

Aprendizagem por Projetos Integrados 2025-1

Parceiro:	IT-01	
Período / Curso:	6º DSM	
Professor M2:	Prof. Leônidas	leonidas.alvarez@fatec.sp.gov.br
Professor P2:	Prof. Dawilmar	dawilmar.araujo@fatec.sp.gov.br
Contato do Parceiro:	Rubsney Nascimento	rubsney.nascimento@ito1.com.br

Tema do Semestre

Sumarização de texto. Extração de dados de grandes massas de dados e identificação dos principais pontos. Evitar: abordar aspectos generativos da linguagem pois requer estudo avançado de estrutura e dependência linguística.

Conhecimentos ensinados no semestre

Listar todos os conhecimentos e tecnologias ensinadas no semestre – Baseado na Matriz de competências do semestre

- Identificar etapas necessárias em um pipeline de PLN, selecionar corpora de texto para serem utilizados no treinamento e validação de sistemas de PLN, identificar e aplicar ferramentas de inteligência artificial e matemática computacional a resolução de problemas de PLN;
- Implementar um software aplicando conhecimentos de ES, programação e gerência de projetos. Utilizar desenvolvimento front-end e back-end integrando as aplicações desktop, web e mobile em projetos que atendam aos requisitos de transparência das aplicações;
- Empregar metodologias que visem garantir critérios de qualidade no desenvolvimento de uma solução computacional;
- Identificar as necessidades dos projetos buscando adaptá-las aos mais diversos meios de hospedagem, compartilhado, cloud, virtual private server, collocation e servidor dedicado, utilizando o recurso adequado de cada um. Implantar sistemas nas diversas infraestruturas de Redes de Computadores, buscando a melhor performance.

Título do Desafio

Definir o problema em uma Frase

Automação e Extração de Informações em Atendimento Inteligente

Descrição do Desafio

Definir entre 2 e 3 parágrafos

Empresas que utilizam sistemas de atendimento automatizado frequentemente lidam com grandes volumes de dados textuais gerados por interações entre clientes e sistemas de IA (*chatbots*, assistentes virtuais e *CRMs*). A análise dessas interações pode fornecer insights valiosos, mas a extração de dados relevantes e a sumarização eficiente das informações ainda são desafios.

Este desafio visa desenvolver uma solução que permita a extração automática dos principais pontos de conversas, classificação das interações e geração de insights diários sobre os atendimentos. O

objetivo é melhorar a eficiência do atendimento, reduzir a sobrecarga manual na análise e facilitar a tomada de decisão com base nos dados extraídos.

Requisitos Funcionais e Não Funcionais

Listar entre 5 e 7 Itens

Requisitos Funcionais:

- Implementar um mecanismo de extração de informações baseado em técnicas de Processamento de Linguagem Natural (PLN).
- Oferecer funcionalidade de busca semântica (ex: Quanto tempo a solicitação XPTO levou para ser executada?)
- Criar uma funcionalidade de sumarização automática de interações para destacar os principais pontos das conversas.
- Desenvolver um módulo de classificação automática de atendimentos para categorização eficiente (exemplo: reclamação, dúvida, solicitação de suporte).
- Integrar insights diários sobre as interações realizadas, fornecendo métricas como tempo médio de atendimento, principais dúvidas e padrões de interação. Estes insights devem ser apresentados em *DASHBOARDS*.
- Oferecer funcionalidade de cadastro de usuários. Usuários têm dois níveis de acesso: Admin e Viewer. Viewers não podem criar usuários e não têm acesso a dados pessoais (PII).

Requisitos Não Funcionais:

- Garantir que a solução seja escalável e possa lidar com grandes volumes de dados textuais sem comprometer o desempenho.
- Implementar padrões de segurança para proteger os dados dos usuários e manter a conformidade com a LGPD/GDPR.
- A solução deve funcionar mesmo que o volume de dados seja alto, sem exceder o máximo de tokens das APIs de LLM utilizadas (RAG deve ser aplicado para garantir que apenas informações relevantes são incluídas nos prompts)
- Onde necessário, a aplicação deve remover dados pessoais usando técnicas de NLP.

Outras informações fornecidas pelo Parceiro

Informações relevantes ao projeto

- Fonte de Dados Externos: API do Jira e arquivos CSV (para compatibilidade com outros sistemas)
- API Application Programming Interface, existentes: OpenAI, Microsoft Presidio (PII removal), Pinecone (Vector Database para busca semântica), LangChain (orquestração de solução LLM. N8N também é uma alternativa).
- Aplicações similares:
 - <https://marketplace.atlassian.com/apps/1224025/ai-insights-for-jira-service-management?tab=overview&hosting=cloud>
 - <https://www.domrock.net/>