



**UFC**

**UNIVERSIDADE FEDERAL DO CEARÁ**

**CENTRO DE TECNOLOGIA**

**DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA**

**CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO**

**FRANCISCO THIERRY OLIVEIRA SOUSA**

**USO DE MODELO DE LINGUAGEM LLM PARA ANÁLISE DE SENTIMENTO E  
RESUMO DE LIGAÇÕES EM CALL CENTER**

**FORTALEZA**

**2025**

FRANCISCO THIERRY OLIVEIRA SOUSA

USO DE MODELO DE LINGUAGEM LLM PARA ANÁLISE DE SENTIMENTO E  
RESUMO DE LIGAÇÕES EM CALL CENTER

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Orientador: Prof. Edilson Rocha Porfírio Filho.

FORTALEZA

2025

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- S696u Sousa, Francisco Thierry Oliveira.  
Uso de modelo de linguagem LLM para análise de sentimento e resumo de ligações em call center /  
Francisco Thierry Oliveira Sousa. – 2025.  
53 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia,  
Curso de Engenharia de Computação, Fortaleza, 2025.  
Orientação: Prof. Esp. Edilson Rocha Porfírio Filho.
1. Inteligência artificial. 2. Generative AI. 3. Análise de Sentimento. 4. Google Cloud Platform. 5.  
Call center. I. Título.

CDD 621.39

---

FRANCISCO THIERRY OLIVEIRA SOUSA

USO DE MODELO DE LINGUAGEM LLM PARA ANÁLISE DE SENTIMENTO E  
RESUMO DE LIGAÇÕES EM CALL CENTER

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Aprovada em: 08/08/2025

BANCA EXAMINADORA

---

Prof. Edilson Rocha Porfírio Filho (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Carlos Alexandre Rolim Fernandes  
Universidade Federal do Ceará (UFC)

---

Prof. Me. Ricardo Jardel Nunes da Silveira  
Universidade Federal do Ceará (UFC)

A Deus.

Aos meus pais, família e amigos por me apoiarem nessa trajetória.

## **AGRADECIMENTOS**

À minha família que sempre me apoiou e amigos. Ao Prof. Edilson Rocha Porfírio Filho, pela excelente orientação.

“Any sufficiently advanced technology is indistinguishable from magic.” (Arthur C. Clarke)

## RESUMO

O presente trabalho propõe uma solução de Inteligência Artificial para análise de ligações telefônicas em uma central de atendimento de uma companhia elétrica, utilizando exclusivamente recursos do *Google Cloud Platform* (GCP). A motivação surge da necessidade de otimizar processos, reduzir custos e melhorar o relacionamento com o cliente em um cenário de alta demanda por suporte e reclamações. Para tanto, são exploradas técnicas de Processamento de Linguagem Natural (PLN) e Modelos de Linguagem de Grande Porte (*Large Language Models* – LLMs), com foco na conversão de áudio em texto, classificação de sentimento e geração de resumos automáticos das interações. Os dados de áudio são armazenados inicialmente no *Cloud Storage*; em seguida, são transcritos para texto e importados no *BigQuery*, onde se empregam recursos do *BigQuery ML* (BQML) e do Vertex AI para desenvolvimento e integração dos modelos de análise. A abordagem proposta permite identificar o teor principal do problema, avaliar o sentimento do cliente (positivo, negativo ou neutro) e resumir o conteúdo de forma concisa. Dessa forma, a companhia pode priorizar atendimentos críticos, monitorar a satisfação do consumidor e obter insights valiosos para melhoria contínua do serviço. Os resultados obtidos confirmam a eficácia da aplicação de LLMs para apoiar a tomada de decisão nas centrais de atendimento, demonstrando que as soluções em nuvem do GCP oferecem escalabilidade, confiabilidade e segurança para lidar com grandes volumes de dados de forma eficiente.

**Palavras-chave:** Inteligência Artificial, Large Language Models (LLMs), Processamento de Linguagem Natural (PLN), Análise de Sentimento, Google Cloud Platform (GCP), BigQuery ML (BQML), Vertex AI

## ABSTRACT

The present work proposes an Artificial Intelligence solution for analyzing phone calls in a call center of an electric utility company, using exclusively Google Cloud Platform (GCP) resources. The motivation arises from the need to optimize processes, reduce costs, and improve customer relationships in a scenario of high demand for support and complaints. To this end, techniques of Natural Language Processing (NLP) and Large Language Models (LLMs) are explored, focusing on audio-to-text conversion, sentiment classification, and automatic summarization of interactions. The audio data is initially stored in Cloud Storage; then, it is transcribed into text and imported into BigQuery, where BigQuery ML (BQML) and Vertex AI resources are employed for the development and integration of analytical models. The proposed approach allows for identifying the main issue, assessing customer sentiment (positive, negative, or neutral), and summarizing the content concisely. In this way, the company can prioritize critical cases, monitor customer satisfaction, and obtain valuable insights for continuous service improvement. The results confirm the effectiveness of applying LLMs to support decision-making in call centers, demonstrating that GCP cloud solutions offer scalability, reliability, and security to handle large volumes of data efficiently.

**Keywords:** Artificial Intelligence, Large Language Models (LLMs), Natural Language Processing (NLP), Sentiment Analysis, Google Cloud Platform (GCP), BigQuery ML (BQML), Vertex AI.

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>9</b>
1.1 Contexto e Motivação.....	9
1.2 Problemas de Pesquisa.....	11
1.3 Objetivos.....	12
1.3.1 Objetivo Geral.....	12
1.3.2 Objetivos Específicos.....	12
1.4 Justificativa.....	12
1.5 Estrutura do Trabalho.....	14
<b>2 FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>15</b>
2.1 Análise e Processamento de Áudio (Speech-to-Text).....	16
2.1.1 Processamento de Linguagem Natural.....	19
2.1.2 Análise Multimodal Combinando Texto e Características Acústicas.....	20
2.2 Modelos de Linguagem de Grande Porte (LLMs).....	22
2.2.1 Evolução e Arquitetura.....	22
2.2.2 Aplicações em Cenários de Atendimento.....	23
2.2.3 Desafios e Considerações Técnicas.....	23
2.3 Análise de Sentimento.....	24
2.3.1 Metodologias Tradicionais.....	25
2.3.2 Abordagens Baseadas em Aprendizado Profundo.....	25
2.3.3 Importância do Contexto e do Léxico Setorial.....	26
2.4 Computação em Nuvem e Google Cloud Platform (GCP).....	26
2.4.1 Armazenamento no Google Cloud Storage (GCS).....	27
2.4.2 BigQuery(BQ) para Análises em Larga Escala.....	28
2.4.3 BigQuery ML (BQML).....	28
2.4.4 Vertex AI.....	29
2.4.5 Visualização de Dados com Looker Studio.....	30
2.4.6 Benefícios para o Projeto Proposto.....	30
<b>3 METODOLOGIA.....</b>	<b>32</b>
<b>4 RESULTADOS E DISCUSSÃO.....</b>	<b>41</b>
4.1 Apresentação dos resultados experimentais.....	41
4.2 Protocolo e Resultados da Validação do Modelo.....	42
4.3 Discussão dos Achados.....	42
4.3 Implicações, Limitações e Trabalhos Futuros.....	43
<b>5 CONCLUSÃO.....</b>	<b>44</b>
<b>REFERÊNCIAS.....</b>	<b>46</b>
<b>APÊNDICE A – SCRIPTS E PROMPTS DE ANÁLISE.....</b>	<b>49</b>
A.1 Introdução.....	49
<b>APÊNDICE B – DICIONÁRIO DE DADOS DAS TABELAS GERADAS.....</b>	<b>50</b>

B.1 Introdução.....	50
---------------------	----

## LISTA DE FIGURAS

Figura 1 – Variáveis Observadas no Questionário de Pesquisa do IASC	16
Figura 2 – Diagrama de Blocos para Análise Multimodal de Áudio	21
Figura 3 – Arquitetura do projeto no <i>Google Cloud Storage</i> (GCS)	34
Figura 4 – Interface do GCS	34
Figura 5 – Criação do <i>bucket</i>	35
Figura 6 – Local de registro do <i>datacenter</i>	36
Figura 7 – Registro de ligações	37
Figura 8 – Criação do Dataset	38
Figura 9 – Conexão do Vertex AI	39
Figura 10 – Tabela Recordings	40
Figura 11 – Dashboard LookerStudio	41
Figura 12 – Output Vertex	50

## 1 INTRODUÇÃO

A crescente demanda por serviços de qualidade e agilidade no fornecimento de energia elétrica, aliada à evolução constante das tecnologias de Inteligência Artificial (IA), tem impulsionado companhias do setor a repensarem seus processos de atendimento ao consumidor. Entre os canais de suporte disponíveis, destaca-se a central de ligações, ou *call center*, na qual os clientes relatam problemas, solicitam informações ou manifestam insatisfações. No entanto, lidar com o grande volume de chamadas diárias, muitas vezes repletas de vocabulário técnico e expressões de insatisfação, torna-se uma tarefa complexa e dispendiosa quando realizada manualmente.

Diante desse cenário, este Trabalho de Conclusão de Curso (TCC) propõe o desenvolvimento de uma solução automatizada, baseada em Modelos de Linguagem de Grande Porte (LLMs) e hospedada integralmente no *Google Cloud Platform* (GCP), para análise e classificação de ligações telefônicas de uma companhia elétrica. A proposta abrange desde a conversão de áudio em texto, passando pela identificação do problema relatado e classificação do sentimento do cliente, até a geração de resumos que facilitem a tomada de decisão gerencial. Seguindo um enfoque prático, a pesquisa busca demonstrar como a computação em nuvem e as técnicas de Processamento de Linguagem Natural (PLN) podem se alinhar às necessidades de negócios, otimizando o atendimento e reduzindo custos operacionais. O principal elemento que influencia a possibilidade de automação e otimização desses processos é a qualidade dos métodos de transcrição de chamadas, pois eles formam a base para o reconhecimento de intenções e a implementação de assistentes virtuais eficazes (PLAZA; PAWLIK; DENIZIAK, 2021).

### 1.1 Contexto e Motivação

A expansão do setor elétrico brasileiro, marcada pelo crescimento do consumo de energia e pela complexidade dos sistemas de distribuição, exige das concessionárias um atendimento cada vez mais ágil e de qualidade aos consumidores. Nesse cenário, as centrais de atendimento (*call centers*) tornam-se a principal interface entre a empresa e o cliente, recebendo diariamente grandes fluxos de ligações relacionadas a interrupções no fornecimento, reclamações, dúvidas, solicitações de religação, denúncias e outras demandas inerentes à prestação de serviços de energia. No entanto, lidar com esse elevado volume de chamadas de forma manual implica em desafios significativos, sobretudo em termos de tempo

de espera, qualidade do registro das informações e priorização das solicitações. Nesse sentido, a análise automática das emoções nessas conversas surge como uma ferramenta valiosa para avaliar a qualidade das interações entre clientes e agentes, com o objetivo final de aumentar a satisfação do cliente (FENG; DEVILLERS, 2023).

Ao mesmo tempo, a transformação digital tem favorecido a adoção de tecnologias capazes de aprimorar o relacionamento com o cliente e otimizar processos internos. Mais especificamente, o surgimento de técnicas avançadas de Inteligência Artificial (IA) — em particular o Processamento de Linguagem Natural (PLN) — permite a análise automatizada de grandes volumes de dados textuais e de voz, viabilizando *insights* sobre o teor das demandas, a satisfação do cliente e a eficiência do atendimento. Entre as aplicações de maior destaque, encontram-se os Modelos de Linguagem de Grande Porte (*Large Language Models* – LLMs), que possibilitam uma compreensão contextual mais profunda do conteúdo das interações e auxiliam na classificação de sentimentos, identificação de temas e geração de resumos.

Nesse sentido, a proposta de analisar automaticamente as ligações de um *call center* de uma companhia elétrica torna-se especialmente atrativa. Esse tipo de aplicação oferece diversos benefícios, como a redução de custos operacionais (ao minimizar a necessidade de escuta manual de gravações), a melhoria da satisfação do cliente (mediante identificação de demandas críticas) e a geração de inteligência de negócios (por meio da análise sistemática das principais causas de reclamações). E, de fato, a literatura aponta que a sumarização de diálogos possui grande valor comercial, pois permite que outros colaboradores ou sistemas subsequentes compreendam rapidamente os pontos principais da interação, melhorando a eficiência de todo o processo de atendimento (MA et al., 2022).

No que se refere à infraestrutura, o *Google Cloud Platform* (GCP) oferece um ecossistema de serviços robusto e escalável para dar suporte a aplicações de IA. Por meio de ferramentas como o *Cloud Storage*, o *BigQuery*, o *Vertex AI*, *BigQuery ML* (BQML) e o *Looker DataStudio*, é possível construir uma solução “de ponta a ponta”, iniciando pelo armazenamento de dados de áudio, passando pela transcrição e processamento textual, até a criação de modelos de análise de sentimento e geração de sumários. Ao se basear inteiramente no GCP, a companhia elétrica diminui a complexidade de integração entre diferentes plataformas, aproveitando as sinergias oferecidas por uma nuvem unificada.

Desse modo, este trabalho é motivado tanto pelas demandas específicas do setor elétrico, que exigem maior eficiência e assertividade na relação com o consumidor, quanto pelo potencial das novas tecnologias de IA e computação em nuvem, que propiciam soluções

automatizadas e escaláveis para análise de grande volume de dados. A aliança entre esses fatores fortalece a relevância do projeto e sustenta as premissas de melhoria contínua na prestação de serviços de energia.

## 1.2 Problemas de Pesquisa

Ainda que existam inúmeras soluções para *call centers*, a maioria dos processos de identificação de temas e classificação de sentimento ocorre de maneira manual ou semiautomática, o que pode gerar inconsistências e atrasos na tomada de decisão. Adicionalmente, a transcrição de ligações costuma ser um gargalo para muitas empresas que precisam lidar com áudios em grande quantidade e em diferentes qualidades de gravação, por vezes com ruídos ou vocabulário específico do setor. Essa dificuldade é corroborada pela literatura especializada, que aponta a resiliência a ruídos e a variabilidade dos ambientes acústicos como barreiras persistentes para o processamento de áudio em larga escala, mesmo com os recentes avanços na área (SAILAJA et al., 2025).

Assim, o problema de pesquisa que este trabalho se propõe a investigar pode ser formulado da seguinte maneira:

“Como analisar automaticamente grandes quantidades de ligações telefônicas de uma central de atendimento de uma companhia elétrica, identificando o problema principal, resumindo o teor do chamado e classificando o sentimento do cliente de modo a otimizar a triagem e a resolução de demandas?”

Dessa questão principal, derivam-se também os seguintes questionamentos:

Como garantir a precisão na transcrição de áudio em textos em ambientes sujeitos a ruídos, diferentes sotaques e uso de terminologia técnica?

Quais modelos de linguagem são mais adequados para realizar a extração de informações relevantes e a classificação de sentimento em interações relacionadas ao setor elétrico?

De que forma a integração entre as ferramentas do GCP (*Cloud Storage*, *BigQuery*, *Vertex AI*, *BigQuery ML*, *Looker DataStudio*) pode ser organizada para criar um *pipeline* consistente e escalável?

Quais métricas são mais adequadas para avaliar a qualidade dos resumos gerados e a precisão da análise de sentimento, garantindo a utilidade prática do sistema?

A resolução desses questionamentos orienta o desenvolvimento metodológico e a construção de uma prova de conceito que alicerça a contribuição deste trabalho, tanto para a academia quanto para o mercado.

## **1.3 Objetivos**

### ***1.3.1 Objetivo Geral***

Desenvolver e avaliar uma solução de Inteligência Artificial, hospedada no *Google Cloud Platform*, para analisar gravações de ligações de uma central de atendimento de uma companhia elétrica, identificando o conteúdo principal, resumindo as interações e classificando o sentimento dos clientes.

### ***1.3.2 Objetivos Específicos***

- Implementar um processo de armazenamento e organização dos arquivos de áudio no Google Cloud Storage, assegurando escalabilidade e segurança.
- Realizar a conversão de voz em texto (Speech-to-Text) com alta taxa de acerto, considerando cenários de ruído e variabilidade de fala.
- Aplicar técnicas de Processamento de Linguagem Natural (PLN) e LLMs para extrair palavras-chave, identificar o problema principal e categorizar o teor dos chamados.
- Desenvolver um modelo de análise de sentimento (positivo, neutro ou negativo), de modo a detectar variações de humor ou frustração do cliente ao longo da chamada.
- Implementar a geração de resumos automatizados, sintetizando as informações mais relevantes de cada ligação em um painel de BI no Looker DataStudio.
- Documentar as boas práticas de integração entre os serviços do GCP e as lições aprendidas, de forma a facilitar a adoção e manutenção do sistema no ambiente corporativo.

## **1.4 Justificativa**

A justificativa para esta pesquisa repousa em três pilares principais: o valor estratégico para o negócio, a contribuição acadêmico-científica e o impacto social. No que tange ao pilar de negócios, a análise de sentimento de chamadas já se estabeleceu como uma ferramenta crítica para a melhoria tanto da satisfação do cliente quanto da eficiência

operacional, permitindo que as empresas ajam de forma proativa e empática (RIZVI et al., 2025). Em relação a este valor, a possibilidade de automatizar parte do processo de análise de ligações traz benefícios diretos para a operação de uma companhia elétrica. A otimização do fluxo de atendimento e a agilidade na identificação de problemas críticos podem reduzir custos, mitigar riscos de imagem em casos de atendimento insatisfatório e, ainda, elevar a satisfação do cliente. Isso tem impacto direto na competitividade e no cumprimento das regulações do setor, que muitas vezes impõem metas de qualidade e prazos de resposta. Além disso, a análise sistematizada das chamadas gera dados estruturados que servem de insumo para estratégias de manutenção preditiva, política de investimentos em infraestrutura e aprimoramento de campanhas de comunicação com o público.

Do ponto de vista acadêmico-científico, o trabalho avança o debate sobre a aplicação prática de técnicas de PLN e LLMs em cenários de alto volume de dados e alta criticidade. Embora existam pesquisas voltadas à transcrição de áudio e análise de sentimento, a integração completa em um único *pipeline* na nuvem, com ênfase em modelos de linguagem capazes de resumir e identificar contexto em chamadas do setor elétrico, agrega novidade e relevância ao tema. A adoção exclusiva de recursos do GCP também representa uma contribuição, pois avalia empiricamente a efetividade e a flexibilidade de serviços como *Cloud Storage*, *BigQuery*, *BQML*, *Vertex AI* e *Looker DataStudio* em uma aplicação industrial real.

Outro aspecto a salientar é a dimensão social. A melhora na qualidade do atendimento em serviços essenciais, como a distribuição de energia elétrica, reflete diretamente na vida das pessoas. Um atendimento mais ágil e eficaz tende a diminuir transtornos em situações de interrupção no fornecimento de energia, especialmente em regiões mais vulneráveis, onde a eletricidade é crucial para preservar itens como medicamentos e alimentos, bem como para o funcionamento de escolas, hospitais e sistemas de água e esgoto.

Assim, o conjunto de motivações econômicas, técnicas e sociais fundamenta a necessidade e a pertinência deste estudo, indicando que a adoção de tecnologias de IA e computação em nuvem se configura como um caminho promissor para organizações que buscam elevar a eficiência de seus processos de atendimento e, ao mesmo tempo, prover um serviço de melhor qualidade ao consumidor final.

## 1.5 Estrutura do Trabalho

A fim de abordar de maneira sistemática todos os aspectos que envolvem o desenvolvimento e a validação da solução proposta, este trabalho está organizado da seguinte forma:

Capítulo 1 – Introdução: apresenta o contexto em que se insere o tema, as motivações para a pesquisa, o problema a ser investigado, os objetivos gerais e específicos, bem como a justificativa para a realização do estudo. Finaliza com a descrição da estrutura do trabalho.

Capítulo 2 – Fundamentação Teórica: dedica-se à revisão bibliográfica dos principais conceitos envolvidos, tais como análise de áudio, conversão voz-texto (Speech-to-Text), Processamento de Linguagem Natural, Modelos de Linguagem de Grande Porte (LLMs), análise de sentimento e computação em nuvem (GCP). Também apresenta trabalhos correlatos, discutindo pontos fortes, limitações e oportunidades de pesquisa.

Capítulo 3 – Metodologia: descreve em detalhes o design metodológico adotado, bem como a arquitetura da solução, os processos de coleta e pré-processamento de dados, o fluxo de treinamento e validação dos modelos de IA e a estratégia de integração dos serviços no GCP. Traz também os critérios de avaliação e as métricas utilizadas para mensurar o desempenho da proposta.

Capítulo 4 – Resultados e Discussões: exhibe os resultados experimentais obtidos, incluindo as métricas de acurácia e efetividade dos modelos, gráficos e tabelas comparativas, além de uma discussão crítica sobre os achados, relacionando-os com os objetivos e hipóteses iniciais. Analisa possíveis limitações e aponta caminhos para otimização.

Capítulo 5 – Conclusões e Trabalhos Futuros: retoma os objetivos propostos e verifica em que medida foram atingidos, apresentando uma síntese dos principais resultados e contribuições do trabalho. Em seguida, discute limitações encontradas e propõe linhas de pesquisas futuras que possam expandir ou aprofundar os aspectos tratados.

Ao longo de cada capítulo, busca-se aliar rigor teórico-científico à aplicabilidade prática, de forma a oferecer um trabalho que seja útil tanto para pesquisadores interessados em aplicações de IA no setor elétrico quanto para profissionais que almejam implementar soluções semelhantes em seus respectivos contextos organizacionais.

## 2 FUNDAMENTAÇÃO TEÓRICA

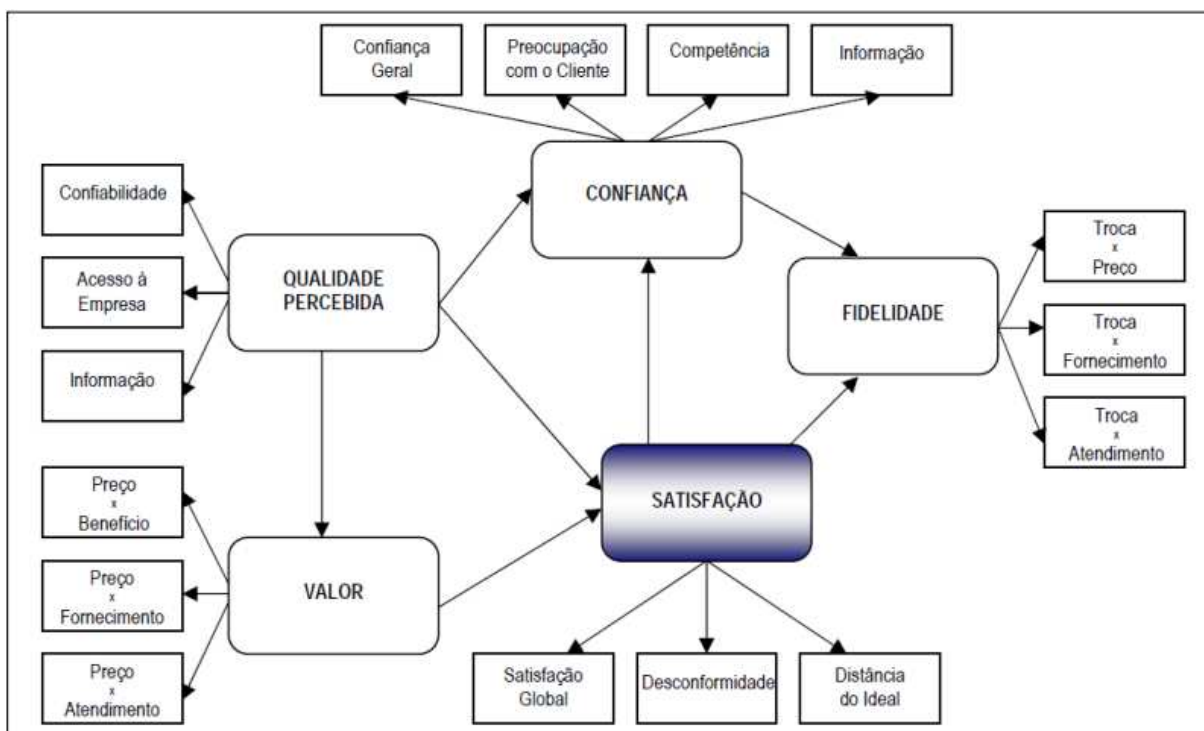
A qualidade de um serviço pode ser definida como o conjunto de atributos que atendem às necessidades explícitas e implícitas dos consumidores. Essa definição abrange parâmetros como desempenho, segurança e, de forma crucial, a ausência de falhas que gerem insatisfação. De maneira pragmática, a qualidade pode ser associada ao conceito de "utilidade para o uso", refletindo o quão bem um serviço cumpre sua função na perspectiva do cliente (SILVA, 2021). Nesse escopo, destacam-se a confiabilidade, que representa a consistência e a ausência de falhas do serviço, e a manutenibilidade, que se refere à facilidade com que um sistema pode ser reparado para restaurar sua funcionalidade (MUELLER, 2014).

A confiabilidade é uma parte indispensável da disponibilidade que mostra a falta de falha. O termo manutenção é usado como uma expressão que mostra os serviços de facilidade de reparo que podem ser feitos (programados ou não) (MUELLER, 2014).

No setor de distribuição de energia elétrica brasileiro, a qualidade é regulada e avaliada sob duas óticas principais. A primeira é a qualidade técnica, focada na continuidade do serviço. A Agência Nacional de Energia Elétrica (ANEEL) utiliza indicadores coletivos para essa medição, como a Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e a Frequência Equivalente de Interrupção por Unidade Consumidora (FEC). O descumprimento das metas estabelecidas para esses indicadores pode resultar em penalidades severas para as concessionárias (BARBOSA, CARVALHO E LOPES, 2005). A segunda ótica é a qualidade percebida, que reflete a satisfação do consumidor com o serviço como um todo. Ciente de sua importância, a ANEEL também promove o “Índice ANEEL de Satisfação do Consumidor (IASC)”, um prêmio concedido anualmente desde 2002 às concessionárias mais bem avaliadas pelos próprios clientes. O IASC serve como um mecanismo de incentivo para que as empresas aprimorem o atendimento e o relacionamento com o público, complementando os indicadores puramente técnicos (ANEEL, 2010).

O modelo conceitual do IASC, apresentado na Figura 1, detalha as variáveis que compõem essa avaliação, como Qualidade Percebida, Confiança, Valor e Fidelidade. A Qualidade Percebida, por exemplo, é desmembrada em subcritérios como acesso à empresa e confiabilidade dos serviços (ANEEL, 2010). Embora a fidelidade do cliente no setor elétrico seja cativa, visto que não há escolha de fornecedor, a satisfação registrada por meio desses instrumentos impacta diretamente a reputação e o desempenho regulatório da companhia.

Figura 1: Variáveis Observadas no Questionário de Pesquisa do IASC



Fonte: Agência Nacional de Energia Elétrica – ANEEL

É na central de atendimento (*call center*) que a maior parte da "qualidade percebida" é forjada. Este canal recebe um volume massivo de interações diárias, carregadas de informações críticas sobre falhas técnicas e, principalmente, sobre o sentimento do cliente. A análise manual sistemática dessas chamadas para extrair insights é operacionalmente inviável. Diante desse desafio, a aplicação de tecnologias de automação para analisar dados de voz não estruturados em larga escala emerge como uma solução estratégica fundamental.

Para tanto, este trabalho se aprofunda nos fundamentos teóricos que viabilizam tal solução. As seções seguintes abordarão os conceitos de Análise e Processamento de Áudio (*Speech-to-Text*), Processamento de Linguagem Natural (PLN), Modelos de Linguagem de Grande Porte (LLMs), Análise de Sentimento e a arquitetura de Computação em Nuvem que sustenta o projeto.

## 2.1 Análise e Processamento de Áudio (*Speech-to-Text*)

A análise de áudio compreende um conjunto de métodos e técnicas que permitem extrair informações úteis de sinais sonoros, como reconhecimento de fala, identificação de locutor (*speaker identification*), detecção de emoções, entre outros. Em aplicações de *call*

*center*, o foco recai predominantemente sobre a conversão de voz em texto (*Speech-to-Text*), viabilizando o subsequente emprego de técnicas de Processamento de Linguagem Natural (PLN).

A identificação do locutor é um pré-requisito para a Compreensão de Linguagem Natural (NLU), pois a mesma declaração pode ter significados muito diferentes quando dita pelo agente ou pelo cliente. Uma frase como 'o problema não foi resolvido', por exemplo, tem um peso e uma implicação distinta dependendo de quem a pronuncia. A falha em atribuir corretamente a autoria de cada trecho pode levar a uma interpretação de sentimento e intenção completamente equivocada (DO; MAI, 2020).

De acordo com Almeida e Haddad (2018), o processamento de áudio, e em particular a tecnologia de conversão de fala em texto (STT), provou ser uma ferramenta transformadora em muitas áreas, do atendimento ao cliente à acessibilidade e automação de atividades rotineiras. Essa tecnologia automatizada de conversão de áudio para texto é baseada em avanços significativos em inteligência artificial e aprendizado de máquina, alcançando níveis de precisão cada vez mais impressionantes. No entanto, apesar do seu potencial, a STT ainda enfrenta desafios que precisam ser superados para que seja aplicada de forma eficaz e abrangente. O processo de conversão de fala em texto começa com o pré-processamento de áudio, que envolve a remoção de ruído do sinal de áudio e sua preparação para análise.

Segundo Martins e Santos (2022), características relevantes, como frequência e intensidade, são então extraídas para ajudar a distinguir entre sons e palavras. Métodos como os Coeficientes Cepstrais em Escala Mel (MFCC) são comumente usados para modelar a percepção sonora humana. Com base nessas características, modelos de aprendizado de máquina, como redes neurais profundas, são usados para reconhecer padrões e associá-los a palavras ou frases.

Modelos modernos, como aqueles baseados na arquitetura Transformer, se distinguem por sua capacidade de capturar contextos complexos e dependências temporais na fala.

O texto gerado passará por pós-processamento, onde erros serão corrigidos, pontuação será adicionada e o conteúdo será ajustado para se adequar ao contexto específico da aplicação. As aplicações da STT são muito amplas e têm impacto direto na vida cotidiana de pessoas e organizações. Por exemplo, no atendimento ao cliente, a gravação automatizada de chamadas permite a análise em tempo real do sentimento do cliente, a detecção de

palavras-chave e a geração automática de respostas, ajudando a aumentar a eficiência e a personalização do serviço.

Assistentes virtuais como *Siri*, *Google Assistant* e *Alexa* usam essa tecnologia para entender comandos de voz e realizar tarefas, desde pesquisar na web até controlar dispositivos inteligentes. Em ambientes profissionais, o STT é usado para gerar automaticamente atas de reuniões, palestras e entrevistas, ajudando a documentar e verificar o conteúdo. Além disso, a tecnologia desempenha um papel fundamental na acessibilidade, permitindo que pessoas surdas ou com deficiência de fala se comuniquem usando legendas em tempo real ou conversão de texto em fala. As empresas também usam o STT para análise de sentimentos e pesquisa de mercado, extraindo informações valiosas dos dados dos clientes (MARTINS E SANTOS, 2022).

Apesar das suas vantagens, o método STT ainda enfrenta sérios problemas que limitam a sua precisão e utilidade em alguns contextos. Um dos principais obstáculos são as diferenças na fala, como diferentes sotaques, velocidades de fala e entonações, que podem dificultar o reconhecimento preciso das palavras. Ruído ambiente e sons de fala sobrepostos, comuns em reuniões ou ambientes barulhentos, também reduzem a qualidade da gravação. A linguagem humana, repleta de ambiguidades e expressões idiomáticas, cria outro problema porque a mesma palavra pode ter significados completamente diferentes dependendo do contexto. Além disso, os modelos STT baseados em aprendizado de máquina exigem grandes quantidades de dados de treinamento, o que pode ser desafiador para idiomas menos comuns ou dialetos locais. (SOUZA E RIBEIRO, 2020).

Há também preocupações com privacidade e segurança, pois a gravação e o processamento de áudio envolvem o manuseio de dados confidenciais, exigindo conformidade com regulamentações como a LGPD (LIMA e FERREIRA, 2022).

Atrasos no processamento de áudio podem prejudicar a experiência do usuário em aplicativos que exigem respostas em tempo real, como assistentes virtuais. O desenvolvimento adicional da tecnologia STT promete resolver muitos desses problemas, tornando a tecnologia mais precisa, mais eficiente e melhor integrada à vida cotidiana.

Com o desenvolvimento de tecnologias mais avançadas e o aumento do poder computacional, a STT tem o potencial de transformar ainda mais a maneira como nos comunicamos e usamos a tecnologia. No entanto, isso requer investimento em pesquisa para remover as limitações atuais, bem como garantir que a tecnologia seja acessível, segura e adequada às diversas necessidades dos usuários. O processamento de áudio, especialmente o processamento de fala para texto, não é apenas uma ferramenta tecnológica, mas também um

meio de facilitar relacionamentos interpessoais e um catalisador para mudanças e inovação em muitos campos (SOUZA E RIBEIRO, 2020).

### ***2.1.1 Processamento de Linguagem Natural***

O processamento de linguagem natural (PLN), é um campo da inteligência artificial que se concentra na interação entre computadores e linguagens humanas. O objetivo é permitir que as máquinas entendam, interpretem e produzam texto ou fala assim como os humanos. No Brasil, o PLN se popularizou graças ao desenvolvimento da tecnologia digital e à necessidade de soluções tecnológicas que levassem em conta as especificidades da língua portuguesa.

Autores como Almeida e Haddad (2018) destacam que o processamento da linguagem natural é a base para o desenvolvimento de sistemas inteligentes capazes de processar grandes volumes de dados textuais, como mídias sociais, reuniões, documentos e gravações. Uma das principais aplicações da PLN é a análise de sentimentos, que permite identificar as emoções e opiniões expressas no texto.

Segundo Costa e Silva (2020), esse método é muito utilizado em pesquisas de *marketing* e monitoramento de mídias sociais, onde as empresas buscam entender como os consumidores se comportam frente a produtos e serviços prestados. No Brasil, o processamento de linguagem natural é usado para analisar comentários em plataformas como o X (antigo *Twitter*) e ReclameAqui, ajudando organizações a identificar tendências e tomar decisões estratégicas.

Além disso, o processamento de linguagem natural é essencial para criar assistentes virtuais como *chatbots*, que usam técnicas de compreensão de linguagem para se comunicar com os usuários de forma natural e eficaz. Outra aplicação atual da PLN é a tradução automática, que ajuda a facilitar a comunicação entre falantes de diferentes idiomas. Entretanto, como Freitas e Souza (2019) apontaram, o português brasileiro apresenta diversos problemas, como diferenças regionais e expressões idiomáticas, que dificultam a criação de modelos precisos. Entretanto, avanços recentes em redes neurais e modelos de linguagem como BERT (*Bidirectional Encoder Representations from Transformers*) melhoraram significativamente a capacidade de traduzir e interpretar textos em português.

O processamento de linguagem natural também desempenha um papel importante na área da saúde, onde é usado para processar registros médicos e artigos científicos. Martins

e Santos (2022) enfatizam que extrair informações relevantes de textos médicos pode auxiliar no diagnóstico de doenças e na personalização de tratamentos.

No Brasil, as pesquisas nessa área se concentram no desenvolvimento de sistemas que possam identificar termos técnicos e relacioná-los a sintomas e diagnósticos, contribuindo assim para melhorar a qualidade da assistência à saúde. Apesar dos progressos realizados, o processamento da linguagem natural ainda enfrenta desafios significativos, especialmente quando se trata do português.

Lima e Ferreira (2022) enfatizam que a falta de grandes conjuntos de dados portugueses anotados dificulta o treinamento de modelos de aprendizado de máquina. Além disso, a diversidade linguística do Brasil, com seus muitos sotaques e regiões, exige uma abordagem especial para garantir a precisão do sistema. Outra preocupação é a privacidade dos dados, já que o processamento de texto geralmente envolve informações confidenciais, como registros médicos ou conversas privadas. Portanto, o cumprimento da Lei Geral de Proteção de Dados (LGPD) é um aspecto importante no desenvolvimento de soluções baseadas no processamento de linguagem natural.

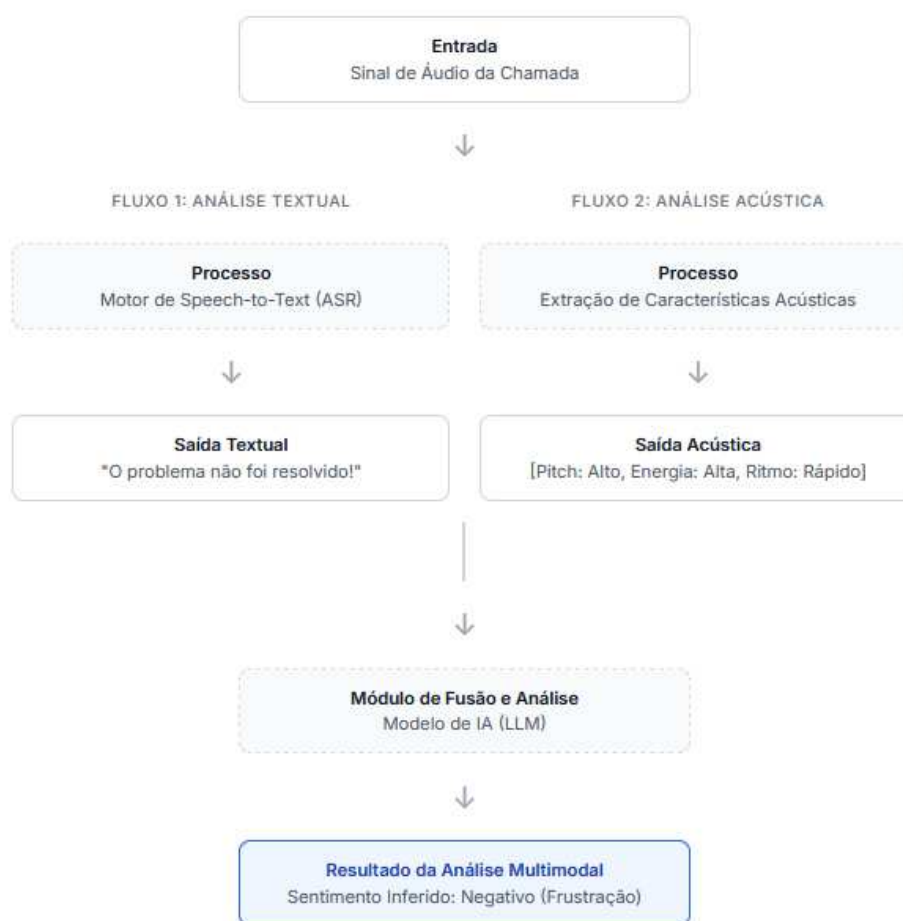
### ***2.1.2 Análise Multimodal Combinando Texto e Características Acústicas***

Embora a análise do conteúdo textual transcrito seja fundamental, uma abordagem mais completa para a compreensão da interação em um *call center* envolve a análise multimodal, que considera não apenas o conteúdo linguístico (o que foi dito), mas também as características paralinguísticas ou acústicas da fala (o como foi dito). Essas características, como o tom de voz, a frequência, o ritmo e a energia, carregam informações emocionais ricas que podem não estar explícitas no texto.

Conforme aponta a literatura, a incorporação de características acústicas extraídas diretamente do sinal de áudio enriquece significativamente a análise de sentimento e a detecção de emoções. Atributos como tom, *pitch* e níveis de energia fornecem pistas cruciais sobre o estado emocional do locutor, que muitas vezes são perdidas em uma análise puramente textual (RIZVI et al., 2025). Em cenários de alta criticidade, como chamadas de emergência, a análise de características acústicas, como os Coeficientes Cepstrais de Frequência Mel (MFCCs), pode ser utilizada para inferir a severidade de uma situação, complementando a análise da transcrição (ABI KANAAN et al., 2023).

Essa abordagem combinada é particularmente poderosa em um ambiente de *call center*, onde a variabilidade emocional é grande. Um cliente pode usar palavras neutras, mas expressar frustração através de um tom de voz elevado ou uma fala mais rápida. Da mesma forma, a empatia de um agente pode ser percebida não apenas por suas palavras, mas pela suavidade e calma de sua voz (FENG; DEVILLERS, 2023).

Figura 2: Diagrama de Blocos para Análise Multimodal de Áudio



Fonte: elaborado pelo próprio autor

Portanto, um sistema de análise de ponta a ponta ideal não se limita a transcrever o áudio para depois analisá-lo, mas processa o áudio em paralelo para extrair tanto o conteúdo textual quanto um vetor de características acústicas. Esses dois fluxos, ilustrados na Figura 2, de informação podem, então, ser fundidos para alimentar um modelo de IA, resultando em uma classificação de sentimento e intenção muito mais precisa e contextualmente ciente.

## 2.2 Modelos de Linguagem de Grande Porte (LLMs)

Os Modelos de Linguagem de Grande Porte, conhecidos pela sigla LLM (*Large Language Models*), têm revolucionado a forma como as máquinas compreendem e geram texto. Esses modelos são treinados em vastos corpos de dados textuais e utilizam arquiteturas baseadas em transformadores, que se destacam por sua capacidade de capturar relações complexas e contextos extensos entre palavras e frases.

### 2.2.1 Evolução e Arquitetura

Historicamente, os métodos de PLN evoluíram de abordagens baseadas em regras e modelos estatísticos para técnicas de aprendizado profundo, culminando no surgimento dos transformadores. Essa arquitetura, introduzida com o modelo Transformer, baseia-se em mecanismos de *self-attention*, que permitem ao modelo ponderar a importância de cada palavra em relação a todas as outras na sequência para aprender o contexto do texto (ABI KANAAN et al., 2023). Esse mecanismo é crucial para entender nuances contextuais, como ironia, negação e ambiguidade semântica.

Entre os LLMs mais conhecidos, destacam-se:

- **BERT (*Bidirectional Encoder Representations from Transformers*)**: Focado na compreensão bidirecional do contexto, BERT tem se mostrado altamente eficaz em tarefas de classificação e extração de informações.
- **GPT (*Generative Pre-trained Transformer*)**: Especializado na geração de texto, o GPT, em suas várias iterações, demonstra uma capacidade notável para produzir textos coerentes e contextualmente relevantes, sendo a base de muitos assistentes de voz e *chatbots* modernos (GOBINATH et al., 2024).
- **T5 (*Text-to-Text Transfer Transformer*)**: Este modelo unifica diferentes tarefas de PLN em um formato único de transformação de texto para texto, facilitando o ajuste fino para tarefas específicas.

### 2.2.2 Aplicações em Cenários de Atendimento

No contexto de centrais de atendimento, os LLMs desempenham papéis essenciais em diversas frentes:

1. **Geração de Resumos:** Uma das aplicações mais valiosas é a capacidade de sintetizar longas interações telefônicas em resumos concisos e informativos. Isso facilita a rápida compreensão do conteúdo das chamadas por gestores e equipes de suporte, permitindo uma triagem eficiente e o direcionamento adequado das demandas.
2. **Classificação e Extração de Temas:** Os LLMs podem identificar automaticamente os principais tópicos abordados durante uma ligação, diferenciando, por exemplo, entre reclamações sobre interrupção no fornecimento e dúvidas sobre a conta de energia. Essa categorização automatizada é fundamental para a priorização de ações e para a identificação de padrões recorrentes que podem indicar problemas sistêmicos.
3. **Análise de Sentimento Aprimorada:** Ao integrar técnicas de LLM com algoritmos de análise de sentimento, é possível obter uma interpretação mais refinada do estado emocional dos clientes. A compreensão contextual profunda permite identificar nuances, como a mudança gradual do sentimento ao longo da chamada, o que pode indicar se um atendimento conseguiu reverter uma situação de insatisfação.

Além desses pontos, uma vertente de pesquisa emergente é o uso de LLMs para a correção de erros de sistemas de ASR (*Automatic Speech Recognition*). Nessa abordagem, as transcrições geradas pelo ASR, mesmo que contenham falhas, são utilizadas como entrada para que o LLM, com seu vasto conhecimento linguístico e contextual, possa prever a transcrição correta, superando os métodos tradicionais de reclassificação de hipóteses (ANH; HO, 2024).

### 2.2.3 Desafios e Considerações Técnicas

Apesar das inúmeras vantagens, a utilização de LLMs também apresenta desafios. O treinamento desses modelos requer grandes quantidades de dados e poder computacional, o que pode elevar os custos e a complexidade do desenvolvimento, sendo um desafio notável para treinar modelos para idiomas com poucos recursos ou domínios muito específicos (LIU

et al., 2024). Além disso, é fundamental considerar questões éticas e de viés, uma vez que os dados utilizados para treinamento podem refletir preconceitos existentes na sociedade.

Para contornar esses desafios, é comum recorrer a técnicas de *fine-tuning* em que um modelo pré-treinado é ajustado com um conjunto de dados específico do domínio, no caso deste projeto, dados relativos às interações de uma companhia elétrica. Essa abordagem não só reduz os custos computacionais, como também melhora a precisão e a relevância dos resultados para o contexto aplicado (GOOGLE, 2025).

A integração dos LLMs com os demais componentes da solução é facilitada pelas plataformas de computação em nuvem, como o *Google Cloud Platform* (GCP), que são cruciais para aplicações de negócio que exigem segurança, escalabilidade e ferramentas de monitoramento (PLAZA; PAWLIK; DENIZIAK, 2021). No projeto, os LLMs podem ser empregados em conjunto com serviços como o *BigQuery ML* e o *Vertex AI* para:

- Automatizar a geração de resumos e a classificação de chamadas;
- Realizar ajustes contínuos dos modelos a partir de dados reais de interações, garantindo que a ferramenta evolua conforme as demandas do atendimento;
- Facilitar a análise e a visualização dos resultados através de dashboards interativos no Looker Studio.

Em suma, os Modelos de Linguagem de Grande Porte representam uma ferramenta poderosa para transformar dados não estruturados em insights estratégicos, permitindo que as centrais de atendimento se tornem mais responsivas, precisas e orientadas a dados.

### **2.3 Análise de Sentimento**

A análise de sentimento, também conhecida como *opinion mining*, consiste na aplicação de técnicas computacionais para identificar e classificar o teor emocional ou a polaridade de um texto, geralmente categorizando-o em positivo, negativo ou neutro. Em cenários mais avançados, podem-se atribuir graduações de intensidade (como “muito positivo” ou “ligeiramente negativo”) ou até detectar emoções específicas, como alegria, raiva, tristeza e surpresa. Seu principal objetivo é compreender a percepção subjetiva do autor de um texto em relação a um tema, produto, serviço ou evento.

### **2.3.1 Metodologias Tradicionais**

No início, a análise de sentimento era comumente realizada por métodos *éxicon-based*, que utilizam dicionários com termos positivos e negativos e, por meio de contagens, inferem o sentimento predominante de um texto. Embora simples de implementar, essas abordagens enfrentam limitações, pois não levam em conta o contexto em que as palavras aparecem. Uma frase que contenha termos “positivos” e “negativos” pode mudar completamente de sentido ao se inserir uma negação ou ironia (RIZVI et al., 2025).

Para superar essa limitação, passaram a ser empregados algoritmos de aprendizado de máquina tradicional, como *Naive Bayes*, *Support Vector Machines* (SVM) e Regressão Logística, onde um conjunto de textos previamente rotulados em categorias de sentimento servia de base para o treinamento do modelo. Nessas abordagens, o texto normalmente é convertido em vetores de atributos (por exemplo, usando *bag-of-words* ou *tf-idf*), e o algoritmo aprende a relacionar padrões de vocabulário com as classes de sentimento. Embora mais robustos do que os métodos *éxicon-based*, esses modelos ainda têm dificuldades para lidar com ambiguidades e expressões sutis, como sarcasmo ou gírias regionais (YADAV; VISHWAKARMA, 2020).

### **2.3.2 Abordagens Baseadas em Aprendizado Profundo**

Com o advento das redes neurais profundas, a análise de sentimento alcançou novos patamares de desempenho. Modelos como redes neurais recorrentes (RNN), LSTMs (*Long Short-Term Memory*) ou GRUs (*Gated Recurrent Units*) foram capazes de capturar melhor o contexto ao longo de sequências de palavras, preservando dependências de longo alcance em um texto. Dessa forma, expressões negativas que aparecem no começo de uma frase podem ser relacionadas adequadamente a termos positivos do final, refletindo uma mudança de tom ou sentido (RIZVI et al., 2025).

Mais recentemente, a adoção de arquiteturas baseadas em Transformadores vem demonstrando resultados ainda superiores. Esses modelos incorporam mecanismos de *self-attention*, os quais identificam como cada termo do texto se relaciona com os demais termos, tanto anteriores quanto posteriores. Essa capacidade de análise contextual global é altamente relevante para a análise de sentimento, uma vez que ironias, negações ou palavras ambíguas podem ser interpretadas de forma mais precisa quando o modelo compreende o texto como um todo (ABI KANAAN et al., 2023).

### 2.3.3 Importância do Contexto e do Léxico Setorial

A análise de sentimento exige atenção ao vocabulário específico do domínio. Em uma companhia elétrica, termos como “relição”, “poste”, “disjuntor” e “queda de fase” podem assumir conotações distintas, influenciando a percepção do cliente. É imprescindível que o modelo seja treinado ou ajustado levando em conta esse léxico particular, uma vez que modelos de reconhecimento de fala genéricos apresentam fraco desempenho em domínios com terminologia técnica especializada (NITU; CATRUNA; RADOI, 2024). No caso de um cliente que menciona “falta de luz” repetidamente, é provável que o sentimento geral seja negativo, embora a palavra “luz” isoladamente não indique nem positividade nem negatividade.

Além disso, é fundamental considerar o **contexto de uso** das palavras. Uma frase como “O atendimento foi ótimo, mas resolveram apenas metade do problema” traz indícios de um tom misto. Um modelo menos robusto poderia classificar esse texto como “positivo” por conta de termos como “ótimo”, ignorando o descontentamento subsequente, enquanto outro modelo poderia considerar todo o texto como “negativo” ao focalizar a insatisfação final. Modelos baseados em Transformadores conseguem equilibrar ambas as nuances, capturando a leve contradição expressa na sentença.

## 2.4 Computação em Nuvem e Google Cloud Platform (GCP)

A computação em nuvem se consolidou como um modelo de uso de recursos de TI que dispensa a aquisição e a manutenção de infraestrutura física própria. Em vez disso, as empresas podem provisionar servidores, armazenamento e serviços de forma elástica, pagando apenas pelo que consomem. Essa abordagem oferece vantagens como escalabilidade rápida, alta disponibilidade, segurança gerenciada e integração simplificada entre diversos componentes.

No contexto do presente projeto, o *Google Cloud Platform* (GCP) foi escolhido como provedor de nuvem devido à sua infraestrutura global robusta e ao amplo ecossistema de serviços que abrange desde o armazenamento de dados e análise em larga escala até a criação e implantação de modelos de *Machine Learning* (GOOGLE, 2025). A seguir, detalham-se os principais componentes do GCP empregados nesta solução.

### 2.4.1 Armazenamento no Google Cloud Storage (GCS)

O *Cloud Storage* é o serviço de armazenamento de objetos do GCP, projetado para alta durabilidade, disponibilidade e escalabilidade (GOOGLE, 2025). Nesse sistema, os dados são organizados em *buckets*, que são os contêineres básicos que armazenam as informações na nuvem. Cada *bucket* funciona como um repositório único, com nome global e configurações próprias de localização e acesso. No projeto em questão, o *Cloud Storage* desempenha o papel de repositório central de dados de áudio das ligações telefônicas. Assim que as gravações são concluídas, elas são enviadas para um *bucket* específico configurado na plataforma, onde ficam prontamente acessíveis para processamento.

#### 1. Classes de Armazenamento

O *Cloud Storage* oferece diferentes classes (*Standard, Nearline, Coldline, Archive*), cada uma com características de custo e frequência de acesso específicas. Para gravações de chamadas que precisam de análises frequentes, a classe *Standard* costuma ser indicada por oferecer acesso rápido e baixo tempo de latência (GOOGLE, 2025).

#### 2. Segurança e Controle de Acesso

A segurança dos dados é um aspecto crucial, sobretudo em aplicações que envolvem informações sensíveis de clientes. O *Cloud Storage* possibilita configurar permissões detalhadas (IAM), criptografia em repouso e em trânsito, além de ferramentas de auditoria, garantindo conformidade com regulações como a LGPD (Lei Geral de Proteção de Dados) (GOOGLE, 2025).

#### 3. Integração com Outros Serviços

Uma vantagem do *Cloud Storage* é sua integração nativa com o restante do ecossistema GCP, permitindo orquestrar *pipelines* automáticos. Por exemplo, assim que um arquivo de áudio é armazenado no *bucket*, um serviço de *Speech-to-Text* pode ser acionado para efetuar a transcrição e enviar o resultado diretamente para o *BigQuery* (GOOGLE, 2025).

### 2.4.2 *BigQuery(BQ) para Análises em Larga Escala*

O *BigQuery* é o *data warehouse* totalmente gerenciado e sem servidor do Google, projetado para agilizar a análise de dados em larga escala por meio de consultas SQL (GOOGLE, 2025). Isso significa que, mesmo sem conhecimentos profundos em programação, equipes de análise podem realizar consultas complexas sobre um volume massivo de registros em um curto espaço de tempo. No escopo deste projeto, o *BigQuery* cumpre as seguintes funções principais:

#### 1. **Armazenamento e Estruturação de Dados**

Após a transcrição dos áudios, cada ligação gera um registro textual contendo campos como data, horário, duração, ID do cliente, transcrição do diálogo e, posteriormente, classificações de sentimento e tópicos identificados pelos modelos. Esses dados são carregados em tabelas do *BigQuery*, que passam a servir como base para consultas, agregações e relatórios (GOOGLE, 2025).

#### 2. **Consultas Analíticas**

Com comandos SQL, é possível agrupar as chamadas por tipo de problema (por exemplo, “falta de energia”, “dúvidas sobre conta”, “relição”), calcular estatísticas de tempo médio de atendimento ou identificar termos mais frequentes em queixas. Essa praticidade facilita a geração de insights gerenciais que orientam melhorias no atendimento (GOOGLE, 2025).

#### 3. **Escalabilidade e Otimização de Custos**

O *BigQuery* aloca recursos de processamento conforme a complexidade das consultas e o volume de dados, dispensando configuração manual de servidores ou *clusters*. O modelo de cobrança *pay-as-you-go* assegura que a empresa pague apenas pelas consultas executadas, eliminando custos fixos elevados (GOOGLE, 2025).

### 2.4.3 *BigQuery ML (BQML)*

O *BigQuery ML* (BQML) adiciona a capacidade de construção e implantação de modelos de aprendizado de máquina diretamente no ambiente SQL do *BigQuery*. Isso representa um diferencial importante, pois elimina a necessidade de extrair dados para outras plataformas, reduzindo tempo e complexidade na preparação e na movimentação de dados (GOOGLE, 2025).

### 1. Construção e Treinamento de Modelos

Por meio de instruções SQL, é possível criar modelos de regressão, classificação e *time series forecasting*. No projeto, a principal aplicação do BQML envolve a **classificação de sentimento** (positivo, negativo ou neutro) e **categorização** dos chamados por assunto. Assim, basta especificar a tabela de treinamento e os parâmetros do modelo, e o BQML cuida do processo de aprendizado internamente (GOOGLE, 2025).

### 2. Integração Direta com Relatórios

Por estar no mesmo ambiente do *BigQuery*, os resultados das previsões ficam prontamente disponíveis para consultas adicionais ou para visualizações em ferramentas de BI, como o *Looker Studio*, o que agiliza a criação de *dashboards* e relatórios analíticos (GOOGLE, 2025).

#### 2.4.4 Vertex AI

O *Vertex AI* é a plataforma unificada de Inteligência Artificial do *Google Cloud* que abrange todo o ciclo de vida de um projeto de *machine learning*, desde a preparação e rotulagem dos dados, passando pelo treinamento e ajuste fino dos modelos, até a implantação e monitoramento contínuo em produção (GOOGLE, 2025). Essa plataforma facilita a orquestração de *pipelines* de IA, permitindo que desenvolvedores e cientistas de dados se concentrem na melhoria dos modelos, sem se preocupar com a infraestrutura subjacente.

No contexto deste projeto, o *Vertex AI* desempenha um papel central na integração e gerenciamento dos modelos de análise de sentimentos e na geração de resumos das ligações. Em particular, utilizei o modelo **Gemini-1.5-Pro**, um dos mais avançados oferecidos pela Google, adaptado para compreender e processar o vocabulário e as especificidades das interações em centrais de atendimento de uma companhia elétrica. O Gemini 1.5 Pro, incorporado via *Vertex AI*, demonstra alta capacidade de entendimento contextual e precisão na extração de informações relevantes, contribuindo significativamente para a classificação dos sentimentos e para a síntese dos diálogos transcritos (GOOGLE, 2025).

Além disso, o *Vertex AI* permite automatizar processos como o ajuste de hiperparâmetros, o monitoramento de desempenho e a atualização dos modelos com dados novos. Essa integração facilita a manutenção de um sistema de IA robusto e escalável, garantindo que a solução se adapte às variações de volume e complexidade das chamadas ao longo do tempo. Dessa forma, o uso do modelo Gemini-1.5-Pro via *Vertex AI* não só

potencializa a qualidade das análises realizadas, mas também promove a agilidade na implantação de atualizações e na geração de insights que suportam a tomada de decisão estratégica (GOOGLE, 2025).

#### **2.4.5 Visualização de Dados com Looker Studio**

Para que os resultados gerados por BQ e BQML sejam efetivamente utilizados na tomada de decisão, faz-se essencial a apresentação de informações de forma clara e intuitiva. Nesse sentido, o *Looker Studio* (antigo *Google Data Studio*) é uma ferramenta de *Business Intelligence* (BI) que permite a criação de *dashboards* e relatórios personalizados por meio de uma interface amigável e conectores diretos com o ecossistema GCP (GOOGLE, 2025).

##### **1. Conexão e Fonte de Dados**

A ferramenta se integra de forma nativa ao *BigQuery*, possibilitando a criação de relatórios dinâmicos em que os dados são atualizados quase em tempo real. Dessa forma, as métricas de ligações atendidas, tempo de chamada, tendências de sentimento e outras estatísticas podem ser visualizadas imediatamente após o processamento.

##### **2. Criação de Paineis Interativos**

Por meio de gráficos, tabelas e filtros customizáveis, o *Looker Studio* permite que gestores e analistas naveguem pelas informações, obtendo insights detalhados. É possível, por exemplo, criar um painel que mostra o volume de chamadas negativas por dia.

##### **3. Compartilhamento e Colaboração**

Como é uma plataforma de nuvem, o *Looker Studio* permite compartilhar *dashboards* com membros de diferentes equipes ou até mesmo com clientes externos, definindo permissões de exibição ou edição. Esse recurso é valioso em organizações que desejam promover a tomada de decisão baseada em dados em todos os níveis hierárquicos.

#### **2.4.6 Benefícios para o Projeto Proposto**

A adoção integral do GCP permite que todo o *pipeline* de dados e IA seja hospedado em um único ambiente, simplificando a arquitetura e reduzindo possíveis pontos de falha. Além disso, a fácil integração entre os serviços acelera o processo de

desenvolvimento e diminui custos de manutenção, pois não é necessário gerenciar infraestrutura de *clusters* de processamento ou configurar plataformas de *machine learning* do zero.

Para a empresa em questão, uma companhia do setor elétrico com alto volume de chamadas diárias, isso significa responder rapidamente a mudanças na demanda (por exemplo, épocas de grande incidência de temporais ou picos de reclamações). A solução consegue escalar virtualmente de forma automática, evitando indisponibilidades e permitindo análises imediatas das gravações mais recentes.

### 3 METODOLOGIA

A metodologia de Estudo de Caso foi adotada para a realização e abordagem da presente temática. Esta metodologia de pesquisa é amplamente utilizada nas ciências sociais e aplicadas, que permite a investigação detalhada de um fenômeno específico dentro de seu contexto real.

De acordo com Yin (2015), o estudo de caso é uma investigação empírica que estuda um fenômeno contemporâneo dentro de seu contexto da vida real, especialmente quando os limites entre o fenômeno e o contexto não estão claramente definidos. Quando aplicado de forma quantitativa, o estudo de caso utiliza dados numéricos e métodos estatísticos para analisar e interpretar os resultados, permitindo uma abordagem mais objetiva e generalizável.

O objetivo principal deste estudo de caso quantitativo é analisar a qualidade do atendimento telefônico em uma empresa do setor de distribuição elétrica, com foco na experiência positiva do cliente. A pesquisa busca entender como a empresa pode manter uma imagem positiva mesmo em situações em que o cliente relata problemas. Para isso, foram utilizados dados fornecidos pela empresa, que incluem métricas como tempo de espera, resolução de problemas no primeiro contato, satisfação do cliente e feedback pós-atendimento.

Além disso, a pesquisa tem como objetivos específicos: identificar os principais fatores que influenciam a percepção do cliente sobre a qualidade do atendimento; avaliar a eficácia dos mecanismos de medição de qualidade atualmente utilizados pela empresa; e por último, propor melhorias nos processos de atendimento ao cliente, com base nos dados coletados e analisados.

Para realizar a análise de dados dentro desta abordagem, a metodologia técnica para a sumarização e extração de informações das chamadas pode ser conceitualmente enquadrada como uma tarefa de Compreensão de Texto por Máquina (*Machine Reading Comprehension* - MRC). Conforme proposto por Ma et al. (2022), em vez de uma tarefa de sumarização genérica, a análise é transformada em um processo de pergunta e resposta. O sistema recebe a transcrição do diálogo como 'contexto' e é instruído com 'perguntas' predefinidas (via *prompt*) para extrair os pontos-chave, como 'Qual o problema relatado?' ou 'Qual a solução oferecida?'. Esta técnica permite a extração precisa dos segmentos de informação mais relevantes, evitando ruído no resultado final.

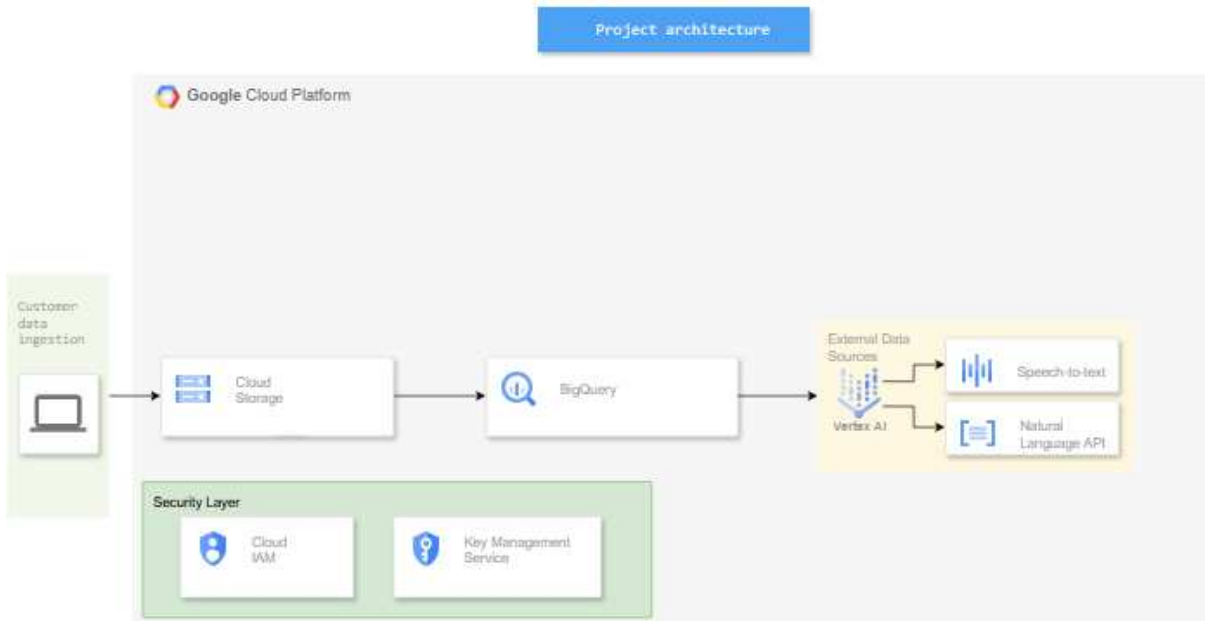
Para implementar esta abordagem de pergunta e resposta, foi utilizada a técnica de **Engenharia de Prompt** (*prompt engineering*). Conforme validado pela literatura, a engenharia de *prompt* é um método eficaz para integrar informações de múltiplas fontes e guiar um modelo de linguagem pré-treinado (como o Gemini) a executar tarefas complexas com resultados satisfatórios, sem a necessidade de um custoso processo de ajuste fino (*fine-tuning*) (XU et al., 2024). O prompt deste trabalho foi cuidadosamente construído para fornecer o contexto da chamada e instruir o modelo a extrair os insights desejados.

Em um mundo no qual a voz do cliente ecoa como um termômetro preciso da qualidade dos serviços, este estudo mergulhou fundo no universo das ligações de um *call center* de uma empresa de energia elétrica, desvendando nuances que, muitas vezes, passam despercebidas no dia a dia, conforme supracitado. Utilizando ferramentas de ponta, como a *API Gemini do Google*, o estudo não só escutou, mas interpretou cada palavra, cada pausa, cada entonação, buscando entender sobre o que realmente importa para quem está do outro lado da linha. O objetivo era claro: entender os fatores que moldam a percepção do cliente sobre o atendimento, avaliar se os métodos atuais de medição de qualidade estão à altura do que a empresa busca, por fim, trazer propostas que transformem o serviço em uma experiência positiva, para o cliente e para a empresa.

A pesquisa começou com a coleta de gravações de ligações (conforme Figura 3), com um mosaico de vozes, reclamações, dúvidas e soluções. Essas interações, que antes eram apenas registros armazenados, ganharam vida ao serem processadas pela *API Gemini*. Não se tratava apenas de transcrever palavras, mas de capturar o que estava entre elas: a frustração de um cliente que esperou com expectativa um serviço prestado, o alívio de outro cujo problema foi resolvido rapidamente, a irritação diante de respostas evasivas, dentre outras.

Abaixo, apresentam-se todas as métricas pré-estabelecidas para a captação e interpretação dos dados, como um roteiro desenhado para guiar a jornada de análise. Esses indicadores revelam os caminhos percorridos para decifrar o que importa nas interações. Cada figura interpretada revela um retrato claro e detalhado da experiência do cliente.

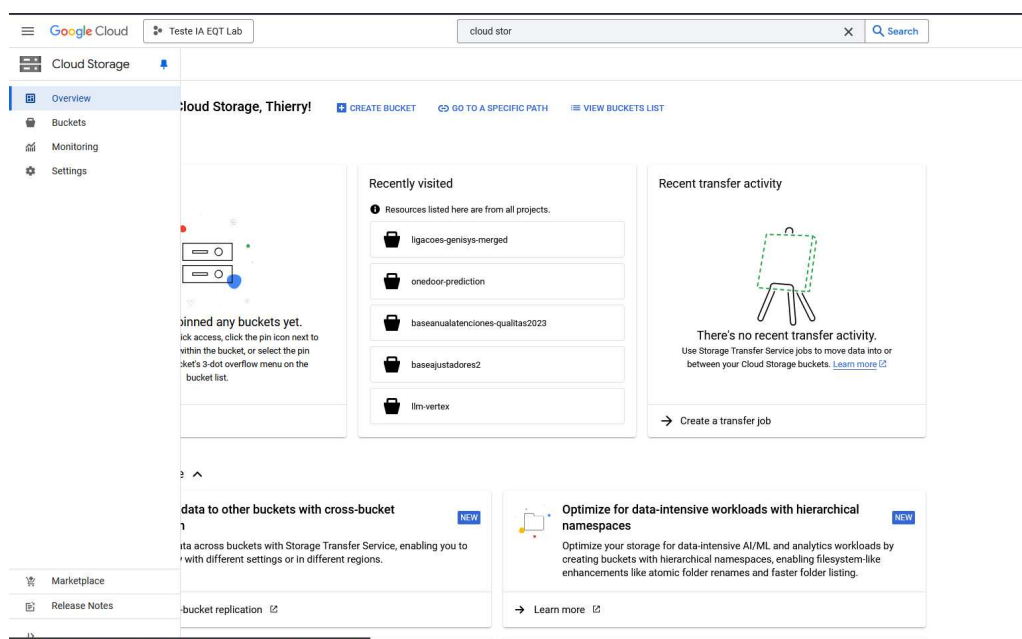
Figura 3: Arquitetura do projeto no *Google Cloud Storage (GCS)*



Fonte: elaborado pelo próprio autor

Primeiramente, todas as ligações foram registradas e atualizadas no *Google Cloud Storage (GCS)*, ferramenta segura e acessível. Esse foi o ponto de partida, o momento em que as vozes dos clientes, carregadas de expectativas, frustrações e soluções, encontraram um “lar temporário” na nuvem, prontas para serem transformadas em insights valiosos.

Figura 4: Interface do GCS



Fonte: elaborado pelo próprio autor

A Figura 4 ilustra a interface principal do *Google Cloud Storage* (GCS), que funciona como o repositório central do projeto. A tela exibe a interface principal do GCS, onde armazenamos arquivos das gravações (objetos) de forma organizada, onde cada um é identificado por um endereço único (URI), garantindo que estejam seguros e permanentemente acessíveis para as etapas seguintes de análise.

Figura 5: Criação do Bucket

**Get Started**

Pick a **globally unique, permanent name**. [Naming guidelines](#)

Ex: 'example', 'example\_bucket-1', or 'example.com'

Tip: Don't include any sensitive information

Optimize storage for data-intensive workloads

Labels (optional)

**CONTINUE**

**Choose where to store your data**

Location: us (multiple regions in United States)  
Location type: Multi-region

**Choose a storage class for your data**

Default storage class: Standard

**Choose how to control access to objects**

Public access prevention: On  
Access control: Uniform

**Choose how to protect object data**

Soft delete policy: Default  
Object versioning: Disabled  
Bucket retention policy: Disabled  
Object retention: Disabled  
Encryption type: Google-managed

**Good to know**

**Location pricing**

Storage rates vary depending on the storage class of your data and location of your bucket. [Pricing details](#)

Current configuration: Multi-region / Standard

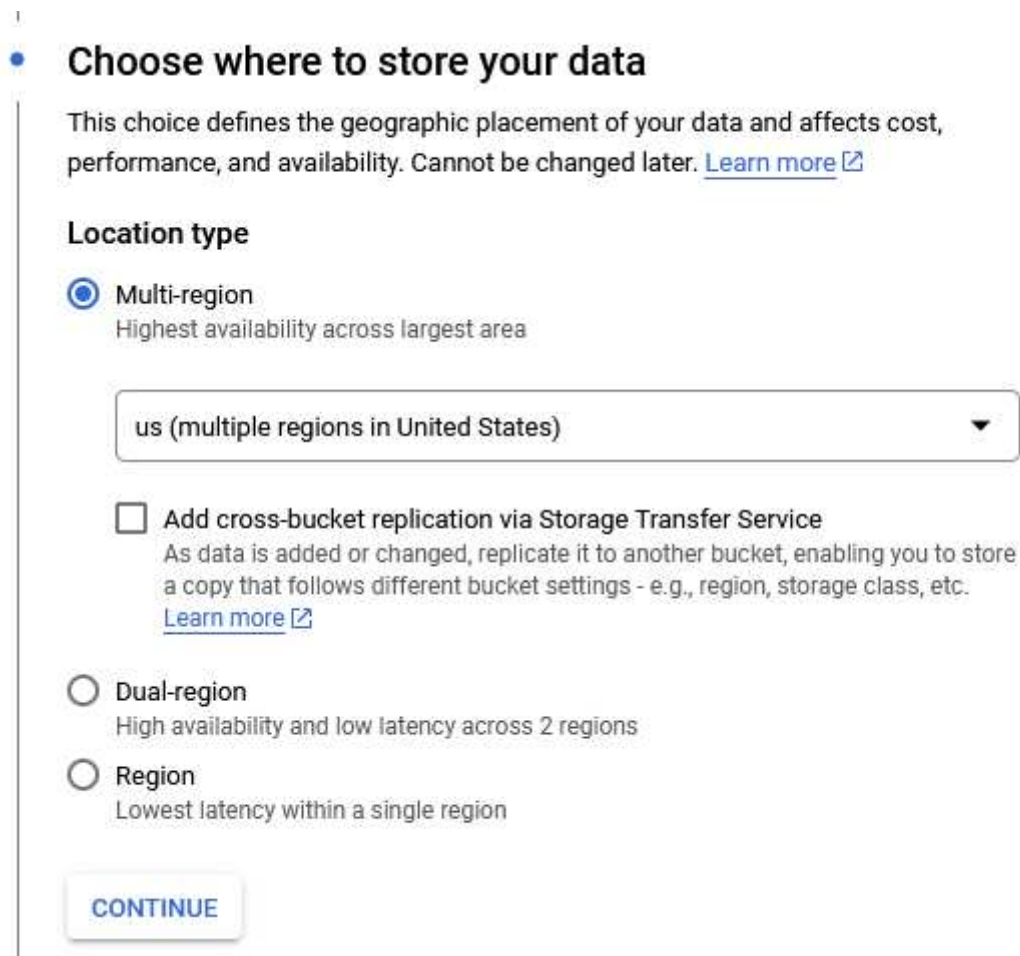
Item	Cost
us (multiple regions in United States)	\$0.026 per GB-month
With default replication	\$0.020 per GB written

**ESTIMATE YOUR MONTHLY COST**

**CREATE** **CANCEL**

Fonte: elaborado pelo próprio autor

A Figura 5 mostra que no GCS é possível criar o que se denomina de *bucket*, um conceito que vai muito além de um simples repositório. É um local digital, customizável e versátil, onde foram definidas características específicas, como permissões de acesso, localização e regras de armazenamento. Cada *bucket* funcionou como uma caixa organizadora, pronta para guardar e proteger os dados, é nele que tudo começou. Este repositório representa o ponto de partida para a análise, armazenando os dados brutos que, subsequentemente, são processados para a geração de indicadores valiosos.

Figura 6: Local de registro do *datacenter*

**Choose where to store your data**

This choice defines the geographic placement of your data and affects cost, performance, and availability. Cannot be changed later. [Learn more](#)

**Location type**

**Multi-region**  
Highest availability across largest area

us (multiple regions in United States)

**Add cross-bucket replication via Storage Transfer Service**  
As data is added or changed, replicate it to another bucket, enabling you to store a copy that follows different bucket settings - e.g., region, storage class, etc. [Learn more](#)

**Dual-region**  
High availability and low latency across 2 regions

**Region**  
Lowest latency within a single region

**CONTINUE**

Fonte: elaborado pelo próprio autor

Neste projeto, foi escolhido o *Multi-Region*, como mostrado na figura 6, com foco principal no *US (multiple regions in United-States)*, uma escolha que se mostrou não apenas viável, mas estratégica. Essa configuração é como uma rede de segurança ampla e resistente, capaz de oferecer alta disponibilidade e uma tolerância a falhas que inspira confiança. Ao distribuir os dados em múltiplas regiões, o sistema cria réplicas e redundâncias, como se fossem *backups* invisíveis espalhados por diferentes locais, garantindo que, mesmo diante de imprevistos, as informações permaneçam seguras e acessíveis.

Adicionalmente, outras configurações foram aplicadas ao *bucket* para garantir a eficiência e a segurança da solução, seguindo as melhores práticas da plataforma. A classe de armazenamento selecionada foi a *Standard*, por ser a mais adequada para dados que necessitam de acesso frequente para as análises propostas. Para o controle de acesso, optou-se pelo modo **Uniforme**, que centraliza as permissões no nível do *bucket* (IAM) e previne o

acesso público aos dados. Por fim, as políticas de proteção foram mantidas no padrão recomendado pelo GCP, utilizando a **exclusão reversível (*soft delete*)** para recuperação de desastres e a **criptografia com chaves gerenciadas pelo Google**, assegurando a integridade e a confidencialidade das informações.

Figura 7: Registros de ligações

Name	Size	Type	Created	Storage class	Last modified
002ce2cc-7805-4c0b-b26c-ccf5a3...	3.1 MB	audio/mpeg	Sep 24, 2024, 9:40:44 AM	Standard	Sep 24, 2024, 9:40:44 AM
00322372-bbf7-4d69-be52-7c1a7...	4.8 MB	audio/mpeg	Sep 24, 2024, 9:40:51 AM	Standard	Sep 24, 2024, 9:40:51 AM
00534e8e-e288-4e82-8301-3da25...	4.3 MB	audio/mpeg	Sep 24, 2024, 9:40:56 AM	Standard	Sep 24, 2024, 9:40:56 AM
005dec6b-bb69-44c7-83f4-7d4f9f...	4.1 MB	audio/mpeg	Sep 24, 2024, 9:40:46 AM	Standard	Sep 24, 2024, 9:40:46 AM
007d4752-aa7e-4bba-b408-37eb3...	6.7 MB	audio/mpeg	Sep 24, 2024, 9:41:01 AM	Standard	Sep 24, 2024, 9:41:01 AM
00870e78-e6a7-4fc9-812d-09d120...	4.7 MB	audio/mpeg	Sep 24, 2024, 9:40:57 AM	Standard	Sep 24, 2024, 9:40:57 AM
00aba829-3393-4364-8846-322c0...	6.6 MB	audio/mpeg	Sep 24, 2024, 9:41:02 AM	Standard	Sep 24, 2024, 9:41:02 AM
014fc135-c70b-4120-8552-8606fd...	5.4 MB	audio/mpeg	Sep 24, 2024, 9:40:59 AM	Standard	Sep 24, 2024, 9:40:59 AM
01516434-ae6d-4ca7-8e77-8b1a2...	3 MB	audio/mpeg	Sep 24, 2024, 9:40:46 AM	Standard	Sep 24, 2024, 9:40:46 AM
018e12af-e16b-4281-84d0-c74d49...	5.3 MB	audio/mpeg	Sep 24, 2024, 9:40:55 AM	Standard	Sep 24, 2024, 9:40:55 AM
01a0da2a-e1a4-4e40-a578-7bbc7...	4.4 MB	audio/mpeg	Sep 24, 2024, 9:41:33 AM	Standard	Sep 24, 2024, 9:41:33 AM
01d2b470-1650-4059-b8fa-08d59...	3.9 MB	audio/mpeg	Sep 24, 2024, 9:41:31 AM	Standard	Sep 24, 2024, 9:41:31 AM
01e25c46-e4ff-4584-ae2b-cae05c...	4.5 MB	audio/mpeg	Sep 24, 2024, 9:41:31 AM	Standard	Sep 24, 2024, 9:41:31 AM
01e38781-d59d-4d55-90f5-67a61...	7.8 MB	audio/mpeg	Sep 24, 2024, 9:41:37 AM	Standard	Sep 24, 2024, 9:41:37 AM
01e6cb35-bdccc-4409-9148-c2791f...	4.1 MB	audio/mpeg	Sep 24, 2024, 9:41:32 AM	Standard	Sep 24, 2024, 9:41:32 AM

Fonte: elaborado pelo próprio autor

Após criar o *Bucket*, o repositório estava pronto para receber as ligações. O cliente enviou um total de **1.774 ligações**, mas, para o projeto, foram **utilizadas 208**, cada uma contendo pelo menos 2 minutos de áudio, selecionadas criteriosamente para garantir a qualidade da análise. Todas as 1.774 ligações ficaram armazenadas no *Google Cloud Storage*, protegidas por uma camada adicional de segurança: os nomes dos arquivos e das gravações foram borrados, substituídos por pseudônimos, para preservar a privacidade e evitar a exposição de dados pessoais. Essa medida reforça o compromisso com a ética e a confidencialidade, garantindo que as informações sejam tratadas com o máximo de responsabilidade.

Cada linha ou faixa acima (Figura 7), representa um arquivo de áudio .mp3, capturando a interação entre atendente e cliente, no qual exibe fragmentos de conversas e seus significados. Após essa etapa no GCS, foi necessário configurar o *BigQuery* (BQ) para

receber esses arquivos e dar continuidade à análise. A transferência para o BQ não foi apenas uma escolha, mas uma necessidade estratégica, já que ele é, por excelência, uma plataforma ideal para análise de dados, um *data warehouse*, onde as informações são organizadas, processadas e transformadas em insights.

Com os dados brutos armazenados no GCS, a etapa seguinte foi a estruturação do ambiente de análise no *BigQuery*. Foi criado um *Dataset* para organizar as tabelas do projeto, como ilustra a **Figura 8**. A localização deste *Dataset* foi mantida na mesma multirregião do *bucket* de origem para garantir a integração e o desempenho entre os serviços.

Figura 8: Criação do Dataset

The image shows a screenshot of the 'Create dataset' form in the Google Cloud console. The form is titled 'Create dataset' and is divided into several sections. The 'Project ID' is 'teste-ia-eqt-lab' with a 'CHANGE' link. The 'Dataset ID' is 'DatasetTeste' with a note that letters, numbers, and underscores are allowed. The 'Location type' is set to 'Multi-region' (selected with a radio button), with a note that it allows BigQuery to select a region within a group to achieve higher quota limits. The 'Multi-region' dropdown is set to 'US (multiple regions in United States)'. The 'External Dataset' section indicates that the selected region supports Cloud Spanner and has a checkbox for 'Link to an external dataset' which is unchecked. There are expandable sections for 'Tags' and 'Advanced options'. At the bottom, there are two buttons: 'CREATE DATASET' (highlighted in blue) and 'CANCEL'.

**Create dataset**

Project ID \*  
teste-ia-eqt-lab [CHANGE](#)

Dataset ID \*  
DatasetTeste  
Letters, numbers, and underscores allowed

Location type ?

Region  
Specify a region to colocate your datasets with other Google Cloud services.

Multi-region  
Allow BigQuery to select a region within a group to achieve higher quota limits.

Multi-region \*  
US (multiple regions in United States) ▼

**External Dataset**  
The selected region supports the following external dataset types: Cloud Spanner

Link to an external dataset ?

**Tags** ▼

**Advanced options** ▼

**CREATE DATASET** CANCEL

Fonte: elaborado pelo próprio autor

O passo fundamental nesta fase foi o estabelecimento de uma conexão externa com o *Vertex AI*, um requisito técnico que habilita o *BigQuery* a invocar os modelos de IA remotamente. A configuração desta conexão, detalhada na Figura 9, permitiu a orquestração do *pipeline* de processamento e a execução de análises complexas diretamente sobre os dados referenciados

Figura 9: Conexão do Vertex AI

The image shows a screenshot of the Google Cloud console interface for configuring an external data source. On the left, there is a sidebar with options like 'Local file', 'Auto-create external and BigLake tables from Cloud Storage', and 'Additional sources'. The main area is titled 'External data source' and contains the following configuration fields:

- Connection type:** A dropdown menu set to 'Vertex AI remote models, remote functions and BigLake (Cloud Resource)'.
- Connection ID \*:** A text input field containing 'vertex-ai'.
- Location type:** Radio buttons for 'Region' and 'Multi-region'. 'Multi-region' is selected.
- Multi-region \*:** A dropdown menu set to 'US (multiple regions in United States)'.
- Friendly name:** An empty text input field.
- Description:** An empty text input field.

Fonte: elaborado pelo próprio autor

Uma vez configurado o ambiente de nuvem, os arquivos de áudio armazenados no GCS foram mapeados no *BigQuery* por meio da criação de uma tabela externa denominada *recordings* (Figura 10). Essa técnica viabiliza a análise direta dos metadados dos áudios sem a necessidade de movê-los, funcionando como uma ponte entre o repositório de arquivos e o ambiente analítico. O script SQL utilizado para esta etapa está documentado no Apêndice A (Código SQL 1).

Figura 10: Tabela Recordings

1 SELECT \* FROM `teste-ia-eqt-lab.genesis.recordings200` LIMIT 1000

Query results

JOB INFORMATION	RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH		
Row	uri	generation	content_type	size	md5_hash	updated	metada
1	gs://ligacoes-genisys-merged/0...	17271816445800...	audio/mpeg	3091821	c59c84ac77653912cafd9ee1ee...	2024-09-24 12:40:44.617000 UTC	null
2	gs://ligacoes-genisys-merged/0...	17271816515155...	audio/mpeg	4801389	06d25928fb0df832b4f3e34d7a...	2024-09-24 12:40:51.553000 UTC	null
3	gs://ligacoes-genisys-merged/0...	17271816564157...	audio/mpeg	4311405	6656afe9e25a12f4d768a307b8...	2024-09-24 12:40:56.455000 UTC	null
4	gs://ligacoes-genisys-merged/0...	17271816468797...	audio/mpeg	4102125	9607d4ec6fc7f3f50a2ce6ca47...	2024-09-24 12:40:46.918000 UTC	null
5	gs://ligacoes-genisys-merged/0...	17271816609654...	audio/mpeg	6651885	005d73464046d4aae697b4c55...	2024-09-24 12:41:01.004000 UTC	null
6	gs://ligacoes-genisys-merged/0...	17271816570308...	audio/mpeg	4737261	05ab159c1d3715c773ac2bda4...	2024-09-24 12:40:57.069000 UTC	null
7	gs://ligacoes-genisys-merged/0...	17271816621299...	audio/mpeg	6605421	6a692e5ea8693aeb9381b1397...	2024-09-24 12:41:02.167000 UTC	null
8	gs://ligacoes-genisys-merged/0...	17271816593255...	audio/mpeg	5375085	dcf583aea69b18539b7f3c64d3...	2024-09-24 12:40:59.363000 UTC	null
9	gs://ligacoes-genisys-merged/0...	17271816464454...	audio/mpeg	3018477	3e636a318cada1b5ebd430403...	2024-09-24 12:40:46.483000 UTC	null
10	gs://ligacoes-genisys-merged/0...	17271816553178...	audio/mpeg	5305581	71cddec7e5e2b2fa0e194d91db...	2024-09-24 12:40:55.356000 UTC	null
11	gs://ligacoes-genisys-merged/0...	17271816930472...	audio/mpeg	4439661	4528c824d302f0a5e2192e0ebe...	2024-09-24 12:41:33.086000 UTC	null
12	gs://ligacoes-genisys-merged/0...	17271816912749...	audio/mpeg	3889389	7447dac41f58d86f103c7df8a3...	2024-09-24 12:41:31.312000 UTC	null
13	gs://ligacoes-genisys-merged/0...	17271816912776...	audio/mpeg	4469613	54470793a0d95704e53f8e844...	2024-09-24 12:41:31.332000 UTC	null
14	gs://ligacoes-genisys-merged/0...	17271816970474...	audio/mpeg	7765485	f39d87cf8eb6951f0db0f103324...	2024-09-24 12:41:37.086000 UTC	null
15	gs://ligacoes-genisys-merged/0...	17271816924538...	audio/mpeg	4138605	0763cb11eb3736d170d453ae8...	2024-09-24 12:41:32.492000 UTC	null
16	gs://ligacoes-genisys-merged/0...	17271816960099...	audio/mpeg	3511917	6aea3779aeeca839c048deeb7...	2024-09-24 12:41:36.048000 UTC	null
17	gs://ligacoes-genisys-merged/0...	17271816932814...	audio/mpeg	4146669	ad8b69cf7d98e64f295aa77659...	2024-09-24 12:41:33.320000 UTC	null
18	gs://ligacoes-genisys-merged/0...	17271816938156...	audio/mpeg	5445357	04baf1e317f2233dd0ebce197e...	2024-09-24 12:41:33.853000 UTC	null
19	gs://ligacoes-genisys-merged/0...	17271816925723...	audio/mpeg	4163181	7bd79fc613c1652f8394aa4d96...	2024-09-24 12:41:32.610000 UTC	null
20	gs://ligacoes-genisys-merged/0...	17271817161262...	audio/mpeg	4193133	642f89a6ffd9380595cfff23132...	2024-09-24 12:41:56.165000 UTC	null

Fonte: elaborado pelo próprio autor

Com as informações dos áudios devidamente organizadas e a tabela *recordings* criada para referenciar todos os arquivos armazenados no GCS, chegou o momento de rodar a consulta que irá enviar esses arquivos para o *Vertex AI*. Cada gravação será processada com base nos *inputs* e *prompts* definidos. O *Vertex AI* atuou como um especialista, examinando cada detalhe das interações e retornando insights precisos, transformando áudios brutos em informações estruturadas e acionáveis. Essa etapa é o coração da análise, onde a tecnologia se encontra com a necessidade de entender, de forma profunda, o que cada ligação revela sobre a experiência do cliente. Para essa etapa utilizamos o script detalhado no Apêndice A (Código SQL 2). A execução desta consulta gerou a tabela intermediária “*output\_vertex*”, que armazena a saída bruta do modelo em formato JSON. O

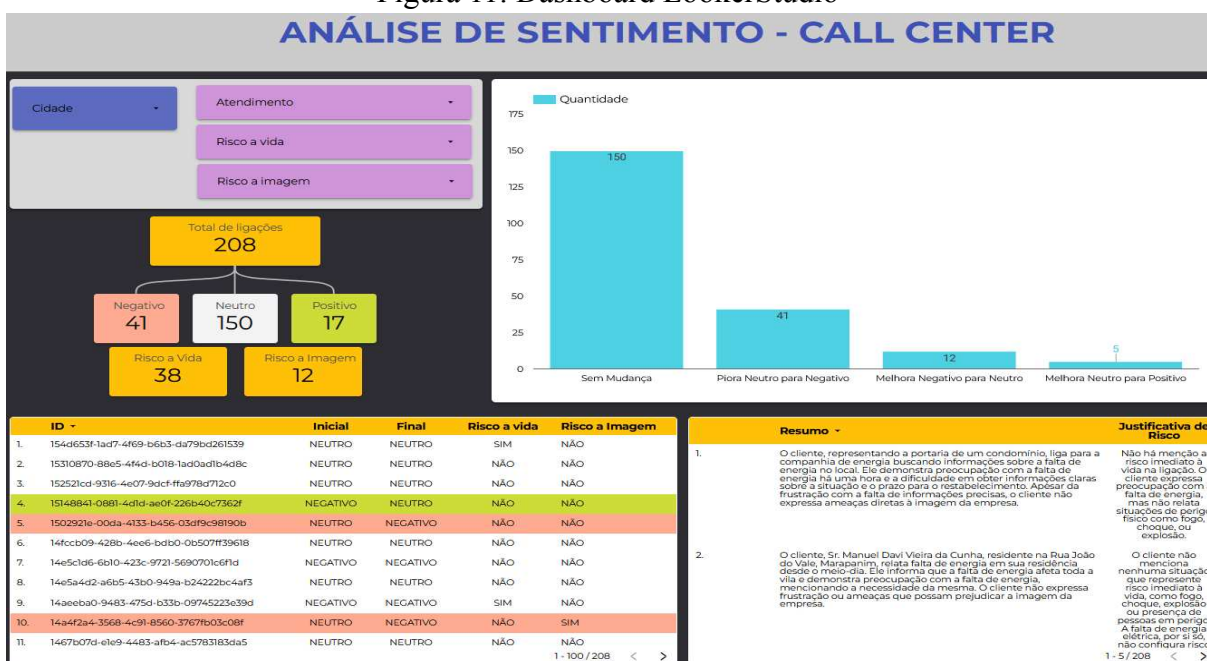
dicionário de dados detalhado para esta tabela encontra-se no Apêndice B (Tabela Output Vertex).

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Apresentação dos resultados experimentais

A análise do conjunto de 208 ligações resultou na seguinte distribuição de sentimentos: 150 chamadas (72,1%) foram classificadas como 'Neutro', 41 (19,7%) como 'Negativo' e 17 (8,2%) como 'Positivo'. Adicionalmente, o sistema identificou 38 interações com potencial 'Risco à Vida' e 12 com 'Risco à Imagem' da empresa. Esses dados quantitativos, consolidados no painel de BI detalhado na figura 11, oferecem uma visão macro da natureza das interações no *call center*.

Figura 11: Dashboard LookerStudio



Fonte: elaborado pelo próprio autor

As predições do modelo e os dados extraídos foram organizados em uma tabela analítica final, permitindo a exploração detalhada de cada interação. O dicionário de dados completo, descrevendo cada campo desta tabela e das tabelas intermediárias, está disponível para consulta no **Apêndice B**.

#### 4.2 Protocolo e Resultados da Validação do Modelo

Para validar a eficácia e a aplicabilidade prática da solução, foi conduzida uma análise de concordância entre os resultados gerados pela IA e a avaliação realizada por analistas humanos. O protocolo de validação foi executado sobre um conjunto de 200 ligações, amostradas aleatoriamente do universo total.

O processo ocorreu em duas frentes paralelas:

1. O conjunto de 200 áudios foi submetido à ferramenta de IA desenvolvida, que realizou a classificação de sentimento e a sumarização de forma automática.
2. Independentemente, o mesmo conjunto de 200 áudios foi distribuído a analistas da própria empresa, que aplicaram seus conhecimentos de negócio para classificar o sentimento de cada chamada manualmente.

Após a conclusão das duas análises, foi realizada uma etapa de **conferência manual**, na qual os resultados da IA foram comparados, registro por registro, com os resultados fornecidos pelos analistas humanos. A métrica de avaliação utilizada foi a **taxa de concordância (ou acurácia)**, que representa a porcentagem de casos em que a classificação de sentimento da IA foi idêntica à classificação do especialista humano. Nesta análise comparativa, o sistema de IA obteve uma **taxa de concordância de 90%** com as classificações dos analistas. Este resultado expressivo indica que, em 9 de cada 10 chamadas, a ferramenta foi capaz de interpretar o sentimento da interação da mesma forma que um profissional treinado. Este alto índice de alinhamento valida a ferramenta não apenas como um modelo tecnicamente preciso, mas como um suporte operacional confiável, capaz de replicar com grande fidelidade a avaliação humana no contexto específico do *call center* da empresa.

#### 4.3 Discussão dos Achados

A acurácia de 90% demonstra a viabilidade da aplicação de LLMs, via engenharia de prompt, para a análise de interações em um ambiente complexo de *call center*. Este resultado é particularmente relevante no contexto discutido por (PLAZA; PAWLIK; DENIZIAK, 2021), que ressalta os desafios na transcrição e análise de áudio de centrais de atendimento. A solução proposta consegue superar parte desses desafios, entregando análises consistentes e com alto grau de confiabilidade, gerando valor direto para o negócio. Adicionalmente, a análise qualitativa dos resumos gerados revelou que a percepção de qualidade do cliente transcende a simples resolução do problema. A falta de personalização e as respostas protocolares, identificadas em diversas chamadas com sentimento negativo, emergiram como fatores críticos. Este achado reforça que a qualidade do atendimento está no 'como' a interação é conduzida, e não apenas nos indicadores quantitativos tradicionais, como o tempo de espera.

A distribuição de sentimentos (Figura 11) revela que a vasta maioria das chamadas (72,1%) é de natureza neutra, enquanto uma parcela significativa (19,7%) é negativa. Isso sugere que, embora muitas interações sejam rotineiras, há um volume considerável de clientes insatisfeitos cujo atendimento pode ser priorizado pela ferramenta, gerando valor direto para o negócio.

### ***4.3 Implicações, Limitações e Trabalhos Futuros***

As implicações práticas desta solução incluem a otimização da auditoria de qualidade, a identificação proativa de problemas sistêmicos e a melhoria na priorização de atendimentos críticos. Como principal limitação, a validação de 90% foi baseada em uma métrica de acurácia global fornecida pela equipe de analistas, não sendo possível, no escopo deste trabalho, detalhar a performance (precisão e recall) para cada classe de sentimento individualmente. Adicionalmente, o modelo não foi ajustado para nuances complexas como sarcasmo ou trotes.

Para trabalhos futuros, sugere-se o aprofundamento da análise de validação com a construção de uma matriz de confusão detalhada e o *fine-tuning* do modelo com um léxico específico da empresa para aprimorar ainda mais a precisão em casos de nicho.

## 5 CONCLUSÃO

Este estudo demonstrou a viabilidade e a eficácia de uma metodologia de análise de ponta a ponta, utilizando um Modelo de Linguagem de Grande Porte (LLM) na plataforma GCP. Ao enquadrar a análise como uma tarefa de Compreensão de Texto por Máquina (MRC), conforme proposto por Ma et al. (2022), e ao empregar Engenharia de Prompt, este trabalho não apenas transcreveu, mas interpretou com 90% de acurácia o sentimento, o problema principal e o resumo de chamadas em um ambiente de *call center*, transformando dados não estruturados de voz em inteligência de negócio acionável.

A arquitetura proposta com os áudios armazenados no *Cloud Storage*, transcritos e analisados no *BigQuery*, e enriquecidos pelo *BigQuery ML* (BQML) e *Vertex AI*, mostrou-se eficiente. Cada etapa foi pensada para garantir que os dados fluíssem sem obstáculos. A criação de um painel no *Looker Studio*, transformando números e textos em visualizações claras, prontas para guiar a tomada de decisões estratégicas, foi de extrema importância.

Outro fato importante contido no estudo, foi a validação da precisão do modelo de IA, pois quando comparado à análise manual de 200 ligações, o modelo alcançou incríveis 90% de concordância. Esse resultado não só confirmou a eficácia da IA, mas também mostrou que ela pode ser uma ferramenta confiável, complementando e até aprimorando o trabalho humano.

A análise revelou que a qualidade do atendimento não se resume a números frios, como tempo de espera ou duração da chamada. Ela é, na verdade, uma teia delicada, onde a clareza na comunicação, a empatia no tom de voz e a personalização do atendimento se entrelaçam para criar a conexão de confiança do cliente. Adicionalmente, conclui-se que respostas robotizadas afastam e frustram o cliente, enquanto um atendimento humanizado traz conforto e conexão, frente ao cliente que muitas vezes está nervoso ou com algum problema grave.

Com base nesses insights, é possível propor melhorias que podem transformar o *call center*. Treinamentos focados em comunicação e empatia, ferramentas de análise em tempo real e métricas que vão além do quantitativo são alguns dos caminhos sugeridos. Essas mudanças têm o potencial de não apenas melhorar o atendimento, mas também de fortalecer o vínculo com o cliente, mostrando que a empresa não apenas ouve, mas entende e se importa.

Adicionalmente, a manutenção e o aprimoramento de sistemas conversacionais em um ambiente corporativo envolvem um extenso 'trabalho de articulação' (*articulation work*), que frequentemente é invisível. Esse trabalho inclui a colaboração entre curadores de

conteúdo e equipes técnicas, a negociação com gestores de produtos e a validação de respostas para garantir a consistência com as políticas da empresa (CANDELLO et al., 2022). Portanto, o sucesso de uma ferramenta de análise de chamadas depende não apenas de sua precisão técnica, mas também de sua integração nos fluxos de trabalho e colaboração humanos existentes.

Desta forma, conclui-se que o estudo atingiu seus objetivos. Este trabalho, portanto, não é apenas um diagnóstico técnico, mas um convite para que a organização encare o desafio de transformar cada ligação em uma oportunidade de surpreender e extrair insights essenciais para a melhoria contínua de seus serviços.

## REFERÊNCIAS

- ALMEIDA, T. A.; HADDAD, D. B. **Processamento de Sinais de Áudio: Técnicas e Aplicações**. São Paulo: Editora Blucher, 2018.
- BARBOSA, Alisson S.; CARVALHO, Paulo L.; LOPES, Paulo H. S. **Procedimentos para Aplicação de Penalidade por Violação dos Padrões dos Indicadores de Continuidade DEC e FEC**. In: SEMINÁRIO BRASILEIRO SOBRE A QUALIDADE DA ENERGIA ELÉTRICA, VI, Belém, 2005.
- COSTA, M. C.; SILVA, J. M. **Inteligência Artificial e Aprendizado de Máquina Aplicados ao Reconhecimento de Fala**. In: CONGRESSO BRASILEIRO DE INTELIGÊNCIA COMPUTACIONAL (CBIC), 2020, Porto Alegre. Anais. Porto Alegre: Sociedade Brasileira de Inteligência Computacional, 2020. p. 123-134.
- FREITAS, L. G.; SOUZA, R. M. Desafios e Avanços no Reconhecimento Automático de Fala para o Português Brasileiro. **Revista de Tecnologia da Informação e Comunicação**, v. 10, n. 2, p. 45-58, 2019.
- LIMA, A. R.; FERREIRA, P. J. S. Privacidade e Segurança em Sistemas de Reconhecimento de Fala: Uma Análise sob a Perspectiva da LGPD. **Revista Brasileira de Segurança da Informação**, v. 8, n. 3, p. 22-35, 2022.
- MUELLER, Suzana Pinheiro Machado; PERUCCHI, Valmira. Universidades e a produção de patentes: tópicos de interesse para o estudioso da informação tecnológica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 2, p. 15-36, 2014.
- SOUZA, A. M.; RIBEIRO, V. H. A Evolução das Técnicas de Processamento de Áudio e seu Impacto na Indústria 4.0. **Revista de Engenharia e Tecnologia**, v. 12, n. 4, p. 33-47, 2020.
- PLAZA, M.; PAWLIK, Ł.; DENIZIAK, S. Call transcription methodology for Contact Center systems. **IEEE Access**, v. 9, p. 114623-114635, 2021. DOI: 10.1109/ACCESS.2021.3102502.
- SAILAJA, N. V. et al. Advancing Audio Processing and Emotion Recognition through Deep Learning Techniques. **SSRG International Journal of Electrical and Electronics Engineering**, v. 12, n. 4, p. 298-307, abr. 2025.
- MA, B. et al. Extractive Dialogue Summarization Without Annotation Based on Distantly Supervised Machine Reading Comprehension in Customer Service. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 30, p. 87-97, 2022. DOI: 10.1109/TASLP.2021.3133206.
- ANH, N. M. T.; HO, T. S. Improving Speech Recognition with Prompt-based Contextualized ASR and LLM-based Re-predictor. In: INTERSPEECH 2024, 2024, Kos. **Proceedings...** p. 737-741. DOI: 10.21437/Interspeech.2024-1762.
- XU, Y. et al. Multimodal Emotion Captioning Using Large Language Model with Prompt Engineering. In: 2ND INTERNATIONAL WORKSHOP ON MULTIMODAL AND

RESPONSIBLE AFFECTIVE COMPUTING (MRAC '24), 2024, Melbourne. **Proceedings...** New York: ACM, 2024. p. 104-109. DOI: <https://doi.org/10.1145/3689092.3689403>.

RIZVI, M. A. H. et al. Sentiment Analysis of Incoming Calls on Helpdesk. In: 2025 3RD INTERNATIONAL CONFERENCE ON DISRUPTIVE TECHNOLOGIES (ICDT), 2025. **Proceedings...** p. 1188-1193. DOI: 10.1109/ICDT63985.2025.10986341.

NITU, N.; CATRUNA, A.; RADOI, E. Romanian Speech-to-Text Transcription for Medical Applications. In: IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT COMPUTER COMMUNICATION AND PROCESSING (ICCP), 20., 2024, [S.l.]. **Proceedings...** [S.l.]: IEEE, 2024.

YADAV, A.; VISHWAKARMA, D. K. Sentiment analysis using deep learning architectures: a review. **Artificial Intelligence Review**, v. 53, p. 4335–4385, 2020.

DO, V. H.; MAI, V. T. Agent/Client Speech Identification for Mixed-Channel Conversation in Customer Service Call Centers. In: 2020 INTERNATIONAL CONFERENCE ON ASIAN LANGUAGE PROCESSING (IALP), 2020, Kuala Lumpur. **Proceedings...** p. 197-200. DOI: 10.1109/IALP51396.2020.9310484.

CANDELLO, H. et al. Unveiling Practices of Customer Service Content Curators of Conversational Agents. **Proceedings of the ACM on Human-Computer Interaction**, v. 6, n. CSCW2, art. 348, p. 1-33, nov. 2022. DOI: <https://doi.org/10.1145/3555768>.

FENG, Y.; DEVILLERS, L. End-to-End Continuous Speech Emotion Recognition in Real-life Customer Service Call Center Conversations. In: 2023 11TH INTERNATIONAL CONFERENCE ON AFFECTIVE COMPUTING AND INTELLIGENT INTERACTION WORKSHOPS AND DEMOS (ACIIW), 2023, Boston. **Proceedings...** DOI: 10.1109/ACIIW59127.2023.10388120.

ABI KANAAN, M. et al. A methodology for emergency calls severity prediction: from pre-processing to BERT-based classifiers. In: ARTIFICIAL INTELLIGENCE APPLICATIONS AND INNOVATIONS CONFERENCE, 2023, Leon, Spain. **Proceedings...** Springer Nature Switzerland, 2023. p. 331-345.

LIU, R. et al. Generative Expressive Conversational Speech Synthesis. In: ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA, 32., 2024, Melbourne. **Proceedings...** New York: ACM, 2024.

GOOGLE. Introdução aos Ajustes. [S.l.]: Google, 2025. Disponível em: <http://cloud.google.com/vertex-ai/generative-ai/docs/models/tune-models>. Acesso em: 24 jul. 2025.

GOOGLE. Documentação do Google Cloud. [S.l.]: Google, 2025. Disponível em: <https://cloud.google.com/docs>. Acesso em: 25 jul. 2025.

GOOGLE. Visão geral do Cloud Storage. [S.l.]: Google, 2025. Disponível em: <https://cloud.google.com/storage/docs/overview>. Acesso em: 25 jul. 2025.

GOOGLE. O que é o BigQuery?. [S.l.]: Google, 2025. Disponível em: <https://cloud.google.com/bigquery/docs/introduction>. Acesso em: 25 jul. 2025.

GOOGLE. Introdução ao BigQuery ML. [S.l.]: Google, 2025. Disponível em: <https://cloud.google.com/bigquery-ml/docs/introduction>. Acesso em: 25 jul. 2025.

GOOGLE. Visão geral da Vertex AI. [S.l.]: Google, 2025. Disponível em: <https://cloud.google.com/vertex-ai/docs/start/introduction-unified-platform>. Acesso em: 26 jul. 2025.

GOOGLE. Este é o Looker Studio. [S.l.]: Google, 2025. Disponível em: <https://cloud.google.com/looker/docs/studio>. Acesso em: 27 jul. 2025.

## APÊNDICE A – SCRIPTS E PROMPTS DE ANÁLISE

### A.1 Introdução

Este documento tem por propósito documentar todos os scripts utilizados para as análises e trabalho feito.

Código SQL 1.

```
CREATE OR REPLACE EXTERNAL TABLE `teste-ia-eqt-lab.DatasetTeste.recordings`
WITH CONNECTION `projects/teste-ia-eqt-lab/locations/us/connections/vertex-ai`
OPTIONS(
  object_metadata = 'SIMPLE',
  uris = ['gs://ligacoes-genisys-merged/*'],
  max_staleness = INTERVAL 1 DAY,
  metadata_cache_mode = 'MANUAL'); ##Set to AUTOMATIC for the metadata cache to
be refreshed at a system-defined interval, usually somewhere between 30 and 60 minutes.
```

Código SQL 2.

```
-- resultado do modelo pro
create or replace table `teste-ia-eqt-lab.DatasetTeste.output_vertex` AS
SELECT *
FROM ML.GENERATE_TEXT(
  MODEL `teste-ia-eqt-lab.DatasetTeste.vertexaigemini15pro`,
  TABLE `teste-ia-eqt-lab.DatasetTeste.recordings`,
  STRUCT(
    0.2 AS temperature,
    2048 AS max_output_tokens,
    0.2 AS top_p,
    15 AS top_k,
    TRUE AS flatten_json_output,
    """"
    Role: You are a contact center manager.
    Task: Task: Transcribe the following audio file, clearly distinguishing between the agent
and the customer by labeling each speaker accordingly.
    Respond in Portuguese.
    Output your response strictly in JSON format with the following keys:
    {
    "transcription": [
```

```

{"agent": "Primeira fala do agente."},
{"customer": "Resposta do cliente."},
// Continue for each exchange
],
"positive_feedback": "Comentários positivos identificados.",
"negative_feedback": "Comentários negativos identificados.",
"sentiment": "Sentimento geral da conversa.",
"summary": "Resumo da interação."
}
Do not include any additional text outside of the JSON structure.
"" AS prompt
)
)
;

```

## APÊNDICE B – DICIONÁRIO DE DADOS DAS TABELAS GERADAS

### B.1 Introdução

Este apêndice tem por propósito documentar o dicionário de dados das tabelas geradas no trabalho.

Tabela Output Vertex:

Figura 12: Output Vertex

Row	ml_generate_text_llm_result	ml_generate_text_llm_result	ml_generate_text_status	url	generation	content_type	size	md5_hash	updated	metadata name	metadata value
10	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182064...	audio/mpeg	2888645	9c8a77852ab226fac932380...	2024-09-24 12:47:44.512000 UTC	mlfile	mlfile
11	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182105...	audio/mpeg	2899821	6ca57882243a52a6684205d...	2024-09-24 12:48:26.018000 UTC	mlfile	mlfile
12	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182172...	audio/mpeg	2905197	26068676956f6c627ac8002...	2024-09-24 12:49:32.345000 UTC	mlfile	mlfile
13	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181975...	audio/mpeg	2918637	21f313965492982dbbdaa3d21...	2024-09-24 12:46:15.875000 UTC	mlfile	mlfile
14	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182004...	audio/mpeg	2942829	62f4ada81b54874bc0ff5d92e...	2024-09-24 12:46:44.087000 UTC	mlfile	mlfile
15	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182271...	audio/mpeg	2999277	8e45a4164585921bc5024487...	2024-09-24 12:52:51.673000 UTC	mlfile	mlfile
16	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182368...	audio/mpeg	3004653	02671c4a226ec3877623031b...	2024-09-24 12:52:48.115000 UTC	mlfile	mlfile
17	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181646...	audio/mpeg	3018477	3e636a318cda1b5eb430403...	2024-09-24 12:40:46.480000 UTC	mlfile	mlfile
18	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182210...	audio/mpeg	3063255	57a295a1fb71f4e7720c996...	2024-09-24 12:50:10.748000 UTC	mlfile	mlfile
19	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181870...	audio/mpeg	3075693	6a75464efc263a70b08f55ab...	2024-09-24 12:44:30.840000 UTC	mlfile	mlfile
20	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182035...	audio/mpeg	3084641	e4313a4876a3b70b9b0c0f9...	2024-09-24 12:47:15.414000 UTC	mlfile	mlfile
21	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181644...	audio/mpeg	3091821	c59c84ac77653913ca1f9ee...	2024-09-24 12:40:44.617000 UTC	mlfile	mlfile
22	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182012...	audio/mpeg	3095277	f788a07204c12483d3ac969...	2024-09-24 12:46:53.026000 UTC	mlfile	mlfile
23	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182322...	audio/mpeg	3106797	6004fa7246608f3a5c54c0e7...	2024-09-24 12:52:02.799000 UTC	mlfile	mlfile
24	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727182321...	audio/mpeg	3118317	964fa9b8a4be575aa9cd1eac8...	2024-09-24 12:52:01.480000 UTC	mlfile	mlfile
25	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181770...	audio/mpeg	3172845	7f93a36500ca4f89a35166ca8e...	2024-09-24 12:42:50.947000 UTC	mlfile	mlfile
26	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181734...	audio/mpeg	3175149	51be965640227936f7922393...	2024-09-24 12:48:54.480000 UTC	mlfile	mlfile
27	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181740...	audio/mpeg	3222381	6a48bca788a6651a1e16a1ef3...	2024-09-24 12:42:20.777000 UTC	mlfile	mlfile
28	["json ( "sentimento_inicial":...	["category": "probabilidade": "pro...		gs://lgacces-genisys-emerged.0...	1727181969...	audio/mpeg	3241581	93a8c48f11ba196642d8a896b...	2024-09-24 12:46:09.771000 UTC	mlfile	mlfile

Fonte: elaborado pelo próprio autor

- **Row**
  - Indica o número (ou ID) da linha/tabela. Geralmente é um índice que facilita a identificação de cada registro no BigQuery.
- **ml\_generate\_text\_llm\_result**
  - Armazena o resultado textual gerado pelo modelo de linguagem (LLM), em formato JSON.
  - No seu caso, esse JSON inclui campos como:
    - **sentimento\_inicial**: sentimento do cliente no início da ligação (e.g., NEGATIVO, NEUTRO, POSITIVO).

- **sentimento\_final**: sentimento do cliente ao final da ligação.
- **risco\_a\_vida**: se a situação envolve ou não risco de vida.
- **risco\_vida\_medida**: um valor numérico (por exemplo, “1”) que pode representar um nível de risco (pode ser parte da lógica do modelo).
- **risco\_imagem**: se há risco para a imagem da empresa.
- **summary**: um resumo da chamada.
- **ml\_generate\_text\_rai\_result**
  - Traz um array de objetos JSON, provavelmente relacionado a “Responsible AI” (RAI) ou alguma forma de classificação adicional.
- **ml\_generate\_text\_status**
  - Indica o status do processamento de texto gerado pelo modelo de linguagem (por exemplo, se foi concluído, se deu erro, etc.).
- **uri**
  - Endereço (URI) do arquivo de áudio no Google Cloud Storage (GCS).
  - Geralmente no formato `gs://nome-do-bucket/nome-do-arquivo.mp3`.
- **generation**
  - Pode indicar o ID ou a versão da “geração” do arquivo, ou alguma referência do processo de ingestão ou processamento.
- **content\_type**
  - Tipo de conteúdo do arquivo. No caso, “audio/mpeg” indica que são arquivos de áudio no formato MP3.
- **size**
  - Tamanho do arquivo em bytes.
  - Útil para verificação de integridade ou para fins de auditoria e consumo de storage.
- **md5\_hash**
  - Hash MD5 do arquivo, geralmente utilizado para verificação de integridade (confirma que o arquivo não foi corrompido durante o upload/download).
- **updated**
  - Data e hora em que o arquivo (ou o registro) foi atualizado no GCS ou no BigQuery.
  - Formato padrão UTC, permitindo saber quando aquele registro foi modificado pela última vez.
- **metadata.name**
  - Campo que pode armazenar chaves de metadados adicionais do arquivo ou do processo de ingestão.
- **metadata.value**
  - Complementa o `metadata.name`, armazenando o valor correspondente ao nome do metadado.