cotée à la BRVM à le devoir chaque année de publiée ces états financiers pour le compte des actionnaires, des Etats, des investisseurs, de potentiels investisseurs ou n'importe quelles personnes physiques ou morales. Leurs états financiers sont aussi publiés dans le [site de la BRVM](#) ainsi, il est possible de les télécharger et de faire notre travail, on choisit une entreprise et c'est bon.

Une fois téléchargés, les états financiers d'une seule année se présentent comme suit :

- Le rapport d'activité
- Le résultat financier
- La synthèse des rapports de gestion

Mais ses données ne sont pas pour le moment exploitables, il va falloir faire un certain nombre de transformations c'est-à-dire extraire les données qui nous intéressent, les mettre sous format CSV, JSON ou autres avant de pouvoir passer au Feature Engineering. Avec des états financiers on peut faire une analyse financière, vérifier la rentabilité de l'entreprise, prédire des valeurs etc.

4.2.1.2 Les données texte

De tous les types de données que l'on va utiliser pour des modèles intelligents, bien qu'étant les plus difficiles à exploiter, les données textes sont les plus faciles à collecter. Nous aurons besoin des données texte pour le développement du Chatbot. Notre principal objectif ici c'est de comprendre les questions que l'utilisateur du Chatbot peut poser. Bien sûr, le Chatbot que nous allons développer va être spécialisé dans nos états financiers, donc nous aurons besoin de texte qui traite de la finance des entreprises. Il y a plusieurs manières de collecter ce genre de texte :

- Les sites de finance

Il y a un certain nombre de sites web traitant de la finance en générale, et représentent une aubaine pour la collecte de données. Car labà, on peut trouver des définitions, des formules et calculs, des textes relatifs à la finance, chose qui va nous permettre d'avoir un contexte c'est-à-dire de savoir dans quelle mesure un mot va apparaître sachant qu'un autre est apparu.

- Les sites d'informations

A côté des textes relatif à la finance, il va nous falloir évidemment du texte généraliste qui traite de tout et de rien. Et les sites d'information représentent une bonne ressource pour trouver du texte de qualité. Ceci va être important pour comprendre les objets d'une question que l'utilisateur peut éventuellement posée au Chatbot.

- Les bases de données

Il y a beaucoup de *dataset* disponible dans le web, mais aussi dans les bases de données. C'est d'autre développeur qui crée ses *dataset* et qui les stockent dans des bases de données, il y a des site web comme [kaggle](https://www.kaggle.com/) qu'on peut prendre comme exemple. Dans certaines organisations comme les GAFAM (Google, Amazon, Facebook, Apple, Microsoft), elles font la collecte de données et puis paient des personnes pour faire l'étiquetage.

- Les sites de génération de textes

Les sites de génération de textes vont nous être d'utilité importante pour notre texte, ils vont nous permettre de d'avoir des données qui sont spécifiques pour notre travail. Dans le cas d'un Chatbot spécialisé, il est très difficile de trouver dans données qui nous conviennent pour le Intent classification pour des raison évidente. Donc, l'un des meilleurs moyens d'avoir des données qui sont spécifique à notre travail, c'est de générer voire même créer nos propres données.

- Le Web Scraping

Le web scraping est une technique en informatique ou l'on développe des programmes qui vont aller recueillir des informations sur le web sans l'intervention direct d'une personne. Ces programmes ressemblent dans leur fonctionnement aux robots des moteur de recherche même s'ils n'ont rien à voir l'un de l'autres, mais l'analogie était pertinente. Il est nécessaire de préciser que certaines grandes organisations comme Google n'aime vraiment pas le web scraping mais ce dernier n'est pas illégal.

Nous avons utilisé un condensé de tout cela pour avoir les données texte dont nous aurons besoin. Il y aura d'abord les textes qui vont nous permettre de faire le Intent classification, ensuite il va utiliser des techniques pour reconnaître les entités d'une question posée, pour pouvoir répondre avec la plus grande des précisions. Si vous n'avez pas compris cette dernière phrase, ne vous inquiétez pas car nous allons revenir en détail sur tout cela dans la partie où l'on parle du développement du Chatbot.

4.2.2 Prédiction des valeurs

La prédiction est un vaste domaine qui ne relève pas seulement de l'IA. En effet beaucoup de domaines scientifiques essaient de faire des prédictions sur les valeurs avec lesquelles ils travaillent, c'est notamment le cas de l'économie, la météo, la bourse etc.

Les intérêts de faire des prédictions dépendent du domaine où elles sont faites, pour ce qui nous concerne qui est le domaine de la finance d'entreprise, cela nous permet d'approximer la future santé financière de l'entreprise en se basant sur les valeurs actuelles et passées.

Dans le chapitre passé, nous avons montré les techniques intelligentes qui permettent de prédire des valeurs ici nous allons voir comment cela fonctionne en pratique.

D'abord il nous faut des données, elles sont collectées, et nous avons vu comment dans la précédente partie. Nous allons faire la prédiction de chaque élément de chaque état financier. Ce qui nous fait une centaine de prédiction à faire.

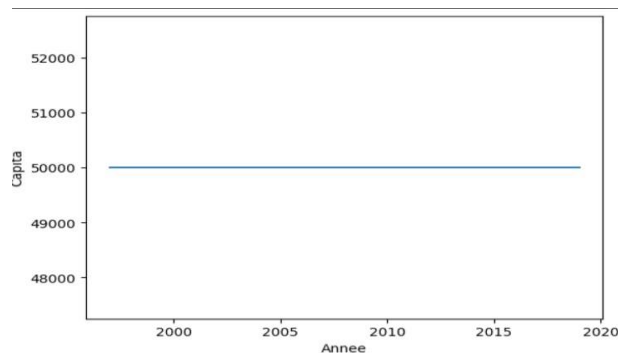
Ce que nous allons faire par la suite c'est mettre toutes les valeurs dans un fichier CSV, il faut rappeler que les documents téléchargés dans le site de la BRVM, sont sous format PDF et donc non exploitables. Une fois sous le format CSV, nous aurons trois (3) fichiers à savoir les bilans, les comptes de résultats et les tableaux des flux de trésorerie.

elements	annee 1	annee 2	annee 3	annee 4	annee 5	annee 6	annee 7	annee 8	annee 9	annee 10	annee 11
annee	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
charges immobilisees	2428	2474	2482	1877	1287	1015	102	73	84	287	1199
immobilisations incorporelles	488	628	2569	2940	8028	39169	40218	36369	42342	42633	60285
immobilisations corporelles	88698	121013	159338	189503	204902	215730	224943	246391	278246	333088	399672
immobilisations financieres	13661	13997	13680	14319	14516	12524	13147	14386	13001	13951	17901
amortissements et provisions											
total actif immobilises	105275	138065	178071	208639	228733	268438	278410	297220	333674	389960	479059
stock	2114	2119	3189	3670	6177	4445	3649	7600	5698	7264	8313
fournisseurs avances versees	2531	2896									
creances et emplois assimiles											
clients	12699	23401	29647	35347	40902	44797	37205	40202	89137	92928	83203
autres creances	37255	30080	27917	45499	46844	48647	59369	74976	38414	51425	57791
total actif circulant	52483	56378	60753	84516	93923	97889	100223	122599	133250	151617	149308
total tresorerie actif	48171	32115	26451	15756	30925	29673	43325	72576	98924	113732	114807
ecart de conversion actif	0	1059									
total actif	205929	227617	265274	308911	353581	396000	421958	492394	565849	655308	743175
capital	50000	50000	50000	50000	50000	50000	50000	50000	50000	50000	50000
primes et reserves	61619	42000	82277	93945	106190	126177	143464	154827	170239	206086	247969
report a nouveau	0	0	0	0							
ecart de conversion											
resultat net	29981	36246	40783	42512	47451	46682	55504	76312	105548	130628	140967
autres capitaux propres	38839	33126	0	0							

Figure 13 Fichier CSV des données

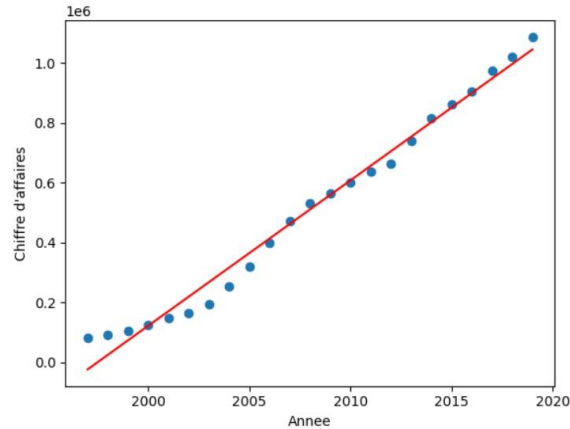
Du fait que nous avons plusieurs prédictions à faire, nous allons seulement en présenter quatre (4) du plus simple au plus intéressant.

- Le capital



Pour ce qui est du capital de cette société, nous voyons qu'il n'a pas évolué de 1997 à 2019, donc ici il n'y a pas de prédiction à faire, puisque nous savons que le capital de cette de société ne varie pas, sur toute cette période il reste à 50 000 000 000 de F CFA. Pour les valeurs futures et pour les calculs futurs qui vont faire intervenir le capital nous allons choisir cette même valeur.

- Le chiffre d'affaires

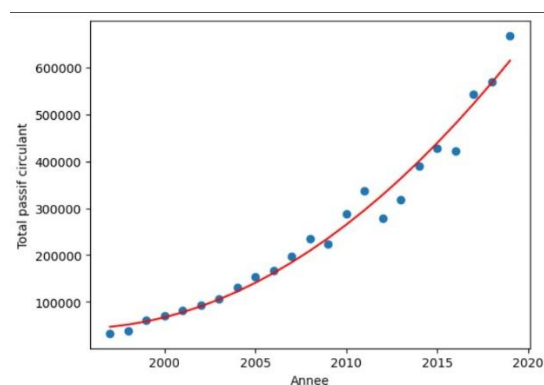


Il y a deux choses à voir ici, les points en bleu qui représentent les valeurs réelles du chiffre d'affaires en fonction des années et la droite linéaire en rouge qui est la droite de régression. La régression linéaire a bien fonctionné ici puisqu'elle épouse à la presque perfection les données. Cette droite va renvoyer le coefficient et la constante qui représentent nos w_0 et w_1 que nous avons déjà expliqué pour faire des prédictions pour les années à venir.

Le tableau ci-dessous nous donne pour la droite de régression du chiffre d'affaires son coefficient, sa constante et la prédiction de quelques valeurs futures.

Coefficient	Constante	2025	2030	2035
48578.23616	-97034585.22134	1336343.0148221	1579234.1956521	1822125.376482

- Le passif circulant

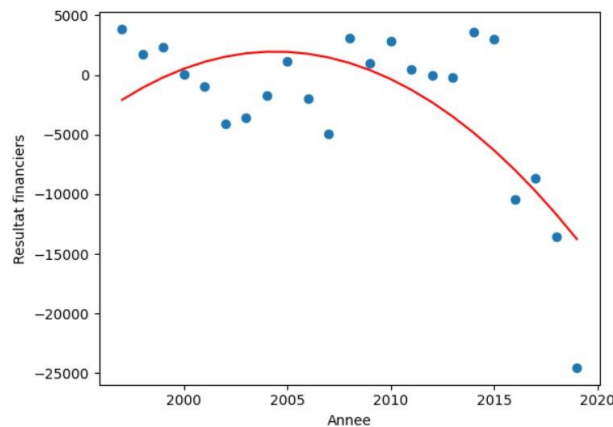


Ce cas est un tantinet plus intéressant que le précédent les données ne suivent pas une forme linéaire, alors il nous faut autre chose. Cette autre chose c'est évidemment la régression

polynomiale que nous avons aussi déjà vue. Pour ce cas de figure nous n'aurons pas seulement w_0 et w_1 mais aussi w_2 et w_3 (le biais), puisque le degré de notre polynôme est égal à 2.

W0	W1	W2	W3	Année A
0	-4.00030285e+06	1.00251237e+03	3.99062354e+09	$W_0 + W_1 * A + W_2 A^2 + W_3$

- Le résultat financier



Le modèle des résultats financiers est l'un des plus intéressants pour trois (3) raisons : d'abord les valeurs sont très dispersées par rapport aux autres, ensuite les valeurs décroissent de manière exponentielle et enfin il y a des valeurs négatives. C'est normal qu'il y ait des valeurs négatives pour le résultat financier dans la mesure où les charges financières sont supérieures aux produits financiers. Nous pourrions être tentés de penser que ce modèle va être plus compliqué que les autres, mais il n'en est rien. Nous allons simplement le faire passer dans une régression polynomiale et la magie va opérée. Le résultat pour ce modèle est la suivante.

W0	W1	W2	W3	Année A
0	2.95998776e+05	-7.3836900e+01	-2.9664942e+08	$W_0 + W_1 * A + W_2 A^2 + W_3$

Il faut noter que pour tous les modèles, nous allons suivre exactement les mêmes procédures que ces quatre, tous les autres vont se classer dans l'un de ces cas de figure.

Maintenant que nous avons prédit nos modèles, il est possible de les déployer pour utiliser dans une interface graphique. Pour ce modèle, le déploiement peut être simple du fait qu'il y a seulement des nombres. On peut les stocker dans un fichier et l'appeler depuis le frontend.

4.2.3 Le développement du Chatbot

L'année 1964 à 1967, fut le temps nécessaire pour développer le premier Chatbot nommé ELIZA par un scientifique de MIT. Ce Chatbot était capable de tenir une conversation en utilisant un algorithme de reconnaissance de mot, ce qui implique qu'il ne comprenait pas vraiment le texte mais rechercher des mots clés qu'il utilisait pour répondre.

Mais c'est quoi un Chatbot ? C'est un programme informatique comme les autres mais qui a des aspects bien particulier. Il sait mener une conversation qui se rapproche de celle des humains. Au début les Chatbots étaient créés pour reproduire le comportement humain chez les machines mais aussi faire avancer le NLP mais aujourd'hui les Chatbots sont là pour nous aider dans nos vies.

Comment marche les Chatbots ? Le fonctionnement des Chatbot est simple, c'est comme une conversation entre deux personnes, l'un parle l'autre écoute puis répond. Cependant dans le cas des Chatbot, très souvent c'est, je pose une question et le Chatbot réponds, mais il est possible d'avoir des banalités avec les Chatbot. Tout dépend de comment les scientifiques ont décidé de développer le Chatbot. Il faut préciser que même si le fonctionnement est simple, le développement ne l'est pas et nous allons voir pourquoi.

Mettons-nous à la place d'un Chatbot qui reçoit une question, comment ferions-nous pour répondre, oui c'est évident pour nous humains, mais pour un programme informatique c'est très compliqué. C'est là que l'on va introduire les deux (2) concepts de base de tous les Chatbot : les *intents* et le *entity detection*.

4.2.3.1 Les Intents

Ils représentent l'objet de la question ou la phrase ou tout ce que l'utilisateur donne comme input au Chatbot : c'est la phase de compréhension de la question. Ci-après quelques phrases et leurs objets :

Comment vas-tu ? **Salutation**

Quel temps va-t-il faire demain ? **Prédiction**

Pourquoi le gouvernement ne peut-il pas imprimer plus de monnaie ? **Explication**

Nous avons trois questions et trois objets différents, c'est-à-dire que chaque question s'attend à une réponse différente. Encore une fois pour nous humains cela paraît évident mais ce n'est pas le

cas pour les machines. Maintenant comment les machines font pour comprendre l'objet d'une question, il faut appeler quelqu'un que nous connaissons déjà : le Machine Learning. Tout à l'heure nous avons parlé de la collecte de données texte, maintenant nous allons les utiliser en faisant ce que l'on appelle un *intent classification*.

Nous allons faire un apprentissage supervisé, avec les données texte qui vont être étiquetées de leur objet. Les trois phases avec leur objet que nous avons donné représentent aussi des exemples d'à quoi va ressembler le *dataset*. Ce qui suit c'est la représentation de texte, transformer le texte en un format compréhensible par la machine. Après avoir fait cela, nous le faisons passer dans un algorithme de Machine Learning, si nous avons une bonne *accuracy*, donc le modèle est bon, notre programme est capable maintenant de comprendre le texte. L'*intent classification* est un moyen très puissant qui nous permet de comprendre l'objet d'une question posée par un utilisateur.

4.2.3.2 L'*entity detection*

Maintenant que l'on comprend le sens des questions il faut bien répondre, mais pour répondre il nous faut des arguments, c'est qu'y a-t-il dans la question qui peut nous permettre de répondre, deux cas de figure se présentent. Il y a des questions où nous pouvons répondre directement sans recherche d'arguments comme quand l'utilisateur salue ou remercie le Chatbot. S'il dit « Salut » ou « merci beaucoup », si le Chatbot parvient à bien détecter que l'un a pour objet une salutation et l'autre un remerciement, alors nous pouvons renvoyer la réponse associée. Cependant si l'utilisateur pose la question suivante : « Quel sera notre chiffre d'affaires de 2026 », et que le Chatbot a bien compris que l'objet de cette question est une prédiction, c'est là, que va intervenir l'*entity detection*.

« Quel sera notre chiffre d'affaires de 2026 », quelle sont les éléments de cette question que nous allons utiliser pour y répondre. D'abord il y a le chiffre d'affaires, c'est l'élément que l'on veut prédire et aussi l'année 2026, c'est la période sur lequel on veut faire une prédiction. C'est cela l'*entity detection*, maintenant une fois que nous avons relevé chiffre d'affaires et 2026 de la question, on appelle la fonction qui permet de faire les prédictions en lui passant les arguments, et après on envoie la réponse à l'utilisateur.

Il peut y arriver que l'utilisateur pose une question ou il n'y a pas tous les arguments mais le modèle parvient quand même à comprendre l'objet, exemple : « quelle est la prédiction de l'actif total », c'est une prédiction mais il n'y a pas de période. Dans ce cas, le Chatbot va lui envoyer un

message pour préciser la période ou l'élément sur lequel va être faite la prédiction si c'est ce dernier qui manque.

Parlons un peu des réponses, ce n'est vraiment pas la partie la plus compliquée dans le processus de développement. Une fois que l'on a passé avec brio les étapes d'*intents* et le *entity detection* et que nous avons aussi appelé les fonctions appropriées, on renvoie des réponses fixes préalablement définies. Maintenant il est possible d'avoir plusieurs réponses et renvoyer une au hasard pour avoir cet aspect moins robotique. Les réponses fixes sont appropriées pour notre programme sinon il faut utiliser des algorithmes générateurs de textes qui va rendre la chose beaucoup plus compliquée.

4.2.3.3 La modélisation du texte

Maintenant, il y a une phase qui précède le *intent classification* et le *entity detection*, c'est la représentation de texte. L'ordinateur ne comprend pas le texte et ne peut pas faire des calculs sur du texte, or nos données sont de type texte et devons les comprendre, comment faire ? il y a un certain nombre de méthodes qui nous permet de modéliser le texte en nombre, c'est-à-dire sous un format que l'ordinateur va comprendre.

- Label encoding

Le *label encoding* est une technique de représentation de texte où l'on va associer chaque mot de tout de le *dataset* a un nombre bien particulier. Ce qui implique que nous aurions autant de *label* que de mot.

Si nous prenons la phrase « L'argent fait le bonheur », nous aurons : L' => 1, argent => 2, fait => 3, le => 4, bonheur => 5. Par conséquent notre phrase devient « 1 2 3 4 5 » et nous pouvons faire des calculs sur ce résultat.

- One hot encoding

Pour ce qui est du *one hot encoding*, il est plus sophistiqué que le *label encoding* mais un peu compliqué. Ici nous allons créer un tableau a deux dimensions, autant de ligne que de mots. Pour ce qui est de la colonne, elle sera égale au nombre de mot dans le texte, pour un mot, il va être noté de 1 à la place qui lui est représenté dans la colonne et toutes les autres valeurs de 0. Prenons encore la phrase « L'argent fait le bonheur » comme exemple.

Tableau 5 One hot encoding

	<i>L'</i>	<i>Argent</i>	<i>Fait</i>	<i>Le</i>	<i>Bonheur</i>
<i>L'</i>	1	0	0	0	0
<i>Argent</i>	0	1	0	0	0
<i>Fait</i>	0	0	1	0	0
<i>Le</i>	0	0	0	1	0
<i>Bonheur</i>	0	0	0	0	1

Le problème qui se pose ici, c'est ce qu'on appelle le *curse of dimensionality* ou la malédiction de la dimension, ce qui veut dire les dimensions vont facilement augmenter si le texte augmente. Imaginons un texte de 5000 mots, nous aurions 5000 éléments dans la colonne.

- Bag of Word

Le *bag of word* s'inscrit dans le même cadre que le OHE mais pour apporter des améliorations à ce dernier. Ici la taille de la colonne ne varie pas c'est les lignes qui vont changer. La différence OHE c'est ici on ne modélise pas un mot mais une phrase. Le fonctionnement c'est qu'après avoir mis tous les mots en colonne, on compte le nombre de fois que ce mot apparaît dans une phrase.

Exemple : « L'argent fait le bonheur », « Le bonheur fait le bonheur », « L'argent est le pouvoir »

Tableau 6 Bag of word

	<i>L'</i>	<i>Argent</i>	<i>Fait</i>	<i>Le</i>	<i>Bonheur</i>	<i>Est</i>	<i>Pouvoir</i>
<i>L'argent fait le bonheur</i>	1	1	1	1	1	0	0
<i>Le bonheur fait le bonheur</i>	0	0	1	1	2	0	0
<i>L'argent est le pouvoir</i>	1	0	1	0	0	1	1

Il y a moins d'entrée est plus de signification, le grand plus que l'on a ici c'est on modélise par phrase mais pas par mot. Le BOW connaît aussi le *curse of dimensionality*, la taille de la colonne est proportionnelle aux données. En plus cela il y a aussi le problème du OOV (Out Of Vocabulary) pour des mots qui peut arriver dans les données testées et qui étaient absents dans les données d'entraînement.

- Bag of n-grams

Le *bag of n-grams* est une variation voire une amélioration du BOW, dans ce dernier on comptait mot par mot mais avec le *bag of n-grams*, on compte deux mots par deux mots, ou trois mots par trois mots, enfin n mots par n mots : on parle de n grams. Chose qui nous permet de donner du sens au texte : « je suis » est plus compréhensible que « je » et « suis ».

Tableau 7 Bag og n-grams

	<i>L'argent</i>	<i>Argent</i> <i>fait</i>	<i>Fait</i> <i>le</i>	<i>Le</i> <i>bonheur</i>	<i>Argent</i> <i>est</i>	<i>Est le</i>	<i>Le</i> <i>pouvoir</i>
<i>L'argent fait le</i> <i>bonheur</i>	1	1	1	1	0	0	0
<i>L'argent est le</i> <i>pouvoir</i>	1	0	0	0	1	1	1

Nous voyons ici, on compte par un mot cela fait plus de sens que de compter par mots. Mais il n'est pas interdit de mettre plusieurs grams. Si le gram est égal à 1 donc on revient au BOW.

- TF-IDF

Le TF-IDF qui signifie *Term Frequency – Inverse Document Frequency* est une méthode qui calcule la fréquence d'un mot sur le document total. Pour ce faire, nous allons faire deux calcul TF et le IDF, ensuite nous allons prendre le produit des deux.

$$TF = \frac{\text{occurrence mot}}{\text{nombre total mot}}$$

$$IDF = \frac{\text{nombre total document}}{\text{nombre document qui contient le mot}}$$

$$TF - IDF = TF * IDF$$

A la suite de cela nous obtenons une seule valeur pour chaque mot, ce qui est plus facile à travailler avec. Le TF-IDF quant à lui souffre seulement du OOV.

- Word Embedding

Le *word embedding* est un ensemble de méthode qui propose des moyens de transformer le texte en vecteur. Le plus populaire d'entre eux c'est le word2vec qui à partir d'un mot lui trouve son vecteur assimilé. Le bénéfice du word2vec c'est que des mots qui se ressemblent vont avoir des vecteurs qui se ressemblent, il est possible de trouver la ressemblance entre deux en calculant la similarité cosinus.

Exemple : père : [1, 1, 0, 0], mère : [0.9, 1, 0.01, 0], maison : [0.2, 0.54, 2, -0.5], bâtiment : [0.21, 0.5, 1.9, -0.49]. Nous voyons ici que les vecteurs de père et de mère se ressemblent et ceux de maison et bâtiment se ressemblent.

Maintenant les valeurs de ces vecteurs ne tombent pas du ciel, ils sont trouvés avec le Machine Learning en faisant ce qu'on appelle le *Self Supervises Learning*.

- Fast text

Dans le *bag of n-grams*, on fait on comptait l'occurrence des mots dans une phrase, cela peut bien marcher mais souffre du OOV. Le *fast text* est venu pour palier à cela, ce qu'il fait par contre c'est de diviser le mot en *grams* et faire la représentation de chaque *gram*.

Si nous prenons le bonheur, avec 3 *grams* : {bo, bon, onh, nhe, heu, eur, ur} et le mot bonheur lui-même. Ce qui aura pour effet, si un mot qui ressemble à bonheur et qui n'était pas dans le *dataset*, nous pouvons travailler avec.

Il peut y arriver des fois qu'on est besoin de faire l'inverse, c'est-à-dire après avoir fait la représentation de texte et nos calculs, on obtient le résultat en nombre, de ce fait il faut faire l'opération inverse : transformer les nombres en texte.

4.2.3.1 La gestion des réponses

Une fois toutes ces parties gérées, il faut bien répondre à l'utilisateur, c'est le sens même d'un Chatbot. Pour ce qui concerne notre Chatbot, nous l'avons évoquée tout a l'heure, ce sera des réponses fixes et aléatoire. Pour chaque *intent*, nous aurons un certain nombre de réponses qui seront associée, c'est rendre plus agréable l'expérience de l'utilisateur.

Comment marche les réponses fixes mais aléatoire ? Dans le cas d'une salutation par exemple, nous allons avoir 3 réponses comme suit : « bonjour comment allez-vous », « bien le bonjour mon très cher », « je vous salue ». A partir de là, une fois une salutation détectée, nous lui renvoyons une de ces trois (3) réponses au hasard.

4.2.3.2 Exemple de fonctionnement du Chatbot

Maintenant donnons l'exemple de fonctionnement du Chatbot, ce que nous allons faire c'est d'entrer dans la "tête" du Chatbot qui reçoit une question, la traite et renvoie la réponse.

Utilisateur : quelle est la valeur présente du total actif ?

- La première chose à faire c'est la représentation de texte, utilisons pour cette phrase le BOW et nous aurons un vecteur qui compte les mots qui se trouve dans cette phrase.

- Après avoir obtenu le vecteur nous allons le faire passer dans un algorithme de Machine Learning pour obtenir l'*intent*, le Naive Bayes en l'occurrence, et la sortie un « get » c'est recevoir une valeur dans la base de données.
- Après l'*intent*, on passe au *entity detection*, ici nous comme entité l'actif total et « présente » qui signifie l'année actuelle, nous aurons (actif total, 2024), ce sera nos paramètres pour la fonction de « get ».
- Ensuite on appelle la fonction avec les paramètres qui nous renvoie la valeur dont on a besoin, disons 5000.
- Nous avons tout ce qui nous faut pour répondre, de toutes les réponses de « get » fixe on en choisi une par hasard, que l'on concatène avec la valeur et la réponse est prête.

Chatbot : la valeur actuelle du total actif est de 5000.

Ci-après des exemple chat de notre Chatbot avec un utilisateur.

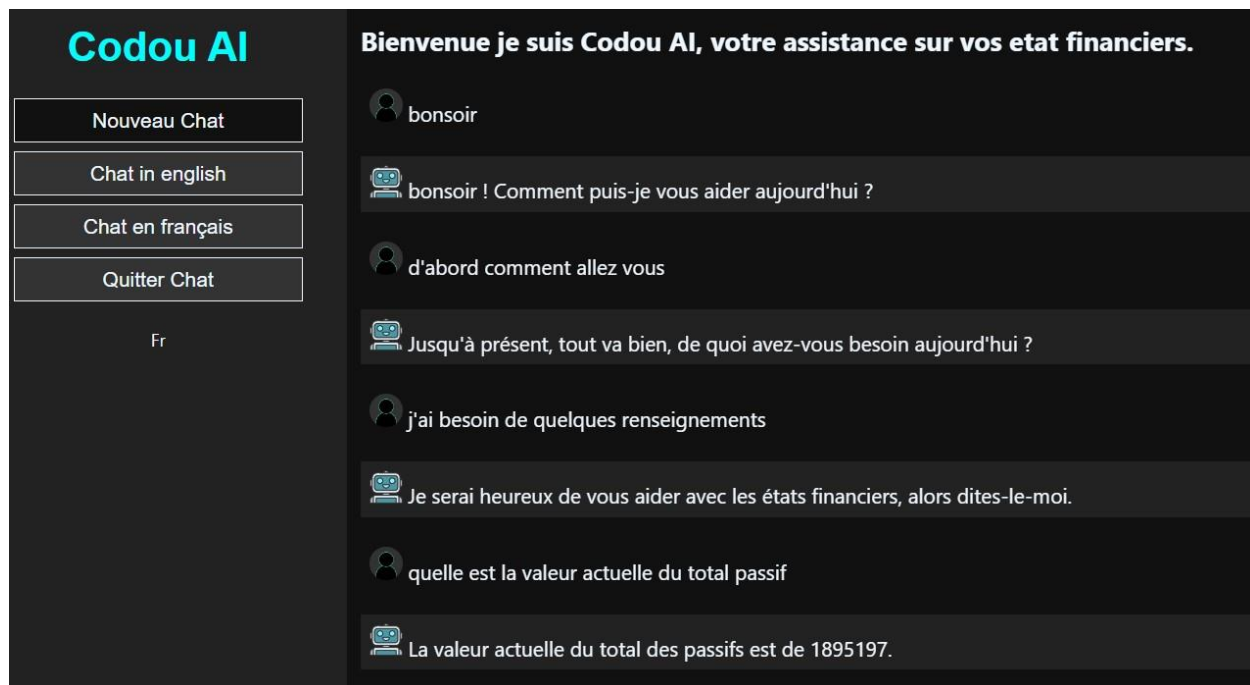


Figure 14 Conversation 1

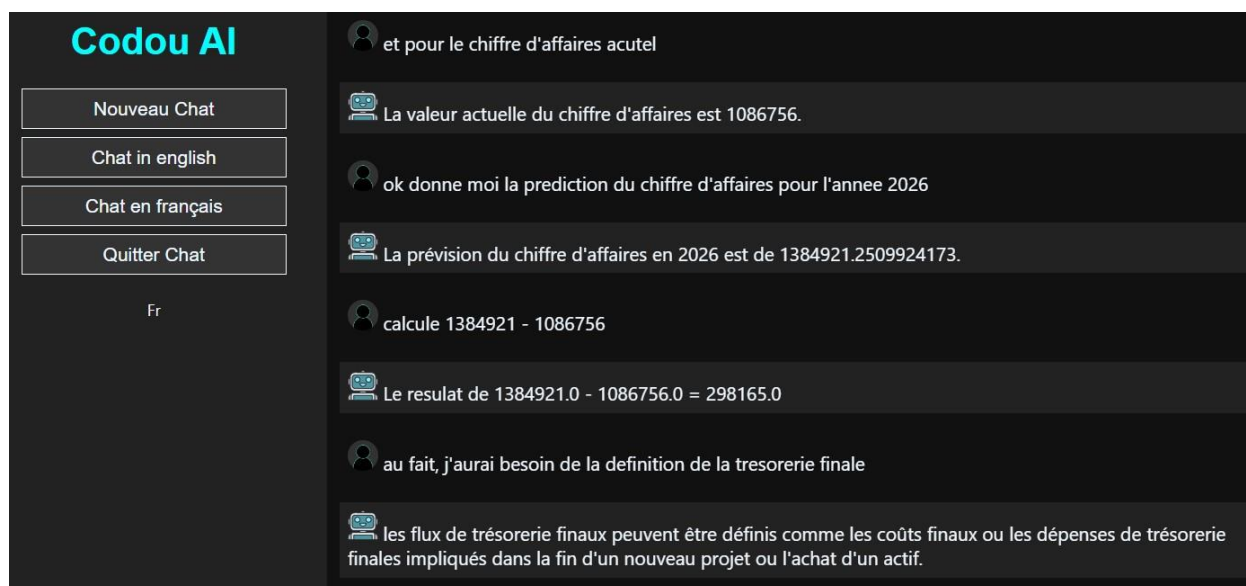


Figure 15 Conversation 2

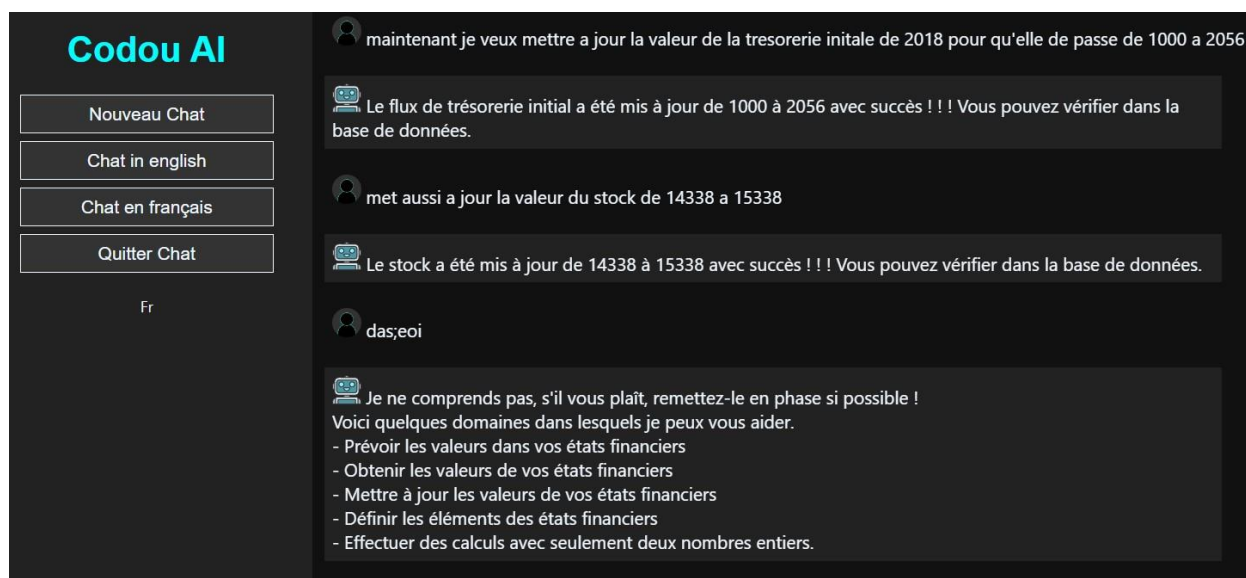


Figure 16 Conversation 3

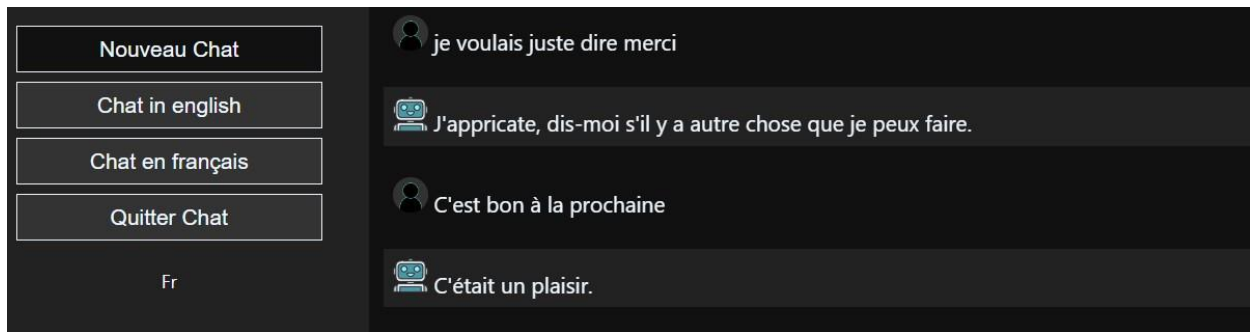


Figure 17 Conversation 4

Nous pouvons conclure cette partie sur le Chatbot en disant pour son développement la question est plus importante que la réponse, si nous parvenons à comprendre parfaitement la question, nous parviendrons à donner la bonne réponse.

4.3 L'interface graphique

A ce moment de notre travail, tout a déjà été fait, ou presque. Oui les prédictions ont été faites, oui le chatbot a été développé mais il subsiste un grand problème : à moins être un ingénieur en IA, vous ne pouvez pas interagir ou utiliser les modèles. C'est ce problème que nous allons essayer de résoudre ici en développant une interface graphique que n'importe quel utilisateur peut utiliser pour interagir avec les modèles. Pour ce faire il y aura deux phases : faire de Graphical User Interface (GUI) et lier les modèles avec le GUI à travers un web service.

4.3.1 Développement de l'interface graphique

L'interface graphique va être développée avec la bibliothèque Qt/C++, qui est un Framework qui nous facilite cette tâche en utilisant le langage C++. Nous avons utilisé cette bibliothèque car l'application n'a pas pour vocation d'être déployé sur internet, c'est une organisation qui va utiliser une instance de celle-ci.

4.3.1.1 Le frontend

Qt nous donne un certain nombre de fonctionnalités pour réaliser des logiciels robustes et très avancés mais nous allons en prendre quelques-uns, qui nous ont été utiles dans notre travail, pour les expliquer et faire comprendre comment nous avons utilisé dans nos modèles.

- Widget

Les Widgets font partie des éléments de base sur Qt, au fait tout est un widget et il y a des Widgets pour tout faire. Pour mieux comprendre ce concept, un Widget est un élément visible que l'utilisateur peut interagir avec. Nous pouvons donner l'exemple d'un bouton, une zone de texte, un label ou voir même une page entière.

L'un des Widget les plus importante c'est ce que Qt appelle QMainWindow qui est le Widget principale sur lequel tous des autres vont s'accoler. Il faut préciser chaque Widget a un parent sauf le Widget principale. Ce qui nous donne une structure sous forme d'arbre.

Dans notre logiciel nous allons utiliser toute sorte de Widget commençant bien évidemment par le QMainWindow en passant par les autres qui permettent de présenter le résultat de nos modèles d'une manière simple.

- Layout

Le deuxième élément sur lequel nous allons un peu nous attarder ce sont les Layout. Ils représentent tous les manières de disposer nos Widgets. Par défaut, un Widget va empiler tous ses enfants sur son point (0, 0) en haut en gauche dans un ordinateur.

Avec l'aide des Layouts, nous pouvons avoir plusieurs manières de présenter les Widgets, ceci aura pour effet d'avoir un logiciel plus sérieux. Il y a trois sortes de Layout sur Qt : QVBoxLayout (disposition verticale), QHBoxLayout (disposition horizontale) et QGridLayout (disposition en grille). Ses trois héritent tous de la classe mère QLayout.

Bien que nous ayons plus utilisé les dispositions verticales et horizontales dans notre logiciel, la disposition en grille peut être très pertinente par moments. C'est impressionnant de se dire quand même que tous les frontends que nous voyons dans les logiciels ne sortent pas de ces trois Layout.

- Signal and Slot

Les Signal and Slot sont les deux concepts de la programmation événementielle que Qt met à notre disposition. Un Signal c'est un signal c'est-à-dire un déclencheur, cela peut être une clique, un focus, un survol un changement quelconque. Il y a des Signal prédéfinis comme ceux que nous venons de citer, mais nous avons la possibilité d'en créer des nouveaux. Néanmoins, les prédéfinis vont nous suffire dans la plupart des cas.

Les Slot ce sont les fonctions que nous allons utiliser pour réagir par rapport au Signal. Mais pourquoi ce ne sont pas des méthodes comme les autres ? C'est que les Signal n'acceptent que les Slot comme fonction, une méthode qui ne portent pas la mention Slot ne va pas être acceptée dans la programmation événementielle de Qt.

Nous allons illustrer cela en donnant l'exemple de l'utilisateur qui veut ouvrir les paramètres. S'il clique sur le bouton paramètre (Signal), l'application va ouvrir la boîte de dialogue paramètre (Slot). Ce qui intéressant avec Qt, c'est qu'il nous permet de faire cela en une seule ligne de code avec la fonction « connect » de la classe QObject.

4.3.1.2 Le Backend

Le backend d'une application c'est tout ce que l'utilisateur ne voit pas, et qui fait toute la puissance d'une application. Tout ce qui est base de données, classes, appelle de fonctions sont classer dans le backend. Pour ce qu'il s'agit de nos applications nous aurons une base de données MySQL, des classes qui vont interagir avec la base de données.

- La base de données

Nous allons avoir une base de données locale qui va nous servir à stocker les valeurs des états financiers. Il y aura trois tables dans la base de données à savoir le bilan, le compte de résultat et aussi le tableau des flux de trésorerie et chaque enregistrement de chaque table va représenter une année. C'est une base de données simple, nous allons rarement y insérer des données ou même modifier. Dans la plupart des cas il sera seulement fait des requêtes SELECT pour obtenir des données.

- Le contrôleur

Même si nous ne sommes pas vraiment dans un modèle MVC (Model, View, Controller), car nous ne pouvons-nous permettre de laisser les requêtes directement accéder à notre base de données, cela nous permet aussi une bonne gestion des données. Ce modèle va se présenter sous forme de classe, ici aussi nous aurons trois (3) classes pour nos trois états financiers. Pour chaque classe, nous aurons les méthodes qui renvoient les états financiers, les analyses verticales, les getters et les setters pour chaque élément des états financiers.

- Serveur

Avant de pouvoir utiliser une base de données, il nous faut un SGBD (système de gestion de base de données) qui va nous permettre d'interagir avec la base, nous l'avons déjà dit c'est MySQL que nous allons utiliser avec XAMPP, qui est un serveur qui nous permet la connexion avec une base de données MySQL, Apache etc.

4.3.1.3 *Le Web server*

Un web server est un programme informatique qui nous permet une encapsulation de données avec les protocoles http ou https. Avec un web server on peut déployer une fonction qui peut être invoquée depuis plusieurs Endpoint.

De là, on peut voir comment cela peut nous être utile, nous avons déjà développé nos modèles intelligents et ce qui suit c'est de les lier avec l'interface graphique. Les modèles ont été développés en langage Python et l'interface graphique avec C++, il nous faut les lier. C'est là qu'intervient la puissance des web server. Il est possible d'écrire un Application Programming Interface (API) dans un langage A et de l'invoque dans un langage B avec l'aide d'un navigateur.

Il y a deux types de web server : étendu avec SOAP et REST, nous allons utiliser REST avec la bibliothèque de Python FastAPI.

Nous pouvons donner l'exemple d'un cas d'utilisateur du chatbot. L'utilisateur pose sa question depuis l'application Qt, une requête REST va être invoquée avec la question comme paramètre, le web server va le prendre le traiter et renvoyer la réponse sous format JSON ou XML.

Cette méthode est aussi appelée le développement multi-tiers il y a deux programmes différents (deux tiers) mais ils peuvent communiquer tout en étant indépendant l'un de l'autre. Ci-après une image qui illustre l'architecture du processus global.

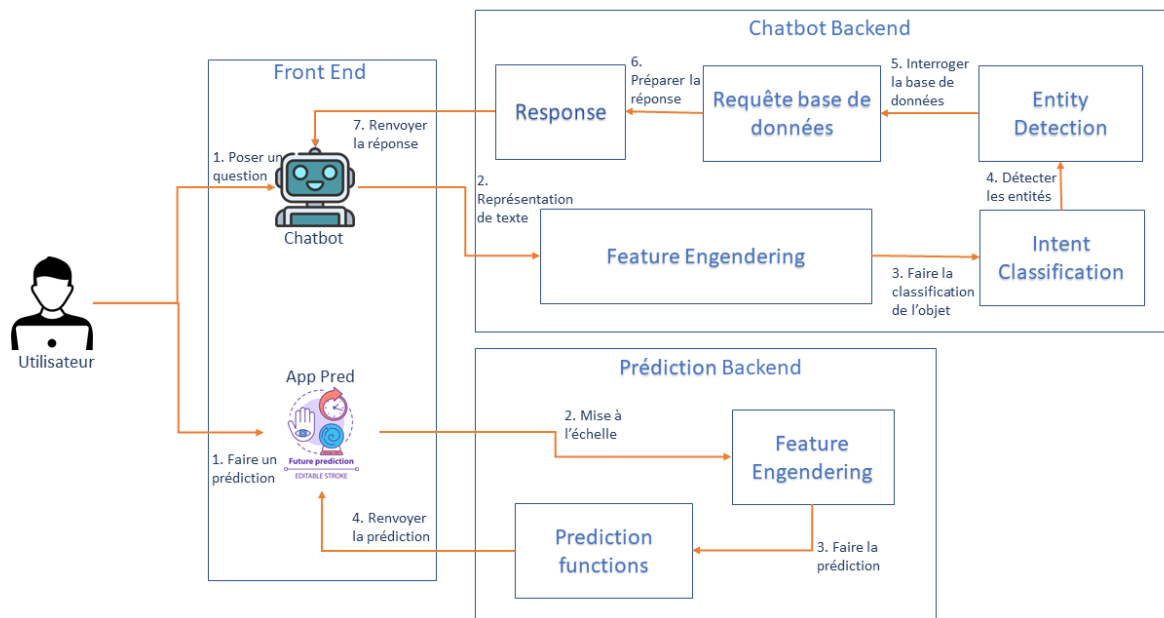


Figure 18 Architecture globale

4.3.2 Présentation de l'application

Toute l'explication a été faite donc ici nous allons juste montrer à quoi ressemble l'application, l'interface graphique de l'utilisateur. Toutes les pages ne seront pas montrées ici mais seulement les plus pertinentes, celle que l'utilisateur va ouvrir très souvent.

- La page d'accueil

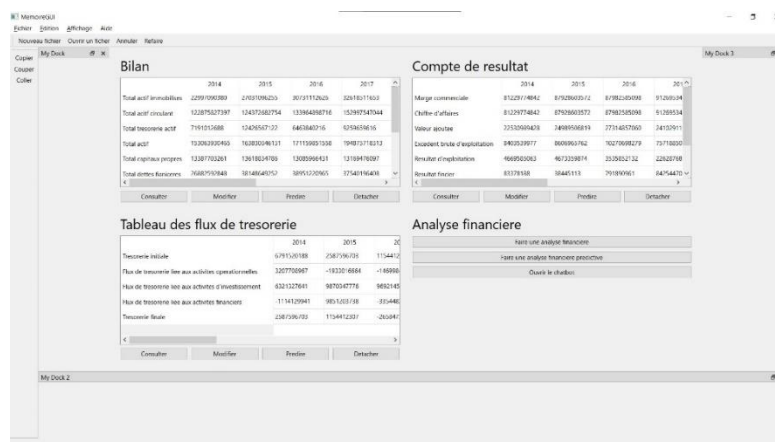


Figure 19 Page d'accueil de l'application

- La page de consultation du bilan

Bilan

Consultation le bilan

	2014	2015	2016	2017	2018	2019
Charges immobilisées	0	0	0	0	0	0
Immobilisations incorporelles	286952160	192546166	439667421	397422185	252255188	318452127
Immobilisations corporelles	20564739901	24755667996	28306320625	30371654656	33761301118	36047194228
Immobilisations financières	2145398319	2082882093	1985124580	1849434812	1715752598	1562352707
Amortissements et provisions	0	0	0	0	0	0
Total actif immobilisé	22997090380	27031096255	30731112626	32618511653	35729308904	37927999062
Stock	9407671633	12298331651	12869667591	14549513975	16037718059	15948665233
Fournisseurs avances versées	6572848542	7751081969	7738661652	5185483044	6212187549	9334294925
Créances et emplois assimilés	113468155764	112074351103	121095231125	138448033069	152671403410	178155854030
Clients	103127713047	100397198170	108915111440	114901946159	126280091310	127642859261
Autres créances	3767594175	3926070964	4441458033	18360603866	20179124551	41178699844
Total actif circulant	122875827397	124372682754	133964898716	152997547044	168709121469	194104519263
Total trésorerie actif	7191012688	12426567122	6463840216	9259659616	6890452333	6459808043
Total actif	153063930465	163803046131	171159851558	194875718313	211328882706	238492326368
Capital	4500000000	4500000000	4500000000	4500000000	4500000000	4500000000
Primes et réserves	255906342	255906342	255906342	255906342	255906342	255906342

Figure 20 La page consultation du bilan de l'application

- La page de modification du tableau des flux de trésorerie

Tableau des flux de trésorerie

Modification du tableau des flux de trésorerie

	2014	2015	2016	2017	2018	2019
TRESORERIE INITIALE	6791520188	2587596703	1154412307	-26584733756	-42778772996	-42138213035
Capacité d'autofinancement	7001397269	7482326992	9616314007	8281469494	9841388668	9632364195
Variation actif circulant HAO	0	0	0	0	0	0
Variation des stocks	2426496245	2890660018	571335940	1679046304	1488204084	-89052826
Variation des créances	1262329938	-1393304661	9020880022	17352801944	14223370341	25484450620
Variation du passif circulant	-104862119	-7918487399	-14723940911	6686987628	16351016628	22054153868
Variation du BFR	-3793688302	-9415342756	-24316156873	-12345660700	639442203	-3341243926
FLUX DE TRESORERIE DES ACTIVITES OPERATIONNELLES	3207708967	-1933016664	-14699842866	-4064191206	10490831071	6291120269
Décaissements liés aux acquisitions d'immobilisations incorporelles	83874778	43743069	448360065	53489849	40573334	172397066
Décaissements liés aux acquisitions d'immobilisations corporelles	6470102885	9906937006	9355144972	8070054614	9303251118	8254193576
Décaissements liés aux acquisitions d'immobilisations financières	12914761	139399776	104158489	66226234	669548829	422305239
Encaissements liés aux acquisitions d'immobilisations incorporelles et corporelles	32148781	17816073	13602521	42301189	38845956	16741133
Encaissements liés aux acquisitions d'immobilisations financières	213416002	201916002	201916002	201916002	201916002	201916002

Valider

Figure 21 La page modification du TFT de l'application

- La partie analyse des états financiers

Faire l'analyse financière

Analyse financière simple

Verification des états financiers | Analyse des états financiers | Analyse de l'activité et des relations de trésorerie | Analyse tendancielle et la méthode des ratios

Bilan
Compte de résultat
Tableau des flux de trésorerie

Analyse du bilan

Analyse verticale | Analyse horizontale

	2014	2015	2016	2017	2018	2019
Charges immobilisées	0%	0%	0%	0%	0%	0%
Immobilisations incorporelles	0.187472%	0.117528%	0.256875%	0.203936%	0.119366%	0.133527%
Immobilisations corporelles	13.4354%	15.1106%	16.5379%	15.5851%	15.9757%	15.1146%
Immobilisations financières	1.40164%	1.27137%	1.15361%	0.949033%	0.811887%	0.655096%
Amortissements et provisions	0%	0%	0%	0%	0%	0%
Total actif immobilisé	15.0245%	16.4994%	17.9546%	16.7381%	16.907%	15.9032%
Stock	6.14624%	7.50675%	7.51909%	7.46605%	7.58899%	6.68729%
Fournisseurs avances versées	4.29419%	4.73116%	4.52131%	2.66092%	2.93958%	3.91388%
Créances et emplois assimilés	74.1312%	68.4008%	70.7498%	71.0443%	72.2435%	74.7009%
Clients	67.3756%	61.2812%	63.6336%	58.9617%	59.7552%	53.5207%
Autres créances	2.46145%	2.39642%	2.59492%	9.4217%	9.54868%	17.2663%

Figure 22 La page de l'analyse financière de l'application

- La partie analyse prédictive de la rentabilité sur les 3 prochaines années

4.4 Résultats et discussions

4.4.1 Présentation des résultats

4.4.2 Limites et défis

4.4.3 Améliorations possibles

Conclusion partielle

Nous laissons à l'appréciation du lecteur de juger si le travail de l'implémentation a été bien fait ou pas. Ce qui est sûr c'est nous avons utilisé toutes les armes que nous avons en main pour mener à bien cette partie. Bien que n'étant pas la partie qui nous a posé le plus de problèmes mais c'est définitivement d'une des plus intéressantes, car nous savions ce que nous faisons et là il n'y a rien de mieux que de coder.

Dans cette partie nous avons développé des modèles intelligents qui nous ont permis de prédire les états financiers avec un bon degré de confiance mais surtout le chatbot qui était le grand challenge de ce travail de mémoire. Il est intéressant de préciser qu'il n'est pas obligatoire pour un data scientist de développer des GUI, car très souvent le travail de ce dernier s'arrête au développement des modèles mais aussi le web server. Néanmoins, c'est toujours intéressant de maîtriser ces genres de chose.

Ce travail de mémoire a été nécessaire car il nous a appris tellement de choses, nous nous sommes confrontés à la réalité du métier chose qui nous a forgés. Nous avons aussi découvert le trajet complet que les scientifiques prennent dans leur travail, à savoir de la collecte de données jusqu'au premier test de validité.

Conclusion générale et perspectives

Pour synthétiser ce travail dans lequel nous nous étions lancés, il faut rappeler la question que ce mémoire a pour but de répondre à savoir dans quelle mesure l'intelligence artificielle va pouvoir aider les entreprises cotées à la Bourse régionale des Valeurs Mobilières (BRVM) à faciliter leur analyse financière et tout ce qui tourne au tour de leurs états financiers. Cette question centrale de

recherche nous a amené à explorer les dimensions de l'IA qui pourrons nous permettre d'améliorer la performance des entreprises de l'UEMOA en leur prédisant leur santé future.

Par conséquent, répondre à cette question va, à notre sens, être évident car l'IA peut aider la finance à bien des égards. C'est ainsi que nous avons choisi la prédiction de valeur qui est un domaine de prédilection de l'IA et de l'appliquer à la finance, mais surtout le NLP pour le développement de Chatbot. Il y a bien d'autres domaines d'application de l'IA sur la finance mais pour un début ses deux peuvent s'avérer être suffisants.

Parlons des résultats, au début nous avions une centaine de prédiction à faire à savoir tous les éléments du bilan, du compte de résultat et du tableau des flux de trésorerie. Et chaque élément nécessitait une attention particulière, la raison est que les données ne se ressemblent jamais. Nous avons utilisé la régression linéaire, mais aussi la régression polynomiale par moments pour faire les prédictions. Comme nous l'avons vu certains éléments ont été faciles de travailler avec, mais d'autres plus compliquées. Il faut aussi préciser que les données avec lesquelles nous avons travaillé sur la prédiction des modèles nous viennent du site de la BRVM.

Pour ce qui s'agit du chatbot, il y avait trois (3) niveau de difficulté d'abord il fallait trouver des données avec lesquelles il faut travailler, puis faire ce qu'on appelle un *Intent Classification*, enfin finir de faire le *Entity Detection*. La collecte de données de type texte ne fut pas un challenge de taille puisque le texte est disponible en quantité et en qualité, là où cela devient intéressant c'est quand il faut transformer le texte en un format compréhensible par l'ordinateur et par ailleurs un modèle d'IA. Nous avons vu qu'il y avait plusieurs moyens de faire cette représentation de texte, et une fois ce travail fait on peut le faire passer au *Intent Classification*. Là aussi plusieurs modèles de Machine Learning s'offrent à nous, nous avons fait la recherche en grille et choisi celui qui donne le plus grand *accuracy*. L'*Entity Detection* n'est pas encore une fois bien compliqué, nous avons une liste de tous les éléments des états financiers, nous faisons juste une recherche de ses éléments et aussi de la période si nécessaire.

Et bien évidemment, après tout ce travail, il faut donner une interface de communication aux utilisateurs. C'est ce que nous avons fait pour terminer le travail pratique de développement, l'interface a été fait sous forme de logiciel qui regroupe les deux applications, c'est-à-dire celle des prédictions et le chatbot.

En guise de perspectives, nous avons descellé plusieurs mais ici, va être listé les plus importants. D'abord parlons de la gestion des fichiers, si cela venait à être implémentée, il serait possible d'avoir des fichiers pour plusieurs années. Il serait aussi intéressant d'avoir la possibilité de faire une rédaction complète et rigoureuse de rapport d'analyse. Pour que le chatbot soit moins robotique, augmenter les *Intents* va s'avérer être une bonne idée, ce serait vraiment intéressant si le chatbot pouvait nous répondre sur différent domaine de la finance. Il serait aussi intéressant de ne plus avoir des réponses fixes comme l'état actuel des choses, mais générer des réponses, c'est le travail de la *Generative AI*, il y a les LSTM (Long Short Term Memory) qui peuvent nous aider par rapport à cela. Actuellement, pour entrer ou modifier des valeurs il faut le faire manuellement, il est possible d'automatiser cela en implémentant des CNN qui vont aller directement récupérer les informations sur des fichiers PDF voir même des images. Le travail que nous avons fait ici s'applique à l'analyse financière, nous pouvons aussi intégrer d'autres domaines de la finance pour avoir un produit encore plus puissant, d'autres domaines comme la finance de marché, la finance d'entreprise, l'analyse du risque etc. Voici de manière résumée les perspectives de l'application.

- La gestion des fichiers
- Rédaction complète et rigoureuse d'un rapport d'analyse (RNN)
- Utiliser les CNN pour extraire les informations directement sur les fichiers PDF
- Pousser le travail avec du Deep Learning sur toute l'étendue de la finance d'entreprise
- Augmenter les *Intents* pour avoir un Chatbot plus performant
- Générateur de texte pour les réponses (LSTM)

Ce qui met fin à ce travail passionnant de mémoire, nous avons appris tellement de choses en rapport avec l'IA et la finance. Ceci nous motive à poursuivre nos études de recherche pour pouvoir creuser encore plus profond dans le domaine de l'IA.

Bibliographies

- Piaget, J. (2005). *The psychology of intelligence*. Routledge.
- Howard, W. R. (1993). On What Intelligence Is. *British Journal of Psychology*, 84(1), 27-37.
- Legg, S., & Hutter, M. (2007). A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and applications*, 157, 17.
- Bianchini, F. A. (2023). New Definition of “Artificial” for Two Artificial Sciences. *Found Sci* **28**, 401–417. <https://doi.org/10.1007/s10699-021-09799-w>
- Simmons, A. B., & Chappell, S. G. (1988). Artificial intelligence-definition and practice. *IEEE journal of oceanic engineering*, 13(2), 14-42.
- Ginsberg, M. (2012). *Essentials of artificial intelligence*. Newnes.
- Sheikh, H., Prins, C., Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In: Mission AI. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2
- Kutyniok, G. (2022). The mathematics of artificial intelligence. arXiv preprint arXiv:2203.08890.
- Brette, R. (2003). *Modeles impulsionnels de réseaux de neurones biologiques* (Doctoral dissertation, Université Pierre et Marie Curie-Paris VI).
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- Collins, G. S., & Moons, K. G. M. (2019). Reporting of artificial intelligence prediction models. *The Lancet*, 393(10181), 1577–1579. doi:10.1016/s0140-6736(19)30037-6
- Miah, M. B. A., Hossain, M. Z., Hossain, M. A., & Islam, M. M. (2015). Price prediction of stock market using hybrid model of artificial intelligence. *International Journal of Computer Applications*, 111(3).
- Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy*, 47, 1-6.

- Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In IFIP international conference on artificial intelligence applications and innovations (pp. 373-383). Springer, Cham.
- Okuda, T., & Shoda, S. (2018). AI-based chatbot service for financial industry. *Fujitsu Scientific and Technical Journal*, 54(2), 4-8.
- Muhammad Farman, M. A. (2023, Novembre). Artificial Intelligence for fraud detection and prevention. ResearchGate.
- Choi, D., & Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. *Security and Communication Networks*, 2018(1), 5483472.
- Song, M., Xing, X., Duan, Y., Cohen, J., & Mou, J. (2022). Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention. *Journal of Retailing and Consumer Services*, 66, 102900.
- Buhalis, D. and Moldavska, I. (2022), "Voice assistants in hospitality: using artificial intelligence for customer service", *Journal of Hospitality and Tourism Technology*, Vol. 13 No. 3, pp. 386-403. <https://doi.org/10.1108/JHTT-03-2021-0104>
- Alex Avelar, E., & Jordão, R. V. D. (2024). The role of artificial intelligence in the decision-making process: a study on the financial analysis and movement forecasting of the world's largest stock exchanges. *Management Decision*.
- Yang, N. (2022). Financial big data management and control and artificial intelligence analysis method based on data mining technology. *Wireless Communications and Mobile Computing*, 2022(1), 7596094.
- Mbuli Landu, A. (2021). Chapitre 3. La trésorerie actif. Dans : , A. Mbuli Landu, *Le mémo d'un comptable: Approche par le SYSCOHADA révisé* (pp. 301-324). Paris: L'Harmattan.
- Altman, E. I. (2005). An emerging market credit scoring system for corporate bonds. *Emerging Markets Review*, 6(4), 311–323. doi:10.1016/j.ememar.2005.09.007
- Ahmad, B. (2020, Juillet). Basic design of a neural network.

Cayla, B. (2021, Mars 7). The Stochastic Gradient Descent (SGD) & Learning Rate.

Friedman, D. (2020). Relu Activation.

Graph of polynomial functions. (2020, Juin 20).

Kanade, V. (2022, Setembre 29). A Pictorial Representation of the Reinforcement Learning Model.

Kumar, V. (2020, Mai 4). Neural Network.

mathisfun. (n.d.). Introduction to Derivatives.

Oman, S. (2016, Janvier 11). Xor representation.

Poudel, S. (2023, Aout 28). RNN Unfolded.

Reche, J. (2019, Octobre). 7 : Schématisation d'un neurone biologique. Paris.

Rooij, P. V. (2020, Decembre 23). The four key questions for comprehensive knowledge management.

Saleem, M. (2023, Juillet). The standard logistic function.

Shahriar, N. (2023, Fevrier 1). Convolutional Neural Network — CNN architecture.

Zhao, O. (2021, Mars 24). What is a Decision Tree.

Webographies

<https://www.lafinancepourtous.com/decryptages/entreprise/gestion-et-comptabilite/comptes-de-l-entreprise/comprendre-le-bilan-le-compte-de-resultat-et-l-annexe/le-bilan/>

Tout savoir sur la trésorerie passive de l'entreprise. (2024, 5 2). Récupéré sur AGICAP: <https://agicap.com/fr/article/tresorerie-passive-definition-utilisation/>

Marge commerciale : définition simple, exemple, calcul. (2024, 2 6). Retrieved from JDN: <https://www.journaldunet.fr/business/dictionnaire-comptable-et-fiscal/1198457-marge-commerciale-definition-exemple-formule/>

Marge commerciale : définition simple, exemple, calcul. (2024, 2 6). Récupéré sur JDN: <https://www.journaldunet.fr/business/dictionnaire-comptable-et-fiscal/1198457-marge-commerciale-definition-exemple-formule/>

MARCHAL, J. (2024, 2 8). *Comment élaborer un tableau de flux de trésorerie ? Intérêts et analyse.* Retrieved from L'expert comptable: <https://www.l-expert-comptable.com/a/6312-comment-elaborer-un-tableau-de-flux-de-tresorerie-interets-et-analyse.html#:~:text=flux%20de%20tr%C3%A9sorierie%20%3F-D%C3%A9finition,ann%C3%A9e%20etc...>

Annexes

Annexe 1 : Exemple d'analyse financière sur Excel

Analyse verticale du compte de résultat

Analyse du compte de resultat						
Analyse verticale						
Ventes de marchandises	0.03%	0.04%	0.04%	0.05%	0.06%	0.09%
Achats de marchandises	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Variation de stocks	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Marge brute sur marchandises	0.03%	0.04%	0.04%	0.05%	0.06%	0.09%
Marge brute sur matières	82.13%	82.49%	81.06%	0.00%	0.00%	0.00%
Marge commerciale	0.03%	0.04%	0.04%	0.05%	0.06%	0.09%
Ventes de produits fabriqués	54.71%	56.31%	61.39%	60.26%	58.00%	64.50%
Travaux, services vendus	43.44%	41.73%	36.31%	37.55%	39.84%	32.70%
Produits accessoires	1.82%	1.92%	2.26%	2.15%	2.10%	2.71%
Chiffre d'affaires	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Production stockée (ou déstockage)	-0.11%	0.06%	0.15%	1.10%	0.79%	-0.88%
Production immobilisée	3.03%	2.39%	1.95%	2.17%	3.88%	5.57%
Subventions d'exploitation	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Autres produits	3.10%	3.03%	1.93%	0.46%	0.58%	0.50%
Transferts de charges d'exploitation	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Achats des matières premières et fournitures	15.57%	20.69%	18.75%	20.26%	23.79%	22.95%

Analyse horizontale du compte de résultat

Analyse du compte de resultat						
Analyse verticale						
Analyse horizontale						
Ventes de marchandises		50.29%	-12.19%	27.17%	50.99%	29.46%
Achats de marchandises						
Variation de stocks						
Marge brute sur marchandises	↑	50.29%	↓	-12.19%	↑	27.17%
Marge brute sur matières	→	8.72%	↓	-1.67%	↓	-100.00%
Marge commerciale	↑	50.29%	↓	-12.19%	↑	27.17%
Ventes de produits fabriqués		11.42%	9.08%	1.83%	3.80%	9.01%
Travaux, services vendus		3.99%	-12.93%	7.26%	14.43%	-19.56%
Produits accessoires		13.98%	18.33%	-1.51%	5.44%	26.55%
Chiffre d'affaires	→	8.25%	→	0.06%	→	3.74%
Production stockée (ou déstockage)		-161.43%	135.33%	674.47%	-22.78%	-209.02%
Production immobilisée		-14.84%	-18.17%	15.49%	92.56%	40.67%
Subventions d'exploitation						
Autres produits		5.84%	-36.23%	-75.35%	35.31%	-15.41%
Transferts de charges d'exploitation						
Achats des matières premières et fournitures liées		43.78%	-9.33%	12.14%	26.63%	-5.44%

Analyse verticale du bilan

Stocks et encours	6.15%	7.51%	7.52%	7.47%	7.59%	6.69%
Fournisseurs, avances versées	4.29%	4.73%	4.52%	2.66%	2.94%	3.91%
Clients	67.38%	61.28%	63.63%	58.96%	59.76%	53.52%
Autres créances	2.46%	2.40%	2.59%	9.42%	9.55%	17.27%
Créances et emplois assimilés	74.13%	68.41%	70.75%	71.04%	72.24%	74.70%
Total actif circulant	80.28%	75.92%	78.27%	78.51%	79.83%	81.39%
Titres de placement	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Valeurs à encaisser	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Banques, chèques postaux, caisses et ass	4.70%	7.59%	3.78%	4.75%	3.26%	2.71%
Total trésorerie actif	4.70%	7.59%	3.78%	4.75%	3.26%	2.71%
Ecart de conversion actif	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total Actif	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Dettes fiscales et sociales	0.00%	0.00%	0.00%	0.00%	8.44%	7.19%
Dettes fiscales	30.42%	22.06%	12.15%	5.80%	0.00%	0.00%
Dettes sociales	2.26%	2.27%	1.78%	1.98%	0.00%	0.00%
Autres dettes	3.09%	2.91%	3.15%	2.80%	2.60%	2.97%
Provision pour risques à court terme (risque)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total Passif circulant	70.68%	61.52%	50.29%	47.28%	51.33%	54.73%
Banques, crédits d'escompte	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Banques, établissements financiers et crédits de trésorerie	3.01%	6.88%	19.31%	26.70%	23.20%	22.24%
Total Trésorerie passif	3.01%	6.88%	19.31%	26.70%	23.20%	22.24%
Ecart de conversion - Passif	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total passif	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Analyse horizontale du bilan

Dettes fiscales et sociales	0.00%	0.00%	0.00%	0.00%	8.44%	7.19%
Dettes fiscales	30.42%	22.06%	12.15%	5.80%	0.00%	0.00%
Dettes sociales	2.26%	2.27%	1.78%	1.98%	0.00%	0.00%
Autres dettes	3.09%	2.91%	3.15%	2.80%	2.60%	2.97%
Provision pour risques à court terme (risque)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total Passif circulant	70.68%	61.52%	50.29%	47.28%	51.33%	54.73%
Banques, crédits d'escompte	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Banques, établissements financiers et crédits de trésorerie	3.01%	6.88%	19.31%	26.70%	23.20%	22.24%
Total Trésorerie passif	3.01%	6.88%	19.31%	26.70%	23.20%	22.24%
Ecart de conversion - Passif	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total passif	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Total Passif circulant	↓	-6.84%	↓	-14.60%	↑	7.03%	↑	17.75%	↑	20.33%
Banques, crédits d'escompte										
Banques, établissements financiers et crédits de trésorerie		144.87%	↑	193.19%	↑	57.46%	↓	-5.78%	→	8.18%
Total Trésorerie passif	↑	144.87%	↑	193.19%	↑	57.46%	↓	-5.78%	→	8.18%
Ecart de conversion - Passif										
Total passif	→	7.03%	→	4.47%	↑	13.86%	→	8.44%	↑	12.85%

Analyse verticale du tableau des flux de trésorerie

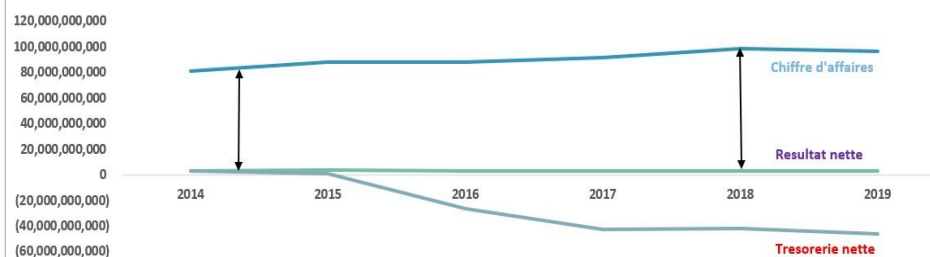
Flux de trésorerie provenant des capitaux	-93.91%	-265.07%	12.19%	6.31%	6.41%	5.80%
(+) Emprunts	22.15%	958.02%	-6.18%	0.00%	0.00%	0.00%
(+) Autres dettes financières	57.03%	236.46%	-9.79%	-9.85%	-9.04%	-4.95%
(-) Remboursement des emprunts et autres dettes financières	28.33%	76.06%	-16.40%	-11.84%	-2.79%	-3.67%
Flux de trésorerie provenant des emprunts	50.85%	1118.42%	0.43%	1.99%	-6.25%	-1.28%
Flux de trésorerie provenant des activités	-43.06%	853.35%	12.62%	8.30%	0.16%	4.51%
Variation de la trésorerie nette de l'exercice	-162.46%	-124.15%	104.34%	37.86%	-1.52%	9.53%
Trésorerie nette au 31 décembre	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Analyse horizontale du tableau des flux de trésorerie

Flux de trésorerie provenant des capitaux propres	25.93%	5.88%	-16.67%	0.00%	0.00%
(+) Emprunts	1829.42%	-85.15%	-100.00%		
(+) Autres dettes financières	84.98%	-4.61%	61.84%	-9.63%	-39.42%
(-) Remboursement des emprunts et autres dettes financières	19.78%	396.61%	16.16%	-76.78%	45.32%
Flux de trésorerie provenant des emprunts	881.19%	-100.89%	643.20%	-409.39%	-77.29%
Flux de trésorerie provenant des activités de financement	-984.21%	-134.05%	5.85%	-98.09%	3006.95%
Variation de la trésorerie nette de l'exercice	-65.91%	1835.49%	-41.62%	-103.96%	-793.34%
Trésorerie nette au 31 décembre	-55.39%	-2402.88%	60.91%	-1.50%	10.54%

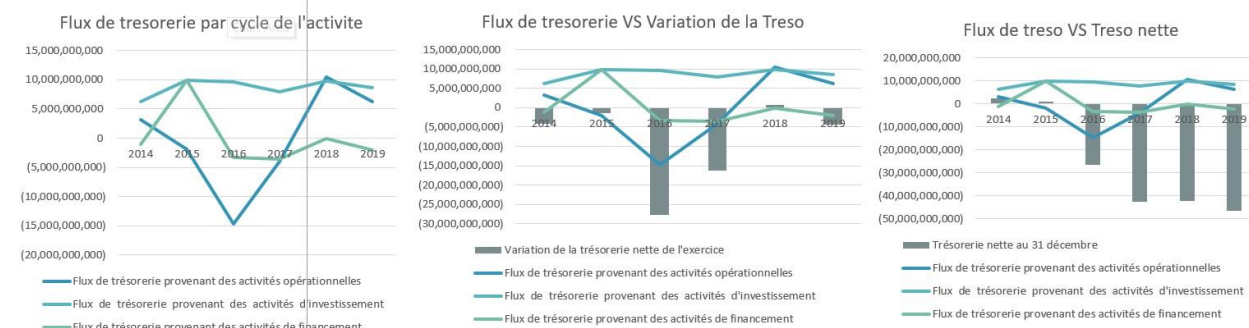
Analyse du cycle de vie de l'activité

Chiffre d'affaires	81,229,774,842	87,928,603,572	87,982,585,098	91,269,534,900	98,432,703,111	96,477,917,195
Resultat net	3,117,919,358	3,291,131,525	2,707,131,645	2,722,631,753	2,863,674,115	2,952,343,608
Tresorerie nette	2,587,596,703	1,154,412,307	(26,584,733,756)	(42,778,772,996)	(42,138,213,035)	(46,579,488,294)



Analyse du comportement des flux de trésorerie

Flux de trésorerie provenant des activités opérat	3,207,708,967	(1,933,016,664)	(14,699,842,866)	(4,064,191,206)	10,480,831,071	6,291,120,2
Flux de trésorerie provenant des activités d'in	6,321,327,641	9,870,347,776	9,692,145,003	7,945,553,506	9,772,611,323	8,630,238,7
Flux de trésorerie provenant des activités de fin	(1,114,129,941)	9,851,203,738	(3,354,482,017)	(3,550,829,906)	(67,659,787)	(2,102,156,7
Variation de la trésorerie nette de l'exercice	(4,203,923,485)	(1,433,184,396)	(27,739,146,063)	(16,194,039,240)	640,559,961	(4,441,275,2
Tresorerie nette au 31 décembre	2,587,596,703	1,154,412,307	(26,584,733,756)	(42,778,772,996)	(42,138,213,035)	(46,579,488,2



Analyse des équilibres financiers et la relation de trésorerie

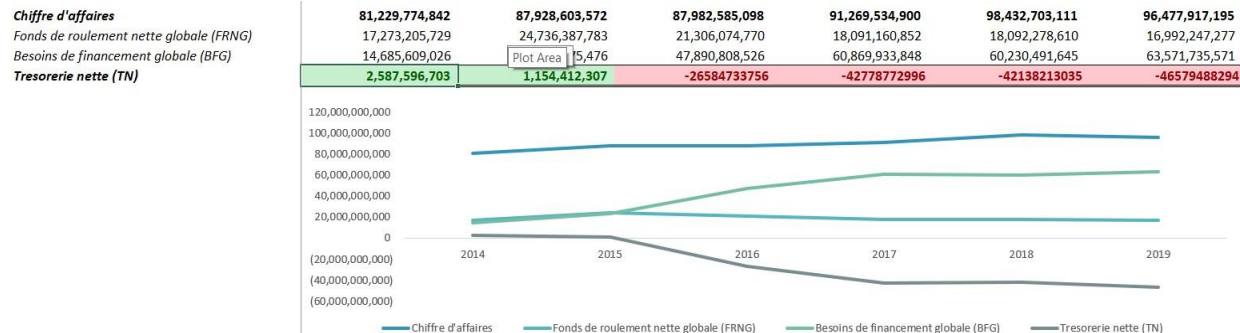
Bilan économique

Actif immobilisee (valeur nette)	22,997,090,380	27,031,096,255	30,731,112,626	32,618,511,653	35,729,308,904	37,927,999,062
Besoin de financement global (BFG)	14,685,609,026	23,581,975,476	47,890,808,526	60,869,933,848	60,230,491,645	63,571,735,571
Actif économique	37,682,699,406	50,613,071,731	78,621,921,152	93,488,445,501	95,959,800,549	101,499,734,633
Capitaux propres et ressources assimilées	13,387,703,261	13,618,834,786	13,085,966,431	13,169,476,097	13,727,894,920	14,461,067,669
Endettement nette	24,294,996,145	36,994,236,945	65,535,954,721	80,318,969,404	82,231,905,629	87,038,666,964
Capitaux investis	37,682,699,406	50,613,071,731	78,621,921,152	93,488,445,501	95,959,800,549	101,499,734,633
Verification	OK	OK	OK	OK	OK	OK

Les équilibres financiers

Fonds de roulement nette globale (FRNG)	17,273,205,729	24,736,387,783	21,306,074,770	18,091,160,852	18,092,278,610	16,992,247,277
Besoins de financement globale (BFG)	14,685,609,026	23,581,975,476	47,890,808,526	60,869,933,848	60,230,491,645	63,571,735,571
Tresorerie nette (TN)	2,587,596,703	1,154,412,307	-26584733756	-42778772996	-42138213035	-46579488294

La relation de trésorerie



Analyse tendancielle et la méthode des ratios

Profitabilité

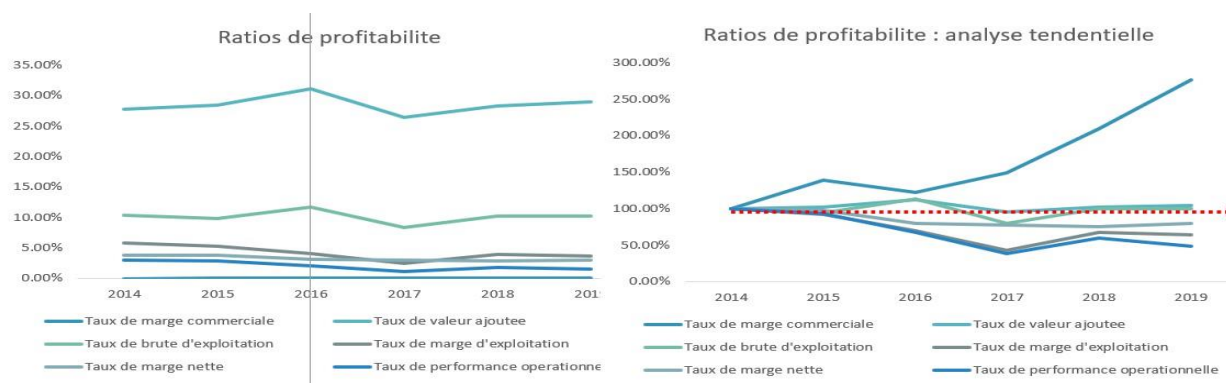
Calcul des ratios

Taux de marge commerciale	0.03%	0.04%	0.04%	0.05%	0.06%	0.09%
Taux de valeur ajoutée	27.74%	28.42%	31.05%	26.41%	28.27%	28.99%
Taux de brute d'exploitation	10.35%	9.79%	11.67%	8.30%	10.28%	10.29%
Taux de marge d'exploitation	5.75%	5.31%	4.02%	2.48%	3.89%	3.67%
Taux de marge nette	3.84%	3.74%	3.08%	2.98%	2.91%	3.06%
Taux de performance operationnelle	3.05%	2.85%	2.07%	1.16%	1.81%	1.48%

Analyse tendancielle

Taux de marge commerciale	100.00%	138.84%	121.83%	149.36%	209.11%	276.20%
Taux de valeur ajoutée	100.00%	102.46%	111.93%	95.21%	101.92%	104.51%
Taux de brute d'exploitation	100.00%	94.62%	112.84%	80.19%	99.35%	99.43%
Taux de marge d'exploitation	100.00%	92.46%	69.91%	43.13%	67.60%	63.76%
Taux de marge nette	100.00%	97.51%	80.16%	77.72%	75.79%	79.72%
Taux de performance operationnelle	100.00%	93.50%	67.72%	38.06%	59.33%	48.60%

Représentation graphique



Rentabilité

Calcul des ratios

Rentabilite economique (methode directe)	8.36%	6.80%	2.87%	1.85%	2.83%	2.64%
Rentabilite economique (methode indirecte)	8.36%	6.80%	2.87%	1.85%	2.83%	2.64%
Taux de marge d'exploitation	3.88%	3.91%	2.56%	1.89%	2.76%	2.77%
Taux de rotation de l'actifs economique	2.16x	1.74x	1.12x	.98x	1.03x	.95x

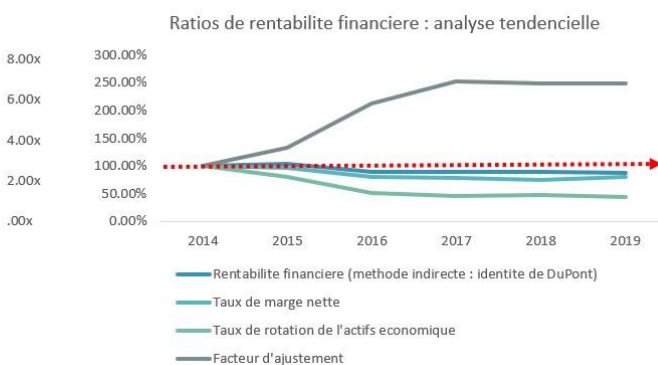
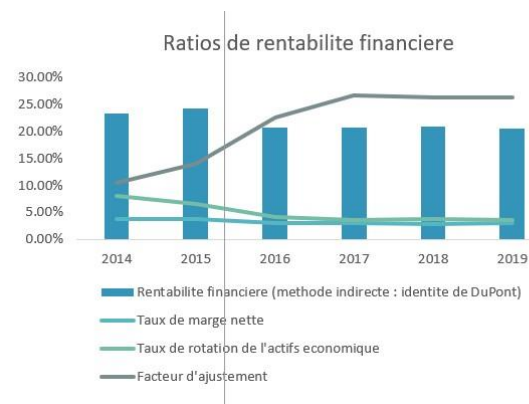
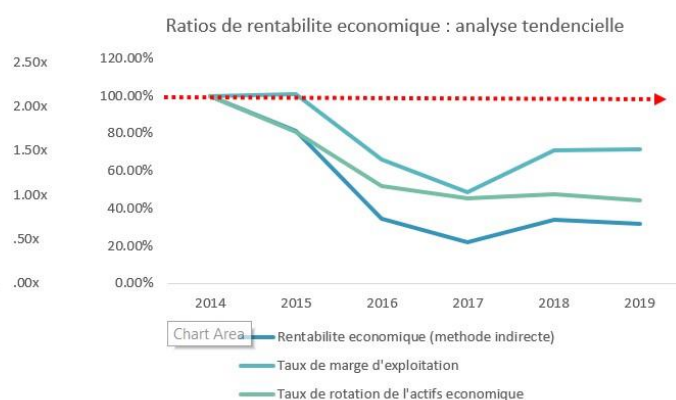
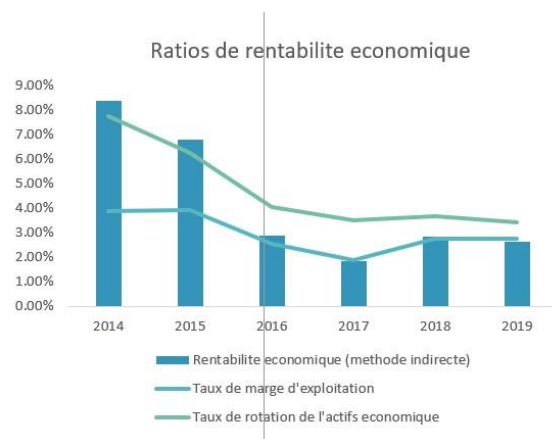
Rentabilite financiere (methode directe)	23.29%	24.17%	20.69%	20.67%	20.86%	20.42%
Rentabilite financiere (methode indirecte : identite de DuPont)	23.29%	24.17%	20.69%	20.67%	20.86%	20.42%
Taux de marge nette	3.84%	3.74%	3.08%	2.98%	2.91%	3.06%
Taux de rotation de l'actifs economique	2.16x	1.74x	1.12x	.98x	1.03x	.95x
Facteur d'ajustement	2.81x	3.72x	6.01x	7.10x	6.99x	7.02x

Analyse tendancielle

Rentabilite economique (methode directe)	100.00%	81.25%	34.26%	22.12%	33.80%	31.53%
Rentabilite economique (methode indirecte)	100.00%	81.25%	34.26%	22.12%	33.80%	31.53%
Taux de marge d'exploitation	100.00%	100.82%	65.99%	48.84%	71.03%	71.51%
Taux de rotation de l'actifs economique	100.00%	80.59%	51.91%	45.29%	47.59%	44.10%

Rentabilite financiere (methode directe)	100.00%	103.76%	88.83%	88.77%	89.57%	87.66%
Rentabilite financiere (methode indirecte : identite de DuPont)	100.00%	103.76%	88.83%	88.77%	89.57%	87.66%
Taux de marge nette	100.00%	97.51%	80.16%	77.72%	75.79%	79.72%
Taux de rotation de l'actifs economique	100.00%	80.59%	51.91%	45.29%	47.59%	44.10%
Facteur d'ajustement	100.00%	132.03%	213.45%	252.20%	248.34%	249.36%

Représentation graphique



Politique comptable

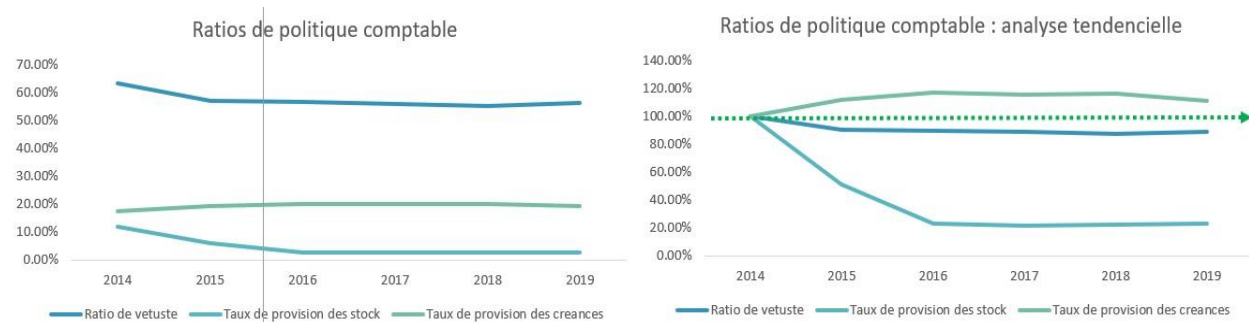
Calcul des ratios

Ratio de vetuste	63.11%	57.07%	56.57%	55.93%	55.17%	56.09%
Taux de provision des stock	11.94%	6.11%	2.77%	2.58%	2.71%	2.76%
Taux de provision des creances	17.29%	19.26%	20.16%	19.92%	20.14%	19.17%

Analyse tendancielle

Ratio de vetuste	100.00%	90.43%	89.63%	88.62%	87.42%	88.89%
Taux de provision des stock	100.00%	51.13%	23.20%	21.58%	22.65%	23.12%
Taux de provision des creances	100.00%	111.43%	116.63%	115.22%	116.53%	110.88%

Représentation graphique



Liquidité

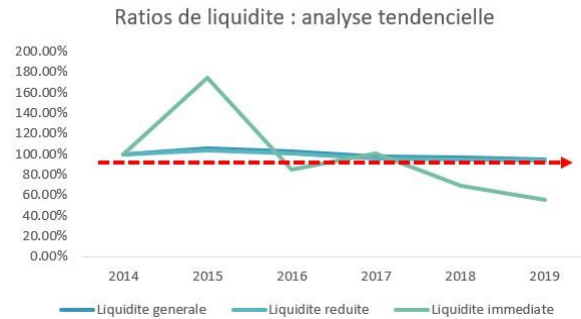
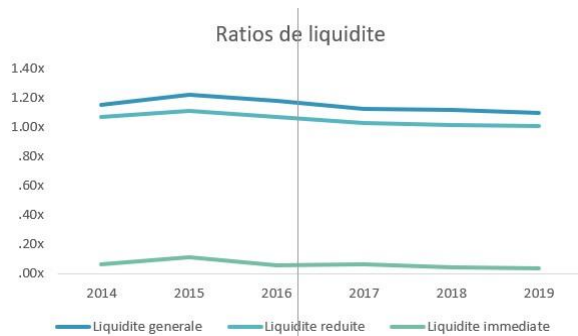
Calcul des ratios

Liquidite generale	1.15x	1.22x	1.18x	1.13x	1.11x	1.09x
Liquidite reduite	1.07x	1.11x	1.07x	1.02x	1.01x	1.01x
Liquidite immediate	.06x	.11x	.05x	.06x	.04x	.04x

Analyse tendancielle

Liquidite generale	100.00%	105.86%	102.23%	97.60%	96.68%	94.75%
Liquidite reduite	100.00%	103.86%	100.10%	95.78%	94.70%	94.01%
Liquidite immediate	100.00%	173.93%	85.11%	100.75%	68.62%	55.20%

Représentation graphique



Analyse de la probabilité de défauts

Tableau 8 Ratios de probabilité de défauts

Ratios	Poids
$\frac{\text{Constante}}{\text{besoin de financement globale}}$	1
$\frac{\text{total actifs}}{\text{report à nouveau}}$	6,56
$\frac{\text{total actifs}}{\text{resultat d'exploitation}}$	3,26
$\frac{\text{total actifs}}{\text{capitalisation boursiere}}$	6,72
$\frac{\text{total actifs}}{\text{total actifs}}$	1.05

$$zscore = \sum_{i=1}^n XiWi$$

Équation 44 Zscore d'Altman

Tableau 9 Zones d'Altman

Zscore	Rating	Zone
> 8,15	AAA	Zone de sécurité
7,60 - 8,15	AA+	
7,30 - 7,60	AA	
7,00 - 7,30	AA-	
6,85 - 7,00	A+	
6,65 - 6,85	A	
6,40 - 6,65	A-	
6,25 - 6,40	BBB+	
5,85 - 6,25	BBB	
5,65 - 5,85	BBB-	Zone d'incertitude
5,25 - 5,65	BB+	

		4,95 - 5,25	BB				
		4,75 - 4,95	BB-				
		4,50 - 4,75	B+				
		4,15 - 4,50	B				
		3,75 - 4,15	B-				
		3,20 - 3,75	CCC+				
		2,50 - 3,20	CCC			Zone de détresse	
		1,75 - 2,50	CCC-				
		< 1,75	D				
Parametre	Coef						
Constante	3.25	1	1	1	1	1	1
X1	6.56	0.096	0.144	0.280	0.312	0.285	0.267
X2	3.26	0.004	0.004	0.004	0.004	0.004	0.004
X3	6.72	0.031	0.029	0.021	0.012	0.018	0.015
X4	1.05	0.537	0.571	0.444	0.273	0.165	0.106
Zscore		4.662	5.000	5.705	5.676	5.427	5.222
Implication		Zone d'incertitude	Zone d'incertitude	Zone d'incertitude	Zone d'incertitude	Zone d'incertitude	Zone d'incertitude
Rating		B+	BB	BBB-	BBB-	BB+	BB

Evaluer une entreprise

Eu égard à tout ce qui a été dit et fait plus haut, on peut donner un avis général sur l'entreprise à travers son évaluation. L'évaluation de l'entreprise est l'exercice par lequel l'analyste va donner son opinion le plus objectif possible sur la santé financière de l'entreprise. Cette opinion va être naturellement basée sur les résultats de l'analyse financière mais pas que. Oui, trop souvent il y a des facteurs macroéconomiques qui entrent en jeu dans le processus de création de richesse d'une entreprise. L'analyste financier doit être en mesure de trouver cela et le prendre en compte dans l'évaluation de l'entreprise.

Durant la crise des surprimes, beaucoup d'entreprises avaient des résultats douteux mais ce n'est pas pour cela que l'on accordait une mauvaise note à cette dernière.

Les enjeux macroéconomiques peuvent s'avérer être déterminants, surtout dans le monde d'échange où nous vivons. Donc cet aspect doit toujours attiser l'attention de l'analyste avant d'émettre un quelconque jugement sur une entreprise.

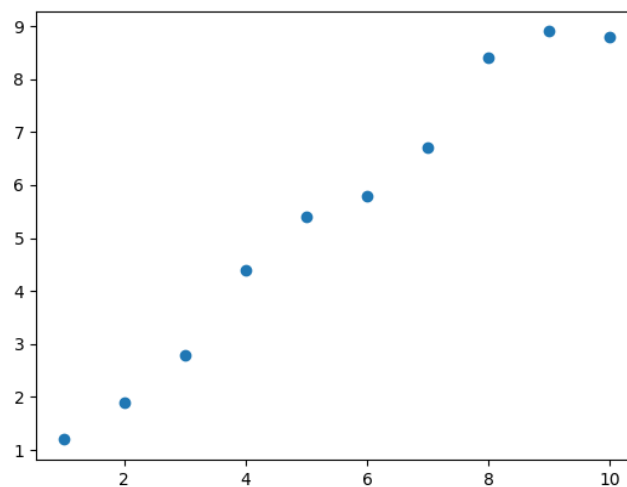
Ceci étant dit, cette évaluation faite à travers un rapport que l'analyste va rédiger et qui va ensuite être soumis au dirigeant, aux investisseurs, ou tout autres personnes physique ou morale s'intéressant à la structure analysée. L'analyste va reprendre les rubriques de l'analyse, l'explique d'une manière claire et concise à ce que des non financiers puissent le comprendre.

Si l'analyse a été bien faite, les bonnes décisions vont être prises. C'est là que réside tout l'intérêt de l'analyse financière.

Annexe 2 : Régression exemple

Linéaire

X	1	2	3	4	5	6	7	8	9	10
Y	1.2	1.9	2.8	4.4	5.4	5.8	6.7	8.4	8.9	8.8



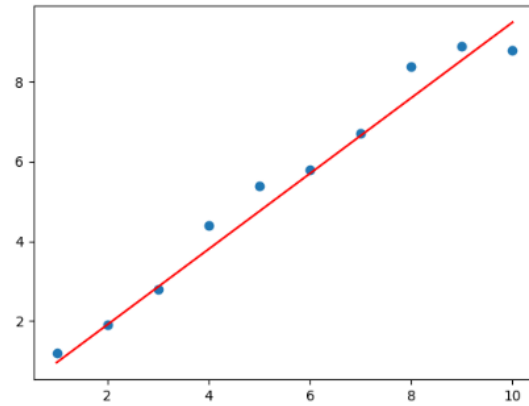
$$w_0 = 1 \text{ et } w_1 = 1$$

$$\text{Pour } x = 1, \quad y = 1 * 1 + 1, \quad y = 3$$

$$y = 3, \quad \hat{y} = 1.2, \quad MSE = (1.2 - 3)^2, \quad MSE = 3.24$$

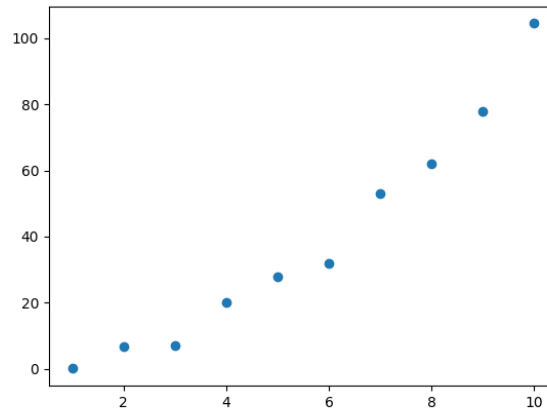
$$\frac{\partial MSE}{\partial Y} = -2 * 1(1.2 - 3), \quad \frac{\partial y}{\partial w_0} = 1$$

$$\frac{\partial MSE}{\partial w_0} = 3.6$$



Polynomiale

X	1	2	3	4	5	6	7	8	9	10
Y	0.1	6.9	7.2	20	28	32	53	62	78	104.5



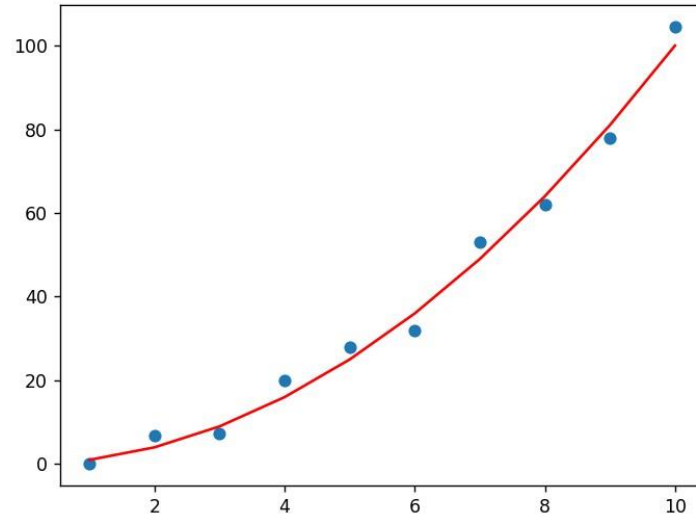
$$y = w0 + w1 * x + w2 * x^2$$

$$w0 = 1, \quad w1 = , \quad w2 = 1$$

$$\text{pour } x = 3, \quad y = 1 + 1 * 3 + 1 * 3^2, \quad y = 13$$

$$MSE = (7.2 - 3)^2, MSE = 17.64$$

$$\frac{\partial MSE}{\partial w2} = -2 * (x^2) * (y - (w0 + w1 * x + w2 * x^2))$$



Logistique

X1	X2	OU
0	0	0
0	1	1
1	0	1
1	1	1

Pour $x_1 = 1$, $x_2 = 0$, $y = 1 * 1 + 1 * 1 + 1 * 0 = 2$

$$\sigma(2) = 0.12$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2$$

$$MSE = (1 - 0.12)^2$$

$$MSE = 0.77$$

Tableau 10 Resultat XOR

X0	X1	X2	W0	W1	W2	Y	$\sigma(Y)$	OU
1	0	0	-2.2121	5.41528	5.41528	-2.2121	0.099	0
1	0	1				3.20318	0.961	1
1	1	0				3.20318	0.961	1
1	1	1				8.61846	0.999	1

Annexe 3 : Classification exemple

Arbre de décision

Tableau 11 Exemple données pour arbre de décision

Numéro	Plat	Teint	Taille	Si sénégalais
1	Riz	Sombre	Grande	Oui
2	Attiéké	Claire	Petite	Non
3	Mafé	Sombre	Grande	Non
4	Riz	Sombre	Grande	Oui
5	Attiéké	Sombre	Petite	Non
6	Mafé	Claire	Grande	Oui
7	Riz	Sombre	Grande	Oui

Calculons l'entropie générale

$$Entropy(S) = -\frac{4}{7} * \log_2\left(\frac{4}{7}\right) - \frac{3}{7} * \log_2\left(\frac{3}{7}\right)$$

$$Entropy(S) = 0.985$$

Gain d'information de l'attribut plat

$$Entropy(S_{Riz}) = -\frac{3}{3} * \log_2\left(\frac{3}{3}\right) - \frac{0}{3} * \log_2\left(\frac{0}{3}\right) = 0$$

$$Entropy(S_{Attieke}) = -\frac{2}{2} * \log_2\left(\frac{2}{2}\right) - \frac{0}{2} * \log_2\left(\frac{0}{2}\right) = 0$$

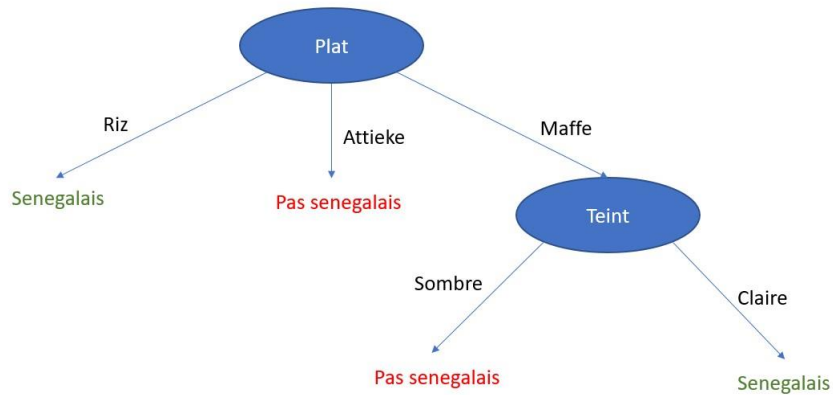
$$Entropy(S_{Mafé}) = -\frac{1}{2} * \log_2\left(\frac{1}{2}\right) - \frac{1}{2} * \log_2\left(\frac{1}{2}\right) = 1$$

$$GI(S, Plat) = 0.985 - \frac{3}{7} * 0 - \frac{2}{7} * 0 - \frac{2}{7} * 1 = 0.7$$

Si nous répétons les calculs avec les attributs nous allons trouver que

$$GI(S, Teint) = 0.006$$

$$GI(S, Taille) = 0.249$$



Naive Bayes

Numéro	Plat	Teint	Taille	Si sénégalais
1	Riz	Sombre	Grande	Oui
2	Attiéké	Claire	Petite	Non
3	Mafé	Sombre	Grande	Non
4	Riz	Sombre	Grande	Oui
5	Attiéké	Sombre	Petite	Non
6	Mafé	Claire	Grande	Oui
7	Riz	Sombre	Grande	Oui

Probabilité des valeurs cibles

$$P(\text{oui}) = \frac{4}{7} = 0.57, \quad P(\text{non}) = \frac{3}{7} = 0.43$$

Les probabilités des valeurs d'attributs

Plat	Oui	Non	Teint	Oui	Non
Riz	$\frac{3}{4}$	$\frac{0}{3}$	Sombre	$\frac{3}{4}$	$\frac{2}{3}$
Attiéké	$\frac{0}{4}$	$\frac{2}{3}$	Claire	$\frac{1}{4}$	$\frac{1}{3}$
Maffé	$\frac{1}{4}$	$\frac{1}{3}$			
Taille	Oui	Non			
Grande	$\frac{4}{4}$	$\frac{1}{3}$			
Petite	$\frac{0}{4}$	$\frac{2}{3}$			

Avec ce tableau nous avons tout ce qu'il nous faut pour classer un nouvel individu. D'ailleurs c'est

ce que nous allons faire, classons I1 (Plat = riz, Teint = sombre, Taille = Grande) et I2 (Plat = Attiéké, Teint = claire, Taille = Petite).

I1 :

$$P(oui|I1) = P(oui) * P(Plat = riz|oui) * P(Teint = sombre|oui) \\ * P(Taille = petite|oui)$$

$$P(oui|I1) = \frac{4}{7} * \frac{3}{4} * \frac{3}{4} * \frac{4}{4} = 0.32$$

$$P(non|I1) = P(non) * P(Plat = riz|non) * P(Teint = sombre|non) \\ * P(Taille = petite|non)$$

$$P(non|I1) = \frac{3}{7} * \frac{0}{4} * \frac{2}{3} * \frac{2}{3} = 0$$

Annexe 4 : Unsupervised Learning exemple

Clustering (k-means)

Tableau 12 Exemple donnees pour K-means

	P1	P2	P3	P4	P5	P6	P7	P8	P9
X	1	1	2	5	5	6	1	1	2
Y	1	2	1	5	6	5	9	10	9

C1 = {P1},

C2 = {P2},

C3 = {P3, P4, P5, P6, P7, P8, P9}.

Cg(C1) = (1, 1), Cg(C2) = (1, 2), Cg(C3) = (3.42, 6.42)

	P1	P2	P3	P4	P5	P6	P7	P8	P9
C1	0.0	1.0	1.0	5.66	6.40	6.40	8.0	9.0	8.06
C2	1.0	0.0	1.41	5.0	5.66	5.83	7.0	8.0	7.07
C3	5.94	5.04	5.60	2.12	1.63	2.94	3.54	4.21	2.94

C1 = {P1, P2},

C2 = {P3},

C3 = {P4, P5, P6, P7, P8, P9}.

Règles d'association

T1 = {A, B, C},

T2 = {E, F},

T3 = {A, C, E},

T4 = {A, E,}

T5 = {B}.

$$\sup(A) = \frac{3}{5} \quad \sup(B) = \frac{2}{5} \quad \sup(C) = \frac{2}{5} \quad \sup(E) = \frac{2}{5} \quad \sup(F) = \frac{1}{5} \quad \text{minsup} = \frac{2}{5}$$

$$\begin{aligned} \sup(AB) &= \frac{1}{5} & \sup(AC) &= \frac{2}{5} & \sup(AE) &= \frac{2}{5} & \sup(BC) &= \frac{1}{5} & \sup(BE) \\ &= \frac{0}{5} & \sup(CE) &= \frac{1}{5} \end{aligned}$$

$A \rightarrow C$, conf = 2/3, sup = 2/5 : cette règle n'est pas intéressante car son support > minsup et sa confiance < minconf.

$C \rightarrow A$, conf = 1, sup = 2/5 : cette règle est intéressante car son support > minsup et sa confiance > minconf.

Annexe 5 : Les outils utilisés

C++

Le C++ est un langage de programmation créé en 1985 par l'informaticien danois Bjarne Stroustrup pour pallier aux manquements du langage C qui n'est pas orienté objet. Le C++ est un langage de programmation très utilisé par les développeurs, notamment en ce qui concerne les applications. Il permet d'aborder le développement sous plusieurs paradigmes : programmation générique, procédurale et orientée objet. C'est un langage compilé, ce qui signifie que le code source est traduit en code objet ou binaire pour que la machine puisse l'exécuter. (C++ : présentation du langage de programmation, 2024)

Ce langage de programmation est un langage orienté objet ce qui veut dire il permet de créer des classes. Il est si populaire, on peut donner l'exemple de Google qui l'utilise pour son moteur de recherche, Microsoft qui l'utilise pour Word, Excel ou PowerPoint et aussi Autodesk qui l'utilise pour Maya. Pour ce qui est de l'IA, avant l'avènement de Python, les ingénieurs l'utilisaient pour écrire les codes mais son impact est toujours présent. Car derrière presque tous les Framework de Python, qui nous aident dans l'IA, il y a le C++ ou C, le cas de Numpy, Pandas ou Matplotlib.

Les avantages de C++ :

- La performance, rapidité
- La popularité
- La portabilité dans les OS
- L'abondance de bibliothèques
- La programmation orienté objet

Les inconvénients :

- Syntaxe compliquée
- Langage pas du tout pour les débutants

Python

Python est un langage de programmation créé par Guido Van Rossum. La première version publique du langage est sortie en 1991. Son nom provient de la troupe de comiques anglais les Monty Python.

Python est un langage de programmation dit de "très haut niveau". Cela signifie qu'il possède un haut niveau d'abstraction par rapport au langage machine. Pour le dire très simplement : plus un langage de programmation est de "haut niveau", plus sa syntaxe se rapproche de notre langage (l'anglais) plutôt que du langage machine. Un langage de haut niveau est donc plus facile à comprendre et à utiliser qu'un langage de plus bas niveau.

Certains langages (comme Python) utilisent un interpréteur comme traducteur tandis que d'autres utilisent un compilateur.

Un interpréteur se distingue d'un compilateur par le fait que, pour exécuter un programme, les opérations d'analyse et de traductions sont réalisées à chaque exécution du programme (par un interprète) plutôt qu'une fois pour toutes (par un compilateur). (Introduction à Python, 2024)

Les avantages de Python :

- Facile à utiliser
- Sécuriser
- Très populaire
- Compatibilité avec d'autres langages
- Possède beaucoup de bibliothèque pour le Machine Learning

Les limites de Python

- Temps d'interprétation très lent
- Mauvaise présentation des erreurs

SQL

Structured Query Language (SQL) est un langage de gestion de données sous forme de base de données. Il est utilisé pour gérer des bases de données relationnelles avec ces quatre (4) actions principales que sont le CRUD (CREATE, RETREIVE, UPDATE ET DELETE).

De manière simple SQL va nous permettre de créer des bases de données en utilisant un système de gestion de base de données comme PostgreSQL, Oracle, Maria DB mais nous allons utiliser MySQL. Il va être créer une base de données locale pour stocker les états financiers avec lesquels nous allons travailler.

Les bibliothèques et Framework

Un Framework est en ensemble de fonction prédéfinie dans un langage de programmation nous permettant de faire une action bien précise. Pour une tache bien définie, si nous avons un Framework, il n'est pas nécessaire de commencer de zéro puisque certaines fonctionnalités sont déjà implémentées. Les Framework ont été créer pour les tâches complexes qui nécessite beaucoup de compétence, ainsi même les développeurs de niveau moyen peuvent créer des programmes avancés, ce qui va servir à la productivité.

Il y a différents Framework pour différents domaines informatiques (développement web, mobile, logiciel...), mais nous allons présenter les Framework qui vont nous aider dans le développement de modèle intelligent et le développement d'interface graphique, ils sont tous liés soit à Python ou à C++.

Scikit-learn

Scikit-learn est une bibliothèque de Python qui a commencé en 2007 avec le Google Summer of Code Project par David Carpaneau.

Ce Framework s'est spécialisé dans le Machine Learning (supervisé et non supervisé) et nous donne des fonctions pratiques pour le développement de modèles. Scikit-learn supporte parfaitement des domaines comme la classification, la régression, le clustering ...

Pour ce qui est de nos modèles, nous allons l'utiliser pour faire la prédiction des éléments des états financiers, la représentation de texte et bien d'autres.

Spacy

Spacy est un Framework de python qui est utilisée dans le NLP, il nous permet tout simplement de raffiner du texte avant sa modélisation. La raison est que du texte écrit par un humain, avant de faire sa représentation en chiffre, va contenir du bruit. On entend par bruit tous les éléments, mots, vocabulaire qui ne nous sert pas. Exemple dans le langage français, il y a une conjugaison très lourde, un verbe comme « faire » peut donner fait, faite, furent etc. Spacy peut nous aider à trouver la racine de ses mots. Il peut aussi nous aider à tout mettre minuscule, supprimer les ponctuation...

Gensim

Gensim est un Framework Python avec lequel, il est possible de faire directement de la classification de texte. Dans son fonctionnement, ce Framework utilise les modèle Word2vec et aussi Fasttext. Ces derniers sont des techniques qui sont appelé Self Supervided Learning ou l'on utilise un texte pour générer les inputs et les outputs.

Avec ce Framework on peut calculer la similarité entre deux mots, voir faire des calculs sur des mots, oui des calculs des sur des mots.

Tensorflow

Si Scikit-Learn est une bibliothèque de Machine Learning, Tensorflow en est une spécialisée sur le Deep Learning. Cette bibliothèque a été développée par Google dans le but de permettre aux experts mais aussi de débutant d'avoir un environnement pour travailler dans le Machine Learning en général.

De tous les Framework que nous avons présentée, Tensorflow est très probablement le plus puissant car nous permettant de faire ce que tous les autres font. C'est un outil tout en un avec ses forces et ses faiblesses.

En plus de tout cela, il nous permet de faire du Computer Vision avec les CNN, du NLP avancée avec la RNN, les LSTM, de créer des API pour le déploiement et bien d'autres.

Pandas

Pandas est un Framework Python très pratique dans le développement de modèle. Il est utilisé dans le travail a priori, le Feature Engineering. Avec pandas, nous pouvons importer des fichiers CSV, vérifier les données manquantes, les outliers...

Avant chaque développement de modèle, Pandas va certainement intervenir, ce Framework supporte les statistiques qui pourront nous permettre de mettre les données dans un format acceptable par l'ordinateur.

Numpy

Nativement, les structure de type tableau n'existe pas en Python, il y a des listes en Python pour le remplacer. La différence entre ces deux c'est que les tableaux acceptent un seul type de donnée et taille ne varie pas or, les listes acceptent plusieurs types et sa taille peut varier.

C'est là qu'intervient Numpy pour permettre d'utiliser des tableaux des Python, qui sont bien plus rapide à exécuter. En plus de cela, Numpy a un excellent support de l'algèbre linéaire, les matrices, les vecteurs et autres domaines mathématiques.

Il est nécessaire d'ajouter que Numpy, bien qu'utilisé en Python est écrit en langage C qui est plus puissant et plus rapide que le Python.

Matplotlib

Matplotlib est un Framework de visualisation avec Python, il sert à tracer des courbes en utilisant Numpy ou Pandas. La visualisation peut intervenir avant et après le modèle, soit pour les comprendre les données brutes, soit pour vérifier les résultats.

La visualisation est en train de devenir une science à part entière, donc Matplotlib est utilisé dans des domaines autres que le Machine Learning, notamment dans le développement d'interface graphique que nous allons voir.

Qt

Qt est une bibliothèque de C++ cross plateforme lancée en 1995, et qui est complètement gratuit. Qt nous permet de créer des interfaces très avancées et dans un IDE et simple à comprendre et à utiliser. Avec cette bibliothèque, il n'est pas nécessaire de savoir coder pour créer des interfaces graphiques car il y a la possibilité de créer des widgets avec du glisser-déposer.

Pour ce qui nous concerne nous allons bien évidemment l'utiliser pour l'interface qui va accueillir les clients. Mais un logiciel mais pas du web.

Puisque cette application n'a pas pour vocation d'être déployé dans le cloud, le web n'est pas nécessaire. De plus, développer l'interface graphique de cette manière nous donne une certaine sécurité car il n'y aura pas de brèche que des personnes extérieures à l'organisation peuvent utiliser pour accéder aux données sensibles.

Table des matières

Remerciements.....	I
Sommaire	II
Liste des figures	IV
Liste des tableaux.....	V
Liste des équations	VI
Liste des sigles et des acronymes.....	VIII
Introduction générale	1
1 Chapitre I : Revue de la littérature scientifique	3
Introduction de chapitre	3
1.1 Généralités l'intelligence artificielle	3
1.1.1 Définition de l'intelligence artificielle.....	3
1.1.1.1 C'est quoi l'intelligence	3
1.1.1.2 C'est quoi artificielle	4
1.1.1.3 Proposition de définitions de l'intelligence artificielle	5
1.1.2 Des autres sciences à l'intelligence artificielle	6
1.1.2.1 Les mathématiques.....	6
1.1.2.2 La biologie	6
1.1.2.3 L'informatique.....	7
1.1.2.4 La cybernétique.....	7
1.1.3 Domaines d'application de l'intelligence artificielle.....	8
1.1.3.1 Les prédictions	8
1.1.3.2 La robotique.....	8
1.1.3.3 La finance	8
1.1.3.4 Voitures autonomes.....	8
1.1.3.5 Les Chatbots.....	8
1.2 Etat de la recherche sur l'IA appliquée à la finance.....	9
1.2.1 Les recherche sur l'IA.....	9

1.2.1.1	Genèse de l'IA : le premier neurone artificiel	9
1.2.1.2	L'IA dans la prédiction.....	9
1.2.1.3	Les Chatbots.....	10
1.2.2	L'IA appliquée à la finance.....	11
1.2.2.1	La détection de fraude	12
1.2.2.2	Service client	12
1.2.2.3	Analyse financière.....	13
1.2.3	Les limites sur les travaux actuels	14
1.2.3.1	Qualités des données financières	14
1.2.3.2	L'éthique	14
1.2.3.3	L'insuffisance des études dans le contexte africain.....	15
	Conclusion partielle	15
2	Chapitre II : L'analyse financière	16
	Introduction de chapitre	16
2.1	Présentation des états financiers.....	16
2.1.1	Le bilan	16
2.1.1.1	Les actifs.....	17
2.1.1.2	Les passifs.....	18
2.1.1.3	Exemple de bilan.....	20
2.1.2	Le compte de résultat	21
2.1.2.1	Les charges.....	21
2.1.2.2	Les produits.....	21
2.1.2.3	Les soldes intermédiaires de gestion (SIG)	22
2.1.2.4	Exemple de compte de résultat	25
2.1.3	Les flux de trésorerie (TFT).....	26
2.1.3.1	La trésorerie initiale (au 1 janvier de l'année N).....	26
2.1.3.2	Les flux de trésorerie liées aux activités opérationnels (FTAO)	26
2.1.3.3	Les flux de trésorerie liées aux activités d'investissements (FTAI)	27
2.1.3.4	Les flux de trésorerie liées aux activités de financement (FTAF).....	27
2.1.3.5	La trésorerie finale (au 31 décembre de l'année N)	28

2.1.3.6	Exemple de tableau des flux de trésorerie	29
2.2	L'analyse financière proprement dite.....	30
2.2.1	Vérification des états financiers.....	30
2.2.2	Analyse des états financiers	30
2.2.2.1	Analyse du compte de résultat	30
2.2.2.2	Analyse du bilan	31
2.2.2.3	Analyse du tableau des flux de trésorerie	31
2.2.3	Analyse de l'activité et des relations de trésorerie.....	32
2.2.3.1	Analyse du cycle de vie de l'activité.....	32
2.2.3.2	Analyse du comportement des flux de trésorerie	32
2.2.3.3	Analyse des équilibres financiers et la relation de trésorerie	32
2.2.4	Analyse tendancielle et la méthode des ratios	33
2.2.4.1	Les familles de ratios.....	33
2.2.4.2	L'analyse par familles de ratios.....	35
2.2.4.2.1	Calcul des ratios	36
2.2.4.2.2	Analyse tendancielle des ratios	36
2.2.4.2.3	La représentation graphique	36
	Conclusion partielle	37
3	Chapitre III : Théories derrière les algorithmes d'intelligence artificielle.....	38
	Introduction de chapitre	38
3.1	Les prérequis	38
3.1.1	Les mathématiques.....	38
3.1.1.1	Les statistiques et probabilités	39
3.1.1.2	L'algèbre linéaire.....	40
3.1.1.3	L'analyse.....	40
3.1.2	L'informatique	41
3.1.2.1	L'algorithme	41
3.1.2.2	Les structure de données.....	42
3.1.2.3	Les langages de programmation.....	43
3.2	Les algorithmes d'intelligence artificielle.....	43

3.2.1	Machine Learning	43
3.2.1.1	Supervised learning.....	43
3.2.1.2	La régression	44
3.2.1.2.1	La régression linéaire	44
3.2.1.2.2	La régression logistique	46
3.2.1.2.3	La régression polynomiale	48
3.2.1.3	La classification	49
3.2.1.3.1	Support Vector Machine (SVM)	49
3.2.1.3.2	L'arbre de decision.....	51
3.2.1.3.3	Naive Bayes.....	53
3.2.1.4	Unsupervised learning	54
3.2.1.4.1	Clustering	54
3.2.1.4.2	Règles d'associations	55
3.2.2	Deep Learning.....	55
3.2.2.1	Artificial neuron network (ANN)	56
3.2.2.2	Convolutional neuron network (CNN)	56
3.2.2.3	Recurrent neuron network (RNN).....	57
	Conclusion partielle	57
4	Chapitre IV : Implémentation des modèles	58
	Introduction de chapitre	58
4.1	Spécifications des besoins.....	59
4.1.1	Contexte de l'application	59
4.1.2	L'expression des besoins	60
4.1.3	Les fonctionnalités de l'application.....	61
4.2	Développement des modèles intelligents	62
4.2.1	La collecte de données	62
4.2.1.1	Les données d'états financiers.....	63
4.2.1.2	Les données texte	63
4.2.2	Prédiction des valeurs	65

4.2.3	Le développement du Chatbot	69
4.2.3.1	Les Intents	69
4.2.3.2	L'entity detection	70
4.2.3.3	La modélisation du texte.....	71
4.2.3.1	La gestion des réponses	74
4.2.3.2	Exemple de fonctionnement du Chatbot.....	74
4.3	L'interface graphique	77
4.3.1	Développement de l'interface graphique	77
4.3.1.1	Le frontend.....	77
4.3.1.2	Le Backend	79
4.3.1.3	Le Web server	80
4.3.2	Présentation de l'application.....	81
4.4	Résultats et discussions	83
4.4.1	Présentation des résultats	83
4.4.2	Limites et défis.....	83
4.4.3	Améliorations possibles	83
	Conclusion partielle	83
	Conclusion générale et perspectives	83
	Bibliographies	A
	Webographies	D
	Annexes.....	E
	Table des matières.....	Y