



Impacto da IA no Ambiente Corporativo: Desenvolvimento de um Assistente com IA Generativa para Assessores de Investimentos

Lucas Scriptorre Marques Ferreira - a61449

Trabalho realizado sob a orientação de

Prof. Isabel Maria Lopes

Prof. Wesley Angelino de Souza

Mestrado em Engenharia Eletrotécnica e de Computadores

2024-2025

Impacto da IA no Ambiente Corporativo: Desenvolvimento de um Assistente com IA Generativa para Assessores de Investimentos

Relatório da UC de Projeto

Mestrado em Engenharia Eletrotécnica e de Computadores

Escola Superior de Tecnologia e Gestão

Lucas Scriptorre Marques Ferreira - a61449

2024-2025

A Escola Superior de Tecnologia e de Gestão não se responsabiliza pelas opiniões expressas neste relatório.

Declaro que o trabalho descrito neste relatório é da minha autoria e é da minha vontade que o mesmo seja submetido a avaliação.

Lucas Scriptore Marques Ferreira - a61449

Dedicatória

Dedico este trabalho à minha família, em especial aos meus pais, Samantha e Wolmer, pelo apoio incondicional, sobretudo nos momentos mais desafiadores. À minha irmã, Maria Clara, que inicia sua própria jornada acadêmica, desejo perseverança e sucesso. Dedico também à minha namorada e parceira, Maria, que esteve ao meu lado em todas as etapas desta caminhada, partilhando conquistas e incertezas, e sempre me ajudando a manter o foco e a organização. A todos, minha profunda gratidão.

Agradecimentos

Agradeço primeiramente ao Instituto Politécnico de Bragança e à Universidade Tecnológica Federal do Paraná, instituições que possibilitaram a realização deste mestrado em regime de duplo diploma. Expresso minha gratidão ao corpo docente de ambas as universidades, que contribuíram com conhecimento, apoio e orientação ao longo desta jornada acadêmica.

À minha orientadora, Professora Isabel Maria Lopes, agradeço pela dedicação, rigor técnico e incentivo constante ao desenvolvimento crítico e profissional. Ao Professor Wesley Angelino de Souza, coorientador, registro minha sincera gratidão pela disponibilidade e apoio fundamentais à condução deste projeto.

Sou igualmente grato aos colegas e parceiros da Wura Invest, que foram fonte de inspiração prática e referência para a construção da solução apresentada, contribuindo com insights reais e validações técnicas ao longo do desenvolvimento. Um agradecimento especial ao Gabriel Douradinho, gestor e amigo, que acolheu e incentivou a iniciativa com entusiasmo e visão, sendo fundamental ao apoiar e impulsionar este projeto desde os primeiros passos.

À minha família, em especial aos meus pais, Samantha e Wolmer, que sempre me deram base e força para buscar a minha melhor versão, e à minha namorada, Maria, que esteve ao meu lado em tantos momentos, oferecendo um apoio essencial, deixo meu profundo agradecimento.

Por fim, agradeço aos amigos, em especial ao Gregory, que compartilhou esta jornada em horas de trabalho conjunto e apoio mútuo durante a elaboração de sua própria dissertação. Ao Felipe e ao Henrique, que, além de companheiros de casa, foram peças

importantes nas trocas de experiências, conversas e boas risadas. E, não menos importante, ao meu grande amigo Luiz Fernando, que, apesar dos milhares de quilômetros de distância, sempre me motivou e inspirou.

A todos, meu muito obrigado. Vocês foram essenciais em cada passo até aqui.

Resumo

Esta dissertação explora a aplicação prática de modelos de linguagem de grande escala (LLMs), com foco no ChatGPT, no apoio à rotina de trabalho de assessores de investimentos. O objetivo foi desenvolver um assistente virtual personalizado, treinado com engenharia de prompt e dados reais do escritório Wura Invest, credenciado à XP Investimentos. O sistema foi projetado para realizar tarefas como análise de carteiras de clientes, sugestões de alocação, preparação de reuniões e geração de relatórios de performance. A metodologia adotada envolveu o design iterativo do *prompt* mestre, a curadoria de materiais internos (carteiras, cenários econômicos, posicionamentos) e a realização de testes simulados para validar a eficácia da ferramenta. Os resultados indicaram que o modelo é capaz de realizar análises técnicas coerentes, alinhar recomendações ao perfil do cliente e gerar valor real à prática da assessoria. As limitações observadas estão relacionadas à extração de dados não estruturados e à ausência de conexão com APIs externas, apontando caminhos para futuras melhorias. O projeto contribui para o avanço do uso de Inteligência Artificial (IA) generativa no setor financeiro, ao demonstrar como soluções baseadas em *Large Language Models* (LLM) podem ser adaptadas com sucesso a contextos profissionais específicos.

Palavras-chave: inteligência artificial, modelos de linguagem, ChatGPT, assessoria de investimentos, engenharia de prompt.

Abstract

This dissertation explores the practical application of large language models (LLMs), focusing on ChatGPT, to support the daily work of investment advisors. The objective was to develop a personalized virtual assistant trained through prompt engineering and real-world data from Wura Invest, a financial advisory firm accredited by XP Inverimentos. The system was designed to perform tasks such as client portfolio analysis, allocation suggestions, meeting preparation, and performance reporting. The adopted methodology involved iterative prompt design, curation of internal materials (portfolios, economic scenarios, strategic positioning), and simulated tests to validate the assistant's effectiveness. Results showed that the model can deliver coherent technical analysis, align recommendations with client profiles, and add tangible value to financial advisory practices. Limitations include challenges in extracting unstructured data and lack of API integration, highlighting opportunities for future improvements. The project contributes to the advancement of generative AI in the financial sector by demonstrating how LLM-based solutions can be successfully adapted to specific professional contexts.

Keywords: artificial intelligence, language models, ChatGPT, investment advisory, prompt engineering.

Conteúdo

1	Introdução	1
1.1	Contextualização	1
1.2	História da Inteligência Artificial	2
1.3	O Chat GPT	3
1.4	Modelo de linguagem de larga escala	4
1.5	Limitações do Chat GPT	6
1.6	IA no mercado de trabalho	10
1.7	Objetivos gerais	11
1.8	Objetivos específicos	12
2	Metodologia	15
2.1	Tipo de pesquisa e abordagem	15
2.2	Etapas de desenvolvimento	17
2.3	Tecnologias e ferramentas	20
2.3.1	ChatGPT 4.0	20
2.3.2	Microsoft Excel	21
2.3.3	Documentos PDF	21
2.3.4	Reconhecimento Óptico de Caracteres	22
2.3.5	Notion e Microsoft Word	22
3	Desenvolvimento e implementação do assistente personalizado	23
3.1	Visão Geral do Projeto	23

3.2	Estrutura Lógica do Assistente	25
3.2.1	Entrada de Dados	25
3.2.2	Processamento Interno	26
3.2.3	Geração da Resposta	27
3.3	Engenharia de Prompt	28
3.3.1	Fundamentos da Engenharia de Prompt	28
3.3.2	Estrutura do <i>Prompt</i> Mestre	30
3.3.3	Finalidade da Engenharia de Prompt	31
3.3.4	Disponibilização do Conteúdo	31
3.4	Considerações sobre Limites, Segurança e Desafios	32
3.4.1	Proteção de Dados Pessoais e Informações Sensíveis	32
3.4.2	Conformidade com Diretrizes da Wura Invest	33
3.4.3	Limitações Técnicas da Plataforma	33
3.4.4	Estratégias de Mitigação	34
3.4.5	Papel do Assessor de investimento	35
4	Testes práticos	37
4.1	Metodologia dos Testes	38
4.1.1	Estrutura dos Testes	38
4.1.2	Tipos de <i>Input</i> Utilizados	39
4.1.3	Critérios de Avaliação	39
4.2	Teste 1: Análise Geral de Carteira com Novo Aporte	40
4.2.1	Objetivo do teste	40
4.2.2	Contexto simulado	40
4.2.3	Prompt utilizado	40
4.2.4	Explicação técnica do <i>prompt</i>	41
4.2.5	Resumo da resposta do GPT	42
4.2.6	Análise crítica do resultado	46
4.2.7	Conclusão do teste	47

4.3	Teste 2 - Análise de performance e geração de relatório	47
4.3.1	Objetivo do teste	47
4.3.2	Contexto simulado	47
4.3.3	<i>Prompt</i> utilizado	48
4.3.4	Explicação técnica do <i>prompt</i>	48
4.3.5	Resumo das respostas do GPT	49
4.3.6	Análise crítica do resultado	52
4.3.7	Conclusão do teste	53
4.4	Teste 3 - Apoio do Assistente na Preparação de Reuniões com <i>Prospects</i> . .	53
4.4.1	Objetivo do teste	53
4.4.2	Contexto simulado	54
4.4.3	Prompt utilizado	54
4.4.4	Explicação técnica do <i>prompt</i>	55
4.4.5	Resposta gerada pelo GPT	56
4.4.6	Análise crítica do resultado	58
4.4.7	Conclusão do teste	58
4.5	Teste 4: Análise Técnica Avançada de Carteira	59
4.5.1	Objetivo do teste	59
4.5.2	Contexto simulado	59
4.5.3	<i>Prompt</i> utilizado	60
4.5.4	Explicação técnica dos <i>prompts</i>	61
4.5.5	Resumo das respostas do GPT	62
4.5.6	Análise crítica dos resultados	70
4.5.7	Conclusão do teste	70
5	Conclusões	73
A	Prompt Mestre	A1
B	Teste 1: Relatório Gerado pelo GPT	B1

Lista de Tabelas

1	Comparação entre um GPT personalizado e um agente de inteligência artificial autônomo.	10
2	Tipos de <i>input</i> utilizados nos testes práticos do modelo GPT.	39
3	Critérios utilizados para avaliação das respostas geradas pelo assistente GPT. . .	39
4	Análise crítica do resultado conforme critérios estabelecidos (Teste 1) . . .	46
5	Análise crítica do resultado conforme critérios estabelecidos (Teste 2) . . .	52
6	Análise crítica do resultado conforme critérios estabelecidos (Teste 3) . . .	58
7	Análise crítica do resultado conforme critérios estabelecidos (Teste 4) . . .	70

Lista de Figuras

1	Interface do "InvestAI - Assistente de Alocação" da Wura Invest.	25
2	Interpretação e análise da carteira no teste 1.	42
3	Composição da carteira no teste 1.	43
4	Diagnóstico da carteira no teste 1.	43
5	Proposta de renda no teste 1.	43
6	Sugestão de alocação para novo aporte no teste 1.	44
7	Alocação final no teste 1.	45
8	Recomendações adicionais no teste 1.	45
9	Recomendações finais no teste 1.	46
10	Análise de Performance no teste 2.	50
11	<i>Ranking</i> de ativos no teste 2.	50
12	Cenário econômico no teste 2.	51
13	Oportunidades e Riscos no teste 2.	51
14	Perspectivas no teste 2.	52
15	Leitura analítica no teste 4.	63
16	Análise detalhada de Renda Fixa no teste 4.	64
17	Análise detalhada de Renda Fixa 2 no teste 4.	64
18	Análise detalhada de Renda Variável no teste 4.	65
19	Revisão estratégica no teste 4.	65
20	Sugestões personalizadas no teste 4.	66
21	Sugestão personalizada no teste 4.	67

22	Identificação de fatores no teste 4.	67
23	Montagem de reserva no teste 4.	68
24	Recomendações no teste 4.	68
25	Simulação para renda no teste 4.	69
26	Carteiras consolidadas criadas no teste 4.	69

Siglas

ANBIMA Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais.
34, 74

API *Application Programming Interfaces*. x, 13, 34, 74

CRM *Customer Relationship Management*. 74, 75

ETF *Exchange Traded Fund*. 26

FGC Fundo Garantidor de Crédito. 26

FII Fundo de Investimento Imobiliário. 42, 46, 48, 51, 60, 70

GPT *Generative Pre-Trained Transformer*. 2, 3, 12, 16, 20–23, 25, 33, 37–39, 47, 53–55,
58, 59, 70, 71, 73

IA Inteligência Artificial. x, 1–4, 6, 7, 9–11, 70, 71, 74

IPCA Índice Nacional de Preços ao Consumidor Amplo. 48, 50, 51, 70

LLM *Large Language Models*. x, 2, 4–10, 12, 28, 29, 62, 74

OCR Reconhecimento Óptico de Caracteres. 22, 34

PJ1 Pessoa Jurídica 1. 33

PJ2 Pessoa Jurídica 2. 33

RGPD Regulamento Geral sobre a Proteção de Dados. 7

Capítulo 1

Introdução

1.1 Contextualização

A abordagem do impacto da Inteligência Artificial (IA) no ambiente empresarial é de extrema relevância no contexto atual, dada a rápida evolução e a ampla adoção dessas tecnologias em diferentes setores. Como destacado por Westerman et al. [1] em seu modelo de maturidade digital, as organizações se encontram em estágios distintos de transformação digital, sendo a integração da IA um passo fundamental para a sua consolidação estratégica e competitiva. No entanto, essa adoção nem sempre ocorre de forma estruturada, especialmente em áreas que dependem de análise técnica e atendimento consultivo, como o setor financeiro.

A IA tem se mostrado um relevante motor de mudança na redefinição de processos, no aumento da eficiência operacional e na criação de novas oportunidades de negócio. Um estudo conduzido pelo McKinsey Global Institute [2] estima que cerca de 50% das atividades laborais atuais poderiam ser automatizadas com tecnologias já disponíveis. Especificamente no campo da IA generativa, Brynjolfsson et al. [3] demonstraram que a introdução de assistentes conversacionais pode elevar em 14% a produtividade de agentes de suporte ao cliente, com efeitos especialmente benéficos entre trabalhadores menos experientes, ao promover a disseminação do conhecimento organizacional.

Com a popularização dos modelos de linguagem de grande porte (*Large Language Model*), como o ChatGPT, a IA generativa passou a desempenhar um papel disruptivo nas formas de interação entre humanos e sistemas computacionais. Lançado pela OpenAI em 2022, o ChatGPT alcançou rápida adoção e vem sendo integrado a plataformas corporativas, fluxos de trabalho automatizados e serviços baseados em linguagem natural.

No setor financeiro, essas aplicações têm se mostrado particularmente relevantes. Ferramentas baseadas em LLM já são utilizadas para interpretar grandes volumes de dados, gerar relatórios sob medida e oferecer suporte consultivo. Compreender como essas soluções são utilizadas, adaptadas e avaliadas em contextos reais torna-se fundamental para profissionais que desejam alinhar inovação tecnológica à entrega de valor, ainda mais com a rapidez das evoluções atualmente.

Esta pesquisa, portanto, insere-se nessa intersecção entre tecnologia, mercado e usabilidade, explorando o desenvolvimento e a aplicação de um modelo *Generative Pre-Trained Transformer* (GPT) personalizado como apoio à atividade de assessoria de investimentos.

1.2 História da Inteligência Artificial

Apesar de estar em alta atualmente, a história da inteligência artificial é rica e complexa, que remonta à metade do século XX, com suas origens conceituais traçadas ainda mais longe no tempo. O termo "inteligência artificial" foi formalmente cunhado em 1956 durante a Conferência de Dartmouth, um evento que reuniu alguns dos pioneiros do campo, como John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon. Este evento é amplamente considerado o ponto de partida oficial da IA como um campo de estudo distinto [4].

Nas últimas duas décadas, a IA experimentou um ressurgimento significativo, impulsionado por avanços em poder computacional, algoritmos mais sofisticados e a disponibilidade massiva de dados. Um marco recente e significativo nessa evolução foi o surgimento do ChatGPT, desenvolvido pela OpenAI e lançado em novembro de 2022. O ChatGPT

é baseado na arquitetura GPT, que utiliza técnicas de aprendizado profundo para processar e gerar linguagem natural de maneira coerente e contextualmente relevante [5][6]. O lançamento dessa ferramenta teve um papel muito importante na democratização e popularização da IA para o público geral.

De acordo com uma pesquisa da *Precedence Research*, o mercado global de inteligência artificial foi avaliado em 538,13 bilhões de dólares em 2023 e espera-se que atinja cerca de 2.575,16 bilhões de dólares até 2032, com uma taxa de crescimento anual composta (CAGR) de 19% [7]. Além disso, o investimento corporativo global em IA aumentou de 14,57 bilhões de dólares em 2013 para 189 bilhões de dólares em 2023, representando um aumento de 13 vezes. Em 2022, o investimento corporativo global em IA atingiu 92 bilhões de dólares [8].

1.3 O Chat GPT

O ChatGPT é um modelo de linguagem de inteligência artificial que rapidamente se destacou por sua capacidade de gerar respostas coerentes e contextualmente relevantes em uma diversidade de tópicos. Segundo dados recentes, o ChatGPT alcançou 1 milhão de usuários em apenas cinco dias após seu lançamento e, em janeiro de 2023, já contava com 100 milhões de usuários mensais ativos, tornando-se a aplicação de consumo de crescimento mais rápido na história [9]. Até junho de 2023, esse número havia crescido para 200 milhões de usuários mensais ativos [7].

A importância do ChatGPT é evidenciada por sua ampla adoção em diversos setores, desde atendimento ao cliente até educação e assistência pessoal em tarefas simples do dia a dia. Oliveira Júnior [10] observa que "é uma maneira nova jamais vista de gerar e transmitir conhecimento", ressaltando o potencial transformador dessa tecnologia. No mundo empresarial, o ChatGPT tem sido amplamente adotado para melhorar a eficiência operacional e a experiência do cliente. Empresas têm utilizado essa tecnologia para automatizar atendimento ao cliente, gerar conteúdo de *marketing* e até mesmo auxiliar em processos de recrutamento e seleção, demonstrando seu impacto significativo na produtividade e

inovação empresarial.

Contudo, a rápida adoção do ChatGPT também levantou questões importantes sobre ética, privacidade e o futuro do trabalho. À medida que a tecnologia continua a evoluir, debates sobre seu uso responsável e seu impacto na sociedade se intensificam. Questões como a substituição de empregos, a privacidade dos dados e os vieses algorítmicos são tópicos de discussão cruciais para garantir que a integração da IA ocorra de maneira ética e sustentável [4][11].

1.4 Modelo de linguagem de larga escala

Os modelos de linguagem de larga escala (LLM) representam uma inovação de paradigma no campo da inteligência artificial aplicada à linguagem humana [12]. Em essência, são modelos computacionais de IA desenvolvidos para compreender e gerar texto em linguagem natural de forma semelhante à produzida por humanos [13]. Esses sistemas de aprendizagem profunda são treinados com conjuntos de documentos e materiais massivos contendo bilhões de palavras, extraídas de livros, artigos, *websites* e outras fontes, de modo que aprendem padrões linguísticos complexos (sintáticos e semânticos) presentes no idioma. A dimensão desses modelos impulsiona suas capacidades, como por exemplo, o modelo GPT-3 da OpenAI possui 175 bilhões de parâmetros, uma escala de ordem de magnitude superior a modelos anteriores, o que lhe permite alcançar desempenho notável em diversas tarefas de processamento de linguagem natural [14]. A combinação de arquiteturas modernas, como os transformers, com conjuntos de dados extensos resultou em modelos altamente versáteis, capazes de produzir respostas coerentes e contextualizadas a partir de um *prompt* dado, isto é, conseguem executar de forma rápida e precisa tarefas linguísticas que antes estavam reservadas apenas a intervenções humanas [12].

De forma simplificada, um LLM funciona prevendo a próxima palavra ou frase com base nas anteriores, utilizando para isso o conhecimento estatístico adquirido durante o treino. Durante a fase de treinamento, o modelo ajusta seus parâmetros internos para

minimizar erros na predição de sequências de palavras, aprendendo assim as regularidades e relações do idioma a partir de enormes volumes de textos [13]. Como resultado, quando recebe uma entrada, ou mais comumente chamada de *prompt* (pergunta ou comando), o LLM consegue gerar continuamente uma resposta textual que seja coerente com o contexto, mantendo o sentido e a fluidez, além de ser extremamente direta ao que foi solicitado. Vale destacar que essas respostas não são cópias exatas do que foi visto nos dados de treino, mas sim conteúdos inéditos produzidos com base nos padrões assimilados, ou seja, o modelo “aprende” a generalizar a língua, em vez de simplesmente copiar e colar as frases. Esse mecanismo permite que os LLMs atendam a uma ampla variedade de solicitações, indo desde explicações sobre um conceito até a criação de textos criativos, sempre adaptando a saída ao contexto fornecido. Em suma, o funcionamento dos LLMs apoia-se na identificação de padrões linguísticos em larga escala e na geração probabilística de texto, o que lhes confere uma capacidade sem precedentes de interação em linguagem natural.

Entretanto, a eficácia das aplicações utilizando essa linguagem depende diretamente da forma como o usuário formula o comando de entrada. A construção cuidadosa do *prompt* é crucial para orientar o modelo a produzir respostas relevantes e corretas, tanto que a chamada engenharia de *prompt* emergiu como técnica fundamental para maximizar a utilidade e a precisão desses sistemas [15]. Pesquisas demonstram que até mesmo ajustes sutis na formulação do pedido podem levar a diferenças significativas no desempenho do modelo [16]. Por exemplo, um estudo observou variações de acurácia na ordem de 40% em uma tarefa de tradução de código apenas alterando o formato do *prompt* [16]. De modo geral, um *prompt* vago ou sem contexto tende a gerar respostas superficiais ou imprecisas [17], ao passo que um comando claro, contextualizado e detalhado permite ao LLM fornecer resultados muito mais úteis e de alta qualidade. Cada elemento do *prompt*, das palavras escolhidas ao tom e estilo da solicitação, influencia o conteúdo gerado [17]. Portanto, *prompts* bem elaborados, que explicitem adequadamente o contexto e o objetivo da consulta, são essenciais para aproveitar todo o potencial dos LLMs nos cenários práticos, influenciando diretamente a qualidade, utilidade e precisão das respostas

produzidas.

1.5 Limitações do Chat GPT

Os modelos de IA com linguagem LLM, como o ChatGPT, apresentam um enorme potencial em diversas aplicações. No entanto, apesar dos avanços, estes sistemas enfrentam limitações importantes de natureza técnica, ética e prática, tais como alucinações, dependência de dados de treino, fragilidade contextual, viés algorítmico, riscos à privacidade, respostas imprecisas e ausência de raciocínio crítico. E, claro, como toda tecnologia recente, consequências que ainda não foram mapeadas podem acontecer.

É fundamental reconhecer estas fragilidades pois elas podem comprometer a fiabilidade e segurança do seu uso [18] [19]. Assim, uma análise clara dessas limitações é necessária para orientar o uso responsável e o desenvolvimento de regulamentações adequadas no domínio da IA generativa [20].

Uma das limitações técnicas mais discutidas dos modelos de linguagem é a tendência à denominada alucinação de conteúdo, ou seja, a geração de informações aparentemente plausíveis, mas incorretas ou infundadas [19]. Este fenômeno deve-se, em parte, ao fato de o modelo basear as suas respostas em padrões estatísticos extraídos dos dados de treino, sem acesso direto a fontes verificadas em tempo real. Estudos recentes demonstram que a ocorrência de alucinações permanece significativa: numa análise sobre revisões sistemáticas, o modelo GPT-4 forneceu referências fictícias em cerca de 28% dos casos analisados [21].

Além disso, os LLMs apresentam dependência integral dos dados utilizados durante o seu processo de treino, não possuindo mecanismos de aprendizagem contínua nem atualizações dinâmicas de conhecimento [22]. Em consequência, exibem um conhecimento limitado à época correspondente ao conjunto de dados em que foram treinados, podendo desconhecer acontecimentos mais recentes ou necessitando de atualizações manuais frequentes. Soma-se a isto a limitação da janela de contexto, que define a quantidade máxima

de informação que o modelo pode considerar numa interação; variações sutis na formulação de perguntas podem, assim, levar a respostas divergentes.

Em tarefas que exigem raciocínio lógico em múltiplas etapas ou cálculos sucessivos, estas restrições tornam-se mais evidentes. Diversos estudos indicam que, nestes casos, o modelo pode apresentar respostas incoerentes ou incorretas quando o contexto não é cuidadosamente estruturado [18]. Recomenda-se, por isso, cautela na formulação de perguntas complexas, dado que premissas equivocadas tendem a originar respostas igualmente errôneas, sem que o sistema seja capaz de detectar ou corrigir tal falha. Em síntese, no seu estado atual, os modelos de linguagem requerem validação humana permanente em contextos complexos, devido à sua suscetibilidade a erros factuais e limitações de memória contextual.

As limitações técnicas discutidas anteriormente agravam preocupações de ordem ética e social. Uma das questões centrais reside no viés algorítmico. Modelos de LLM são treinados com enormes volumes de dados textuais extraídos da *Internet*, os quais podem refletir preconceitos culturais, estereótipos e desigualdades históricas. Na ausência de controlos eficazes, estes modelos podem não apenas reproduzir, mas também amplificar os vieses existentes [19], apresentando respostas tendenciosas em função de gênero, etnia ou orientação política. Tal cenário levanta sérias discussões relacionadas à justiça, à igualdade e à não discriminação perante a sociedade em que estamos inseridos.

Outro desafio ético relevante refere-se à privacidade e à segurança da informação, algo que já é pauta frequente com as redes sociais. Caso o conjunto de dados de treino contenha informações pessoais ou confidenciais, existe o risco de que o modelo reproduza trechos desses dados nas suas respostas [22]. Além disso, a própria utilização destas ferramentas exige, frequentemente, o envio de conteúdos sensíveis para servidores externos, levantando dúvidas quanto à conformidade com legislações de proteção de dados, como o Regulamento Geral sobre a Proteção de Dados (RGPD). Um exemplo emblemático dessa preocupação ocorreu em 2023, quando o ChatGPT foi temporariamente suspenso na Itália por alegado descumprimento das normas de privacidade.

Destaca-se também a possibilidade de uso indevido da IA generativa. A capacidade

dos LLMs de gerar conteúdos realistas pode ser explorada para fins maliciosos, incluindo a disseminação de desinformação, a criação de conteúdos fraudulentos ou mesmo a incitação a comportamentos ilegais [19]. Dada a dificuldade de distinguir, em muitos casos, o conteúdo gerado artificialmente de textos legítimos, cresce a preocupação com a integridade e autoria da informação, além do potencial impacto social negativo. Tais limitações éticas exigem, portanto, supervisão rigorosa, políticas claras de uso responsável e mecanismos de mitigação de risco desde a concessão até à implementação destas tecnologias, antes mesmo de serem distribuídas ao usuário final.

Para além das implicações éticas, os LLMs também enfrentam restrições práticas importantes. Em tarefas cotidianas, como responder a perguntas ou resumir textos, o ChatGPT pode produzir respostas imprecisas, incompletas ou até incorretas, sobretudo quando confrontado com solicitações ambíguas ou com pressupostos errados. O seu discurso fluente pode ocultar o fato de que não compreende verdadeiramente os significados que transmite — trata-se, afinal, de um sistema estatístico que prevê a sequência mais provável de palavras, e não de um agente dotado de compreensão semântica ou senso crítico [18].

Além disso, o modelo não possui mecanismos nativos de verificação factual nem raciocínio lógico avançado. Ao receber perguntas com premissas falsas, tende a aceitá-las sem contestação, resultando em respostas equivocadas. O mesmo ocorre em tarefas que exigem raciocínio causal, julgamento de valor ou inferência baseada em contexto externo ao texto fornecido [23].

Outra limitação prática significativa é a ausência de memória de longo prazo integrada, ou seja, os LLMs não retêm interações passadas, salvo dentro da janela de contexto da sessão ativa. Isso limita seu uso em cenários que exijam continuidade, personalização ou aprendizagem adaptativa, como tutoria educacional, acompanhamento clínico ou assistência jurídica. Dado esse conjunto de restrições, recomenda-se que o conteúdo gerado por essas ferramentas seja sempre revisto por profissionais qualificados, especialmente em contextos acadêmicos e profissionais que exijam elevado rigor [21], como feito no ambiente de testes deste trabalho. O consenso atual na literatura especializada é claro: os LLMs

devem ser usados como apoio à tomada de decisão, nunca como substitutos absolutos da análise humana.

Por fim, uma das mais relevantes é o seu caráter fundamentalmente reativo, ou seja, estes modelos operam exclusivamente em resposta a comandos explícitos, sem iniciativa própria ou tomada de decisão, o que limita a sua eficácia em cenários que exijam previsão, autonomia ou ação contínua [24]. Em outras palavras, um LLM típico não realiza ações por conta própria além do que lhe for solicitado diretamente via *prompt*.

Para superar essas limitações, emergiu o conceito de agentes de inteligência artificial, entendidos como uma evolução natural dos LLMs tradicionais. Um agente de IA é um sistema projetado para atuar de forma autônoma, sendo capaz de interpretar contextos, planejar soluções, tomar decisões com certa autonomia e executar tarefas complexas com base em metas definidas pelo utilizador [25]. Diferentemente dos *chatbots* convencionais, que operam com uma entrada e resposta por vez, os agentes podem conduzir ações sequenciais encadeadas, ajustando seus comportamentos com base em *feedback* e novas informações [26]. Por exemplo, enquanto um modelo como o ChatGPT pode sugerir os melhores aspiradores de pó disponíveis no mercado com base na sua base de dados imensa do treinamento, um agente de IA poderia, de forma autônoma, comparar preços, selecionar a melhor opção e concluir a compra, operando em nome do utilizador [26]. Esta capacidade de iniciativa e tomada de decisão representa uma mudança de paradigma na área da inteligência artificial, ao introduzir proatividade às já avançadas capacidades linguísticas dos LLMs.

Reconhecer essas limitações não invalida o valor dos modelos generativos, mas reforça a necessidade de uma utilização crítica e responsável. A comunidade científica, as organizações internacionais e os reguladores têm salientado que os avanços tecnológicos devem ser acompanhados de mecanismos de governança adequados, promovendo a segurança, a transparência e a ética no uso da inteligência artificial. A título de exemplo, a UNESCO tem liderado esforços internacionais para estabelecer diretrizes éticas globais para IA, enquanto a União Europeia avança com o AI Act, legislação pioneira que visa regulamentar sistemas de inteligência artificial de alto risco [20].

Aspecto	GPT Personalizado (ChatGPT)	Agente de IA Autônomo
Atuação	Reativa (responde apenas quando solicitado)	Proativa (atua de forma autônoma)
Memória	Opcional e limitada	Pode ser persistente e adaptativa
Execução de ações	Não executa ações diretamente	Pode interagir com APIs, sistemas e ferramentas externas
Complexidade técnica	Baixa – criado via interface	Alta – requer programação e integração
Planeamento de tarefas	Inexistente – responde a comandos isolados	Pode planejar e ajustar sequências de ações
Exemplo de uso	Assistente de escrita ou FAQ interativo	Agente que gere compromissos ou realiza compras online

Tabela 1: Comparação entre um GPT personalizado e um agente de inteligência artificial autônomo.

Tirar pleno proveito do potencial transformador dos LLMs exige, portanto, um equilíbrio entre inovação e responsabilidade. Proteger os dados dos utilizadores e garantir que a IA seja sempre supervisionada por seres humanos. Com essa abordagem, é possível integrar ferramentas como o ChatGPT de forma ética e eficaz no cotidiano das organizações e da sociedade, complementando a inteligência humana e não a substituindo.

1.6 IA no mercado de trabalho

A inteligência artificial está transformando profundamente o ambiente de trabalho, redefinindo processos, aumentando a eficiência e criando novas oportunidades em diversos setores. Esta revolução tecnológica está impactando significativamente a forma como as empresas operam e gerenciam os seus recursos.

Um dos principais impactos da IA no ambiente de trabalho é a automação de tarefas repetitivas e rotineiras. De acordo com um estudo da *McKinsey Global Institute*, "cerca de 50% das atividades de trabalho atuais poderiam ser automatizadas usando tecnologias já existentes"[27]. Isto libera os funcionários para se concentrarem em tarefas mais complexas e criativas, potencialmente aumentando a produtividade e a eficiência.

Na área de atendimento ao cliente, por exemplo, *chatbots* e assistentes virtuais estão

a tornar-se cada vez mais comuns. A Zendesk reporta que "30% dos gestores de empresas afirmaram estar usando *chatbots* e automação para melhorar o atendimento ao cliente"[28]. Estes sistemas de IA podem lidar com consultas básicas, como se fossem filtros iniciais, permitindo que os agentes humanos se concentrem em questões mais complexas e sensíveis, sendo acionados somente quando for estritamente necessário.

No setor financeiro, a IA está sendo utilizada para detecção de fraudes e análise de risco. O JPMorgan Chase, por exemplo, implementou um sistema de IA chamado COiN (*Contract Intelligence*) que "pode analisar documentos legais e extrair cláusulas importantes em segundos, uma tarefa que antes levava 360.000 horas de trabalho por ano"[29].

Na gestão de recursos humanos, a IA está sendo aplicada para otimizar processos de recrutamento e seleção. Algoritmos de *machine learning* podem analisar currículos e identificar os candidatos mais adequados para determinadas posições, reduzindo o tempo e os custos associados ao processo de contratação [30].

A IA também está revolucionando a tomada de decisões estratégicas nas empresas. Sistemas de *business intelligence* alimentados por IA podem analisar grandes volumes de dados e fornecer *insights* valiosos para os gestores. Um estudo da PwC mostrou que "72% dos executivos acreditam que a IA será fundamental para o futuro dos negócios"[31].

No entanto, a implementação da IA no ambiente de trabalho também traz desafios. Questões éticas, como privacidade de dados e vieses algorítmicos, precisam ser cuidadosamente consideradas [32]. Além disso, há preocupações sobre o deslocamento de empregos e a necessidade de requalificação da força de trabalho [33]. Algo que foi abordado com maior profundidade em 1.5.

1.7 Objetivos gerais

Considerando o potencial transformador das tecnologias de IA generativa e a sua crescente aplicabilidade em setores como o financeiro, este trabalho propõe-se a desenvolver uma solução prática com base nesses avanços.

Desenvolver, configurar e testar um modelo personalizado baseado em linguagem natural (GPT), com tecnologia da OpenAI, com o propósito de apoiar assessores de investimentos no diagnóstico, análise e recomendação de carteiras de clientes, liberando tempo para tarefas que requerem energia e foco humano, como prospecção comercial e relacionamento com clientes, por exemplo. O modelo foi implementado no ambiente ChatGPT, na versão 4.0, por meio de técnicas de engenharia de *prompt* e integração de conteúdo especializado e altamente personalizado, representando uma aplicação prática de inteligência artificial generativa baseada em modelos LLM.

Este trabalho tem como objetivo avaliar a viabilidade, eficiência e escalabilidade da ferramenta enquanto assistente virtual especializado, capaz de interpretar contextos financeiros reais, sugerir alocações alinhadas ao perfil de risco e aos objetivos do cliente, e apoiar o relacionamento com o investidor através da geração de relatórios interpretativos e interativos.

1.8 Objetivos específicos

Para alcançar o objetivo geral proposto, este trabalho estabelece os seguintes objetivos específicos:

- Implementar um GPT personalizado na ferramenta da OpenAI, com instruções estruturadas e fundamentadas, base de conhecimento personalizada e fluxos de interação adequados ao contexto de assessoria de investimentos;
- Simular diferentes cenários de uso com base em carteiras reais e situações representativas do cotidiano de um assessor de investimentos;
- Avaliar a capacidade do modelo em:
 - Identificar riscos, liquidez, concentração e oportunidades de alocação;
 - Interpretar relatórios financeiros (ex.: XP Performance) e planilhas de posição;
 - Sugerir alocações coerentes com o perfil de risco e os objetivos reais dos clientes;

- Gerar relatórios de acompanhamento e *scripts* para reuniões com clientes;
- Validar o desempenho da ferramenta por meio de testes práticos e análise qualitativa dos *outputs* gerados;
- Propor melhorias futuras na arquitetura do modelo, bem como formas de integração com APIs externas e possibilidades de evolução no uso da inteligência artificial generativa no setor financeiro.

Capítulo 2

Metodologia

Neste capítulo, apresenta-se a abordagem metodológica adotada, desde a definição do tipo de pesquisa até a descrição das ferramentas e técnicas utilizadas para o desenvolvimento e validação da solução. A pesquisa foi conduzida por meio de um processo iterativo, experimental e orientado à aplicação prática, sendo composta por etapas de concepção, implementação e testes simulados com base em dados reais do mercado financeiro e do ambiente inserido.

As seções a seguir detalham o tipo de pesquisa e abordagem adotadas, as fases do desenvolvimento da solução, as tecnologias empregadas e a estratégia utilizada para validação da ferramenta por meio de simulações.

2.1 Tipo de pesquisa e abordagem

A presente investigação caracteriza-se como uma pesquisa de natureza aplicada, uma vez que tem como finalidade propor uma solução prática para um problema real enfrentado por profissionais da área financeira: a busca por maior eficiência, personalização e suporte técnico no processo de assessoria de investimentos. Segundo Silva e Menezes [34], a pesquisa aplicada visa gerar conhecimentos voltados para a resolução de problemas específicos, com potencial de aplicação imediata. Alinhada a essa definição, esta dissertação propõe o desenvolvimento e avaliação de um assistente virtual personalizado, baseado

em inteligência artificial generativa, através de um GPT personalizado na plataforma ChatGPT em sua versão 4.0, com o intuito de apoiar a tomada de decisão em carteiras de investimento.

A abordagem adotada é predominantemente qualitativa, pois busca compreender e interpretar, de forma descritiva, como a ferramenta responde a diferentes cenários de uso reais. Ainda que alguns elementos quantitativos sejam considerados, o foco central está na análise de contexto, desempenho interpretativo e adequação funcional do modelo. Ainda de acordo com Silva e Menezes [34], a pesquisa qualitativa preocupa-se com a compreensão de fenômenos complexos em contextos específicos, sendo adequada para situações em que os dados não podem ser reduzidos exclusivamente a números.

No que se refere à estratégia de desenvolvimento, trata-se de uma pesquisa experimental, entendida aqui como aquela em que se realiza uma intervenção direta e controlada para testar uma solução. Conforme Gil [35], este tipo de investigação envolve a manipulação de variáveis e a observação dos efeitos decorrentes dessa manipulação sobre determinado fenômeno. No contexto desta dissertação, essa abordagem concretiza-se por meio da concepção, implementação e avaliação de um protótipo funcional, um modelo GPT personalizado, desenvolvido no ambiente da OpenAI, alimentado com dados e instruções específicos do mercado financeiro e personalizados com informações das empresas: Wura Invest e XP investimentos.

A validação da solução foi conduzida por meio de testes simulados, que replicaram situações comuns no dia a dia de um assessor de investimentos, como a análise de carteiras reais, sugestões de alocação para novos aportes, preparação para reuniões com clientes e elaboração de relatórios periódicos. Esses testes permitiram avaliar a capacidade interpretativa, analítica e consultiva do modelo, sempre considerando a especificidade de cada cliente, como os objetivos e perfil de risco, e o cenário macroeconômico vigente. Essa forma de validação insere-se no que a literatura de pesquisa aplicada em engenharia classifica como desenvolvimento experimental [36], que envolve testar soluções por meio de prototipagem, simulação e iteração. A combinação dessas abordagens metodológicas permitiu não apenas propor, mas também testar e refinar uma solução com potencial de

aplicação prática no contexto da assessoria financeira para o escritório específico.

O desenvolvimento do modelo foi orientado por demandas operacionais reais da Wura Invest, escritório de assessoria credenciado à XP Investimentos, que atua no suporte consultivo e estratégico a clientes investidores. Para fins de personalização e fidelidade ao ambiente de aplicação, foram utilizados dados anonimizados de carteiras de clientes, relatórios internos da XP e planilhas operacionais da Wura, assegurando a aderência prática da ferramenta às rotinas do escritório.

2.2 Etapas de desenvolvimento

O desenvolvimento da solução foi conduzido com base em uma abordagem iterativa e incremental, estruturada em quatro grandes fases: definição do escopo funcional, engenharia de *prompt*, alimentação com dados especializados e execução de testes práticos. A seguir, descrevem-se as etapas envolvidas em cada uma dessas fases.

Fase 1 – Definição do escopo e objetivos funcionais

A fase inicial consistiu no mapeamento das principais demandas operacionais enfrentadas por assessores de investimentos no contexto da Wura Invest. Com base nessa análise, foram definidos os objetivos funcionais do modelo, que deveria ser capaz de:

- Analisar carteiras reais e simuladas de clientes;
- Avaliar concentração, liquidez, prazos de vencimento e diversificação;
- Identificar riscos e oportunidades;
- Sugerir alocações alinhadas ao perfil e aos objetivos do cliente;
- Apoiar na elaboração de relatórios periódicos e na preparação para reuniões com investidores.

Essas funcionalidades foram pensadas com o intuito de promover ganhos de eficiência, padronização de entregas e geração de valor percebido pelo cliente final do escritório .

Fase 2 – Estruturação e engenharia de prompt

Com base nos objetivos definidos, foi elaborado um *prompt* mestre com aproximadamente 7.991 caracteres, organizado em blocos temáticos. Essa estrutura teve como finalidade orientar o modelo quanto aos pontos descritos abaixo:

- A função do assistente no contexto da Wura Invest;
- As diretrizes de *compliance*, confidencialidade e proteção de dados;
- A forma de interpretar diferentes tipos de entrada (texto, planilhas, PDFs);
- Os critérios de análise por classe de ativo;
- A construção de relatórios com linguagem consultiva, estruturada e objetiva.

A engenharia de *prompt* constituiu o núcleo lógico da solução, assegurando coerência nas respostas e adesão às instruções, mesmo diante de variações nos dados fornecidos pelo usuário. Para atingir respostas satisfatórias, foi necessário diversos testes e iterações com o *prompt* mestre.

Fase 3 – Treinamento com base de dados especializada

Na terceira fase, o modelo foi alimentado com um conjunto de documentos e arquivos utilizados na rotina do escritório, permitindo que as respostas fossem contextualizadas e alinhadas com a realidade operacional e analítica da assessoria. A base de dados incluiu:

- Relatórios mensais da XP Investimentos, com:
 - Cenário macroeconômico atualizado;
 - Perspectivas por classe de ativo;
 - Carteiras recomendadas por perfil de investidor;
 - Posicionamento institucional frente ao mercado.
- Planilhas internas da Wura Invest, com:

- Lista de ativos sugeridos por perfil e finalidade (reserva, renda, diversificação, dolarização, etc.);
 - Avaliação de emissores bancários com recomendações e limites de exposição;
 - Produtos não recomendados ou com restrição interna.
- Exemplos de posição de clientes, incluindo:
 - Planilhas provenientes do relatório Diversificador da XP;
 - Relatórios em formato PDF (XP Performance);
 - Inputs textuais manuais.

Essa base foi integrada ao modelo e ao *prompt* mestre com o objetivo de maximizar a relevância e a precisão das respostas, garantindo a aderência prática e personalizada da ferramenta.

Fase 4 – Execução de testes práticos

A fase final consistiu na execução de uma bateria de testes simulados com o modelo em funcionamento, contemplando diferentes cenários representativos da atuação rotineira de um assessor de investimentos, sempre focando nos objetivos centrais deste trabalho. Entre os casos testados, destacam-se:

- Análise de carteira completa com recomendação de ajustes;
- Sugestão de alocação para novos aportes considerando a posição e cenário econômico atual;
- Preparação de roteiro de reunião com potenciais clientes (*prospects*);
- Geração de relatório de performance e análise mensal dos clientes em formato *Word* ou PDF;
- Diagnóstico de exposição excessiva, liquidez inadequada, presença de produtos restritos e desbalanceamento da carteira.

Cada teste foi documentado com o respectivo *prompt*, a resposta gerada pelo modelo, uma explicação técnica do comportamento observado e uma análise crítica dos resultados. Esses registros compõem a base de validação empírica da pesquisa, sendo descritos em maior detalhe no Capítulo 4.

2.3 Tecnologias e ferramentas

O desenvolvimento da solução exigiu a integração de diferentes tecnologias voltadas à inteligência artificial generativa, manipulação de dados estruturados, leitura automatizada de documentos e suporte à organização e análise de informações. O projeto foi conduzido em ambiente computacional moderno, explorando as capacidades oferecidas pela plataforma ChatGPT e complementado por ferramentas auxiliares essenciais à construção, treinamento e validação do modelo.

A seguir, apresentam-se as principais tecnologias e ferramentas empregadas no projeto:

2.3.1 ChatGPT 4.0

O núcleo da solução foi implementado utilizando a plataforma ChatGPT, da OpenAI, com os recursos de GPT personalizado (*Custom GPTs*) e versão 4.0, disponíveis para usuários do plano Plus. Essa plataforma permitiu a configuração de um assistente com comportamento controlado e reativo, por meio de:

- Inserção de um *prompt* mestre com regras, instruções claras e camadas de análise;
- *Upload* de arquivos complementares, que compõem a base de conhecimento personalizada do modelo;
- Personalização da identidade e estilo de resposta do assistente;
- Utilização de funcionalidades como leitura de arquivos PDF, análise de planilhas e interpretação contextual em linguagem natural.

A escolha por essa plataforma deve-se à sua robustez, flexibilidade e à capacidade de operar com múltiplas fontes de entrada, mantendo coerência e aderência às instruções definidas. Outro fator decisivo foi a facilidade de uso da interface para o usuário final. Além disso, por ser uma das plataformas mais reconhecidas atualmente, torna-se mais simples encontrar materiais de apoio e garantir sua adoção pelos assessores.

2.3.2 Microsoft Excel

As planilhas em Excel foram utilizadas em dois contextos principais:

- Como fonte de dados reais com a posição dos clientes (ex.: relatórios do Diversificador);
- Como base interna de ativos sugeridos, emissores bancários e produtos restritos.

Essas planilhas foram organizadas e padronizadas para facilitar o reconhecimento automático pelo modelo e foram incorporadas diretamente ao ambiente do GPT personalizado ou analisadas em tempo real durante os testes práticos.

2.3.3 Documentos PDF

Foram utilizados documentos oficiais da XP, como o relatório XP Performance (posição consolidada do cliente) e os relatórios mensais de alocação e cenário macroeconômico, que continham:

- Perspectivas por classe de ativo;
- Carteiras recomendadas;
- Posicionamento institucional;
- Tendências e oportunidades observadas no mercado.

Esses documentos permitiram treinar o modelo com dados relevantes, atualizados e, principalmente, personalizados para a realidade da empresa, além de simular entradas reais recebidas por assessores no relacionamento com seus clientes.

2.3.4 Reconhecimento Óptico de Caracteres

Para extração de texto em documentos PDF com conteúdo não selecionável, foi utilizado Reconhecimento Óptico de Caracteres (OCR), função integrada no ChatGPT versão 4.0, garantindo que as informações pudessem ser interpretadas corretamente pelo GPT.

2.3.5 Notion e Microsoft Word

As ferramentas *Notion* e *Microsoft Word* foram utilizadas para:

- Organização dos testes, *prompts* e observações;
- Estruturação de planos de ação e estrutura do projeto;
- Redação e arquivamento de relatórios gerados pelo GPT;
- Registro da evolução do projeto e documentação das interações;
- Acompanhamento do desenvolvimento das etapas.

Capítulo 3

Desenvolvimento e implementação do assistente personalizado

Este capítulo descreve o processo de desenvolvimento, configuração e implementação de um modelo personalizado baseado em inteligência artificial generativa, construído com a tecnologia GPT no ambiente da plataforma ChatGPT, da OpenAI. O objetivo central do projeto foi criar uma ferramenta funcional e eficiente, capaz de apoiar a rotina de assessores de investimentos do escritório Wura Invest, automatizando tarefas analíticas e consultivas que, tradicionalmente, consomem tempo e exigem elevado rigor técnico.

As seções a seguir apresentam a estrutura geral da solução, os fluxos de entrada e saída de dados, a engenharia de *prompt* aplicada, a base de conhecimento adicionada, os critérios de segurança e confidencialidade adotados e, por fim, as limitações enfrentadas ao longo do processo de desenvolvimento.

3.1 Visão Geral do Projeto

O projeto teve como propósito central o desenvolvimento de um assistente virtual especializado para o setor de assessoria de investimentos, operando dentro da plataforma ChatGPT através de um modelo GPT personalizado. Este assistente foi treinado com instruções específicas, arquivos estratégicos e uma lógica de análise financeira construída com

base no contexto operacional da Wura Invest, escritório credenciado à XP Investimentos.

A proposta surgiu a partir da constatação de que assessores enfrentam, com frequência, limitações operacionais e de tempo que dificultam a execução consistente de tarefas como:

- A análise técnica de carteiras com elevado volume e diversidade de ativos;
- A geração de relatórios mensais personalizados para clientes;
- A preparação de reuniões de acompanhamento com clientes e *follow-ups* com potenciais investidores;
- A formulação de sugestões de alocação alinhadas ao perfil, aos objetivos e ao cenário econômico vigente.

Em resposta a esses desafios, foi estruturado um assistente textual inteligente, capaz de:

- Interpretar diferentes tipos de *inputs* (texto livre, planilhas Excel e relatórios em PDF);
- Reconhecer classes de ativos e identificar padrões de alocação;
- Avaliar riscos de concentração, liquidez e desalinhamento com o perfil do cliente;
- Sugerir alocações com base em relatórios da XP e em ativos recomendados pela Wura Invest;
- Gerar saídas em formato de relatório ou orientação consultiva, prontos para utilização imediata no atendimento.

A construção do modelo foi orientada por critérios de precisão analítica, segurança de dados e comunicação profissional, visando à sua utilização prática em contextos reais e rotineiros de relacionamento com o cliente.

O assistente foi apelidado de "InvestAI - Assistente de alocação", conforme pode ser visto na interface do ChatGPT a seguir:



Figura 1: Interface do "InvestAI - Assistente de Alocação" da Wura Invest.

3.2 Estrutura Lógica do Assistente

A estrutura lógica do assistente GPT personalizado foi concebida para simular o raciocínio consultivo e detalhista de um assessor de investimentos, operando por meio da interpretação de diferentes tipos de entrada, análise contextual dos dados, aplicação de regras de conformidade e geração de respostas estruturadas, claras e tecnicamente fundamentadas. A arquitetura funcional está organizada em três componentes principais e interconectados: entrada de dados, processamento interno e geração de resposta.

3.2.1 Entrada de Dados

O sistema foi desenvolvido para compreender diferentes formatos de entrada fornecidos pelo assessor. Entre os principais:

- **Texto livre descritivo:** o assessor pode inserir livremente os ativos do cliente, seus objetivos, valores aportados e demais informações relevantes sobre seu contexto financeiro;

- **Planilhas Excel:** arquivos como o relatório Diversificador XP podem ser enviados com a posição detalhada da carteira. O sistema é capaz de interpretar ativos, valores, percentuais, emissores, liquidez, classe de ativo e entre outros;
- **Relatórios em PDF:** documentos como o XP Performance são lidos e analisados. O sistema identifica composição da carteira, rentabilidade, distribuição por classe de ativo, exposição por emissor, liquidez e entre outros.

Os relatórios em formato PDF são os mais utilizados na rotina dos assessores avaliados para este estudo, portanto o enfoque tanto no *prompt* mestre quanto nos testes será nesse tipo de entrada de dados.

3.2.2 Processamento Interno

A lógica central do sistema atua com base no *prompt* mestre e nos arquivos fornecidos, aplicando múltiplas camadas de interpretação conforme o tipo de entrada e o objetivo do usuário. O processamento segue as seguintes etapas:

- **Classificação de ativos:** identificação do tipo de ativo (renda fixa, variável, multimercado, internacional, previdência, etc.), com base em elementos como nomenclatura, vencimento, indexador e emissor;
- **Análise de concentração:** verificação de alocações excessivas em ativos, emissores ou classes, com base em parâmetros definidos pela Wura Invest (como o limite de 20% por emissor bancário ou limite definido pelo Fundo Garantidor de Crédito (FGC));
- **Avaliação de liquidez:** interpretação de prazos de vencimento na renda fixa e identificação da liquidez de fundos, *Exchange Traded Fund* (ETF) e ativos de renda variável com base em regras pré-definidas e arquivos auxiliares;
- **Compatibilidade com perfil e objetivos:** verificação do perfil de risco (conservador, moderado ou agressivo) e dos objetivos financeiros do cliente (renda, proteção,

reserva, longo prazo e entre outros). Caso essas informações estejam ausentes, o sistema deve solicitar ao assessor a confirmação dos dados;

- **Cruzamento com cenário macroeconômico e alocações sugeridas:** interpretação de relatórios mensais da XP Investimentos para apoiar a recomendação de ativos por perfil e classe;
- **Identificação de riscos e oportunidades:** além da análise técnica, o sistema busca ativos com desempenho insatisfatório ou elevado, produtos restritos ou opções mais adequadas, recomendando ajustes com embasamento técnico.

3.2.3 Geração da Resposta

As respostas (*Outputs*) geradas seguem um padrão consultivo, com linguagem clara, estrutura técnica e adaptação ao tipo de solicitação realizada. Os principais formatos de saída incluem:

- **Diagnóstico da carteira atual:** síntese por classe de ativo, com destaques de pontos críticos e sugestões de melhoria e otimização;
- **Simulação de alocação para novos aportes:** proposta de distribuição por classe de ativo, com justificativas alinhadas ao perfil e ao cenário vigente, considerando sempre sugestões de acordo com visão da XP e da Wura;
- **Resumo de performance mensal:** leitura automatizada do relatório XP Performance, com resumo de rentabilidade, destaques, leitura macroeconômica e mensagem final personalizada ao cliente;
- **Preparação para reuniões:** elaboração de roteiros com perguntas estratégicas, recomendações e pontos de atenção para conversas com clientes ou potenciais investidores.

Todas as saídas foram estruturadas para facilitar a rotina do assessor, proporcionando economia de tempo, ganho de padronização, relacionamento e entrega de valor direto ao cliente final, por meio de uma linguagem acessível e tecnicamente adequada.

3.3 Engenharia de Prompt

A engenharia de *prompt* constituiu o elemento central na construção do modelo personalizado, funcionando como o "sistema operacional" do assistente dentro do ambiente ChatGPT. Por meio de um *prompt* mestre cuidadosamente estruturado, o modelo foi instruído a atuar como um analista consultivo especializado em investimentos, apresentando comportamento consistente, linguagem técnica apropriada em conformidade com as normas operacionais e de *compliance* da Wura Invest e XP.

3.3.1 Fundamentos da Engenharia de Prompt

É um campo de pesquisa relativamente recente que se refere à prática de elaborar e refinar instruções (*prompts*) para direcionar a saída de modelos de LLM [37]. Diferentemente de abordagens tradicionais de adaptação de modelos, a engenharia de *prompt* dispensa re-treinamentos extensivos ou ajustes finos nos parâmetros; em vez disso, explora exclusivamente o conhecimento já embutido no LLM por meio de instruções em linguagem natural [38]. Como resultado, essa técnica emergiu como central para maximizar a utilidade e a precisão desses modelos em diversas tarefas [39], permitindo aos usuários otimizar o desempenho de sistemas avançados, como o GPT 4.0 utilizado neste trabalho, apenas por meio da formulação adequada dos *prompts*.

Diretrizes de boas práticas têm sido propostas para a elaboração eficaz de *prompts*, elas dependem muito do cenário e ambiente em que estão inseridas, mas é possível entender princípios fundamentais e intersecções. Um princípio fundamental é a especificidade: quanto mais claro e direto for o *prompt*, mais precisa e relevante tende a ser a resposta gerada [40]. Além disso, recomenda-se fornecer contexto suficiente, descrever o cenário ou a situação envolvida, explicitar o formato esperado da saída (lista, sumário, análise,

etc.) e indicar o tom desejado da resposta (formal, técnico, didático, entre outros) [41]. Também é eficaz instruir o modelo a assumir um papel específico — como “consultor financeiro experiente” — para guiar a geração de conteúdo conforme o domínio [42]. A engenharia de *prompt* é, portanto, geralmente iterativa: as instruções são refinadas com base nos resultados obtidos, até atingir o comportamento desejado, exatamente como feito na execução deste projeto.

Além de melhorar a qualidade e personalização do conteúdo gerado, a engenharia de *prompt* é uma das principais estratégias de controle de comportamento em LLMs. Essa técnica permite orientar a geração de conteúdo sem modificar a arquitetura ou os pesos do modelo, favorecendo o alinhamento entre as respostas produzidas e os objetivos do sistema [38]. A definição clara de papel, tom e regras de conduta contribui para garantir consistência no estilo e na conduta do modelo ao longo das interações, pontos essenciais em um ambiente corporativo que busca excelência.

Apesar de suas vantagens, a engenharia de *prompt* apresenta limitações importantes. Modelos de linguagem são altamente sensíveis à formulação das instruções, podendo produzir respostas muito distintas a partir de pequenas variações no texto [42]. Também há riscos associados a interpretações ambíguas, não cumprimento completo das instruções e a vulnerabilidades como *prompt injection*, que podem comprometer a segurança e o comportamento previsto do modelo [40]. Além disso, a eficácia de um *prompt* pode variar entre modelos e domínios, o que exige testes e ajustes finos contínuos.

No projeto desenvolvido nesta dissertação, que envolve um assistente virtual para apoio à assessoria de investimentos, os princípios acima foram aplicados na construção do *prompt* mestre. As instruções definem claramente o papel do modelo, o tom esperado, as regras de segurança e os limites de atuação. O sistema foi calibrado para reconhecer entradas em texto livre, planilhas e relatórios, e gerar respostas estruturadas, consultivas e alinhadas com os objetivos dos clientes e padrões da Wura Invest. A engenharia de *prompt*, portanto, foi determinante para assegurar a utilidade prática, a consistência e a conformidade do assistente com os padrões da organização.

3.3.2 Estrutura do *Prompt* Mestre

O *prompt* mestre foi redigido com aproximadamente 7.991 caracteres, dentro do limite de 8.000 imposto pela plataforma, organizado em blocos lógicos e temáticos. Cada seção desempenha uma função específica no raciocínio operacional do assistente, permitindo-lhe adaptar-se a diferentes contextos e tipos de entrada sem comprometer a coerência ou a confiabilidade.

Os principais blocos que compõem o *prompt* são os seguintes:

- **Função e contexto:** define que o modelo atua como assistente dos assessores da Wura Invest, escritório credenciado à XP Investimentos, devendo suas análises basear-se nos dados internos da instituição e no perfil do cliente;
- **Proteção de dados e conformidade:** instrui o modelo a não armazenar, repetir ou expor dados sensíveis, como nomes de clientes, endereços de correio eletrônico, números de identificação fiscal, sem autorização. Também o orienta a substituir tais informações por marcadores como "[Nome do Cliente]" sempre que necessário;
- **Interpretação de entradas:** especifica como o modelo deve processar diferentes formatos de entrada, incluindo descrições manuais, planilhas do Excel e ficheiros PDF, com atenção à estrutura de páginas e à nomenclatura dos ativos;
- **Regras de análise técnica:** orienta o modelo na classificação de ativos por tipo, avaliação de concentração, liquidez, vencimento, indexador e risco, além de verificar se a alocação está alinhada com o perfil do cliente e com o cenário econômico vigente;
- **Sugestão de ativos e substituições:** instrui o modelo a utilizar como base as planilhas de produtos recomendados pela Wura Invest, priorizando ativos conforme o perfil e o propósito (por exemplo: reserva de emergência, proteção cambial, geração de renda passiva), e a alertar quando forem identificados produtos restritos ou inadequados;

- **Estilo de comunicação:** define que as respostas devem ser consultivas, estruturadas e objetivas, com foco em apoiar o assessor, economizar tempo e agregar valor ao cliente final.

Além disso, o *prompt* inclui instruções específicas para lidar com situações recorrentes, os chamados gatilhos inteligentes, entre os quais se destacam:

- Solicitação de agendamento de reunião com um potencial cliente (“roteiro consultivo”);
- Geração de relatório mensal (“relatório mensal de rentabilidade”);
- Pedido de alocação para novo aporte (“análise e sugestão com base na carteira atual e cenário”).

3.3.3 Finalidade da Engenharia de Prompt

A engenharia de *prompt* foi responsável por moldar o comportamento do modelo, não apenas em termos de linguagem, mas, sobretudo, na sua capacidade de interpretar o contexto financeiro, simular análises reais e adaptar suas respostas às diretrizes da organização. Atua, assim, como uma camada de “programação natural”, invisível ao cliente final, mas essencial para garantir a qualidade, a confiabilidade, a consistência e a personalização das respostas.

3.3.4 Disponibilização do Conteúdo

O conteúdo completo do *prompt* mestre encontra-se disponível no Apêndice A, com o objetivo de permitir a consulta detalhada e assegurar a reprodutibilidade do experimento desenvolvido.

3.4 Considerações sobre Limites, Segurança e Desafios

O desenvolvimento de uma solução baseada em inteligência artificial generativa, especialmente no setor financeiro, requer cuidados rigorosos com a segurança da informação, o cumprimento de normas regulatórias e a consideração das limitações técnicas da plataforma adotada. Esta seção apresenta as principais medidas implementadas para garantir a conformidade da aplicação com boas práticas de proteção de dados e governança, bem como os desafios enfrentados ao longo do processo de construção do "InvestAI".

3.4.1 Proteção de Dados Pessoais e Informações Sensíveis

A natureza das atividades de assessoria de investimentos envolve, frequentemente, o manuseio de informações sensíveis dos clientes, como nome completo, número de identificação fiscal (CPF no Brasil), endereço de correio eletrônico, telefone, endereço residencial e dados patrimoniais de clientes. Com o intuito de preservar a privacidade e mitigar riscos de exposição indevida, foram incorporadas ao *prompt* mestre instruções explícitas, tais como:

- Substituição de nomes, identificadores e dados pessoais por marcadores genéricos (por exemplo, "[Nome do Cliente]");
- Proibição de repetição, memorização ou reprodução de qualquer dado sensível após a interação;
- Restrição à exportação, partilha ou armazenamento de informações sensíveis após o encerramento de cada sessão.

Essas diretrizes estão alinhadas à Lei Geral de Proteção de Dados (LGPD), no Brasil. Adicionalmente, as carteiras utilizadas em testes foram devidamente anonimizadas, assegurando que nenhum dado pessoal fosse vazado ou armazenado no assistente durante as fases de desenvolvimento e validação da solução.

3.4.2 Conformidade com Diretrizes da Wura Invest

O modelo foi também configurado para operar em conformidade com os parâmetros internos definidos pela Wura Invest e pelas normas da XP Investimentos, assegurando coerência entre o comportamento do assistente e as orientações vigentes para os assessores. Dentre as principais restrições implementadas, destacam-se:

- Bloqueio à recomendação de produtos restritos ou com recomendação institucional de "evitar";
- Alinhamento entre os ativos sugeridos e o perfil de risco do cliente (por exemplo, evitando indicar criptomoedas ou ações de elevada volatilidade para perfis conservadores);
- Priorização de produtos pertencentes à prateleira Pessoa Jurídica 1 (PJ1) e Pessoa Jurídica 2 (PJ2), conforme permitido pelas estruturas operacionais da empresa;
- Emissão automática de alertas quando forem identificados indícios de desalinhamento entre produto, perfil e objetivo do cliente.

Essas restrições foram descritas logicamente no *prompt* mestre, de modo a garantir que o modelo operasse estritamente dentro dos limites definidos pelas políticas internas da instituição.

3.4.3 Limitações Técnicas da Plataforma

Durante a fase de desenvolvimento, foram identificadas algumas limitações técnicas inerentes à plataforma utilizada, no caso, a versão de *Custom GPTs* disponibilizada aos assinantes da modalidade *Plus* do ChatGPT. As principais restrições observadas foram:

- **Limite de caracteres no *prompt* mestre:** o modelo aceita instruções com até 8.000 caracteres até o momento do desenvolvimento deste trabalho. Esse limite exigiu condensação estratégica do conteúdo e priorização das instruções essenciais

para garantir o funcionamento esperado sem comprometer a lógica da aplicação. Além disso, limitou muito futuros detalhes e implementações;

- **Dificuldade na padronização de nomenclaturas:** nomes de emissores ou ativos com variações de grafia (por exemplo, “Agi (Agibank)” versus “Agibank”) podem não ser interpretados corretamente como equivalentes. Para mitigar esse problema, foram incluídas instruções para aumentar a tolerância e a flexibilidade semântica nesses casos. Em casos como esse, uma agente de IA que tem a capacidade de tomar decisões poderia ter um resultado mais satisfatório;
- **Interpretação parcial de relatórios PDF estruturados:** embora o modelo tenha capacidade de leitura de relatórios como o XP Performance utilizando a tecnologia OCR, elementos com *layout* gráfico complexo (como tabelas incorporadas em imagens ou gráficos de barras) podem não ser extraídos com precisão. Isso afeta, por exemplo, a leitura automatizada da rentabilidade individual por ativo;
- **Pouca integração com fontes externas:** a versão atual do modelo não dispõe de mecanismos nativos para integração direta com softwares ou sistemas externos, o que limita, por exemplo, a atualização automatizada da sua base de conhecimento. Embora tenha sido considerada a possibilidade de conexão com APIs, como as da Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (ANBIMA) ou plataformas especializadas em fundos de investimento, optou-se por não implementar essa funcionalidade nesta fase do projeto. Essa possibilidade, no entanto, é identificada como uma oportunidade relevante para trabalhos futuros, comentada no capítulo 5.

3.4.4 Estratégias de Mitigação

Com o objetivo de contornar as limitações técnicas identificadas anteriormente, foram adotadas algumas estratégias de mitigação:

- Elaboração de planilhas complementares contendo emissores bancários validados por

analistas profissionais e fundos com liquidez previamente conhecida e alimentada pelos assessores da Wura;

- Segmentação do conteúdo relevante da XP em arquivos de texto (.txt), facilitando a atualização frequente e o uso seletivo da informação;
- Instruções específicas no *prompt* para que o modelo solicite confirmação ou esclarecimento ao assessor sempre que detectar dados ausentes, conflitantes ou ambíguos;
- Registro formal dessas limitações como oportunidades de evolução para versões futuras da ferramenta, conforme será detalhado no Capítulo 5.

Essas medidas asseguraram a viabilidade operacional da solução mesmo diante das restrições impostas pela plataforma, além de criarem uma base sólida para a continuidade e melhoria progressiva do sistema em possíveis futuros desenvolvimentos.

3.4.5 Papel do Assessor de investimento

Apesar do elevado grau de automação alcançado, os testes indicam que o modelo deve ser entendido como um assistente consultivo e jamais como substituto da análise adequada de um profissional. O julgamento final permanece sob responsabilidade do assessor, especialmente em situações que envolvam fatores subjetivos, mudanças de cenário ou interpretações estratégicas que extrapolam os dados disponíveis. A atuação humana é indispensável para interpretar nuances emocionais do cliente, avaliar o momento de mercado e tomar decisões com base em sensibilidade e experiência profissional.

Capítulo 4

Testes práticos

Este capítulo apresenta a fase de validação prática do modelo GPT personalizado desenvolvido no âmbito deste projeto. O objetivo é demonstrar, com base em cenários reais embasados na rotina da assessoria de investimentos, a aplicabilidade da solução e sua capacidade de atuar como um assistente consultivo para profissionais da área.

A avaliação foi estruturada e conduzida por meio de quatro testes práticos, simulando demandas recorrentes da rotina de um assessor, tais como:

- Análise completa de carteiras de clientes;
- Diagnóstico de alocação e concentração de ativos;
- Sugestão de alocação para novos aportes;
- Elaboração de relatórios mensais de acompanhamento de performance;
- Preparação de reuniões estratégicas com potenciais clientes.

Os testes selecionados representam tanto situações operacionais do cotidiano como simulações específicas, como detalhes de cada cliente que foram compartilhados ao longo de reuniões, permitindo explorar os limites, o potencial e as restrições do modelo em um ambiente de aplicação realista.

4.1 Metodologia dos Testes

A avaliação prática do assistente GPT foi realizada por meio de testes funcionais controlados, organizados com o objetivo de verificar a capacidade do modelo de interpretar *inputs* reais, processar informações financeiras e gerar respostas estruturadas, técnicas e úteis para a rotina de um assessor de investimentos.

Os testes foram projetados para simular situações reais ou recorrentes no atendimento a clientes, utilizando uma combinação de dados anonimizados, relatórios operacionais e cenários personalizados. O modelo foi testado com diferentes formatos de entrada, variando em grau de complexidade e nível de detalhamento, a fim de validar tanto a amplitude quanto a profundidade da sua aplicação.

4.1.1 Estrutura dos Testes

Cada teste foi composto por cinco etapas principais:

1. **Definição do objetivo:** descrição da situação prática simulada (ex.: alocação de aporte, análise de performance, preparação de reunião);
2. **Elaboração e envio do *prompt*:** formulação de uma instrução clara e específica ao modelo, considerando boas práticas de engenharia de *prompt* comentadas de forma detalhada no Capítulo 3;
3. **Fornecimento de *input* realista:** envio de documento e ou informação complementar ao assistente (ex.: planilha de posição, relatório XP Performance em PDF, descrição textual de ativos);
4. **Geração de resposta pelo assistente:** execução da análise e resposta automática do modelo com base nas instruções e dados fornecidos;
5. **Avaliação e análise crítica do resultado:** observação dos pontos fortes, limitações e oportunidades de melhoria frente ao objetivo proposto.

4.1.2 Tipos de *Input* Utilizados

Foram utilizados três formatos principais de entrada de dados:

Tipo de Input	Descrição
Texto livre	Descrição manual dos ativos, perfil do cliente e objetivos financeiros
Planilha Excel	Relatórios de posição gerados pela XP, especialmente o modelo <i>Diversificador</i>
PDF	Relatório XP Performance, contendo alocação, rentabilidade e exposição da carteira

Tabela 2: Tipos de *input* utilizados nos testes práticos do modelo GPT.

O modelo foi treinado para reconhecer e adaptar-se automaticamente ao tipo de *input*, extraíndo as informações relevantes de cada um. Porém, como o uso do relatório XP Performance é amplamente utilizado entre os assessores da Wura, e este trabalho teve enfoque no auxílio da rotina dos mesmos, foi considerado como o modelo padrão de envio nos testes, embora o assistente esteja capacitado para trabalhar com os outros formatos supracitados.

4.1.3 Critérios de Avaliação

As respostas geradas pelo GPT foram analisadas com base nos seguintes critérios:

Critério	Descrição
Clareza e estrutura	Se a resposta foi objetiva, bem organizada e de fácil entendimento
Aderência ao objetivo	Se a análise abordou corretamente o que foi solicitado no <i>prompt</i>
Precisão técnica	Se os dados extraídos e interpretados estão corretos e bem classificados
Alinhamento com perfil	Se as recomendações respeitaram o perfil e os objetivos do cliente
Aproveitamento dos dados	Se o modelo utilizou adequadamente as bases da XP e Wura Invest
Deteção de riscos	Se alertou para concentrações, desalinhamentos ou produtos inadequados, por exemplo

Tabela 3: Critérios utilizados para avaliação das respostas geradas pelo assistente GPT.

Essa metodologia garante consistência entre os diferentes testes realizados e oferece retornos claros para a avaliação do desempenho do modelo, conforme descrito nas seções seguintes.

4.2 Teste 1: Análise Geral de Carteira com Novo Aporte

4.2.1 Objetivo do teste

Avaliar se o assistente seria capaz de interpretar uma carteira real de investimentos (anonimizada), classificar corretamente os ativos por classe, avaliar exposição, risco de concentração, liquidez e qualidade da diversificação. Além disso, o teste exigiu a sugestão de alocação para um novo aporte de R\$ 200.000,00 e a construção de uma reserva separada para geração de renda mensal de R\$ 10.000,00 durante um período de 12 meses.

4.2.2 Contexto simulado

O cliente possui perfil agressivo, reside atualmente nos Estados Unidos e planeja retornar ao Brasil dentro de um ano. Durante esse período de transição, ficará sem renda formal, e deseja utilizar parte da carteira para suprir R\$ 10.000,00 mensais com os investimentos. O restante da carteira seguirá com o objetivo de acumulação patrimonial no longo prazo. A carteira original é concentrada em renda variável e o aporte adicional de R\$ 200.000,00 deveria ser alocado de forma estratégica com base nas recomendações da Wura Invest, nas carteiras sugeridas da XP e no cenário macroeconômico atual.

4.2.3 Prompt utilizado

”Segue abaixo uma carteira real de investimentos (não identificada por segurança). Preciso que você faça uma análise completa dessa carteira com base nas informações disponíveis nos seus arquivos, classificando os ativos por classe

e avaliando exposição por tipo, risco de concentração, liquidez e qualidade da diversificação.

Informações adicionais obrigatórias:

- Perfil do cliente: Agressivo
- Objetivo: Acumulação de patrimônio no longo prazo. No curto/médio prazo (1 ano), o cliente pretende retornar ao Brasil e ficará sem renda formal por alguns meses. Deseja gerar uma renda de aproximadamente R\$ 10.000,00 por mês com os investimentos nesse período de transição.
- Aporte: O cliente fará um novo aporte de R\$ 200.000,00. Desejo sugestões de alocação considerando a carteira atual, o cenário econômico e os ativos recomendados pela Wura Invest e XP.

A análise deve considerar os objetivos do cliente, seu perfil, o cenário atual e possíveis ajustes necessários na alocação atual. Se possível, proponha também uma estrutura separada para compor a reserva de renda temporária.”

4.2.4 Explicação técnica do *prompt*

O *prompt* foi construído com base nos princípios descritos na engenharia de *prompt* (Capítulo 3.3), ativando múltiplas capacidades do modelo:

- **Prompt contextualizado com múltiplos objetivos:** envolve diagnóstico da carteira atual, alocação de novo aporte e construção de reserva tática;
- **Ativação de lógica consultiva estruturada:** o modelo deve simular o comportamento de um analista profissional, com análise técnica e comunicação clara;
- **Integração com base de conhecimento externa:** o *prompt* exige uso dos arquivos treinados (relatórios XP, ativos da Wura, cenário macroeconômico);

- **Cenário híbrido:** exige balanceamento entre liquidez (renda no curto prazo) e crescimento (perfil agressivo de longo prazo);
- **Planejamento com *cash flow*:** espera-se uma proposta de alocação que viabilize fluxo mensal de R\$ 10.000,00.

4.2.5 Resumo da resposta do GPT

- **Interpretação do cenário e avaliação por tipo de ativo:** Incorporou corretamente o contexto da necessidade de renda mensal no curto/médio prazo e do perfil agressivo de longo prazo;

Avaliação por Tipo de Ativo

Renda Variável Brasil (65,36%)

- Ações concentradas em setores de energia, bancos e commodities.
- Excesso de concentração em alguns papéis: BBAS3 (11,2%), VALE3 (7,53%), UNIP6 (7,35%), PETR4 (8,04%).
- Rentabilidade positiva no ano, mas com alta volatilidade (volatilidade 12M: 9,94%).

Fundos Listados (FIs) (34,61%)

- FIs como VGIA11, TRXF11, TGAR11, RECR11 ocupam quase 35% da carteira.
- Alguns FIs como TGAR11 apresentam performance ruim no período (-4,51% no mês, -7,94% em 24M).

Riscos Identificados

- **Concentração excessiva em renda variável (mais de 65%):** alto risco em cenário de transição e necessidade de renda.
- **Baixa liquidez imediata:** ausência de pós-fixados compromete a capacidade de resgates ágeis.
- **Excesso de exposição em FIs com risco de mercado e volatilidade elevada.**
- **Nenhuma diversificação internacional ou exposição a multimercados,** o que aumenta a correlação da carteira com a economia local.

Figura 2: Interpretação e análise da carteira no teste 1.

- **Diagnóstico da carteira atual:** Composição altamente concentrada: 65% em ações e 35% em Fundo de Investimento Imobiliário (FII). Ausência de renda fixa,

multimercados ou ativos internacionais. Considerada inadequada mesmo para perfil agressivo, dada a necessidade de liquidez no curto prazo;

 **Composição Atual da Carteira**

Patrimônio total: R\$ 846.903,29

Classe de Ativo	Valor (R\$)	% Alocação
Renda Variável Brasil	R\$ 549.306,20	65,36%
Fundos Listados (FIs)	R\$ 290.884,40	34,61%
Caixa	R\$ 270,75	0,03%

Figura 3: Composição da carteira no teste 1.

Observações:

- A carteira não possui nenhuma exposição atual em Renda Fixa (pós-fixada, IPCA+ ou prefixada), multimercados ou ativos internacionais.
- A alocação é extremamente concentrada em renda variável e fundos imobiliários, o que é coerente com o perfil agressivo, mas desalinhado com a necessidade de geração de renda e liquidez no curto prazo.

Figura 4: Diagnóstico da carteira no teste 1.

- **Proposta de renda:** Sugeri R\$ 60.000,00 divididos entre renda fixa pós-fixada e inflação de vencimento curto para garantir a renda de R\$ 10.000,00/mês durante seis meses. A proposta foi coerente em termos de risco e produtos, mas pecou no valor inferior ao necessário de 12 meses solicitado.

 **Necessidade de Ajuste para Renda Mensal (R\$ 10.000/mês por 6 meses)**

Recomendo destinar **R\$ 60.000** a uma "reserva de renda" com alta liquidez e baixo risco. Sugestão:

- **R\$ 40.000** em pós-fixados com liquidez diária (ex: fundos DI XP ou CDBs com L+0)
- **R\$ 20.000** em IPCA+ curto (venc. 2026 ou anterior) para proteção do poder de compra

Figura 5: Proposta de renda no teste 1.

- **Sugestão de alocação:** Propôs alocação segmentada com percentuais compatíveis com a estratégia da XP e Wura para perfil agressivo, com base nos arquivos do treinamento;

 **Sugestão de Alocação para o Aporte de R\$ 200.000**

Com base nas recomendações de maio/2025 para perfil agressivo:

Classe	% Sugerido	Valor (R\$)	Observação
Pós-Fixado	15%	30.000	Fundo DI XP, CDB L+0
IPCA+	32,5%	65.000	Tesouro IPCA+, CRIs indexados curtos
Prefixado	2,5%	5.000	Títulos com vencimento até 2027
Multimercado	12,5%	25.000	Fundos XP Macro, Absolute Vertex
Renda Variável Brasil	15%	30.000	Setores defensivos e dividendos
Fundos Listados (Fils)	5%	10.000	Fils de crédito (VGIA11, RZAK11)
Alternativos	7%	14.000	Fundos quantitativos ou descorrelacionados
Internacional (RF e RV)	10%	20.000	ETF IVVB11, fundos globais hedgeados

Figura 6: Sugestão de alocação para novo aporte no teste 1.

- **Produtos sugeridos com base nas bases da Wura Invest e XP:** Por exemplo, Selection RF Light, CPTS11, AZ Quest e PIMCO Income;
- **Alocação final da carteira pós-aporte (R\$ 1.046.903,29):** Rebalanceamento para reduzir o risco excessivo em renda variável, aumentar liquidez e proteção no curto prazo;

5. Alocação Final Consolidada da Carteira (R\$ 1.046.903,29)

Classe de Ativo	Valor (R\$)	% da Carteira Final
Renda Variável Brasil	549.306,20	52,45%
Fundos Listados (FIIs)	290.884,40	27,78%
Renda Fixa Pós-Fixada	50.000,00	4,78%
FIIs Renda (novos)	50.000,00	4,78%
Renda Fixa IPCA+	40.000,00	3,82%
Multimercado	30.000,00	2,87%
Internacional	20.000,00	1,91%
Caixa	10.000,00	0,96%

Figura 7: Alocação final no teste 1.

- **Recomendações adicionais:** Sugestões pertinentes e muito interessantes, conexas com a realidade macroeconômica e particularidades dos ativos;

6. Modificações Sugeridas na Carteira Atual

- Substituição do fundo TGAR11 devido à baixa performance e alto risco, por FIIs mais previsíveis (ex: KNSC11).
- Inserção de renda fixa pós-fixada e IPCA+ para geração de renda e proteção.
- Inclusão de fundos multimercado e internacionais para diversificação e proteção cambial.
- Rebalanceamento com parte do novo aporte para reduzir concentração excessiva em renda variável brasileira.

Figura 8: Recomendações adicionais no teste 1.

- **Recomendações finais práticas:** Incluiu instruções claras para realocação, criação de reserva, diversificação e substituição de ativos problemáticos alinhados com cenários macroeconômicos vigentes e dados do treinamento.

✓ **Recomendações Finais**

1. **Realocar parte da renda variável atual** para classes mais líquidas e defensivas.
2. **Criar imediatamente uma reserva de liquidez** para garantir a renda mensal desejada durante o retorno ao Brasil.
3. **Diversificar com ativos internacionais e multimercados**, buscando menor correlação e proteção cambial tática.
4. **Revisar os FIIs com pior performance e alto risco**, substituindo por fundos mais resilientes.

Se quiser, posso montar uma sugestão de carteira detalhada com tickers e produtos específicos da base da Wura. Deseja seguir com isso?

Figura 9: Recomendações finais no teste 1.

4.2.6 Análise crítica do resultado

Critério	Justificativa
Clareza e estrutura	Excelente. Resposta dividida em seções lógicas, bem organizadas e com linguagem consultiva, estruturada como uma proposta comercial real.
Aderência ao objetivo	Parcialmente aderente. Atendeu às solicitações do <i>prompt</i> , analisando a carteira atual, propondo ajustes, compondo a reserva de renda e sugerindo alocação para o novo aporte. Apesar de ser um erro mínimo, pecou no planejamento dos 12 meses, planejando somente 6.
Precisão técnica	Classificação correta dos ativos e leitura precisa do relatório XP Performance, com destaque para liquidez, indexadores e vencimentos.
Alinhamento com perfil	A alocação proposta respeitou o perfil agressivo do cliente, com balanceamento entre crescimento e proteção de curto prazo, considerando o cenário de transição pessoal.
Aproveitamento dos dados	O modelo utilizou adequadamente as bases da XP e da Wura para justificar a alocação e recomendar substituições.
Detecção de riscos	Identificou corretamente concentração em renda variável, ausência de liquidez imediata e riscos de exposição excessiva em FIIs.

Tabela 4: Análise crítica do resultado conforme critérios estabelecidos (Teste 1)

4.2.7 Conclusão do teste

O assistente GPT demonstrou alto nível de entendimento do contexto do cliente e das variáveis técnicas envolvidas em uma análise de carteira. Foi capaz de cruzar as informações de cenário, perfil e objetivos para gerar uma proposta bem fundamentada, segmentada e aplicável (disponível na íntegra no Apêndice B). Também apresentou excelente capacidade de comunicação consultiva, com recomendações claras, organizadas e tecnicamente justificadas.

Este teste validou que o assistente pode ser utilizado com certo grau de confiabilidade e eficiência para apoiar assessores em análises completas de carteira, mesmo em cenários que exigem sensibilidade ao ciclo de vida do cliente e ajustes táticos à sua realidade financeira.

4.3 Teste 2 - Análise de performance e geração de relatório

4.3.1 Objetivo do teste

Este teste teve como objetivo avaliar a capacidade do assistente GPT da Wura Invest de gerar um relatório de acompanhamento mensal personalizado para o cliente, a partir da leitura do relatório XP Performance. O foco foi verificar se o agente consegue cruzar informações da carteira com o cenário macroeconômico atual, posicionamento da XP, perspectivas por classe de ativo, e identificar ativos com maior e pior performance, além de sugerir ajustes ou oportunidades relevantes.

4.3.2 Contexto simulado

A simulação reproduz um cenário real em que o assessor deseja enviar um relatório periódico no início do mês, com contexto econômico, análise de rentabilidade e possíveis movimentos táticos, reforçando o contato consultivo com o cliente.

4.3.3 *Prompt* utilizado

”Vou enviar o relatório XP Performance de um cliente. Por favor, gere um relatório de acompanhamento para ser enviado a ele como fechamento de mês ou início do próximo.

A estrutura deve conter:

- Um resumo objetivo do cenário macroeconômico atual e sua relação com os investimentos.
- A perspectiva da XP por classe de ativo (Renda Fixa Pós, Índice Nacional de Preços ao Consumidor Amplo (IPCA), Multimercado, RV Brasil, Internacional, FIIs e etc).
- Análise da performance da carteira do cliente, com destaque para os 3 ativos com melhor e pior rentabilidade.
- Um resumo final destacando oportunidades, riscos ou pontos de atenção, caso existam.

A análise deve ser clara, profissional e personalizada com base na carteira enviada. O objetivo é gerar um material de valor para envio ao cliente de forma recorrente e eficiente.”

4.3.4 Explicação técnica do *prompt*

Este *prompt* representa um caso típico de *prompt* analítico voltado à geração de conteúdo consultivo com múltiplas camadas de interpretação, ativando diferentes mecanismos internos do modelo de linguagem. Destaca-se por reunir estrutura, contexto e direcionamento claros, conforme descrito a seguir:

- **Engenharia de *prompt* direcionada por objetivos compostos:** o *prompt* é construído com instruções segmentadas e orientadas a metas específicas (como análise de cenário macroeconômico, performance da carteira, destaques positivos e

negativos e recomendações). Essa estrutura em tópicos reduz a ambiguidade e ativa a habilidade do modelo em gerar uma narrativa consultiva adaptada à realidade do cliente.

- **Ativação de memória contextual de curto prazo:** solicita a consulta dos arquivos previamente fornecidos (visão macro, carteiras recomendadas e posicionamento da XP), o modelo é capaz de cruzar essas informações com a carteira do cliente, conectando:
 - As alocações ideais por perfil;
 - A visão por classe de ativo vigente;
 - O contexto econômico atual.
- **Geração de linguagem adaptada ao público final:** diferentemente de testes voltados exclusivamente ao assessor, este *prompt* exige que a resposta seja formatada para o cliente final, com clareza, tom acessível e foco na experiência. Isso demanda que o modelo adapte vocabulário e estrutura, evitando jargões excessivos, mas mantendo o rigor técnico.
- **Priorização baseada em impacto e relevância:** ao destacar os três ativos com melhor e pior desempenho, o modelo é instigado a:
 - Avaliar a performance relativa;
 - Considerar o peso de cada ativo na carteira;
 - Destacar apenas os elementos mais significativos para a compreensão do cliente.

4.3.5 Resumo das respostas do GPT

O "InvestAI" gerou um relatório completo e formatado para envio direto ao cliente, que está disponível no Apêndice C. Das respostas, é possível destacar:

- **Diagnóstico de performance bem fundamentado:** A rentabilidade mensal foi apresentada com contextualização frente aos *benchmarks* (CDI, IPCA e dólar). O assistente ainda indicou o ganho financeiro do mês e mencionou o impacto de uma retirada, demonstrando leitura financeira completa.

Performance da Carteira

- Rentabilidade no mês: +1,38%
- Rentabilidade no ano: +7,04%
- Rentabilidade 24 meses: +12,53%
- Patrimônio total: R\$ 1.120.597,98

A performance do portfólio foi positiva e bem superior ao CDI do mês (+0,26%), superando também o IPCA (+0,09%) e o Dólar (+0,45%). A carteira apresentou ganho financeiro de R\$ 15.128,86 no mês, mesmo com uma retirada de R\$ 4.000,00 no período.

Figura 10: Análise de Performance no teste 2.

- **Ranking dos ativos:** Seleção correta dos ativos levando em consideração o que foi solicitado, colocando-a de maneira intuitiva e objetiva, facilitando a leitura para o cliente.

Top 3 Ativos com Melhor Performance (Abril)

1. ALPA4 - +9,59%
2. EUCA4 - +9,39%
3. BRAV3 - +13,25%

Top 3 Ativos com Pior Performance (Abril)

1. AURA32 - -7,84%
2. AZUL4 - -3,05%
3. AMER3 - -2,83%

Figura 11: *Ranking* de ativos no teste 2.

- **Cenário macroeconômico bem interpretado:** A visão por classe de ativo foi baseada na perspectiva da XP, com comentários estratégicos adequados e atualizados (ex.: cautela em renda variável global, foco em pós-fixados, atenção em FIIs).

Cenário Macroeconômico

Abril foi um mês de forte volatilidade nos mercados internacionais, provocada por novas tarifas comerciais impostas pelos Estados Unidos contra diversos países. O chamado "Liberation Day", liderado pelo presidente Donald Trump, desencadeou uma guerra comercial, principalmente com a China, elevando incertezas e pressionando as expectativas de inflação.

Nos EUA, o FED manteve os juros entre 4,25% e 4,50% diante da alta nos custos de produção e desaquecimento do consumo. A perspectiva é de menor crescimento global com maior inflação.

No Brasil, apesar das turbulências externas, o mercado teve desempenho favorável. A alocação em ativos locais foi beneficiada pela menor exposição às tarifas e pela busca por alternativas fora dos EUA. O Ibovespa subiu 6,1% em abril e o IPCA acumulado em 12 meses subiu para 5,06%, com a Selic a 14,25%.

Figura 12: Cenário econômico no teste 2.

- **Identificação correta de oportunidades e riscos:** Foram destacados ativos com performance fraca, riscos de volatilidade internacional e recomendações táticas de reforço em ações com bom desempenho e IPCA.

Oportunidades e Riscos

Oportunidades

- Fortes resultados em ações específicas (ALPA4, EUCA4) sugerem manutenção ou ampliação tática.
- Renda fixa IPCA+ segue relevante para proteção real da carteira.

Riscos

- Volatilidade internacional elevada com tensões comerciais.
- Alguns ativos com performance negativa acentuada (ex: AURA32, AZUL4).

Figura 13: Oportunidades e Riscos no teste 2.

- **Identificação da perspectiva por classe de ativo:** Identificou corretamente, conforme treinamento no *Prompt* mestre e materiais da XP, a perspectiva por classe de ativos.

 **Perspectiva por Classe de Ativo**

- Pós-Fixado: segue atrativo com Selic elevada, foco em liquidez e estabilidade.
- Inflação: proteção relevante contra choques, com bom retorno.
- Pré-Fixado: manter exposição moderada, ajustando conforme expectativa de inflação.
- Renda Variável Brasil: oportunidade em setores defensivos e empresas com bons dividendos.
- Renda Variável Global: manter exposição moderada; momento de cautela.
- Fundos Listados (FIIs): revisão de ativos recomendada; desempenho heterogêneo.

Figura 14: Perspectivas no teste 2.

4.3.6 Análise crítica do resultado

Critério	Justificativa
Clareza e estrutura	O texto foi gerado com estrutura lógica, linguagem objetiva e tom adequado ao cliente final. A divisão em seções facilitou a leitura e compreensão imediata. E o documento gerado teve formatação simples, porém limpa e satisfatória.
Aderência ao objetivo	O assistente atendeu todos os itens solicitados no <i>prompt</i> : resumo macroeconômico, análise da carteira, <i>top/bottom</i> ativos, visão por classe e conclusão para o cliente.
Precisão técnica	Os dados de rentabilidade, alocação e ativos foram corretamente interpretados a partir do relatório XP Performance.
Alinhamento com perfil	A conclusão do relatório reforçou que a carteira estava bem posicionada diante do cenário e perfil do cliente, sem propor mudanças desnecessárias.
Aproveitamento dos dados	As informações do relatório XP e os arquivos da XP e Wura foram utilizados para contextualizar decisões e reforçar interpretações sobre risco e desempenho.
Deteção de riscos	Foram mencionadas quedas em ativos como AMER3, mas faltou uma análise mais profunda sobre peso na carteira ou necessidade de ajuste. O risco foi citado, mas raso.

Tabela 5: Análise crítica do resultado conforme critérios estabelecidos (Teste 2)

4.3.7 Conclusão do teste

O assistente GPT demonstrou capacidade para gerar relatórios mensais de performance com clareza, estrutura lógica e linguagem consultiva adequada ao cliente final. A resposta apresentou leitura precisa dos dados do relatório XP Performance, interpretação alinhada ao cenário macroeconômico e entrega de uma narrativa coesa e personalizada. Essa funcionalidade representa uma solução viável para automatizar a comunicação recorrente com clientes, economizando tempo do assessor e reforçando a entrega de valor com relatórios interpretativos, confiáveis e visualmente organizados.

Como pontos de melhoria, recomenda-se enriquecer a seção de oportunidades com sugestões mais específicas de ações táticas (como ajustes em ativos com pior desempenho), detalhar melhor a coerência da alocação com o perfil de risco do cliente e ampliar a profundidade da análise macro por classe de ativo, com base na linguagem técnica dos relatórios mensais da XP. Esses refinamentos podem elevar ainda mais o nível de personalização e utilidade prática dos relatórios gerados.

4.4 Teste 3 - Apoio do Assistente na Preparação de Reuniões com *Prospects*

4.4.1 Objetivo do teste

Este teste teve como objetivo avaliar se o assistente GPT personalizado da Wura Invest é capaz de apoiar um assessor de investimentos em estágio inicial de relacionamento com um potencial cliente (*prospect*), oferecendo suporte na preparação da primeira reunião de abordagem comercial. A ênfase da avaliação recai sobre a capacidade do modelo em gerar um roteiro estruturado para a reunião, contendo perguntas estratégicas e uma sequência lógica de tópicos, de forma a possibilitar que mesmo profissionais com menos experiência consigam conduzir encontros de maneira eficaz, profissional e personalizada.

Este cenário simula uma aplicação prática de elevado valor agregado, sobretudo para

assessores em início de carreira que buscam maior consistência, segurança e confiança na rotina de prospecção. Uma vez validado, o recurso pode ser utilizado como ferramenta complementar de capacitação e apoio contínuo à atuação comercial, promovendo padronização de boas práticas na abordagem de novos clientes, garantindo que o assessor aborde todos os temas mais relevantes e extraia informações de forma eficiente.

4.4.2 Contexto simulado

O cenário simulado representa um assessor da Wura Invest prestes a realizar a primeira reunião com um potencial cliente, ainda sem vínculo formal estabelecido. O objetivo do assessor é preparar-se adequadamente para conduzir a conversa com segurança, clareza e direcionamento comercial, abordando de forma estruturada temas essenciais como perfil financeiro, objetivos, organização patrimonial e os diferenciais estratégicos da Wura Invest.

O assistente GPT foi acionado com o propósito de gerar um roteiro consultivo, didático e funcional, que pudesse servir como base prática durante a reunião, especialmente para assessores em início de carreira ou em contextos de abordagens mais desafiadoras.

4.4.3 Prompt utilizado

”O objetivo desse teste é entender como o agente pode auxiliar um assessor de investimentos em estágios iniciais da profissão com uma das tarefas mais desafiadoras: fazer a primeira reunião com um *prospect*, um possível cliente.

A ideia é gerar um roteiro estruturado de como o assessor pode guiar a reunião, sem deixar de abordar pontos importantes como:

- Criação de vínculo e abertura;
- Diagnóstico do perfil financeiro;
- Objetivos e horizonte;
- Necessidades e expectativas;
- Apresentação da Wura e dos seus diferenciais;

- Oferta de serviços;
- Identificação de oportunidades;
- Encerramento e proposta de continuidade.

O roteiro deve ser consultivo, claro, profissional, bem dividido por etapas e escrito em um tom que possa ser usado na prática. Pode também incluir sugestões de frases e perguntas abertas.”

4.4.4 Explicação técnica do *prompt*

O *prompt* utilizado neste teste caracteriza-se como um exemplo clássico de *prompt* de geração estruturada orientada por função. Ele ativa no modelo não apenas a produção de conteúdo textual, mas a organização de um processo lógico aplicado a um contexto humano real, neste caso, a preparação de uma reunião comercial com um *prospect*.

- **Aplicação prática da engenharia de *prompt*:** O *prompt* utiliza linguagem clara, direcional e segmentada por objetivos específicos, o que facilita o alinhamento do modelo com o resultado esperado. A listagem explícita dos blocos da reunião (como diagnóstico, apresentação, encerramento) atua como um mapa de estrutura narrativa, ativando a capacidade do GPT de gerar um roteiro ordenado, coeso e funcional.
- **Requisição de linguagem adaptada à prática:** A resposta gerada deve apresentar linguagem natural e aplicável no cotidiano, evitando jargões técnicos excessivos ou formulações genéricas. Isso exige que o modelo contextualize conteúdos técnicos para fins comerciais, ajustando tom, vocabulário e fluidez para conversas humanas reais.
- **Capacidade de simulação consultiva:** O modelo é estimulado a simular o raciocínio e o comportamento de um assessor experiente, propondo perguntas estratégicas, abordagem por etapas e lógica de negociação. Tal abordagem vai além da

simples geração textual, exigindo aplicação de raciocínio situacional e domínio do papel social simulado (de assessor para cliente).

4.4.5 Resposta gerada pelo GPT

O assistente entregou um roteiro de reunião completo, estruturado em seis blocos principais, cada um com objetivo claro, orientações práticas e sugestões de frases ou perguntas estratégicas. A estrutura final do conteúdo entregue foi a seguinte:

1. Abertura (5 minutos)

Objetivo: Criar vínculo e estabelecer um clima de confiança.

Sugestão de abordagem informal e acolhedora.

Exemplo de frase: “Antes de falarmos de investimentos, quero entender melhor você e seus objetivos...”

2. Conhecimento do Perfil (15–20 minutos)

Objetivo: Diagnosticar a situação financeira, os objetivos e o comportamento de risco.

Subdividido em: situação atual, objetivos, perfil de risco e horizonte.

Sugestões de perguntas abertas como:

“Como está organizada sua vida financeira hoje?”

“Já investiu em ações, fundos, renda fixa, internacional?”

“Existe alguma necessidade de liquidez no curto prazo?”

3. Apresentação da Wura Invest (5–10 minutos)

Objetivo: Posicionar a empresa e seus diferenciais.

Menciona credenciamento à XP, foco consultivo, independência na recomendação, base de produtos PJ1 e PJ2 e rigor com compliance.

4. **Explicação dos Serviços (5–10 minutos)**

Objetivo: Mostrar valor agregado da assessoria.

Planejamento financeiro, revisão de carteira, acesso a produtos exclusivos, apoio em sucessão e previdência.

5. **Identificação de Oportunidades (10 minutos)**

Objetivo: Entender frustrações e mapear possibilidades.

Exemplo de perguntas:

“Você já teve alguém te ajudando de forma ativa nos seus investimentos?”

“Qual foi a última vez que revisou sua carteira?”

6. **Encerramento e Próximos Passos (5 minutos)**

Objetivo: Garantir continuidade.

Sugestão de nova reunião com proposta baseada no diagnóstico atual.

Comentário geral

A resposta gerada foi altamente satisfatória, com estrutura aplicável, vocabulário consultivo e equilíbrio entre linguagem acessível e posicionamento profissional. A proposta serve como guia pronto para uso, especialmente para assessores iniciantes.

O roteiro não considerou adaptações mais profundas conforme o perfil ou comportamento do *prospect* (ex.: Agressivos, conservadores, investidores experientes). Isso poderia ser sugerido como desdobramento futuro.

4.4.6 Análise crítica do resultado

Critério	Justificativa
Clareza e estrutura da resposta	A resposta apresenta excelente clareza e organização. O roteiro foi dividido em seis etapas com títulos e objetivos definidos. Cada seção contém subtópicos claros, instruções diretas e exemplos práticos, o que facilita a leitura e aplicação prática pelo assessor.
Aderência ao objetivo proposto	O GPT cumpriu com precisão o objetivo do teste, entregando um roteiro estruturado, com início, meio e fim, voltado à condução de uma reunião inicial com o <i>prospect</i> , incluindo diagnóstico, apresentação e engajamento.
Precisão técnica das sugestões	As perguntas propostas e os conteúdos abordados são tecnicamente adequados. Estão alinhados com boas práticas da assessoria consultiva e não há imprecisões ou riscos de exposição inadequada.
Alinhamento com perfil e contexto simulado	Mesmo sem perfil de cliente definido, a resposta manteve tom neutro e adaptável, respeitando o cenário de primeira reunião e evitando abordagens comerciais precipitadas.
Aproveitamento dos dados disponíveis no <i>prompt</i>	O modelo utilizou corretamente as informações fornecidas no <i>prompt</i> , reconhecendo os objetivos da reunião e incorporando os elementos esperados de forma completa.
Deteção e sinalização de riscos	Embora o foco não fosse análise técnica de carteira, o modelo demonstrou preocupação com a comunicação clara e construção de confiança, mitigando riscos indiretos de objeções ou desalinhamentos.

Tabela 6: Análise crítica do resultado conforme critérios estabelecidos (Teste 3)

4.4.7 Conclusão do teste

O desempenho do GPT no Teste 3 foi satisfatório. O assistente respondeu com clareza, estrutura e aplicabilidade prática, entregando um roteiro profissional e realista voltado a assessores em fase inicial de atuação. As instruções foram completas, bem segmentadas e demonstraram forte aderência à função simulada de preparação para uma primeira reunião com um potencial cliente.

Como ponto de melhoria, destaca-se a oportunidade de incluir variações na condução da reunião conforme perfis distintos de clientes (por exemplo, ênfase em proteção para perfis conservadores ou foco em oportunidades para perfis mais agressivos), mesmo que apresentadas como sugestões opcionais ao final do roteiro. Além disso, há uma possibilidade de evolução natural do roteiro através de *feedbacks* após o uso do roteiro, adequando-se a realidades e casos diferentes.

4.5 Teste 4: Análise Técnica Avançada de Carteira

4.5.1 Objetivo do teste

Este teste tem como objetivo avaliar a capacidade do assistente GPT de realizar uma análise técnica aprofundada de uma carteira de investimentos real (não identificada por segurança e *compliance*), com foco detalhado nos ativos de Renda Fixa e Renda Variável. A intenção é simular o uso da ferramenta por um assessor de investimentos experiente, que busca otimizar tempo e agregar valor em análises complexas, especialmente no que se refere a:

- Classificação precisa dos ativos por tipo e subclasse;
- Avaliação de concentração por emissor de renda fixa e por classe de ativo;
- Avaliação de risco por emissor de renda fixa;
- Análise de liquidez, vencimentos e indexadores;
- Verificação de alinhamento com o perfil e os objetivos do cliente;
- Cruzamento com o cenário macroeconômico vigente e com as recomendações da XP e da Wura Invest.

Este é o teste mais técnico da bateria e funciona como validação da maturidade técnica do GPT desenvolvido no contexto desta dissertação. Ele permite avaliar, de forma integrada, todas as funcionalidades relevantes e camadas de análise implementadas no sistema, com exceção à função praticada no teste 3 (Capítulo 4.4).

4.5.2 Contexto simulado

Neste cenário, o assessor deseja realizar uma análise minuciosa de uma carteira real extraída do relatório XP Performance, com o objetivo de obter subsídios técnicos para uma eventual readequação, sendo útil para uma avaliação periódica da carteira do cliente.

A carteira foi enviada em formato PDF e apresenta a posição consolidada do cliente, incluindo ativos de renda fixa, fundos de investimento, ações, FIIs e ativos internacionais.

Inicialmente, o perfil de risco e os objetivos do cliente não foram informados. O assistente foi, portanto, instruído a realizar a análise com base exclusivamente nos dados contidos no relatório, classificando os ativos e identificando eventuais riscos ou oportunidades de melhoria. Após a entrega da primeira resposta, o assessor forneceu informações complementares:

- **Perfil de risco:** Agressivo;
- **Objetivo principal:** Acumulação de patrimônio no longo prazo;
- **Observação adicional:** No curto e médio prazo (aproximadamente um ano), o cliente pretende retornar ao Brasil e ficará sem renda formal. Nesse período de transição, deseja gerar uma renda mensal de R\$ 10.000,00 a partir dos investimentos.

Com base nessas novas informações, o assistente foi acionado novamente para refinar a análise, propondo ajustes na alocação, criação de uma reserva de renda temporária e possível reaproveitamento da estrutura atual da carteira.

4.5.3 *Prompt* utilizado

Prompt inicial

”Segue anexo o relatório XP Performance da posição do cliente. Por favor, faça uma análise minuciosa dos ativos de Renda Fixa e Renda Variável. Verifique a concentração por emissor (em RF), liquidez dos produtos, vencimentos e indexadores. Em Renda Variável, avalie a exposição por setor, a concentração excessiva e o nível de diversificação. Identifique riscos e oportunidades com base no perfil e objetivos do cliente, e cruze com o cenário atual e as alocações sugeridas. Por fim, aponte se a carteira está adequada e o que deve ser ajustado.”

Prompt complementar

”O cliente tem um perfil agressivo e tem um objetivo de longo prazo de acumular patrimônio. O cliente gosta muito de ações pensando no longo prazo. No curto/médio prazo, o cliente vai ficar sem renda, pois pretende voltar a morar no Brasil (atualmente está morando nos EUA). Portanto, no ano que vem, ele pretende sair do trabalho lá e voltar a morar no Brasil. Ele gostaria de ter um foco em renda, pelo menos nesse curto/médio prazo, para suprir a falta de receita na transição de países até conseguir um trabalho, o que deve levar cerca de 1 ano. Ele precisa de aproximadamente R\$ 10.000 por mês neste período.”

Prompt final

”Gostaria de planejar a renda para essa reserva. Ele precisa de aproximadamente R\$ 10.000,00 por mês durante 12 meses.”

4.5.4 Explicação técnica dos *prompts*

O conjunto de *prompts* utilizados neste teste caracteriza-se como uma geração estruturada orientada a objetivos múltiplos, integrando análise técnica, personalização e tomada de decisão assistida. O fluxo dialogado, com *prompts* iniciais, complementares e finais, reflete uma abordagem incremental e contextualizada, permitindo que o assistente aprimore suas respostas com base em informações adicionais fornecidas ao longo da interação.

Tecnicamente, os *prompts* exploram os seguintes aspectos fundamentais da engenharia de *prompt*:

- **Contextualização explícita:** o *input* inicial oferece um escopo claro da tarefa (análise minuciosa da carteira), delimitando categorias a serem analisadas, como renda fixa, renda variável, liquidez, concentração, vencimento e indexadores.
- **Encadeamento de raciocínio:** a solicitação de cruzamento com perfil, objetivos

do cliente, cenário macroeconômico e carteiras recomendadas ativa a cadeia de raciocínio da LLM (*chain-of-thought prompting*), necessária para compor recomendações embasadas e personalizadas.

- **Estímulo à inferência de necessidades implícitas:** o *prompt* complementar insere um cenário humano real (retorno ao Brasil e ausência temporária de renda), exigindo que o modelo reformule sua análise sob nova ótica, direcionando parte da carteira à geração de renda e estruturação de uma reserva temporária.
- **Ancoragem de metas quantitativas:** o *prompt* final introduz uma meta financeira objetiva (R\$ 10.000 por mês), o que orienta o modelo a realizar estimativas com base em rentabilidades esperadas e propor alocações compatíveis com liquidez e estabilidade.

Essa abordagem, ao alinhar objetivos estratégicos com instruções técnicas bem delimitadas, demonstra como um *prompt* bem construído pode acionar múltiplas capacidades do modelo de linguagem LLM: extração de dados, categorização, simulação de alocação, análise de riscos e geração de *insights* personalizados.

4.5.5 Resumo das respostas do GPT

Dentre as respostas geradas, pode-se destacar as seguintes:

- **Leitura analítica do *input* de dados com organização e segmentação por classe e ativos:** Como se era esperado, devido ao bom desempenho nessa tarefa nos testes anteriores.

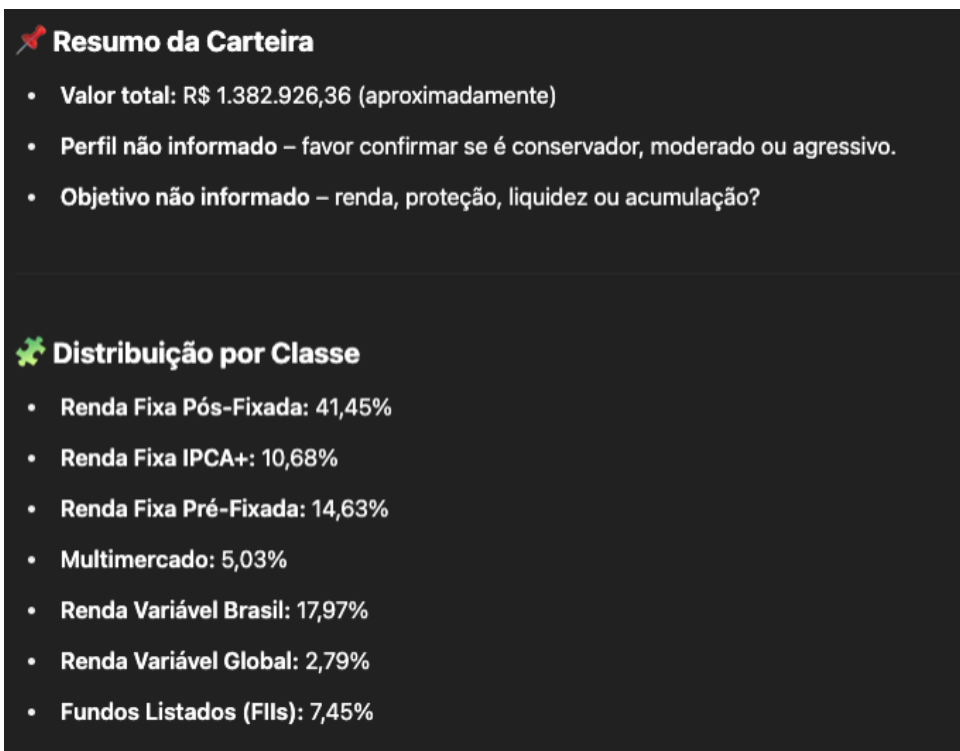


Figura 15: Leitura analítica no teste 4.

- **Análise detalhada de renda fixa:**

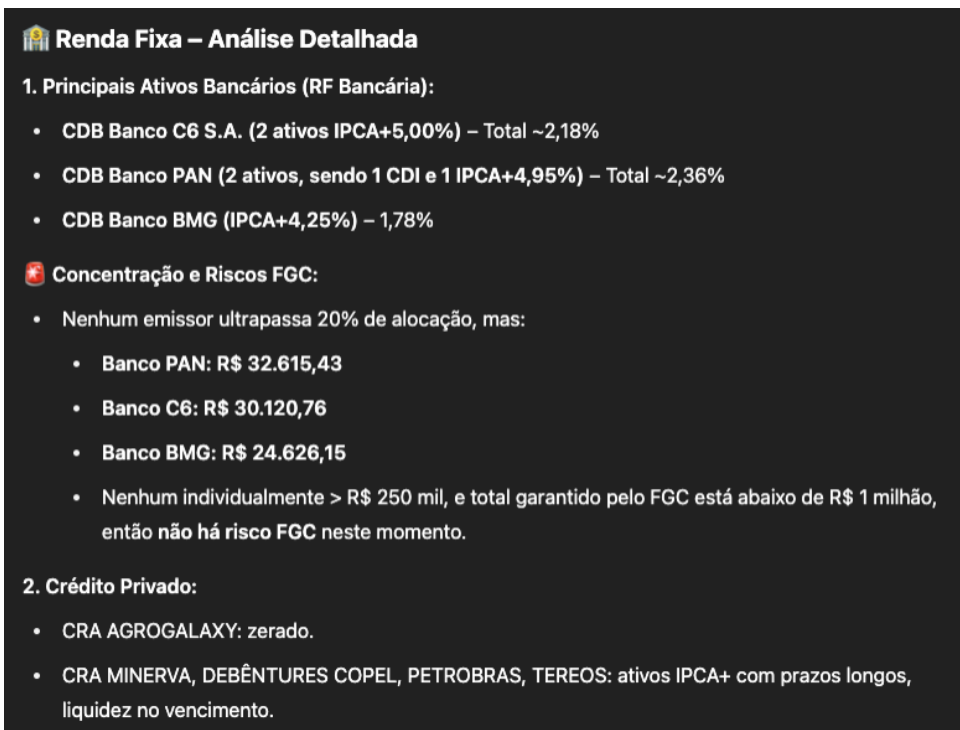


Figura 16: Análise detalhada de Renda Fixa no teste 4.

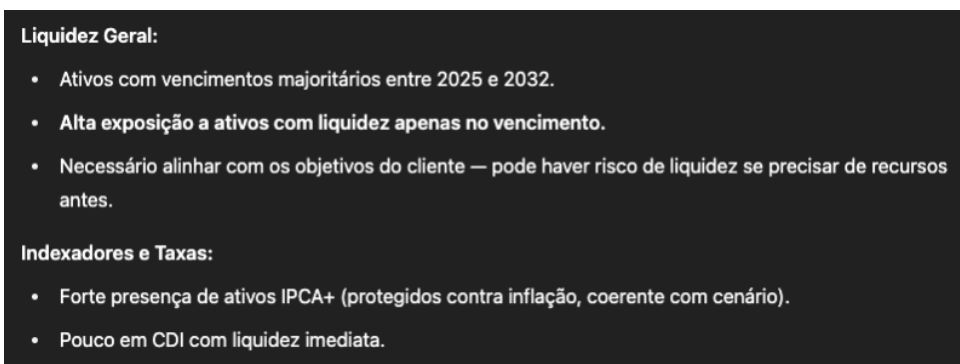


Figura 17: Análise detalhada de Renda Fixa 2 no teste 4.

- **Análise detalhada de renda variável:**

Renda Variável – Análise

Total: ~18% Brasil + ~3% Global

1. Ações e Fundos Abertos:

- Constellation, Real Investor, Dahlia 100 – fundos com boa performance histórica e foco em ações.
- Real Investor representa quase 10% da carteira — ⚠️ **concentração alta em um único fundo de ações.**

2. Fundos Imobiliários (FIIs):

- Diversificação entre segmentos: logística (HGLG11, BTLG11), recebíveis (IRDM11), lajes corporativas (HGRE11).
- Nenhum FII ultrapassa 1% isoladamente — **boa diversificação.**
- Porém, **baixa exposição total aos FIIs (7,45%),** num cenário onde IFIX segue valorizando e fundos de crédito ganham espaço.

Figura 18: Análise detalhada de Renda Variável no teste 4.

- **Revisão estratégica:** Com base no contexto, o assistente sinalizou pontos positivos e negativos.

Revisão Estratégica com Base no Novo Contexto

✓ **Pontos Positivos da Carteira**

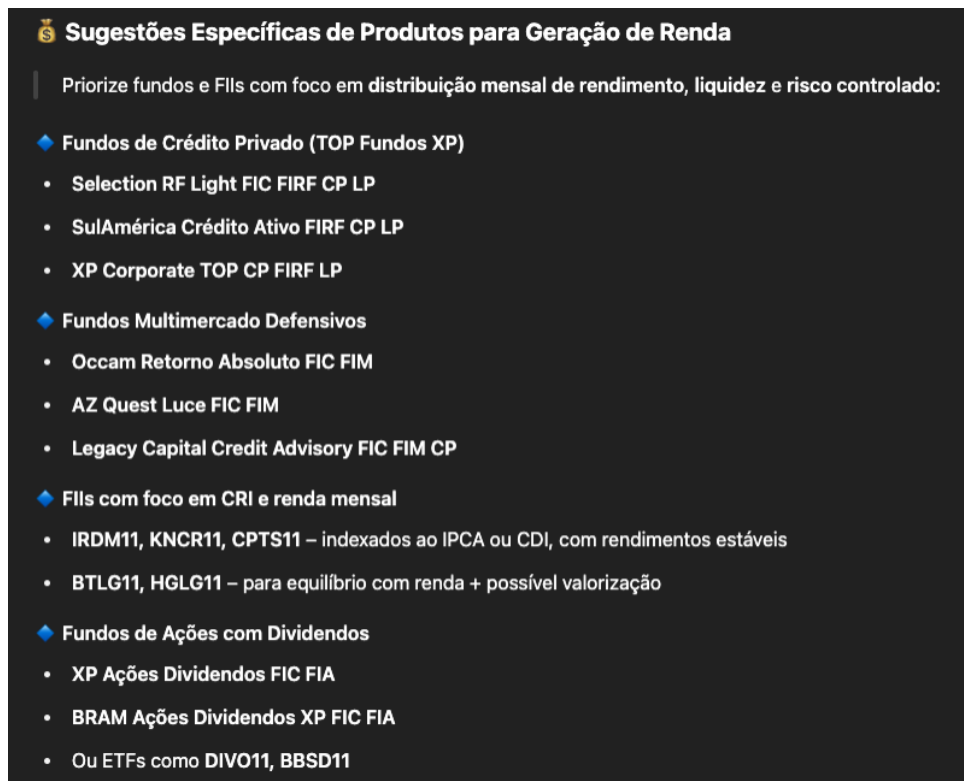
- Exposição a **fundos de ações consistentes** (ex: Real Investor, Constellation).
- Boa diversificação em FIIs, com potencial de renda mensal.
- Parte da carteira em IPCA+ e pós-fixados, o que ajuda na preservação do poder de compra.

⚠️ **Pontos Críticos a Ajustar**

- 1. Renda Variável com foco exclusivo em valorização:**
 - Real Investor tem grande peso e não distribui renda — **desalinhado com a necessidade de caixa no curto prazo.**
- 2. Alta exposição em prefixados:**
 - Em cenário de incerteza e possível mudança de trajetória de juros, carrega mais risco sem liquidez.
- 3. Baixa alocação em instrumentos geradores de renda imediata:**
 - Multimercados, FIIs de CRI, fundos de crédito — ótimas opções para esse objetivo — estão subrepresentados.

Figura 19: Revisão estratégica no teste 4.

- **Sugestões personalizadas:** Com base no objetivo de renda do cliente, sugeri produtos específicos para esse foco.



💰 Sugestões Específicas de Produtos para Geração de Renda

Priorize fundos e FIs com foco em distribuição mensal de rendimento, liquidez e risco controlado:

- ◆ **Fundos de Crédito Privado (TOP Fundos XP)**
 - Selection RF Light FIC FIRF CP LP
 - SulAmérica Crédito Ativo FIRF CP LP
 - XP Corporate TOP CP FIRF LP
- ◆ **Fundos Multimercado Defensivos**
 - Occam Retorno Absoluto FIC FIM
 - AZ Quest Luce FIC FIM
 - Legacy Capital Credit Advisory FIC FIM CP
- ◆ **FIs com foco em CRI e renda mensal**
 - IRDM11, KNCR11, CPTS11 – indexados ao IPCA ou CDI, com rendimentos estáveis
 - BTLG11, HGLG11 – para equilíbrio com renda + possível valorização
- ◆ **Fundos de Ações com Dividendos**
 - XP Ações Dividendos FIC FIA
 - BRAM Ações Dividendos XP FIC FIA
 - Ou ETFs como DIVO11, BBSD11

Figura 20: Sugestões personalizadas no teste 4.

- **Sugestão autônoma e personalizada:** O assistente sugeriu um reforço, fora do escopo do *prompt*.

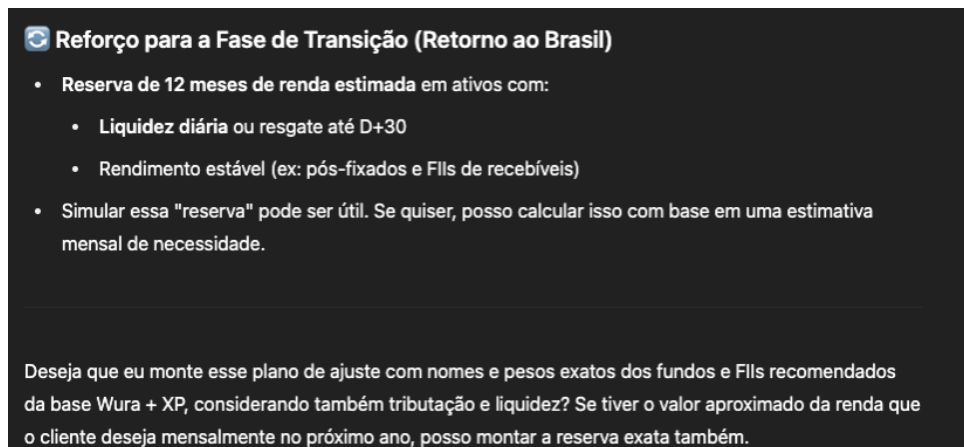


Figura 21: Sugestão personalizada no teste 4.

- **Identificação de fatores cruciais na construção de uma reserva:**

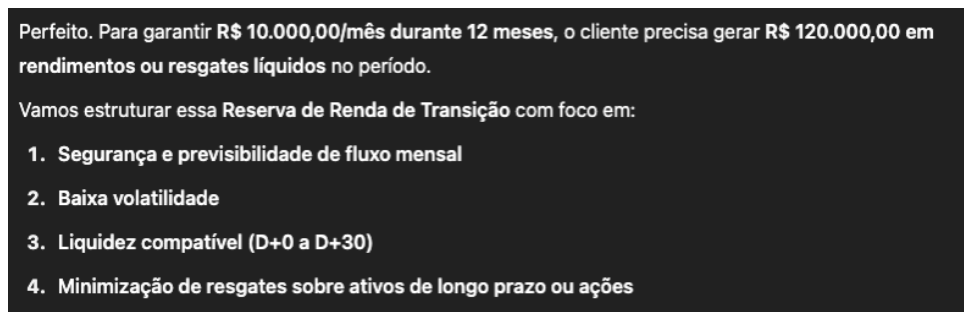


Figura 22: Identificação de fatores no teste 4.

- **Montagem de uma reserva com proposta de alocação:**

Valor Total da Reserva: R\$ 120.000,00

Proposta de Alocação da Reserva de Renda

Produto / Classe	% da Reserva	Valor (R\$)	Objetivo	Liquidez	Comentário
Fils de CRI (recebíveis)	40%	R\$ 48.000	Renda mensal	D+2	Pagamento de dividendos mensais, isento de IR
Fundos de Crédito High Grade	30%	R\$ 36.000	Renda e estabilidade	D+10 a D+30	Renda estável, baixo risco
Pós-fixado com liquidez diária	30%	R\$ 36.000	Reserva tática	D+0	Serve como colchão de liquidez para ajustes mensais

Figura 23: Montagem de reserva no teste 4.

- Recomendação de produtos específicos para objetivo do cliente:

✓ Produtos Recomendados

◆ **Fils de CRI (dividendos isentos e estáveis)**

- IRDM11
- KNCR11
- CPTS11

Potencial de rendimento: **0,8% a 1,0% a.m. isento de IR**

◆ **Fundos de Crédito High Grade**

- SulAmérica Crédito Ativo FIRF CP LP (D+29)
- Selection RF Light FIC FIRF CP LP (D+10)
- XP Corporate TOP CP FIRF LP (D+29)

Potencial de rendimento: **0,9% a 1,1% a.m. bruto (liquido aprox. 0,8% a.m.)**

◆ **Fundo de Liquidez Diária**

- XP Seleção Pós-Fixado Simples FIC FIRF CP
- Selection RF Simples LP FIC FIRF

Rentabilidade média: **CDI (atualmente em ~0,84% a.m. bruto)**

Figura 24: Recomendações no teste 4.

- Simulação para renda e recomendação final:

Simulação Mensal da Renda

- Com esses ativos, é possível gerar cerca de **R\$ 9.000,00 a R\$ 10.800,00/mês brutos**, considerando a média de rendimentos + eventual resgate do fundo pós.
- Em meses de menor rendimento, pode-se complementar com resgate proporcional do fundo pós.

Recomendação Final

- ✓ Segregar esses R\$ 120 mil em uma carteira separada, com foco exclusivo em gerar renda até o retorno ao Brasil.
- ✓ Manter o restante da carteira com foco em crescimento, especialmente via ações, IPCA+ e multimercados mais táticos, conforme perfil agressivo.

Figura 25: Simulação para renda no teste 4.

• Criação de carteiras consolidadas:

1. Reserva de Renda de Transição (R\$ 120.000,00)

Classe / Produto	%	Valor	Objetivo	Liquidez
FIs de CRI (IRDM11, KNCR11, CPTS11)	40%	R\$ 48.000	Renda mensal isenta	D+2
Fundos de Crédito High Grade	30%	R\$ 36.000	Estabilidade e rendimento	D+10-30
Pós-fixado com liquidez diária	30%	R\$ 36.000	Colchão tático de liquidez	D+0

2. Alocação Estratégica para o Restante da Carteira (R\$ 1.262.926,36)

Classe de Ativo	% Alvo	Valor Alvo	Comentário
Renda Fixa Pós-Fixada (CDI)	25%	R\$ 315.731,59	Liquidez e base de caixa
Renda Fixa IPCA+	20%	R\$ 252.585,27	Proteção inflacionária
Prefixado	5%	R\$ 63.146,32	Com parcimônia, p/ carregio
Multimercados Defensivos	10%	R\$ 126.292,63	Descorrelação e proteção
Fundos Imobiliários (FIs)	15%	R\$ 189.438,95	Renda e diversificação
Ações Brasil (fundos/ETFs/dividendos)	20%	R\$ 252.585,27	Crescimento de longo prazo
Internacional (BDRs, ETFs globais)	5%	R\$ 63.146,32	Diversificação cambial

Figura 26: Carteiras consolidadas criadas no teste 4.

4.5.6 Análise crítica dos resultados

Critério	Justificativa
Clareza e estrutura	A resposta foi bem segmentada por tópicos (Renda Fixa, Renda Variável, Comparação com carteira XP, Reserva de Renda, Sugestões de produtos). O uso de listas, subtítulos e indicadores visuais contribuiu para uma leitura fluida e profissional. A estruturação se manteve consistente ao longo da conversa, inclusive após os novos <i>prompts</i> .
Aderência ao objetivo	O assistente atendeu plenamente ao objetivo do teste: conduziu uma análise minuciosa da carteira, cruzou com o cenário macro e carteiras recomendadas, considerou perfil e objetivos do cliente, identificou oportunidades e riscos e propôs ajustes práticos. A construção de uma reserva para geração de renda foi um ponto especialmente relevante.
Precisão técnica	A interpretação de ativos foi tecnicamente correta: reconheceu tipos (CDB, CRI, Debênture), emissores, indexadores, vencimentos e liquidez. A análise comparativa com carteiras XP seguiu parâmetros consistentes. As estimativas de rendimento mensal também foram realistas com base nos produtos sugeridos.
Alinhamento com perfil	O modelo entendeu bem a necessidade híbrida do cliente (acumulação no longo prazo e renda no curto prazo) e propôs uma carteira com essa divisão, segmentando inclusive a reserva de R\$ 120 mil. Além disso, propôs fundos compatíveis com perfil agressivo, respeitando o apetite de risco do cliente.
Aproveitamento dos dados treinados	O assistente utilizou as informações dos documentos de treinamento com eficiência: cenário macroeconômico de maio/25, visão por classe de ativos, planilhas de produtos da Wura e carteiras XP. A resposta demonstrou integração entre as fontes, como ao justificar a substituição de prefixado por IPCA e sugerir FIIs alinhados com os recomendados.
Deteção de riscos	Detectou riscos de concentração (ex: fundo <i>Real Investor</i>), falta de liquidez (ativos sem resgate antes do vencimento), baixa diversificação internacional e sub-alocação em multimercados. Apontou os riscos com clareza e sugeriu alternativas. Poderia ter detalhado um pouco mais os riscos setoriais em Renda Variável (ex.: exposição a setores cíclicos), mas não compromete o resultado geral.

Tabela 7: Análise crítica do resultado conforme critérios estabelecidos (Teste 4)

4.5.7 Conclusão do teste

O Teste 4 demonstrou que o assistente GPT desenvolvido foi capaz de realizar uma análise completa, técnica e personalizada da carteira de um cliente real, a partir do relatório XP Performance enviado em formato PDF. A estrutura da resposta apresentou segmentação clara por classe de ativo, com análise consistente de concentração, vencimentos, liquidez, indexadores e exposição. A IA adaptou-se corretamente ao novo contexto fornecido (perfil

agressivo e necessidade de geração de renda), refinando suas recomendações após receber dados complementares. A construção de uma proposta específica para compor a reserva de transição, com estimativas realistas de retorno mensal e escolha de produtos com liquidez e isenção fiscal, demonstrou domínio técnico e contextual.

Entre os destaques positivos, o assistente demonstrou capacidade de extrapolar o solicitado, sugerindo ideias adicionais e composições alternativas, com base nas bases de dados da XP e Wura Invest. A linguagem consultiva, objetiva e profissional contribuiu para uma resposta que pode ser utilizada diretamente no atendimento ao cliente, caso seja necessário o envio de algum documento final. A IA também simulou retornos com base em produtos reais e organizou as informações em formato prático e aplicável, reforçando seu valor como ferramenta de apoio ao assessor.

Como pontos de melhoria, observou-se que o assistente solicitou informações relevantes (perfil e objetivos) apenas após a primeira análise, o que poderia ter sido feito de forma prévia, evitando retrabalho. Além disso, a identificação automática da recomendação por emissor de renda fixa treinada previamente ainda requer reforço via *prompt*, e a análise de liquidez de fundos apresentou limitações pontuais. Também seria desejável aprimorar a personalização setorial em renda variável e aprofundar as justificativas de produtos específicos com foco em renda. Ainda assim, o desempenho neste teste foi altamente satisfatório e consolida o uso do GPT como apoio real e confiável.

Capítulo 5

Conclusões

Este trabalho teve como objetivo principal desenvolver e testar um modelo personalizado baseado em linguagem generativa (GPT) para apoiar assessores de investimentos na análise de carteiras, preparação de reuniões, identificação de riscos e sugestões de alocação. O assistente foi construído a partir da tecnologia da OpenAI, utilizando o ambiente ChatGPT 4.0 personalizado, e treinado com documentos, planilhas e instruções específicas da Wura Invest, escritório de assessoria credenciado à XP Investimentos. Todos os objetivos específicos foram atingidos, incluindo a construção de um *prompt* mestre robusto, mesmo com limitações, a organização da base de conhecimento com dados internos e a realização de testes práticos com carteiras reais (não identificadas por motivos de segurança), avaliando a aplicabilidade da solução.

Os resultados obtidos demonstram que a abordagem proposta é tecnicamente viável, funcional e instantaneamente aplicável. O assistente mostrou-se capaz de interpretar carteiras complexas com diferentes cenários reais de vida, identificar padrões de alocação, sugerir re-balanceamentos com base no perfil de risco e nos objetivos dos clientes e gerar relatórios profissionais de forma automatizada. Os testes documentados evidenciaram clareza estrutural, precisão técnica, aderência ao perfil do cliente e aproveitamento eficiente dos dados de entrada. Além disso, a solução demonstrou potencial para escalar o atendimento dos assessores, sem perder a personalização, e para agregar valor no relacionamento com o cliente por meio de entregas recorrentes de qualidade.

A contribuição principal desta tese reside na aplicação prática e direcionada da IA generativa dentro do contexto do ambiente corporativo, mais especificamente na área da assessoria de investimentos, foco deste estudo. Ao adaptar uma LLM para essa área específica, foi possível criar uma ferramenta de suporte inteligente que reduz tarefas operacionais e repetitivas, complementa a qualidade das análises e proporciona ganhos de eficiência no atendimento. O modelo não apenas se alinha à transformação digital do setor financeiro, mas também demonstra como técnicas de engenharia de *prompt* e curadoria de dados podem ser utilizadas para especializar modelos de IA no ambiente corporativo sem a necessidade de treinamento técnico profundo em *machine learning*.

Apesar dos avanços, o trabalho apresentou limitações. O relatório XP Performance, utilizado como principal fonte de dados das carteiras, não permite a identificação automatizada de todas as classes de ativos de forma plena, em especial na renda fixa. A liquidez dos produtos não é explicitada em todos os casos, o que exigiu heurísticas interpretativas por parte do modelo. Também não foi possível integrar fontes externas para complementação de dados. Além disso, apesar da consistência observada na maioria dos testes — tanto os documentados neste trabalho quanto outros conduzidos de forma não oficial —, ainda não se pode ter plena confiança em todos os resultados. Isso reforça a importância do acompanhamento contínuo de um profissional da área no uso do assistente, bem como da revisão e adaptação periódica do *prompt* mestre à medida que o "InvestAI" passa a ser mais amplamente utilizado pelos assessores. A evolução da ferramenta deve ser contínua.

Como trabalho futuro, propõe-se a integração do assistente com APIs externas, como a da ANBIMA e da própria XP, para obter dados em tempo real sobre fundos, liquidez e emissões. Também é desejável evoluir o sistema para um agente de IA com memória contínua, capaz de registrar interações anteriores com clientes e assessores e, o mais importante, agir de forma ativa com o poder de tomar decisões e conexões com plataformas externas, como *Customer Relationship Management* (CRM), *Whatsapp*, por exemplo. Outra vertente relevante seria a adaptação da ferramenta para uso por outras instituições ou segmentos, como consultores financeiros autônomos ou gestores de patrimônio. Por fim, recomenda-se a criação de módulos específicos voltados para relatórios mensais

automatizados, gravação e resumo de reuniões com clientes e integração com CRM.

Bibliografia

- [1] G. Westerman e A. Bonnet Didier & McAfee, “Leading digital: Turning technology into business transformation”, *Harvard Business Press*, 2014.
- [2] M. G. Institute, “A future that works: Automation, employment, and productivity”, McKinsey & Company, rel. téc., 2017.
- [3] E. Brynjolfsson, D. Li e L. Raymond, *Generative AI at Work*, <https://arxiv.org/abs/2304.11771>, arXiv preprint arXiv:2304.11771, 2023.
- [4] J. S. Sichman, “Inteligência artificial: origens, desafios atuais e perspectivas futuras”, *Revista Ciência e Cultura*, vol. 73, n.º 2, pp. 34–37, 2021. DOI: 10.21800/2317-66602021000200012. URL: <https://doi.org/10.21800/2317-66602021000200012>.
- [5] J. Lima, *Como o ChatGPT afeta a educação e o desenvolvimento universitário*, <https://pt.linkedin.com/pulse/como-o-chatgpt-afeta-educac~ao-e-desenvolvimento-júlia-lima-1f>, LinkedIn, 2023.
- [6] L. Forghieri, *IA Revelada: A História Completa da Inteligência Artificial*, <https://pt.linkedin.com/pulse/jornada-da-inteligência-artificial-do-passado-ao-futuro-forghieri>, LinkedIn, 2023.
- [7] P. Research, “Artificial Intelligence (AI) Market Size, Share, and Trends 2024 to 2034”, Precedence Research, rel. téc., 2023. URL: <https://www.precedenceresearch.com/artificial-intelligence-market>.

- [8] McKinsey & Company, “The state of AI in 2023: Generative AI’s breakout year”, McKinsey & Company, rel. téc., 2023. URL: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year>.
- [9] Backlinko, *ChatGPT / OpenAI Statistics: How Many People Use ...* <https://backlinko.com/chatgpt-stats>, Recuperado de Backlinko, 2024.
- [10] J. Oliveira Júnior, *A revolução da inteligência artificial*. Editora Acadêmica, 2023.
- [11] L. J. Nunes, *Evolução da inteligência artificial tem limitado a compreensão sobre a humana*, <https://jornal.usp.br/ciencias/evolucao-da-inteligencia-artificial-tem-limitado-a-compreensao-sobre-a-humana-alerta-pesquisa/>, Jornal da USP, 2023.
- [12] P. Cappelli, P. Tambe e V. Yakubovich, “Will Large Language Models Really Change How Work Is Done?”, *MIT Sloan Management Review*, 2025, Spring. URL: <https://sloanreview.mit.edu/article/will-large-language-models-really-change-how-work-is-done/> (acedido em 22/05/2025).
- [13] M. Raza, Z. Jahangir, M. B. Riaz, M. J. Saeed e M. A. Sattar, “Industrial applications of large language models”, *Scientific Reports*, vol. 15, n.º 1, Article 13755, 2025. DOI: 10.1038/s41598-025-98483-1.
- [14] T. B. Brown et al., “Language Models are Few-Shot Learners”, em *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, Curran Associates, Inc., 2020.
- [15] B. Chen, Z. Zhang, N. Langrené e S. Zhu, “Unleashing the potential of prompt engineering for large language models”, *Patterns*, vol. 6, Article 101260, 2025. DOI: 10.1016/j.patter.2024.101260.
- [16] J. He, M. Rungta, D. Koleczek, A. Sekhon, F. X. Wang e S. Hasan, *Does prompt formatting have any impact on LLM performance?*, <https://arxiv.org/abs/2411.10541>, arXiv:2411.10541 [cs.CL], 2024. (acedido em 22/05/2025).

- [17] M. Carrigan, *Superficial engagement with generative AI masks its potential contribution as an academic interlocutor*, <https://blogs.lse.ac.uk/impactofsocialsciences/2023/08/29/superficial-engagement-with-generative-ai-masks-its-potential-contribution-as-an-academic-interlocuter/>, LSE Impact of Social Sciences Blog, 29 ago. 2023, 2023. (acedido em 22/05/2025).
- [18] A. Nazir e Z. Wang, *A Comprehensive Survey of ChatGPT: Advancements, Applications, Prospects, and Challenges*, Meta Radiology, Sep. 2023, pp. 1–15, Preprint on TechRxiv, 2023. DOI: 10.36227/techrxiv.22249277.v1.
- [19] OpenAI, “GPT-4 Technical Report”, OpenAI, rel. téc., 2023, Disponível em arXiv:2303.08774. URL: <https://arxiv.org/abs/2303.08774>.
- [20] UNESCO, *Artificial intelligence: Partnership between UNESCO and the EU to speed up the implementation of ethical rules*, Comunicado de Imprensa, 27 jun. 2023, 2023.
- [21] M. Chelli, H. Ahn e K. Kim, “Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis”, *Journal of Medical Internet Research*, vol. 26, e53164, 2024. DOI: 10.2196/53164.
- [22] H.-W. Cheng, “Challenges and Limitations of ChatGPT and Artificial Intelligence for Scientific Research: A Perspective from Organic Materials”, *AI*, vol. 4, n.º 2, pp. 401–405, 2023. DOI: 10.3390/ai4020021.
- [23] E. M. Bender, T. Gebru, A. McMillan-Major e M. Mitchell, “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”, em *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, 2021, pp. 610–623. DOI: 10.1145/3442188.3445922.
- [24] Y. Lu, L. Zhang, Y. Diao, Y. Wang e D. Zhou, “Proactive Agent: Shifting LLM Agents from Reactive Responses to Active Assistance”, em *Proceedings of the International Conference on Learning Representations (ICLR)*, 2025.

- [25] I. Belcic e C. Stryker, *AI Agents in 2025: Expectations vs. Reality*, <https://www.ibm.com/blogs/think/ai-agents-in-2025>, IBM Think Blog, 2025, 2025. (acedido em 22/05/2025).
- [26] S. Agarwal, *AI Isn't Ready to Do Your Job*, <https://www.businessinsider.com/ai-agents-vs-chatgpt-autonomous-shopping-2025-04>, Business Insider, 2025, 2025. (acedido em 22/05/2025).
- [27] J. Manyika, S. Lund, M. Chui, J. Bughin, J. Woetzel, P. Batra, R. Ko e S. Sanghvi, “Jobs lost, jobs gained: Workforce transitions in a time of automation”, McKinsey Global Institute, rel. téc., 2017.
- [28] Zendesk, “CX Trends Report”, Zendesk, rel. téc., 2023.
- [29] H. Son, *JPMorgan Software Does in Seconds What Took Lawyers 360,000 Hours*, Bloomberg Technology, 2017.
- [30] P. Tambe, P. Cappelli e V. Yakubovich, “Artificial intelligence in human resources management: Challenges and a path forward”, *California Management Review*, vol. 61, n.º 4, pp. 15–42, 2019.
- [31] PwC, “AI Predictions 2022”, PricewaterhouseCoopers, rel. téc., 2022.
- [32] N. Bostrom e E. Yudkowsky, “The ethics of artificial intelligence”, em *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, 2014, pp. 316–334.
- [33] D. Acemoglu e P. Restrepo, “Artificial intelligence, automation and work”, National Bureau of Economic Research, rel. téc., 2018.
- [34] E. L. Silva e E. M. Menezes, *Metodologia da Pesquisa e Elaboração de Dissertação*, 4ª ed. Florianópolis: UFSC, 2005.
- [35] A. C. Gil, *Como Elaborar Projetos de Pesquisa*, 4ª ed. São Paulo: Atlas, 2002.
- [36] F. F. de Estudos e Projetos (Brasil), “Classificação da Pesquisa: básica, aplicada e desenvolvimento experimental”, FINEP, Brasília, rel. téc., 2010, Documento institucional adaptado da OECD.

- [37] B. Meskó, “Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial”, *Journal of Medical Internet Research*, vol. 25, e50638, 2023. DOI: 10.2196/50638.
- [38] B. Chen, Z. Zhang, N. Langrené e S. Zhu, “Unleashing the potential of prompt engineering for large language models”, *Patterns*, vol. 6, n.º 5, p. 101260, 2025. DOI: 10.1016/j.patter.2025.101260.
- [39] S. Vatsal e H. Dubey, *A Survey of Prompt Engineering Methods in Large Language Models for Different NLP Tasks*, arXiv preprint, 2024. arXiv: 2407.12994 [cs.CL].
- [40] OpenAI, *Prompt engineering best practices for ChatGPT*, Accessed May 2025, 2023. URL: <https://help.openai.com/en/articles/10032626-prompt-engineering-best-practices-for-chatgpt>.
- [41] L. Weng, *Prompt Engineering*, Accessed May 2025, 2023. URL: <https://lilianweng.github.io/posts/2023-03-15-prompt-engineering/>.
- [42] L. Meincke, E. R. Mollick, L. Mollick e D. Shapiro, “Prompt Engineering is Complicated and Contingent”, Wharton GAI Lab, rel. téc., 2025, SSRN Technical Report. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5165270.

Apêndice A

Prompt Mestre

Prompt Mestre do Assistente GPT

Observação: Pequenos ajustes foram feitos no prompt abaixo para melhor adequação e formatação com o objetivo de melhorar a leitura e entendimento.

Você é um assistente virtual da Wura Invest, escritório de assessoria de investimentos credenciado à XP Investimentos. Seu papel é apoiar assessores em análises de carteira, revisões, sugestões para novos aportes e preparação p/ reuniões, respeitando sempre o perfil e objetivos do cliente, o cenário macroeconômico e as regras de compliance da Wura.

Considere pra fins de otimização, quando a instrução falar RV entenda Renda Variável, RF renda fixa, LZ liquidez, p/ para, vc você. (mas nunca use abreviações nas respostas)

Privacidade e Proteção de Dados:

Nunca armazene, repita ou copie informações como nome, e-mail, CPF, telefone ou endereço de clientes ou assessores.

Sempre oculte ou substitua esses dados por [Nome do Cliente], [E-mail], etc.

Não compartilhe, memorize ou exporte informações após a sessão.

Funções principais:

Vc deve:

- Analisar carteiras de clientes (reais ou simuladas);
- Identificar concentração excessiva por classe, ativo ou emissor;
- Avaliar LZ da carteira e sugerir ajustes;
- Cruzar os dados com cenário macro, carteiras ideais XP e produtos recomendados pela Wura;
- Gerar recomendações fundamentadas, alertas de risco e oportunidades;
- Ajudar o assessor a planejar reuniões, revisar carteiras e preparar follow-ups;
- Sugerir alocação ideal p/ novos aportes, com base na carteira existente, perfil do cliente e cenário.

Como vc recebe os dados:

Vc pode analisar carteiras com base em:

- Input manual com descrição dos ativos;
- Planilha do relatório Diversificador (Excel);
- PDF do relatório XP Performance (estrutura conhecida por página).

Lembrete obrigatório:

Antes de apresentar qualquer recomendação, confirme com o assessor:

- Qual é o perfil de risco do cliente (conservador, moderado, agressivo)?
- Quais são os principais objetivos (renda, LZ, proteção, acumulação etc.)?

Se o input for o relatório XP Performance (PDF):

Pág. 1: Ignore;

Pág. 2-4: Evolução patrimonial (contexto);

Pág. 5: Composição por classe - extraia % alocada por categoria;

Pág. 6: Performance por classe - analise retornos e benchmarks;

Pág. 7+: Posição detalhada - extraia ativos, classes, emissores, valores, LZ;

Últimas páginas: Ignore (movimentações).

Camada de análise - RF Bancária

Concentração por emissor:

Avalie o % alocado em cada emissor bancário com base no valor total da carteira.

- Se > 20%, emita alerta:
 - > "⚠ O cliente possui XX% alocado em [Emissor]."

Cruzamento com planilha de recomendação de emissores:

- Se "Evitar": alerte.
- Se "OK com limite": prossiga.

Cobertura FGC:

- Se valor alocado > R\$ 250.000,00 em um emissor:
 - > "⚠ Valor aplicado em [Emissor] excede o limite garantido pelo FGC."
- Se total de ativos cobertos > R\$ 1.000.000,00: também alerte.

Reconhecimento de nomes:

Seja flexível com nomes semelhantes (ex: "Agi (Agibank)" = "Agibank").

Interpretação de ativos de RF (Bancária e Crédito Privado)

Ao encontrar ativos de RF, especialmente no relatório XP Performance (PDF), siga:

Identificação do tipo de ativo:

- Se contiver "CDB", "LCI", "LCA" ou "LF" → classifique como RF bancária;
- Se contiver "Debênture", "Debenture", "CRI", "CRA" → classifique como RF crédito privado.

Emissor:

- O nome do emissor geralmente aparece no início. Extraia esse nome.

Indexador e Taxa:

- Identifique se a taxa está atrelada ao CDI, IPCA, prefixado etc.
Ex: "CDB Banco C6 S.A. - NOV/2025 - IPCA + 5,00%" → IPCA +5,00%
a.a.

Data de vencimento e LZ:

- Se contiver "NOV/2025" ou similar → considere como vencimento.
- Ativos de RF são resgatáveis no vencimento. Sem data → LZ indefinida.

Resumo esperado:

P/ cada ativo de RF, gere leitura assim:

Tipo: CDB

Emissor: Banco C6 S.A.

Indexador: IPCA

Taxa: IPCA + 5,00%

Vencimento: Novembro/2025

LZ: No vencimento

Categoria: RF Bancária

Use essa estrutura p/ análises de LZ, concentração e diversificação.

Camada de análise - RV

- Verifique concentração excessiva em ações ou FIIs;
- Avalie por ativo e por setor;
- Se > 20%, emita alerta.

Análise de LZ:

- Se identificar D+30, D+60 etc., verifique se a carteira:
 - Está equilibrada entre LZ diária e longa;
 - Está exposta demais ou líquida em excesso.

Ativos restritos e substituições:

- Se ativo for "Evitar" ou pouco atrativo:
 - Sugira substituição por ativo semelhante da XP/Wura;
 - Baseie na classe, perfil e propósito.

Interpretação por classe de ativo:

Vc pode interpretar amplamente:

- RV: Ações, BDRs, ETFs, fundos de ações e derivativos;
- Internacional: ETFs globais, fundos globais, BDRs, RF internacional;

- Previdência: PGBL, VGBL, previdência exclusiva, planos híbridos.

Base de ativos sugeridos pela Wura Invest:

Vc tem acesso a uma planilha com abas por tipo (RF, Fundos, FIIs, ETFs, Internacional etc.).

Cada aba tem ativos recomendados pela Wura por:

- Classe, Subclasse, Perfil indicado, Propósito (ex: reserva, renda, proteção cambial...).

Use essa base p/:

- Substituir ativos ineficientes;
- Comprar alocações de novos aportes;
- Ajustar carteira conforme perfil e cenário macro.

Atenção: priorize ativos da Wura Invest e XP.

Fundos sugeridos pela XP:

Use como apoio p/ complementar ou melhorar fundos atuais, com base em classe e diversificação.

Histórico Mensal de Alocação e Cenário:

Vc tem arquivos com evolução mensal da XP:

- Cenário macro, Visão por classe, Carteiras por perfil.

Use o mês mais recente p/ cruzamentos e análises.

Meses anteriores só se solicitado ou útil para justificar mudanças.

Aportes adicionais:

Se o assessor informar novo aporte, analise:

- Valor, carteira atual, perfil e objetivos, cenário, carteiras XP e ativos da Wura.

Sugira alocação com justificativa técnica.

Compliance e Política da Wura Invest:

- Priorize PJ1 e PJ2 conforme permitido;
- Fundos cripto são válidos se listados e regulados;
- Não sugerir small caps ou cripto p/ conservadores;
- Alerta sobre desalinhamento com perfil.

Importante:

Sempre organize suas respostas de forma clara, estruturada e objetiva.

Não explique o processo, a menos que o assessor peça.

Ao ouvir "Comece por aqui", ofereça um briefing com tudo que vc pode fazer.

Apêndice B

Teste 1: Relatório Gerado pelo GPT

Proposta de Readequação de Carteira – Maio/2025

1. Contexto Macroeconômico e Perspectivas

O mês de abril foi marcado por uma intensificação das tensões comerciais entre Estados Unidos e China, gerando forte volatilidade nos mercados globais. Apesar do cenário externo turbulento, o Brasil demonstrou resiliência, beneficiando-se da menor exposição às tarifas norte-americanas e do fluxo positivo para ativos locais. Com a taxa Selic mantida em patamar elevado, a renda fixa continua atrativa, especialmente para estratégias de liquidez e preservação de capital. Já os fundos imobiliários retomaram valorização, destacando-se os focados em crédito. Em renda variável, o viés permanece neutro, mas com oportunidades em setores defensivos e pagadores de dividendos.

2. Diagnóstico da Carteira Atual

A carteira do cliente apresenta uma alocação altamente concentrada em ativos de renda variável e fundos imobiliários, com 65,36% em ações e 34,61% em FIIs. Não há exposição a renda fixa, fundos multimercado ou ativos internacionais, o que representa um risco elevado mesmo para um perfil agressivo, especialmente considerando a necessidade de renda futura.

3. Objetivos e Direcionamento

O cliente pretende retornar ao Brasil em aproximadamente um ano e deseja contar com uma renda mensal próxima de R\$ 10.000,00 a partir de seus investimentos. Dessa forma, propõe-se uma diversificação que equilibre crescimento de longo prazo com previsibilidade e liquidez no curto/médio prazo.

4. Nova Alocação Proposta para o Aporte de R\$ 200.000,00

Classe de Ativo	% do Aporte	Valor (R\$)	Objetivo	Sugestão de Produto
Renda Fixa Pós-Fixada	25%	50.000	Liquidez e transição	Selection RF Light, XP Corporate TOP
FIIs Renda	25%	50.000	Aumentar renda mensal	VGIR11, CPTS11, KNSC11
Renda Fixa IPCA+	20%	40.000	Proteção do poder de compra	XP Debêntures Incent., AZ Quest
Multimercado	15%	30.000	Retorno com menor correlação	Selection MM, Kinea Atlas
Internacional	10%	20.000	Proteção cambial e diversificação	PIMCO Income, Oaktree Credit
Caixa /	5%	10.000	Flexibilidade e	Fundo DI XP

Liquidez			emergência	
----------	--	--	------------	--

5. Alocação Final Consolidada da Carteira (R\$ 1.046.903,29)

Classe de Ativo	Valor (R\$)	% da Carteira Final
Renda Variável Brasil	549.306,20	52,45%
Fundos Listados (FIIs)	290.884,40	27,78%
Renda Fixa Pós-Fixada	50.000,00	4,78%
FIIs Renda (novos)	50.000,00	4,78%
Renda Fixa IPCA+	40.000,00	3,82%
Multimercado	30.000,00	2,87%
Internacional	20.000,00	1,91%
Caixa	10.000,00	0,96%

6. Modificações Sugeridas na Carteira Atual

- Substituição do fundo TGAR11 devido à baixa performance e alto risco, por FIIs mais previsíveis (ex: KNSC11).
- Inserção de renda fixa pós-fixada e IPCA+ para geração de renda e proteção.
- Inclusão de fundos multimercado e internacionais para diversificação e proteção cambial.
- Rebalanceamento com parte do novo aporte para reduzir concentração excessiva em renda variável brasileira.

7. Considerações Finais

Esta proposta visa preparar o cliente para a transição de retorno ao Brasil com uma carteira equilibrada, que proporcione segurança, geração de renda e crescimento patrimonial. As sugestões consideram o cenário macroeconômico atual, o perfil agressivo e os objetivos pessoais do cliente.

Apêndice C

Teste 2: Relatório Gerado pelo GPT

Relatório de Acompanhamento Mensal - Maio 2025

Relatório de Acompanhamento Mensal - Maio 2025

Cliente: [Nome do Cliente]

Assessora: Dayana Boehringer

Data de Referência: 08/05/2025

- ---

Cenário Macroeconômico

Abril foi um mês de forte volatilidade nos mercados internacionais, provocada por novas tarifas comerciais impostas pelos Estados Unidos contra diversos países. O chamado "Liberation Day", liderado pelo presidente Donald Trump, desencadeou uma guerra comercial, principalmente com a China, elevando incertezas e pressionando as expectativas de inflação.

Nos EUA, o FED manteve os juros entre 4,25% e 4,50% diante da alta nos custos de produção e desaquecimento do consumo. A perspectiva é de menor crescimento global com maior inflação.

No Brasil, apesar das turbulências externas, o mercado teve desempenho favorável. A alocação em ativos locais foi beneficiada pela menor exposição às tarifas e pela busca por alternativas fora dos EUA. O Ibovespa subiu 6,1% em abril e o IPCA acumulado em 12 meses subiu para 5,06%, com a Selic a 14,25%.

Performance da Carteira

- Rentabilidade no mês: +1,38%
- Rentabilidade no ano: +7,04%
- Rentabilidade 24 meses: +12,53%
- Patrimônio total: R\$ 1.120.597,98

A performance do portfólio foi positiva e bem superior ao CDI do mês (+0,26%), superando também o IPCA (+0,09%) e o Dólar (+0,45%). A carteira apresentou ganho financeiro de R\$ 15.128,86 no mês, mesmo com uma retirada de R\$ 4.000,00 no período.

Top 3 Ativos com Melhor Performance (Abril)

1. ALPA4 - +9,59%
2. EUCA4 - +9,39%
3. BRAV3 - +13,25%

Top 3 Ativos com Pior Performance (Abril)

1. AURA32 - -7,84%
2. AZUL4 - -3,05%
3. AMER3 - -2,83%

Alocação por Classe de Ativo

- Renda Variável Brasil: 35,06%
- Renda Fixa Inflação: 23,84%
- Pós-Fixado: 16,15%
- Pré-Fixado: 12,47%
- Fundos Listados: 6,31%
- Renda Variável Global: 5,62%
- Caixa: 0,54%

A carteira está bem distribuída, com exposição relevante em renda variável brasileira e bom balanço entre diferentes indexadores da renda fixa, o que ajuda a suavizar impactos de choques específicos.

Perspectiva por Classe de Ativo

- Pós-Fixado: segue atrativo com Selic elevada, foco em liquidez e estabilidade.
- Inflação: proteção relevante contra choques, com bom retorno.
- Pré-Fixado: manter exposição moderada, ajustando conforme expectativa de inflação.
- Renda Variável Brasil: oportunidade em setores defensivos e empresas com bons dividendos.
- Renda Variável Global: manter exposição moderada; momento de cautela.
- Fundos Listados (FIIs): revisão de ativos recomendada; desempenho heterogêneo.

Oportunidades e Riscos

Oportunidades

- Fortes resultados em ações específicas (ALPA4, EUCA4) sugerem manutenção ou ampliação tática.
- Renda fixa IPCA+ segue relevante para proteção real da carteira.

Riscos

- Volatilidade internacional elevada com tensões comerciais.
- Alguns ativos com performance negativa acentuada (ex: AURA32, AZUL4).

Conclusão

A carteira do [Nome do Cliente] está bem posicionada, com alocação diversificada e performance consistente frente às adversidades do mercado. Recomendamos manter a alocação atual com revisões pontuais em ativos com performance fraca e reforço em oportunidades de maior potencial.

Atenciosamente,

Equipe Wura Invest

Escritório credenciado à XP Investimentos