

Sistema de processamento de linguagem natural com IA generativa

Módulo 11: Ciência da Computação

- RESUMO EXECUTIVO

Neste módulo, equipes de alunos desenvolvem soluções de processamento de linguagem natural (NLP) e IA generativa para resolver problemas organizacionais onde sistemas computacionais precisam compreender, interpretar e gerar linguagem humana. O trabalho usa Large Language Models (LLMs) e técnicas de transfer learning em contextos reais do parceiro.

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.

- O DESAFIO

A empresa traz um problema onde grandes volumes de texto, conversas ou documentos em linguagem natural precisam ser processados, compreendidos ou gerados de forma automatizada. O desafio pode envolver atendimento, análise de conteúdo, geração de relatórios ou extração de informações de fontes textuais.

- OBJETIVO DO MÓDULO

1. Desenvolver soluções de NLP usando LLMs, RNNs e técnicas de transfer learning.
2. Aplicar álgebra linear em word embeddings e análise de séries temporais de texto.
3. Implementar computação paralela em cluster para processamento em larga escala.
4. Analisar o impacto da IA generativa nos negócios, estratégias de sustentabilidade e frameworks de ESG.
5. Desenvolver capacidade de liderança com foco em impacto positivo e gestão de experimentos com IA.

- PROBLEMAS QUE PODEMOS RESOLVER

- Atendimento ao cliente com volumes que tornam respostas manuais inviáveis ou lentas demais.
- Extração e sumarização de informações relevantes de grandes bases de documentos não estruturados.
- Classificação automática de textos, solicitações ou feedbacks por categoria, urgência ou sentimento.
- Geração assistida de relatórios, contratos ou comunicações padronizadas a partir de dados estruturados.
- Tradução, adaptação ou personalização de conteúdo textual em escala para diferentes audiências.

- POSSÍVEIS CASOS DE APLICAÇÃO

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

- **Mobilidade e Tecnologia:** Trabalhamos com a **Uber** no desenvolvimento de um sistema de classificação e roteamento automático de solicitações de suporte em linguagem natural, focado na redução drástica do tempo de resolução de chamados.
- **Telecomunicações e Serviços:** Criamos para a **Brastel** uma solução de PLN para o processamento e a categorização inteligente de comunicações internacionais e contratos em múltiplos idiomas.
- **Mídia e Conteúdo:** Estruturamos com a **Curadobia** uma plataforma de geração assistida de conteúdo curatorial, utilizando IA generativa para a produção de textos hiperpersonalizados em escala.

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

A IA Generativa e o Processamento de Linguagem Natural (PLN) transformam a maneira como as empresas lidam com o conhecimento. Nossos alunos podem construir aplicações avançadas de IA para ler, interpretar e redigir textos, resolvendo gargalos críticos em setores como:

- **Atendimento ao Cliente:** Criação de assistentes virtuais de última geração (*chatbots* inteligentes) que realmente entendem a intenção do cliente, resolvendo problemas complexos sem intervenção humana e reduzindo filas de espera.
- **Jurídico e Compliance:** Automatização do processo de revisão de contratos e documentos legais, identificando instantaneamente cláusulas-chave, riscos, ambiguidades e oportunidades.
- **Saúde e Auditoria:** Processamento de textos médicos não estruturados para extrair informações de prontuários clínicos, apoiando a decisão médica e otimizando a auditoria de faturamentos.
- **Recursos Humanos:** Ferramentas automatizadas para análise inteligente de currículos (*matching* de talentos) e mensuração do engajamento e moral da equipe a partir da análise de sentimentos em *feedbacks* internos em texto.
- **Finanças e Mercado:** Monitoramento inteligente de notícias financeiras, relatórios econômicos e redes sociais para identificar tendências de mercado de forma consolidada e apoiar decisões de investimentos.

- POSSÍVEIS ENTREGÁVEIS

- Aplicação funcional de NLP com pipeline de processamento documentado.
- Modelo avaliado com métricas de desempenho (BLEU, ROUGE, F1, acurácia por contexto).
- Relatório técnico com análise de erros e estratégias de melhoria.
- Apresentação executiva com análise de impacto operacional, viabilidade de escala e considerações de ESG.

- **PREMISSAS E REQUISITOS**

1. Da Empresa Parceira:

- Fornecer base de dados textual representativa do problema (conversas, documentos, feedbacks) com volume suficiente para fine-tuning ou avaliação.
- Definir critérios mensuráveis de qualidade para a solução de NLP.
- Disponibilizar responsável de negócio para validar exemplos de entrada e saída esperada ao longo do módulo.

2. Do Projeto:

- O dataset textual deve estar disponível antes do início ou nas primeiras semanas do módulo.
- A solução será um protótipo funcional de NLP, não necessariamente integrado a todos os sistemas de produção.
- Restrições de privacidade e conformidade de dados (LGPD) devem ser documentadas antes do início.

- **ESCOPO E RESTRIÇÕES**

1. Escopo Incluído:

- Análise exploratória do dataset e definição da estratégia de NLP.
- Implementação do pipeline de processamento com LLMs ou modelos de linguagem especializados.
- Avaliação do modelo com métricas relevantes ao problema do parceiro.
- Interface funcional de uso e documentação técnica completa.

2. Fora do Escopo:

- Treinamento de LLMs do zero (uso de modelos pré-treinados com fine-tuning)
- Integração com sistemas de produção em tempo real sem supervisão do parceiro.
- Desenvolvimento de infraestrutura de GPU própria para o projeto.
- Contratação de ferramentas ou plataformas específicas condicionadas ao ambiente do parceiro.

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

- Python
- Hugging Face Transformers / LangChain
- OpenAI API / Gemini API (LLMs)
- spaCy / NLTK (NLP clássico)
- PyTorch / TensorFlow
- FastAPI (deployment)
- Docker / Ray (paralelização)
- PostgreSQL / Pinecone (armazenamento vetorial)