

Processamento e Interação em Linguagens Naturais em Texto, Vídeo e Áudio

Módulo 7: Engenharia de Software

– RESUMO EXECUTIVO

Este módulo capacita estudantes a desenvolver sistemas inteligentes que processam e interpretam linguagem natural em diferentes formatos: texto, áudio e vídeo. As soluções incluem reconhecimento de voz, análise de sentimentos e processamento de linguagem natural (PLN) e interação assíncrona (UX), bem representado na indústria por aplicativos de mensagens como por exemplo o Whatsapp(MR) e Telegram(MR).

Durante 10 semanas, uma turma dedicada de alunos e seus professores irá entender o seu desafio e propor soluções utilizando metodologias ágeis. Ao final do cronograma, sua empresa receberá até 5 protótipos funcionais, permitindo que você acelere e valide ideias.

Com o boom das IAs Generativas, implementar a automação de análise de sentimentos, transcrição e chatbots inteligentes resolve o problema de alto custo operacional em atendimento. Neste projeto, os alunos aprimoram esses serviços aliando competências de gestão de processos, modelagem financeira e estratégia de inovação, formando profissionais capazes de entregar soluções completas com impacto organizacional.

– O DESAFIO

Empresas de todos os setores geram e recebem volumes crescentes de informações em formato de texto, áudio e vídeo: chamadas de atendimento, gravações de reuniões, avaliações de clientes, documentos internos, interações em redes sociais. A capacidade de processar, interpretar e agir sobre esse conteúdo de forma automatizada tornou-se um diferencial competitivo.

Sem ferramentas adequadas, organizações perdem oportunidades de identificar padrões, antecipar demandas e personalizar interações. O processamento manual desses conteúdos é lento, caro e sujeito a inconsistências, limitando a escalabilidade das operações.

– OBJETIVO DO MÓDULO

Desenvolver uma aplicação funcional que utilize processamento de linguagem natural para interpretar, classificar e gerar insights a partir de dados em texto, áudio ou vídeo, atendendo a uma necessidade real da empresa parceira e demonstrando impacto mensurável no negócio.

– PROBLEMAS QUE PODEMOS RESOLVER

1. Automação da Análise de Conteúdo

- Análise automática de sentimentos em avaliações, comentários e interações de clientes.
- Classificação inteligente de documentos, chamados e comunicações internas.
- Transcrição e interpretação automática de áudio e vídeo (speech-to-text).

2. Interação Inteligente com Usuários

- Desenvolvimento de interfaces conversacionais (chatbots, assistentes virtuais).
- Reconhecimento de comandos de voz e intenções do usuário.

- Personalização de respostas e recomendações baseadas em linguagem natural.

3. Inteligência Estratégica e Operacional

- Detecção de tendências e padrões em grandes volumes de texto e áudio.
- Monitoramento de reputação e percepção de marca em múltiplos canais.
- Automação de processos de compliance e auditoria documental.

– POSSÍVEIS CASOS DE APLICAÇÃO

1. Casos de Sucesso Reais (O que já entregamos)

- **Serviços Financeiros:** Desenvolvemos com o **Bank of America** um sistema avançado de análise de sentimentos focado nas interações dos clientes, com o objetivo de identificar riscos reputacionais e oportunidades de melhoria em produtos financeiros.
- **Tecnologia:** Criamos para a **IBM** uma plataforma de Processamento de Linguagem Natural (PLN) que realiza a classificação automatizada de chamados de suporte, permitindo o direcionamento inteligente e imediato do atendimento.
- **Energia e Utilidades:** Estruturamos para a **Comgás** um sistema de análise de *feedback* de clientes e transcrição de chamadas, detectando padrões de insatisfação para gerar oportunidades de melhoria operacional.
- **Mídia e Entretenimento:** Implementamos com a **Sambatech** uma plataforma de análise de conteúdo audiovisual que automatiza a transcrição, categorização e geração de metadados para otimizar catálogos.

2. Outras Aplicações Possíveis (Onde mais podemos gerar valor)

- **Saúde e Medicina:** Sistemas de transcrição e análise de consultas médicas em áudio, oferecendo apoio automatizado à documentação clínica e detecção de eventos adversos nos prontuários.
- **Jurídico e Compliance:** Automação da leitura e análise de grandes volumes de contratos e documentos legais, identificando rapidamente cláusulas críticas e potenciais riscos de conformidade.
- **Varejo e Atendimento (E-commerce):** Criação de interfaces conversacionais (assistentes virtuais e chatbots) mais humanos, que reconhecem comandos de voz e intenções reais do usuário para personalizar respostas e recomendações.
- **Gestão de Marca e Marketing:** Monitoramento automatizado da percepção da marca (*Social Listening*) e detecção de tendências extraíndo inteligência de grandes volumes de menções em redes sociais e avaliações de produtos.

– POSSÍVEIS ENTREGÁVEIS

Os grupos desenvolvem uma aplicação funcional de PLN validada, podendo incluir:

- Sistema de análise de sentimentos com classificação automática de textos.
- Motor de processamento de linguagem natural com reconhecimento de entidades e intenções.
- Interface conversacional (chatbot ou assistente virtual) integrada à solução.
- Pipeline de transcrição de áudio para texto com análise automatizada.
- Dashboard com visualização de insights extraídos dos dados processados.
- Documentação técnica e de negócios com relatório de validação.

– **PREMISSAS E REQUISITOS**

1. Da Empresa Parceira:

- Disponibilização de dados textuais, de áudio ou vídeo para análise (anonimizados quando necessário).
- Acesso a especialistas do negócio e, quando necessário, especialistas técnicos para reuniões quinzenais.
- Definição clara do problema de negócio e dos critérios de sucesso esperados.
- Abertura para testar e validar as soluções propostas ao longo do desenvolvimento.

2. Do Projeto:

- Foco em problemas que envolvam processamento de linguagem natural em texto, áudio ou vídeo.
- Volume mínimo de dados suficiente para treinar e validar os modelos.
- Potencial de impacto mensurável no negócio ou na experiência do usuário.
- Alinhamento com práticas éticas, LGPD e políticas de segurança da informação.

– **ESCOPO E RESTRIÇÕES**

1. Escopo Incluído:

- Desenvolvimento de modelos de processamento de linguagem natural.
- Construção de interfaces para interação e visualização de resultados.
- Integração de componentes de reconhecimento de áudio (speech-to-text).
- Documentação técnica completa e documentação de negócios.
- Validação funcional com dados reais.

2. Fora do Escopo:

- Implementação em ambiente de produção (deployment final).
- Integração definitiva com sistemas legados da empresa.
- Manutenção e suporte pós-projeto.
- Processamento em tempo real de streaming de vídeo em larga escala.

– **POSSÍVEIS TECNOLOGIAS / ARQUITETURA / FERRAMENTAS DE NEGÓCIOS**

Tecnologias e Arquitetura:

- Python, bibliotecas de NLP (NLTK, spaCy, Hugging Face Transformers).
- Arquitetura orientada a serviços e Design Patterns.
- APIs de speech-to-text e text-to-speech.
- Modelos de aprendizado de máquina (classificação, regressão, agrupamento).
- Frameworks web (Flask, FastAPI).
- Ferramentas de visualização de dados.

Ferramentas de Negócios:

- Ferramentas de BPM (Business Process Management).
- Modelos de negociação e análise de cenário.
- Frameworks de gerenciamento financeiro (balanço, custos, rentabilidade).
- Ferramentas de prototipação e modelagem de sistemas.