**1.1 Les robo-advisors *(version enrichie et humanisée)***

Les **robo-advisors**, ou conseillers financiers automatisés, sont des plateformes numériques qui permettent aux investisseurs, souvent particuliers, d’accéder à des services de gestion de portefeuille sans avoir à passer par un conseiller humain traditionnel. Ils utilisent principalement des algorithmes pour analyser le profil de l’utilisateur et proposer une **allocation d’actifs optimisée**, généralement via des fonds indiciels cotés (ETF).

Leur apparition remonte à la fin des années 2000, dans un contexte de remise en question des modèles bancaires classiques, notamment après la crise financière de 2008. À cette période, la défiance envers les établissements financiers traditionnels, combinée à l’essor des technologies web, a favorisé l’émergence d’une nouvelle génération d’acteurs comme **Betterment** ou **Wealthfront** aux États-Unis. Ces plateformes ont permis d’automatiser une partie du conseil financier, en le rendant plus accessible, plus transparent, et surtout **moins coûteux** pour les petits investisseurs (Nexialog, 2023).

Le fonctionnement d’un robo-advisor repose sur un processus en plusieurs étapes :

1. **Recueil du profil de l’investisseur** via un questionnaire (âge, objectifs, horizon d’investissement, tolérance au risque) ;
2. **Proposition d’un portefeuille type** basé sur des modèles quantitatifs et des indices de marché ;
3. **Suivi automatique** du portefeuille, avec parfois un rééquilibrage périodique.

La plupart des robo-advisors reposent sur des **stratégies passives**, s’appuyant sur des allocations larges et diversifiées, comme le fameux modèle 60/40 (60 % actions, 40 % obligations). Cette approche limite les frais et évite les prises de position trop spéculatives.

En France, le marché a vu émerger des plateformes comme **Yomoni**, **Nalo** ou **Advize**, qui ont adapté le concept au cadre réglementaire européen. Conformément à la directive **MIFID II**, ces acteurs doivent garantir la traçabilité des recommandations, la transparence des frais, et une personnalisation suffisante pour répondre aux exigences du “devoir de conseil”.

Du point de vue de l’utilisateur, les robo-advisors offrent plusieurs avantages. Ils permettent une **entrée sur les marchés financiers simple et encadrée**, même avec un faible capital de départ. Ils assurent également une **gestion dépersonnalisée**, donc moins sujette aux biais émotionnels typiques des investisseurs non professionnels (MNA Robo-Advisor, 2021).

Mais ces plateformes ne sont pas exemptes de limites. Leur approche repose sur des modèles prédéfinis, qui peinent parfois à intégrer des cas particuliers (patrimoine atypique, objectifs non standards, évolutions personnelles). De plus, elles ne proposent pas, sauf exception, **d’accompagnement humain réel**. Cela peut poser problème lorsqu’il s’agit de répondre à des enjeux plus complexes, comme la transmission, la fiscalité internationale, ou encore l’investissement responsable.

Des études récentes ont aussi pointé une **forme d’illusion de simplicité** : l’expérience utilisateur fluide et les interfaces épurées peuvent masquer la complexité réelle des décisions financières, voire inciter certains profils à se lancer sans compréhension suffisante des risques encourus (Robo-advisors and Wealth Management, 2020).

**Citation à insérer (si besoin) :**  
*« While robo-advisors offer cost efficiency, human advisors still hold advantages in trust and emotional intelligence. »*  
(*Robo-advisors and Wealth Management, ~2020*)

L’intégration croissante des technologies dans les services d’investissement, notamment via les robo-advisors, témoigne d’un changement de paradigme dans la manière dont les conseils financiers sont conçus et délivrés. Ce glissement vers des solutions automatisées n’est que le début d’un mouvement plus profond, aujourd’hui renforcé par l’essor des modèles de langage à grande échelle. Ces nouveaux outils, portés par l’intelligence artificielle générative, promettent d’élargir encore les capacités des systèmes de conseil numérique. C’est dans cette dynamique que s’inscrit l’usage croissant des **LLM** dans le domaine financier.

**1.2 Les modèles de langage (LLM) en finance**

L’émergence des **modèles de langage de grande taille**, ou **LLM** (*Large Language Models*), constitue l’un des tournants technologiques majeurs de ces dernières années. Ces modèles, entraînés sur d’immenses volumes de textes, sont capables de **comprendre, générer et reformuler du langage naturel** avec un niveau de cohérence impressionnant. Le plus connu d’entre eux est sans doute **GPT-4**, développé par OpenAI, mais d’autres modèles comme **Claude**, **LLaMA**, ou encore **BloombergGPT** commencent à s’imposer dans des secteurs spécifiques.

Dans le secteur financier, ces outils ne sont plus cantonnés à de simples fonctions de service client ou de génération de contenu marketing. Ils sont aujourd’hui testés pour des usages beaucoup plus ambitieux : **lecture de rapports financiers**, **extraction d’informations clés**, **synthèse d’analyses de marché**, voire **génération de recommandations d’investissement**.

Un bon exemple est **BloombergGPT** (2023), un modèle de langage entraîné spécifiquement sur des données financières, mêlant textes issus d’articles, de rapports d’entreprise, de données économiques et de publications réglementaires. Selon les auteurs du projet, ce modèle surpasse GPT-3 sur plusieurs tâches financières, en particulier dans l’interprétation de textes spécialisés et la classification d’informations critiques (*NLP : Natural Language Processing*¹).

Les avantages des LLM dans le domaine financier sont multiples :

* Capacité à **traiter des volumes massifs d’informations textuelles** en un temps record
* Possibilité de produire des **résumés intelligents de rapports financiers**, d’analyses de portefeuille ou d’actualités de marché
* Facilité d’intégration dans des interfaces utilisateurs (chatbots, assistants de conseil automatisés)

Cependant, ces modèles posent aussi des **questions cruciales de fiabilité**. Par exemple, **les hallucinations**² sont encore fréquentes : le modèle peut produire des affirmations fausses ou inventées mais formulées avec aplomb. Ce phénomène soulève un risque majeur lorsqu’il s’agit d’applications critiques comme le conseil en investissement ou la gestion de risque.

Par ailleurs, la **manque de transparence** de ces modèles — souvent qualifiés de *black boxes* — rend leur usage délicat dans des secteurs soumis à des régulations strictes, comme la finance. Les institutions doivent pouvoir justifier les décisions algorithmiques, notamment dans le cadre des directives comme **MIFID II**³.

Des études récentes comme **“A Survey of Generative AI in Finance”** (Liu et al., 2023) montrent que les LLM ont un potentiel énorme dans des domaines tels que :

* L’analyse ESG (*environnement, social, gouvernance*)
* La conformité réglementaire (*regulatory compliance*)
* L’assistance à la gestion de portefeuille
* La veille concurrentielle

Mais elles insistent également sur le fait qu’un **encadrement rigoureux** est indispensable pour éviter des usages à risque, notamment dans les phases de génération automatique de recommandations financières.

En résumé, les LLM sont en train de **transformer les capacités analytiques** des institutions financières. Leur intégration dans des outils comme les **robo-advisors** ouvre des perspectives très prometteuses, à condition de maîtriser les limites techniques et éthiques de ces technologies.

Si les capacités techniques des modèles de langage ont été largement démontrées, notamment dans le traitement de textes financiers complexes, une question essentielle reste en suspens : **dans quelle mesure ces technologies sont-elles déjà appliquées concrètement dans le domaine du conseil financier ?** Pour répondre à cela, il est pertinent d’observer les travaux de recherche récents ainsi que les premiers produits et expérimentations mis en œuvre par les acteurs du marché.

**1.3 Travaux et produits existants liant IA et conseil financier**

L’application de l’intelligence artificielle, et plus récemment des **modèles de langage génératifs**, au conseil financier, n’est plus une idée théorique. De nombreuses initiatives, à la fois académiques et industrielles, cherchent à **évaluer l’apport réel de ces technologies** dans la prise de décision d’investissement.

Côté recherche, plusieurs travaux se sont penchés sur l’usage des LLM pour **générer des conseils financiers**, évaluer des portefeuilles ou anticiper des mouvements de marché. Par exemple, dans leur étude publiée en 2023, **Chen et al.** se sont posé la question suivante : *“ChatGPT peut-il prédire l’évolution des prix des actions ?”* Leurs résultats montrent que, bien que le modèle soit capable de capter certains signaux faibles à partir de textes économiques, il n’est pas encore fiable pour produire des recommandations d’investissement direct (*Can ChatGPT Forecast Stock Price Movements ?*, 2023).

Dans une optique plus opérationnelle, d’autres chercheurs se sont intéressés à l’automatisation du **processus de conseil**. C’est le cas du projet **FinGPT**, qui cherche à créer une interface ouverte de LLM entraînée sur des données financières à grande échelle. L’objectif est de proposer une alternative open source à des modèles comme GPT-4, mais spécifiquement orientée vers les cas d’usage bancaires et assurantiels (FinGPT, 2023)⁴.

Du côté des acteurs économiques, plusieurs entreprises testent ou intègrent déjà ces technologies. Le plus emblématique est **Bloomberg**, qui a développé **BloombergGPT**, un modèle interne conçu pour répondre aux besoins très spécifiques de ses utilisateurs professionnels : génération de résumés financiers, classification de documents, réponse à des requêtes en langage naturel, etc. Contrairement aux modèles généralistes, BloombergGPT est optimisé pour éviter les hallucinations et les approximations propres aux modèles généralistes utilisés hors contexte.

Certaines fintechs, de leur côté, expérimentent des **chatbots intelligents** pour compléter leurs interfaces de robo-advisor. Ces agents conversationnels permettent d’expliquer des allocations, de répondre aux questions de base sur les produits proposés, voire de simuler des scénarios d’investissement en langage naturel. Bien que ces interfaces restent limitées à un rôle informatif — pour des raisons légales et techniques —, elles montrent le potentiel d'une **expérience utilisateur enrichie par l’IA générative**.

En revanche, peu d’acteurs vont jusqu’à **générer automatiquement une recommandation personnalisée intégrant un raisonnement financier complet**, comme un conseiller humain pourrait le faire. Cela tient à plusieurs facteurs : d’une part, les LLM manquent encore de précision sur les données quantitatives (*comme les ratios financiers ou les historiques de performance*⁵), et d’autre part, la régulation limite fortement la possibilité de fournir un conseil sans supervision humaine.

Ainsi, les travaux et produits existants montrent un intérêt croissant pour l’IA générative dans le conseil financier, mais également **un usage encore encadré**, prudent, et souvent limité à des fonctions de support ou d’assistance.

Bien que les expérimentations en cours démontrent un réel potentiel d’intégration des modèles de langage dans les services d’investissement, leur adoption à grande échelle reste conditionnée par une série de **défis majeurs**. Ces défis sont à la fois d’ordre technique, éthique et réglementaire, et constituent aujourd’hui les principaux freins à une automatisation complète du conseil financier.

**1.4 Défis : hallucinations, explicabilité, biais, cadre réglementaire**

**1.4.1 Hallucinations et fiabilité des modèles**

L’un des problèmes les plus documentés dans l’usage des LLM est le phénomène d’**hallucination** : le modèle peut générer des informations factuellement fausses, tout en les présentant avec une grande assurance. Dans le cadre du conseil financier, ce risque est particulièrement problématique, car il peut **induire l’investisseur en erreur** ou le pousser à prendre des décisions inadaptées à son profil ou à la réalité du marché.

Même des modèles spécialisés comme BloombergGPT, bien qu’optimisés pour la finance, ne sont pas totalement à l’abri de ce type de comportement. Cela s’explique par le fonctionnement probabiliste des LLM, qui repose sur la **prédiction de la suite la plus probable d’un texte donné**, et non sur une vérification systématique de la véracité des données mobilisées (*cf. note 6*).

**1.4.2 Explicabilité et boîte noire**

Un autre défi majeur réside dans le **manque de transparence des modèles**. Les LLM sont souvent qualifiés de *boîtes noires* : même leurs concepteurs ont parfois du mal à expliquer pourquoi tel output a été produit pour tel input. Cette opacité est incompatible avec certaines exigences réglementaires dans la finance, où il est impératif de pouvoir **justifier les conseils fournis** au client, notamment dans le cadre de MIFID II⁷.

Des pistes comme l’analyse des *tokens*, les *attention maps* ou les techniques de **post-hoc explainability** (ex. LIME, SHAP) sont explorées, mais restent encore peu opérationnelles dans un cadre grand public.

**1.4.3 Biais et discrimination algorithmique**

Les LLM peuvent reproduire, voire amplifier, des **biais présents dans les données d’entraînement**. En finance, cela peut se traduire par une sur-représentation de certains types d’actifs, une préférence implicite pour des stratégies de marché dominantes, ou même — dans les pires cas — une forme de discrimination involontaire envers certains profils d’investisseurs (jeunes, seniors, personnes à faible capital...).

C’est pourquoi plusieurs institutions recommandent de **tester systématiquement les modèles sur la base de critères d’équité** avant tout déploiement opérationnel. Cela rejoint les principes de la **finance responsable** et de l’**IA éthique**, aujourd’hui largement promus par les régulateurs et les grands acteurs du secteur.

**1.4.4 Environnement réglementaire**

Enfin, l’encadrement réglementaire représente un obstacle — mais aussi un garde-fou essentiel — à l’utilisation massive des LLM dans le conseil en investissement. En Europe, la directive **MIFID II** impose une série d’obligations en matière de transparence, de connaissance client et de justification des décisions. Le règlement **RGPD** encadre quant à lui l’usage des données personnelles et pose des exigences strictes en matière de consentement et de droit à l’explication.

La législation à venir sur l’**IA Act européen** prévoit également un classement des IA utilisées en finance comme présentant un **haut niveau de risque**, impliquant des audits réguliers, une supervision humaine obligatoire et une documentation approfondie.

L’état de l’art a montré que les modèles de langage ont un potentiel réel pour assister ou automatiser certaines tâches de conseil financier. Cependant, leur intégration dans un outil utilisable repose sur des choix techniques précis, une architecture cohérente, et des critères d’évaluation rigoureux.  
La suite de ce mémoire décrit la méthodologie retenue pour concevoir et expérimenter un prototype de robo-advisor s’appuyant sur un LLM.

**2.1 Choix du modèle de langage et justification**

Pour ce projet, j’ai choisi d’utiliser **GPT-4**, développé par **OpenAI**. C’est aujourd’hui l’un des modèles de langage les plus performants accessibles via une API.

Il a été retenu pour plusieurs raisons simples :

* **Qualité des réponses** : GPT-4 produit des textes clairs, structurés et adaptés au contexte posé. C’est essentiel pour générer des recommandations compréhensibles par un utilisateur non expert.
* **Souplesse d’utilisation** : grâce à l’API d’OpenAI, il est possible d’intégrer le modèle dans une application Python en quelques lignes de code. Cela permet de construire un prototype rapidement.
* **Expérimentation libre** : contrairement à des modèles fermés (comme BloombergGPT), GPT-4 permet de tester différents prompts, scénarios ou formats de sortie.
* **Documentation abondante** : GPT-4 est bien documenté. De nombreux exemples, tutoriels ou guides sont disponibles, ce qui facilite la mise en œuvre.

Il faut cependant noter certaines limites. GPT-4 est un **modèle généraliste**. Il n’a pas été entraîné spécifiquement sur des textes financiers, même s’il comprend bien ce domaine grâce à la masse de données utilisées pour son entraînement. Il peut aussi générer des **réponses imprécises ou approximatives**, notamment si le prompt est mal formulé.

Enfin, GPT-4 est un modèle **hébergé** (cloud), ce qui impose une connexion Internet et dépend du respect de conditions d’utilisation propres à OpenAI.

Malgré ces limites, GPT-4 reste aujourd’hui l’un des meilleurs choix pour tester la génération de conseils automatisés dans un cadre exploratoire comme celui de ce mémoire.

Le modèle GPT-4 ayant été retenu pour ses performances et sa facilité d’intégration, il a ensuite été intégré dans une architecture simple conçue pour tester son usage dans un contexte de conseil financier simulé.

**2.2 Architecture fonctionnelle du prototype *(version enrichie)***

Le prototype conçu dans le cadre de ce mémoire repose sur une architecture simple, modulaire et adaptée à un usage exploratoire. Il s’agit d’un système capable de générer des recommandations d’investissement à partir d’un profil utilisateur, en s’appuyant sur un **modèle de langage (LLM)**, ici **GPT-4**, accessible via une API.

L’architecture suit **quatre étapes fonctionnelles clés**, décrites ci-dessous.

**1. Entrée utilisateur – Définition du profil d’investisseur**

L’utilisateur commence par indiquer son **profil de risque**. Trois profils types sont proposés :

* **Prudent** : préférence pour la stabilité, faible tolérance au risque
* **Équilibré** : recherche d’un bon compromis entre sécurité et rendement
* **Dynamique** : tolérance élevée au risque, objectif de performance long terme

Cette étape correspond à un **profilage utilisateur simplifié**, équivalent au processus de connaissance client (KYC) utilisé par les robo-advisors professionnels. Les réponses sont sélectionnées à l’aide d’un menu déroulant dans l’interface.

(*Profils types : catégories d’allocation de portefeuille standardisées selon la tolérance au risque. Ex. : 80 % actions pour un profil dynamique.)*

**2. Génération de la recommandation avec GPT-4**

Une fois le profil sélectionné, un **prompt structuré** est envoyé à GPT-4 via l’API OpenAI.

Exemple de prompt :

*Tu es un conseiller financier professionnel. Propose un portefeuille simple et diversifié, adapté à un investisseur prudent. Utilise des ETF ou indices représentatifs. Justifie brièvement chaque choix.*

Ce prompt suit une logique de **prompt engineering** : il précise le rôle de l’IA (*« Tu es un conseiller »*), la tâche attendue (*« proposer un portefeuille »*), les contraintes (*« ETF uniquement, justification brève »*), et le profil utilisateur.

La réponse de GPT-4 est un **texte structuré**, contenant généralement :

* Une liste d’ETF ou classes d’actifs
* Les pourcentages alloués
* Un ou deux arguments par ligne

**3. Extraction et structuration des données**

Le texte généré par GPT-4 est ensuite **traité automatiquement**. L’objectif est de passer d’un paragraphe brut à une **structure lisible et exploitable** :

* Portefeuille sous forme de tableau
* ETF identifiés ligne par ligne
* % d’allocation convertis en valeurs numériques

Extrait typique à analyser :

« 60 % – iShares Core US Aggregate Bond ETF (AGG) : stabilité et faible volatilité  
40 % – SPDR S&P 500 ETF Trust (SPY) : exposition aux grandes capitalisations US »

Ce traitement est réalisé en Python, via une fonction de **parsing personnalisé**. Elle recherche les lignes contenant des pourcentages, détecte les noms d’ETF et extrait les commentaires associés.

(*Parsing : transformation d’un texte non structuré en données exploitables. Essentiel pour visualiser les résultats.)*

**4. Affichage des résultats dans Streamlit**

Les données extraites sont ensuite affichées dans une **interface web développée avec Streamlit**.

L’utilisateur visualise :

* Le portefeuille proposé sous forme de tableau
* Un graphique circulaire interactif (Pie chart)
* Le texte original généré par le LLM

Cette interface permet de tester plusieurs profils rapidement et de **comparer les recommandations**.

Voici un aperçu du flux :

Fonctionnement général de l’application (schéma logique)

**Limites techniques**

* GPT-4 peut parfois produire des formats incohérents ou ambigus (ex. : noms d’ETF incomplets).
* Il peut aussi oublier de respecter certaines consignes (ex. : ne pas dépasser 100 % d’allocation).
* Le parsing repose donc sur une logique de “meilleur effort”, avec des cas non couverts.

Ces limites sont importantes à garder à l’esprit dans l’analyse critique à venir.

**2.3 Sources de données utilisées**

Le prototype repose volontairement sur des **données publiques et ouvertes**. Cela permet d’éviter les problématiques liées à la confidentialité, tout en garantissant un certain réalisme dans la simulation des recommandations.

Trois types de données sont mobilisées dans ce projet.

**1. Données de marché (ETF, indices)**

Les recommandations générées par GPT-4 mentionnent souvent des **ETF classiques** (ex. : SPY, AGG, QQQ) ou des **indices de référence** (ex. : S&P 500, MSCI World). Pour enrichir le prototype, il est utile de récupérer :

* Les **noms exacts** des fonds ou indices
* Leurs **performances passées**
* Leurs **caractéristiques principales** (volatilité, exposition géographique...)

Ces données sont récupérées via la bibliothèque Python **yfinance**, qui permet d’interroger gratuitement les données boursières de Yahoo Finance.

(*ETF : fonds indiciels cotés qui répliquent la performance d’un indice. Ex. : SPY réplique le S&P 500.)*

**2. Profils types d’investisseurs**

Les profils utilisateur ne sont pas collectés depuis une base réelle. Ils sont **définis manuellement** selon trois catégories classiques :

* Prudent (risque faible, forte part obligataire)
* Équilibré (répartition actions/obligations)
* Dynamique (majorité actions)

Ces profils permettent de simuler un questionnaire de robo-advisor, sans exposer de données personnelles.

(*Profils types : schémas standards utilisés par les conseillers pour adapter une stratégie aux préférences de l’investisseur.)*

**3. Réponses générées par le modèle (texte brut)**

Les réponses de GPT-4 sont considérées comme une **source de données à part entière**. Chaque réponse contient :

* Des titres d’actifs (souvent approximatifs)
* Des pourcentages d’allocation
* Des arguments qualitatifs (ex. : “diversification”, “stabilité”, “croissance à long terme”)

Ces éléments sont ensuite **extraits et restructurés**, puis comparés à des allocations types.

**Limites des sources**

* Les noms d’ETF cités par GPT-4 ne sont pas toujours précis ou existants : une vérification manuelle ou semi-automatisée est parfois nécessaire.
* Les données de performance historique ne sont pas directement intégrées dans le système, mais peuvent servir pour l’évaluation (section 2.5).
* Le modèle ne dispose pas d’accès en temps réel à l’actualité financière, donc les recommandations sont de nature générique.

**2.4 Développement de l’interface utilisateur via Streamlit**

Pour permettre à un utilisateur de tester facilement le prototype, une **interface graphique simple** a été développée à l’aide de **Streamlit**, un outil open source permettant de créer des applications web en Python.

L’objectif n’est pas d’avoir un produit final commercialisable, mais plutôt une interface :

* Intuitive
* Rapide à mettre en œuvre
* Adaptée à l’expérimentation

**Choix de Streamlit**

Streamlit a été retenu pour plusieurs raisons :

* Il ne nécessite pas de connaissance en HTML ou JavaScript.
* Il permet de créer une interface interactive avec très peu de code.
* Il s’intègre facilement avec l’API OpenAI et les bibliothèques de data science comme pandas, matplotlib ou yfinance.

**Fonctionnalités de l’interface**

L’application offre une expérience simple et fluide :

1. L’utilisateur choisit son **profil** dans un menu déroulant.
2. Il clique sur un bouton pour **générer une recommandation**.
3. Le **texte de GPT-4** s’affiche.
4. La composition du portefeuille est **extraite automatiquement**.
5. Un **graphique circulaire** illustre la répartition proposée.

Chaque étape est déclenchée automatiquement dans l’interface, sans rechargement de page, ce qui rend l’outil agréable à utiliser même en phase de test.

**Fonctionnalités supplémentaires intégrées**

Pour enrichir l’expérience utilisateur et aller au-delà du simple affichage, deux fonctionnalités ont été ajoutées au prototype :

* La **mémorisation des recommandations générées**, via une variable de session. L’utilisateur peut ainsi consulter les portefeuilles précédemment proposés, sans avoir à relancer GPT-4.
* La possibilité de **comparer plusieurs profils d’investisseurs en parallèle**. Chaque profil (prudent, équilibré, dynamique) peut être visualisé côte à côte, avec ses allocations spécifiques, permettant une lecture comparative immédiate.

Ces ajouts permettent à l’utilisateur de mieux comprendre la logique des recommandations, tout en gardant une interface simple et rapide à utiliser.

**Limites actuelles de l’interface (mise à jour)**

* L’application ne propose pas de personnalisation fine (âge, revenu, horizon d’investissement).
* Les recommandations ne sont pas sauvegardées sur le long terme (pas de base de données, seulement en session).
* Le traitement automatique du texte généré par GPT-4 peut échouer si la réponse est mal structurée.
* L’utilisateur n’est pas guidé pour interpréter la qualité d’un portefeuille (pas d’indicateur de performance intégré pour l’instant).

**2.5 Critères d’évaluation des recommandations**

Pour juger de la qualité et de la pertinence des recommandations générées par le prototype, plusieurs critères ont été définis. Ils permettent d’évaluer le système de façon simple, mais structurée, à la fois sur le **fond financier** et sur la **forme** de la réponse produite.

**1. Cohérence avec le profil utilisateur**

La première vérification concerne l’adéquation entre le portefeuille proposé et le profil sélectionné.  
Exemples attendus :

* Pour un profil **prudent** : majorité d’obligations, peu d’actions
* Pour un profil **dynamique** : forte exposition aux actions, diversification internationale possible

👉 Cela permet de vérifier si GPT-4 a bien “compris” le niveau de risque attendu.

**2. Diversification du portefeuille**

Un bon portefeuille contient généralement :

* Plusieurs classes d’actifs (actions, obligations…)
* Une exposition à différentes zones géographiques
* Une certaine variété sectorielle

Une recommandation proposante uniquement 2 ETF très proches est jugée comme **peu diversifiée**, donc moins robuste.

**3. Lisibilité et justification**

On évalue ici :

* La **clarté** de la recommandation (phrases compréhensibles, structure lisible)
* La **capacité à justifier les choix** (ex. : “ce fonds est choisi pour sa stabilité”)

Une recommandation sans explication ou mal rédigée est considérée comme faible, même si les ETF choisis sont cohérents.

**4. Structure et exactitude de la réponse**

Le texte généré est aussi évalué sur :

* La **présence de toutes les informations attendues** (titres des ETF, % d’allocation)
* Le **respect des contraintes** : total à 100 %, formats lisibles

On vérifie également si les ETF cités **existent réellement**, grâce à des outils comme Yahoo Finance. Les réponses contenant des noms inventés ou incorrects sont signalées comme “hallucinées”.

👉 \*(*Hallucination : réponse générée par le modèle mais factuellement fausse ou inventée.)*

**5. Éléments quantitatifs optionnels (selon le temps disponible)**

Si les données de marché sont intégrées à temps (via yfinance), on peut :

* Calculer le **rendement historique** du portefeuille simulé
* Estimer une **volatilité** simple
* Visualiser l’évolution dans le temps (graphique)

Cela permet une **analyse plus poussée**, mais reste optionnel si la partie technique n’est pas prioritaire dans l’évaluation finale du mémoire.

**Méthode d’évaluation prévue**

Un tableau d’analyse sera construit pour noter chaque recommandation selon :

* Cohérence profil
* Diversification
* Clarté / justification
* Exactitude / respect des contraintes

Chaque critère pourra être noté **sur 5**, avec des commentaires à l’appui.

Une fois l’architecture définie, les sources identifiées, et les critères d’évaluation établis, le prototype a été développé puis testé. Cette partie présente les principales fonctionnalités de l’application, les cas d’usage simulés, ainsi que les résultats concrets obtenus.

**3.1 Fonctionnalités de l’application**

L’application développée dans le cadre de ce mémoire est un prototype fonctionnel, accessible via une interface web simple. Elle permet à un utilisateur de recevoir une **recommandation d’investissement personnalisée** à partir d’un profil de risque, en s’appuyant sur un modèle de langage (GPT-4) intégré par API.

Le parcours utilisateur est volontairement court, afin de faciliter l’expérimentation. L’objectif n’est pas de reproduire une plateforme de gestion de patrimoine complète, mais de démontrer qu’un modèle comme GPT-4 peut générer des conseils financiers simples, compréhensibles et cohérents avec un profil type.

**Fonctionnalités principales**

1. **Choix du profil d’investisseur**  
   L’utilisateur sélectionne un des trois profils standards :  
   – Prudent  
   – Équilibré  
   – Dynamique

Chaque profil correspond à une tolérance au risque différente, et déclenche un type de réponse adapté.

1. **Génération automatique de la recommandation**  
   Un prompt spécifique est généré et envoyé à GPT-4, via l’API d’OpenAI. Le modèle retourne un texte contenant :

* Une liste d’ETF ou d’indices recommandés
* Des pourcentages d’allocation
* Une justification brève pour chaque choix

1. **Affichage clair et visuel**  
   Le texte est affiché directement dans l’application. Il est ensuite **interprété** pour en extraire :

* Les allocations proposées (en tableau)
* Un graphique circulaire représentant la répartition

1. **Mémorisation des résultats**  
   Chaque recommandation générée est conservée dans la session en cours. Cela permet de consulter l’historique sans avoir à relancer le modèle.
2. **Comparaison multi-profils (expérimentale)**  
   L’utilisateur peut visualiser côte à côte les portefeuilles générés pour les trois profils types. Cette fonctionnalité est utile pour observer les différences d’allocation selon la tolérance au risque.

**Interface et expérience utilisateur**

L’interface a été développée avec **Streamlit**, ce qui permet :

* Un accès via navigateur web local
* Une réactivité sans rechargement de page
* Une simplicité d’utilisation sans configuration préalable

L’ensemble est pensé pour des tests rapides, répétés, et facilement exploitables dans le cadre d’une analyse manuelle ou semi-automatisée.

**Accessibilité du prototype**

Le prototype a été conçu pour être **utilisable au-delà de l’environnement de développement local**. Il est déployé sur la plateforme **Streamlit Cloud**, ce qui permet d’y accéder depuis n’importe quel navigateur web, sans installation préalable.

Un **lien cliquable** sera fourni en annexe ou à la fin de ce mémoire, afin de permettre à un lecteur ou évaluateur de tester l’application directement. Cette mise en ligne vise à faciliter l’expérimentation, la démonstration, et à illustrer concrètement le fonctionnement du système développé.

**3.2 Scénarios testés**

Afin d’évaluer le fonctionnement du prototype dans différentes situations, trois scénarios utilisateurs ont été définis. Chacun correspond à un **profil d’investisseur** distinct, couramment utilisé dans les services de gestion d’actifs.

Ces scénarios permettent d’observer comment le modèle (GPT-4) s’adapte aux préférences de risque exprimées par l’utilisateur, et de vérifier si les recommandations sont cohérentes avec les standards du conseil financier.

**Scénario 1 : Investisseur prudent**

* **Objectif du profil** : préserver le capital, limiter l’exposition aux marchés volatils.
* **Attentes** : forte proportion d’actifs à revenu fixe (obligations), exposition minimale aux actions.
* **Ce qu’on cherche à vérifier** :
  + Si GPT-4 comprend la logique de sécurisation
  + Si les ETF proposés correspondent à des fonds obligataires ou défensifs
  + Si la part en actions est faible

**Scénario 2 : Investisseur équilibré**

* **Objectif du profil** : équilibre entre sécurité et croissance.
* **Attentes** : répartition typique de type 50/50 ou 60/40 entre actions et obligations.
* **Ce qu’on cherche à vérifier** :
  + Si la répartition est raisonnablement partagée
  + Si le portefeuille reste diversifié
  + Si le langage utilisé reflète bien l’idée d’un “compromis” de stratégie

**Scénario 3 : Investisseur dynamique**

* **Objectif du profil** : maximiser la croissance à long terme, acceptation d’une forte volatilité.
* **Attentes** : majorité d’actions, exposition à des marchés plus risqués (ex. : technologique, émergents).
* **Ce qu’on cherche à vérifier** :
  + Si GPT-4 propose une allocation orientée actions
  + Si des ETF à fort potentiel ou spécifiques sont proposés
  + Si le discours reflète une recherche de performance

**Méthode de test**

Pour chaque scénario :

* Une recommandation est générée via l’application
* Les résultats sont sauvegardés (session Streamlit)
* Chaque réponse est évaluée selon les critères définis en section 2.5 (cohérence, lisibilité, diversification, etc.)