

APLICAÇÃO DO PROCESSAMENTO DE LINGUAGEM NATURAL PARA AUXILIAR O GERENCIAMENTO DE REQUISITOS EM UMA EMPRESA MULTINACIONAL¹

Andrea Chen Gao²

RESUMO

Este estudo investiga os desafios enfrentados por uma empresa multinacional no setor de desenvolvimento de produtos, focando na gestão da biblioteca de requisitos. Existe a necessidade de melhorar o processo de gestão de requisitos para evitar falhas e otimizar recursos. Com base teórica em Grandes Modelos de Linguagem (LLMs), uma aplicação do Processamento de Linguagem Natural (PLN) voltada para a análise de escrita dos requisitos, o estudo propõe uma metodologia composta por duas etapas: aplicação do Google Gemini com dados fictícios e aplicação do Google Gemini para avaliação de dados reais da empresa. O objetivo dessa aplicação é verificar a eficiência do Google Gemini em agrupar requisitos para auxiliar na redução de redundâncias na biblioteca de requisitos. Espera-se que a implementação dessa metodologia resulte em agrupamentos de requisitos que tornem a biblioteca de requisitos mais organizada e precisa, facilitando a entrega pontual e adequada dos produtos aos clientes, além de promover uma gestão mais eficaz dos recursos corporativos.

Palavras-chave: Engenharia de Requisitos; Grande Modelo de Linguagem (LLM); Google Gemini; Gerenciamento de requisitos.

ABSTRACT

This study investigates the challenges faced by a multinational company in the product development sector, focusing on the management of the requirements library. There is a need to enhance the requirements management process to prevent failures and optimize resources. Based on the theory of Large Language Models (LLMs), an application of Natural Language Processing (NLP) focused on the analysis of requirement writing, the study proposes a methodology consisting of two stages: the application of Google Gemini with fictitious data and the application of Google Gemini with real company data. The objective of this application is to evaluate the efficiency of Google Gemini in grouping requirements to assist in reducing redundancies in the requirements library. It is expected that the implementation of this methodology will result in groupings of requirements that make the requirements library more organized and accurate, facilitating the timely and proper delivery of products to clients, as well as promoting more effective management of corporate resources.

Keywords: Requirements Engineering; Large Language Models (LLMs); Google Gemini; Requirements Management.

¹ Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do grau de bacharel no Curso de Bacharelado em Ciência e Tecnologia, do Centro Tecnológico de Joinville (CTJ), da Universidade Federal de Santa Catarina (UFSC), sob orientação do Dr. Benjamin Grando Moreira.

² Graduanda do Bacharelado em Ciência e Tecnologia

1. INTRODUÇÃO

Em um mundo cada vez mais tecnológico e competitivo, as empresas de desenvolvimento de produto tendem a ter dificuldades em entregar o *produto certo no momento certo* ao cliente e, para obter isso, é necessário estabelecer os requisitos adequados para chegar ao resultado esperado. Os requisitos são a base de todo projeto, uma vez que relacionam o que os stakeholders³ (consumidores, negociadores, usuários, entre outros) querem para que o objetivo final seja atingido, além disso, direcionam as atividades do processo, portanto, quando não são definidas, qualquer imprevisto pode acontecer, inclusive a falha ou cancelamento do projeto (Dick *et al.*, 2017).

Projetos complexos podem ser desenvolvidos com o auxílio de engenharia de sistemas, que trabalham com o cotidiano, relacionam requisitos com a realidade, têm uma perspectiva de olhar o sistema como um todo. Um desses papéis é gerenciar requisitos, e a partir disso, tomar decisões antes mesmo que o ciclo de vida do projeto comece. Quando as consequências não são compreendidas desde o início do ciclo, a falta de gerenciamento pode resultar em complicações no final do sistema (INCOSE, 2006).

No ciclo de vida dos requisitos, existem as vozes dos stakeholders, conhecidos como requisitos de stakeholders, que representam papel fundamental no processo de desenvolvimento de produto uma vez que fornecem os pontos principais e diferenciais que o produto precisa ter. A falta de gerenciamento desses requisitos é tão arriscada quanto em qualquer outro tipo de requisito dentro de projetos (Dick *et al.*, 2017). Uma das negligências desse cuidado é a criação de requisitos sem gestão adequada, o que pode resultar em uma biblioteca de requisitos sem controle apropriado.

Neste contexto, o presente trabalho aborda um desafio enfrentado por uma empresa multinacional do setor de desenvolvimento de produtos: o gerenciamento da biblioteca de requisitos dos stakeholders. Um dos problemas identificados é a criação desorganizada de requisitos, muitos dos quais não seguem o padrão exigido pela empresa. Além disso, destaca-se a duplicação de requisitos, quando múltiplos requisitos atendem ao mesmo objetivo, resultando em uma ocupação desnecessária da biblioteca. Nesse cenário, o objetivo é propor uma metodologia que facilite o agrupamento e a análise de requisitos, permitindo identificar e eliminar duplicidades posteriormente, com base em uma pesquisa de estudo de caso.

O Processamento de Linguagem Natural (PLN) é uma área de pesquisa atrelada à inteligência artificial, que tem como foco o desenvolvimento de técnicas para que os computadores possam entender, interpretar e gerar uma linguagem humana de forma natural. O PLN possui diversos campos de atuação, como reconhecimento de fala, análise sintática, análise semântica, correção ortográfica ou gramatical, geração de texto e tradução automática (Caseli *et al.*, 2023).

Os Grandes Modelos de Linguagem (LLMs, sigla do inglês *Large Language Models*) representam um avanço significativo no campo do PLN. Segundo Paes *et al.* (2023), esses modelos se destacam por sua escala, contando com bilhões de parâmetros, o que os torna capazes de lidar com frases complexas, interpretar nuances contextuais e gerar respostas de forma fluida e coerente. Exemplos notáveis incluem o GPT-3 e o LaMDA, que têm contribuído para ampliar as possibilidades de interação humano-máquina e automatizar tarefas linguísticas de

³ Stakeholder pode ser traduzido livremente como *parte interessada*. Optou-se por manter o termo em inglês, pois é amplamente utilizado e reconhecido na área.

maneira mais eficaz. Com isso, LLMs podem auxiliar no processo de gerenciamento de requisitos através da busca de soluções para as atividades de agrupamento e análises dos mesmos dentro da biblioteca.

Este trabalho tem como objetivo avaliar a capacidade do Google Gemini em agrupar requisitos, facilitando a análise e a identificação de duplicidades na biblioteca de requisitos. A proposta é investigar como um modelo avançado de linguagem pode ser aplicado ao gerenciamento de requisitos, verificando sua eficácia na organização e estruturação das informações de forma eficiente. Busca-se, com isso, não apenas mitigar problemas existentes, mas também oferecer uma abordagem inovadora e tecnológica para os desafios do gerenciamento de requisitos em ambientes corporativos. Vale destacar que o escopo deste estudo não inclui o treinamento de um recurso de PLN ou o ajuste fino de um LLM. A pesquisa é fundamentada em um estudo de caso real, permitindo avaliar o impacto prático da solução proposta e contribuir para o avanço das práticas de engenharia de requisitos no desenvolvimento de produtos.

2. REFERENCIAL TEÓRICO

Esta seção aborda assuntos relacionados à problemática apresentada no trabalho, em que há relação entre o gerenciamento de requisitos e o LLM para auxiliar na atividade de agrupar requisitos.

2.1. Engenharia de Requisitos

Requisitos são responsáveis por motivar as atividades dos projetos de uma empresa. Através deles, é possível entender as necessidades não só do consumidor, mas também de outros stakeholders, como fornecedores, negócios, desenvolvedores, entre outros. A base dos requisitos são planejamento de requisitos, gerenciamento de riscos, trade-off, controle de mudanças e teste de aceitação. Pela sua relevância, os requisitos devem ser bem definidos desde o começo de qualquer projeto, para que a principal meta seja atingida. Sem uma definição precisa, os resultados podem ser totalmente diferentes dos esperados (Dick *et al.*, 2017).

Segundo a ISO/IEC 29148 (2018), a Engenharia de Requisitos é uma disciplina interdisciplinar que atua como ponte entre as necessidades do comprador e as soluções propostas pelo fornecedor, garantindo a definição e a manutenção dos requisitos que devem ser atendidos pelo sistema, software ou serviço em questão. Essa área tem como foco, o gerenciamento e o desenvolvimento dos requisitos, que envolve identificação, análise, especificação e verificação dos requisitos (Dick *et al.*, 2017).

O relatório CHAOS Report, produzido pelo Standish Group, identifica três fatores principais que contribuem para o fracasso de projetos: requisitos mal definidos, gerenciamento inadequado e questões políticas. Requisitos mal elaborados, incompletos ou desnecessários frequentemente levam a falhas, enquanto a falta de planejamento e uma gestão deficiente agrava o problema. Além disso, questões políticas funcionam como um catalisador, intensificando os impactos negativos dos dois primeiros fatores (Dick *et al.*, 2017).

Por outro lado, o sucesso dos projetos está associado à otimização de processos, ao uso eficiente de ferramentas e, sobretudo, à elaboração de uma declaração clara e bem estruturada dos requisitos. Esses elementos ressaltam a

importância de uma gestão cuidadosa e de requisitos de alta qualidade como pilares fundamentais para alcançar resultados esperados em projetos (Dick *et al.*, 2017).

Segundo Vazquez e Simões (2016), a abordagem tradicional da Engenharia de Requisitos, limitada a uma fase inicial com documentação extensa, é inadequada para projetos modernos devido à sua rigidez, dificultando adaptações e a colaboração com o cliente. Em contraste, as metodologias ágeis propõem uma visão flexível e interativa, com coleta, análise e validação de requisitos ocorrendo continuamente ao longo do ciclo de desenvolvimento. A documentação é simplificada e focada nas necessidades do cliente, promovendo maior colaboração, melhor comunicação e entrega de valor mais ágil.

A Engenharia de Requisitos é fundamental em diversas áreas, e, segundo Lopes (2024), ela se integra a outras disciplinas como a Engenharia de Software ao longo do ciclo de vida de um projeto de software. Esta última se concentra na aplicação de princípios científicos para transformar um problema em uma solução de software e garantir sua manutenção. Diferentes modelos de ciclo de vida, como o em cascata, em espiral e o Unified Process, abordam a Engenharia de Requisitos de maneiras distintas, mas todos convergem para a importância de uma análise detalhada e contínua dos requisitos para o sucesso do projeto.

2.2. Engenharia de Sistemas

Segundo Hitchin (2008), Engenharia de sistemas é uma abordagem que trata sistemas tecnológicos como entidades dinâmicas e complexas, interagindo com seu ambiente e com outros sistemas. Ao contrário das disciplinas de engenharia tradicionais, a engenharia de sistemas não se limita ao estudo de componentes isolados, mas sim à compreensão do comportamento global do sistema e de suas interações. Essa abordagem enfatiza a dinâmica, a adaptabilidade e as propriedades emergentes dos sistemas, ou seja, características que surgem da interação de suas partes e que não podem ser previstas pela análise de cada componente isoladamente.

Essa disciplina baseia-se no *pensamento sistêmico*, uma perspectiva que analisa os elementos de um sistema de forma integrada, identificando como suas partes se relacionam e contribuem para o objetivo geral. Ela também lida com questões técnicas, organizacionais e humanas, promovendo uma visão ampla para resolver problemas e projetos complexos. Entre os processos envolvidos estão a definição de requisitos, a modelagem e simulação para prever o comportamento do sistema, a análise de trade-offs para equilibrar soluções e a gestão de riscos ao longo do ciclo de vida (INCOSE, 2006).

A Engenharia de Sistemas é essencial em projetos de grande escala ou alta complexidade, onde a integração eficiente de diferentes disciplinas e tecnologias é necessária. Ela também desempenha um papel crucial na validação e verificação dos sistemas, garantindo que o produto final esteja alinhado com as expectativas e requisitos inicialmente estabelecidos (INCOSE, 2006).

A Engenharia de Sistemas e a Engenharia de Requisitos são disciplinas complementares que trabalham juntas para assegurar o sucesso de projetos, especialmente em ambientes complexos. A primeira fornece a estrutura geral para entender, projetar e gerenciar sistemas, enquanto a segunda garante que as necessidades e expectativas dos stakeholders sejam capturadas e traduzidas em soluções viáveis e efetivas. Ambas desempenham papéis cruciais no desenvolvimento de sistemas eficazes e alinhados aos objetivos.

2.3. Processamento de Linguagem Natural (PLN)

Processamento de Linguagem Natural (PLN) é uma tecnologia capaz de desenvolver programas computacionais que podem oferecer soluções para o ser humano. A palavra *natural* da sigla remete à linguagem falada por seres humanos, como linguagens de programação, matemática, visuais, entre outros. Através dessa área, é possível aplicar inúmeras funções, como tradução automática, sumarizador automático, sistemas de auxílio à escrita, corretor ortográfico ou gramatical, dicionários monolíngues ou bilíngues, entre outros (Caseli *et al.*, 2023).

O PLN possui duas áreas divididas em: *Natural Language Understanding* (NLU) e *Natural Language Generation* (NLG). O primeiro termo está relacionado a tudo que necessita de interpretação, mais especificamente, análise e entendimento da língua utilizada, ou seja, é a possibilidade da tecnologia de entender a frase que o usuário fornece, avaliando cada palavra, para saber o que executar após o pedido. Já o NLG, está atrelado à geração de linguagem humana, cujo objetivo é gerar uma linguagem natural para cumprir com o que foi solicitado (Caseli *et al.*, 2023).

Nesse contexto, há perguntas e respostas (PR, sigla do inglês *question answering*) que é uma área que estuda a criação de sistemas que possibilitam respostas diretas, apropriadas e precisas, além disso, é capaz de fornecer respostas de forma automática que respondem perguntas de linguagem humana. Esse sistema está diretamente relacionado ao PLN, uma vez que para que uma dúvida seja respondida, é necessária a compreensão da língua perguntada, e a geração de linguagem natural que é a resposta da pergunta. Os sistemas PR, pelo fato de serem projetados para responder de forma isolada, diferentemente dos chatbots que envolvem conversação, fornecem resultados informativos e rápidos, tornando o processo focado na eficiência (Cortes *et al.*, 2024).

No PLN existe a tradução automática (TA), em que qualquer língua pode ser traduzida de maneira adequada para outra totalmente diferente, para isso, existem diversas abordagens, das simples às complexas, como tradução direta, TA baseada em regras, TA neural, entre outras. Para avaliar a qualidade da tradução, existem as *métricas automáticas* (MAAS), que fornecem um sistema de avaliação com a finalidade de testar quão precisa foi a tradução, porém, apesar de ser eficiente é quantitativa e menos recomendada para verificar todas as opções de resposta em uma tradução. O ideal é, portanto, a intervenção humana juntamente com as MAAS para validar a TA e garantir sua confiabilidade (Castilho; Caseli, 2023).

Uma das aplicações do PLN é a correção automática de redação (CAR), em que há atribuição de nota a partir da leitura, reconhecendo e considerando possíveis erros dentro do texto. Esse sistema de correção abrange inúmeros estilos textuais (dissertação, carta, resenha, entre outros), tanto que pode ser amplamente utilizado para corrigir escritas de vestibulares (Unesp, Fuvest, Unicamp, entre outros). Toda correção tem critérios dependendo do estilo textual, porém, em todos os casos há avaliação da língua escrita, tema, gênero, coerência e coesão (Leal; Aluísio, 2024). Dessa forma, o CAR demonstra que o PLN tem inúmeras aplicações quanto à escrita, como a análise textual de requisitos, sendo confiável e eficiente.

2.4. Large Language Models (LLMs)

Os Large Language Models (LLMs), ou Grandes Modelos de Linguagem, são modelos de linguagem baseados em redes neurais que se destacam por sua escala e capacidades avançadas. Sua principal característica é a enorme quantidade de parâmetros que possuem, frequentemente ultrapassando 1 bilhão e podendo alcançar centenas de bilhões. Essa complexidade confere sua notável habilidade para processar e gerar texto com alta sofisticação, fluidez e coerência (Paes *et al.*, 2023).

No contexto do PLN, os LLMs representam um avanço significativo. Segundo Caseli *et al.* (2023), o PLN abrange uma ampla gama de tarefas, como tradução automática, correção gramatical e sistemas de auxílio à escrita. Para Paes *et al.* (2023), os LLMs se destacam nesse campo pela capacidade de combinar habilidades de compreensão e geração de texto, integrando as subáreas NLU e NLG, sendo que essa integração permite que os LLMs analisem frases complexas fornecidas pelos usuários, compreendam seus significados e gerem respostas contextualizadas, atendendo às demandas de forma eficiente e precisa.

Além disso, os LLMs estão enquadrados na categoria de Inteligência Artificial Generativa. Sua principal função é a geração de conteúdo textual, podendo ser aplicados em tarefas que vão desde a escrita criativa até a resolução de problemas complexos. Uma característica notável desses modelos é a presença de habilidades emergentes, que surgem à medida que os modelos aumentam em escala. Entre essas habilidades, destaca-se o aprendizado em contexto (em inglês, *in-context learning*), que permite aos LLMs resolver tarefas apenas a partir de instruções ou exemplos fornecidos em linguagem natural, sem a necessidade de ajustes adicionais nos parâmetros do modelo. Outra habilidade emergente relevante é a estratégia de cadeia de pensamento (em inglês, *chain-of-thought*), que possibilita aos modelos resolver problemas em múltiplos passos ao seguir prompts que incentivam raciocínios estruturados (Paes *et al.*, 2023).

O funcionamento dos LLMs é fortemente baseado em *prompts*, que direcionam as saídas geradas pelo modelo. Esses *prompts* podem variar em formato e complexidade, sendo ajustados conforme a tarefa. Essa dependência torna os LLMs ferramentas versáteis e adaptáveis a uma ampla gama de aplicações, como análise de sentimentos, resolução de problemas matemáticos ou inferências complexas. No entanto, as respostas desses modelos são probabilísticas, o que significa que a mesma entrada pode gerar diferentes saídas em execuções distintas (Paes *et al.*, 2023).

Exemplos de LLMs incluem o GPT-3, o PaLM, o LaMDA, o LLaMA e o modelo Sabiá, este último desenvolvido especificamente para a língua portuguesa. Esses modelos são frequentemente usados como base para agentes de conversação, como o ChatGPT (alimentado pelo GPT) e o Bard (baseado no LaMDA). Eles não só ilustram a evolução tecnológica na área de PLN, como também representam um marco importante no desenvolvimento da inteligência artificial generativa, sendo amplamente utilizados em diversas aplicações práticas (Paes *et al.*, 2023).

Neste trabalho, utilizou-se o Google Gemini, reconhecido por suas capacidades multimodais, que incluem a compreensão e geração de texto. Com versões que vão desde a eficiente Nano, ideal para dispositivos móveis, até a Ultra, destinada a tarefas altamente complexas, o Gemini apresenta uma janela de contexto ampla, alcançando até 2 milhões de tokens na versão 1.5 Pro (Carraro,

2024). Essas características tornam o modelo ideal para aplicações que demandam precisão e flexibilidade, posicionando-o como uma ferramenta revolucionária para desafios como a organização e análise de requisitos .

2.5. Trabalhos Similares

A pesquisa de Krishna *et al.* (2024) investiga a aplicação de LLMs, como o GPT-4 e o CodeLlama, na geração e validação de documentos de Especificação de Requisitos de Software (SRS). Tradicionalmente, a elaboração de um SRS é uma tarefa que exige grande esforço humano, sendo essencial para o sucesso de projetos de software, mas também demorada e propensa a erros. A pesquisa propõe a automação de partes desse processo com o uso dos LLMs, com o objetivo de reduzir o tempo e o esforço necessários, além de melhorar a qualidade dos documentos. O estudo também mediu o tempo economizado na geração do SRS pelos LLMs em comparação ao tempo gasto por um engenheiro novato, demonstrando o potencial desses modelos para otimizar o processo e melhorar a qualidade da documentação.

Em outro estudo, Chen *et al.* (2024) propõe uma abordagem para melhorar a precisão e a confiabilidade de LLMs (ChatGLM2-6B) em cenários clínicos, como o diagnóstico de amiloidose, por meio da combinação de *engenharia de conhecimento* e *engenharia de qualidade*. A metodologia envolve integrar informações especializadas sobre a doença a um grafo de conhecimento, utilizando verificações de similaridade para garantir que o conteúdo gerado pelo LLM seja consistente com esse conhecimento. Além disso, implementa um controle de qualidade baseado na probabilidade de tokens gerados, selecionando as respostas mais confiáveis.

O estudo realizado por Fran *et al.* (2024) explora o uso de LLMs para geração de código e testes automatizados de software, abordando desafios como a dificuldade em avaliar a correção do código gerado e garantir sua funcionalidade. Ele propõe novas abordagens para melhorar a avaliação da funcionalidade do código, integrando geração de testes automatizados e otimizando explicações fornecidas pelos LLMs. Também investiga métodos híbridos que combinam LLMs com técnicas tradicionais de teste, como fuzzing, para aumentar a cobertura de código e a detecção de falhas, visando aprimorar a eficácia e precisão dessas ferramentas na engenharia de software.

Embora os três estudos mencionados envolvam o uso de LLMs, eles adotam abordagens distintas: o primeiro foca na geração e validação de SRS, o segundo aplica LLMs em contextos clínicos, como diagnóstico de amiloidose, e o terceiro no desenvolvimento de código e testes automatizados de software. Nenhum desses trabalhos aborda diretamente a aplicação de LLMs na engenharia de requisitos, o que torna o foco deste trabalho um diferencial.

3. METODOLOGIA PROPOSTA PARA UTILIZAR LLM NO GERENCIAMENTO DE REQUISITOS

Este trabalho consiste em um estudo de caso realizado em parceria com uma empresa multinacional do setor de produção de eletrodomésticos. A escolha da empresa foi motivada pela sua relevância no mercado e pela necessidade de melhorias em sua biblioteca de requisitos.

Todos os requisitos são armazenados em uma plataforma conhecida como biblioteca de requisitos, que reúne todas as informações necessárias sobre eles. Essa ferramenta mantém um histórico dos requisitos utilizados em cada projeto realizado pela empresa, o que a torna altamente relevante. É por meio dela que os engenheiros de sistemas realizam o gerenciamento de requisitos. No entanto, quanto maior e mais complexa for a biblioteca, mais desafiador se torna seu gerenciamento.

O objetivo deste estudo é trabalhar com os requisitos da empresa para otimizar o trabalho da equipe de engenharia de sistemas, reduzindo o volume da biblioteca, que apresenta problemas de duplicação e necessita ser simplificada. A aplicação do LLM será utilizada para otimizar esse processo e, ao final, contribuir para o aprimoramento do gerenciamento de requisitos da empresa. A organização forneceu acesso a dados operacionais, que foram analisados para o desenvolvimento do estudo.

A empresa lida com diversas vozes de stakeholders, mas este estudo tem como foco os consumidores. Segundo Griffin e Hauser (1993), os requisitos da voz do consumidor (VOC, sigla em inglês para *voice of the customer*) enfoca as necessidades e expectativas dos consumidores em relação aos produtos, buscando capturar suas demandas de forma estruturada. Em essência, o VOC traduz essas demandas em requisitos que são gerenciados pelos engenheiros de sistemas, permitindo a criação de soluções que atendam de forma eficaz às expectativas do mercado.

Essa voz se baseia em técnicas de pesquisa qualitativa, como entrevistas individuais ou visitas ao local de uso do produto, para entender a experiência do cliente. Essa abordagem envolve a coleta de feedback detalhado que é, posteriormente, estruturado em hierarquias de necessidades - primárias, secundárias e terciárias. Essas categorias ajudam a alinhar estratégias de marketing, desenvolvimento de produto e outros esforços para atender às demandas do mercado (Griffin; Hauser, 1993).

O LLM selecionado para o estudo foi o Gemini, devido à sua disponibilidade na empresa, sendo essa ferramenta a única disponibilizada. Além dessa ferramenta, foram consideradas outras opções, como LLM Studio, Chat GPT e Copilot, mas nenhuma delas possui autorização para uso no ambiente de trabalho da empresa. Como o estudo envolve dados reais da organização, optou-se por utilizar uma ferramenta que permitisse o uso desses dados com segurança e confiabilidade.

A Figura 1 mostra um diagrama com a metodologia empregada no trabalho, sendo que essa consiste em 2 etapas com 3 sub etapas cada:

- 1) Aplicação do Gemini com dados fictícios: elaboração dos dados; elaboração e testes de prompts; avaliação de resultados.
- 2) Aplicação do Gemini com dados reais: coleta de dados; elaboração e testes de prompts; avaliação de agrupamentos dos requisitos.

A primeira etapa envolve a aplicação do LLM com dados fictícios, com o objetivo de avaliar como o Gemini responde às solicitações feitas. Na primeira sub etapa, foi realizada a elaboração dos dados, que envolveu uma pesquisa com indivíduos sem conhecimento prévio sobre requisitos. A segunda sub etapa consistiu nos testes dos prompts, visando coletar informações a serem processadas pelo Gemini. A última sub etapa envolveu a avaliação das respostas fornecidas pelo LLM, considerando os prompts e a avaliação de especialista da área.

Figura 1 - Diagrama da metodologia com suas duas etapas e sub etapas



Fonte: Autora (2024).

Após a compreensão do funcionamento do LLM, a segunda etapa foi a avaliação do Gemini com dados reais. A primeira sub etapa foi a coleta de dados da empresa, focando no VOC, onde os dados foram filtrados para determinar os mais relevantes para o estudo. Em seguida, realizaram-se os testes dos prompts, com o objetivo de analisar as diferentes respostas geradas conforme o prompt utilizado.

Por último, a terceira sub etapa consistiu na comparação dos agrupamentos de requisitos definidos pelos engenheiros de sistemas com aqueles fornecidos pelo Gemini. A análise dos resultados será feita tanto de forma qualitativa, levando em consideração os critérios (*aprovados, consideráveis e não aprovados*), quanto quantitativa, com a obtenção de um valor numérico representando a porcentagem de cumprimento de cada critério na análise.

4. DESENVOLVIMENTO

Nesta seção, é mostrada a aplicação do PLN com dois diferentes conjuntos de dados: dados fictícios e dados reais. Abordagens distintas foram utilizadas para cada conjunto de dados. Em ambos os experimentos, o resultado esperado é o mesmo: dado uma quantidade de requisitos que podem conter duplicidade entre si, o Gemini deve retornar uma resposta que indique quais requisitos são duplicados/similares para que possam ser agrupados, e posteriormente, excluídos da base de dados.

4.1. Aplicação do Gemini com dados fictícios

Os dados fictícios foram criados através de uma pesquisa para coletar informações de diferentes perspectivas do consumidor em relação à compra e o uso de fogão. As perguntas para a coleta foram: quando você compra um fogão, quais requisitos são levados em conta para comprá-lo? O que importa para você quando se tem esse eletrodoméstico para valer a pena tê-lo?

O público geral entrevistado foi de faixa etária entre 20 a 40 anos, totalizando 15 pessoas que contribuíram para a pesquisa. Os entrevistados não têm vínculo com a empresa, portanto, nenhum deles possui conhecimento formal sobre requisitos. As respostas foram obtidas em conversas informais e, quando necessário, a autora as reescreveu para garantir clareza, porém, sem alterar o sentido dos requisitos. Em alguns casos, foi preciso solicitar que os participantes reformulassem suas respostas para maior precisão. Além disso, como este trabalho foca nos requisitos voltados para consumidor, o objetivo foi coletar informações a partir da perspectiva do consumidor, tanto como comprador quanto como usuário.

Ao final, os dados fictícios foram de 30 requisitos, como mostra o Quadro 1. Em seguida, todos os requisitos de mesmo tópico foram agrupados e separados para se comparar os resultados com o teste no Gemini. Esse experimento foi apenas para entender como a ferramenta funciona considerando o prompt e a forma como o LLM trabalhou com os dados fornecidos.

Quadro 1 - Requisitos coletados para experimento fictício

Identificação	Descrição do requisito
1	Quero um fogão com, no mínimo, 2 bocas.
2	O fogão deve ser fácil de limpar.
3	Quero que eu possa cozinhar mais de 3 comidas diferentes ao mesmo tempo.
4	Todas as minhas panelas devem ocupar as bocas do fogão sem que uma bata na outra ou interfira no cozimento das comidas.
5	O fogo deve acender rapidamente.
6	O fogão deve durar mais de 4 anos.
7	Meu produto não pode vir danificado.
8	Meu fogão deve ter uma interface digital bonita.

Fonte: Autora (2024).

Percebe-se que, devido à ausência de uma formulação padrão para os requisitos e ao fato de não ter sido redigido por especialistas, cada requisito possui uma estrutura diferente. Alguns começam com verbos no presente, como *quero*, enquanto outros utilizam o futuro do pretérito, como *gostaria*, e também há requisitos que não iniciam com verbos. Esse fato evidencia que, por não terem sido escritos com um direcionamento claro e por diferentes pessoas, esses requisitos não possuem um padrão consistente, similar aos requisitos criados para a biblioteca da

empresa. Isso dificulta o agrupamento de requisitos semelhantes, uma vez que pode haver mais de uma interpretação para o mesmo requisito.

4.2. Aplicação do Gemini com dados reais da empresa

A empresa disponibilizou 27 requisitos extraídos da biblioteca de requisitos, sendo todos classificados no mesmo atributo e mantidos em sua forma original, ou seja, sem modificar sua escrita. A companhia trabalha com mais de um atributo, porém, no presente trabalho, foi focado em apenas um, o de Estética. Vale ressaltar que nem todos os requisitos foram escritos pela mesma pessoa, momento e padrão. A sua criação desordenada foi o principal fator para a necessidade deste trabalho, uma vez que atualmente, há muitos requisitos escritos de maneiras distintas e em diferentes momentos, mas que transmitem a mesma mensagem. Isso gera uma ocupação desnecessária na biblioteca de requisitos da empresa, o que causa uma gerência complexa para os engenheiros de sistemas.

Todos os 27 requisitos foram avaliados e classificados pelo time de engenharia de sistemas. A Figura 2 ilustra as etapas realizadas para determinar quais requisitos seriam mantidos ou excluídos da biblioteca da empresa. A primeira etapa envolve a leitura de todos os requisitos de forma individual, com atenção ao conteúdo de cada um. Na sequência, é feita uma análise para identificar os requisitos comuns entre si, com o intuito de sugerir possíveis agrupamentos. A terceira etapa consiste no alinhamento com os stakeholders, a fim de esclarecer dúvidas e validar os agrupamentos propostos. Na quarta etapa, realiza-se uma análise detalhada de cada grupo para decidir quais requisitos devem ser mantidos ou excluídos da biblioteca. Por fim, a última etapa é a validação dos requisitos selecionados, realizada por meio de uma reunião da equipe de engenharia de sistemas.

Como as três últimas etapas dependem de fatores externos, como conversas com outros times e validações por profissionais, este trabalho concentra-se na execução das etapas 1 e 2, etapas manuais, que têm como objetivo identificar requisitos com temas semelhantes e agrupá-los, facilitando a análise pela equipe de engenharia.

Figura 2 - Etapas para definir conjunto de requisitos



Fonte: Autora (2024).

Por questões de confidencialidade em relação aos dados da empresa, nenhuma informação será divulgada ou exibida. O trabalho apresentará os requisitos com uma numeração atribuída a cada um.

4.2.1. Avaliação do prompt

Antes de chegar na etapa de avaliação do resultado do Gemini, foi necessário encontrar um prompt adequado que o LLM entendesse o que foi pedido e retornar uma resposta justa à solicitação. Os primeiros prompts criados para o Gemini foram mais genéricos, sem a utilização da palavra *requisito* e focando em *duplicação*.

Um exemplo de prompt desse experimento foi: *Estou verificando algumas frases e preciso que você encontre aquelas que são duplicadas. Se você achar algum duplicado, me fale o número da frase que está duplicado. Verifique na seguinte lista. Ao final do texto, foi enviado a lista contendo todos os requisitos de ordem aleatória. A resposta do Gemini foi: Para encontrar as frases duplicadas, vou comparar cada frase com todas as outras frases da lista. Se encontrar frases que têm o mesmo significado, vou te dizer qual o número da frase em que o significado está duplicado. Vou tentar usar outra biblioteca para verificar as frases duplicadas. Não consegui responder ao seu comando. Reformule a pergunta e tente de novo.* Essa resposta revela que o Gemini não compreendeu completamente o que foi solicitado.

Foram necessárias várias tentativas até encontrar um prompt utilizável, como este: *Estou trabalhando com redução de requisitos focado em produtos como fogões, fornos, cooktops, microondas e airfryers para uma indústria de eletrodomésticos. A problemática é a quantidade de requisitos duplicados/similares que ocupam muito espaço na biblioteca. Tenho aqui uma amostra de 30 requisitos misturados, me fale quais deles são duplicados ou similares para que possamos diminuir a quantidade total de requisitos.* Neste caso, a resposta do Gemini foi bem-sucedida: ele forneceu os agrupamentos, atribuiu nomes a cada um e indicou os números que identificavam os requisitos similares.

A abordagem para a criação desse prompt foi detalhada e específica, considerando o tipo de indústria, o produto em questão, a problemática e a evidência de que se tratam de requisitos e não de frases comuns. Essa estratégia proporcionou um direcionamento claro para o Gemini, minimizando ambiguidades e garantindo respostas mais precisas.

Além disso, vale ressaltar que todos os requisitos fornecidos pela empresa estão em inglês, enquanto os prompts foram formulados em português. Espera-se que essa diferença de idioma não afete a qualidade das respostas, pois no PLN é utilizada a Tradução Automática (TA). Segundo Castilho e Caseli (2023), embora possa haver variações entre os idiomas, o Gemini é capaz de identificar a língua, realizar a tradução e fornecer respostas de alta qualidade no idioma solicitado.

4.2.2. Agrupamentos de requisitos pelos Engenheiros de Sistemas

Todos os requisitos foram descarregados no Google Planilha para poder manipulá-los, já que não havia uma ordem. Depois da etapa 1 da Figura 2, os requisitos foram deslocados de linhas conforme a sugestão do grupo, como mostra o Quadro 2.

Após as 4 primeiras etapas vistas na Figura 2, o time de engenharia de sistemas definiu que os 27 requisitos foram divididos em 6 grupos. Para não divulgar dados reais da empresa, os grupos foram denominados como A, B, C, D, E e F, sendo o último, requisitos que não possuem correlação com nenhum outro da lista, ou seja, requisitos únicos.

Quadro 2 - Exemplo de um agrupamento

Agrupamento	Identificação do requisito
Grupo A	Requisito 1
	Requisito 17
	Requisito 23
	Requisito 14

Fonte: Autora (2024).

Além disso, os requisitos são identificados com números de 1 a 27. Os requisitos de número 1, 17, 23 e 14 tratam do mesmo tema, então, fazem parte do mesmo grupo denominado grupo A, e assim por diante, conforme ilustra o Quadro 3.

Quadro 3 - Resultado final de agrupamento do time

Agrupamento	Requisitos similares
Grupo A	1, 17, 23 e 14
Grupo B	2, 4 e 6
Grupo C	7, 10, 13, 18, 19, 20, 24, 25 e 27
Grupo D	22 e 11
Grupo E	5, 8, 9 e 12
Grupo F	3, 15, 16, 21 e 26

Fonte: Autora (2024).

Os requisitos 2, 4, 6, 11 e 22 foram considerados como mal classificados, pois segundo o time de engenharia de sistemas, não pertencem ao atributo de Estética, portanto, tiveram que ser reclassificados para serem avaliados em outro momento. Como o foco deste trabalho é verificar o desempenho dos agrupamentos do Gemini, este fator não foi considerado na avaliação da ferramenta.

4.2.2.1. Exemplo com utilização dos dados fictícios

O Quadro 6 exibe o agrupamento dos requisitos similares, baseados nos dados fictícios apresentados anteriormente (seção 4.1, Quadro 1). Esse agrupamento foi realizado por uma engenheira de sistemas da empresa. Por exemplo, os requisitos 4, 11, 21, 22, 23 e 27 foram agrupados por tratarem do mesmo tema: a capacidade do fogão/forno.

Este caso demonstra que, embora os requisitos não possuam as mesmas palavras-chave exatas, é possível agrupá-los pela semelhança de significado. O requisito 21, por exemplo, menciona que o cliente deseja um produto com *grande capacidade*, mas a expressão *grande capacidade* é vaga e subjetiva. O conceito de

grande pode variar conforme a percepção do consumidor: para alguns, um espaço grande seria o suficiente para uma panela estilo wok (panela tradicional chinesa) de 35 cm, enquanto outros podem considerar grande o espaço apenas para uma frigideira de 26 a 28 cm. Como o requisito foi formulado de maneira ampla, ele foi agrupado com o requisito 23, que também trata de espaço para panelas, embora de forma diferente, e assim por diante.

Quadro 6 - Agrupamento dos dados fictícios

Agrupamento	Requisitos Similares
Grupo A: Capacidade do produto	Requisito 4: Todas as minhas panelas devem ocupar as bocas do fogão sem que uma bata na outra ou interfira no cozimento das comidas.
	Requisito 11: Meu fogão deve ser grande o suficiente para caber qualquer tipo de panela.
	Requisito 21: Quero meu produto com capacidade grande.
	Requisito 22: Quero que o interior do meu produto seja grande o suficiente para caber mais de uma travessa de comida.
	Requisito 23: Gostaria que meu fogão coubesse o máximo de panelas possível.
	Requisito 27: Quero um produto de grande capacidade.

Fonte: Autora (2024).

Os seis requisitos poderiam ser consolidados em um único requisito, bastando reescrevê-los de maneira que englobasse todos os aspectos necessários. No entanto, também seria possível optar por manter dois requisitos, dependendo da estratégia da empresa e do uso previsto em projetos futuros. Contudo, essa decisão não é simples, pois é importante considerar diversos fatores, como a análise de sua aplicação em projetos anteriores ou o que outros stakeholders pensam disso, por exemplo. Cabe à equipe de engenharia de sistemas avaliar os critérios da empresa e sua estratégia para determinar a abordagem mais adequada.

4.2.3. Avaliação do agrupamento de requisitos

O Gemini forneceu os grupos de acordo com os requisitos similares e nomeou cada um deles, diferentemente do resultado do agrupamento validado pelo time de engenharia. O LLM retornou 7 grupos em vez de 6, conforme mostra o Quadro 5.

O Quadro 6 apresenta uma comparação entre os requisitos considerados similares tanto pelos engenheiros quanto pelo Gemini. Embora muitos requisitos tenham sido classificados como similares, apenas um grupo - formado pelos grupos B e 3 - foi 100% idêntico, com os requisitos 2, 4 e 6 agrupados. Por exemplo, os grupos A e 1 compartilham os requisitos 1, 17 e 23, assim como os grupos E e 2,

que têm em comum os requisitos 5, 8 e 9. No entanto, o requisito de número 12 não foi incluído na resposta do Gemini.

Quadro 5 - Resposta de agrupamento do Gemini

Agrupamento sugerido	Nome sugerido aos grupos	Requisito similares
Grupo 1	Alinhamento estético geral	1, 3, 17, 23 e 27
Grupo 2	Personalização e intercambiabilidade	5, 8 e 9
Grupo 3	Alinhamento estético entre aparelhos	2, 4 e 6
Grupo 4	Integração com a cozinha	7, 18, 24, 25 e 20
Grupo 5	Qualidade da aparência	22, 11 e 19
Grupo 6	Requisitos específicos de design	10, 12 e 13
Grupo 7	Requisitos isolados	14, 15, 16, 21 e 26

Fonte: Autora (2024).

Devido a essa correspondência parcial e a diferença no número de grupos entre os dois resultados, a abordagem para avaliar a aplicação do LLM foi realizada por meio de uma análise individual dos requisitos, verificando se faz sentido que cada requisito esteja no grupo correspondente e se os requisitos agrupados com ele são os mesmos determinados pelos engenheiros da empresa.

O Quadro 7 mostra os agrupamentos sugeridos pelo Gemini, vistos nas duas primeiras colunas. Por exemplo, os requisitos 1, 3, 17, 23 e 27 foram classificados no grupo 1, assim como os requisitos 5, 8 e 9 foram juntados ao grupo 2, e assim por diante. A terceira coluna se refere à avaliação da resposta do LLM em comparação com a análise feita pelos engenheiros de sistemas.

Os três critérios utilizados para essa avaliação são: *aprovado*, *não aprovado* e *considerável*. Um requisito é classificado como *aprovado* quando se encaixa no nome do grupo ao qual foi atribuído e compartilha um tema em comum com outros requisitos. Por exemplo, os requisitos 1, 17 e 23 foram alocados no grupo A pelos engenheiros, e o Gemini também os agrupou dessa forma, mas adicionou dois requisitos adicionais - 3 e 27 - que não correspondem ao tema do grupo. Já os requisitos são classificados como *não aprovado* quando não atendem a nenhum dos critérios mencionados para a categoria *aprovado*, como ocorreu com os requisitos 3 e 27.

O critério *considerável* se aplica a situações em que até mesmo os engenheiros tiveram dificuldade para definir ou entender o que o requisito realmente gostaria de transmitir, seja por estar mal redigido ou por poder se encaixar em mais de um grupo simultaneamente, como ocorreu com os requisitos 10, 13 e 14. Alguns requisitos são difíceis de entender à primeira vista, especialmente quando não seguem o padrão usual de escrita. Além disso, como foram escritos por diferentes pessoas, muitos requisitos acabam sendo ambíguos, gerando interpretações diversas até mesmo entre os próprios engenheiros.

Quadro 6 - Agrupamentos: Equipe de engenharia x Gemini

Gemini		Equipe de engenharia		Requisitos agrupados iguais
Grupo 1	1, 3, 17, 23 e 27	Grupo A	1, 17, 23 e 14	1, 17 e 23
Grupo 2	5, 8 e 9	Grupo B	2, 4 e 6	2, 4 e 6
Grupo 3	2, 4 e 6	Grupo C	7, 10, 13, 18, 19, 20, 24, 25 e 27	7, 18, 24, 20 e 25
Grupo 4	7, 18, 24, 25 e 20	Grupo D	22 e 11	22 e 11
Grupo 5	22, 11 e 19	Grupo E	5, 8, 9 e 12	5, 8 e 9
Grupo 6	10, 12 e 13	Grupo F	3, 15, 16, 21 e 26	15, 16, 21 e 26
Grupo 7	14, 15, 16, 21 e 26	-	-	-

Fonte: Autora (2024).

A última coluna do Quadro 7 teve um propósito diferente, sendo utilizada para verificar a resposta do Gemini e avaliar o quão influenciável ele poderia ser. Para cada requisito que ele classificou, foi feita a mesma pergunta: o motivo pelo qual o requisito foi agrupado em determinado grupo. Essa abordagem foi aplicada a todos os requisitos, independentemente da classificação - *aprovado, não aprovado ou considerável*.

Por exemplo, a primeira pergunta foi: *Por que o requisito 1 foi alocado no grupo 1?* A resposta do Gemini forneceu o motivo pelo qual o requisito 1 foi classificado neste grupo. No caso do requisito 3, ele também justificou sua alocação no grupo 1. No entanto, o esperado era que o Gemini corrigisse essa classificação, indicando que o requisito 3 deveria, na verdade, estar no grupo 7, pois se tratava de um requisito isolado. Por esse motivo, o requisito 1 obteve a classificação *aprovado*, e o de número 3, *não aprovado*, na segunda avaliação.

O requisito 28 não foi aprovado inicialmente, mas, na segunda avaliação, ao ser questionado sobre o motivo de sua classificação, a resposta do Gemini (adaptada pela autora por questão de confidencialidade) foi: *Este requisito é bem abrangente e pode gerar dúvidas sobre seu posicionamento. A princípio, ele parece similar ao requisito 18, que alocamos no Grupo 5. No entanto, o requisito 28 é mais genérico devido a uma expressão que não foi totalmente clara na escrita*. Essa resposta foi apresentada aos especialistas, que concordaram que o significado do requisito poderia ser confundido com os requisitos do grupo 1. Por essa razão, o requisito foi classificado como *considerável*.

Quadro 7 - Avaliação das respostas do Gemini

Agrupamento	Identificação	Avaliação	Reavaliação
Grupo 1	Requisito 1	Aprovado	Aprovado
	Requisito 3	Não aprovado	Não aprovado
	Requisito 17	Aprovado	Aprovado
	Requisito 23	Aprovado	Aprovado
	Requisito 28	Não aprovado	Considerável
Grupo 2	Requisito 5	Aprovado	Aprovado
	Requisito 8	Aprovado	Aprovado
	Requisito 9	Aprovado	Aprovado
Grupo 3	Requisito 2	Aprovado	Aprovado
	Requisito 4	Aprovado	Aprovado
	Requisito 6	Aprovado	Aprovado
Grupo 4	Requisito 7	Aprovado	Aprovado
	Requisito 18	Aprovado	Aprovado
	Requisito 24	Aprovado	Aprovado
	Requisito 25	Aprovado	Aprovado
	Requisito 20	Aprovado	Aprovado
Grupo 5	Requisito 22	Aprovado	Aprovado
	Requisito 11	Aprovado	Aprovado
	Requisito 19	Não aprovado	Aprovado
Grupo 6	Requisito 10	Considerável	Aprovado
	Requisito 12	Não aprovado	Aprovado
	Requisito 13	Considerável	Aprovado
Grupo 7	Requisito 14	Considerável	Aprovado
	Requisito 15	Aprovado	Aprovado
	Requisito 16	Aprovado	Aprovado
	Requisito 21	Aprovado	Aprovado
	Requisito 26	Aprovado	Aprovado

Fonte: Autora (2024).

Os requisitos *não aprovados* 12 e 19 foram aceitos na segunda avaliação, quando o Gemini reconheceu o erro e explicou que o requisito 19 deveria ter sido alocado no grupo 4, em vez do grupo 5. O mesmo ocorreu com o requisito 12, quando o LLM pediu desculpas e afirmou que ele se encaixava melhor no grupo 2, e não no grupo 6. A mesma lógica se aplicou aos requisitos 10, 13 e 14, que, inicialmente, foram classificados como *consideráveis*, mas, na segunda análise, foram reclassificados e *aprovados* ao final.

5. RESULTADOS

Os resultados são apresentados de forma separada, seguindo as abordagens utilizadas na obtenção dos dados. Inicialmente, são exibidos os resultados da aplicação do Gemini com dados fictícios. Em seguida, apresentam-se os resultados obtidos com dados reais da empresa.

5.1. Experimento no Gemini com dados fictícios

Neste experimento, o critério para avaliar a resposta do Gemini foi comparar os requisitos agrupados pela autora e pelo LLM. Dos 30 requisitos avaliados, 13 podem ser considerados classificados corretamente pelo Gemini, ou seja, menos da metade do resultado condiz com o que era esperado. O fato desse valor ter sido baixo se deve pela construção de requisitos generalistas, que foram criados por pessoas sem conhecimento prévio acerca de requisitos. Além disso, o experimento teve o intuito de conhecer a maneira como o LLM trabalha, então a expectativa não era o resultado quantitativo, mas o entendimento de como o Gemini respondeu ao prompt e construiu suas respostas.

5.2. Experimento no Gemini com dados reais

A Tabela 1 revela que 74,07% dos requisitos avaliados obtiveram resultado positivo e foram *aprovados*, enquanto 11,12% foram considerados como *consideráveis* e 14,81% foram *não aprovados*. Isso indica que mais da metade dos requisitos foi classificada de forma consistente com a avaliação dos engenheiros. Na análise da segunda avaliação, apresentada na Tabela 2, onde ocorreu um teste de reafirmação das respostas do Gemini, os resultados mostram-se ainda mais precisos: 92,60% dos requisitos foram *aprovados*, enquanto apenas 3,7% (equivalente a um único requisito para cada classificação) foram classificados como *consideráveis* e *não aprovados*.

Vale destacar que o processo de agrupamento dos requisitos pelos engenheiros da empresa multinacional demandou cerca de uma semana. Segundo os próprios engenheiros, muitos requisitos careciam de informações ou eram ambíguos, dificultando sua interpretação. Foram necessárias diversas discussões e alinhamentos para compreender cada requisito, permitindo agrupá-los e decidir quais seriam mantidos ou excluídos da biblioteca de requisitos. Esse cuidado foi essencial, considerando que a biblioteca seria uma ferramenta frequente para futuros projetos da equipe.

Tabela 1 - Resultado quantitativo da primeira avaliação

Avaliação	Quantidade de cada classificação	Porcentagem em relação ao total
Aprovados	20	74.07%
Consideráveis	3	11.12%
Não aprovados	4	14.81%
Total	27	100,00%

Fonte: Autora (2024).

Tabela 2 - Resultado quantitativo da reavaliação

Avaliação	Quantidade de cada classificação	Porcentagem em relação ao total
Aprovados	25	92.60%
Consideráveis	1	3.70%
Não aprovados	1	3.70%
Total	27	100,00%

Fonte: Autora (2024).

Por outro lado, com a utilização de LLM, o tempo necessário para realizar o mesmo agrupamento foi reduzido para alguns minutos. Com prompts bem elaborados e direcionados, o tempo de execução pode ser ainda mais otimizado. Os resultados apresentados nas Tabelas 1 e 2 reforçam a eficácia do uso do Gemini para o agrupamento de requisitos, demonstrando seu potencial como uma solução eficiente para auxiliar na redução e no gerenciamento de requisitos em projetos futuros da empresa.

6. CONCLUSÃO

A análise da aplicação do LLM para o agrupamento de requisitos demonstrou que, embora o modelo Gemini tenha apresentado alguns desafios na primeira avaliação, ele foi capaz de fornecer resultados consistentes e valiosos na reavaliação, alcançando uma taxa de aprovação de 92,60% dos requisitos avaliados. Esse desempenho é significativo, considerando que o agrupamento manual realizado pelos engenheiros levou cerca de uma semana, enquanto o processo com o LLM foi concluído em alguns minutos. Além disso, o uso do Gemini ajudou a acelerar a tomada de decisões, oferecendo uma ferramenta útil para otimizar o gerenciamento de requisitos em projetos futuros. Apesar da diferença de língua entre o prompt em português e os requisitos em inglês, a eficácia do modelo não foi comprometida.

Entretanto, a análise também revelou algumas limitações do modelo, como a alocação incorreta de requisitos em alguns grupos, o que foi ajustado após a segunda avaliação. Esses ajustes destacam a importância de uma interação contínua entre o modelo de PLN e especialistas humanos para garantir a precisão e a relevância dos agrupamentos. A flexibilidade do Gemini em revisar suas respostas, após receber feedback, também evidenciou o potencial de evolução dos LLMs para tarefas mais complexas e específicas, como o agrupamento de requisitos em ambientes de engenharia.

Além disso, destaca-se também como resultado do trabalho a elaboração de uma base de dados fictícia que pode ser utilizada por outros trabalhos. Embora os requisitos não sejam reais em relação aos de uma empresa do setor, eles foram obtidos a partir de entrevistas com possíveis consumidores. Dessa forma, a base oferece um cenário inicial para experimentos futuros, seja no aprimoramento do LLM ou no desenvolvimento de novas técnicas de gestão de requisitos.

Por fim, o estudo mostrou que o Gemini tem um grande potencial como ferramenta de apoio na análise e organização de requisitos, especialmente em contextos onde a agilidade e a redução de tempo são essenciais. No entanto, para maximizar sua eficiência, é necessário aprimorar o treinamento do modelo e sua capacidade de lidar com a complexidade e ambiguidade dos requisitos, além de considerar o auxílio contínuo de profissionais qualificados para garantir a aplicabilidade dos resultados.

Para trabalhos futuros, é possível explorar a solicitação ao Gemini de gerar palavras-chave para cada requisito, aprimorar sua capacidade de reescrever requisitos para seguir um padrão específico, além de desenvolver a habilidade do LLM para reconhecer requisitos já presentes na biblioteca, mesmo quando redigidos de maneira diferente. Também seria interessante investigar a integração do Gemini com os VOEs (do inglês *Voice of Engineering*), estabelecendo uma conexão mais direta com o conhecimento e as necessidades da engenharia.

AGRADECIMENTOS

Gostaria de expressar, primeiramente, minha profunda gratidão ao meu orientador, Benjamin Grando Moreira, por compartilhar seu conhecimento e me ajudar a transformar ideias em resultados concretos durante o desenvolvimento deste trabalho. Minha gratidão se estende aos professores dos cursos de Ciência e Tecnologia, Engenharia de Transportes e Logística, e Engenharia Mecatrônica, que desempenharam um papel fundamental na minha formação acadêmica, assim como à Universidade Federal de Santa Catarina (UFSC), pelo suporte técnico e acadêmico oferecido durante toda esta jornada.

Agradeço imensamente à empresa parceira deste trabalho, não apenas por disponibilizar os dados essenciais para sua realização e permitir o uso de um LLM, mas também por proporcionar uma experiência enriquecedora em um ambiente profissional. Agradeço também aos meus colegas de trabalho que fizeram parte dessas vivências adquiridas durante meu período na empresa que foram cruciais para o desenvolvimento das competências e que tornaram este estudo possível.

Por fim, sou eternamente grata à minha família, especialmente aos meus pais, pelo amor e apoio incondicional ao longo de toda a minha trajetória acadêmica. Este trabalho reflete os valores de persistência e dedicação que sempre me ensinaram. Aos meus amigos, que estiveram ao meu lado com palavras de incentivo e compreensão, tornando essa jornada mais leve, deixo minha mais sincera gratidão. Mesmo aqueles que não me acompanharam desde o início do meu percurso universitário, compartilharam momentos preciosos que, de alguma forma, contribuíram para o meu crescimento pessoal e acadêmico.

REFERÊNCIAS

CARRARO, F. **O que é o Google Gemini e o que esse modelo de IA é capaz de fazer** - com exemplo prático. Alura. Disponível em: <https://www.alura.com.br/artigos/google-gemini>. Acesso em: 27 nov. 2024.

CASELI, H.; NUNES, M.; PAGANO, A. O que é PLN? *In*: CASELI, H. M.; NUNES, M. G. V. (org.) **Processamento de linguagem natural**: conceitos, técnicas e aplicações em português. 2. ed. São Carlos: BPLN, 2024. p. 10-14.

CASTILHO, S.; CASELI, H. Tradução Automática. *In*: CASELI, H. M.; NUNES, M. G. V. (org.) **Processamento de linguagem natural**: conceitos, técnicas e aplicações em português. 2. ed. São Carlos: BPLN, 2024. p. 492-517.

CHEN, W.; YAN-YI, L.; TIE-ZHENG, G.; DA-PENG, L.; TAO, H.; ZHI, L.; QING-WEN, Y.; WANG, H.; YING-YOU, W. **Systems engineering issues for industry applications of large language models**. *Applied Soft Computing*, v. 151, p. 111165, 2024.

CORTES, E.; VIEIRA, R.; BARONE, D. Perguntas e respostas. *In*: CASELI, H. M.; NUNES, M. G. V. (org.) **Processamento de linguagem natural**: conceitos, técnicas e aplicações em português. 2. ed. São Carlos: BPLN, 2024. p. 368-370.

DICK, J.; HULL, E.; JACKSON, K. **Requirements engineering**. 4. ed. Springer, 2017.

FAN, A.; GOKKAYA, B.; HARMAN, M.; LYUBARSKIY, M.; SENGUPTA, S.; YOO, S.; ZHANG, J. **Large Language Models for Software Engineering**: Survey and Open Problems. *IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE)*. 2023.

HITCHINS, D. **Systems engineering**: a 21st century systems methodology. John Wiley & Sons, 2008.

INCOSE. **Systems engineering handbook**: A guide for system life cycle processes and activities. 3. ed. John Wiley & Sons, v.3, 2006.

ISO/IEC/IEEE. **Systems and software engineering** - Life cycle processes - Requirements engineering 29148. v. 2018, 2018.

GRIFFIN, A.; HAUSER, J. **The voice of the customer**. *Marketing science*, v. 12, n. 1, p. 1-27, 1993.

KRISHNA, M.; GAUR, B.; VERMA, A.; Pankaj Jalote. **Using LLMs in Software Requirements Specifications**: An Empirical Evaluation. *arXiv preprint arXiv: 2404.17842*, 2024.

LEAL, S.; ALUÍSIO, S. Complexidade textual e suas tarefas relacionadas. *In*: CASELI, H. M.; NUNES, M. G. V. (org.) **Processamento de linguagem natural**: conceitos, técnicas e aplicações em português. 2. ed. São Carlos: BPLN, 2024. p. 537-546.

LOPES, P. **Uma taxonomia da pesquisa na área de engenharia de requisitos.** 2002. Dissertação (Mestrado em Ciência da Computação) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2002.

PAES, A.; VIANNA, D.; RODRIGUES, J. Modelos de linguagem. *In*: CASELI, H. M.; NUNES, M. G. V. (org.) **Processamento de linguagem natural: conceitos, técnicas e aplicações em português.** 3. ed. São Carlos: BPLN, 2024. p. 365-412.

VAZQUEZ, C., & SIMÕES, G. **Engenharia de requisitos: software orientado ao negócio.** Rio de Janeiro: Brasport, 2016.