

UNIVERSIDADE PRESBITERIANA MACKENZIE

Faculdade de Computação e Informática

Prof. Dr. Leandro Zerbinatti

Inteligência Artificial – 7ºJ SI - Noite

Projeto da Disciplina Aplicação de Inteligência Artificial em Problema Real

1) Título: Sistema de transcrição e categorização de áudio

2) Integrantes: Gustavo Detoni, 10403471; Felipe Wurcker Goe Mazuca, 10401979

3) Resumo

O projeto busca aprimorar a eficiência e a qualidade do atendimento telefônico em setores como Help Desk, Service Desk, Call Centers e Inside Sales. Para isso, serão empregadas tecnologias de Inteligência Artificial voltadas para a transcrição de áudios em texto e categorização das interações, permitindo um monitoramento detalhado das ligações. O objetivo é gerar insights que contribuam para a otimização dos processos, aumentando a produtividade das equipes e aprimorando a experiência do cliente.

4) Introdução

a. Contextualização;

O atendimento ao cliente desempenha um papel essencial na experiência do consumidor e na eficiência operacional das empresas. Setores como Help Desk, Service Desk, Call Centers e Inside Sales lidam diariamente com um grande volume de interações, onde a qualidade do atendimento pode impactar diretamente a satisfação do cliente e a produtividade das equipes. Com o avanço da tecnologia, novas abordagens têm sido exploradas para aprimorar esses serviços, tornando-os mais ágeis e personalizados.

b. Justificativa;

As novas soluções inovadoras para otimizar o atendimento telefônico se tornam cada vez mais necessárias, considerando os desafios enfrentados por essas equipes, como a alta demanda, dificuldades na comunicação e a necessidade de melhorias contínuas nos processos. O uso de Inteligência Artificial oferece oportunidades para monitorar e analisar as interações de forma automatizada, fornecendo insights valiosos para aprimorar a qualidade do serviço prestado.

c. Objetivo;

O projeto tem como objetivo implementar tecnologias de Inteligência Artificial para transcrição e categorização de interações telefônicas, permitindo um acompanhamento mais eficiente das ligações. A partir disso, busca-se identificar padrões, otimizar processos e promover melhorias contínuas no atendimento, aumentando a satisfação dos clientes e a produtividade das equipes envolvidas.

d. Opção do projeto;

Para alcançar esse objetivo, serão utilizadas ferramentas de Inteligência Artificial especializadas na conversão de áudio em texto e na análise automatizada das interações. Isso permitirá um monitoramento detalhado das ligações, fornecendo informações estratégicas para a otimização do atendimento. Dessa forma, o projeto visa não apenas aprimorar a eficiência dos serviços, mas também impulsionar a inovação no setor de atendimento ao cliente.

5) Descrição do Problema

Os serviços de atendimento telefônico desempenham um papel fundamental na comunicação entre empresas e clientes, sendo amplamente utilizados em setores como Help Desk, Service Desk, Call Centers e Inside Sales. No entanto, esses atendimentos frequentemente apresentam desafios que comprometem sua eficiência e qualidade. Problemas como falhas na comunicação, demora na resolução das demandas e falta de personalização são recorrentes e impactam negativamente tanto a experiência do cliente quanto a produtividade das equipes.

Outro fator crítico é a dificuldade em monitorar e avaliar de forma eficaz as interações realizadas por telefone. Muitas empresas ainda utilizam métodos manuais ou avaliações subjetivas para acompanhar a qualidade do atendimento, o que limita a identificação de padrões e a implementação de melhorias. Diante da crescente necessidade de um atendimento mais ágil e assertivo, torna-se essencial a adoção de tecnologias que permitam a análise automatizada dessas interações, viabilizando aprimoramentos contínuos e garantindo um serviço mais eficiente e satisfatório.

6) Discutir a respeito dos aspectos Éticos do uso da IA e a sua Responsabilidade no desenvolvimento da solução

O uso da Inteligência Artificial no nosso projeto, voltado para o monitoramento e análise de ligações telefônicas em um ambiente de atendimento ao cliente, levanta questões éticas fundamentais que precisam ser cuidadosamente consideradas. Como estamos lidando com interações reais entre clientes e atendentes, a privacidade e a proteção de dados são aspectos cruciais. Garantir que as gravações e transcrições respeitem normas como a LGPD é essencial para evitar o uso indevido dessas informações, garantindo que todos os envolvidos estejam cientes e tenham consentido com essa análise.

Conforme destacado por Trindade e Oliveira (2024), a implementação de sistemas de monitoramento automatizado baseados em inteligência artificial apresenta desafios significativos, especialmente no que diz respeito à privacidade e à segurança de dados. Eles ressaltam que "é crucial equilibrar o uso de tecnologias avançadas com o respeito aos direitos individuais, evitando a vigilância excessiva e invasões de privacidade".

Além disso, a transparência sobre o funcionamento da IA no monitoramento das ligações é indispensável para que a tecnologia seja percebida como um recurso de melhoria contínua e não como uma ferramenta de vigilância opressiva.

Outro ponto importante no nosso trabalho é o risco de viés algorítmico na categorização das interações. Se o modelo for treinado com um conjunto de dados enviesado, ele pode reforçar padrões de julgamento inadequados, comprometendo a precisão da análise e impactando negativamente a avaliação do atendimento.

7) Dataset, se for o caso (anonimizados quando necessário), descrição detalhada do seu conteúdo/origem, análise exploratória e preparação dos dados em Python

Nosso dataset consiste em um conjunto de 30 áudios provenientes de uma distribuidora de um escritório de investimentos, contendo gravações de cold calls realizadas pelos atendentes, esses dados representam interações reais entre vendedores e potenciais clientes. Os dados estão organizados em uma tabela simples com duas colunas: uma coluna contendo o identificador único da gravação (ID) e outra contendo o caminho do arquivo de áudio correspondente.

Para a análise exploratória e preparação dos dados em Python, inicialmente realizamos a conversão dos áudios em texto por meio do modelo de transcrição automática Whisper da OpenAI. O Whisper já executa uma transcrição bastante limpa, removendo pausas prolongadas, ruídos e palavras de preenchimento, além de lidar bem com diferentes sotaques e entonações.

De acordo com Silva (2023), o modelo Whisper tem capacidade robusta para transcrição de áudios, mesmo em condições adversas. No entanto, dependendo da qualidade do áudio, aplicamos um pós-processamento para ajustes adicionais, incluindo correção ortográfica, remoção de repetições e identificação de segmentos mais relevantes da conversa.

8) Metodologia e Resultados Esperados: apresentar a abordagem que pretende empregar na resolução do problema e quais são os resultados esperados.

Para a resolução do problema proposto, será utilizada a metodologia KDD (Knowledge Discovery in Databases), que consiste em um processo estruturado para a descoberta de conhecimento a partir de dados. O KDD é composto por cinco etapas principais: Seleção, Preparação dos dados, Transformação, Mineração de Dados e Avaliação.

Na etapa de Seleção, o projeto utilizará um conjunto de 30 áudios provenientes de um escritório de investimentos, contendo gravações de cold calls realizadas pelos

atendentes. Esses dados serão armazenados em uma tabela contendo um identificador único para cada gravação e o caminho do respectivo arquivo de áudio. Na fase de Preparação dos dados, os áudios serão convertidos em texto por meio do modelo de transcrição automática Whisper da OpenAI, que já realiza uma limpeza inicial ao remover pausas prolongadas, ruídos e palavras de preenchimento.

Em seguida, na etapa de Transformação, serão empregadas técnicas de Processamento de Linguagem Natural (NLP) para estruturar e organizar os dados textuais. Entre as transformações aplicadas, destaca-se a extração de palavras-chave, análise de sentimentos e a identificação de padrões linguísticos que possam indicar interações bem-sucedidas ou problemáticas.

Na fase de Mineração de Dados, serão utilizados algoritmos de aprendizado de máquina não supervisionado, especialmente técnicas de agrupamento (clustering), para categorizar os atendimentos com base em suas características. O objetivo dessa etapa é identificar padrões recorrentes nas interações, como as abordagens mais eficazes, dificuldades comuns enfrentadas pelos atendentes e possíveis objeções dos clientes.

Por fim, na etapa de Interpretação e Avaliação, os resultados obtidos serão analisados para extração de insights estratégicos. A partir dos padrões identificados, será possível sugerir melhorias nos atendimentos, direcionar treinamentos específicos para os atendentes e otimizar os scripts utilizados nas chamadas. Além disso, os modelos de agrupamento serão validados para garantir que os resultados sejam representativos e úteis para a tomada de decisão.

Utilizando essa metodologia o resultado esperado é alcançar uma análise automatizada e precisa das interações telefônicas, eliminando avaliações manuais e subjetivas. Além disso, busca-se identificar padrões de comunicação que contribuam para a melhoria dos atendimentos, permitindo a otimização de abordagens e scripts utilizados pelos atendentes. Outro resultado esperado é a segmentação das ligações, possibilitando uma categorização eficiente das interações com base em características como perfil do cliente e sentimento predominante. Por fim, o projeto visa gerar insights estratégicos para aprimoramento contínuo do atendimento, promovendo uma experiência mais assertiva e eficiente tanto para os atendentes quanto para os clientes.

9) Referências: citadas dentro do texto do projeto

Trindade, R., & Oliveira, M. (2024). Inteligência Artificial e Monitoramento Automatizado: Desafios e Soluções para Privacidade e Segurança de Dados. Recuperado de <https://revistaft.com.br/inteligencia-artificial-e-monitoramento-automatizado-desafios-e-solucoes-para-privacidade-e-seguranca-de-dados>

PUCRS. (s.d.). Mineração de Dados Aplicada à Análise de Interações Telefônicas. Recuperado de <https://tede2.pucrs.br/tede2/bitstream/tede/3044/1/388093.pdf>

Silva, J. (2023). *Um estudo de caso do modelo de reconhecimento de voz Whisper para transcrição de Conferências TEDx via aprendizado fraco*. Recuperado de <http://repositorio.ufc.br/handle/riufc/73941>

10) Bibliografia: na qual o grupo se baseará para criar e desenvolver o projeto e/ou Referências citadas dentro do texto do projeto.

OpenAI. (2022). *Whisper: Robust Speech Recognition via Large-Scale Weak Supervision*. Disponível em: <https://openai.com/research/whisper>. Acesso em: 26 mar. 2025.

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4ª ed.). Pearson.

Jurafsky, D., & Martin, J. H. (2020). *Speech and Language Processing* (3ª ed.). Pearson.

Brasil. (2020). *Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709/2018*. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 26 mar. 2025.

Silva, J. (2023). *Um estudo de caso do modelo de reconhecimento de voz Whisper para transcrição de Conferências TEDx via aprendizado fraco*. Disponível em: <http://repositorio.ufc.br/handle/riufc/73941>. Acesso em: 26 mar. 2025.