



Pró-Reitoria de Pesquisa e Pós-Graduação Coordenadoria Geral de Pós-Graduação Stricto Sensu

PLANO DE ENSINO

Unidade Acadêmica: Faculdade de Computação e Informática – FCI				
Programa de Pós-Graduação em Computação Aplicada - PPGCA Mestrado Profissional em Computação Aplicada				
Curso: ☐ Mestrado Acadêmico ☐ Mes	strado Pro	fissional	utorado Profissional	
Modelos de Linguagem e Generativos				Código: ENST XXXXX
Professor(es): Rogério de Oliveira Leandro Augusto da Silva				DRT : 11.1566-5
Carga horária: 48	Crédito	s 04	Obrigatória Optativa Eletiva	
Ementa: O curso oferece uma introdução teórica e prática dos fundamentos dos modelos de linguagem (que envolvem o processamento de linguagem natural) e generativos (que geram novos dados, sejam de linguagem ou outras mídias), e que encontram grande aplicabilidade na geração de conteúdo, Chatbots, agregadores de notícias, sistemas de recomendação, gerenciamento de ativos de mídia etc. Os alunos aprenderão os principais conceitos e técnicas associadas ao pré-processamento de linguagem, principais modelos de representação de linguagem para processamento automático e o aprendizado de máquina (bag of words, TF-IDF, word embedding, transformers), geração de imagens (redes CNN, auto-encoders variacionais, GANs), ferramentas para construção de aplicações e agentes de linguagem natural (APIs, bancos de dados vetoriais, orquestradores), técnicas de fine-tuning e transfer-learning para diferentes tarefas, e modelos multimodais, com ênfase na implementação com o uso de modelos abertos, bibliotecas como NLTK e SpaCy, e frameworks como TensorFlow e PyTorch.				
Objetivos: Proporcionar ao aluno um de modernos modelos de				definição, construção e avaliação a diferentes problemas.
Conceitos		Habi	lidades	Valores



Pró-Reitoria de Pesquisa e Pós-Graduação Coordenadoria Geral de Pós-Graduação Stricto Sensu



- Conhecer os principais conceitos e princípios que envolvem o Processamento da Linguagem Natural, e os atuais modelos de Linguagem e Generativos.
- Entender a aplicação e o potencial do uso de Modelos de Linguagem e Generativos em diferentes domínios e tarefas (como Classificação Automática, ChatBots, Sumarização, Tradução, Geração de Conteúdo etc.)
- Estar apto a empregar as principais ferramentas para o desenvolvimento de aplicações de modelos de linguagem e generativos (modelos abertos, APIs, orquestradores, implementação de finetuning de modelos prétreinados etc.).
- Reconhecer as
 possibilidades de uso
 dessas técnicas e
 ferramentas para a
 solução de diferentes
 problemas, e os modelos
 e técnicas mais
 adequados a serem
 empregados em cada
 caso.

- Valorizar o conhecimento científico e a interdisciplinaridade que envolvem o uso adequado dos modelos de Linguagem e Generativos.
- Estar atento e apto a entender a evolução e o estado da arte desses modelos.
- Ter e promover a Ética e o uso Responsável dessas técnicas na construção de modelos e aplicações.



Pró-Reitoria de Pesquisa e Pós-Graduação Coordenadoria Geral de Pós-Graduação Stricto Sensu



Conteúdo Programático

1. Introdução aos Modelos de Linguagem

 Definição de modelos de linguagem, suas aplicações e importância; Evolução dos modelos de linguagem.

2. Ténicas de pré-Processamento de Textos

 Técnicas fundamentais como normalização, tokenização, stemming e lematização; Bibliotecas de processamento de linguagem natural: NLTK e SpaCy.

3. Modelos Simples: Bag of Words e TF-IDF

 Introdução aos métodos Bag of Words e Term Frequency-Inverse Document Frequency (TF-IDF); Medidas de Similaridade (distância cosseno).

4. Modelos Baseados em N-gramas e Word Embeddings

 Uso de n-gramas para modelagem de linguagem e suas aplicações; Introdução a embeddings de palavras: Word2Vec e GloVe.

5. Transformers e Atenção

 Modelos de redes neurais recorrentes e suas limitações; Estrutura e funcionamento de modelos transformers e o mecanismo de atenção.

6. Transfer Learning

 Abordagens de transfer learning em processamento de linguagem natural e implementação prática do uso de modelos pré-treinados, como BERT, GPT, Llama, em tarefas de aprendizado de máquina; Técnicas de Contexto.

7. Tarefas e Métricas de Modelos de Linguagem

 Métodos de avaliação para modelos de linguagem, incluindo perplexidade, precisão, recall e F1-score; Avaliação e interpretar resultados de diferentes modelos; Prompts; Classificação de Textos; Tradução e Sumarização Automática;

8. APIs, Bancos de Dados Vetoriais e Agentes de IA

 APIs disponíveis (como OpenAI, Hugging Face) e seu uso; Criação de agentes de conversação utilizando modelos de linguagem; Uso de orquestradores (LangChain, LangGraph, MCP); Bancos de Dados Vetoriais (Chroma DB, Faiss).

9. IA Generativa: Introdução e Aplicações

 Conceitos fundamentais sobre IA generativa e suas aplicações em diferentes mídias; Geração Aumentada de Recuperação (RAG); Discussão sobre modelos de geração de texto, imagem e som; Viés e Aspectos Éticos.

10. Modelos de Imagem

 Modelos neurais CNN (redes convolucionais) para imagens; Extensões para vídeos; Frameworks e modelos pré-treinados (Yolo).

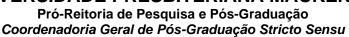
11. Modelos Generativos

 Auto-encoder Variacional; Generative Adversarial Networks (GANs); outros modelos; Modelos pré-treinados e Fine Tunning.

12. Modelos Multimodais

 Integração de múltiplas fontes de dados (texto, imagem, som); Aprendizado Multi-modal







Metodologia:

Aulas expositivas com atividades práticas de desenvolvimento de modelos de linguagem e generativos, além de um projeto final empregando dados de fontes primárias.

Critério de Avaliação:

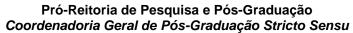
Haverá uma Nota Final (NF) de 0.00 a 10.0, média aritmética da média atividades práticas (mínimo 4 atividades) e do projeto.

NF = 0.5 média das atividades + 0.5 nota do projeto final

O **Grau Final** segue o Regulamento Geral da Pós-Graduação Stricto Sensu, Art. 98, que diz: "Será considerado aprovado o aluno que obtiver, em cada disciplina obrigatória, optativa e nas atividades programadas o conceito final "A", "B" ou "C", conforme relação de conceitos a seguir:

I - A - EXCELENTE: corresponde a NF entre 9,0 e 10,0 corresponde a NF entre 8,0 e 8,9 corresponde a NF entre 7,0 e 7,9 corresponde a NF abaixo de 6,9.







Bibliografia:

Básica

- 1. Jurafsky, D., & Martin, J. H. (2020). Speech and Language Processing (3rd ed.). Pearson.
- 2. Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly Media.
- 3. Goyal, P., & Gupta, A. (2020). Deep Learning for Natural Language Processing: A Hands-on Guide. Springer.
- 4. Ba, Y., et al. (2017). *Neural Networks for Natural Language Processing*. Addison-Wesley.
- 5. Tunstall, L., von Werra, L., & Wolf, T. (2022). *Natural Language Processing with Transformers: Building Language Applications with Hugging Face and PyTorch*. O'Reilly Media.
- 6. Foster, D. (2021). Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media.

Complementar:

- 1. Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). Dive into deep learning. Cambridge University Press.
- 2. Bengio, Y., Goodfellow, I., & Courville, A. (2016). Deep Learning. MIT Press.
- 3. Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media.